MASTER'S THESIS

Rapid and Robust Fitting of Cole-Cole Models to Electrical Permittivity Spectra

Robin Mäki
2014

Master of Science in Engineering Technology
Engineering Physics and Electrical Engineering

Luleå University of Technology
Department of Engineering Sciences and Mathematics

LULEÅ UNIVERSITY OF TECHNOLOGY

Rapid and Robust Fitting of Cole-Cole Models to Electrical Permittivity Spectra

Robin Mäki

Luleå University of Technology
Departement of Engineering Sciences and Mathematics

Abstract

The dielectric properties of materials plays an important role in many fields, these properties can be described by the Cole-Cole parameters. In this thesis an automated method for fitting of the Cole-Cole model to relative permittivity spectrum has been developed. Two different iterative estimation techniques were implemented and analyzed based on their performance. The implemented techniques were the Levenberg-Marquardt method and the variable projection method.

One of the main difficulties with implementing an automated estimation method is the need for good initial values. The variable projection method factorizes the Cole-Cole equation into a linear and non-linear part. By doing this the problem can be reduced down from a five dimensional problem into a problem only depending on the two non-linear parameters. The reduction in dimensionality makes it viable to find initial values by performing a grid search.

A Monte-Carlo simulation was part in the validation and comparison of the methods. The result from this was that the variable projection method was more time consuming but were able to obtained tighter uncertainty bounds for all of the parameters and a smaller mean square error for the fitted curves.

The methods were applied to a data set containing measurements of the relative permittivity of 15 crude oils. The result from this was consistent with that of the Monte-Carlo simulation. The variable projection method obtained better fitting curves and smaller mean square error for all oils.

Preface

The making of this thesis concludes the fifth and final year of study for my Master's degree in Engineering Physics and Electrical Engineering at Luleå University of Technology.

I would like to take the opportunity to thank Prof. Johan Carlson and Prof. Inge Söderkvist for their support and encouragement throughout the project. I am also very thankful to Kjetil Folgerø and Christian Michelsen Research AS for providing the measurement data and valuable feedback.

Finally, a special thanks to my family for their encouragement and support during these five years.

Contents

CHAPTER 1

Introduction

The Cole-Cole equation [1], is a relaxation model often used to model relative permittivity, ε_r. The model parameters describe dielectric properties that, in combination with other metrics, are used when characterizing crude oils, [2]. The measured relative permittivity is a complex valued quantity that is dependent on the (angular) frequency, ω [rad/s]. Current research aims at developing methods for automatic measurement of the dielectric parameters, using a wide range of frequencies. Fitting the Cole-Cole model to the measured relative permittivity spectrum is done by adjusting the unknown parameters. Since the size of the Cole-Cole parameters varies by several orders of magnitude the estimation problem is ill-conditioned. Hence, the parameter estimation is sensitive to experimental noise. In order to assure decent to global minimum good initial guesses of the parameters needs to be available. Furthermore, in order for the method to be of practical use in an online setup it has to be automated, robust and reasonably fast.

1.1 Goal

The primary objective for this thesis is to:

- Develop and implement a rapid and robust estimation technique for fitting of Cole-Cole models to relative permittivity measurements.
- The developed algorithm should be fully automated, i.e. no manual adjustments or initialization sequences should be required.

1.2 Related work/Literature study

The Cole-Cole model [1], is a equation used to describe the dielectric relaxation in different materials. The model is used in a wide area of fields, most notably in the field of biology, where it is used for characterizing the electrochemical properties of biological tissues and biochemical materials [3], and geology, where the method of spectral induced polarization is widely used in geophysical surveys and the interpretation of the results are often based on the Cole-Cole model [4],[5],[6]. In all of those cases it is used in order to describe the impedance or permittivity in dielectric materials.

Similarly to other fields, dielectric properties are used in the oil industry where they, in combination with other metrics, are used in the quality control and characterization of crude oils, [2]. They are also used in several Multiphase Flow Metering systems [7], which are used for online monitoring of the flow rates of gas, water and oil in the pipelines.

Since the introduction and spread of the Cole-Cole model different techniques have been developed for estimating the dielectric parameters.

A estimation technique based on least absolute deviation (LAD) was developed, [8], for extraction of the Cole-Cole parameters.

But commonly the estimation problem is treated as a least squares problem and solved by using iterative Gauss-Newton-based schemes, [9], [10]. The damped least squares algorithm developed by Levenberg, [11], and Marquardt, [12], called the Levenberg-Marquardt algorithm is a Gauss-Newton base algorithm that uses a trust region strategy. It is more robust than the traditional Gauss-Newton algorithm, but has similar to other iterative techniques it is dependant on good initial guesses in order to assure decent to the global minima. Difficulties with these approaches.

Since the estimation results strongly depend on the choice of starting values a method for finding initial values to these algorithms is important. A method where purposed in [13] which gave the suggestion that a Markov-chain Monte-Carlo based method can be used first to obtain the medians of unknown parameters by starting from an arbitrary set of initial values. The median could then be used as a initial guess.

The Cole-Cole equation can be separated in to two parts that depends linearly and nonlinearly on the parameters, respectively. Golub and Pereyra [14] developed a Gauss-Newton algorithm for solving separable nonlinear least squares problem. Borges later extended the algorithm into a full-Newton algorithm [15]. The advantage of separable least squares problem approach is that it allows the problem to be reduced to only depend on the nonlinear parameter which will increase the robustness.

1.3 Thesis structure

The outline of this thesis will be as follows: Chapter 2 contain the theoretical framework of the thesis. The chapter begins with giving an introduction to permittivity and the Cole-Cole model. This is followed by derivation of the Levenberg-Marquardt algorithm and the variable projection method. Chapter 3 contain some of the key aspects of the numerical implementation of the methods. The result of the validation and comparison of the methods can be found in Chapter 4. In the last two chapters the results are discussed and recommendations for future improvements to the algorithms are provided.

1.4 Frequently used variables

$\boldsymbol{a}$ - The linear model parameters.
$\boldsymbol{A}$ - The model matrix.
$\boldsymbol{b}$ - The non-linear model parameters.
$\boldsymbol{D}$ - The electric displacement field.
D_0 - The amplitude of the electric displacement field.
$\boldsymbol{E}$ - The electric field.
E_0 - The amplitude of the electric field.
f - The model function.
$\boldsymbol{g}$ - The gradient.
$\boldsymbol{h}$ - The search direction.
$\boldsymbol{H}$ - The Hessian matrix.
$\boldsymbol{I}$ - The identity matrix.
j - The imaginary unit.
$\boldsymbol{J}$ - The Jacobian matrix.
k - Iteration index.
m - Number of data points.
n - Number of parameters.
$\boldsymbol{p}$ - The model parameters.
$\boldsymbol{p}_0$ - The initial guess of the model parameters.
$\boldsymbol{P}_d$ - The polarization density.
$\boldsymbol{P}$ - The projection matrix. Projects vectors onto the range of A.
$\boldsymbol{P}^{\perp}$ - The projection matrix. Projects vectors onto the null space of A^T.
$\boldsymbol{r}$ - The residual, i.e. the difference between the measure data and the model.
S - The sum of the square residual, i.e. the cost-function.
t - Time.
$\boldsymbol{x}$ - Independent variable of the model function, e.g. frequency or time.
$\boldsymbol{X}$ - A Matrix of linear equations, with more equations than unknown coefficients.
$\boldsymbol{y}$ - The measure data.
α - The distribution factor.
γ - The step length.
Δ - The radius of the trust-region.
ε - The permittivity.
ε_0 - The permittivity in vacuum.
ε_r - The Relative permittivity.
ε_s - The static permittivity.
ε_∞ - The high-frequency permittivity.
λ - The damping parameter.
ρ - The gain ratio.
σ - The conductivity.

τ - The relaxation time.
ϕ - The quadratic model function.
ω - The angular frequency.

^{T} - Matrix transpose.
† - The Moore-Penrose generalized inverse.
* - Indicates the solution to an equation.
∇ - The Nabla operator.
$^\perp$ - Denotes projection to an orthogonal subspace.

Chapter 2

Theory

This section will give the reader an introduction to the theoretical framework used in this thesis. A introduction of permittivity, [16, ch. 5], and the Cole-Cole model, [1], will be presented followed by theory regarding techniques and methods used to perform the curve fitting, [17, 18]. The main focus will lay on two different curve fitting techniques, the Levenberg-Marquardt method, [11, 12], and the Variable projection method [14].

2.1 Permittivity and dielectric materials

A dielectric medium is an electrical insulator, a material whose internal electrical charges do not flow freely and that can be polarized when an electric field is applied. Absolute permittivity is a measure of how an electric field affects and is affected by a dielectric medium. In the simple case the dielectric can be reduced down to an electric dipole, a molecule or atom where the positive and negative electric charges are separated by a small distance, in Figure 2.1 a schematic of a hydrogen fluoride molecule is shown.

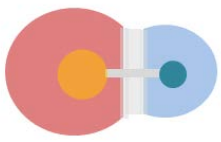

Figure 2.1: Schematic description of an electric dipole, in this case a hydrogen fluoride molecule. A separation of charge is present with negative charge to the right (red shade), and positive charge to the left (blue shade).

If an electric field were to be applied to a dielectric, the electric charges in the material would shift from their average equilibrium position causing dielectric polarization. An easy way to visualize this is by using the classical approach to the dielectric model. That is, each material is made up of atoms and each atom has a positive charge at its core. The positive charge is surrounded by a negatively charged cloud consisting of electrons. By applying an electric field the negatively charged cloud will be temporarily distorted as long as the field is applied, see Figure 2.2.

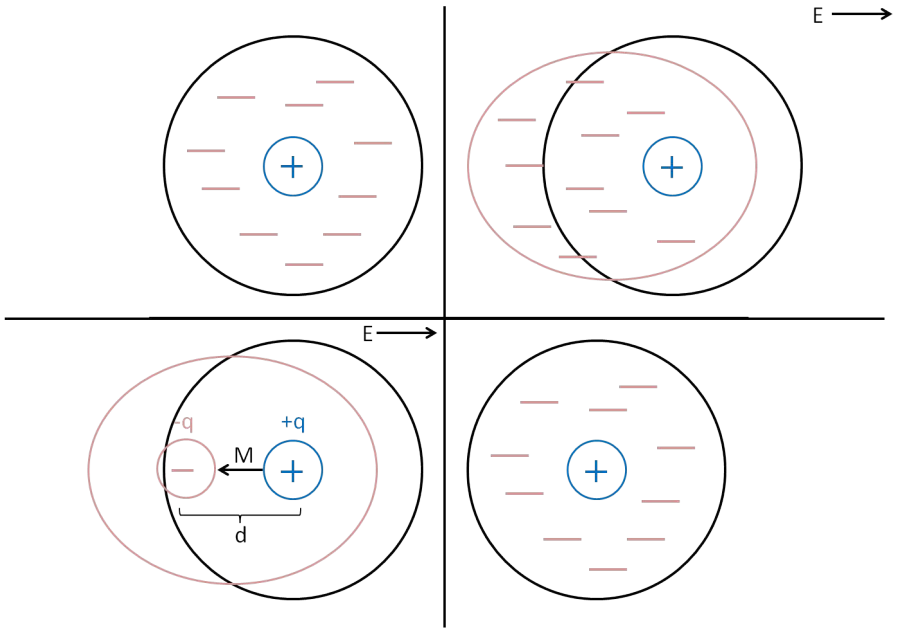

Figure 2.2: Schematic description of the interaction between an applied electric field and an atom.

The Electrical dipole momentum, represented by M in Figure 2.2, is a measure of the systems overall polarity, i.e. the separation of the positive $(+q)$ and negative $(-q)$ electrical charges. The applied electric field is represented by E. It is the relationship between the electrical dipole moment and the electric field that is responsible for the behavior of the dielectric. If the electric field is removed the atom will return to its original state, however a momentary lag will by percent due to the delay in molecular polarization with respect to a changing electric field in a dielectric medium, this lag is call relaxation time. If a dipole, see Figure 2.1, is placed in an electric field the dielectric polarization will cause the positive and negative charges to rotate and align with the field, see Figure 2.3.

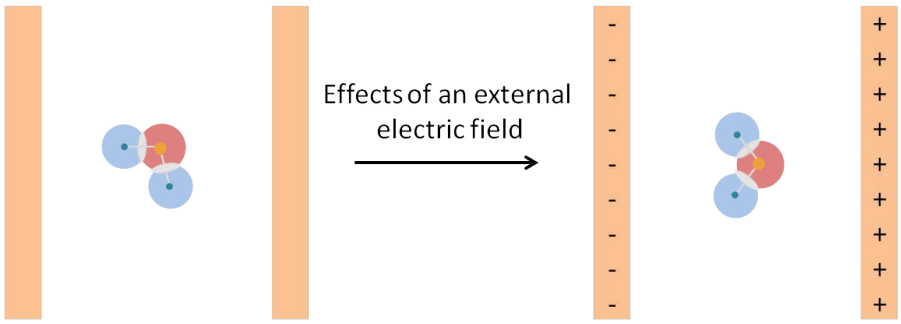

Figure 2.3: Schematic of the polarization effects when an electric field is applied to a dipole.

Similarly the same thing will happen when considering a material consisting of more than one single dipole.

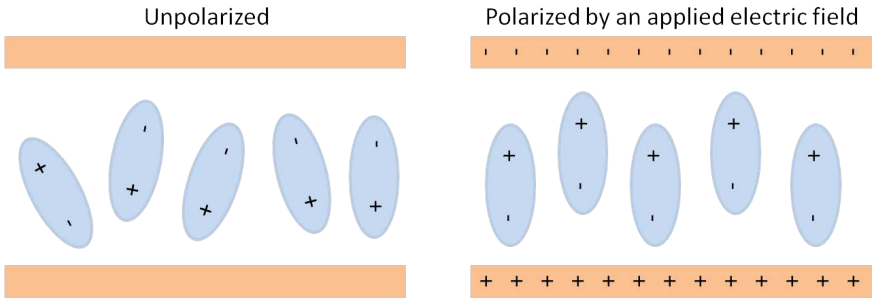

Figure 2.4: Schematic of the polarization effects when an electric field is applied to a dielectric.

The shift of positive charges will be towards the field and the negative charges will shift in the opposite direction. This shift of the charged particles will also create an internal electric field that reduces the overall field within the dielectric. The dielectric medium both affect and is affected by the electric field and a measure of these affects is the absolute permittivity.

As shown, the presence of an electric field in a dielectric material causes the bound charges in the material to slightly separate, see Figure 2.2. The electric displacement field, $\mathbf{D}$, describes how the presence of an electric field, $\mathbf{E}$, affects the orientation of the electric charges and is defined as

$$\mathbf{D} \equiv \varepsilon_0 \mathbf{E} + \mathbf{P}_d, \tag{2.1}$$

where $\mathbf{P}_d$ is the polarization density, an entity that describes the induced dipole moment in the material. The constant ε_0 is the permittivity in vacuum and is defined as

$$\varepsilon_0 \overset{def}{=} \frac{1}{c_0^2 \mu_0} \approx 8.854 \cdot 10^{-12} \qquad [F/m], \tag{2.2}$$

where c_0 is the speed of light in vacuum and μ_0 is the permeability in vacuum, i.e. the ability of the material to support the formation of a magnetic field within itself. In a dielectric that is linear, isotropic, nondispersive, and homogeneous, and respond instantaneous to a change in the electric field, the polarization density depends linearly on the electric field,

$$\mathbf{P}_d = \varepsilon_0 \chi \mathbf{E}. \tag{2.3}$$

Here χ is a dimensionless constant called electric susceptibility and indicates the dielectrics degree of polarization in response to an applied electric field. The electric susceptibility is directly related to the relative permittivity, $\varepsilon_r = \varepsilon/\varepsilon_0$, by

$$\chi = \varepsilon_r - 1 = \frac{\varepsilon}{\varepsilon_0} - 1. \tag{2.4}$$

Hence, the electric displacement field can be expressed as

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}_d = \varepsilon_0 \mathbf{E} + \varepsilon_0 \chi \mathbf{E} = \varepsilon_0 \mathbf{E}\,(1 + \chi) = \varepsilon_0 \mathbf{E} \left(1 + \frac{\varepsilon}{\varepsilon_0} - 1\right) = \varepsilon \mathbf{E}. \tag{2.5}$$

In dispersive dielectric mediums the polarization of the charged particles, due to an applied electric field, cannot happen immediately. Hence the relation between $\boldsymbol{P}_d$ and $\boldsymbol{E}$ is dynamic. If an arbitrary electric field is applied at time $t = 0$, the polarization density can be expressed as a superposition of the effects of $\boldsymbol{E}(t')$ for all $t' \leq t$. That is

$$\mathbf{P}_d(t) = \varepsilon_0 \int_{-\infty}^{t} \chi(t - t') \mathbf{E}(t') dt'. \tag{2.6}$$

The polarization density is time dependent and obtained by calculating the convolution of the impulse response susceptibility and the electric field. The convolution theorem, [16, p.1129], makes it possible to move a convolution in time domain to a point wise product of Fourier transforms in frequency domain it is therefore convenient to take the Fourier transform and write the function as a function of frequency as.

$$\mathbf{P}_d(\omega) = \varepsilon_0 \chi(\omega) \mathbf{E}(\omega), \tag{2.7}$$

where ω in this case is the angular frequency. Since the susceptibility, χ, now is frequency dependent, the permittivity will also be frequency dependent, see (2.4). Usually permittivity is treated as frequency dependent since the response of materials to electric fields in general depend on the frequency of the applied field, an exception being vacuum permittivity which is a constant.

This frequency dependence arrive from the fact that the polarization of the material does not respond immediately to the applied electric field, the response to the field

most be casual. This causality can be represented by a phase difference. Since complex numbers allow specification of magnitude and phase, the permittivity is usually treated as a complex function dependent on angular frequency. This allows permittivity to be defined as

$$D_0 e^{-j\omega t} = \widehat{\varepsilon}(\omega) E_0 e^{-j\omega t}, \tag{2.8}$$

where D_0 and E_0 are the amplitudes of the displacement and electric fields, respectively and j is the imaginary unit, $j^2 = -1$.

Since the permittivity is represented by a complex function, it is natural to separate the real and imaginary part as

$$\widehat{\varepsilon}(\omega) = \varepsilon'(\omega) + j\varepsilon''(\omega), \tag{2.9}$$

where the real part, $\varepsilon'(\omega)$, is related to the energy stored within the medium and the imaginary part, $\varepsilon''(\omega)$, is related to the loss of energy within the medium.

In some mediums, often called lossy mediums, a significant amount of the energy is absorbed. The reason for this is that the medium have free electric charges and an associated electric current density, $\boldsymbol{J}_d$, that has to be taken into account. By including $\boldsymbol{J}_d$ along with the displacement current density, $i\omega\boldsymbol{D}$, the total current density is given by

$$\boldsymbol{J}_{tot} = j\omega\boldsymbol{D} + \boldsymbol{J}_d. \tag{2.10}$$

For dielectric mediums with conductivity σ and linear conductive and dielectric properties, i.e. $\boldsymbol{D} = \varepsilon\boldsymbol{E}$, the electrical current density is proportional to the electric field

$$\boldsymbol{J}_d = \sigma\boldsymbol{E}, \tag{2.11}$$

which is a form of Ohm's law, [16, p.758]. Combining (2.10) with (2.11) gives

$$\boldsymbol{J}_{tot} = j\omega\boldsymbol{D} + \sigma\boldsymbol{E} = j\omega(\varepsilon + \frac{\sigma}{j\omega})\boldsymbol{E}. \tag{2.12}$$

That is, if the medium is lossy the extra term $\frac{\sigma}{j\boldsymbol{\omega}}$ needs to be taken into account. It should be noted that the term varies inversely with frequency and hence the contribution of the conductive component decreases as the frequency increases.

The electromagnetic energy absorbed by dielectrics does in general depend on a few different mechanisms which in turn affect shape of the permittivity spectrum. There are resonance effects that can arise from the vibration or rotation of the electrons, atoms or ions. These can be observed in the area around their absorption frequencies. There are also relaxation effects. The field changes slowly at low-frequencies and this allows the dipoles to reach equilibrium before the field has a measurable change. At higher frequencies and due to the viscosity of the medium, the dipoles cannot always follow the changes in the allied field and an energy loss will occur due to the absorption of the

fields energy. For ideal dipoles the dielectric relaxation can be described by the Debye relaxation model, [19].

The Debye equation is used to describe the relaxation response of a group of ideal noninteracting dipoles to an applied altering electric field. The Debye relaxation equation given below describes the complex relative permittivity of the medium as a function of the frequency, ω, of the applied electric field

$$\widehat{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + j\omega\tau}, \tag{2.13}$$

where ω is the angular frequency, $\varepsilon_s = \lim_{\omega\to 0} \widehat{\varepsilon}(\omega)$ is the static permittivity, $\varepsilon_\infty = \lim_{\omega\to\infty} \widehat{\varepsilon}(\omega)$ is the high-frequency permittivity and τ is the relaxation time of the dipole.

In 1941 the brothers Kenneth and Robert Cole developed a variant of the Debye equation. It is

$$\widehat{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + (j\omega\tau)^{1-\alpha}} + j\frac{\sigma}{\omega\varepsilon_0}, \tag{2.14}$$

which is called the Cole-Cole equation, [1]. The last term in 2.14 is added in the case of a lossy medium. The main differences between the models is that the Cole-Cole model uses a relaxation time *distribution* and that the parameter τ here is the central relaxation time. The model includes an exponent $1 - \alpha$, which describes the broadness of the relaxation time distribution. It is bounded by $0 \leq \alpha \leq 1$. When the exponent becomes smaller the relaxation time distribution becomes broader. This means that the transition between the low- and the high-frequency values becomes wider, and the peak on the imaginary part of the spectrum becomes wider. The Cole-Cole model can be viewed as a superposition of multiple Debye models. The central relaxation time is an inverse of the frequency corresponding to peak position on the imaginary part of spectrum.

It should be noted that the model assumes that the dielectric dispersion curve is symmetric and that both the Debye and the Cole-Cole equation given here describes the complex *relative* permittivity, $\varepsilon_r = \varepsilon/\varepsilon_0$. The models are based on relative permittivity because it is usually what is measured.

Measurements of the relative permittivity can be obtained by measuring the capacitance between two plates separated by a distance l. By first measuring the capacitance with vacuum in between the plates, C_0, and then measure the capacitance with the dielectric in between the plates, C_l. The quota between these two measurements will be the same as the quota between the absolute permittivity and the permittivity in vacuum, that is, the relative permittivity.

$$\frac{C_l}{C_0} = \frac{\varepsilon}{\varepsilon_0} = \varepsilon_r. \tag{2.15}$$

2.2 The method of least squares

When fitting a model to data points a standard approach is the method of least squares [17, ch.10]. This method finds an approximate solution to a over determined system and does so by finding the unknown parameters that minimizes the sum of squared errors. Least squares can fall under two categories

- Linear least squares,
- Nonlinear least squares,

depending on whether the model is linear or nonlinear in unknown parameters. For both the cases the overall objective of the least squares method is to find the unknown parameters $\boldsymbol{p}$ so that the model $\boldsymbol{f}(\boldsymbol{p};\boldsymbol{x})$, where $\boldsymbol{x}$ is an independent variable, best fit the observed data $\boldsymbol{y}$ in a least square sense. Here it will be assumed that $\boldsymbol{x} = [x_1, ..., x_m]$ and $\boldsymbol{y} = [y_1, ..., y_m]$ are vectors. The vector $\boldsymbol{p} = [p_1, ..., p_n]$, contains the adjustable parameters and the subscripts n and m satisfy $n \leq m$.

As stated above the least squares method assumes that the best fit of the model to the data is found when the sum,

$$S(\boldsymbol{p}) = \frac{1}{2}\sum_{i=1}^{m} r_i(\boldsymbol{p})^2 = \frac{1}{2}||\boldsymbol{r}(\boldsymbol{p})||_2^2, \tag{2.16}$$

of the squared residuals is minimized. Here $||\cdot||$ is the Euclidean norm and the residual, $r_i(\boldsymbol{p}) = f(\boldsymbol{p}; x_i) - y_i$, is the difference between the model at the points x_i and the measured spectrum. It is clear that both S and $\boldsymbol{r}$ are functions of $\boldsymbol{p}$ although from now on we will suppress that in the notation for the sake of clarity.

The minimum of the sum S in (2.16) is found when the gradient vector is zero, if it is not zero then there is a direction in which we can move to minimize it further. By differentiating S with respect to the unknown parameters $\boldsymbol{p}$ and setting the result to zero, the minimum of S can be found.

Linear least squares, [17, ch.10.2], is a least squares method used for fitting data to a model which is linear in the unknown parameters. If the function, $\boldsymbol{f}(\boldsymbol{p};\boldsymbol{x})$, is linear then the residual vector can be written as $\boldsymbol{r} = \boldsymbol{X}\boldsymbol{p} - \boldsymbol{y}$, where $\boldsymbol{X}$ is a $m \times n$ matrix of linear equations. Hence, the linear least squares problem is

$$\min_{\boldsymbol{p}} \frac{1}{2}||\boldsymbol{X}\boldsymbol{p} - \boldsymbol{y}||. \tag{2.17}$$

The minimizer, $\boldsymbol{p}^*$, of (2.17) must satisfy

$$\left(\boldsymbol{X}^T\boldsymbol{X}\right)\boldsymbol{p}^* = \boldsymbol{X}^T\boldsymbol{y}, \tag{2.18}$$

which are known as the normal equations of (2.17).

Nonlinear least squares, [17, ch.10.3], fits data to models that are nonlinear in the unknown parameters. But since the model and hence also the residual is nonlinear the derivatives are functions of both the independent variable, $\boldsymbol{x}$, and the parameters, $\boldsymbol{p}$. This means that usually there is no closed solution to the gradient equations,

$$\frac{\partial S}{\partial p_j} = 0. \tag{2.19}$$

A common method used to solve nonlinear least squares problem is the Gauss-Newton method. The method can, given an initial guess, find the parameters by successive iterations. At each iteration the estimated parameters will be refined, accordingly

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)} + \boldsymbol{h}. \tag{2.20}$$

Here k represents the iteration number and $\boldsymbol{h}$ is the refinement of the parameter, i.e. the parameter change between the iteration k and $k+1$. From now on a quantity with superscript k or $k+1$ will denote that it is evaluated at $\boldsymbol{p}^{(k)}$ or $\boldsymbol{p}^{(k+1)}$ respectively.

The method will linearize the residual at each iteration by approximation to a first-order Taylor series expansion, [20, pp.52-53], about $\boldsymbol{p}^{(k)}$. That is

$$\boldsymbol{r}^{(k+1)} \approx \boldsymbol{r}^{(k)} + \boldsymbol{J}^{(k)}\boldsymbol{h}. \tag{2.21}$$

Here $\boldsymbol{J}^{(k)}$ is a $m \times n$ matrix of all the first order partial derivatives of $\boldsymbol{r}^{(k)}$, known as the Jacobian matrix. Note that since the matrix depends on $\boldsymbol{p}^{(k)}$ it will change from one iteration to another. Since the residual is linearized we have the following linear least squares problem

$$\min_{\boldsymbol{h}} \frac{1}{2}||\boldsymbol{J}^{(k)}\boldsymbol{h} + \boldsymbol{r}^{(k)}||. \tag{2.22}$$

The minimizer of (2.22) satisfies

$$\boldsymbol{J}^{(k)T}\boldsymbol{J}^{(k)}\boldsymbol{h} = -\boldsymbol{J}^{(k)T}\boldsymbol{r}^{(k)}. \tag{2.23}$$

Using the Gauss-Newton algorithm the sum S in (2.16) may not decrease at each iteration. Since the search direction $\boldsymbol{h}$ is a descent direction, it should be decreasing at each iteration, as long as $S(\boldsymbol{p}^{(k)})$ not is a stationary point. If a divergence would occur, a solution would be to only use a fraction, γ, of the increment vector, $\boldsymbol{h}$ in the updating formula, i.e.

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)} + \gamma\boldsymbol{h}. \tag{2.24}$$

The fraction γ, $0 < \gamma < 1$, is a scalar known as the step-length. It is useful to include it in the formula since the increment vector might be pointing in the right direction but it is too long, so by only go a part of the way will decrease S. An optimum value of the step-length γ can be found by using a line search algorithm. There are different methods that can be used for this, Wolfe conditions, [21], are common.

2.3 Method

Since the Cole-Cole equation, (2.14), is nonlinear in some of the unknown parameters the method of nonlinear least squares will be used for the curve fitting. Two different curve fitting techniques will be compared. One being the Levenberg-Marquardt method. The other method is the Variable projection method, which takes into account some of the characteristics of the Cole-Cole equation. Both the Levenberg-Marquardt and the variable projection method makes use of the Gauss-Newton algorithm derived in the previous section. It should be noted that quantities with subscript 1 or 2 denotes that the quantity is specific for the Levenberg-Marquardt and variable projection method respectively.

2.4 The Levenberg-Marquardt algorithm

The Levenberg-Marquardt method, [11] and [12], is a standard, iterative technique used to solve nonlinear least squares problems. The Levenberg-Marquardt method is similar to the Gauss-Newton method but uses a trust region strategy. Using a trust region strategy avoids one of the weaknesses of Gauss-Newton, namely, its behavior when the Jacobian $\boldsymbol{J}$ is rank deficient or ill-conditioned.

In line search methods a search direction $\boldsymbol{h}$ is calculated and then the focus is on finding a step-length γ along this direction. The trust region method, [17, ch.4], on the other hand defines a region around the current working point within which a model function is trusted to provide a reasonable approximation of the cost-function. In other words one can say that both the length and direction of the step is chosen simultaneously.

The size of the trust region is usually chosen depending on the performance of the algorithm during the previous iteration. The fact that the size of the trust region can vary helps increase the effectiveness of each step. This also mean that the choice of the size of the trust region is vital. If it is chosen poorly opportunities to take substantial steps, and hence move much closer to the minimizer of the cost-function, will be missed. Similarly if the trust region is too large then the minimizer of the model might be far from the minimizer of the cost function in that region and thus the size of the trust region has to be reduced and a new step has to be calculated.

The basic idea of the trust region approach is to accept the minimum of the model function as long as the model adequately reflects the behavior of the cost-function. The model function $\phi^{(k)}$, where k is the iteration index, used for the Levenberg-Marquardt method, [17, p.258] is given by the second order Taylor series expansion of S around the current working point $\boldsymbol{p}^k$

$$\phi^{(k)}(\boldsymbol{h}) = \frac{1}{2}||\boldsymbol{r}_1^{(k)}||_2^2 + \boldsymbol{g}_1^{(k)T}\boldsymbol{h} + \frac{1}{2}\boldsymbol{h}^T\boldsymbol{H}_1^{(k)}\boldsymbol{h}. \tag{2.25}$$

Here $\boldsymbol{g}_1^{(k)} = \boldsymbol{J}_1^{(k)T}\boldsymbol{r}_1^{(k)}$ is the gradient and $\boldsymbol{H}_1^{(k)}$ is the Hessian matrix which is approxi-

mated as $\boldsymbol{H}_1^{(k)} = \boldsymbol{J}_1^{(k)T}\boldsymbol{J}_1^{(k)}$. In order to minimize the model new search directions, $\boldsymbol{h}$, has to be calculated. The new steps are obtained by at each iteration solving the constrained subproblem,

$$\min_{\boldsymbol{h}\in\Re^m} \frac{1}{2}||\boldsymbol{J}_1^{(k)}\boldsymbol{h} + \boldsymbol{r}_1^{(k)}||_2^2 \qquad s.t.||\boldsymbol{h}||_2 \leq \Delta^{(k)}. \tag{2.26}$$

Here $\Delta^{(k)}$ is the radius of the trust region. It can be shown, [17, ch.4], that the solution, $\boldsymbol{h}^{(k)*}$, to

$$\left(\boldsymbol{J}_1^{(k)T}\boldsymbol{J}_1^{(k)} + \lambda^{(k)}I\right)\boldsymbol{h}^{(k)*} = -\boldsymbol{J}_1^{(k)T}\boldsymbol{r}_1^{(k)}, \tag{2.27}$$

where I is the identity matrix and $\lambda^{(k)}$ is a scalar such that the matrix $\boldsymbol{J}_1^{(k)T}\boldsymbol{J}_1^{(k)} + \lambda^{(k)}I$ is positive semi-definite, solves the subproblem (2.26) if either $||\boldsymbol{h}||_2 = \Delta$ and $\lambda \geq 0$ or $||\boldsymbol{h}||_2 \leq \Delta$ and $\lambda = 0$. This method is known as the Levenberg-Marquardt method.

It should be noted that if Δ is large enough the solution to (2.27) is found when $\lambda = 0$. That is, (2.27) becomes the Gauss-Newton algorithm, see (2.23). It should also be noted that when $\lambda \to +\infty$, $||\boldsymbol{h}||_2 \to 0$ and $\boldsymbol{h}$ becomes parallel to the search direction of the steepest descent method, [20, pp.102-103]. That is, it approaches the direction of the negative gradient

$$\boldsymbol{h}^{(k)*} = -\boldsymbol{J}_1^{(k)T}\boldsymbol{r}_1^{(k)}. \tag{2.28}$$

An updated relationship where the identity matrix I in (2.27) is replaced with the diagonal elements of the $\boldsymbol{J}_1^{(k)T}\boldsymbol{J}_1^{(k)}$ matrix was suggested by Marquardt, [12]. Using this relationship the search direction is computed by solving

$$\left(\boldsymbol{J}_1^{(k)T}\boldsymbol{J}_1^{(k)} + \lambda^{(k)}\text{diag}(\boldsymbol{J}_1^{(k)T}\boldsymbol{J}_1^{(k)})\right)\boldsymbol{h}^{(k)*} = \boldsymbol{J}_1^{(k)T}\boldsymbol{r}. \tag{2.29}$$

The advantage of this is that each component of the gradient will now be scaled according to the curvature. Hence, there will now be larger movement along the directions where the gradient is smaller. This avoids slow convergence in the direction of small gradient.

A good value of the trust region radius, Δ, and the damping parameter λ needs to be chosen in order to ensure descent. Since the two parameters are related, choosing one is equivalent to choosing the other. The choice can either be done by adjusting the damping parameter directly, or, by first choosing an acceptable step size and then finding a λ such that $||\boldsymbol{h}|| \leq \Delta$. Here the focus will be on λ. There are many different methods available that can be used for this, but usually it is adequate to use the simple method suggested by Marquardt, in which λ is either increased by a fixed factor λ_{up} or decreased by a fixed factor λ_{dn}. Different sets of values will be more suitable for different types of problems, but it is often sufficient to use $\lambda_{up} = \lambda_{dn} = 10$. Whether or not λ will be increased or decreased depends in turn on the change between the model function and the cost-function. Given the current and previous step a scalar ρ, [17, p.68], is defined as

$$\rho^{(k)} = \frac{S^{(k)} - S^{(k+1)}}{\phi^{(k)}(0) - \phi^{(k)}(\boldsymbol{h})}. \tag{2.30}$$

Depending on the value of ρ there are a few things that should be noted. The predicted reduction should always be nonnegative since the step $\boldsymbol{h}$ is obtained by minimizing $\phi^{(k)}$ over a region that includes $\boldsymbol{h} = \boldsymbol{0}$. Thus, if $\rho^{(k)}$ is negative, that means that the cost-function is increasing, $S^{(k+1)} > S^{(k)}$, and the step most be rejected. If, on the other hand, $\rho^{(k)}$ is larger than some threshold $\nu > 0$ it means that the model at the current step is consistent with the function and hence the trust region can be expanded for the next iteration. Similarly if $\rho^{(k)}$ is smaller than the threshold it means that the model is not consistent with the function at the current step and the trust region is reduced to the next iteration.

To summarize

- if $\rho < \nu$: $\lambda = \lambda/\lambda_{down}$ and $\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)}$,
- if $\rho > \nu$: $\lambda = \lambda \cdot \lambda_{up}$ and $\boldsymbol{p}^{(k+1)} = \boldsymbol{p}^{(k)} + \boldsymbol{h}$.

2.5 The variable projection method

The variable projection method, [14], is a method used to solve separable nonlinear least squares problems. Consider a function, $f(\boldsymbol{p}; \boldsymbol{x})$, that one wants to fit to observations, $\boldsymbol{y} = [y_1, ..., y_m]^T$. The model function contains an independent variable $\boldsymbol{x} = [x_1, ..., x_m]^T$, which usually represents time or frequency, and the model parameters $\boldsymbol{p}$. The least squares problem is said to be separable when the model parameters can be separated into two sets of parameters, one that enter nonlinearly into the model, $\boldsymbol{b} = [b_1, ..., b_{n_2}]^T$, and another set of parameters that enter the model linearly, $\boldsymbol{a} = [a_1, ..., a_{n_1}]^T$, i.e. $\boldsymbol{p} = [\boldsymbol{a}, \boldsymbol{b}]$. This fitting problem can be written as

$$\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{b})\boldsymbol{a}, \tag{2.31}$$

where $\boldsymbol{A}$ is called the model matrix and is a function of $\boldsymbol{b}$ and $\boldsymbol{x}$. The fitting problem is a nonlinear least squares problem and the goal is to find the values of the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ by solving

$$\min_{\boldsymbol{a},\boldsymbol{b}} \frac{1}{2} ||\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{b})\boldsymbol{a}||_2^2 . \tag{2.32}$$

From (2.32) it is clear that, given a value of $\boldsymbol{b}$, the linear parameters $\boldsymbol{a}$ can be calculated using the linear least squares, (2.18). The estimate of the linear parameter is thus given by

$$\widehat{\boldsymbol{a}} = (\boldsymbol{A}^T(\boldsymbol{b})\boldsymbol{A}(\boldsymbol{b}))^{-1}\boldsymbol{A}^T(\boldsymbol{b})\boldsymbol{y} = \boldsymbol{A}^\dagger(\boldsymbol{b})\boldsymbol{y}, \tag{2.33}$$

where $\boldsymbol{A}^{\dagger}(\boldsymbol{b})$ denotes the Moore-Penrose generalized inverse of $\boldsymbol{A}(\boldsymbol{b})$ which will be used henceforth for notational simplicity. It will from here on, also be assumed that the model matrix $\boldsymbol{A}$ is a full rank $\Re^{m\times n_1}$ matrix, where $m > n_1$, for all $\boldsymbol{b}$.

Let us now define a projection matrix $\boldsymbol{P}(\boldsymbol{b})$ and a projection matrix $\boldsymbol{P}^{\perp}(\boldsymbol{b})$ that projects a vector onto the range of $\boldsymbol{A}(\boldsymbol{b})$ and onto the null space of $\boldsymbol{A}^{T}(\boldsymbol{b})$ respectively. The projection matrix $\boldsymbol{P}(\boldsymbol{b})$ will be defined as

$$\boldsymbol{P}(\boldsymbol{b}) = \boldsymbol{A}(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}). \tag{2.34}$$

It should be noted that the projection matrix is both idempotent, $\boldsymbol{P}^2(\boldsymbol{b}) = \boldsymbol{P}(\boldsymbol{b})$, and symmetric, $\boldsymbol{P}^{T}(\boldsymbol{b}) = \boldsymbol{P}(\boldsymbol{b})$. The projection matrix $\boldsymbol{P}^{\perp}(\boldsymbol{b})$ will be defined as

$$\boldsymbol{P}^{\perp}(\boldsymbol{b}) = \boldsymbol{I} - \boldsymbol{P}(\boldsymbol{b}). \tag{2.35}$$

By using the expression for the estimate of the linear parameters, $\widehat{\boldsymbol{a}}$, the residual, $\boldsymbol{r}_2 = \boldsymbol{y} - \boldsymbol{A}(\boldsymbol{b})\boldsymbol{a}$, can be rewritten using the projection matrix, $\boldsymbol{P}^{\perp}(\boldsymbol{b})$, as

$$\boldsymbol{r}_2 = \boldsymbol{y} - \boldsymbol{A}(\boldsymbol{b})\widehat{\boldsymbol{a}} = \boldsymbol{y} - \boldsymbol{A}(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b})\boldsymbol{y} = \boldsymbol{y} - \boldsymbol{P}(\boldsymbol{b})\boldsymbol{y} = \left(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{b})\right)\boldsymbol{y} = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{y}. \tag{2.36}$$

As seen from the equation above the residual becomes a function of $\boldsymbol{b}$, $\boldsymbol{r}_2 = \boldsymbol{r}_2(\boldsymbol{b}) = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{y}$. Combining this with (2.32) results in

$$\min_{\boldsymbol{b}} \frac{1}{2} \left|\left|\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{y}\right|\right|_2^2. \tag{2.37}$$

The linear parameters has been removed and we are left with a nonlinear least squares problem that needs to be solved in order to find $\boldsymbol{b}$. Introducing the function $q(\boldsymbol{b}) = \frac{1}{2}\boldsymbol{r}_2^T\boldsymbol{r}_2$ the original problem (2.32) can be separated into two least squares problem

$$\min_{\boldsymbol{b}} q(\boldsymbol{b}), \tag{2.38}$$

$$\min_{\boldsymbol{a}} \frac{1}{2} \left|\left|\boldsymbol{y} - \boldsymbol{A}(\widehat{\boldsymbol{b}})\boldsymbol{a}\right|\right|_2^2, \tag{2.39}$$

where (2.38) is a nonlinear least squares problem whose solution, $\widehat{\boldsymbol{b}}$, is needed as input to the linear least squares problem (2.39). The technique developed by Golub and Pereyra, [14], a useful for solving this problem. The technique is based on applying Gauss-Newton method, (2.23) - (2.24), to the variable projection functional, which is in essence the squared norm of the residual, $\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{y}$.

The developed method solves (2.38) by iteratively updating $\boldsymbol{b}$ as

$$\boldsymbol{b}^{(k)} = \boldsymbol{b}^{(k-1)} - \gamma \left(\boldsymbol{H}_2^{(k)}\right)^{-1} \boldsymbol{g}_2^{(k)}, \tag{2.40}$$

where k is the iteration index, γ is the step length, $\boldsymbol{H}_2$ and $\boldsymbol{g}_2$ are the Hessian matrix and the gradient of $q(\boldsymbol{b})$ respectively. Both the vector gradient, defined as $\boldsymbol{g}_2 = \boldsymbol{J}_2\boldsymbol{r}_2(\boldsymbol{b})$,

and the Hessian matrix, approximated as $\boldsymbol{H}_2 = \boldsymbol{J}_2^T\boldsymbol{J}_2$, are evaluated at $\boldsymbol{b}^{(k-1)}$. Here $\boldsymbol{J}_2 = \nabla \boldsymbol{r}_2(\boldsymbol{b})$ represents the Jacobian matrix

As can be seen above, the Jacobian matrix is needed in order to be able to calculate the next step $\boldsymbol{b}^{(k)}$. Knowing that the Jacobian matrix is defined as $\boldsymbol{J}_2 = \nabla \boldsymbol{r}_2$ and that the residual is given by $\boldsymbol{r}_2 = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{y}$, see (2.36), we need to calculate the first partial derivative of $\boldsymbol{P}^{\perp}(\boldsymbol{b})$ in order to calculate the next step, $\boldsymbol{b}^{(k)}$.

The partial derivatives can be obtained by following the derivation done in [14]. Since $\boldsymbol{P}^{\perp}(\boldsymbol{b}) = \boldsymbol{I} - \boldsymbol{P}(\boldsymbol{b})$, see (2.35), the derivative of $\boldsymbol{P}^{\perp}(\boldsymbol{b})$ can be found be finding the derivative of $\boldsymbol{P}(\boldsymbol{b})$. From now on the subscript i, $1 < i < n_2$, will be used to denote the differentiation of a matrix with respect to the variable $\boldsymbol{b}_i$. By taking into account that the projection matrix is idempotent, $\boldsymbol{P}^2(\boldsymbol{b}) = \boldsymbol{P}(\boldsymbol{b})$, differentiation yields

$$\boldsymbol{P}_i(\boldsymbol{b}) = \boldsymbol{P}_i(\boldsymbol{b})\boldsymbol{P}(\boldsymbol{b}) + \boldsymbol{P}(\boldsymbol{b})\boldsymbol{P}_i(\boldsymbol{b}). \tag{2.41}$$

The next step is to note that the projection matrix satisfies

$$\boldsymbol{P}(\boldsymbol{b})\boldsymbol{A}(\boldsymbol{b}) = \boldsymbol{A}(\boldsymbol{b}). \tag{2.42}$$

By differentiating both sides with respect to $\boldsymbol{b}_i$ the following expression is obtained

$$\boldsymbol{P}_i(\boldsymbol{b})\boldsymbol{A}(\boldsymbol{b}) + \boldsymbol{P}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b}) = \boldsymbol{A}_i(\boldsymbol{b}). \tag{2.43}$$

Subtracting $\boldsymbol{P}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})$ from both sides of the equation yields

$$\boldsymbol{P}_i(\boldsymbol{b})\boldsymbol{A}(\boldsymbol{b}) = \boldsymbol{A}_i(\boldsymbol{b}) - \boldsymbol{P}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b}) = (\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{b}))\boldsymbol{A}_i(\boldsymbol{b}) = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b}). \tag{2.44}$$

Right multiplication with $\boldsymbol{A}^{\dagger}(\boldsymbol{b})$ and applying the definition of the projection matrix given by (2.34) gives

$$\boldsymbol{P}_i(\boldsymbol{b})\boldsymbol{P}(\boldsymbol{b}) = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}). \tag{2.45}$$

By transposing and exploiting the symmetry, $\boldsymbol{P}^T(\boldsymbol{b}) = \boldsymbol{P}(\boldsymbol{b})$, on the left yields

$$\boldsymbol{P}(\boldsymbol{b})\boldsymbol{P}_i(\boldsymbol{b}) = (\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}))^T. \tag{2.46}$$

Combining the results given from (2.45) and (2.46) with (2.41) gives that

$$\boldsymbol{P}_i(\boldsymbol{b}) = \boldsymbol{P}_i(\boldsymbol{b})\boldsymbol{P}(\boldsymbol{b}) + \boldsymbol{P}(\boldsymbol{b})\boldsymbol{P}_i(\boldsymbol{b}) = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}) + (\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}))^T. \tag{2.47}$$

Using the result given by the equation above together with (2.35) gives that the partial derivative of the projection matrix, $\boldsymbol{P}^{\dagger}(\boldsymbol{b})$, becomes

$$\boldsymbol{P}_i^{\perp}(\boldsymbol{b}) = -\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}) - (\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{A}^{\dagger}(\boldsymbol{b}))^T. \tag{2.48}$$

Knowing that the Jacobian matrix is given by $\boldsymbol{J}_2 = \nabla \boldsymbol{r}_2(\boldsymbol{b}) = \boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{y}$ and applying (2.33), its jth column becomes

$$\boldsymbol{J}_{2,i} = -\boldsymbol{P}^{\perp}(\boldsymbol{b})\boldsymbol{A}_i(\boldsymbol{b})\boldsymbol{a} - (\boldsymbol{A}^{\dagger}(\boldsymbol{b}))^T \boldsymbol{A}_i(\boldsymbol{b})^T \boldsymbol{r}_2(\boldsymbol{b}). \tag{2.49}$$

Some more algorithmic details about the variable projection method can be seen in [15] and [22].

CHAPTER 3

Numerical implementation

This section covers some of the implementation of the Levenberg-Marquardt and the variable projection method. Both methods were implemented specifically to fit the Cole-Cole model, (2.14), to measured relative permittivity spectrum. In particular the calculation of the Jacobian matrix, the selection of the initial guesses and the convergence criterion will be covered. All code has been written in MATLAB and as in the previous section, quantities with subscript 1 or 2 denotes that they are specific for the Levenberg-Marquardt and variable projection method respectively.

3.1 Calculation of the Jacobian matrix

Both the Levenberg-Marquardt and the variable projection method uses the Jacobian matrix, $\boldsymbol{J}$, in order to calculate the search direction, $\boldsymbol{h}$. But the calculation of the matrix differ between the methods.

3.1.1 The Jacobian matrix for the Levenberg-Marquardt method

For the Levenberg-Marquardt method the calculation is quite straight forward. From Section 2.2 it is given that the Jacobian matrix is a matrix of all the first order partial derivatives of the residual $\boldsymbol{r}_1 = \boldsymbol{f}(\boldsymbol{p}; \boldsymbol{x}) - \boldsymbol{y}$. Since only the model depends on $\boldsymbol{p}$ the Jacobian matrix can be calculated by using the current estimate of the parameters and the partial derivatives of the function $\boldsymbol{f}(\boldsymbol{p}, \boldsymbol{x})$. For the Cole-Cole equation, (2.14), the analytical expression for the partial derivatives is

$$\frac{\partial \widehat{\varepsilon}}{\partial \varepsilon_\infty} = \frac{j\tau\omega}{(j\tau\omega)^\alpha + j\tau\omega}, \tag{3.1}$$

$$\frac{\partial \widehat{\varepsilon}}{\partial \varepsilon_s} = \frac{1}{1 + (j\tau\omega)^{(1-\alpha)}}, \tag{3.2}$$

$$\frac{\partial \widehat{\varepsilon}}{\partial \tau} = j\omega \frac{\alpha - 1}{\left((j\tau\omega)^{(1-\alpha)} + 1\right)^2} \frac{\varepsilon_s - \varepsilon_\infty}{(j\tau\omega)^\alpha}, \tag{3.3}$$

$$\frac{\partial \widehat{\varepsilon}}{\partial \alpha} = \frac{lnj\tau\omega}{\left((j\tau\omega)^{(1-\alpha)} + 1\right)^2} (\varepsilon_s - \varepsilon_\infty)(j\tau\omega)^{(1-\alpha)}, \tag{3.4}$$

$$\frac{\partial \widehat{\varepsilon}}{\partial \sigma} = -\frac{j}{\omega\varepsilon_0}. \tag{3.5}$$

It should be noted that the Jacobian matrix in this case is a matrix of size $m \times 5$, where m is the number of data points.

3.1.2 The Jacobian matrix for the variable projection method

For the variable projection method the calculation of the Jacobian matrix will be somewhat different. As stated in Section 2.5 the method divides the parameter, $\boldsymbol{p}$ into two separate sets of parameters, $\boldsymbol{a}$ and $\boldsymbol{b}$. Where $\boldsymbol{a}$ are parameters that enter the model linearly and $\boldsymbol{b}$ are parameters that enter the model nonlinearly. This results in the fitting problem given by (2.31).

Applying this to the Cole-Cole equation results in the following factorization of the model

$$f(\boldsymbol{a}, \boldsymbol{b}; \omega) = \underbrace{\begin{bmatrix} \left(1 - \frac{1}{1 + (j\omega\tau)^{1-\alpha}}\right) & \left(\frac{1}{1 + (j\omega\tau)^{1-\alpha}}\right) & \frac{-j}{\omega\epsilon_0} \end{bmatrix}}_{\mathbf{A}(\boldsymbol{b})} \underbrace{\begin{bmatrix} \epsilon_\infty \\ \epsilon_s \\ \sigma \end{bmatrix}}_{\boldsymbol{a}}, \tag{3.6}$$

where $\boldsymbol{b} = [\tau, \alpha]$. The linear parameters in $\boldsymbol{a}$ can be calculated given a value of the nonlinear parameters in $\boldsymbol{b}$ according to (2.33). Due to this, the problem has been reduced from a five dimensional to a two dimensional problem. In (2.49) it is shown how the ith column of the Jacobian matrix can be calculated. The only unknown is the term A_i, which is the partial derivative of $A(\boldsymbol{b})$, see (3.6), with respect to the parameter b_i. Since $\boldsymbol{b} = [b_1, b_2] = [\tau, \alpha]$, the derivative of $A(\boldsymbol{b})$ with respect to τ and α is needed. Differentiating $A(\boldsymbol{b})$ with respect, τ, gives the analytical expression

$$\frac{\partial A(\boldsymbol{b})}{\partial \tau} = \begin{bmatrix} -\dfrac{j(\alpha-1)\omega(j\tau\omega)^{\alpha}}{\left((j\tau\omega)^{\alpha}+j\tau\omega\right)^{2}} \\ \dfrac{j(\alpha-1)\omega(j\tau\omega)^{\alpha}}{\left((j\tau\omega)^{\alpha}+j\tau\omega\right)^{2}} \\ 0 \end{bmatrix}, \tag{3.7}$$

whereas differentiating with respect to α gives

$$\frac{\partial A(\boldsymbol{b})}{\partial \alpha} = \begin{bmatrix} -\dfrac{(j\tau\omega)^{\alpha+1} ln(j\tau\omega)}{\left((j\tau\omega)^{\alpha}+j\tau\omega\right)^{2}} \\ \dfrac{(j\tau\omega)^{\alpha+1} ln(j\tau\omega)}{\left((j\tau\omega)^{\alpha}+j\tau\omega\right)^{2}} \\ 0 \end{bmatrix}. \tag{3.8}$$

The current estimate of the parameters τ and α are used as input.

3.2 The initial guess

Since both methods are iterative, initial guesses, $\boldsymbol{p}_0$, will be needed in order for the iterations to start. The importance of the choice of the initial guess differ depending on the characteristics of the problem. If it is known that the cost-function only has one minima a uniform initial guess like $\boldsymbol{p}_0 = [1, ..., 1]$ will work fine. The algorithm will refine the initial guess at each iteration and eventually stop when a convergence criterion is met. A example of this is shown in Figure 3.1. The function in this example is only dependent on one parameter.

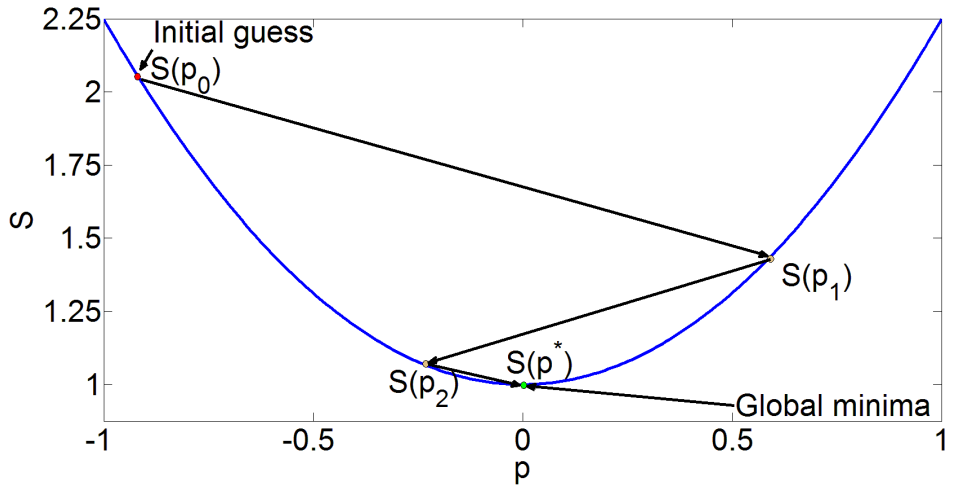

Figure 3.1: Example of initial guess when cost-function only has one minima and the function is only dependant on one parameter.

But using a uniform initial guesses in cases where the cost-function has multiple minima there is a risk that the found minima is only a local and not the global minima. Hence, the correct model parameter $\boldsymbol{p}$ has not been found. An example of this is shown in Figure 3.2, one again the function is only dependant on one parameter.

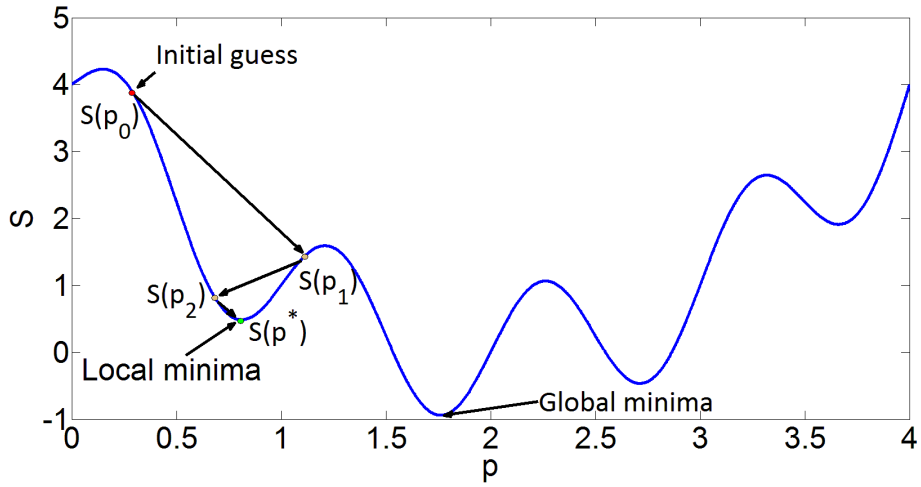

Figure 3.2: Example of initial guess when cost-function has multiple minima and the function is only dependant on one parameter.

As can be seen by the example given by Figure 3.2, good initial guesses are needed to ensure that the global minima is found.

3.2.1 Initial guess for the Levenberg-Marquardt method

The initial guess for the Levenberg-Marquardt method will be chosen by looking at the physical interpretation of the parameters in the Cole-Cole equation,

$$\widehat{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + (i\omega\tau)^{1-\alpha}} + j\frac{\sigma}{\omega\varepsilon_0}. \tag{3.9}$$

By definition it is known that the static permittivity, ε_s, is the permittivity when $\omega \to 0$. Using this, an initial guess $\varepsilon_{s,0}$ can be obtained by looking at the low frequency permittivity of the real part of the measured spectrum. Since measurements contain noise, an average is used to minimize the effects of the noise,

$$\varepsilon_{s,0} = \frac{1}{M}\sum_{i=1}^{M} \Re\left\{\varepsilon(\omega_i)\right\}, \tag{3.10}$$

where $\varepsilon(\omega_i)$ is the measured spectrum at frequency ω_i, $1 < i < M$. Here $M << m$ and m is as before the total number of measurement points.

Similarly, ε_∞ is the high frequency permittivity and is found when $\omega \to \infty$. The initial guess, $\varepsilon_{\infty,0}$, is hence found by taking the average of the high frequency permittivity of the real part of the measured spectrum.

The relaxation time, τ, is related to the (angular) frequency, ω, of the applied electric field. The initial guess, τ_0, can be found by finding at which frequency the imaginary part permittivity spectrum has its maximum value. This frequency is then used to calculate the initial estimate of the relaxation time according to

$$\tau_0 = \frac{j}{\omega_{I_{max}}}, \tag{3.11}$$

where $\omega_{I_{max}}$ here represents the frequency at the point where the imaginary part of the permittivity spectrum has its maximum value.

The parameter α represents the width of the dispersion area and is bounded between $[0, 1]$. The initial guess is chosen as the midpoint in this interval, $\alpha_0 = 0.5$.

Lastly we consider the initial guess for the conductivity parameter σ. In Section 2.1 it was show that the conductivity was part of the extra term, $j\sigma/\omega$, added in order to account for the case of lossy mediums. From this term it is seen that σ will affect the spectrum most at low frequencies. Hence, the initial guess for the conductivity will be chosen as

$$\sigma_0 = \omega_l\varepsilon_0\Im\left\{\varepsilon(\omega_l)\right\}, \tag{3.12}$$

where ω_l represents that the initial guess for the conductivity was calculated at low frequency. The permittivity in vacuum, $\varepsilon_0 \approx 8.854 \cdot 10^{-12}[F/m]$, is part of the expression since it is the relative permittivity that is used.

Figure 3.3 gives a graphical description of how the initial guess are chosen for the Levenberg-Marquardt method.

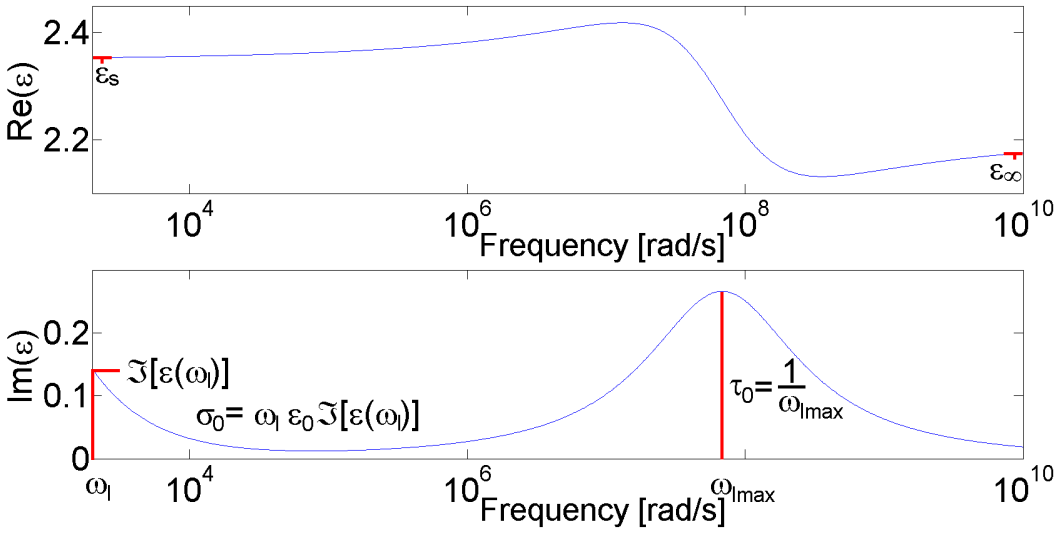

Figure 3.3: Graphical description of how the initial guess is found using the Levenberg-Marquardt method.

Note that the spectrum in Figure 3.3 was generated with the help of the Cole-Cole model and hence does not contain any noise.

3.2.2 Initial guess for the variable projection method

One of the big advantages with this method is that since the linear parameters can be calculated given a value of the nonlinear parameters no initial guesses are needed for the linear parameters. The nonlinear parameters do on the other hand need initial guesses and although the method described in the previous section can be used there is a more robust method for choosing the initial guesses.

By defining an interval of τ and combining it with the fact that α is bounded between 0 and 1, a grid search could be performed. The value of the parameters that results in the smallest value of the cost-function would be chosen as the initial guess. The computation time for the grid search will increase exponentially with the number of calculation points. Obtaining good starting guesses with this method can therefore be time consuming. The computation time can on the other hand be reduced to some extent by using a hierarchic grid search. This means that a coarse gird search will be used in order to find a general area within which the cost-function seems to have its minimum. Another gird search will then be performed on that area, see Figure 3.4 below.

In Figure 3.4 the number of segments the parameter interval is divided into is represented by Q and R. The advantage of a hierarchic grid search is that fever function evaluations has to be performed. If a hierarchic grid search is performed in two steps $2(Q \cdot R) - 1$ function evaluations is needed. To achieve the same accuracy on the whole

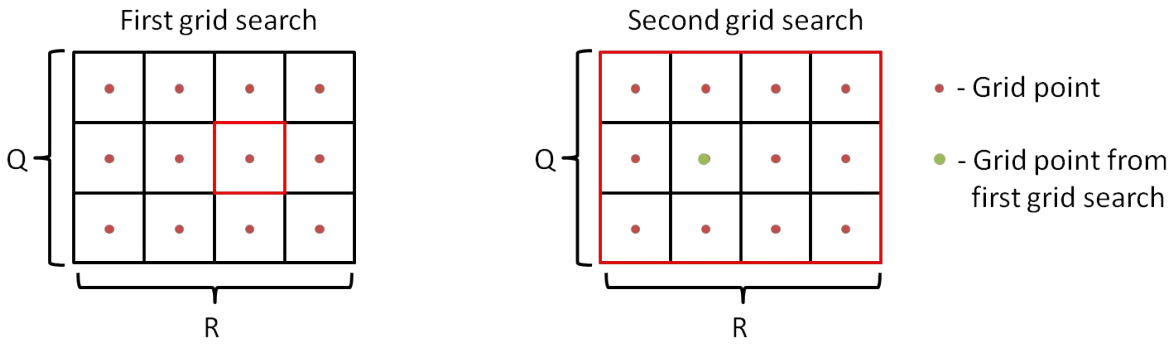

Figure 3.4: Schematic description of a two step hierarchic grid search.

grid $(Q \cdot R)^2$ function evaluations will be needed.

3.3 Convergence criterion

Convergence criterion are used to determine when the algorithm should terminate. We have used the same convergence criterion for both methods.

As stated in Section 2.4 and 2.5 both the Levenberg-Marquardt and variable projection method are built to find the parameter that minimize the cost-function, (2.16). It is also stated that the cost-function should always be decreasing. Hence, the first criteria is

$$S^{(k+1)} < S^{(k)}, \tag{3.13}$$

where S is the cost-function and k is the iteration index.

The other criterion implemented here takes into account what accuracy of the sought after parameter is needed. If the change of the cost-function is smaller than a set tolerance, μ_1,

$$||S^{(k)} - S^{(k+1)}|| < \mu_1, \tag{3.14}$$

then $\boldsymbol{p}$ is deemed to be accurate enough. Similarly if the scaled norm of parameter change is smaller than a set tolerance, μ_2,

$$||\boldsymbol{h}/\boldsymbol{p}|| < \mu_2, \tag{3.15}$$

the parameters are deemed accurate enough. Note that here the division of the parameter change, $\boldsymbol{h}$, and the current estimate of the parameters, $\boldsymbol{p}$, is done element wise. If neither of these criteria's are met the algorithm will terminate after a pre-specified number of iterations, k_{max}.

CHAPTER 4

Comparison between the two methods

In this section the methods will be tested and compared against each other. The comparison will be based on computational speed and accuracy in finding the parameters. Lastly the methods will be applied to real measurement data.

4.1 Validation of the two estimation techniques

All measurements are to a varying degree subject to disturbances. It is therefore of interest to evaluate how sensitive the two estimation methods are to disturbances and how accurately the parameters can be estimated. Here the Monte-Carlo method was used for the evaluation.

For the Monte-Carlo simulations $10,000$ realizations of a generated relative permittivity spectrum was used. The realizations where made to emulate the measurement data in regards of expected parameters, noise and used frequency range. The parameters of the generated spectrum was

$$\boldsymbol{p} = \begin{bmatrix} \varepsilon_\infty \\ \varepsilon_s \\ \tau \\ \alpha \\ \sigma \end{bmatrix} = \begin{bmatrix} 2.2000 \\ 2.3500 \\ 1.4700 \cdot 10^{-8} \\ 0.6000 \\ 2.5000 \cdot 10^{-9} \end{bmatrix}. \tag{4.1}$$

For each realization additive white Gaussian noise (AWGN) was added. The signal-to-noise ratio (SNR) of the added noise was chosen with regards to the SNR of the set of measured spectrum, see appendix. For the Monte-Carlo simulation the SNR was chosen to be similar to the worst part of the worst of the measured spectrum, this resulted in $SNR_{dB} = 40$.

An example of a generated signal where AWGN with a $SNR_{dB} = 40$ has been added can be seen in Figure 4.1 below.

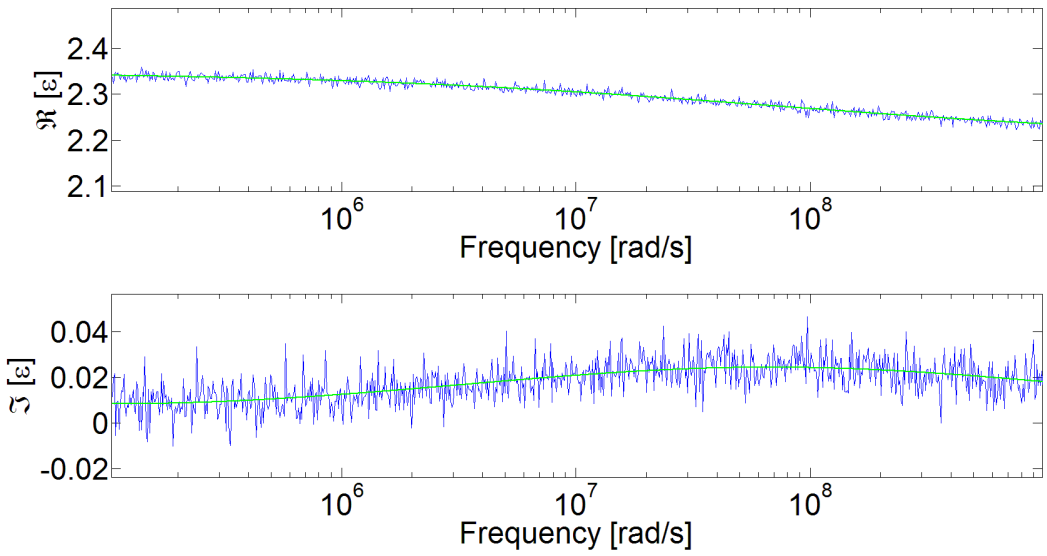

Figure 4.1: Example showing the generated curve (green) and the generated curve when AWGN with SNR = 40 dB has been added (blue).

The Cole-Cole equation was fitted to each realization using both estimation techniques. The grid used in the variable projection method, see Section 3.2.2, divided the parameter interval of τ and α into 15 and 10 segments respectively. The interval for the relaxation time was $\tau \in [10^{-12}, 10^{-5}]$. The methods used the same tolerance values for the convergence criterion, see Section 3.3. If the change in cost-function, (3.14), is smaller than $\mu_1 = 10^{-6}$ or the scaled parameter change, (3.15), is smaller than $\mu_2 = 10^{-4}$ the algorithms will terminate. If neither of these are met the algorithms will terminate after reaching the maximum number of iterations $k_{max} = 1000$.

The Monte-Carlo simulation was used to calculate a two standard deviations ($2SD$) confidence interval for each of the parameters. That is, the confidence interval indicates that the parameters lies within $2SD$ of the mean. In Figure 4.2 the generated permittivity spectrum is shown without noise together with the mean estimate and confidence interval of the spectrum calculated using both the variable projection method and the Levenberg-Marquardt method.

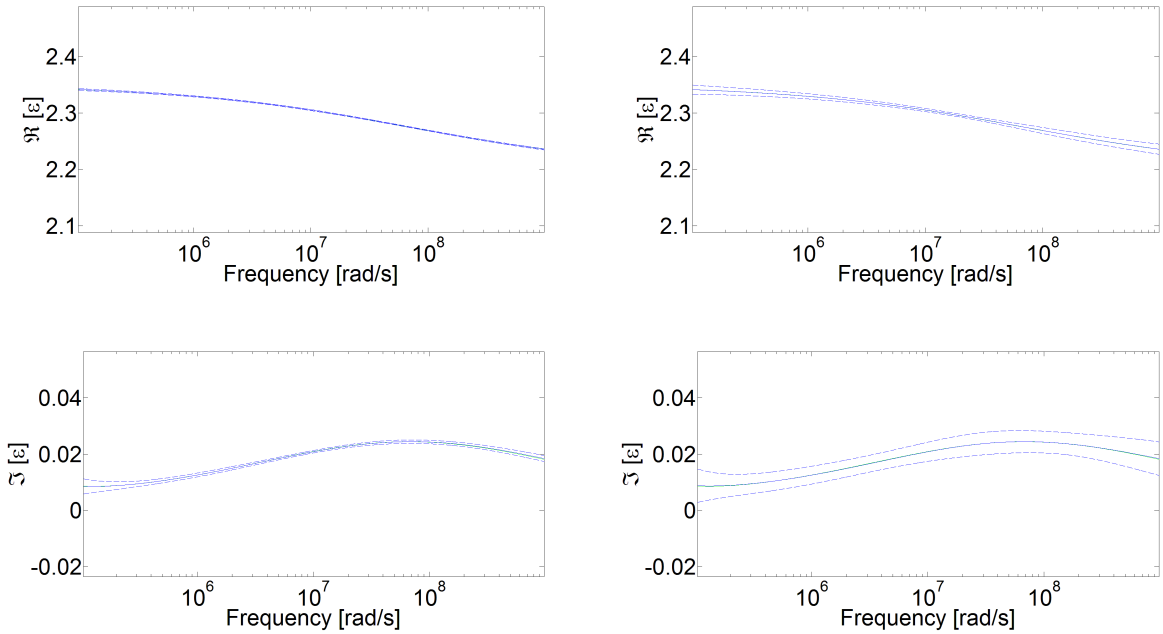

Figure 4.2: Result from the Monte-Carlo simulation. The top and bottom left plots are the real and imaginary part of the permittivity spectrum for the variable projection method, respectively. The top and bottom right plots are the real and imaginary part of the permittivity spectrum for the Levenberg-Marquardt method, respectively. The green and blue solid lines represent the true and mean estimated permittivity spectrum, respectively. The confidence interval is shown as dashed lines.

From the Figure 4.2 it seems like there are more uncertainties in the Levenberg-Marquardt method, due to it having a wider confidence interval, than in the variable projection method. The corresponding parameters and their confidence interval are shown in Table 4.1.

Table 4.1: Monte-Carlo simulation results. The computed parameters $\boldsymbol{p}^*$ are shown with a calculated $2SD$ confidence interval and the mean starting guess $\boldsymbol{p}_0$ for both methods. The generated curve used the parameters displayed in the column *True parameters*.

	Variable projection		True	Levenberg-Marquardt	
	$E[\boldsymbol{p}^*] \pm 2SD$	$E[\boldsymbol{p}_0]$	parameters	$E[\boldsymbol{p}^*] \pm 2SD$	$E[\boldsymbol{p}_0]$
ε_∞	2.1985 ± 0.0063	2.1981	2.2000	2.1952 ± 0.0381	2.2357
ε_s	2.3504 ± 0.0035	2.3498	2.3500	2.3514 ± 0.0152	2.3411
τ $[10^{-8}]$	1.4147 ± 0.2296	1.3947	1.4700	3.1714 ± 5.5305	5.4472
α	0.6041 ± 0.0263	0.6050	0.6000	0.6034 ± 0.1336	0.5000
σ $[10^{-9}]$	2.3456 ± 2.6728	2.1820	2.5000	2.5140 ± 5.1158	9.1158

From Table 4.1 it can be seen that the variable projection method had better (average) initial guesses, i.e. closer to the true values, than the Levenberg-Marquardt method whose initial guesses for especially τ and σ where poor.

The mean estimate of the parameters for both method are on the other hand quite close to the true parameter values. The width of the confidence interval varies between the parameters but the variable projection method had a tighter confidence interval in all cases. For the variable projection method the estimates of ε_s and ε_∞ are good, inside the $2SD$ confidence interval the estimated parameters are within 0.1%, respectively 0.3%, from the true parameter value. The estimates for relaxation time, τ, and the distribution factor, α, both had wider confidence intervals resulting in the estimates being within 16%, respectively 4% from the true values. As seen from Table 4.1 the confidence interval for σ were poor.

In combination with the results from Table 4.1 it is interesting to see how the parameters affect the mean squared error (MSE). This was studied by generating a curve, $\boldsymbol{y}_1$, with the parameters in (4.1). Another curve , $\boldsymbol{y}_2$, was generated using the same parameters with the exception that ε_∞ had been increased with 5%. The MSE between the two curves was then calculated. The same thing was done for the four other parameters. The result can be seen in the Table 4.2. As seen from Table 4.2 a 5% increase of the

Table 4.2: Affect on MSE due to a 5 % change in a parameter.

	MSE
$\varepsilon_\infty + 5\%$	3.5
$\varepsilon_s + 5\%$	5.6
$\tau + 5\%$	$3.5 \cdot 10^{-4}$
$\alpha + 5\%$	$6.9 \cdot 10^{-3}$
$\sigma + 5\%$	$8.8 \cdot 10^{-7}$

parameters τ, α and σ hardly affects the MSE at all. Whereas a 5 % increase of ε_∞ and especially ε_s have a huge affect on the MSE. Combining the result from Table 4.1 and 4.2 the overall conclusion is that the more the parameters affect the MSE the more accurate the estimate of the parameters will be.

As seen from Table 4.2, σ has a small affect on the MSE. This is largely due to the fact that the affect of σ on the spectrum is in the low frequency region of the imaginary part. Extending the frequency range more in the low frequency region should give clearer indication of the affect of σ and hence make it easier to estimate. This was confirmed with a Monte-Carlo simulation where the lower end of the frequency interval was changed from roughly $\omega = 10^5$ to $\omega = 10^4$. The confidence interval for σ became a lot better for both methods $2.4886 \cdot 10^{-9} \pm 0.2812 \cdot 10^{-9}$ and $2.4807 \cdot 10^{-9} \pm 0.5615 \cdot 10^{-9}$ for the variable projection method and the Levenberg-Marquardt method respectively.

The increase of the frequency range had some affect the of the other parameters. For

the variable projection method the estimate for ε_s and α became better, 2.3502 ± 0.0018 and 0.6030 ± 0.0208, whereas the estimate for ε_∞ not experience any significant change. The increase of the frequency range made the grid segments for τ a bit larger which had a negative effect on the estimate $1.4304 \cdot 10^{-8} \pm 0.3408 \cdot 10^{-8}$. To obtain the same coarseness in the grid search the number of segments used in the first level of the hierarchy can be increased.

The Levenberg-Marquardt method obtained overall better estimates after the lower end of the frequency interval was changed than it did before the change. Notably the estimate of τ, which became $1.5322 \cdot 10^{-8} \pm 1.2674 \cdot 10^{-8}$ and the estimate of ε_s, which became 2.3505 ± 0.0091, likely due to the fact that better starting guesses could be obtained, $\varepsilon_{s,0} = 2.3467$.

During the first Monte-Carlo simulations some other quantities where also calculated and can be seen in Table 4.3. Form Table 4.3 it should be noted that $\boldsymbol{h}$ is the parameter

Table 4.3: Quantities used for an extended comparison. All quantities below are averages obtained from the Monte-Carlo simulation.

	Variable projection	Levenberg-Marquardt
MSE	0.0794	0.0937
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	0.0452	0.0795
Number of iterations	4	374
Computation time	0.4660	0.2968

change and $\boldsymbol{p}$ is the current parameter estimate. The quantity $||\boldsymbol{h}/\boldsymbol{p}||$ represents the scaled parameter change during the last iteration and that the division is element wise. As seen the average last step was smaller for the variable projection method. The average MSE is also smaller for the variable projection method. It should be noted that poor initial guesses might result in the methods only finding local minima, which would affect the MSE.

The variable projection method had fewer iterations, better initial guesses is one likely reason. The computation time was on the other hand longer due to the gird search being time consuming. The gird search is responsible for about 40% of the computation time for the variable projection method. As seen the variable projection method is overall more time consuming than the Levenberg-Marquardt method.

The average number of iterations differ a lot between the methods. The Levenberg-Marquardt method use an average of 374 iterations compared to about 4 for the variable projection method. It should also be noted that in about 32% of the cases the Levenberg-Marquardt method is terminated because the maximum number of iterations has been reached. In contrast, the maximum number of iterations was never reached by the variable projection method.

4.2 Applying the models to real measurement data

In this section the two estimation method, the Levenberg-Marquardt method and the variable projection method, will be applied to experimental data. The data set was provided by *Christian Michelsen Research* and contained the measurements of the relative permittivity spectrum for 15 crude oils. The oils originate mainly from the North Sea, but includes one sample from Nigeria/Angola and Brazil respectively. The oils have been classified into two groups, biodegraded oils labeled B and non-biodegraded oils that will be labeled S.

Biodegraded crude oil refers to oils where microorganisms has been degrading certain chemical compounds in the oil and thus altered the crude oil composition. The sample set here contains 6 biodegraded and 9 non-biodegraded oils.

4.2.1 The measurement data

During the storage of the oils waxes may have precipitated. Thus, before the permittivity measurements could be made potential waxes had to be dissolved. This was achieved by placing the oils in an oven, set at a temperature of 60^{o}C, for four hours and afterwards homogenize them by shaking and turning them upside down multiple times.

The permittivity spectra was measured over the frequency range extending from 100 kHz to 1 GHz, all measurements where performed at 20^o C using the experimental set-up and procedure described in [23]. The measurement cell was 20 cm long and the oils where placed between the inner and outer conductor. The measurements were divided into two parts, one for the low frequencies and one for the high frequencies.

In the low frequency range, meaning below 20 MHz, an impedance analyzer (Hewlett Packard 4294) was used to measure the impedance of the measurement cell as a function of frequency. This measurement, combined with a linear, frequency independent model, where the used to calculate the permittivity. Measurements of n-decane, toluene and n-heptane where used to calibrate the model. The n-heptane measurements where used to correct systematic errors in the imaginary part of the measured relative permittivity.

For the high frequency, above 50 MHz, measurements a network analyzer (Hewlett Packard HP8722C and Rohde & Schwarz ZVL-13) was used. The network analyzer measured the reaction and transmission coefficients of the measurement cell. The permittivity was calculate using a bilinear calibration method, see [24] and [25].

Three calibration fluids are required for the calibration of the bilinear method, n-heptane, n-decane and toluene were used. The reaction measurements where used to calculate the permittivity at the frequency range extending from $50 - 150$ MHz. For frequencies higher than 150 MHz the transmission measurements where used.

4.2.2 Result

In this section will only show the results for one of the crude oils, oil $1S$. Results, in form of plots and tables, for the 14 remaining oils can be seen in the appendix. It should be noted that the same settings for the algorithms used in the Monte-Carlo simulation was used here. Since it is unknown if the conductivity parameter, σ, should be included in the Cole-Cole equation, (2.14), both versions of the model were applied and the one resulting in the smallest MSE was chosen. Figure 4.3 below displays the measured permittivity spectrum of the oil (dotted and blue) together with the fitted model (solid and red) for both the variable projection method and the Levenberg-Marquardt method.

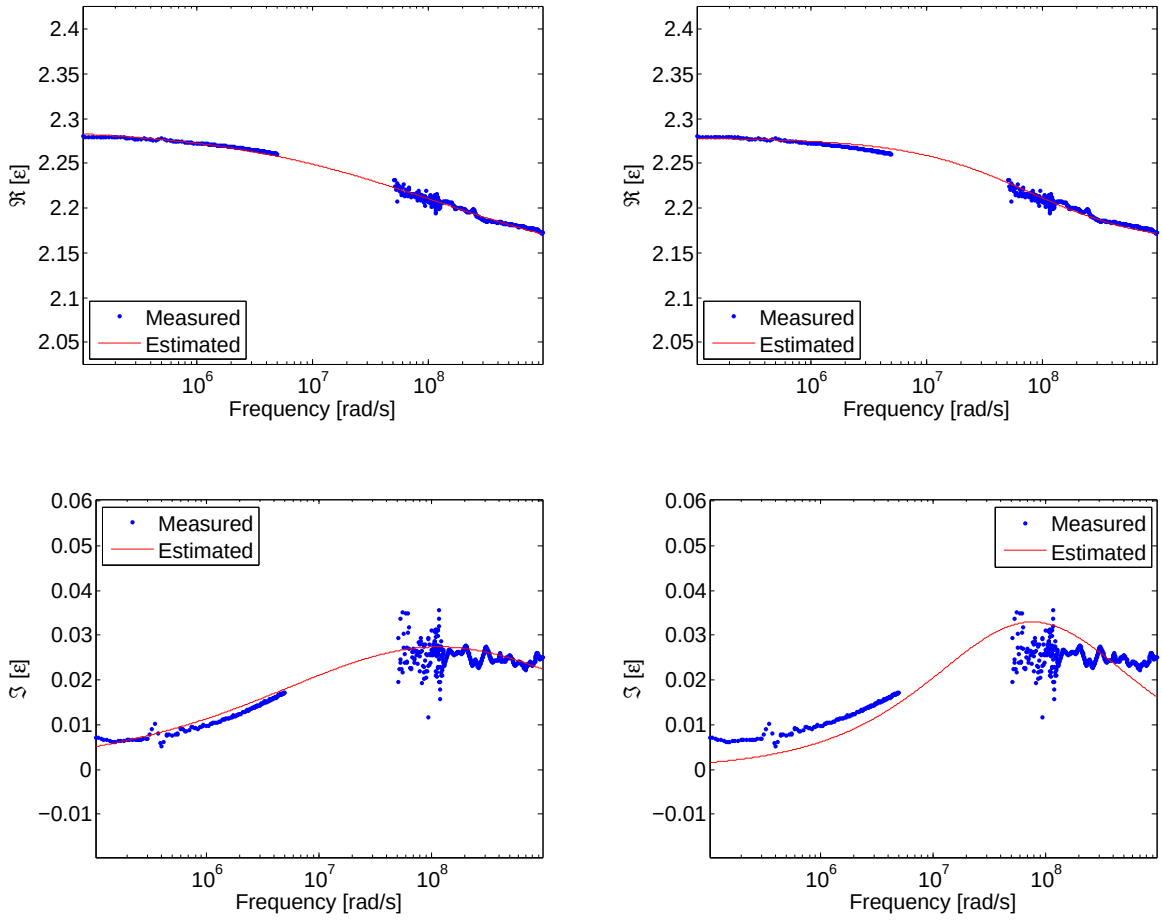

Figure 4.3: Real (top) and imaginary (bottom) part of the estimated and measured relative permittivity spectrum for oil 1S. The estimated spectrum in the left hand plots was obtained with the variable projection method and the estimated spectrum in the right hand plots was obtained with the Levenberg-Marquardt method respectively.

As can be seen from Figure 4.3 there is a gap in the measurement data between the

high and low frequencies which is due to the change of measurement equipment. The gap in the data may at times lead to a poor initial guess of τ for the Levenberg-Marquardt method. This does not seem to be a problem in this case since the initial guesses of τ are quite similar for both methods.

It can also be seen that for the most part the measurement noise is quite limited. The majority of the noise seems to occur after the changing measurement equipment to measure the higher frequencies and then especially at the imaginary part of the spectrum.

From Figure 4.3 it seems like the variable projection method provided a better fit of the Cole-Cole model to the spectrum. Comparing the MSE of the methods, see MSE in Table 4.5, it is seen that there is a significant difference, in favor of the variable projection method. Comparing the initial guess with the final estimate of the parameters, see Table 4.4, for the variable projection method, one can see that the difference is small and hence only one iteration was needed. It is also interesting to note that the variable projection method deemed that the estimate was better when the σ parameter was neglected, hence $\sigma = 0$. If the conduction current is negligible it would mean that the oil 1S not is a lossy medium.

Table 4.4: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 1S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.1249	2.1254	2.1570	2.1720
ε_s	2.2902	2.2899	2.2790	2.2797
τ	$8.4834 \cdot 10^{-9}$	$8.4834 \cdot 10^{-9}$	$1.2900 \cdot 10^{-8}$	$8.6402 \cdot 10^{-9}$
α	0.5924	0.5924	0.3690	0.5000
σ	0	0	0	0

Table 4.5: Quantities used for extended comparison, oil 1S .

	Variable projection	Levenberg-Marquardt
MSE	0.0097	0.0305
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$2.3465 \cdot 10^{-4}$	$5.4406 \cdot 10^{-4}$
Number of iterations	1	70
Computation time [s]	0.355	0.100

The conclusion from this is that the variable projection method did a better job of fitting the Cole-Cole model to the measure spectrum of the oil 1S than the Levenberg-Marquardt method did.

The accuracy of the initial guesses for the variable projection method is dependant on the coarseness of the grid search. To study how the coarseness of the grid affected

the result for the variable projection method the coarseness was both increased and decreased. The change of the coarseness was achieved by dividing the parameter interval for τ and α into a different number of segments. Three different segment pairs were used and the result can be seen in Table 4.6.

Table 4.6: Result from variable projection method with different coarseness of the grid. The segment pair (x,y) means that the interval for τ and α was divided into x and y segments respectively.

	Variable projection		
Segments (τ, α)	$(8, 5)$	$(30, 20)$	$(2000, 200)$
ε_∞	2.1405	2.1288	2.1302
ε_s	2.2860	2.2895	2.2892
τ	$1.1790 \cdot 10^{-8}$	$9.3337 \cdot 10^{-9}$	$9.6452 \cdot 10^{-9}$
α	0.5343	0.5816	0.5772
σ	0	0	0

As can be seen from Table 4.6 changing the coarseness of the grid and hence also the accuracy of the initial guesses results in the method finding different parameters. Table 4.7 shows the various quantities used for the comparison.

Table 4.7: Quantities used for extended comparison, variable projection method with different grid coarseness of the grid.

	Variable projection		
Segments (τ, α)	$(8, 5)$	$(30, 20)$	$(2000, 200)$
MSE	0.0105	0.0095	0.0095
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$6.6303 \cdot 10^{-9}$	$2.8726 \cdot 10^{-4}$	$3.0921 \cdot 10^{-4}$
Number of iterations	2	1	1
Computation time [s]	0.3350	0.7850	392

From Table 4.7 it can be seen that decreasing the number of segments for τ and α from 15 and 10 to 8 and 5 respectively resulted in an increase of the MSE. Whereas doubling the number of segments for τ and α from 15 and 10 to 30 and 20 respectively resulted in a slight decrease of the MSE. This increase resulted in about a doubling of the computation time. Increasing the number of segments further did not seem to have any significant affect on the MSE.

Extending the comparison between the variable projection method and the Levenberg-Marquardt method to include the 14 oils in the appendix the result is similar. The

variable projection method was able to provide a better fit in all cases. The result when comparing the MSE was similarly as above, it is always significantly smaller for the variable projection method. The Levenberg-Marquardt method had at times difficulties with obtaining good initial guesses of τ, mostly in cases when there were no measurement points for the maximum of the imaginary part.

CHAPTER 5

Discussion and conclusions

In this master thesis two estimation techniques, one based on the Levenberg-Marquardt algorithm and the other on the variable projection method, were implemented. The objective was to develop an automated method that in a quick and accurate manner can estimate the parameters of the Cole-Cole model for relative permittivity.

Both methods are iterative and needs good initial guesses to ensure decent to the global minima. An advantage with the variable projection method in this regard is that the quality of the initial guesses is determined by the coarseness of the gird, hence largely depends on what is deemed acceptable in regards of computational time.

When it comes to accurately estimate the parameters it is to be beneficial to factorize the Cole-Cole model into a linear and non-linear part and hence reduce the problem down from a 5-dimensional to a 2-dimensional problem. As seen from the result of the Monte-Carlo simulation the accuracy of the variable projection method were better for all parameters, see Table 4.1. Both methods had some difficulties obtaining accurate estimations for of the relaxation time, τ, and the conductivity of the material, σ.

The difficulties with obtaining accurate estimations of τ were probably due to the small affect the parameter has on the MSE compared to the other parameters, see Table 4.2.

On the other hand, the difficulties with estimating σ stemmed from the fact that the used frequency range did not include lower frequencies. The affect that the conductivity has on the spectrum is largely in the low-frequency region of the imaginary part. If the frequency range does not include low enough frequencies the affects of σ will be difficult to see, hence it will also be difficult to estimate its magnitude.

In regards to time consumption the Levenberg-Marquardt method were faster that the variable projection method, whose the grid search was time consuming. The grid search is largely dependent on the interval for τ. If it can be reduced then the number of segments can be reduced and the computational time will be faster.

The variable projection method were more reliable than the Levenberg-Marquardt method when it came to the real measurement data. It provided smaller MSE:s for every oil and the estimated spectrum seemed to fit well with the measured spectrum.

CHAPTER 6

Future work

Some improvements to the algorithm has been considered and the most significant of them are described here.

- The implementation of complex weights to the algorithm. This will allow the algorithm to weight down noisier regions of the spectrum and thus limit the affect this the noise has on the calculations.
- The implementation of a method that estimates the accuracy of the estimated parameters. This would make it easy for the user to determine if the estimated parameters are good enough.
- Implement the method using QR-decomposition would make it more robust. If e.g. the Jacobian matrix, $\boldsymbol{J}$, is poorly conditioned computing $(\boldsymbol{J}^T\boldsymbol{J})^{-1}$ can be a source of numerical rounding errors.
- If the medium is lossy the conductivity parameter, σ, should be included in the Cole-Cole equation in order to account for the energy loss. Development of a method for determining whether or not it should included in the Cole-Cole equation would be beneficial since it would make sure a more accurate model is used.
- An extension of the algorithm could also be made to include some other relaxation models like e.g. Debye, Cole-Davidson and Havriliak-Negami relaxation.

Appendix A

Cole-Cole model fitted to measured relative permittivity spectrum of crude oils

This appendix will contain the results, in form of plots and tables, of when the variable projection method and the Levenberg-Marquardt method was used to fit the Cole-Cole equation to measured relative permittivity spectrum of crude oils. Two things should be noted. Firstly, the conductivity, parameter, σ, at times can be excluded from the Cole-Cole equation, (2.14). Therefore both versions were applied and the one resulting in the smallest MSE was chosen. For this reason the conductivity parameter will in some cases be $\sigma = 0$. Secondly, all the plots in this appendix contains four images. The top images contain the real part of the permittivity spectrum and the bottom images contain the imaginary part of the permittivity spectrum. The estimated spectrum in the left hand plots were obtained with the variable projection method and the estimated spectrum in the right hand plots were obtained with the Levenberg-Marquardt method respectively.

A.1 Oil 1S

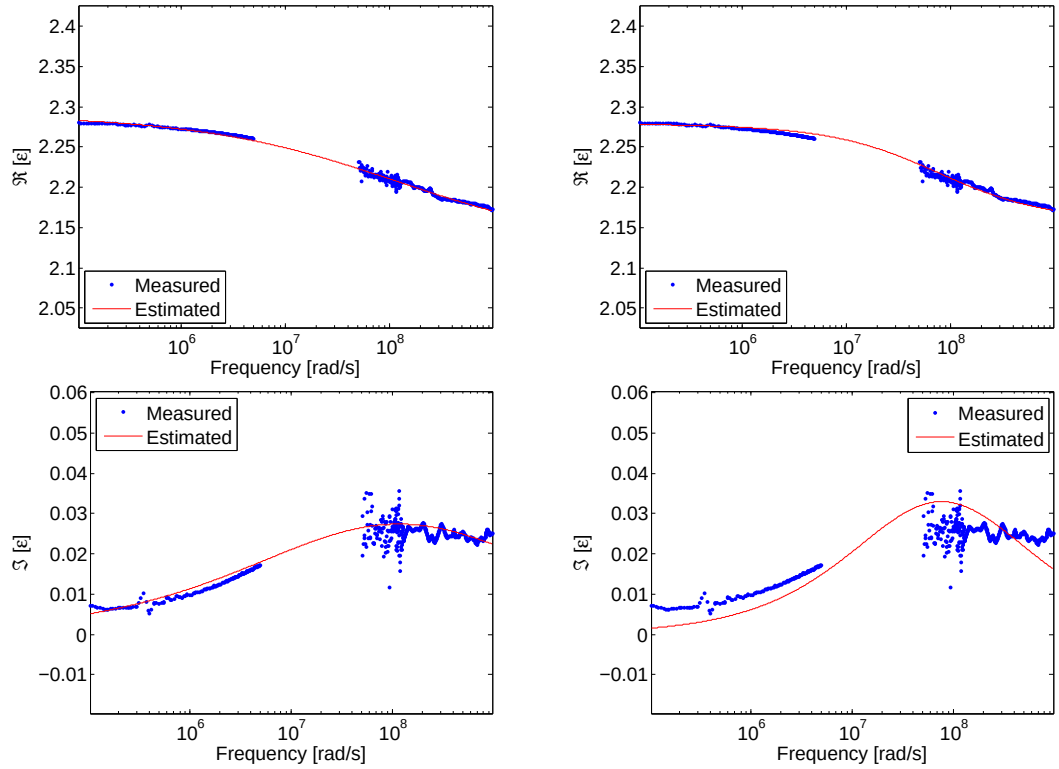

Figure A.1: Measured and estimated relative permittivity spectrum for oil 1S.

Table A.1: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 1S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.1249	2.1254	2.1570	2.1720
ε_s	2.2902	2.2899	2.2790	2.2797
τ	$8.4834 \cdot 10^{-9}$	$8.4834 \cdot 10^{-9}$	$1.2900 \cdot 10^{-8}$	$8.6402 \cdot 10^{-9}$
α	0.5924	0.5924	0.3690	0.5000
σ	0	0	0	0

Table A.2: Quantities used for extended comparison, oil 1S .

	Variable projection	Levenberg-Marquardt
MSE	0.0097	0.0305
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$2.3465 \cdot 10^{-4}$	$5.4406 \cdot 10^{-4}$
Number of iterations	1	70
Computation time [s]	0.355	0.100

A.2 Oil 2S

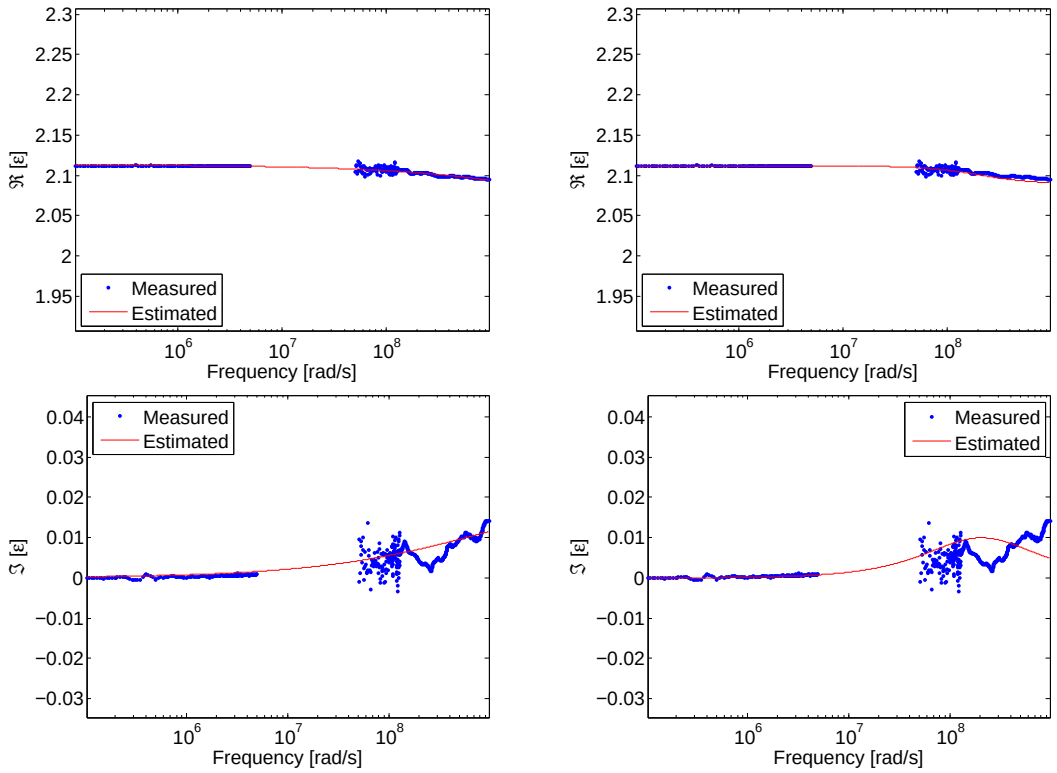

Figure A.2: Measured and estimated relative permittivity spectrum for oil 2S.

Table A.3: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 2S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.0304	2.0310	2.0890	2.0943
ε_s	2.1132	2.1130	2.1120	2.1115
τ	$1.0000 \cdot 10^{-10}$	$1.0000 \cdot 10^{-10}$	$4.8200 \cdot 10^{-9}$	$1.0342 \cdot 10^{-9}$
α	0.5431	0.5431	0.0900	0.5000
σ	0	0	0	0

Table A.4: Quantities used for extended comparison, oil 2S .

	Variable projection	Levenberg-Marquardt
MSE	0.0076	0.0256
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$3.0512 \cdot 10^{-4}$	$5.4152 \cdot 10^{-5}$
Number of iterations	1	71
Computation time [s]	0.327	0.050

A.3 Oil 3S

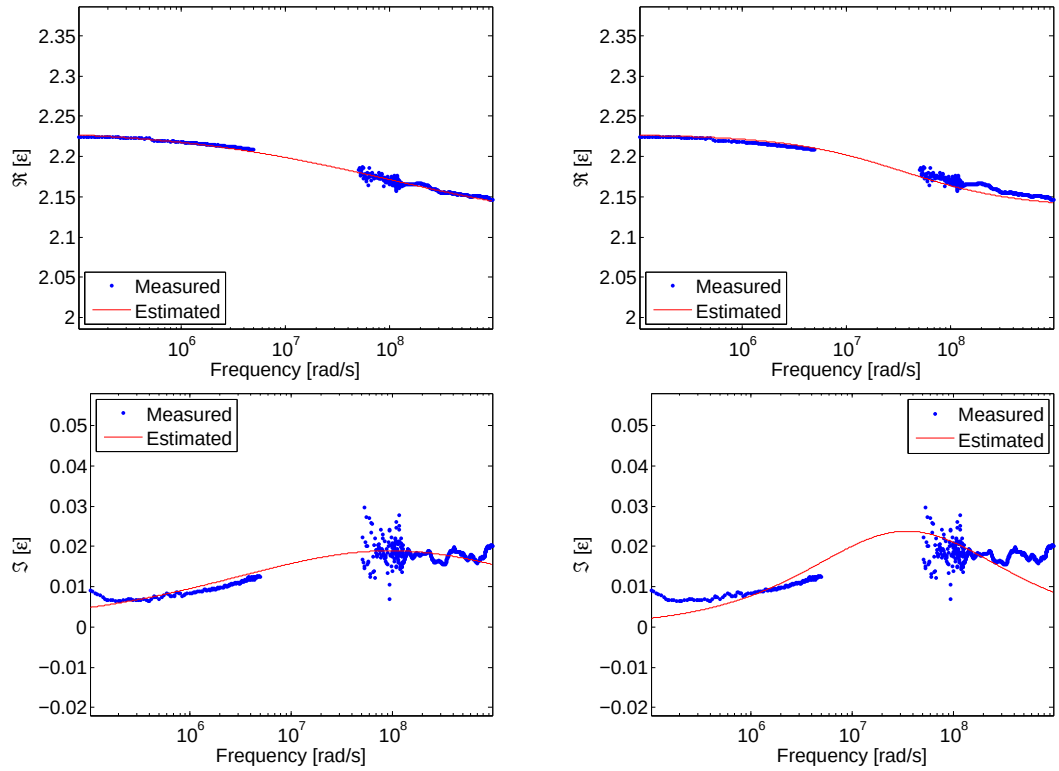

Figure A.3: Measured and estimated relative permittivity spectrum for oil 3S.

Table A.5: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 3S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.1091	2.1023	2.1350	2.1456
ε_s	2.2346	2.2371	2.2280	2.2245
τ	$1.0269 \cdot 10^{-8}$	$8.4834 \cdot 10^{-9}$	$2.8380 \cdot 10^{-8}$	$8.6399 \cdot 10^{-9}$
α	0.6265	0.6540	0.3980	0.5000
σ	0	0	0	0

Table A.6: Quantities used for extended comparison, oil 3S .

	Variable projection	Levenberg-Marquardt
MSE	0.0089	0.0501
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$3.2428 \cdot 10^{-8}$	$9.5957 \cdot 10^{-5}$
Number of iterations	2	33
Computation time [s]	0.429	0.055

A.4 Oil 4S

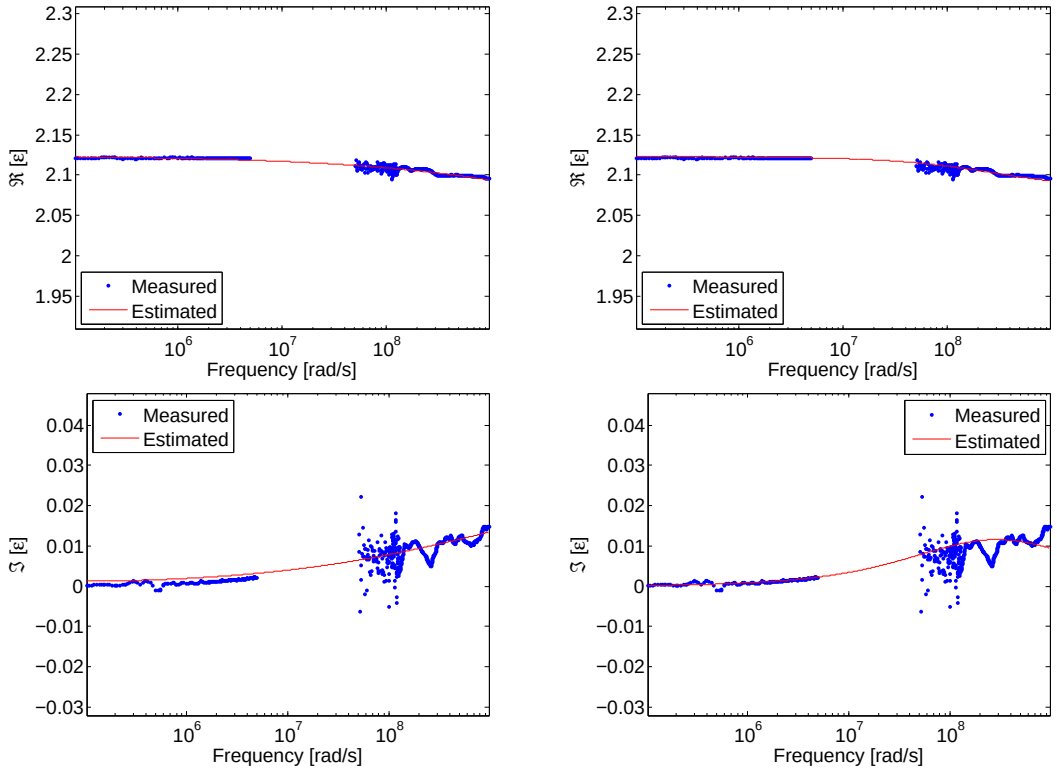

Figure A.4: Measured and estimated relative permittivity spectrum for oil 4S.

Table A.7: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 4S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	1.9816	1.9914	2.0820	2.0952
ε_s	2.1244	2.1243	2.1230	2.1207
τ	$2.7434 \cdot 10^{-11}$	$3.7276 \cdot 10^{-11}$	$3.4210 \cdot 10^{-9}$	$1.0291 \cdot 10^{-9}$
α	0.6595	0.6540	0.3440	0.5000
σ	$3.9062 \cdot 10^{-10}$	$2.3304 \cdot 10^{-10}$	0	0

Table A.8: Quantities used for extended comparison, oil 4S .

	Variable projection	Levenberg-Marquardt
MSE	0.0093	0.0154
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	0.0521	$6.4575 \cdot 10^{-5}$
Number of iterations	2	51
Computation time [s]	0.372	0.031

A.5 Oil 5B

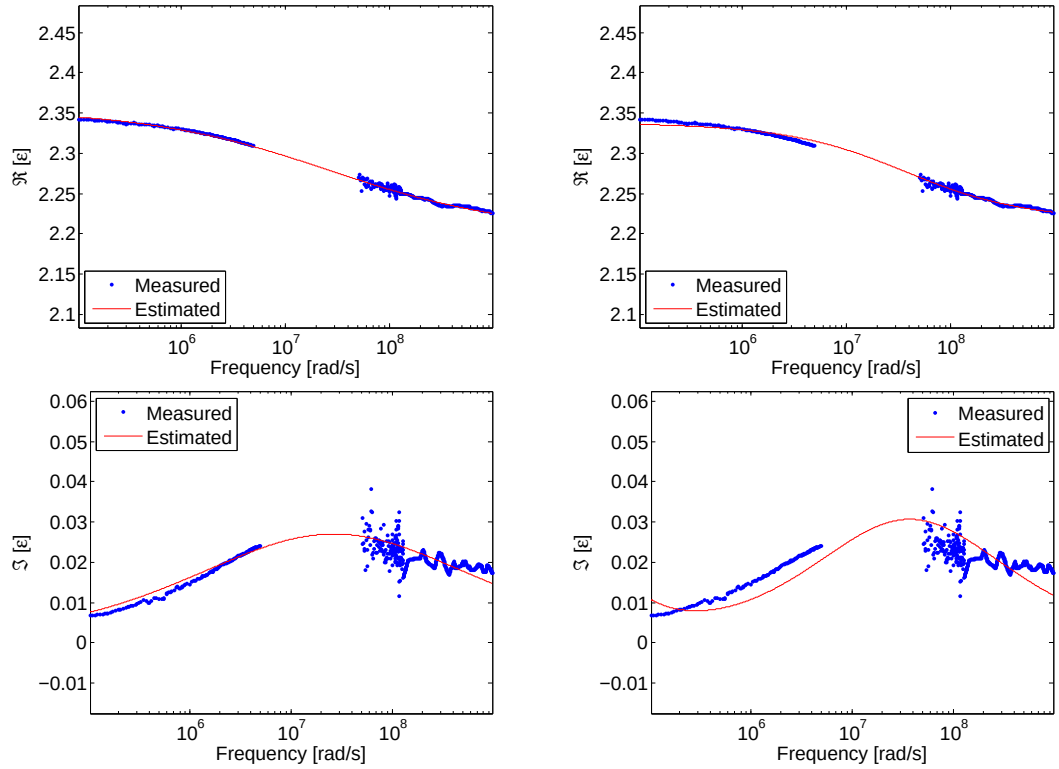

Figure A.5: Measured and estimated relative permittivity spectrum for oil 5B.

Table A.9: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 5B using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.2023	2.2026	2.2160	2.2255
ε_s	2.3553	2.3547	2.3380	2.3415
τ	$3.7276 \cdot 10^{-8}$	$3.7276 \cdot 10^{-8}$	$2.6570 \cdot 10^{-8}$	$1.7504 \cdot 10^{-8}$
α	0.5678	0.5678	0.4070	0.5000
σ	0	0	$7.3160 \cdot 10^{-9}$	$8.4911 \cdot 10^{-9}$

Table A.10: Quantities used for extended comparison, oil 5B .

	Variable projection	Levenberg-Marquardt
MSE	0.0082	0.0202
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$2.8927 \cdot 10^{-4}$	$4.1146 \cdot 10^{-3}$
Number of iterations	1	65
Computation time [s]	0.246	0.047

A.6 Oil 6B

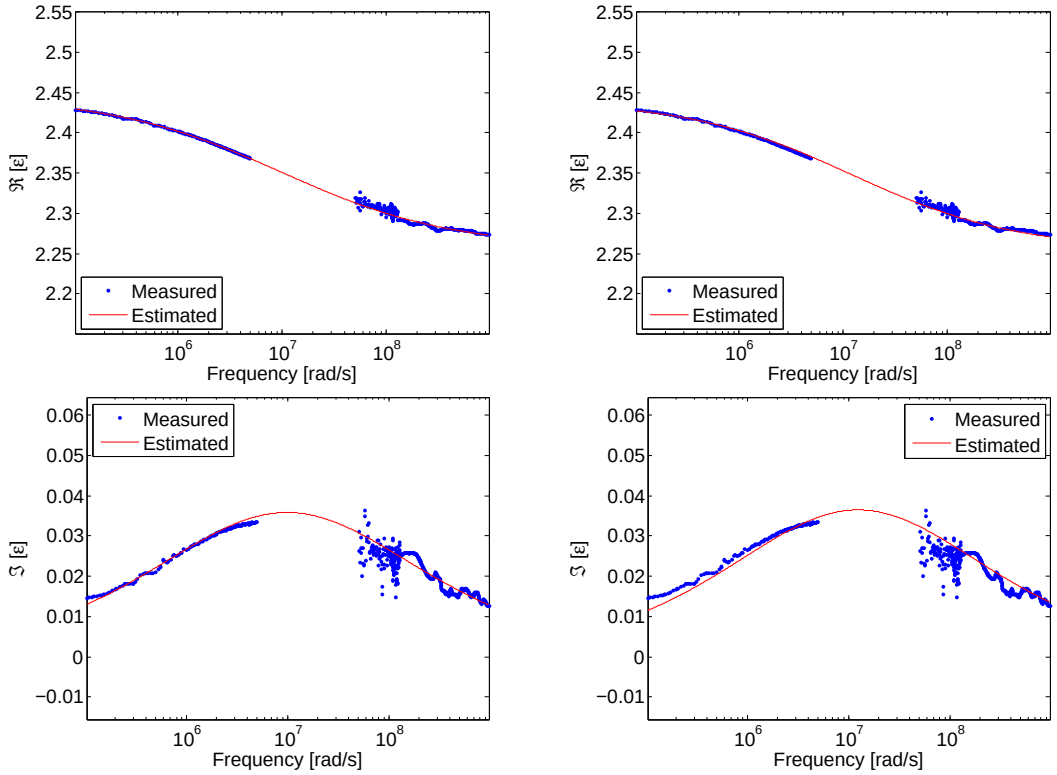

Figure A.6: Measured and estimated relative permittivity spectrum for oil 6B.

Table A.11: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 6B using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.2556	2.2557	2.2530	2.2732
ε_s	2.4466	2.4463	2.4420	2.4275
τ	$1.0000 \cdot 10^{-7}$	$1.0000 \cdot 10^{-7}$	$7.9620 \cdot 10^{-8}$	$2.1975 \cdot 10^{-7}$
α	0.5431	0.5431	0.5310	0.5000
σ	0	0	0	0

Table A.12: Quantities used for extended comparison, oil 6B .

	Variable projection	Levenberg-Marquardt
MSE	0.0105	0.0138
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$1.6624 \cdot 10^{-4}$	$6.0337 \cdot 10^{-4}$
Number of iterations	1	27
Computation time [s]	0.242	0.033

A.7 Oil 7S

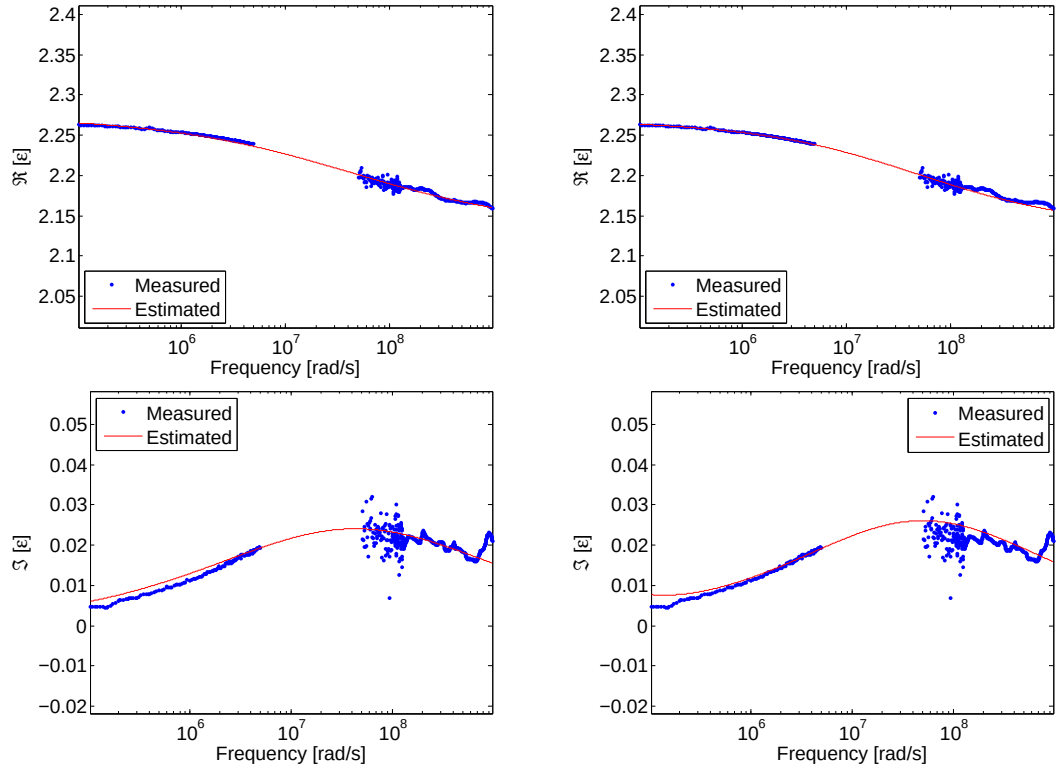

Figure A.7: Measured and estimated relative permittivity spectrum for oil 7S.

Table A.13: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 7S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.1330	2.1334	2.1330	2.1596
ε_s	2.2736	2.2732	2.2690	2.2625
τ	$2.2758 \cdot 10^{-8}$	$2.2758 \cdot 10^{-8}$	$1.9680 \cdot 10^{-8}$	$1.616 \cdot 10^{-8}$
α	0.5801	0.5801	0.5340	0.5000
σ	0	0	$2.8200 \cdot 10^{-9}$	$5.5661 \cdot 10^{-9}$

Table A.14: Quantities used for extended comparison, oil 7S .

	Variable projection	Levenberg-Marquardt
MSE	0.011	0.0203
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$2.3031 \cdot 10^{-4}$	$2.9328 \cdot 10^{-4}$
Number of iterations	1	25
Computation time [s]	0.257	0.029

A.8 Oil 10S

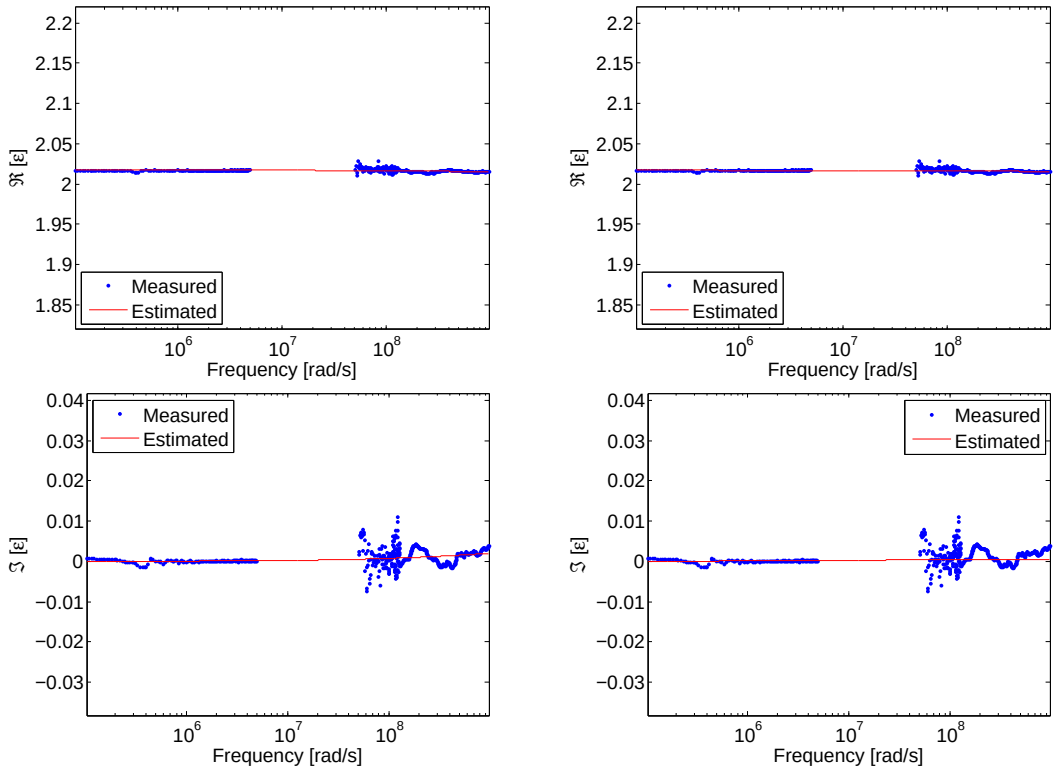

Figure A.8: Measured and estimated relative permittivity spectrum for oil 10S

Table A.15: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 10S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.0055	2.0056	2.0150	2.0152
ε_s	2.0172	2.0172	2.0170	2.0166
τ	$1.3895 \cdot 10^{-10}$	$1.3895 \cdot 10^{-10}$	$5.3480 \cdot 10^{-9}$	$5.3477 \cdot 10^{-9}$
α	0.4322	0.4322	0.5000	0.5000
σ	0	0	0	0

Table A.16: Quantities used for extended comparison, oil 10S .

	Variable projection	Levenberg-Marquardt
MSE	0.0052	0.006
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$2.8871 \cdot 10^{-5}$	$6.1347 \cdot 10^{-5}$
Number of iterations	1	23
Computation time [s]	0.271	0.012

A.9 Oil 11B

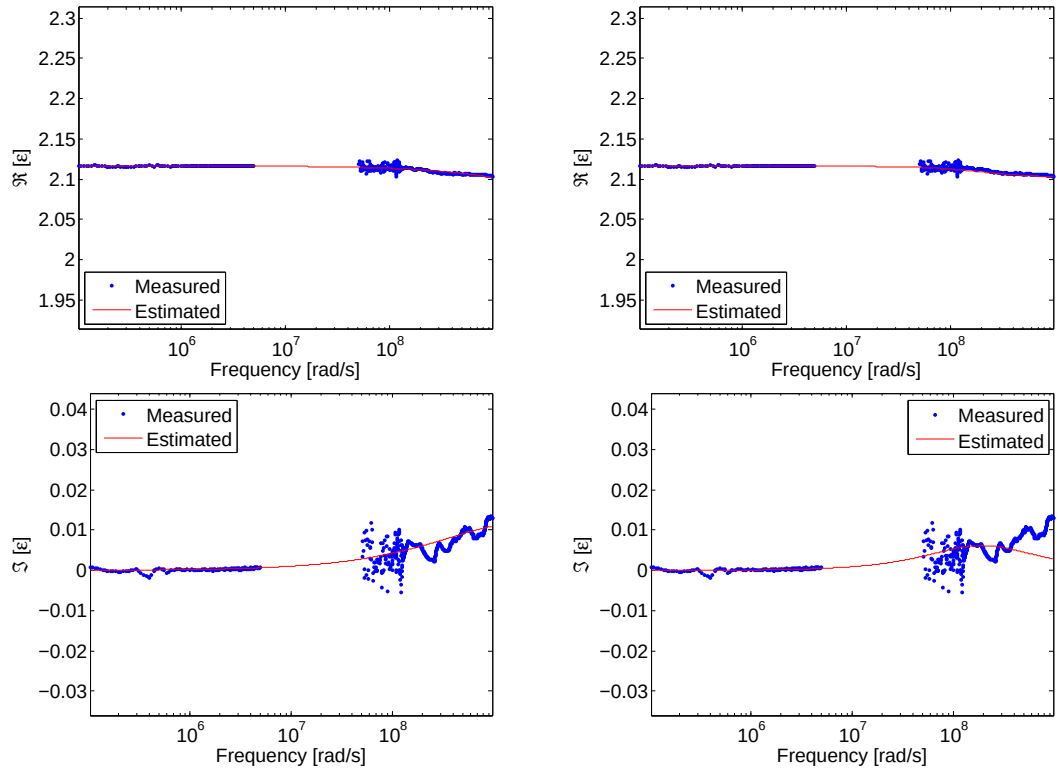

Figure A.9: Measured and estimated relative permittivity spectrum for oil 11B.

Table A.17: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 11B using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.0761	2.0769	2.1020	2.1035
ε_s	2.1165	2.1163	2.1160	2.1162
τ	$5.1795 \cdot 10^{-10}$	$5.1795 \cdot 10^{-10}$	$5.1380 \cdot 10^{-9}$	$1.0457 \cdot 10^{-9}$
α	0.3337	0.3337	0.0840	0.5000
σ	0	0	0	0

Table A.18: Quantities used for extended comparison, oil 11B .

	Variable projection	Levenberg-Marquardt
MSE	0.0069	0.0162
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$3.8911 \cdot 10^{-4}$	–
Number of iterations	1	1000
Computation time [s]	0.254	0.722

A.10 Oil 12S

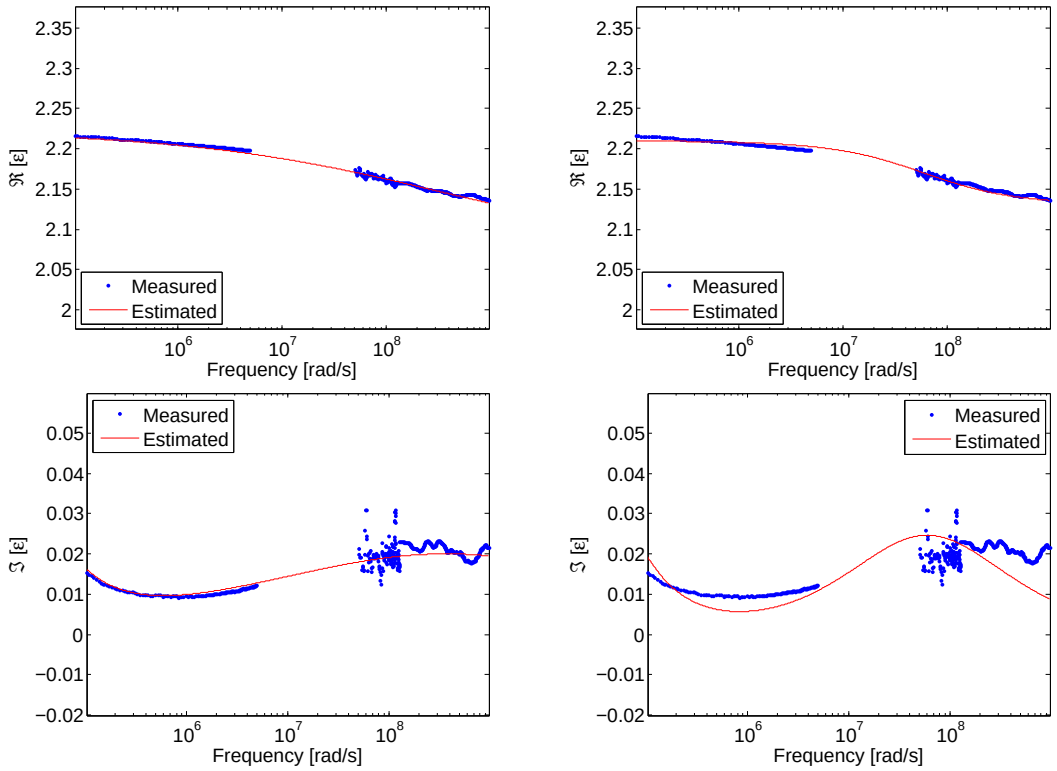

Figure A.10: Measured and estimated relative permittivity spectrum for oil 12S.

Table A.19: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 12S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.0689	2.0678	2.1300	2.1359
ε_s	2.2232	2.2233	2.2100	2.2150
τ	$2.8012 \cdot 10^{-9}$	$2.6827 \cdot 10^{-9}$	$1.6510 \cdot 10^{-8}$	$8.5515 \cdot 10^{-9}$
α	0.6767	0.6787	0.2980	0.5000
σ	$9.3244 \cdot 10^{-9}$	$9.2217 \cdot 10^{-9}$	$1.702, \cdot 10^{-8}$	$1.5817 \cdot 10^{-8}$

Table A.20: Quantities used for extended comparison, oil 12S .

	Variable projection	Levenberg-Marquardt
MSE	0.0088	0.0354
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	0.0049	$6.7963 \cdot 10^{-5}$
Number of iterations	1	115
Computation time [s]	0.254	0.067

A.11 Oil 13B

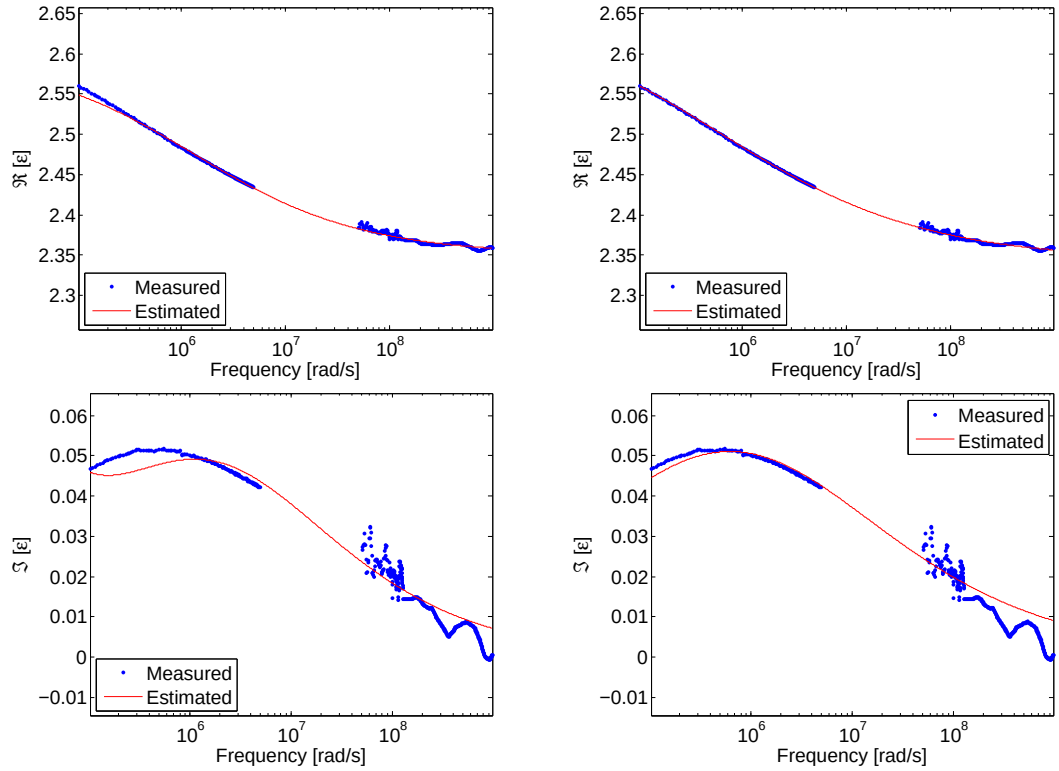

Figure A.11: Measured and estimated relative permittivity spectrum for oil 13B.

Table A.21: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 13B using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.3487	2.3501	2.3430	2.3593
ε_s	2.6180	2.5995	2.6590	2.5563
τ	$9.2086 \cdot 10^{-7}$	$7.1969 \cdot 10^{-7}$	$1.5840 \cdot 10^{-6}$	$2.3904 \cdot 10^{-6}$
α	0.5494	0.5308	0.6020	0.5000
σ	$9.3177 \cdot 10^{-9}$	$1.3998 \cdot 10^{-8}$	0	0

Table A.22: Quantities used for extended comparison, oil 13B .

	Variable projection	Levenberg-Marquardt
MSE	0.0152	0.0196
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	0.0672	$2.1479 \cdot 10^{-3}$
Number of iterations	1	122
Computation time [s]	0.315	0.101

A.12 Oil 15B

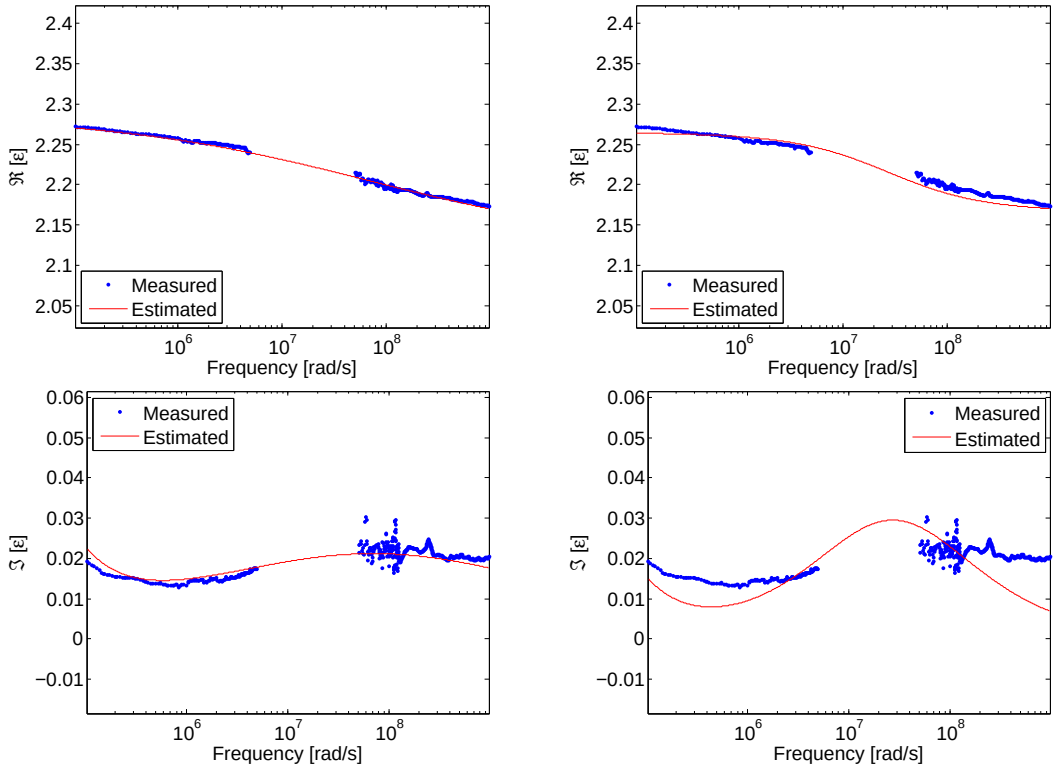

Figure A.12: Measured and estimated relative permittivity spectrum for oil 15B.

Table A.23: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 15B using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.1227	2.1227	2.1660	2.1736
ε_s	2.2870	2.2870	2.2650	2.2714
τ	$1.3895 \cdot 10^{-8}$	$1.3895 \cdot 10^{-8}$	$3.6100 \cdot 10^{-8}$	$8.5098 \cdot 10^{-9}$
α	0.6787	0.6787	0.3160	0.5000
σ	$1.2133 \cdot 10^{-8}$	$1.2133 \cdot 10^{-8}$	$1.2290 \cdot 10^{-8}$	$2.0494 \cdot 10^{-8}$

Table A.24: Quantities used for extended comparison, oil 15B .

	Variable projection	Levenberg-Marquardt
MSE	0.0076	0.0852
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$6.0993 \cdot 10^{-8}$	$8.0109 \cdot 10^{-2}$
Number of iterations	1	1000
Computation time [s]	0.255	0.739

A.13 Oil 16S

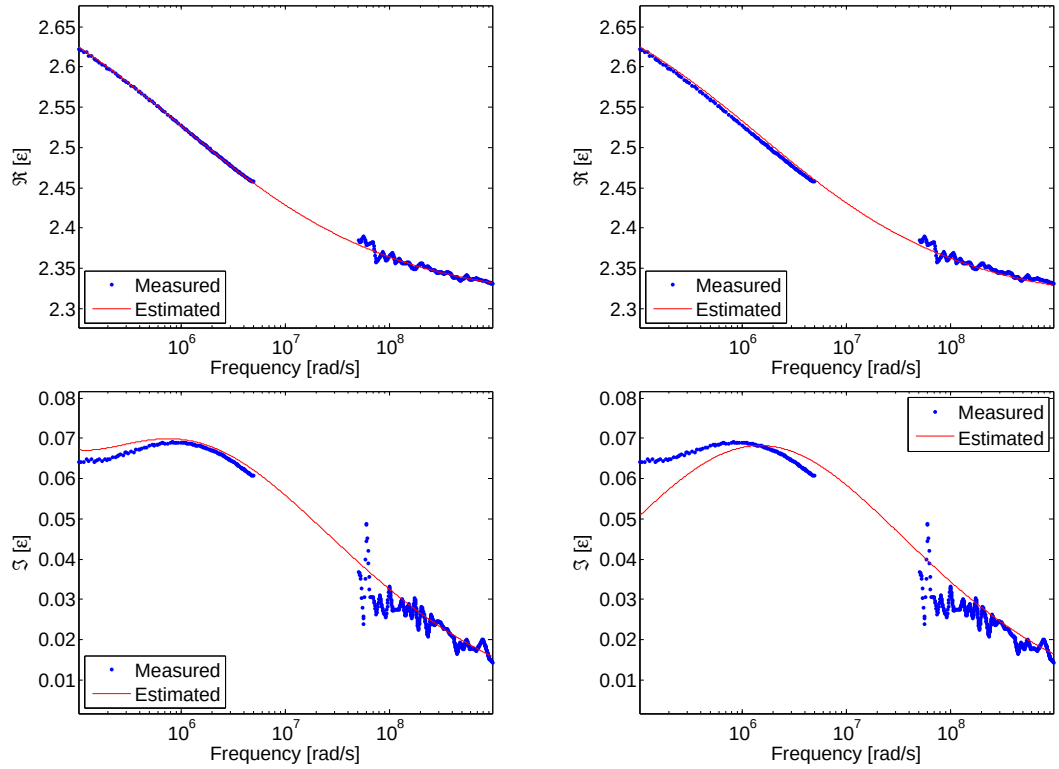

Figure A.13: Measured and estimated relative permittivity spectrum for oil 16S.

Table A.25: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 16S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.3069	2.3069	2.3040	2.3311
ε_s	2.7483	2.7483	2.7250	2.6186
τ	$1.0000 \cdot 10^{-6}$	$1.0000 \cdot 10^{-6}$	$6.6210 \cdot 10^{-7}$	$1.1240 \cdot 10^{-6}$
α	0.6170	0.6170	0.6010	0.5000
σ	$1.0157 \cdot 10^{-8}$	$1.0157 \cdot 10^{-8}$	0	0

Table A.26: Quantities used for extended comparison, oil 16S .

	Variable projection	Levenberg-Marquardt
MSE	0.0123	0.0285
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$2.2833 \cdot 10^{-8}$	$7.7404 \cdot 10^{-5}$
Number of iterations	1	29
Computation time [s]	0.248	0.028

A.14 Oil 17S

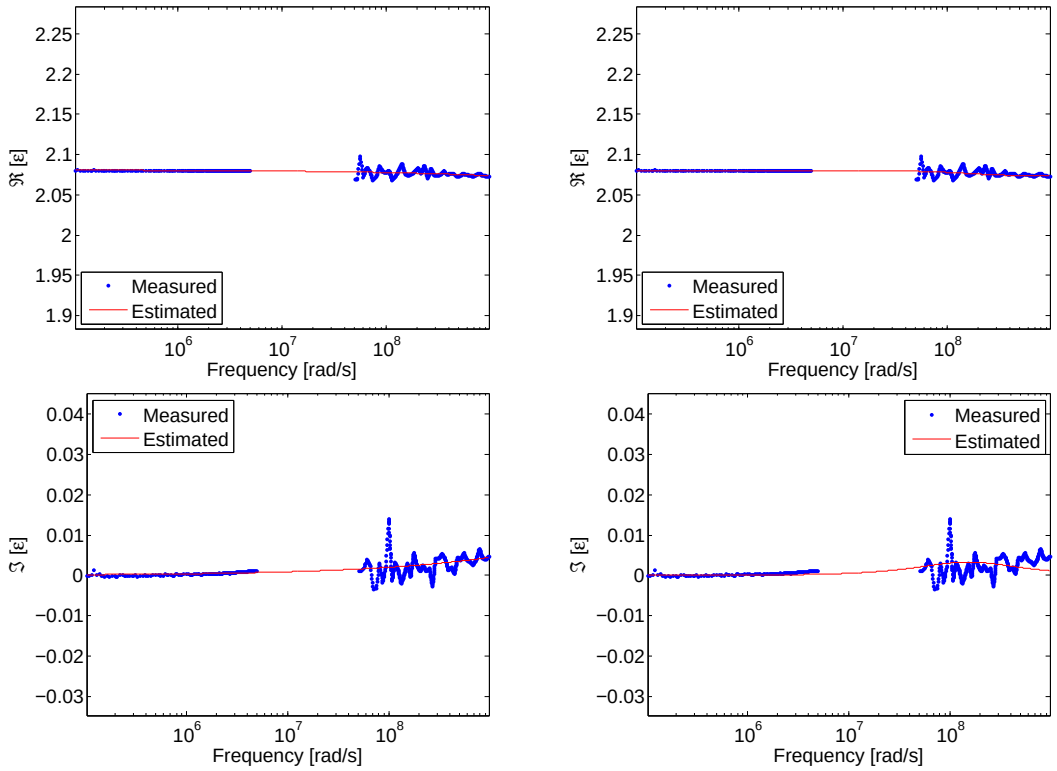

Figure A.14: Measured and estimated relative permittivity spectrum for oil 17S.

Table A.27: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 17S using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	1.9813	1.9815	2.0730	2.0729
ε_s	2.0806	2.0806	2.0800	2.0801
τ	$1.9307 \cdot 10^{-12}$	$1.9307 \cdot 10^{-12}$	$6.496 \cdot 10^{-9}$	$9.9027 \cdot 10^{-9}$
α	0.6170	0.6170	0.0410	0.5000
σ	0	0	0	0

Table A.28: Quantities used for extended comparison, oil 17S .

	Variable projection	Levenberg-Marquardt
MSE	0.013	0.0154
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$7.0101 \cdot 10^{-5}$	$1.0225 \cdot 10^{-3}$
Number of iterations	3	116
Computation time [s]	0.455	0.077

A.15 Oil 18B

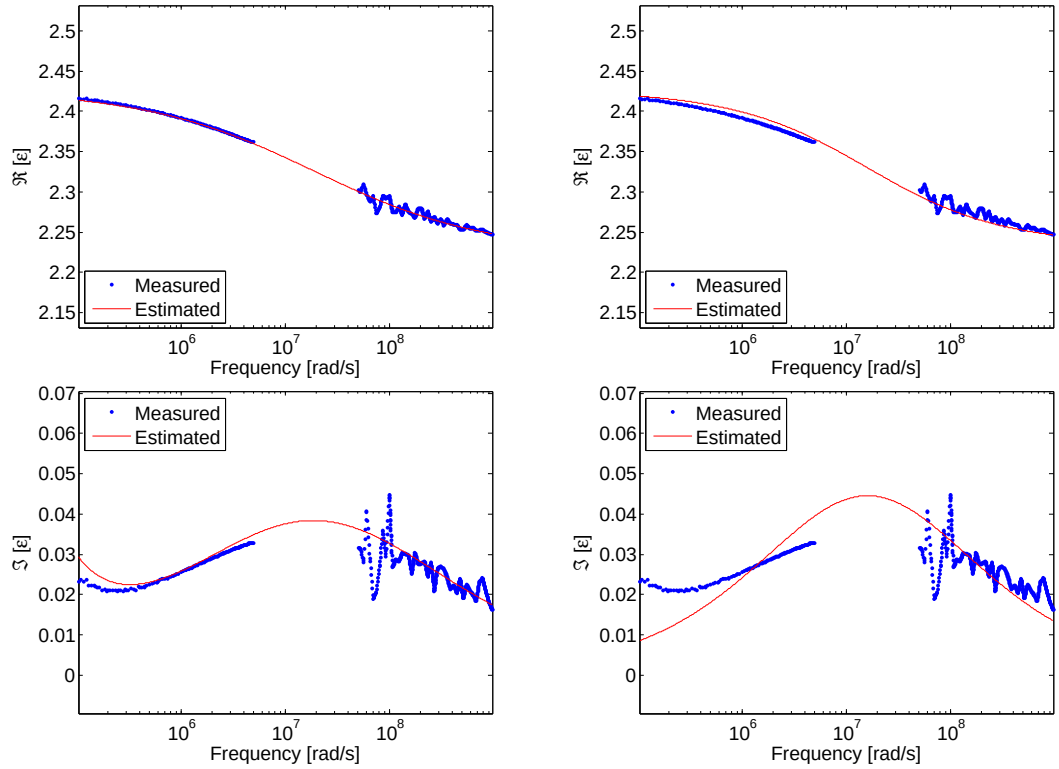

Figure A.15: Measured and estimated relative permittivity spectrum for oil 18B.

Table A.29: The found parameters $\boldsymbol{p}^*$ and the initial guesses $\boldsymbol{p}_0$ for oil 18B using the variable projection method and the Levenberg-Marquardt method.

	Variable projection		Levenberg-Marquardt	
	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$	$\boldsymbol{p}^*$	$\boldsymbol{p}_0$
ε_∞	2.2242	2.2242	2.2330	2.2471
ε_s	2.4268	2.4268	2.4270	2.4153
τ	$5.1795 \cdot 10^{-8}$	$5.1795 \cdot 10^{-8}$	$6.1620 \cdot 10^{-8}$	$9.9027 \cdot 10^{-9}$
α	0.5431	0.5431	0.4520	0.5000
σ	$1.6027 \cdot 10^{-8}$	$1.6027 \cdot 10^{-8}$	0	0

Table A.30: Quantities used for extended comparison, oil 18B .

	Variable projection	Levenberg-Marquardt
MSE	0.0206	0.0595
$\|\|\boldsymbol{h}/\boldsymbol{p}\|\|$	$3.3599 \cdot 10^{-9}$	$8.4423 \cdot 10^{-5}$
Number of iterations	1	10
Computation time [s]	0.273	0.016

Bibliography

[1] K. Cole and R. Cole, "Dispersion and absorption in dielectrics - i alternating current characteristics," *J. Chem. Phys.*, vol. 9, pp. 341–352, 1941.

[2] K. Folgerø, T. Friisø, J. Hilland, and T. Tjomsland, "A broad-band and high-sensitivity dielectric spectroscopy measurement system for quality determination of low-permittivity fluids," *Meas. Sci. Technol*, vol. 6, 1995.

[3] S. Grimnesand and O. Martinsen, *Bioimpedance and bioelectricity basics*. Academic Press, 2000.

[4] W. H. Pelton, S. H. Ward, P. G. Hallof, W. R. Sill, and P. Nelson, "Mineral discrimination and removal ofinductive coupling with multifrequency ip," *Geophysics*, vol. 43, pp. 588–609, 1978.

[5] Y. Luo and Y. Zhang, *Theory and Application of Spectral Induced Polarization (Geophysical Monograph Series, N0.8)*. Society of Exploration, 1998.

[6] J. Milton, *Field Geophysics. Geological Society of London Handbook Series*. Wiley, 1996.

[7] R. Thorn, G. A. Johansen, and B. T. Hjertaker, "Tree-phase flow measurement in the petroleum industry," *Meas. Sci. Technol.*, vol. 24, no. 1, pp. 588–609, 2013.

[8] Y. Yang, W. Ni, Q. Sun, H. Wen, and Z. Teng, "Improved cole parameter extraction based on the least absolute deviation method," *Physiological Measurement*, vol. 34, no. 10, 2013.

[9] W. H. Pelton, W. R. Sill, and B. D. Smith, "Interpretation of complex resistivity and dielectric data," *part II: Geophys. Trans.*, vol. 29, no. 4, pp. 11–45, 1984.

[10] S. R. Jagger and P. A. Fell, "Forward and inverse cole-cole modeling in the analysis of frequency domain electrical impedance data," *Exp. Geophys.*, vol. 19, pp. 463–470, 1988.

[11] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.

[12] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[13] J. Chen, A. Kemna, and S. S. Hubbard, "A comparison between gauss-newton and markov-chain monte-carlo based methods for inverting spectral induced-polarization data for cole-cole parameters," *Geophysics.*, vol. 73, no. 6, pp. 247–259, 2008.

[14] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, no. 2, 1973.

[15] C. Borges, "A full-newton approach to separable nonlinear least squares problems and its application to discrete least squares rational approximation," *Electron. Trans. Numer. Anal.*, vol. 35, pp. 57–68, 2009.

[16] B. E. A. Saleh and M. C. Teich, *The fundamentals of photonics.* Wiley, 2007.

[17] J. Nocedal and S. Wrigth, *Numerical Optimization, 2nd edition.* Springer, 2006.

[18] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization.* Academic Press, 1982.

[19] P. Debye, *Polar Molecules.* The Chemical Catalog Company, Inc., 1929.

[20] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization.* Academic Press, 1982.

[21] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, no. 2, pp. 226–000, 1969.

[22] P. Bergström and I. Söderkvist, "Fitting nurbs using separable least squares techniques," *IJMMNO*, vol. 3, no. 4, pp. 319–334, 2012.

[23] K. Folgerø, "Broad-band dielectric spectroscopy of low-permittivity fluids using one measurement cell," *IEEE Trans. Instrum. Meas.*, vol. 47, no. 4, 1998.

[24] R. H. Cole, J. G. Berberian, S. Mashimo, G. Chryssikos, A. Burns, and E. Tombari, "Time domain reflection methods for dielectric measurements to 10 ghz," *J. Appl. Phys.*, vol. 66, no. 2, 1989.

[25] K. Folgerø, "Bilinear calibration of coaxial transmission/reflection cells for permittivity measurement of low-loss liquids," *Meas. Sci. Tech.*, vol. 7, 1996.