Regularisation Methods for Imaging from Electrical Measurements

by

Andrea Borsic
BSc

A thesis submitted in partial fulfilment of the requirements of Oxford Brookes University for the Degree of

Doctor of Philosophy

School of Engineering
Oxford Brookes University

July 2002

PAGE NUMBERING AS ORIGINAL

ABSTRACT

In Electrical Impedance Tomography the conductivity of an object is estimated from boundary measurements. An array of electrodes is attached to the surface of the object and current stimuli are applied via these electrodes. The resulting voltages are measured. The process of estimating the conductivity as a function of space inside the object from voltage measurements at the surface is called reconstruction. Mathematically the EIT reconstruction is a non-linear inverse problem, the stable solution of which requires regularisation methods.

Most common regularisation methods impose that the reconstructed image should be smooth. Such methods confer stability to the reconstruction process, but limit the capability of describing sharp variations in the sought parameter.

In this thesis two new methods of regularisation are proposed. The first method, Gaussian anisotropic regularisation, enhances the reconstruction of sharp conductivity changes occurring at the interface between a contrasting object and the background. As such changes are step changes, reconstruction with traditional smoothing regularisation techniques is unsatisfactory. The Gaussian anisotropic filtering works by incorporating prior structural information. The approximate knowledge of the shapes of contrasts allows us to relax the smoothness in the direction normal to the expected boundary. The construction of Gaussian regularisation filters that express such directional properties on the basis of the structural information is discussed, and the results of numerical experiments are analysed. The method gives good results when the actual conductivity distribution is in accordance with the prior information. When the conductivity distribution violates the prior information the method is still capable of properly locating the regions of contrast.

The second part of the thesis is concerned with regularisation via the total variation functional. This functional allows the reconstruction of discontinuous parameters. The properties of the functional are briefly introduced, and an application in inverse problems in image denoising is shown. As the functional is non-differentiable, numerical difficulties are encountered in its use. The aim is therefore to propose an efficient numerical implementation for application in EIT. Several well-known optimisation methods are analysed, as possible candidates, by theoretical considerations and by numerical experiments. Such methods are shown to be inefficient. The application of recent optimisation methods called primal-dual interior-point-methods is analysed be theoretical considerations and by numerical experiments, and an efficient and stable algorithm is developed. Numerical experiments demonstrate the capability of the algorithm in reconstructing sharp conductivity profiles.

ACKNOWLEDGEMENTS

I wish to sincerely thank and express my gratitude to my supervisors Chris McLeod, at Oxford Brookes University, and Bill Lionheart, at UMIST. I'm grateful to them for their support and encouragement throughout this project. Both of them have been always enthusiastic about my work and have introduced me in the most positive way to their colleagues in the research community. I have felt that for many aspects the help they provided was beyond their duties.

I'm thankful to Oxford Brookes University for the financial support provided to me. I express my gratitude to the people at the School of Engineering and School of Computing and Mathematical Sciences with whom I have been in contact.

I would like to give my personal thanks to Nick Polydorides, fellow postgraduate, with whom we shared a supervisor, some moments of despair and many positive conversations. I has been a pleasure to find that the EIT research community in UK and abroad was so friendly. I am grateful to the researchers from Kuopio University in Finland; it has been always pleasant to meet them.

I would like to thank the heterogeneous past and present DM25 community: Alex, Amr, Andreas, David, Dimitar, Joshua, Nacer, Pong, Rui, Sveta and Thomas for the pleasant sharing of our office. I am thankful to them for extending their friendship outside work. With many of them we enjoyed pleasant moments, gatherings and even some skiing on the Alps. I would like to express my gratitude for this, as I will remember these years with joy.

Finally, and most of all, I would like to tank my fiancee Elena for sharing the "adventure" in England and my family, that supported me in many ways during the studies, and to whom this thesis is dedicated.

To my parents

Contents

Chapter 1

Introduction

This thesis discusses reconstruction methods in Electrical Impedance Tomography (EIT). EIT is a measurement technique that estimates the conductivity $\sigma(x, y, z)$ of an object from boundary measurements. An array of electrodes is attached to the body under measurement, and some currents are applied to the electrodes. The stimulation results in some electric potentials at the electrodes, which are measured. The process of estimating the conductivity of the object from such measurements is called tomographic reconstruction, or simply reconstruction.

The collected measurements are linked to the object's conductivity by known physical laws [1]. Such a relationship is called a forward operator. The inverse operator, which given the measurements would return the conductivity, is not known in the general case. The widely adopted method for reconstructing the data is to use a non-linear least squares approach. Such an approach is based on a forward model, that given a conductivity distribution calculates the corresponding theoretical measurements. Then, with optimisation techniques, the conductivity distribution for which the corresponding simulated measurements fit best the real measurements (in the least squares sense) is sought.

From the mathematical point of view this estimation task is a non-linear inverse problem. The task is to recover a coefficient (the conductivity) of an elliptic partial differential equation from data on the boundary. The problem is *ill-posed* in the sense that small perturbations in the measured data can cause arbitrarily large errors in the estimated conductivity. The developments of the technique, from the mathematical point of view, began with a publication by Calderón [2], where the uniqueness of the EIT inverse problem was firstly addressed. Other publications on the uniqueness of the inverse problem in the isotropic case followed [3] [4] [5] [6]. On the other hand it was demonstrated that the problem is not unique for anisotropic conductivities [7] [8].

Given the ill-posedness of the inverse problem, the optimisation process used in the reconstruction has to adopt particular techniques in order to obtain a stable solution. In practical terms, such ill-conditioning arises form certain patterns of conductivity for which the corresponding measurements are extremely small [9], and which are therefore affected by measurement noise. Such patterns of conductivities, for which the observations are unreliable, corrupt the reconstructions. Particular techniques, called regularisation techniques, are adopted to discourage the presence of such patterns in the image.

The use of such techniques is equivalent to introducing *a priori* information in the reconstruction process. Most common regularisation techniques impose that the inverse solution should be *smooth*. As the undesired patterns are known to be rich in high spatial frequencies, this is a way of preventing them from participating in the reconstructions. Such techniques therefore confer stability to the inversion process, but affect the capability of describing steep variations in the sought parameter.

In many applications of EIT, the objects under measurement are known to be *non-smooth*. Reconstruction of such conductivity profiles with traditional regularisation techniques is therefore unsatisfactory. In such cases different forms of prior information should be used. Information that can be incorporated into the reconstruction process is, for example, the knowledge of the boundary shape of an expected region of contrast (like an organ in medical imaging), or simply the knoweledge that steep variations are expected in the conductivity (as at the interface of two different media). By using additional information in the reconstruction process, tailored to the specific situation, it may be possible therefore to improve the estimates. Techniques for the incorporation of such forms of structural prior information are fairly recent [10] [11], and are the subject of active research. The principal topic of this thesis is the proposition of novel methods to exploiting such kinds of prior information.

For what concerns the applications of the technique, three principal areas can be identified: medical imaging, industrial process imaging and geophysical surveying [12]. In medicine different organs present different resistivities [13] and so they are distinguishable by electrical means. Secondly, organs can change their resistivities during their physiological activity, allowing functional imaging. Examples of applications in medicine are: monitoring of pulmonary and cardiac functions [14] [15] [16] [17] [18] [19], detection of brain activity [20] [21] [22], detection of haemorrhage [23] [24] [25] and gastric imaging [26] [27], to cite a few.

Applications in industry include the imaging of mixing vessels and of flow of fluids in pipes [28] [29] [30], crack detection [31] [32], and many others. In geophysics the technique is very well developed, and many applications exist at research and commercial levels [33].

Aims and Contents of the Thesis

The aim of this thesis is to develop reconstruction methods that exploit prior information in order to estimate conductivity profiles with sharp variations. In the context of the EIT research group at Oxford Brookes University such methods are desirable for the reconstruction of thoracic images, where the conductivity of the three main imaged organs, lungs, heart and the surrounding tissue, differs considerably [13]: sharp variations are expected at the organs' interfaces. In a broader view, such methods are useful in industrial and geophysical applications, where, for example, large air bubbles need to be imaged against a conducting liquid or a burial chamber must be located against conductive soil.

All the reconstruction methods that are proposed in this thesis are applied to the two–dimensional case. For the anisotropic regularisation method, presented in Chapter 5 the extension to the three–dimensional case is possible, with some modifications. For the total variation regularisation method, presented in Chapter 7, the extension to 3D is straightforward. Details of such extensions are discussed at the end of each chapter.

The approach of the present work to the reconstruction problem is deterministic, the inverse solution is formulated as a non–linear least squares problem, and Tikhonov style regularisation is used to stabilise the algorithms. The equivalence with methods based on the Bayesian approach is often mentioned, especially in Chapter 5, where some of the techniques are motivated from a statistical point of view.

The present work is organised in eight chapters, the first being the introduction. In Chapter 2 a review of the theory of inverse problems is given. The approach is formulated first from the probabilistic point of view, as it constitutes a unifying theory for reconstruction. The deterministic inversion follows and its equivalence to particular statistical methods is shown. The chapter discusses briefly the ill-conditioning of inverse problems and regularisation techniques.

In Chapter 3 the EIT specific topics needed by the following sections of the thesis are introduced. The topics of this chapter regard data acquisition strategies, 2D/3D issues and absolute/differential/dynamic techniques.

Chapter 4 is dedicated to modelling EIT experiments, and to techniques for calculating the forward solution. In the first part of the chapter physical laws that describe the experiments, and models for the electrodes are discussed. The second part of the chapter describes the numerical solution of such models via the Finite Element Method. The realisation of a forward solver is discussed; results from simulations and tests are shown. Results pertaining to these subjects have been published in [34].

Chapter 5 proposes a novel technique for incorporating structural prior information into reconstruction. The technique, called anisotropic Gaussian smoothing, enhances reconstruction of sharp conductivity profiles when the boundary of such profiles is approximately known. The method works by relaxing the smoothness constraints in the direction normal to the expected discontinuities. The approach is justified from the Bayesian point of view. The effect of the introduction of prior information into regularisation is analysed via GSVD decomposition. Results are explained on the basis of the new structure that the generalised singular vectors assume. Results pertaining to the Gaussian anisotropic regularisation were published in [35].

Chapter 6 is dedicated to introducing the use of the total variation as a regularisation functional. The use of such a functional, as a regularisation term, allows reconstruction of discontinuous parameters in inverse problems. Some theoretical properties of the regularisation functional are discussed. Through simulations it is shown that it is possible to solve the denoising inverse problem, and restore images affected by high–frequency perturbations, while retaining the sharp features of the image. Numerical challenges associated with such a method are discussed.

Chapter 7 is dedicated to the application of the total variation (TV) regularisation in EIT. The main issue addressed by the chapter is the efficient and stable solution of the TV regularised problem, as numerical difficulties are associated with such problem. The first part of the chapter analyses, by theoretical considerations and by numerical experiments, the application of well known methods of optimisation. Given the inefficiency of such methods the novel application of new optimisation techniques in the field of EIT is discussed from the theoretical point of view and by simulations. Such new techniques, called primal–dual interior–point–methods, allow the formulation of a framework for efficient EIT regularisation with TV and more in general with ℓ_1–norm functionals. Results pertaining to the use of TV regularisation in EIT have been published in [36].

The last chapter is a discussion section; results of the thesis are discussed as a whole. Suggestions for further development are given.

Chapter 2

Discrete Inverse Problems

2.1 Introduction

In many fields of science and engineering it is desirable to infer properties of a system from experimental observations. Assume that we are able to define a set of *model parameters* that completely describe the properties of the system, to the extent required by the application. These parameters may not be directly measurable, but we might be able to define some experiments that allow us to measure some *observable parameters* whose actual values depend on the values of the model parameters. To solve the *forward problem* is to predict the values of the observable parameters, given arbitrary values of the model parameters. To solve the *inverse problem* is to infer the values of the model parameters from the measured values of the observable parameters.

Methods for solving inverse problems can be categorised as statistical or as deterministic. The statistical approach to inverse problems allows formulation of a unifying *theory of inverse problems* [37], from which deterministic approaches stem. In this chapter some of the main topics on inverse problems are reviewed, introducing first the statistical approach, second showing that deterministic methods can be interpreted from the statistical point of view.

2.2 Statistical Inversion

Probability axioms

In this section we introduce some notations of probability theory that will be used in the following. Let $\mathcal{K} \in \mathbb{R}^K$ be an arbitrary space. A measure over $\mathcal{K}$ is a rule that to any subset $\mathcal{K}_i$ of $\mathcal{K}$ associates a positive real number $P(\mathcal{K}_i)$, named *measure of* $\mathcal{K}_i$, and that

satisfies the following properties

- $P(\emptyset) = 0$
- If $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n$ are disjoint sets in $\mathcal{K}$, then

$$P\left(\sum_i \mathcal{K}_i\right) = \sum_i P(\mathcal{K}_i) \tag{2.1}$$

If $P(\mathcal{K})$ is finite, P is termed *probability*.

Let $\mathbf{k} = (\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_K)^T$ be the generic vector of $\mathcal{K}$. A function $p(\mathbf{k})$ such that for any subset $\mathcal{K}_i \subset \mathcal{K}$

$$P(\mathcal{K}_i) = \int_{\mathcal{K}_i} p(\mathbf{k})\, \mathrm{d}\mathbf{k} \tag{2.2}$$

is called *probability density function*, and it can be normalised by imposing $P(\mathcal{K}) = 1$. The *expectation* of $\mathbf{k}$ is defined as

$$\eta_{\mathbf{k}} = E\{\mathbf{k}\} = \int_{\mathcal{K}} \mathbf{k}\, p(\mathbf{k})\, \mathrm{d}\mathbf{k} \tag{2.3}$$

and the *covariance* of $\mathbf{k}$ as

$$C_{\mathbf{k}} = E\{(\mathbf{k} - \eta_{\mathbf{k}})(\mathbf{k} - \eta_{\mathbf{k}})^T\} \tag{2.4}$$

Let now $\mathbf{x}$ and $\mathbf{y}$ be two vectors of two arbitrary spaces $\mathcal{X}$ *in* $\mathbb{R}^X$ and $\mathcal{Y}$ *in* $\mathbb{R}^Y$. A function $p(\mathbf{x}, \mathbf{y})$ such that for any subset $\mathcal{X}_i \subset \mathcal{X}$ and any subset $\mathcal{Y}_i \subset \mathcal{Y}$

$$P(\mathcal{X}_i, \mathcal{Y}_i) = \int_{\mathcal{X}_i, \mathcal{Y}_i} p(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} \tag{2.5}$$

is called *joint probability density function*, and it can be normalised by imposing $P(\mathcal{X}, \mathcal{Y}) = 1$.

Let $p(\mathbf{x}, \mathbf{y})$ be a normalised joint probability density function of the couple $(\mathbf{x}, \mathbf{y})$. The *marginal probability density* for x is defined as

$$p(\mathbf{x}) = \int_{\mathcal{Y}} p(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{y} \tag{2.6}$$

and the *conditional probability density* for $\mathbf{x}$ given $\mathbf{y} = \mathbf{y}$ is defined as

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{p(\mathbf{x},\mathbf{y})}{\int_{\mathcal{X}} p(\mathbf{x},\mathbf{y})\, \mathrm{d}\mathbf{x}} \tag{2.7}$$

From these definitions follow that the joint probability density equals the conditional probability density times the marginal probability density

$$p(\mathbf{x} \mid \mathbf{y})\, p(\mathbf{y}) = p(\mathbf{x},\mathbf{y}) \tag{2.8}$$

and hence Bayes theorem

$$p(\mathbf{y} \mid \mathbf{x}) p(x) = p(\mathbf{x} \mid \mathbf{y})\, p(\mathbf{y}) \tag{2.9}$$

Bayes interpretation of probability

It is possible to associate more than one meaning to any mathematical theory. The first interpretation of the axioms of Section 2.2 is purely statistical: a *random process* takes place and leads to several *realisations*. The parameters that describe the realisations are called *random variables*. If the number of realisations that have been observed is large, these can be described in terms of probabilities, which follow the laws of 2.2, and no ulterior meaning is associated with them. A second approach, called *Bayesian*, interprets probabilities in terms of *degree of knowledge* of the "true" values of a physical parameter. Let $\mathcal{K}$ be a generic discrete parameter set with a finite number of parameters. The *state of information* on $\mathcal{K}$ is described by defining a probability density over $\mathcal{K}$. If we definitely know that the true value of $\mathbf{k}$ is $\mathbf{k} = \mathbf{k}_0$, then the corresponding probability density is

$$p(\mathbf{k}) = \delta(\mathbf{k} - \mathbf{k}_0) \tag{2.10}$$

which represents the state of *perfect knowledge*, and where δ is the Dirac delta function in $\mathbb{R}^K$. The use of probabilities to describe the degree of knowledge is supported by the fact that real observations are always subject to uncertainties. In practice states of information will be defined by probability densities with some dispersion, which represent the uncertainty of the knowledge. The lowest state of information is the state of *total ignorance* defined by the *non-informative* probability density. The mathematical expression of non informative probability density depends on the system of coordinates in use, for a Cartesian system with $\mathcal{K} \in \mathbb{R}^K$ and finite, it is

$$p_{\text{n.i.}}(\mathbf{k}) = \text{constant} \tag{2.11}$$

Parameter spaces

Model Space Given a physical system under study, after a discretisation, the properties of interest can be described with a vector m of model parameters. Depending on the system under study, the model parameters make take their values in a discrete or in a continuous set. In the following we will assume that the parameters take their values in a continuous set of conceivable model "configurations" $\mathcal{M}$, and that $\mathcal{M} \in \mathbb{R}^M$.

Data Space Given a system under study and having defined a model space $\mathcal{M}$, some experiments can be performed in order to gather information on the model parameters m. The experiments must be designed to carry information about the model parameters. The collected data vector d will take values in a continuous set of conceivable instrument responses $\mathcal{D}$, and $\mathcal{D} \in \mathbb{R}^D$.

Physical Parameters Space The couples $\mathbf{f} = (\mathbf{m}, \mathbf{d})$ define the joint space $\mathcal{F} = \mathcal{M} \times \mathcal{D}$ that holds the model and data parameters.

Sources of information

Three sources provide information in inverse problems: physical laws, results of measurements and a priori information. Physical laws apply to the system under study. They are used to derive a model, called the *forward model*. The forward model holds the relationships that link the model parameters $\mathbf{m}$ to the measured data $\mathbf{d}$. The forward model enables the solution of the forward problem: to predict the error-free values of data, $\mathbf{d}$, that would correspond to a given model $\mathbf{m}$. The predicted values, in general, cannot be identical to the observed values, as uncertainties are acting on the measurements and predictions are affected by modelling errors. These two sources of errors, even if different, produce errors of the same magnitude. Models are developed to match instrumentation errors. It wouldn't be wise to use a poor model for reconstructing data from a good instrument, as it wouldn't be useful to have a very accurate model, but a poor instrument. Following the postulate that the more general way to describe any state of information is to define a probability density function, the forward model and the relative uncertainties can then be described by defining, for given values of the model parameters $\mathbf{m}$, a probability density $\Theta(\mathbf{d} \mid \mathbf{m})$ over $\mathcal{D}$ (called *likelihood*).

Results from experiments are the second source of information in inverse problems. Measurements will give a certain amount of information about the true values of the observable parameters. This state of information can be described with the probability

Chapter 1

Introduction

This thesis discusses reconstruction methods in Electrical Impedance Tomography (EIT). EIT is a measurement technique that estimates the conductivity $\sigma(x, y, z)$ of an object from boundary measurements. An array of electrodes is attached to the body under measurement, and some currents are applied to the electrodes. The stimulation results in some electric potentials at the electrodes, which are measured. The process of estimating the conductivity of the object from such measurements is called tomographic reconstruction, or simply reconstruction.

The collected measurements are linked to the object's conductivity by known physical laws [1]. Such a relationship is called a forward operator. The inverse operator, which given the measurements would return the conductivity, is not known in the general case. The widely adopted method for reconstructing the data is to use a non-linear least squares approach. Such an approach is based on a forward model, that given a conductivity distribution calculates the corresponding theoretical measurements. Then, with optimisation techniques, the conductivity distribution for which the corresponding simulated measurements fit best the real measurements (in the least squares sense) is sought.

From the mathematical point of view this estimation task is a non-linear inverse problem. The task is to recover a coefficient (the conductivity) of an elliptic partial differential equation from data on the boundary. The problem is *ill-posed* in the sense that small perturbations in the measured data can cause arbitrarily large errors in the estimated conductivity. The developments of the technique, from the mathematical point of view, began with a publication by Calderón [2], where the uniqueness of the EIT inverse problem was firstly addressed. Other publications on the uniqueness of the inverse problem in the isotropic case followed [3] [4] [5] [6]. On the other hand it was demonstrated that the problem is not unique for anisotropic conductivities [7] [8].

Given the ill-posedness of the inverse problem, the optimisation process used in the reconstruction has to adopt particular techniques in order to obtain a stable solution. In practical terms, such ill-conditioning arises form certain patterns of conductivity for which the corresponding measurements are extremely small [9], and which are therefore affected by measurement noise. Such patterns of conductivities, for which the observations are unreliable, corrupt the reconstructions. Particular techniques, called regularisation techniques, are adopted to discourage the presence of such patterns in the image.

The use of such techniques is equivalent to introducing *a priori* information in the reconstruction process. Most common regularisation techniques impose that the inverse solution should be *smooth*. As the undesired patterns are known to be rich in high spatial frequencies, this is a way of preventing them from participating in the reconstructions. Such techniques therefore confer stability to the inversion process, but affect the capability of describing steep variations in the sought parameter.

In many applications of EIT, the objects under measurement are known to be *non-smooth*. Reconstruction of such conductivity profiles with traditional regularisation techniques is therefore unsatisfactory. In such cases different forms of prior information should be used. Information that can be incorporated into the reconstruction process is, for example, the knowledge of the boundary shape of an expected region of contrast (like an organ in medical imaging), or simply the knoweledge that steep variations are expected in the conductivity (as at the interface of two different media). By using additional information in the reconstruction process, tailored to the specific situation, it may be possible therefore to improve the estimates. Techniques for the incorporation of such forms of structural prior information are fairly recent [10] [11], and are the subject of active research. The principal topic of this thesis is the proposition of novel methods to exploiting such kinds of prior information.

For what concerns the applications of the technique, three principal areas can be identified: medical imaging, industrial process imaging and geophysical surveying [12]. In medicine different organs present different resistivities [13] and so they are distinguishable by electrical means. Secondly, organs can change their resistivities during their physiological activity, allowing functional imaging. Examples of applications in medicine are: monitoring of pulmonary and cardiac functions [14] [15] [16] [17] [18] [19], detection of brain activity [20] [21] [22], detection of haemorrhage [23] [24] [25] and gastric imaging [26] [27], to cite a few.

Applications in industry include the imaging of mixing vessels and of flow of fluids in pipes [28] [29] [30], crack detection [31] [32], and many others. In geophysics the technique is very well developed, and many applications exist at research and commercial levels [33].

Aims and Contents of the Thesis

The aim of this thesis is to develop reconstruction methods that exploit prior information in order to estimate conductivity profiles with sharp variations. In the context of the EIT research group at Oxford Brookes University such methods are desirable for the reconstruction of thoracic images, where the conductivity of the three main imaged organs, lungs, heart and the surrounding tissue, differs considerably [13]: sharp variations are expected at the organs' interfaces. In a broader view, such methods are useful in industrial and geophysical applications, where, for example, large air bubbles need to be imaged against a conducting liquid or a burial chamber must be located against conductive soil.

All the reconstruction methods that are proposed in this thesis are applied to the two–dimensional case. For the anisotropic regularisation method, presented in Chapter 5 the extension to the three–dimensional case is possible, with some modifications. For the total variation regularisation method, presented in Chapter 7, the extension to 3D is straightforward. Details of such extensions are discussed at the end of each chapter.

The approach of the present work to the reconstruction problem is deterministic, the inverse solution is formulated as a non–linear least squares problem, and Tikhonov style regularisation is used to stabilise the algorithms. The equivalence with methods based on the Bayesian approach is often mentioned, especially in Chapter 5, where some of the techniques are motivated from a statistical point of view.

The present work is organised in eight chapters, the first being the introduction. In Chapter 2 a review of the theory of inverse problems is given. The approach is formulated first from the probabilistic point of view, as it constitutes a unifying theory for reconstruction. The deterministic inversion follows and its equivalence to particular statistical methods is shown. The chapter discusses briefly the ill-conditioning of inverse problems and regularisation techniques.

In Chapter 3 the EIT specific topics needed by the following sections of the thesis are introduced. The topics of this chapter regard data acquisition strategies, 2D/3D issues and absolute/differential/dynamic techniques.

Chapter 4 is dedicated to modelling EIT experiments, and to techniques for calculating the forward solution. In the first part of the chapter physical laws that describe the experiments, and models for the electrodes are discussed. The second part of the chapter describes the numerical solution of such models via the Finite Element Method. The realisation of a forward solver is discussed; results from simulations and tests are shown. Results pertaining to these subjects have been published in [34].

Chapter 5 proposes a novel technique for incorporating structural prior information into reconstruction. The technique, called anisotropic Gaussian smoothing, enhances reconstruction of sharp conductivity profiles when the boundary of such profiles is approximately known. The method works by relaxing the smoothness constraints in the direction normal to the expected discontinuities. The approach is justified from the Bayesian point of view. The effect of the introduction of prior information into regularisation is analysed via GSVD decomposition. Results are explained on the basis of the new structure that the generalised singular vectors assume. Results pertaining to the Gaussian anisotropic regularisation were published in [35].

Chapter 6 is dedicated to introducing the use of the total variation as a regularisation functional. The use of such a functional, as a regularisation term, allows reconstruction of discontinuous parameters in inverse problems. Some theoretical properties of the regularisation functional are discussed. Through simulations it is shown that it is possible to solve the denoising inverse problem, and restore images affected by high-frequency perturbations, while retaining the sharp features of the image. Numerical challenges associated with such a method are discussed.

Chapter 7 is dedicated to the application of the total variation (TV) regularisation in EIT. The main issue addressed by the chapter is the efficient and stable solution of the TV regularised problem, as numerical difficulties are associated with such problem. The first part of the chapter analyses, by theoretical considerations and by numerical experiments, the application of well known methods of optimisation. Given the inefficiency of such methods the novel application of new optimisation techniques in the field of EIT is discussed from the theoretical point of view and by simulations. Such new techniques, called primal–dual interior–point–methods, allow the formulation of a framework for efficient EIT regularisation with TV and more in general with ℓ_1–norm functionals. Results pertaining to the use of TV regularisation in EIT have been published in [36].

The last chapter is a discussion section; results of the thesis are discussed as a whole. Suggestions for further development are given.

Chapter 2

Discrete Inverse Problems

2.1 Introduction

In many fields of science and engineering it is desirable to infer properties of a system from experimental observations. Assume that we are able to define a set of *model parameters* that completely describe the properties of the system, to the extent required by the application. These parameters may not be directly measurable, but we might be able to define some experiments that allow us to measure some *observable parameters* whose actual values depend on the values of the model parameters. To solve the *forward problem* is to predict the values of the observable parameters, given arbitrary values of the model parameters. To solve the *inverse problem* is to infer the values of the model parameters from the measured values of the observable parameters.

Methods for solving inverse problems can be categorised as statistical or as deterministic. The statistical approach to inverse problems allows formulation of a unifying *theory of inverse problems* [37], from which deterministic approaches stem. In this chapter some of the main topics on inverse problems are reviewed, introducing first the statistical approach, second showing that deterministic methods can be interpreted from the statistical point of view.

2.2 Statistical Inversion

Probability axioms

In this section we introduce some notations of probability theory that will be used in the following. Let $\mathcal{K} \in \mathbb{R}^K$ be an arbitrary space. A measure over $\mathcal{K}$ is a rule that to any subset $\mathcal{K}_i$ of $\mathcal{K}$ associates a positive real number $P(\mathcal{K}_i)$, named *measure of* $\mathcal{K}_i$, and that

satisfies the following properties

- $P(\emptyset) = 0$
- If $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n$ are disjoint sets in $\mathcal{K}$, then

$$P\left(\sum_i \mathcal{K}_i\right) = \sum_i P(\mathcal{K}_i) \tag{2.1}$$

If $P(\mathcal{K})$ is finite, P is termed *probability.*

Let $\mathbf{k} = (\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_K)^T$ be the generic vector of $\mathcal{K}$. A function $p(\mathbf{k})$ such that for any subset $\mathcal{K}_i \subset \mathcal{K}$

$$P(\mathcal{K}_i) = \int_{\mathcal{K}_i} p(\mathbf{k})\, \mathrm{d}\mathbf{k} \tag{2.2}$$

is called *probability density function*, and it can be normalised by imposing $P(\mathcal{K}) = 1$. The *expectation* of $\mathbf{k}$ is defined as

$$\boldsymbol{\eta}_\mathbf{k} = E\{\mathbf{k}\} = \int_{\mathcal{K}} \mathbf{k}\, p(\mathbf{k})\, \mathrm{d}\mathbf{k} \tag{2.3}$$

and the *covariance* of $\mathbf{k}$ as

$$C_\mathbf{k} = E\{(\mathbf{k} - \boldsymbol{\eta}_\mathbf{k})(\mathbf{k} - \boldsymbol{\eta}_\mathbf{k})^T\} \tag{2.4}$$

Let now $\mathbf{x}$ and $\mathbf{y}$ be two vectors of two arbitrary spaces $\mathcal{X}$ *in* $\mathbb{R}^X$ and $\mathcal{Y}$ *in* $\mathbb{R}^Y$. A function $p(\mathbf{x}, \mathbf{y})$ such that for any subset $\mathcal{X}_i \subset \mathcal{X}$ and any subset $\mathcal{Y}_i \subset \mathcal{Y}$

$$P(\mathcal{X}_i, \mathcal{Y}_i) = \int_{\mathcal{X}_i, \mathcal{Y}_i} p(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} \tag{2.5}$$

is called *joint probability density function*, and it can be normalised by imposing $P(\mathcal{X}, \mathcal{Y}) = 1$.

Let $p(\mathbf{x}, \mathbf{y})$ be a normalised joint probability density function of the couple $(\mathbf{x}, \mathbf{y})$. The *marginal probability density* for x is defined as

$$p(\mathbf{x}) = \int_{\mathcal{Y}} p(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{y} \tag{2.6}$$

and the *conditional probability density* for $\mathbf{x}$ given $\mathbf{y} = \mathbf{y}$ is defined as

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{\int_{\mathcal{X}} p(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}} \tag{2.7}$$

From these definitions follow that the joint probability density equals the conditional probability density times the marginal probability density

$$p(\mathbf{x} \mid \mathbf{y})\, p(\mathbf{y}) = p(\mathbf{x}, \mathbf{y}) \tag{2.8}$$

and hence Bayes theorem

$$p(\mathbf{y} \mid \mathbf{x}) p(x) = p(\mathbf{x} \mid \mathbf{y})\, p(\mathbf{y}) \tag{2.9}$$

Bayes interpretation of probability

It is possible to associate more than one meaning to any mathematical theory. The first interpretation of the axioms of Section 2.2 is purely statistical: a *random process* takes place and leads to several *realisations*. The parameters that describe the realisations are called *random variables*. If the number of realisations that have been observed is large, these can be described in terms of probabilities, which follow the laws of 2.2, and no ulterior meaning is associated with them. A second approach, called *Bayesian*, interprets probabilities in terms of *degree of knowledge* of the "true" values of a physical parameter. Let $\mathcal{K}$ be a generic discrete parameter set with a finite number of parameters. The *state of information* on $\mathcal{K}$ is described by defining a probability density over $\mathcal{K}$. If we definitely know that the true value of $\mathbf{k}$ is $\mathbf{k} = \mathbf{k}_0$, then the corresponding probability density is

$$p(\mathbf{k}) = \delta(\mathbf{k} - \mathbf{k}_0) \tag{2.10}$$

which represents the state of *perfect knowledge*, and where δ is the Dirac delta function in $\mathbb{R}^K$. The use of probabilities to describe the degree of knowledge is supported by the fact that real observations are always subject to uncertainties. In practice states of information will be defined by probability densities with some dispersion, which represent the uncertainty of the knowledge. The lowest state of information is the state of *total ignorance* defined by the *non-informative* probability density. The mathematical expression of non informative probability density depends on the system of coordinates in use, for a Cartesian system with $\mathcal{K} \in \mathbb{R}^K$ and finite, it is

$$p_{\text{n.i.}}(\mathbf{k}) = \text{constant} \tag{2.11}$$

Parameter spaces

Model Space Given a physical system under study, after a discretisation, the properties of interest can be described with a vector m of model parameters. Depending on the system under study, the model parameters make take their values in a discrete or in a continuous set. In the following we will assume that the parameters take their values in a continuous set of conceivable model "configurations" $\mathcal{M}$, and that $\mathcal{M} \in \mathbb{R}^M$.

Data Space Given a system under study and having defined a model space $\mathcal{M}$, some experiments can be performed in order to gather information on the model parameters m. The experiments must be designed to carry information about the model parameters. The collected data vector d will take values in a continuous set of conceivable instrument responses $\mathcal{D}$, and $\mathcal{D} \in \mathbb{R}^D$.

Physical Parameters Space The couples $\mathbf{f} = (\mathbf{m}, \mathbf{d})$ define the joint space $\mathcal{F} = \mathcal{M} \times \mathcal{D}$ that holds the model and data parameters.

Sources of information

Three sources provide information in inverse problems: physical laws, results of measurements and a priori information. Physical laws apply to the system under study. They are used to derive a model, called the *forward model.* The forward model holds the relationships that link the model parameters $\mathbf{m}$ to the measured data $\mathbf{d}$. The forward model enables the solution of the forward problem: to predict the error-free values of data, $\mathbf{d}$, that would correspond to a given model $\mathbf{m}$. The predicted values, in general, cannot be identical to the observed values, as uncertainties are acting on the measurements and predictions are affected by modelling errors. These two sources of errors, even if different, produce errors of the same magnitude. Models are developed to match instrumentation errors. It wouldn't be wise to use a poor model for reconstructing data from a good instrument, as it wouldn't be useful to have a very accurate model, but a poor instrument. Following the postulate that the more general way to describe any state of information is to define a probability density function, the forward model and the relative uncertainties can then be described by defining, for given values of the model parameters $\mathbf{m}$, a probability density $\Theta(\mathbf{d} \mid \mathbf{m})$ over $\mathcal{D}$ (called *likelihood*).

Results from experiments are the second source of information in inverse problems. Measurements will give a certain amount of information about the true values of the observable parameters. This state of information can be described with the probability

density $\Pi_D(\mathbf{d})$ over the data space.

Lastly, prior information on the model parameters, which is obtained independently from measurements results, is an additional source of information. The probability density that describes this state of information will be denoted $\Pi_M(\mathbf{m})$, and it is defined on the model space. If no prior information is known, then $\Pi_M(\mathbf{m}) = p_{\text{n.i.}}(\mathbf{m})$; the general case is however that some prior information is given. A common form of prior information are constraints on admissible values for the model parameters. In EIT, for example, the conductivity values are known to be positive. Further constraints on model values might come from knowledge of the composition of imaged object. Consider for example a system under study that is composed of a finite number of materials, each with a known conductivity. The model parameters could be constrained to take values in a discrete set of $\mathbb{R}^+$, by defining an opportune probability density over $\mathcal{M}$.

Solution of the inverse problem

All the available sources of information have been described with the three states of information $\Theta(\mathbf{m} \mid \mathbf{d})$, $\Pi_D(\mathbf{d})$ and $\Pi_M(\mathbf{m})$ defined on the spaces $\mathcal{D}$ and $\mathcal{M}$. In general we are interested in "translating" information from $\mathcal{D}$ to $\mathcal{M}$. The Bayesian approach suggests doing so by defining the conditional probability density $\Xi(\mathbf{m} \mid \mathbf{d})$ over the model space, which thanks to (2.9) is

$$\Xi(\mathbf{m} \mid \mathbf{d}) = \frac{\Theta(\mathbf{d} \mid \mathbf{m})\Pi_M(\mathbf{m})}{\Pi_D(\mathbf{d})} \tag{2.12}$$

Equation (2.12) expresses the posterior state of information on the model space as a combination of three sources of information: it can be labelled as "the solution" of the general inverse problem. From $\Xi(\mathbf{m} \mid \mathbf{d})$ it is possible to extract information on the model parameters in many forms, as discussed in the next section.

The analysis of existence and uniqueness of the solution is straightforward : the solution exists if $\Xi(\mathbf{m} \mid \mathbf{d})$ is not identically null over $\mathcal{M}$. If this was the case, it would indicate that the theoretical and prior information are not compatible with the experimental results. The uniqueness of the solution, if by solution we mean the state of information $\Xi(\mathbf{m} \mid \mathbf{d})$, follows from the definition of conditioned probability density, which defines uniquely $\Xi(\mathbf{m} \mid \mathbf{d})$ by (2.12). It may happen of course that $\Xi(\mathbf{m} \mid \mathbf{d})$ is, for example, multimodal or non informative on certain parameters, the state of information itself is however uniquely defined.

Using the information

The most general way of studying the information obtained on the model parameters is by direct study of the probability density $\Xi(\mathbf{m} \mid \mathbf{d})$. A comprehensive way to do so is to study the probability that the true values of the model parameters lie in a given range $\mathcal{M}_1 \in \mathcal{M}$

$$P(\mathbf{m} \in \mathcal{M}_1) = \int_{\mathcal{M}_1} \Xi(\mathbf{m} \mid \mathbf{d}) \mathrm{d}\mathbf{m} \tag{2.13}$$

By experimenting with several ranges, it is then possible to gain a good indication of the parameters. As anticipated, the model's state of information can show a complex behaviour, it could be multimodal or it could present infinite variances.

Very often we are interested in the values of the parameters per se. If the posterior probability density $\Xi(\mathbf{m} \mid \mathbf{d})$ is "well behaved", it is then possible to characterise it by its central estimators and estimators of dispersion.

Central estimators and estimators of dispersion

Central estimators and estimators of dispersion are used for characterising a state of information in an intuitive and concise manner. From the Bayesian point of view, central estimators relate to the question: "what are the values of the parameters ?". Dispersion estimators relate to the question: "what confidence do we have about the values of the parameters ?". There are several central and dispersion estimators. Amongst central estimators the *maximum a posteriori* (MAP) estimate, the *maximum likelihood* (ML) estimate and the the *minimum mean square* (MS) estimate are commonly used in inverse problems. Given the posterior probability density $\Xi(\mathbf{m} \mid \mathbf{d})$ defined for the parameters $\mathbf{m} \in \mathcal{M}$, the maximum a posteriori estimate is given by

$$\mathbf{m}_{MAP} = \max_{\mathbf{m} \in \mathcal{M}} \Xi(\mathbf{m} \mid \mathbf{d}) \tag{2.14}$$

The estimate $\mathbf{m}_{MAP}$ is the most probable configuration of the model, given the outcome of the experiments, the prior information, and the physical laws that apply to the system. The computation of the maximum a posteriori estimate leads to an optimisation problem. The posterior probability density is not guaranteed to show a unique maximum, it could be multimodal, in this case non-trivial numerical methods are required to find the extremum.

The maximum likelihood estimate is defined as

$$\mathbf{m}_{ML} = \max_{\mathbf{m} \in \mathcal{M}} \Theta(\mathbf{d} \mid \mathbf{m}) \tag{2.15}$$

The probability density $\Theta(\mathbf{d} \mid \mathbf{m})$ is called the likelihood. It describes how likely the event is, of the experimental data being $\mathbf{d}$, given that the model configuration is $\mathbf{m}$. The estimate $\mathbf{m}_{ML}$ therefore corresponds to the model configuration which would yield the observation $\mathbf{d}$ with the highest probability. It is important to notice that the ML approach, differently from MAP, does not consider $P_M(\mathbf{m})$. As will be discussed in the following Sections this typically corresponds to solving the inverse problem with no regularisation.

The minimum mean square estimation can be calculated as the expectation of the posterior probability density

$$\mathbf{m}_{MS} = \eta_{\mathbf{m}|\mathbf{d}} = E\{\Xi(\mathbf{m} \mid \mathbf{d})\} \tag{2.16}$$

It can be shown [38] that $\mathbf{m}_{MS}$ minimises the expectation of the mean square error (MSE) of the model parameters, which is

$$\text{MSE} = \|\mathbf{m}_{MS} - \mathbf{m}\|^2 \tag{2.17}$$

In many cases, as for example in the case of Gaussian assumptions about $\Theta(\mathbf{m} \mid \mathbf{d})$, $\Pi_M(\mathbf{m})$, $\Pi_D(\mathbf{d})$, the MS estimation corresponds to the MAP estimation, as the expectation of the Gaussian distribution is numerically equal to the mode.

Having obtained a central estimator for $\mathbf{m}$ it is desirable to have indicators of its reliability. A common dispersion estimator is the *posterior covariance*

$$C_{\mathbf{m}|\mathbf{d}} = \int_{\mathcal{M}} (\mathbf{m} - \eta_{\mathbf{m}|\mathbf{d}})(\mathbf{m} - \eta_{\mathbf{m}|\mathbf{d}})^T \Xi(\mathbf{m} \mid \mathbf{d})\, \mathrm{d}\mathbf{m} \tag{2.18}$$

where $\eta_{\mathbf{m}|\mathbf{d}}$ is the expectation of the posterior probability density $\Xi(\mathbf{m} \mid \mathbf{d})$ as in (2.16). Naturally the posterior covariance gives understandable information only in the case that the posterior probability density can be fitted with good approximation to a Gaussian distribution. A more general way of estimating dispersion is to obtain confidence limits. For a single component $\mathbf{m}_i$ of the model $\mathbf{m}$ the *marginal posterior probability density* is

$$\Xi(\mathbf{m}_i \mid \mathbf{d}) = \int_{\mathcal{M}} \Xi(\mathbf{m} \mid \mathbf{d})\, \mathrm{d}\mathbf{m}_1, \ldots, \mathrm{d}\mathbf{m}_{i-1}, \mathrm{d}\mathbf{m}_{i+1}, \ldots, \mathrm{d}\mathbf{m}_M \tag{2.19}$$

The one-dimensional distribution $\Xi(\mathbf{m}_i \mid \mathbf{d})$ is informative on each single component $\mathbf{m}_i$. Confidence limits, for each parameter of the model, can then be calculated by finding the range around the parameter's estimated value that correspond to the required confidence

level P_c

$$\int_{(\tilde{\mathbf{m}}_i-\delta\mathbf{m}_i)}^{(\tilde{\mathbf{m}}_i+\delta\mathbf{m}_i)} \Xi(\mathbf{m}_i \mid \mathbf{d})\, d\mathbf{m}_i = P_c \tag{2.20}$$

where $\tilde{\mathbf{m}}_i$ is the value of a central estimator of $\mathbf{m}_i$ and $\delta\mathbf{m}_i$ gives the range $[\tilde{\mathbf{m}}_i - \delta\mathbf{m}_i, \tilde{\mathbf{m}}_i + \delta\mathbf{m}_i]$ in which the true value of $\mathbf{m}_i$ falls with a probability P_c (eg. $P_c = 0.9, P_c = 0.99$).

As discussed in the present section, the integration of the posterior conditioned probability $\Xi(\mathbf{m} \mid \mathbf{d})$ is often needed in order to extract information. Examples are the calculation of the expected values and of confidence limits. In the most general case integrating $\Xi(\mathbf{m} \mid \mathbf{d})$ over $\mathcal{M}$ requires numerical techniques. In a few cases, corresponding to particular assumptions on the modelling errors, measurements errors and prior distribution integrations can be calculated analytically. In the next two Sections numerical and analytical methods are briefly overviewed.

Numerical integration

For inverse problems with a small number of model parameters the integration of the posterior probability density can be carried out by traditional integration methods like quadrature formulas. The continuous integral of a function $\phi(\mathbf{m})$ with respect to a probability density $\pi(\mathbf{m})$

$$I = \int_{\mathcal{M}} \phi(\mathbf{m})\, \pi(\mathbf{m})\, d\mathbf{m} \tag{2.21}$$

is approximated by finite sums like

$$I \simeq \frac{1}{Z} \sum_{z=1}^{Z} \phi(\mathbf{m}^{(z)})\, \pi(\mathbf{m}^{(z)}) \tag{2.22}$$

where $\mathbf{m}^{(k)}$ are points over a regular grid defined in the parameter space. As dimensionality of the model space is often large, by exploring the domain in such a way, quadrature methods like (2.22) require an enormous number of evaluations of $\phi(\mathbf{m}^{(z)})\, \pi(\mathbf{m}^{(z)})$ to reach a given accuracy.

Markov Chain Monte Carlo (MCMC) [39] [40] methods are a successful alternative to traditional quadrature formulas for problems with a large number of model parameters. MCMC methods substitute the systematic exploration of the integration domain by a random exploration. The points $\mathbf{m}^{(z)}$ are drawn from an opportune random processes, defined by a probability density $p(\mathbf{m})$ over the space $\mathcal{M}$. If $p(\mathbf{m}) = const.$ the exploration of $\mathcal{M}$ is uniform. The goal of MCMC methods is to explore $\mathcal{M}$ drawing $\mathbf{m}^{(z)}$ from distributions for which a lower number of evaluations of the integrands is required for

a given accuracy, with respect to uniform exploration. This is achieved by sampling $\mathcal{M}$ more densely where significant values of the integrands are expected. MCMC methods use properly built Markov Chains to generate sequences of draws $\mathbf{m}^{(z)}$ that have a stationary distribution equal to $\pi(\mathbf{m})$. Suppose that a series of draws $\{\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \ldots, \mathbf{m}^{(Z)}\} \in \mathcal{M}$ is taken from such a chain, and that the series is a representative ensemble of $\pi(\mathbf{m})$ then (2.21) can be approximated as

$$I \simeq \frac{1}{Z} \sum_{z=1}^{Z} \phi(\mathbf{m}^{(z)}) \tag{2.23}$$

In practice, a certain number h of draws are taken from the Markov Chain. Being $p(\mathbf{m})$ the stationary distribution for the chain, if h is sufficiently large, the distribution of the subsequent $\mathbf{m}^{(h+1)}, \ldots, \mathbf{m}^{(h+Z)}$ draws will be a good approximation to $\pi(\mathbf{m})$. The draws $\mathbf{m}^{(h+1)}, \ldots, \mathbf{m}^{(h+Z)}$ are then used to calculate (2.23).

MCMC methods thus allow one to extract information from the a posteriori probability density $\Xi(\mathbf{m} \mid \mathbf{d})$ in the form of, for example, mean values and confidence limits. It is important to note that no assumptions need to be made on $\Xi(\mathbf{m} \mid \mathbf{d})$. The approach is therefore suitable for dealing with non-differentiable a posteriori densities like the ones deriving from ℓ_1-norm priors (see Chapter 6 and 7). Successful examples of application in Optical Tomography and Electrical Impedance Tomography are the ones by Somersalo [41] and Kolehmainen [11].

Analytic solutions

As seen in the preceding Sections, the sources of information in inverse problems can be described with the three states of information $\Theta(\mathbf{d} \mid \mathbf{m})$, $\Pi_M(\mathbf{m})$ and $\Pi_D(\mathbf{d})$. Each distribution can be modelled by an analytic expression. A common model for states of information is the *generalised Gaussian* distribution, that for a parameter $\mathbf{k}$ is defined as

$$G_n(\mathbf{k}) \propto \exp\left(- \frac{1}{n} \sum_i \frac{|\mathbf{k}_i - \mathbf{k}_{0_i})|^n}{(\sigma_{n_i})^n} \right) \tag{2.24}$$

with $n \geq 1$, and where Γ is the Gamma function and σ_n is the dispersion estimator in the ℓ_n-norm, which is defined as follows

$$(\sigma_n)^n = \int_{\mathcal{K}} |\mathbf{k} - \mathbf{k}_0|^n \, p(\mathbf{k}) \, \mathrm{d}\mathbf{k} \tag{2.25}$$

Plots of the generalised Gaussian distribution, for various values of n, are shown in Figure 2.1. For $n = 2$ the generalised Gaussian is a Gaussian distribution. For $n < 2$ the

generalised Gaussian tends to zero, for $\mathbf{k} \to \infty$, less rapidly than for $n = 2$. The family of distributions $G_n(\mathbf{k})$ with $n < 2$ are therefore called "long tailed" distributions, and are used to model data that presents samples with a possibly very large deviation from the central estimator. Data points with these properties are called *outliers*. Their presence in a data set $\mathbf{d}$ usually does not arise from instrument noise, for which the dispersion is supposed to be "small". Outliers are caused from accidental errors, such as an electrode not making proper contact in electrical impedance tomography. They represent significantly larger errors in comparison to the ones introduced by the measurement system, and they are (hopefully) small in number in comparison to the whole data set. Data sets containing outliers are often modelled with $\Pi_D(\mathbf{d}) = G_1(\mathbf{d})$, which is a symmetric decaying exponential centered in $\mathbf{k}_0$. Reconstruction algorithms that assume such distributions on the input data are called *robust algorithms* for their smaller sensitivity to outliers than algorithms that assume $\Pi_D(\mathbf{d}) = G_2(\mathbf{d})$ (see for example [42]).

For $n > 2$, $G_n(\mathbf{k})$ is called a "short tailed" distribution, as it tends to zero, for $\mathbf{k} \to \infty$, more rapidly than $G_2(\mathbf{k})$. To the limit, for $n \to \infty$, the generalised Gaussian tends to a box function, with support $[\mathbf{k}_0 - \boldsymbol{\sigma}_\infty, \mathbf{k}_0 + \boldsymbol{\sigma}_\infty]$. Such distributions are used to model situations where errors can be strictly controlled, as for example errors arising from rounding numbers.

Other kinds of distributions are also common, as for example the Log Normal [37] and others. In general, each of the three states of information $\Theta(\mathbf{d} \mid \mathbf{m})$, $\Pi_M(\mathbf{m})$ and $\Pi_D(\mathbf{d})$ should be modelled independently with an appropriate distribution. Long tailed distributions will imply that parameters might deviate substantially from the expected value, in other words that our knowledge of the true values of the parameters is scarce. Short tailed distributions will imply "hard bounds" on the errors, in other words the true values of the parameters can be assumed to be within an interval with greater confidence.

In general, given a model for each of the three states of information $\Theta(\mathbf{d} \mid \mathbf{m})$, $\Pi_M(\mathbf{m})$ and $\Pi_D(\mathbf{d})$, the resulting a posteriori probability density $\Xi(\mathbf{m} \mid \mathbf{d})$ could still be difficult to integrate or to be characterised analytically by means of central and dispersion estimators. In few special cases of practical relevance it is possible to do so.

Additive noise and Gaussian hypothesis

An important special case is the one of Gaussian assumptions. The case leads to simplifications for the posterior probability density $\Xi(\mathbf{m} \mid \mathbf{d})$, allowing one to formulate analytic expressions for $\mathbf{m}_{MAP}$ and $\mathbf{m}_{ML}$. Additionally, these point estimates establish an equivalence with least squares deterministic methods.

Assuming that the uncertainties arising from the instrumentation and from modelling

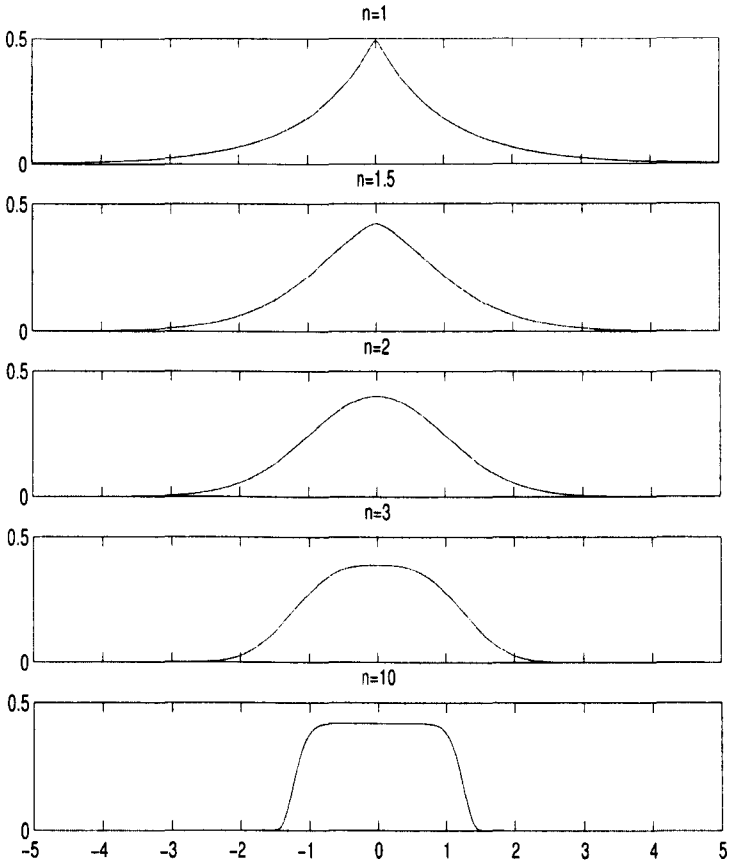


Figure 2.1: Generalised Gaussian distribution for $n = 1, n = 1.5, n = 2, n = 3, n = 10$.

errors are additive

$$\mathbf{d} = h(\mathbf{m}) + \boldsymbol{\nu} \tag{2.26}$$

where $h(\mathbf{m})$ is the forward model, $\boldsymbol{\nu}$ represents the uncertainties, and $p_{\boldsymbol{\nu}}$ is the distribution of $\boldsymbol{\nu}$, the likelihood distribution is

$$\Theta(\mathbf{d} \mid \mathbf{m}) = p_{\boldsymbol{\nu}}(\mathbf{d} - h(\mathbf{m}) \mid \mathbf{m}) \tag{2.27}$$

Additionally, if $\boldsymbol{\nu}$ and $\mathbf{m}$ are statistically independent

$$\Theta(\mathbf{d} \mid \mathbf{m}) = p_{\boldsymbol{\nu}}(\mathbf{d} - h(\mathbf{m})) \tag{2.28}$$

and if $p_{\boldsymbol{\nu}}$ is Gaussian with zero mean, and covariance $C_{\boldsymbol{\nu}}$

$$\Theta(\mathbf{d} \mid \mathbf{m}) \propto \exp\left\{-\frac{1}{2}(\mathbf{d} - h(\mathbf{m}))^T C_{\boldsymbol{\nu}}^{-1} (\mathbf{d} - h(\mathbf{m}))\right\} \tag{2.29}$$

Suppose also now that the prior state of information on $\mathbf{m}$ is Gaussian with mean $\mathbf{m}_0$ and covariance $C_{\mathbf{m}}$, applying 2.12, given that $\mathbf{d}$ is measured and therefore $\Pi_D(\mathbf{d})$ is a constant, the posterior model parameters distribution is

$$\begin{aligned}\Xi(\mathbf{m} \mid \mathbf{d}) \propto & \exp\left\{-\frac{1}{2}(\mathbf{d} - h(\mathbf{m}))^T C_{\boldsymbol{\nu}}^{-1} (\mathbf{d} - h(\mathbf{m}))\right\} \\ & \cdot \exp\left\{-\frac{1}{2}(\mathbf{m} - \mathbf{m}_0)^T C_{\mathbf{m}}^{-1} (\mathbf{m} - \mathbf{m}_0)\right\}\end{aligned} \tag{2.30}$$

As the maximization of $\Xi(\mathbf{m} \mid \mathbf{d})$ and $\Theta(\mathbf{d} \mid \mathbf{m})$ gives respectively the MAP and ML estimates, from the last two equations it follows that

$$\mathbf{m}_{ML} = \operatorname{argmin} \frac{1}{2}(\mathbf{d} - h(\mathbf{m}))^T C_{\boldsymbol{\nu}}^{-1} (\mathbf{d} - h(\mathbf{m})) \tag{2.31}$$

$$\mathbf{m}_{MAP} = \operatorname{argmin} \frac{1}{2}\left[(\mathbf{d} - h(\mathbf{m}))^T C_{\boldsymbol{\nu}}^{-1} (\mathbf{d} - h(\mathbf{m})) + (\mathbf{m} - \mathbf{m}_0)^T C_{\mathbf{m}}^{-1} (\mathbf{m} - \mathbf{m}_0)\right] \tag{2.32}$$

It is important to note that equations (2.31) and (2.32) establish a parallelism with deterministic methods. They are, in their expression, identical to deterministic weighted least square estimates, respectively for the unregularised and regularised case (see Section 2.4). The calculation of $\mathbf{m}_{ML}$ and $\mathbf{m}_{MAP}$ with (2.31) and (2.32) requires solving non linear optimisations problems, as in general the forward model $h(\mathbf{m})$ is so. For the discussion of such optimisations the reader is referred to the methods in Section 2.3.

Additive noise and generalised Gaussian hypothesis

With the same conventions adopted for the Gaussian hypothesis, following the same procedure, and assuming that both the uncertainties ν and the state of prior information on the model parameters $\Pi_M(\mathbf{m})$ are generalised Gaussian distributions respectively of order n_1 and n_2, one obtains for the likelihood

$$\Theta(\mathbf{d} \mid \mathbf{m}) \propto \exp\left\{ -\frac{1}{n_1} \sum_{i=1}^{D} \frac{|\mathbf{d}_i - h(\mathbf{m})_i|^{n_1}}{(\sigma_{n_1})_i^{n_1}} \right\} \tag{2.33}$$

and for the a posteriori density

$$\Xi(\mathbf{m} \mid \mathbf{d}) \propto \exp\left\{ -\frac{1}{n_1} \sum_{i=1}^{D} \frac{|\mathbf{d}_i - h(\mathbf{m})_i|^{n_1}}{(\sigma_{n_1})_i^{n_1}} \right\} \cdot \exp\left\{ -\frac{1}{n_2} \sum_{i=1}^{M} \frac{|\mathbf{m}_i - \mathbf{m}_{0,}|^{n_2}}{(\sigma_{n_2})_i^{n_2}} \right\} \tag{2.34}$$

The ML and MAP approaches therefore become

$$\mathbf{m}_{ML} = \operatorname{argmin} \sum_{i=1}^{D} \frac{|\mathbf{d}_i - h(\mathbf{m})_i|^{n_1}}{(\sigma_{n_1})_i^{n_1}} \tag{2.35}$$

and

$$\mathbf{m}_{MAP} = \operatorname{argmin} \frac{1}{n_1} \sum_{i=1}^{D} \frac{|\mathbf{d}_i - h(\mathbf{m})_i|^{n_1}}{(\sigma_{n_1})_i^{n_1}} + \frac{1}{n_2} \sum_{i=1}^{M} \frac{|\mathbf{m}_i - \mathbf{m}_{0,}|^{n_2}}{(\sigma_{n_2})_i^{n_2}} \tag{2.36}$$

The choice of the model orders n_1 and n_2 should be made independently, in such a way that each distribution models appropriately the respective state of information. Often, for simplicity, model orders are chosen in such a way that $n_1 = n_2$.

For $n1 > 1$ and $n2 > 1$ and finite the objective functions in (2.35) and (2.36) are differentiable functions, the minima are found by traditional optimisation methods. For $n_1 = 1$ or $n_2 = 1$, the objective functions involve sums of absolute values, which are not differentiable at any point for which their argument is null. Special algorithms are required to deal efficiently with the optimisation problem.

Traditionally Gaussian assumptions ($n_1 = 2, n_2 = 2$) are made both on the noise characteristics and on the prior information on model parameters, as this leads to well known least squares problems. As introduced earlier, the choice $n_1 = 1$ is important in the presence of outliers in the data, and leads to robust methods. Recently the choice $n_2 = 1$ is assuming an important role in conjunction with the reconstruction of non-smooth images. The reader is referred to Chapter 7, for a detailed description of the MAP approach for $n_2 = 1$.

2.3 Deterministic Inversion

In this section the deterministic approach to inverse problems is introduced. Central to the approach is the definition of the mapping $h : \mathcal{M} \mapsto \mathcal{D}$, which is the forward model. For a non-linear forward problem

$$\mathbf{d} = h(\mathbf{m}) \tag{2.37}$$

defines the observations corresponding to the model configuration $\mathbf{m}$, if the problem is linear, then the observations are expressed as

$$\mathbf{d} = H\mathbf{m} \tag{2.38}$$

Deterministic approaches seek to determine the value of the parameters $\mathbf{m}$ from the experimental values $\mathbf{d}$ through the deterministic mapping h that links them. In most practical situations, due to modelling errors and to measurement uncertainties, the measured data $\mathbf{d}$ does not belong to the range of h, therefore the classical solution

$$h(\mathbf{m}) - \mathbf{d} = 0 \tag{2.39}$$

does not exist. There are other difficulties are associated with (2.39). For continuous inverse problems, the inversion is often ill-posed in the Hadamard sense:

- (as already seen) The solution might not exist.
- The solution might be non-unique.
- The solution could be a non-continuous function of the data, such that small perturbations on the data cause arbitrarily large errors in the reconstructed parameters.

A common approach is to define the solution by the least squares (LS) method

$$\mathbf{m}_{LS} = \operatorname{argmin} \|h(\mathbf{m}) - \mathbf{d}\|^2 \tag{2.40}$$

which guarantees the existence of a solution even if the measured data is not in the range of the forward model. The problems of uniqueness and stability (continuous dependence of the inverse solution with the measured data) remain.

For finite dimensional problems, strictly speaking, the stability of the inverse solution is not an issue, as errors in the model space are always bounded for bounded errors in the data space [43]. However, most finite dimensional inverse problems might have a similar behaviour to the infinite dimensional counterparts, as the sensitivity to small perturbations

can be very large. For this characteristic, and for the fact the the uniqueness of the inverse might not be granted, these problems are called *discrete ill–posed problems.*

The study the null space of h, the space $\mathcal{N}(h) = \{\mathbf{m} \in \mathcal{M} : h(\mathbf{m}) = 0\}$, is important in order to analyse the ill-posedness of (2.40). As an example, consider a parameter vector $\mathbf{m}_{\mathcal{N}} \in \mathcal{N}(h)$. If $\mathbf{m}_{LS}$ is a solution, then $\mathbf{m}_{LS}+\mathbf{m}_{\mathcal{N}}$ is too. The null space of h is therefore a source of non uniqueness for the inverse. A general characterisation of $\mathcal{N}(h)$ for non-linear problems is difficult. For linear problems, as (2.38), such analysis is possible via *singular value decomposition* (SVD) of the forward operator H. As non-linear inverse problems are usually solved through successive linearisations, the SVD analysis of a linear forward operator is important even in this context.

SVD Decomposition of H and ill–posedness

The linear forward operator $H : \mathcal{M} \mapsto \mathcal{D}$ is represented by the matrix $H \in \mathbb{R}^{D\times M}$, the associated linear inverse problem is

$$\mathbf{m}_{LS} = \operatorname{argmin} \|H\mathbf{m} - \mathbf{d}\|^2 \tag{2.41}$$

The null space $\mathcal{N}(H)$ is defined as

$$\mathcal{N}(H) = \{\mathbf{m} \in \mathcal{M} : H\mathbf{m} = 0\} \tag{2.42}$$

The SVD decomposition of H, is a decomposition of the form

$$H = U\Sigma V^T = \sum_{i=1}^{q} \mathbf{u}_i \lambda_i \mathbf{v}_i^T \tag{2.43}$$

where $U = (\mathbf{u}_1, \ldots, \mathbf{u}_D) \in \mathbb{R}^{D\times D}$ and $V = (\mathbf{v}_1, \ldots, \mathbf{v}_M) \in \mathbb{R}^{M\times M}$ are orthonormal matrices, $q = \min(M, D)$, and $\Sigma = \operatorname{diag}(\lambda_1, \ldots, \lambda_q) \in \mathbb{R}^{D\times M}$ is a diagonal matrix with non negatives entries ordered in such a way that

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q \geq 0 \tag{2.44}$$

The numbers λ_i are called *singular values* of H and the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are called respectively *left singular vectors* and *right singular vectors.* Denoting the matrix H rank with r, if $r < q$, the last $\lambda_{r+1}, \ldots, \lambda_q$ singular values are zero and the corresponding singular vectors $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_q$ are in the null space $\mathcal{N}(H)$.

Consider now the calculated forward measurement $\mathbf{d} = H\mathbf{m}$ of an arbitrary vector

$\mathbf{m} \in \mathcal{M}$. Given the SVD decomposition of H, the vector $\mathbf{m}$ can be expressed as a linear combination of the singular vectors $\mathbf{v}_i$

$$\mathbf{m} = \sum_{i=1}^{r} (\mathbf{v}_i^T \mathbf{m}) \mathbf{v}_i \tag{2.45}$$

we have therefore

$$\mathbf{d} = \sum_{i=1}^{r} \lambda_i (\mathbf{v}_i^T \mathbf{m}) \mathbf{u}_i \tag{2.46}$$

The ill posedness of a problem is showed by the singular values rapidly decaying to zero. This indicates that there are certain model configurations $\mathbf{m}$, parallel to the higher singular vectors $\mathbf{v}_i$, for which the observations $\mathbf{d}$ are strongly attenuated by the small values of λ_i. Such observations are unreliable in the presence of noise, therefore it is difficult to invert the data. In ill-posed problems, the higher-order singular values can become so small, that numerical instability can occur even for simulated data.

It is typical for inverse ill-posed problems to have a singular spectrum that decays to zero without any particular gap in the singular values. We will have the following general situation

$$\lambda_1 \geq \ldots \geq \lambda_z \geq \mu > \lambda_{z+1} \geq \ldots \geq \lambda_w \geq \epsilon > \lambda_{w+1} \geq \ldots \geq \lambda_r > 0 = \lambda_{r+1} = \lambda_q \tag{2.47}$$

where μ is the accuracy threshold set by the instrumentation in use, ϵ is a threshold for which observations are numerically null. Singular values $\{\lambda_1, \ldots, \lambda_z\}$ cause observations which are reliable, singular values $\{\lambda_{z+1}, \ldots, \lambda_w\}$ cause observations which are unreliable, singular values $\{\lambda_{w+1}, \ldots, \lambda_q\}$ cause observation which are in practice null. The space $\mathcal{M}$, which is spanned by the singular vectors $\mathbf{v}_i$, can be categorised therefore as follows:

- The subspace of $\mathcal{M}$ spanned by $\{\mathbf{v}_1, \ldots, \mathbf{v}_z\}$ gives rise to observations that are reliable; no problems are associated with it.
- The subspace of $\mathcal{M}$ spanned by $\{\mathbf{v}_{z+1}, \ldots, \mathbf{v}_w\}$ does not indicate a rank deficiency in H, but it implies that model configurations $\mathbf{m}$ belonging to it give rise to observations that have small amplitudes. The reconstruction of model parameters belonging to this space will therefore be (very) sensitive to observation errors.
- The subspace of $\mathcal{M}$ spanned by $\{\mathbf{v}_{w+1}, \ldots, \mathbf{v}_q\}$ gives rise to observations that are null, at least numerically. The space, if not empty, is a cause of non uniqueness and indicates a rank deficiency in the matrix H. Additional information is needed to invert the data uniquely.

The ill-posedness is therefore a condition for which some models parameters cannot be identified either because they are in the null space of the forward operator, or because they give rise to observations that are too small to be measured with accuracy. An indicator of ill-posedness is the condition number of H

$$\text{cond}(H) = \frac{\lambda_1}{\lambda_q} \tag{2.48}$$

which assumes large values for ill-posed problems, and tends to infinity for rank deficient matrices H.

Linear LS inversion

Suppose, for now, that the matrix H in (2.41) has full column rank and that $D > M$, in other words suppose that $\mathcal{N}(H) = \emptyset$. In this case the solution of the least squares problem is known to be unique, and it is obtained by solving the so called *normal equations*

$$H^T H \mathbf{m}_{LS} = H^T \mathbf{d} \tag{2.49}$$

or equivalently by

$$\mathbf{m}_{LS} = H^\dagger \mathbf{d} \tag{2.50}$$

where the matrix $H^\dagger = (H^T H)^{-1} H^T$ is called the *Moore Penrose inverse* of H. Given the SVD decomposition of H, the LS inverse solution can be expressed from (2.50) as

$$\mathbf{m}_{LS} = \sum_{i=1}^{M} \frac{\mathbf{u}_i^T \mathbf{d}}{\lambda_i} \mathbf{v}_i \tag{2.51}$$

Equation (2.51) shows the difficulties associated with solving inverse ill-posed problems. Terms with larger values of i, for which observations are poor, are strongly amplified by the factor $1/\lambda_i$. As it is common for ill-posed inverse problems to have condition numbers as high as $1 \cdot 10^{20}$, any noise on the data $\mathbf{d}$ which has components parallel to the higher singular vectors $\mathbf{u}_i$ is amplified enormously, leading to a solution that is dominated by noise.

Non–linear LS inversion

Similar conclusions arise from the analysis of the LS inversion of the non–linear inverse problem

$$\mathbf{m}_{LS} = \text{argmin} \, \|h(\mathbf{m}) - \mathbf{d}\|^2 \tag{2.52}$$

The solution for the non-linear problem can be computed using iterative methods such as Newton's Method. The method works by approximating firstly the objective functional $\psi(\mathbf{m}) = \|h(\mathbf{m}) - \mathbf{d}\|^2$ around an initial estimate $\mathbf{m}^{(k)}$ with its second order Taylor series

$$\tilde{\psi}(\mathbf{m}) = \psi(\mathbf{m}^{(k)}) + \frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}^{(k)})(\mathbf{m} - \mathbf{m}^{(k)}) + \frac{1}{2}\,(\mathbf{m} - \mathbf{m}^{(k)})^T \frac{\partial^2 \psi}{\partial \mathbf{m}^2}(\mathbf{m}^{(k)})(\mathbf{m} - \mathbf{m}^{(k)}) \quad (2.53)$$

and then by seeking the minimiser of the approximated functional $\tilde{\psi}(\mathbf{m})$. A new estimate $\mathbf{m}^{(k+1)}$ is found in this way. The iteration is repeated until some convergence criterion is met. By imposing the Euler–Lagrange equations for $\tilde{\psi}(\mathbf{m})$ we find that, at an optimal point $\mathbf{m}^{(k+1)}$, we must have

$$\frac{\partial \tilde{\psi}}{\partial \mathbf{m}}(\mathbf{m}^{(k+1)}) = \frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}^{(k)}) + (\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)})^T \frac{\partial^2 \psi}{\partial \mathbf{m}^2}(\mathbf{m}^{(k)}) = 0 \quad (2.54)$$

from which, assuming that the Hessian matrix $\frac{\partial^2 \psi}{\partial \mathbf{m}^2}(\mathbf{m}^{(k)})$ is invertible, follows

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \left(\frac{\partial^2 \psi}{\partial \mathbf{m}^2}(\mathbf{m}^{(k)})\right)^{-1} \frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}^{(k)}) \quad (2.55)$$

where the gradient $\frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}^{(k)})$ can be expressed in the following form

$$\frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}^{(k)}) = -2\left(\frac{\partial h}{\partial \mathbf{m}}(\mathbf{m}^{(k)})\right)^T (h(\mathbf{m}^{(k)}) - \mathbf{d}) \quad (2.56)$$

and the Hessian can be expressed as

$$\frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}^{(k)}) = 2\left(\frac{\partial h}{\partial \mathbf{m}}(\mathbf{m}^{(k)})\right)^T \left(\frac{\partial h}{\partial \mathbf{m}}(\mathbf{m}^{(k)})\right) + 2\sum_i \left(\frac{\partial^2 h_i}{\partial \mathbf{m}^2}(\mathbf{m}^{(k)})(h_i(\mathbf{m}^{(k)}) - \mathbf{d}_i)\right) \quad (2.57)$$

As in the neighbourhood of a solution the second order terms in (2.57) become negligible with respect to the first order terms, they are often neglected for simplicity. Thus labelling the term $\frac{\partial h}{\partial \mathbf{m}}(\mathbf{m}^{(k)})$ as J_k, the Newton's update equation becomes finally

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - (J_k^T J_k)^{-1} J_k^T (h(\mathbf{m}^{(k)}) - \mathbf{d}) \quad (2.58)$$

or equivalently

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - J_k^{\dagger}(h(\mathbf{m}^{(k)}) - \mathbf{d}) \quad (2.59)$$

In this form, the calculation of the update for a non-linear inverse problem (2.59) is similar to the solution of a linear inverse problem (2.50). The properties of J_k therefore determine

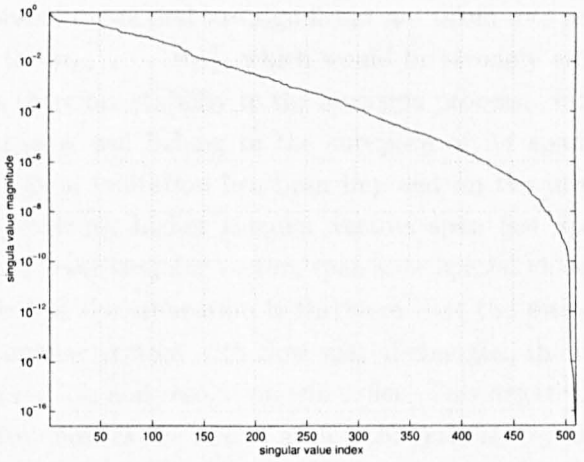


Figure 2.2: Singular spectrum for the linearised EIT forward operator, calculated on a model of a circular tank with 64 electrodes. In this particular simulation $H \in \mathbb{R}^{992\times510}$ as the conductivity is represented with 510 parameters, and 992 observations are collected.

the ill-posedness of each linear step of the Newton's method, and can be analysed by means of SVD analysis: the calculation of each update step suffers from the same difficulties associated with (2.51) if the problem is ill-posed.

As an illustrative example, Figure 2.2 shows the singular spectrum for the linearised EIT forward operator, calculated on a model that reproduces a typical experimental setup. The singular values have been scaled in such a way that $\lambda_1 = 1$. With the typical instrumentation accuracy of $1 \cdot 10^{-2} : 1 \cdot 10^{-3}$ it is possible to gather reliable observations only for model configurations spanned by the first $150 : 250$ right singular vectors. The remaining part of the model space, spanned by $\{\mathbf{v}_{251}, \ldots, \mathbf{v}_{510}\}$ produces observations that are not accurately measurable. Secondly, as some singular values are as small as $1 \cdot 10^{-17}$, measurement noise will be amplified enormously by a simple LS inversion.

It is therefore not possible to solve an inverse ill-posed problem simply by applying (2.49) or (2.52). Specific techniques, called *regularisation techniques* must be adopted to deal with the difficulties. The most immediate way to regularise an inverse problem is called *Truncated SVD* (TSVD) regularisation. Given that there is a threshold μ set by the instrumentation, under which the observations are not reliable, the technique solves (2.50) or (2.59) by expanding the solution in terms of SVD and by stopping the summation at the index z for which $\lambda_z > \mu$. For example, in the linear case, (2.51) becomes

$$\mathbf{m}_{TSVD} = \sum_{i=1}^{z} \frac{\mathbf{u}_i^T \mathbf{d}}{\lambda_i} \mathbf{v}_i \tag{2.60}$$

In such a way only observations that are significant are taken into account. All the noise components parallel to $\{\mathbf{u}_{z+1}, \ldots, \mathbf{u}_D\}$, which would be strongly amplified, are rejected. Regularisation brings therefore stability to the inversion process. Naturally the estimated model parameters $\mathbf{m}_{TSVD}$ will belong to the subspace of $\mathcal{M}$ spanned by the singular vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_z\}$: some limitation has been imposed on the inverted data. In EIT, as in many inverse problems, higher singular vectors span fast spatial changes in the model parameters, and lower singular vectors span slow spatial changes (as illustrated in Figure 2.3). The effect of the truncation is therefore that the estimated model $\mathbf{m}_{TSVD}$ will be spanned by singular vectors with slow spatial changes: the inversion process will produce only smooth models and resolution will suffer. This negative effect is common to all regularisation techniques, as the aim of such techniques is to penalise the presence of the higher singular vectors in the reconstructed images. The TSVD approach completely rejects such vectors; an alternative approach is *Tikhonov Regularisation.*

2.3.1 Tikhonov Regularisation

Tikhonov regularisation [44][45] is popular in many fields of inverse problems. The aim of the method is to dampen the contribution of the higher singular vectors into the reconstructed model $\mathbf{m}_{LS}$. The TSVD achieves this by explicitly truncating (2.60). Tikhonov regularisation is an implicit method that achieves similar results by means of adding a penalty term to the objective function.

Regularised linear LS inversion

For the linear LS inverse problem, the regularised inverse solution is formulated as the minimiser of the following modified functional

$$\mathbf{m}_{(\alpha,L)} = \operatorname{argmin} \|H\mathbf{m} - \mathbf{d}\|^2 + \alpha F(\mathbf{m}) \tag{2.61}$$

where $F(\mathbf{m}) \geq 0$ is a regularisation functional and $\alpha > 0$ is the regularisation parameter, or Tikhonov factor. Very often the regularisation functional $F(\mathbf{m})$ assumes the form

$$F(\mathbf{m}) = \|L\mathbf{m}\|^2 \tag{2.62}$$

or the form

$$F(\mathbf{m}) = \|L(\mathbf{m} - \mathbf{m}_0)\|^2 \tag{2.63}$$

where L is an appropriate regularisation matrix and $\mathbf{m}_0$ is an a priori estimate of $\mathbf{m}$. Of the two forms (2.62) and (2.63) we will be using the second one, for maximum generality.

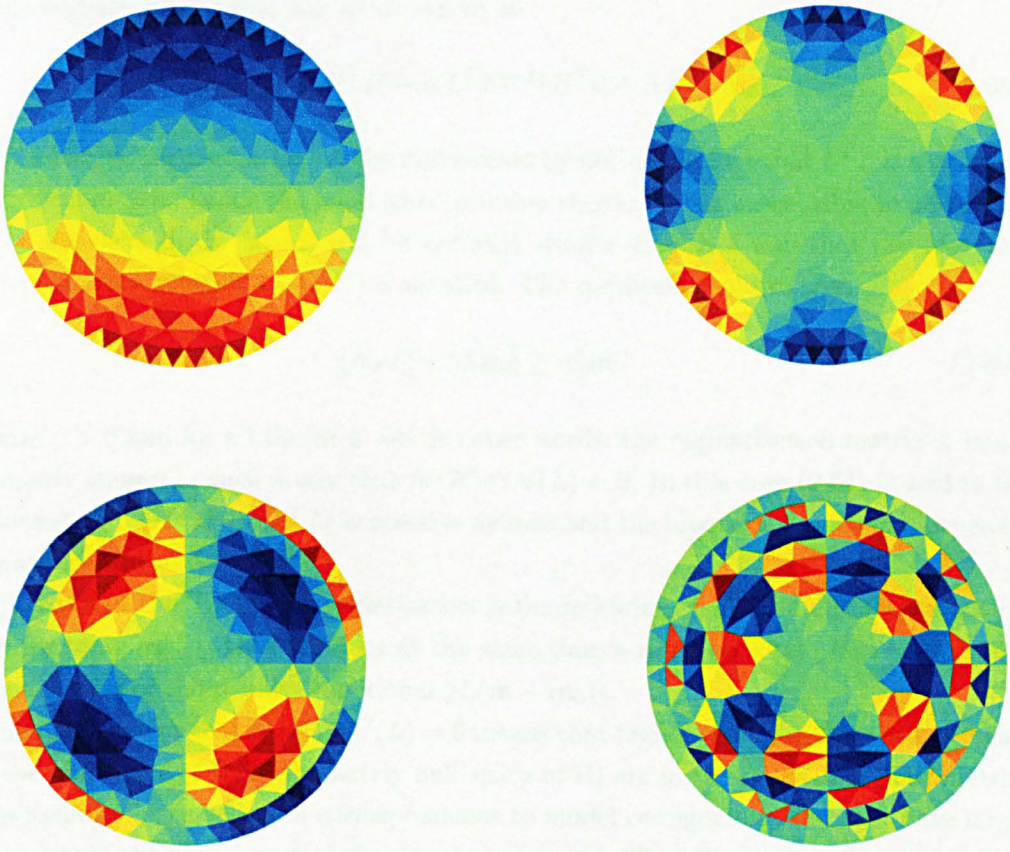


Figure 2.3: From left to right and from top to bottom, plots of the 2^{nd},8^{th},69^{th} and 225^{th} singular vectors for the linearised EIT forward operator, calculated on a model of a circular tank with 64 electrodes. As it is possible to appreciate from the plots, lower singular vectors span slower spatial changes, while the higher ones span fast variations. The 2^{nd} singular vector is associated with a singular value of 0.774, the 8^{th} with a singular value of 0.469, the 69^{th} with a singular value of 0.173 and the 225^{th} with a singular value of 0.001.

The regularised LS inverse problem is therefore

$$\mathbf{m}_{(\alpha,L)} = \text{argmin} \, \|H\mathbf{m} - \mathbf{d}\|^2 + \alpha \, \|L(\mathbf{m} - \mathbf{m}_0)\|^2 \tag{2.64}$$

and the regularised solution can be shown to be

$$\mathbf{m}_{(\alpha,L)} = (H^T H + \alpha \, L^T L)^{-1}(H^T \mathbf{d} + \alpha \, L^T L \mathbf{m}_0) \tag{2.65}$$

The effect of the regularisation can be understood by noting the fact that $L^T L$ is a positive semidefinite matrix, hence the regularised solution should be less susceptible to perturbations in the data $\mathbf{d}$. In fact L will be typically chosen in such a way that the *Morozov complementation condition* [46][47] is satisfied. The condition requires that

$$\|H\mathbf{m}\| + \|L\mathbf{m}\| \geq \gamma \|\mathbf{m}\| \tag{2.66}$$

for some $\gamma > 0$ and for all the $\mathbf{m} \in \mathcal{M}$. In other words, the regularisation matrix L must be properly chosen in such a way that $\mathcal{N}(H) \cap \mathcal{N}(L) = \emptyset$. In this case (2.64) is said to be regularised, and $(H^T H + \alpha \, L^T L)$ is positive defined and the inverse is unique and depends continuously on the data $\mathbf{d}$ [47].

The basis of the Tikhonov regularisation is the following: minimising (2.64) means to search for some $\mathbf{m}_{(\alpha,L)}$ that provides at the same time a small residual $\|H\mathbf{m} - \mathbf{d}\|$ and a moderate value of the penalty functional $\|L(\mathbf{m} - \mathbf{m}_0)\|$.

The requirement that $\mathcal{N}(H) \cap \mathcal{N}(L) = \emptyset$ means that higher right singular vectors (that span the null space of H, or the nearly null space of H) are in the range of L. The matrix L thus assumes large values in correspondence to model configurations $\mathbf{m}$ that have large components in the higher part of the singular spectrum. The effect is therefore to penalise such model configurations in the reconstruction, effectively achieving a similar effect to the TSVD regularisation.

The regularisation L thus introduces some a priori information into the problem, complementing in such a way the information that is not observable because of the null space of H. The role of the Tikhonov factor is to balance the amount of prior information that is used in the inversion process. Several methods for the choice of the parameter exist, the most common ones are the *L-curve Method* [43], the *Discrepancy Principle* [48] and the *Generalised Cross Validation Method* [49]. The reader is referred to [50] [10] for an overview on such methods.

Standard Tikhonov regularisation

One of the simplest forms of Tikhonov regularisation is the *damped least squares* inversion, that corresponds to the choice $L = I$ and $\mathbf{m}_0 = 0$.

With $L = I$, the solution to the linear inverse problem is

$$\mathbf{m}_{(\alpha,I)} = (H^T H + \alpha I)^{-1} H^T \mathbf{d} \tag{2.67}$$

which, by applying the SVD decomposition to H can be written as

$$\mathbf{m}_{(\alpha,I)} = \sum_{i=1}^{M} \frac{\lambda_i^2}{\lambda_i^2 + \alpha} \frac{\mathbf{u}_i^T \mathbf{d}}{\lambda_i} \mathbf{v}_i \tag{2.68}$$

The multiplicative factors $\lambda_i^2/(\lambda_i^2+\alpha)$ tend to the value of 1, for large values of λ_i ($\lambda_i \gg \alpha$), and to the value of 0 for small values of λ_i ($\lambda_i \ll \alpha$). The contribution of the higher singular vectors is therefore rejected from participating to the inverse solution.

NOSER

A second option for the matrix L is the ones suggested by the NOSER algorithm [51] which chooses $L = \text{diag}(\sqrt{J^T J})$. As in the previous choice the regularisation matrix is diagonal, and with SVD decomposition, the regularised solution can be expressed as

$$\mathbf{m}_{(\alpha,NOSER)} = \sum_{i=1}^{M} \frac{\lambda_i^2}{\lambda_i^2 + \alpha l_i^2} \frac{\mathbf{u}_i^T \mathbf{d}}{\lambda_i} \mathbf{v}_i \tag{2.69}$$

where l_i are the diagonal elements of L. Also in this case the multiplicative factors $\lambda_i^2/(\lambda_i^2+\alpha\, l_i^2)$ tend to the value of 1, for large values of λ_i and to the value of 0 for small values of λ_i. The multiplicative terms in front of $\frac{\mathbf{u}_i^T \mathbf{d}}{\lambda_i} \mathbf{v}_i$ are called *filter factors*. Different regularisation techniques can be distinguished by having different filter factors, attenuating (in different) ways the content of higher portion of the singular spectrum in the inversion. For the example the TSVD technique has the following filter factors

$$\begin{array}{l} 1 \text{ for } i < z \\ 0 \text{ for } i \geq z \end{array} \tag{2.70}$$

which correspond to a net truncation.

Regularisation with dense matrices L

If L is a dense matrix, the regularisation is said to be a *Generalised Tikhonov Regularisation.* It is common in many inverse problems to use regularisation matrices L that are discrete representations of first or second order differential operators [52]. Regularisation functionals corresponding to such choices of L assume large values for models $\mathbf{m}$ with large first or second spatial derivatives. As higher singular vectors have sharp spatial variations, their contribution in the inversion is penalised. First order differential operators produce almost constant inversions, while second order differential operators produce smooth inversions.

The analysis of the regularised solution via SVD decomposition is possible only for diagonal matrices L such as the ones discussed in the previous sections. Chapter 5 will discuss the choice of some non-standard matrices L and will introduce the analysis of such inverse solutions by means of GSVD decomposition.

Regularised non-linear LS inversion

The Tikhonov regularised non linear inverse problem is formulated similarly to (2.61)

$$\mathbf{m}_{(\alpha,F)} = \operatorname{argmin} \|h(\mathbf{m}) - \mathbf{d}\|^2 + \alpha \|L(\mathbf{m} - \mathbf{m}_0)\|^2 \tag{2.71}$$

by applying Newton's method, the non-linear problem can be solved iteratively with the following recursion

$$\mathbf{m}_{(\alpha,L)}^{(k+1)} = \mathbf{m}_{(\alpha,L)}^{(k)} - (J_k^T J_k + \alpha L^T L)^{-1}[J_k^T(h(\mathbf{m}_{(\alpha,L)}^{(k)}) - \mathbf{d}) - \alpha L^T L(\mathbf{m} - \mathbf{m}_0)] \tag{2.72}$$

All the considerations made for the linear regularised LS problem apply to each linearised step of the non-linear inversion. Similarly to the linear problem, (2.72) is regularised if $\mathcal{N}(J_k) \cap \mathcal{N}(L) = \emptyset$ for every k.

2.4 Comments on Statistical and Deterministic Inversion

A clear line of separation between statistical and deterministic methods is difficult to be drawn, as deterministic methods make implicit assumptions on the distribution of the measurement errors and model uncertainties. The very common LS inversion is equivalent to assuming that such errors are additive, Gaussian and with zero mean. In particular, the unregularised LS inversion (2.52) is equivalent to the ML estimates (2.31) whenever the distribution of the Gaussian uncertainties has the same variance on all the observations,

and different observations are uncorrelated, i.e $C_{\boldsymbol{\nu}}^{-1} \propto I$.

The equivalence of the two approaches is even more evident for the *weighted least squares inversion* (WLS), where each observation is weighted depending on its reliability by a diagonal matrix W. The unregularised inverse problem becomes thus

$$\mathbf{m}_{WLS} = \operatorname{argmin} \|W(h(\mathbf{m}) - \mathbf{d})\|^2 \tag{2.73}$$

In this case where considerations of probabilistic nature are explicitly incorporated in the deterministic approach, the classification of the method as deterministic or probabilistic is not clear.

Similarly Tikhonov regularised solutions such as (2.71) make implicit assumptions on the prior distribution of model parameters. Such an approach is equivalent to the MAP estimate (2.32) under the same hypothesis for the measurement uncertainties ($C_{\boldsymbol{\nu}}^{-1} \propto I$), and assuming that the model parameters follow a Gaussian distribution such that $L^T L \propto C_{\mathbf{m}}^{-1}$.

The use of the ℓ_2-norm for measuring the data mismatch term and the regularisation term is intimately related to a Gaussian assumption. The choice of different norms corresponds to assuming different models for the distribution either of the measurement errors, or of the model parameters as in (2.36). Deterministic methods can therefore be thought as particular cases of probabilistic methods, where point estimates (ML or MAP) are sought and assumptions about distributions $\Theta(\mathbf{d} \mid \mathbf{m})$ and $\Pi_M(\mathbf{m})$ are made implicitly.

Chapter 3

EIT Techniques

3.1 Introduction

In the preceding chapter a general overview on inverse problems was given, and a general framework for the solution of inverse problems was discussed from the deterministic and statistical point of view. In the present chapter topics specific to EIT are briefly introduced. An exhaustive overview of the field of research would be out of the scope of the present work. The aim is therefore to introduce topics that are relevant to work developed in the following parts of the thesis. Specifically, given that inverse problems are solved by fitting a model to the data, the present chapter is mainly concerned with techniques for gathering such data. Modelling techniques will be discussed in the next chapter.

The present chapter begins by describing the principle of the technique. Methods for energising the object under measurement and for collecting data are then discussed, in particular the concepts of optimal experiments and optimal excitations are discussed. The first part of the chapter, focused on measurement schemes, is followed by a brief discussion of 2D/3D issues, and of Absolute/Differential/Dynamic imaging techniques.

3.2 EIT Principle

Electrical Impedance Tomography is a measurement technique that estimates the conductivity $\sigma(x, y, z)$ of an object from boundary measurements. A set of electrodes is attached to the surface of a body under investigation. Some currents are injected through the electrodes and the resulting potentials are measured. The conductivity is then estimated by processing the measurement data.

A common measurement setup, illustrated in Figure 3.1, comprises a specific data

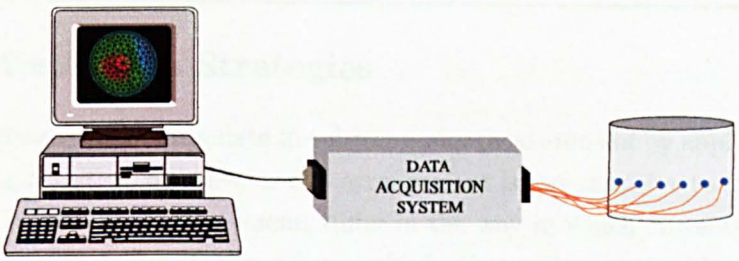


Figure 3.1: Example of Electrical Impedance Tomography experiment.

acquisition system, the tomograph, that is connected to the electrodes applied to the object under measurement and is responsible for setting currents and reading voltages. The instrument is typically controlled by a personal computer that sets the excitations, gathers the data and performs the tomographic reconstruction.

The applied currents are AC, as this reduces the corrosion of the electrodes by electrolytic phenomena. Additionally, in medical applications such stimuli are required to be AC to meet safety standards. The amplitude and frequency of the stimuli vary with the application. Current intensity and frequency can range from several Amperes and few Hz in geophysical applications to 1–10mA and 10kHz–1MHz in medical imaging applications. In the medical filed particular systems apply several AC stimuli of different frequencies at one time [53] [54] [55], such systems are called multifrequency tomographs and are normally used for differential imaging (see Section 3.5).

As the injected currents are AC, the resulting voltages are in principle out–of–phase with respect to them, given that the object's load could be complex. In some settings, as in the industrial field, loads are assumed for simplicity to be mainly resistive or mainly capacitive, and only the in–phase or the out–of–phase component of the voltages is measured. In this case only the real or imaginary parts of σ are reconstructed. In the medical context the loads are not assumed to be purely real, but commonly the interest is focused on their real part and only in–phase components of the voltages are measured. Estimation of complex loads, if needed, is possible [56] [22]. Throughout this work we assume that the measured objects are mainly resistive, or that we are interested in their resistive component.

In the next section the operation of EIT tomographs is discussed, and common data acquisition schemes used in medical and industrial applications are overviewed. For a detailed analysis on the hardware design and analysis the reader is refereed to [12] [57] [58] [59] [60] [61].

3.3 Data Collection Strategies

Common acquisition systems stimulate the object under measurement by applying currents and by measuring resulting voltages, as this arrangement is less sensitive to the degrading effect of contact impedances. Instruments differ in the way in which currents are applied and voltage are measured. Apart from a few methods, that will be covered later, strategies can be categorised as *pair drive* and *multiple drive* collection methods.

3.3.1 Number of Independent Measurements

The number of independent measurements that can be collected with N electrodes applied to a body can be calculated considering the object as an electrical network with N ports. Such a network can be completely characterised by the indefinite admittance matrix $Y \in \mathbb{R}^{N\times N}$ [62]. Given that the matrix Y is rank deficient, as the sum over columns and over rows is zero (the matrix is a description of the N-pole with no ground reference), the number of matrix elements that uniquely identify Y is less then N^2. Considering the rank deficiency, and considering that Y is symmetric due to the reciprocity theorem, the matrix is characterised completely by the diagonal and upper triangular part of the $\mathbb{R}^{(N-1)\times(N-1)}$ matrix obtained by deleting a row and a column from Y. The number of such elements amounts to $N(N-1)/2$. One should therefore expect this to be the maximum number of independent measurements that could be gathered from a body with N electrodes attached to it. Omitting measurements on current carrying electrodes, the total number of independent measurements is $N(N-3)/2$ [63].

3.3.2 Pair Drive Instruments

Pair drive instruments have a single current source whose terminals are connected to the driving electrodes, giving rise to a current flow in the measured object. Voltage differences resulting from the stimulus are then measured between pairs of the remaining electrodes. During the measurement sequence the current source is switched to another pair of electrodes and the voltage measurement process repeated until a complete independent set of measurements has been collected.

The way in which the driving pair is switched and the voltage measurements are collected varies. In geophysics there are several measurement configurations [33], all of them based on a stimulus pair and a measurement pair. In medical applications and process tomography applications, where the imaged object is often similar to a cylinder (thorax, mixing vessel) or a sphere (head), two schemes are prominent: the *adjacent drive method* [12][64][65] where the driving pair are neighboring electrodes and the *polar drive*

method [66] [67] where the driving pair are diametrally opposed electrodes. In both the schemes the driving pairs are rotated in several steps around the object until all the linearly independent measurements are collected.

The adjacent drive method, even if the most common "four electrode" measurement scheme, suffers from the fact that the current is driven mostly in the outer region of the imaged object. The method is therefore very sensitive to conductivity contrasts near the boundary and almost insensitive to central contrasts, worsening the ill-posedness of the problem. Additionally the method is very sensitive to perturbations in the boundary shape of the object and in the positioning of the electrodes. The polar drive method offers a better distribution of the sensitivity, as the current travels with greater uniformity through the imaged body.

The ideal angular position of the driving electrodes has been studied in [68] for a circular object with a circular inclusion of known position and radius. The larger and deeper the expected anomaly is, the larger the angular separation of the electrodes should be. The diametral position optimises the sensitivity for a contrast in the centre of the imaged object.

The problem of optimising the measurements with respect to the measured conductivity profile has been studied by means of *distinguishability* as discussed in the next section. The optimisation of the applied stimuli leads to multiple drive systems.

3.3.3 Distinguishability, Optimal Experiments and Multiple Drive Instruments

Firstly we introduce the concept of *current pattern*. Given the representation of the object under measurement as N-pole, it is natural to arrange the applied currents at each electrode in a vector $I = [I_1, \ldots, I_N]^T$. A generic vector $I \in \mathbb{R}^{N\times 1}$ is said to be a current pattern if $\sum_i I_i = 0$, as imposed by the charge conservation law.

The distinguishability in EIT was defined by Isaacson [69]. Given two conductivity distributions σ_1 and σ_2, they are said to be distinguishable by measurements of precision ϵ if there exists a current pattern I with $\|I\| = 1$ for which

$$\|R(\sigma_1)I - R(\sigma_2)I\| > \epsilon \tag{3.1}$$

where $R(\sigma)$ is the impedance matrix of the N-pole, assumed for simplicity to be real. Based on the definition (3.1) Gisser, Isaacson and Newell [70] defined the *optimal current*

pattern for the distinction of σ_1 from σ_2 as

$$I_{opt} = \text{argmax} \frac{\|R(\sigma_1)I - R(\sigma_2)I\|}{\|I\|} \tag{3.2}$$

The optimal current pattern as defined in [70] maximises therefore the ℓ_2 -norm of the difference of the electrode potentials $V_1 = R(\sigma_1)I$ and $V_2 = R(\sigma_2)I$ with respect to the ℓ_2-norm of the injected current. One problem with such definition is that the ℓ_2 norm of I has no particular physical meaning. In a later publication Gisser [71] considered optimising the distinguishability with respect to the dissipated power, using $I^T RI$ as denominator in (3.2).

Formula (3.2) can be employed in iterative reconstruction algorithms, as Newton's algorithm, to calculate at each step the pattern that maximises the distinguishability of the currently reconstructed conductivity σ_{rec} from the conductivity of the object, enhancing the progress of the algorithm. With this method $R(\sigma_{rec})$ must be calculated numerically on the basis of σ_{rec}, while $R(\sigma)$ is measured.

The optimal current pattern is calculated [69] as the eigenvector of the operator $(R(\sigma_{rec}) - R(\sigma))^2$ associated with the largest eigenvalue. The next best pattern to apply is the current pattern associated with the second largest eigenvalue. Further optimal patterns can be applied until the distinguishability reaches the accuracy threshold of the measurement system. All the applied optimal current patterns gather new information, as they are independent eigenvectors of a matrix and therefore linearly independent stimuli.

As an example the optimal patterns for a centered circular inhomogeneity embedded in a larger circular object are the so-called *trigonometric current patterns* [69].

$$I_i^k = \begin{cases} \cos\left(k\frac{2\pi i}{N}\right) & i = 1,\ldots,N \;\; K = 1,\ldots,N/2 \\ \sin\left((k - N/2)\frac{2\pi i}{N}\right) & i = 1,\ldots,N \;\; K = N/2+1,\ldots,N-1 \end{cases} \tag{3.3}$$

Clearly the application of an arbitrary current vector requires an instrument with multiple current sources. Such systems are called multiple drive tomographs, or adaptive tomographs, as the current patterns can be calculated to be optimal with respect to the measured object with the mentioned procedure. There are two arrangements for such tomographs. The first, proposed by researchers at Rensellaer Polytechnic Institute [57], consists in applying the stimuli to all the electrodes and in measuring potentials on the same electrodes. The tomograph they developed, ACT III, has 32 electrodes in total, they are used both for stimulation and voltage collection. The approach is sensitive to contact impedances, as voltages are collected on driving electrodes. The alternative approach, proposed by researchers at Oxford Brookes University [60], is to split the set of

electrodes: some are used for injecting currents and some solely for voltage measurement. In the developed tomograph, OXBACT-III, 32 injection electrodes (*current electrodes*) are interleaved with a set of 32 passive electrodes (*voltage electrodes*). The arrangement avoids measuring voltages on driving electrodes.

An important side effect in applying optimal current patterns is that they result in faster acquisition rates than traditional pair drive schemes, as there is no need to apply stimuli that would result in measurements under the accuracy threshold of the instrument. As an example, it is common practice with the OXBACT-III system to apply the first 10–12 optimal patterns out of 31 possible linearly independent excitation vectors. This results in a gain of speed of more than two fold.

The rationale behind optimal patterns is to maximise the voltage measurements with respect to the injected currents and, in such a way, to improve the signal to noise ratio in the measured values. A consideration must be taken regarding this conclusion. The use of optimal patterns does not necessarily guarantee a better accuracy over pair drive systems. With the pair drive system the stimuli are generated by a single device therefore they are repeatable through the whole set of patterns. A gain error in the device affects all the measurements simply as a multiplicative factor with no other effect on the reconstructed image than a scaling. The use of multiple current sources requires a perfect matching of the devices, as unmatched gains would distort the applied patterns, and affect the reconstructed image. Accurate calibration techniques are therefore needed for such systems [59] [60].

Some other important considerations for optimal patterns have been raised by several researchers, regarding in particular the definition of optimality. The original definition (3.2) is not based on practical considerations. Eyüboglu and Pilkington [67] argued that in a medical setting, where the safety regulations restrict the total amount of applied current, the distinguishability should be maximised with respect to this constraint. In industrial settings the constraints could arise from the output dynamic range of the current sources, such that $I_i < I_{max} \quad i = 1, \ldots, i = N$. The reader is referred to [72] for a generalisation of optimality criteria.

3.3.4 Other Excitation Methods

Recently Polydorides and McCann [73] proposed an excitation scheme that cannot be classified in the two preceding schemes. They suggest using a single current source instrument, but connecting this source to several neighbouring electrodes at one time. Voltages are collected traditionally, as potential differences between single electrodes. The set of current injection electrodes is then "scrolled" around the object by connecting new electrodes

at one side and leaving some free at the other. Voltage measurements are repeated for each new configuration. Polydorides and McCann have shown that the Jacobian matrix corresponding to measurements gathered with such a scheme shows a slower decay of the singular values, therefore the inverse problem is better conditioned.

3.4 2D and 3D Imaging

When a current is flowing between two electrodes applied to a body, clearly the flow of that current is three–dimensional, in the sense that the current spreads from one electrode both horizontally and vertically for gathering then at the other electrode. If a 2D ring of electrodes is applied to an object, the measurements are sensitive to off-plane conductivity variations. EIT is therefore inherently three–dimensional. On the contrary hard field techniques such as the X–Ray computed tomography, by setting detectors and emitters on a plane, can image just 2D sections as measurements depend only on that portion of the object.

In spite of these considerations, for other reasons, 2D imaging is very common, even though the trend is to go towards 3D in recent years [74] [65] [75] [22]. A first reason for which 2D is often adopted is that 3D imaging involves a number of unknowns much larger then the 2D problem. Modern desktop PCs have just enough resources for dealing with such problems in reasonable times; this was hardly the case just 4 or 5 years ago. A second important reason for which 2D imaging is a widespread technique has to do with data acquisition systems. For a 3D reconstruction electrodes should cover the height of the object. Such a requirement raises quickly the number of electrodes needed, and therefore the complexity of the tomograph. For 3D, a typical configuration is the one with 4 rings of 16 electrodes each [65], requiring 64 electrodes, while most measurement systems have 16 or 32. As an additional consideration, raising the number of electrodes poses performance issues for the data acquisition system. A greater number of measurements translates to a reduced frame rate. In applications such as medical imaging, where the data must be acquired faster than the physiological rates of change, this is an important consideration.

Above the mentioned reasons, and for the fact that in certain cases the user is interested simply in a 2D section, two–dimensional reconstructions are performed. Such reconstructions approximate the 3D problem assuming that currents and conductivities are invariant in the 3rd dimension. The 2D technique has been proven useful at imaging objects, quantitative estimates however suffer from the approximations. Some techniques that account for the vertical spread of the currents, but not for a variability of σ in that dimension have been introduced as a "partial fix" to the problem; they have been called

$2\frac{1}{2}$ D [76].

3.5 Absolute, Differential and Dynamic Imaging

Absolute imaging

Absolute imaging, sometimes referred to as static imaging, is an attempt to quantify in a absolute way the conductivity of an object. The properties of the object are not supposed to change during the measurement cycle, hence the name of static imaging. Most current attempts [22] [56] [75] [77] [78] [79] formulate the reconstruction problem as a non linear least squares inversion. Given the ill-posedness of the problem [9] the reconstruction is regularised with the Tikhonov method. Recalling (2.71), the inverse solution is formulated as

$$\boldsymbol{\sigma}_{rec} = \operatorname{argmin} \|h(\boldsymbol{\sigma}) - \mathbf{d}\|^2 + \alpha F(\boldsymbol{\sigma}) \tag{3.4}$$

where $\boldsymbol{\sigma}$ is the discrete conductivity of the forward model, $\mathbf{d}$ is the measured data and $F(\boldsymbol{\sigma}) \geq 0$ the regularisation functional. Commonly the ℓ_2-norm is used for the regularisation functional. In the discrete setting we have $F(\boldsymbol{\sigma}) = \|L\boldsymbol{\sigma}\|^2$ or $F(\boldsymbol{\sigma}) = \|L(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\|^2$, where L is a regularisation matrix and $\boldsymbol{\sigma}_0$ a prior estimate of the conductivity. With such a choice the non-linear problem (3.4) is easily solved with iterative techniques such as Newton's method (2.72), with the following recursion

$$\boldsymbol{\sigma}_{rec}^{(k+1)} = \boldsymbol{\sigma}_{rec}^{(k)} - (J_k^T J_k + \alpha L^T L)^{-1}[J_k^T(h(\boldsymbol{\sigma}_{rec}^{(k)}) - d) - \alpha L^T L(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)] \tag{3.5}$$

It is common practice [22] [56] [80] to start the reconstruction from a constant conductivity distribution $\boldsymbol{\sigma}_{rec}^{(0)}$. The value of such homogeneous initial guess is found by a least squares fit to the data.

Certain algorithms such as NOSER [51], called one-step reconstructors, perform only the first step of the Newton's method. They calculate $\boldsymbol{\sigma}_{rec}^{(1)}$ from $\boldsymbol{\sigma}_{rec}^{(0)}$. Such algorithms require the computation of the Jacobian J and of the forward solution $h(\boldsymbol{\sigma})$ solely for the initial homogeneous guess $\boldsymbol{\sigma}_{rec}^{(0)}$, which can be done with analytical methods. Such methods assume the inhomogeneities in the investigated medium to differ moderately from the background value, so that the problem can be considered linear (and only one step is needed). It is important to note that with these methods it is possible to precompute the regularised inverse $(J_k^T J_k + \alpha L^T L)^{-1}$ and the forward solution $h(\boldsymbol{\sigma}_{rec}^{(k)})$ for the homogenous initial guess. The reconstruction is then just a matter of matrix multiplication operations.

In EIT, the choice of the regularisation matrix L has been traditionally either an identity matrix [81], or a diagonal matrix [51] or an approximation of differential operators

[22] [56] [78] [82] [79]. In the last decade no-standard choices for L have been proposed in the literature [83] [84] [85] [35], with particular regard to introducing prior information into reconstruction as discussed in greater detail in Chapter 5. Different choices have also been made for the regularisation norm. Particular interest has been placed on the ℓ_1 norm [86] [87] [36] as means of reconstructing non-smooth conductivity profiles. Chapters 6 and 7 discuss such methods in greater detail.

In absolute imaging the accuracy of the measured voltages and of the predicted voltages is crucial, given the ill-posedness of the inverse problem. Apart from the uncertainties arising from the measurement system, and from solving the forward problem with numerical methods, particular consideration must be paid to geometrical uncertainties. Geometrical uncertainties arise from the non-perfect knoweledge of the electrodes' position and of the object's shape. The forward model, that is built on that data, therefore will not match the experiment with precision. These errors are particularly common in medical imaging, given the variability of anatomy and the difficulty in placing electrodes in a repeatable way.

Differential imaging

The aim of differential imaging is to reconstruct a change in conductivity rather then an absolute value. Such methods reconstruct $\Delta\boldsymbol{\sigma} = \boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1$ in the hypothesis that $\Delta\boldsymbol{\sigma}$ is small, and the problem linearisable:

$$\Delta\mathbf{d} = K\Delta\boldsymbol{\sigma} \tag{3.6}$$

where K is a constant matrix and $\Delta\mathbf{d} = \mathbf{d}_2 - \mathbf{d}_1$ is the difference between a reference voltage data $\mathbf{d}_1$ acquired for $\boldsymbol{\sigma}_1$ and a second data set $\mathbf{d}_2$ acquired for $\boldsymbol{\sigma}_2$.

The technique is commonly used for imaging temporal phenomena in medical applications, such as impedance changes during respiration [88] [89]. The data set $\mathbf{d}_1$ is acquired at a time t_1 for which the conductivity $\boldsymbol{\sigma}_1$ is assumed as reference, the data set $\mathbf{d}_2$ is acquired at a later time t_2 for which the conductivity change needs to be calculated. By reconstructing the conductivity change $\Delta\boldsymbol{\sigma}$, parameters of medical interest such as the pulmonary ventilation can be estimated.

In [88] [89] the matrix K is computed as the Jacobian in $\boldsymbol{\sigma}_1$ and the inverse problem formulated as

$$\Delta\boldsymbol{\sigma}_{rec} = \operatorname{argmin} \|J\Delta\boldsymbol{\sigma} - \Delta\mathbf{d}\|^2 \tag{3.7}$$

Problem (3.7) is of course ill-posed. Techniques such as Tikhonov regularisation (2.65) are used, usually assuming no prior information on $\Delta\boldsymbol{\sigma}$, and formulating the reconstruction

as

$$\Delta\sigma = (J^T J + \alpha L^T L)^{-1} J^T \Delta \mathbf{d} \tag{3.8}$$

It is common in differential imaging to use techniques called *backprojection method* and *sensitivity method* in place of (3.7) [65] [90] [91]. Such techniques are based on gross approximations; the reader is referred to the bibliography for further reading on the particular subject.

The rationale behind the use of differential techniques, especially in the medical field, is that errors arising from the uncertain positioning of the electrodes and uncertain knowledge of the patient's shape are lessened by taking the difference of two data sets.

An alternative approach to temporal difference imaging is multifrequency imaging, where with the same intent of acquiring two data sets and lessening geometrical errors by subtraction, voltages are collected at two different frequencies [53] [54] [55]. The reconstructed image depends on how the tissue electrical properties change with frequency.

Dynamic imaging

Absolute imaging techniques assume the conductivity to be constant during the acquisition period. Differential imaging techniques assume the conductivity to be slowly varying, in such a way that two successive data sets differ, but the conductivity can be approximated as constant during the acquisition period. The aim of dynamic imaging is to reconstruct fast conductivity changes. The conductivity is assumed to vary rapidly in comparison with the time period needed by the instrumentation to acquire a single image [10]. The changes are however supposed to be slow with respect to the acquisition period of a single pattern.

Seppänen applied these methods for imaging fast flowing liquids transporting resistive objects [92] [93]. Vauhkonen applied the method in the context of thoracic imaging, with the purpose of imaging cardiac related phenomena.

The reconstruction is stated as a state estimation problem, where the conductivity is not only a function of space but also of discrete time t

$$\sigma^{(t+1)} = F^{(t)}\sigma^{(t)} + \mathbf{w}^{(t)} \tag{3.9}$$

where the matrix $F^{(t)}$ is a transition matrix and $\mathbf{w}^{(t)}$ is a state noise process. The observation model is written linearising the forward operator

$$\mathbf{d}^{(t)}(\sigma^{(t)}) = \mathbf{d}^{(t)}(\sigma_0^{(t)}) + J^{(t)}(\sigma_0^{(t)})(\sigma^{(t)} - \sigma_0^{(t)}) + \nu^{(t)} \tag{3.10}$$

where $\mathbf{d}^{(t)}$ indicates the part of the data vector that has been collected at time t, and $J^{(t)}$ the Jacobian relative to that particular data, $\boldsymbol{\nu}^{(t)}$ is a measurement noise process.

The aim of the method is not to estimate a single conductivity distribution $\boldsymbol{\sigma}$, but instead to estimate a sequence of states $\{\boldsymbol{\sigma}^{(t)}, \boldsymbol{\sigma}^{(t+1)}, \ldots, \boldsymbol{\sigma}^{(t+n)}\}$.

Equations (3.9) and (3.10) constitute what is called state-space representation of the linearised EIT problem. The transition matrix $F^{(t)}$ is used to incorporate prior information on the temporal evolution of the particular phenomena under study. The choice $F^{(t)} = I$ corresponds to the random walk model; other choices can be made considering the physical laws that govern the temporal phenomena [92] [93]. The matrix $F^{(t)}$ has a role of temporal regularisation; spatial regularisation can be included in (3.10) with traditional techniques.

Chapter 4

Forward Model

4.1 Introduction

In inverse problems a forward model is used to predict observations. In the specific case of EIT, a model that predicts the voltage measurements given the current stimuli and the conductivity distribution must be derived. In the first part of this chapter such a mathematical model is derived from Maxwell's equations. It is assumed that the object under measurement has a conductivity which is linear and isotropic, and that the electric and magnetic fields are slowly varying. The presence of the electrodes is taken into account via appropriate boundary conditions. The second part of the chapter describes the numerical solution of the forward problem with the Finite Element Method. Lastly some numerical experiments and a validation of the forward model are briefly described.

4.1.1 Body Electrical Model

The typical experiment in EIT consists in applying some currents to a subset of a set of electrodes attached to a body Ω and in measuring the resulting voltages on the remaining electrodes. The applied currents are sinusoidal with a frequency in the range of 10kHz 1MHz. In order to derive the model, the Faraday and Ampere laws can be formulated in the differential form yielding to the Maxwell equations

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.1a}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{4.1b}$$

where, as in the usual convention, $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic induction, $\mathbf{H}$ is the magnetic field, $\mathbf{J}$ the current density vector and $\mathbf{D}$ the dielectric induction. The

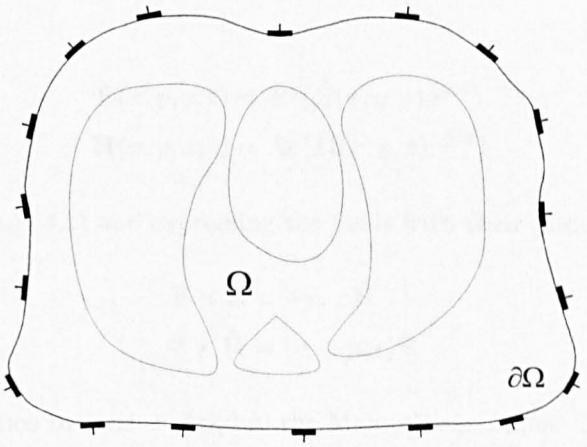


Figure 4.1: Typical EIT experiment.

current density vector **J** in (4.1) is due to purely conductive phenomena: no internal current sources are supposed to be acting on the measurements. The assumption is true for electrically passive imaged objects. In biological experiments, where the imaged object can be a source of bioelectric fields, the rejection of stimuli is achieved by frequency selective voltage detection, as the applied stimuli are at a much higher frequency than physiological signals. As a further simplification, it is general practice to assume the imaged media to be linear and isotopic. The first assumption can be justified by the small amplitude of the stimuli (few mA) and of the resulting voltages (few mV), while the latter is a simplification of the real phenomena which are known to be anisotropic [94]. If the aforementioned hypotheses are however accepted for simplicity, the following constitutive relations hold

$$\mathbf{J} = \sigma \mathbf{E} \tag{4.2a}$$

$$\mathbf{D} = \epsilon \mathbf{E} \tag{4.2b}$$

$$\mathbf{B} = \mu \mathbf{H} \tag{4.2c}$$

where σ is the electric conductivity, ϵ is the dielectric permittivity and μ the magnetic permeability. The applied stimulus in EIT is a single frequency sinusoidal waveform, it is convenient therefore to express all the fields with their phasors, indicated with the

circumflex accent

$$\mathbf{E}(x,y,z,t) = Re\{\hat{\mathbf{E}}(x,y,z)e^{j\omega t}\} \tag{4.3a}$$

$$\mathbf{H}(x,y,z,t) = Re\{\hat{\mathbf{H}}(x,y,z)e^{j\omega t}\} \tag{4.3b}$$

Substituting (4.2) into (4.1) and expressing the fields with their phasors as in (4.3) gives

$$\nabla \times \hat{\mathbf{E}} = -j\omega\mu\hat{\mathbf{H}} \tag{4.4a}$$

$$\nabla \times \hat{\mathbf{H}} = (\sigma + j\omega\epsilon)\hat{\mathbf{E}} \tag{4.4b}$$

It is a common practice to further simplify the Maxwell equations. The injected currents are assumed to be "slowly varying", so that the electric field is conservative and the conduction currents are dominant with respect to the displacement currents, yielding

$$\nabla \times \hat{\mathbf{E}} = 0 \tag{4.5a}$$

$$\nabla \times \hat{\mathbf{H}} = \sigma\hat{\mathbf{E}} \tag{4.5b}$$

Dropping for simplicity, from now on the phasor notation, taking the divergence at both the sides of (4.5b), and substituting (4.5b) in (4.5a) gives

$$\nabla \cdot \sigma\nabla u = 0 \tag{4.6}$$

which is the equation that governs the electric potential u inside the body Ω.

4.1.2 Electrode Models

In order to describe the electrode interactions with the body, appropriate models for the electrodes should be developed. Current injection and voltage measurement are accounted by imposing appropriate boundary conditions. Several electrode models have been developed for use in EIT. In this paragraph we describe briefly four of the most important models, in order of complexity, focusing on the *Complete Electrode Model* that at present is the most sophisticated description of the electrode/body interaction.

Continuum Electrode Model

The Continuum Model is the simplest of the models used in electrical impedance tomography. The electrodes themselves are not modelled as discrete conductors facing the boundary of the object, but the model assumes that the current density can be set arbitrarily

on all of the boundary of the object

$$\mathbf{J}_{int} \cdot \mathbf{n} = -\mathbf{J}_{inj} \cdot \mathbf{n} \qquad \text{on } \partial\Omega \tag{4.7}$$

where $\mathbf{n}$ is the outward pointing normal versor to $\partial\Omega$. Solely normal components have been considered, as electrodes (supposed to be perfect conductors) would shunt tangential components of the electric field. Using Ohm's law and recalling that E is conservative, (4.7) can be expressed as a Neumann boundary condition

$$\sigma \frac{\partial u}{\partial \mathbf{n}} = -\mathbf{J}_{inj} \cdot \mathbf{n} \qquad \text{on } \partial\Omega \tag{4.8}$$

Additionally the conservation of charge must be preserved

$$\int_{\partial\Omega} \mathbf{J}_{inj} \cdot \mathbf{n} = 0 \tag{4.9}$$

and the condition

$$\int_{\partial\Omega} u = 0 \tag{4.10}$$

defines a "ground" or reference voltage, making the model complete. The Continuum Model is a very crude approximation of the electrode/body interaction, forward predictions can differ sensibly from actual measurements [95].

Gap Electrode Model

The Gap Electrode Model is a refinement of the Continuum Model, it accounts for the discreteness of the electrodes. Electrodes are mathematically represented as a set of disjoint subdomains $\{\mathcal{E}_1, \ldots, \mathcal{E}_L\}$ of the boundary $\partial\Omega$. The model sets the injected current density to zero on each interelectrode gap

$$\sigma \frac{\partial u}{\partial \mathbf{n}} = 0 \qquad \text{on } \partial\Omega \setminus \{\mathcal{E}_1 \cup \ldots \cup \mathcal{E}_L\} \tag{4.11}$$

and to a constant underneath each electrode

$$\sigma \frac{\partial u}{\partial \mathbf{n}} = \frac{I_l}{A} \qquad \text{on } \mathcal{E}_l, \quad l = 1, \ldots, L \tag{4.12}$$

where I_l is the injected current at each electrode and A is the area of the electrode. The conservation of charge is stated in this case as

$$\sum_l I_l = 0 \tag{4.13}$$

Equations (4.11), (4.12), (4.13) with (4.10) define completely the Gap Electrode Model, which has been used widely as an improvement over the Continuum Model. However the assumption that the current density is constant on the interface of each electrode is an oversimplification for many practical EIT setups [95]. Secondly the shunting effect of the electrodes and their contact impedances are not accounted for.

Shunt Electrode Model

The Shunt Electrode Model refines the Gap Electrode model in the fact that the current density profile underneath electrodes, far from being constant, is not assumed to be known. Not anticipating any current density profile, the model simply states that the net flux of the current density through the contact surface must be equal to the injected current

$$\int \sigma \frac{\partial u}{\partial \mathbf{n}} = I_l \qquad \text{on } \mathcal{E}_l, \quad l = 1, \ldots, L \tag{4.14}$$

Secondly the model considers the conductive nature of the electrodes, which implies that the potential of each electrode $\mathcal{E}_l$ must be a constant V_l, taking into account in this way the shunting effect

$$u = V_l \qquad \text{on } \mathcal{E}_l, \quad l = 1, \ldots, L \tag{4.15}$$

The model is completed by a condition for defining the ground reference

$$\sum_l V_l = 0 \tag{4.16}$$

The Shunt Electrode Model ignores the contact impedance of the electrodes, it is therefore less accurate than the Complete Electrode Model [1].

Complete Electrode Model

The Complete Electrode Model is a refined description of the interaction between the electrodes and stimulated object [1] [96]. The model has proven able to predict tank measurements within the typical accuracy of a data acquisition system as verified by Cheng *et. al.* [95] and in Section 4.5. The model is a refinement of the Shunt Electrode

Model as the effect of contact impedances is accounted for. As for the Shunt Model we have

$$\int \sigma \frac{\partial u}{\partial \mathbf{n}} = I_l \qquad \text{on } \mathcal{E}_l, \quad l = 1, \ldots, L \tag{4.17}$$

and on the interelectrode gaps (4.11) continues to hold. Considering secondly the conductive nature of the electrodes, and accounting for the drop across the contact impedance z_l, the potential u on $\partial\Omega$ and the potential V_l on the electrode obey the following relationship

$$V_l = u + z_l \sigma \frac{\partial u}{\partial \mathbf{n}} \qquad \text{on } \mathcal{E}_l, \quad l = 1, \ldots, L \tag{4.18}$$

Equations (4.11), (4.17), (4.18) together with a condition for defining the ground reference

$$\sum_l V_l = 0 \tag{4.19}$$

constitute the model, which is well posed and has a unique solution [1]. The Complete Electrode Model is currently the most refined model in use in EIT. However, the accuracy to which the model is able to describe medical experiments has been studied for laboratory phantoms, but not for in-vivo experiments yet.

4.1.3 Modelling the OXBACT-III setup

In this section some characteristics of the data acquisition system are introduced, and a model for the measurement setup is derived from the preceding analysis. Here, we briefly explain the characteristic aspects of the Oxford Brookes Adaptive Current Tomograph - mark III (OXBACT-III) that have a direct influence on the forward model. The reader is referred to consult Section 3.3.3 for an introduction to the adaptive current acquisition strategy. The OXBACT-III tomograph is an 64 channel EIT tomograph. The device uses 32 electrodes for voltage measurements (*voltage electrodes*) and 32 electrodes for current injection (*current electrodes*). The arrangement, for which some electrodes are dedicated solely to gathering measurements and others solely to injecting currents is called hybrid, as opposed to traditional systems where all the electrodes will assume in turns both the functionalities. Each of the current electrodes is connected to an independent current source; the system stimulates the object from different electrodes at the same time, with controlled amplitudes, applying a current pattern.

In Equations (4.20) the complete electrode model is adapted to describe the OXBACT-III measurement setup. The electrodes are split in two sets, indicated by a subscript, the voltage electrodes being $\mathcal{E}_v$ and the current electrodes $\mathcal{E}_c$, the number of voltage electrodes

being L_v and the number of current electrodes L_c. On the voltage electrodes the injected current is zero (4.20d) and on the current ones is equal to $\hat{I}_c$ (4.20e), the conservation of charge law (4.20g) and the reference voltage selection (4.20h) apply separately to the two sets of electrodes, resulting in

$$\nabla \cdot \sigma \nabla u = 0 \quad \text{on } \Omega \tag{4.20a}$$

$$u + z_{v_l} \sigma \frac{\partial u}{\partial \mathbf{n}} = V_{v_l} \quad \text{on } \mathcal{E}_{v_l}, \qquad l = 1, \ldots, L_v \tag{4.20b}$$

$$u + z_{c_l} \sigma \frac{\partial u}{\partial \mathbf{n}} = V_{c_l} \quad \text{on } \mathcal{E}_{c_l}, \qquad l = 1, \ldots, L_c \tag{4.20c}$$

$$\int_{\mathcal{E}_{v_l}} \sigma \frac{\partial u}{\partial \mathbf{n}} = 0 \qquad l = 1, \ldots, L_v \tag{4.20d}$$

$$\int_{\mathcal{E}_{c_l}} \sigma \frac{\partial u}{\partial \mathbf{n}} = I_{c_l} \qquad l = 1, \ldots, L_c \tag{4.20e}$$

$$\sigma \frac{\partial u}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega \setminus \{\{\mathcal{E}_{c_1} \cup \ldots \cup \mathcal{E}_{L_c}\} \cup \{\mathcal{E}_{v_1} \cup \ldots \cup \mathcal{E}_{L_v}\}\} \tag{4.20f}$$

$$\sum_{l=1}^{L_c} I_{c_l} = 0 \tag{4.20g}$$

$$\sum_{l=1}^{L_v} V_{v_l} = 0 \tag{4.20h}$$

Equation (4.20c) states the potentials V_{c_l} assumed by current electrodes during the stimulation, we are concerned usually just with the potential assumed by voltage electrodes V_{v_l}, which is the one the instrument gathers. Nevertheless (4.20c) is required to make the model complete, and to account for the shunt phenomena on current electrodes. In the next sections an implementation of the finite element method that solves (4.20) is discussed.

4.2 Forward Solving

4.2.1 Computational Techniques

For circular geometries of Ω, such as experimental tanks, analytic implementations have been developed [97] [98] [99] [51] [100], although their extension to other geometries is prevented by the complexity involved. Furthermore these methods have been used solely for linear reconstruction algorithms as they handle purely homogeneous conductivities. To treat general domains and to allow non-linear reconstruction, numerical techniques are needed. A review of the methods for bioelectrical field problems is offered by C. R.

Johnson [101]. The three most common methods for solving partial differential equations are the Finite Difference Method (FDM), the Finite Element Method (FEM) and the Boundary Element Method (BEM). The FD method discretises the domain Ω with a regular grid. The coordinates of each point of the grid are of the kind $x = lh$, $y = mh$ and $z = nh$ with l, m, n integers. The electric potential $u(x, y, z)$ is evaluated at each point, and denoted as $u(l, m, n)$. The Taylor's theorem is then used to provide the *difference equation*, for example to express $\frac{\partial}{\partial x}u(l, m, n)$ as a function of u in $(l-1, m, n)$, (l, m, n) and $(l+1, m, n)$, turning the partial differential operator into a linear algebraic operator. For each point in the grid the process is repeated, leading to a system of linear equations of the kind $A\mathbf{u} = \mathbf{b}$, which can be solved for the electric potential at grid points $\mathbf{u}$, given the boundary data specified by $\mathbf{b}$. The FDM is used frequently in geophysical applications of EIT [102] where the imaged underground is often modelled as an half space of $\mathbb{R}^2$. The geometrical simplicity of the domain makes the numerical method suitable for the application. When complex geometries are involved the FD method requires a very fine level of discretisation and special techniques to handle curved boundaries, which make the method less attractive for bioelectrical modelling. The FE and BE methods are often chosen as an alternative to FDM. The BEM operates a reduction of one degree on the dimensionality of the problem, by transforming any differential operator defined on a domain into an integral operator defined on a boundary. The dependent variable is assumed to be constant on each subdomain, so that the object model must be segmented in several regions, depending on its heterogeneity. The BE method is therefore best suited for applications where the domain is highly homogeneous, avoiding excessive levels of discretisation associated to the variations of the material. Additionally the BE method involves dense matrices, and therefore could be computationally disadvantageous.

The FE method consists in discretising the domain Ω in a number of non uniform, non overlapping, elements connected via nodes. The dependent variable is approximated within each element by an interpolating function, defined by the values of the variable at the nodes of the element. The Galerkin principle is then used to turn the original PDE into a set of integral equations for each nodal value. By combining the equations over the domain, the electric potential $\mathbf{u}$ at each node can be expressed in matrix notation $A\mathbf{u} = \mathbf{b}$. The FE method is very versatile with regards to the domain geometry, to boundary conditions and to heterogeneity of the domain, for these reasons it is has been the preferred choice in EIT modelling since the early stages [103] [81][104], and was, for the same reasons, adopted by us.

4.2.2 Finite Element Modelling

In September 1999, when this work began, two FEM forward solvers implementing the complete electrode model were known. A first implementation was proposed by K. Paulson *et. al.* [96]; a similar approach was followed by M. Vauhkonen in his doctoral research [10]. The work of Vauhkonen become publicly available, in the form of MATLAB software libraries, during the year 2000 under the project EIDORS.

It was decided, at Oxford Brookes University, to independently develop a FEM forward solver that would implement the complete electrode model. Several considerations contributed to the choice. The developments of Paulson, former researcher at Oxford Brookes University, could be made available. However the code was written in FORTRAN language, and our aim was to use the MATLAB environment. Secondly, the MATLAB libraries developed by Vauhkonen were not public at the time this work started. Thirdly the experience gained by independently developing the forward solver would contribute some benefit.

As briefly introduced in the previous paragraph, the FE method allows discretisation of the continuous problem (4.20) by subdividing the domain into non overlapping elements of simple shape, which are usually triangles or quadrangles for the bidimensional case. The vertices of the polygons are called nodes for the chosen discretisation. The electric potential is expressed as a linear combination of nodal basis functions. The nodal basis functions are polynomial interpolating functions – they assume the value of 1 at the node i and they decay to 0 on each other node; they are tent-like functions. The support of each nodal basis function is limited to the elements sharing the common node i. The approximated potential will be expressed as

$$\tilde{u} = \sum_{i=1}^{W} n_i \phi_i \tag{4.21}$$

where ϕ_i are the nodal basis functions, n_i are the nodal values assumed by the solution and W is the number of nodes. Given (4.21) the nodal values n that best approximate the continuous solution have to be determined. This can be done by the Galerkin or weighted residuals method, which is briefly described here. The same results can be obtained with the Rayleigh-Ritz or variational method [105]. Equation (4.20a) requires that

$$\nabla \cdot \sigma \nabla u = 0 \quad x \in \Omega \tag{4.22}$$

in general (4.22) will not be satisfied by $\tilde{u}$, and a residual r can be defined as

$$\nabla \cdot \sigma \nabla \tilde{u} = r \neq 0 \quad x \in \Omega \tag{4.23}$$

The method of the weighted residuals determines the W unknowns n_i in such a way that the error r over the entire solution domain is small. This is accomplished by forming a weighted average of the error, and by specifying that it should vanish over the domain. Hence W linearly independent test functions w_i are chosen and the condition

$$\int_{\Omega} [\nabla \cdot \sigma \nabla \tilde{u}] w_i = \int_{\Omega} r\, w_i = 0, \quad i = 1, \ldots, W \tag{4.24}$$

ensures that the residual r is small in some sense. Depending on the weighting functions w_i the error will be distributed differently across the domain. The Galerkin criterion suggests using as weighting functions the same functions used to represent u, that is $w_i = \phi_i$ for $i = 1, \ldots, W$. Thus condition (4.24) becomes

$$\int_{\Omega} [\nabla \cdot \sigma \nabla \tilde{u}] \phi_i = 0, \quad i = 1, \ldots, W \tag{4.25}$$

Equation (4.25) states that the error r is orthogonal to all the basis functions spanning the space of the approximate solutions. Consequently the error is orthogonal to $\tilde{u}$ and must the minimum error given the discretisation. In the preceding discussion we assumed we were dealing with the entire solution domain, however since (4.20a) holds for any region in the domain, it holds if we restrict our attention to a single element E

$$\int_{E_k} [\nabla \cdot \sigma \nabla \tilde{u}] \phi_{i_{loc}} = 0, \quad i_{loc} = 1, \ldots, R \tag{4.26}$$

where k is the element number, i_{loc} is local nodal index relative to element E_k, and R is the number of nodes on the single element. Boundary conditions can be expressed applying Gauss' theorem to (4.26), allowing in this way the description of the electrodes

$$\int_{E_k} \sigma \nabla \phi_{i_{loc}} \cdot \nabla \tilde{u} - \int_{\partial E_k \cap (\mathcal{E}_v \cup \mathcal{E}_c)} \phi_{i_{loc}} \sigma \frac{\partial \tilde{u}}{\partial \mathbf{n}} = 0, \quad i_{loc} = 1, \ldots, R \tag{4.27}$$

The boundary integral needs to be carried out only for elements underneath electrodes. It enables us to take into account the boundary conditions expressed by (4.20b) or (4.20c)

depending on whether E_k is touching a voltage electrode or a current electrode

$$\int_{E_k} \sigma \nabla \phi_{i_{loc}} \cdot \nabla \tilde{u} - \int_{\partial E_k \cap \mathcal{E}_{v_l}} \phi_{i_{loc}} \frac{V_{v_l} - \tilde{u}}{z_{v_l}} = 0, \quad i_{loc} = 1, \ldots, R \tag{4.28a}$$

$$\int_{E_k} \sigma \nabla \phi_{i_{loc}} \cdot \nabla \tilde{u} - \int_{\partial E_k \cap \mathcal{E}_{c_l}} \phi_{i_{loc}} \frac{V_{c_l} - \tilde{u}}{z_{c_l}} = 0, \quad i_{loc} = 1, \ldots, R \tag{4.28b}$$

where l is the electrode number. By applying equations (4.28) to every element, W different conditions on the nodal variables n_i are specified. The model has $W + L_v + L_c$ unknowns: the W nodal potentials n_i, and the potentials of the two sets of electrodes. The remaining $L_v + L_c$ conditions are specified using (4.20d) and (4.20e), substituting (4.20b) or (4.20c) into them, obtaining

$$\int_{\mathcal{E}_{v_l}} \frac{V_{v_l} - \tilde{u}}{z_{v_l}} = \sum_k \int_{\partial E_k \cap \mathcal{E}_{v_l}} \frac{V_{v_l} - \tilde{u}}{z_{v_l}} = 0 \qquad l = 1, \ldots, L_v \tag{4.29a}$$

$$\int_{\mathcal{E}_{c_l}} \frac{V_{c_l} - \tilde{u}}{z_{c_l}} = \sum_k \int_{\partial E_k \cap \mathcal{E}_{c_l}} \frac{V_{c_l} - \tilde{u}}{z_{c_l}} = \hat{I}_{c_l} \qquad l = 1, \ldots, L_c \tag{4.29b}$$

where the contact impedance is supposed to be greater than zero.

Equations (4.28) and (4.29) therefore completely define the model in its discrete form. Substituting the definition of $\tilde{u}$ into them, and arranging all the equations in matrix form, the following linear system in $W + L_v + L_c$ unknowns is obtained

$$Y\mathbf{p} = \mathbf{c} \tag{4.30}$$

where the vector of unknowns $\mathbf{p}$ is arranged in such a way that the topmost entries are the nodal potential values, followed by the voltage electrodes potentials and by the current electrodes potentials

$$\mathbf{p} = \begin{bmatrix} \mathbf{n} \\ V_v \\ V_c \end{bmatrix} \tag{4.31}$$

the Y matrix is partitioned and will assume the following form

$$Y = \begin{bmatrix} Y' + Y'' & Y_v & Y_c \\ Y_v^T & D_v & 0 \\ Y_c^T & 0 & D_c \end{bmatrix} \tag{4.32}$$

and the right hand side will be

$$\mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ I_c \end{bmatrix} \tag{4.33}$$

with the blocks of the Y matrix being

$$Y'(i,j) = \int_\Omega \sigma \nabla\phi_i \cdot \nabla\phi_j \quad i,j = 1,\ldots,W \tag{4.34a}$$

$$Y''(i,j) = \sum_{l=1}^{L_v} \frac{1}{z_{v_l}} \int_{\mathcal{E}_{v_l}} \phi_i \phi_j + \sum_{l=1}^{L_c} \frac{1}{z_{c_l}} \int_{\mathcal{E}_{c_l}} \phi_i \phi_j \quad i,j = 1,\ldots,W \tag{4.34b}$$

$$Y_v(i,j) = -\sum_{l=1}^{L_v} \frac{1}{z_{v_l}} \int_{\mathcal{E}_{v_l}} \phi_i, \quad i = 1,\ldots,W \tag{4.34c}$$

$$Y_c(i,j) = -\sum_{l=1}^{L_c} \frac{1}{z_{c_l}} \int_{\mathcal{E}_{c_l}} \phi_i, \quad i = 1,\ldots,W \tag{4.34d}$$

$$D_v = \operatorname{diag}\left(\frac{\operatorname{meas}(\mathcal{E}_{v_l})}{z_{v_l}}\right) \quad l = 1,\ldots,L_v \tag{4.34e}$$

$$D_c = \operatorname{diag}\left(\frac{\operatorname{meas}(\mathcal{E}_{c_l})}{z_{c_l}}\right) \quad l = 1,\ldots,L_c \tag{4.34f}$$

By choosing the element geometry and nodal basis functions the system matrix Y is completely specified. For the implementation of our two-dimensional forward solver we have adopted triangular elements, with a piecewise representation of the conductivity, and linear nodal basis functions.

4.2.3 Assembly of Y

The most intensive task in assembling the system matrix is to calculate the block Y'. All the remaining integrals in (4.34) are line integrals over electrodes, involving a small number of elements. We recall the formulation of the matrix:

$$Y'(i,j) = \int_\Omega \sigma \nabla\phi_i \cdot \nabla\phi_j \tag{4.35}$$

Since the interpolation functions ϕ is linear, the gradients are constants over each triangle. Similarly the conductivity is chosen to be piecewise constant. The integration over each element therefore reduces to the multiplication of three constant terms times the area of the element itself. Given these considerations, the whole matrix Y' can be calculated with linear algebra operations, if appropriate data structures are precomputed. Two gradient

operators G_x and G_y must be part of this structures. The gradient operators are matrices that return x and y components of the gradient of the field variable, when applied to the vector of nodal values: $\nabla_x \tilde{u} = G_x N$, $\nabla_y \tilde{u} = G_y N$, the elements of the matrices being

$$G_x(i,j) = \nabla_x^{(\mathrm{E}_j)} \phi_i \quad \text{and} \quad G_y(i,j) = \nabla_y^{(\mathrm{E}_j)} \phi_i \tag{4.36}$$

where the superscript $^{(\mathrm{E}_j)}$ is the restriction to the element E_j. The operators ∇_x and ∇_y can be expressed in the form of matrices, since the gradient on each element is a linear combination of the nodal values assumed by the field variable. Given these structures, the whole matrix Y' can be calculated as

$$Y' = G_x^T A \Gamma G_x + G_y^T A \Gamma G_y \tag{4.37}$$

where A is a diagonal matrix containing the area of the triangles and Γ is a diagonal matrix containing the conductivity of each element. For non-linear reconstructions, where the conductivity is updated during iterations, the admittance Y' is recalculated rapidly, as only the matrix Γ needs changes. Formula (4.37) can be derived from the expression of the elements of Y'

$$\begin{aligned} Y'(i,j) = \sum_k \int_{\mathrm{E}_k} \sigma \, \nabla\phi_i \cdot \nabla\phi_j = \\ = \sum_k A_k \, \Gamma_k \, \nabla^{(\mathrm{E}_k)} \phi_i \cdot \nabla^{(\mathrm{E}_k)} \phi_j = \\ = \sum_k A_k \, \Gamma_k \, (\nabla_x^{(\mathrm{E}_k)} \phi_i)^T \, \nabla_x^{(\mathrm{E}_k)} \phi_j + A_k \, \Gamma_k \, (\nabla_y^{(\mathrm{E}_k)} \phi_i)^T \, \nabla_y^{(\mathrm{E}_k)} \phi_j \end{aligned} \tag{4.38}$$

where k is the index over elements belonging to the support of the dot product of gradients.

The matrix Y' is sparse. As illustrated in Figure 4.2.3, the functions ϕ_1 and ϕ_2, basis functions of the nodes N_1 and N_2, have the respective supports $\mathrm{E}_1 \cup \mathrm{E}_2 \cup \mathrm{E}_3 \cup \mathrm{E}_4 \cup \mathrm{E}_5$ and $\mathrm{E}_1 \cup \mathrm{E}_2 \cup \mathrm{E}_6 \cup \mathrm{E}_7 \cup \mathrm{E}_8$, consequently the support of the dot product is: $\mathrm{E}_1 \cup \mathrm{E}_2$. The supports of two nodal basis functions intersect only if the nodes are neighbours; in general, for most of the index couples (i,j) the supports of ϕ_i and ϕ_j will be disjoint, and corresponding elements of Y' null. Assembling the matrix as in (4.37) exploits MATLAB linear algebra, without resorting to `for` loops over elements and nodes, which are slower compared to vector and matrix operations in this particular language. The operation should be performed with sparse matrix routines, since the final product is sparse as well as all the operands. The matrices A and Γ are diagonal and G_x, G_y have only three non-null element per row: the gradient on each triangular element depends on the three nodal

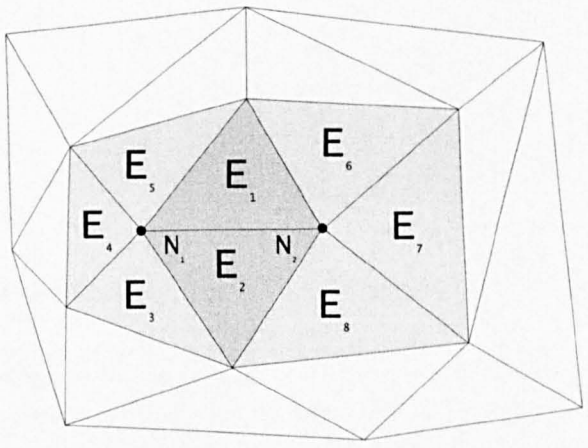


Figure 4.2: The supports of nodal basis functions are in general disjoint, they intersect only for neighbouring nodes as illustrated. This gives rise to the sparsity of the system matrix.

values. In order to complete the assembly of the matrix Y the blocks $Y'', Y_v, Y_c D_v, D_c$ should be calculated. I have not found a straightforward way to express these calculations in terms of linear algebra operations as for Y'.

4.2.4 Solving the forward problem

The vector of unknown potentials p is calculated solving the linear system

$$Y\,\mathbf{p} = \mathbf{c} \tag{4.39}$$

which requires Y to be invertible. A reference potential has not been specified yet, Y is rank deficient, and therefore not invertible. The Complete Electrode Model specifies a reference potential through (4.20h). Condition (4.20h) involves V_{v_l}, which are the unknowns be to be solved for, thus it is not readily applicable. The usual way to proceed is to choose an arbitrary *ground node* and to apply Dirichlet boundary conditions to it. The matrix Y is then symmetric, positive defined and invertible. This ensures that the solution to (4.39) exists and is unique. Each electrode potential is computed applying a measurement operator $\mathbf{m}_l \in \mathbb{R}^{1\times(W+L_c+L_v)}$ to the unknowns vector

$$V_{v_l} = \mathbf{m}_l^T\,\mathbf{p}, \quad l = 1, \ldots, L_v \tag{4.40}$$

(4.20h) will be satisfied by an appropriate choice of $\mathbf{m}_l$, for example

$$\mathbf{m}_l = \begin{pmatrix} 0_W \\ 0_{L_c} \\ \delta(1,l) \\ \vdots \\ \delta(L_v,l) \end{pmatrix} - \begin{pmatrix} 0_W \\ 0_{L_c} \\ 1/L_v \\ \vdots \\ 1/L_v \end{pmatrix} \tag{4.41}$$

where 0_W and 0_{L_c} are column zero vectors of length W and L_c and $\delta(i,j)$ is Kronecker's delta.

Concerning the actual solution of (4.39), in 2D EIT, where the number of unknowns is not excessively large, direct methods such Cholesky factorisation can be used. The system matrix is factorised as $Y = R^T R$, where R is an upper triangular sparse matrix. Equation 4.39 becomes

$$R^T R\,\mathbf{p} = \mathbf{c} \tag{4.42}$$

and using a dummy variable $\mathbf{q} = R\,\mathbf{p}$

$$R^T \mathbf{q} = \mathbf{c} \tag{4.43}$$

firstly (4.43) is solved with respect to $\mathbf{q}$, $\mathbf{p}$ is calculated then as

$$R\,\mathbf{p} = \mathbf{q} \tag{4.44}$$

The steps (4.43) and (4.44) are called back and forward substitution, they are trivial since R^T and R are triangular. The sparsity of R is in general different from the sparsity of Y, the factorisation usually results in some additional non-null elements, called *fill-ins*. Depending on the structure of Y, the number of fill-ins can be considerable. As an example Figure 4.2.4 shows sparsity plots relative to a simulation involving a system matrix $\in \mathbb{R}^{539\times539}$ with 3669 non-null elements, which rose to 37155 after factorisation. To maximise the sparsity of R, reordering techniques such as the MATLAB symmmd routine can be used, which by permuting the columns and rows of Y optimises the number of fill-ins that occur during factorisation.

For small problems, solving 4.39 via Cholesky decomposition is advantageous over iterative methods, as the system be can solved for different RHS computing the factorisation $Y = R^T R$ only once. The repeated solution of 4.39 for several current patterns is computationally inexpensive as just back and forward substitutions are to be computed. Forward solutions $\{V_v^{(1)}, \ldots, V_v^{(P)}\}$ corresponding to different current patterns $\{\mathbf{c}_1, \ldots, \mathbf{c}_P\}$ are ar-

ranged into the measured data vector $\mathbf{d} \in \mathbb{R}^{L_v P \times 1}$

$$\mathbf{d} = \begin{pmatrix} V_v^{(1)} \\ \vdots \\ V_v^{(P)} \end{pmatrix} \tag{4.45}$$

which forms the output of the forward solver.

4.3 Calculation of the Jacobian matrix

The Jacobian matrix, defined as

$$\begin{pmatrix} \frac{\partial \mathbf{d}_1}{\partial \boldsymbol{\sigma}_1} & \cdots & \frac{\partial \mathbf{d}_1}{\partial \boldsymbol{\sigma}_K} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{d}_M}{\partial \boldsymbol{\sigma}_1} & \cdots & \frac{\partial \mathbf{d}_M}{\partial \boldsymbol{\sigma}_K} \end{pmatrix} \tag{4.46}$$

needs to be calculated for use with Gauss–Newton type reconstruction algorithms. The calculation of the matrix can be computationally very intensive, as it requires several forward solutions to be computed. Three calculation methods are known: the *perturbation method* (numerical derivation based on the definition of derivative), *Yorkey's method* [81] and the *lead field method* [22] [106][103]. Both the Perturbation and Yorkey's Methods require one to perform a number of forward solutions equal to the number of model elements, which is an expensive requirement. We adopted the lead field method, as fewer forward solutions are needed. The method requires at most $L_v + L_c - 2$ forward solutions, which, in practical situations, is much smaller than the number of model elements.

The lead field method is explained in the following on the basis of the linear algebra and for the FEM context, an analytic derivation can be found in [106]. Recalling firstly (4.40) and (4.45), each element of the Jacobian can be expressed as

$$J_{i,j} = \frac{\partial \mathbf{d}_i}{\partial \boldsymbol{\sigma}_j} = \frac{\partial (\mathbf{m}_l^T Y^{-1} \mathbf{c}_p)}{\partial \boldsymbol{\sigma}_j} = \mathbf{m}_l^T \frac{\partial (Y^{-1} \mathbf{c}_p)}{\partial \boldsymbol{\sigma}_j} = -\mathbf{m}_l^T Y^{-1} \frac{\partial Y}{\partial \boldsymbol{\sigma}_j} Y^{-1} \mathbf{c}_p \tag{4.47}$$

where $\mathbf{c}_p$, with $p = p(i)$, and $\mathbf{m}_l$, with $l = l(i)$, are the current pattern and the measurement operator from which results the observation $\mathbf{d}_i$. Using the symmetry of Y, the last term on the right of (4.47) becomes

$$-\mathbf{m}_l^T Y^{-1} \frac{\partial Y}{\partial \boldsymbol{\sigma}_j} Y^{-1} \mathbf{c}_p = -((Y^{-1})^T \mathbf{m}_l)^T \frac{\partial Y}{\partial \boldsymbol{\sigma}_j} Y^{-1} \mathbf{c}_p = -(Y^{-1}\, \mathbf{m}_l)^T \frac{\partial Y}{\partial \boldsymbol{\sigma}_j} Y^{-1} \mathbf{c}_p \tag{4.48}$$

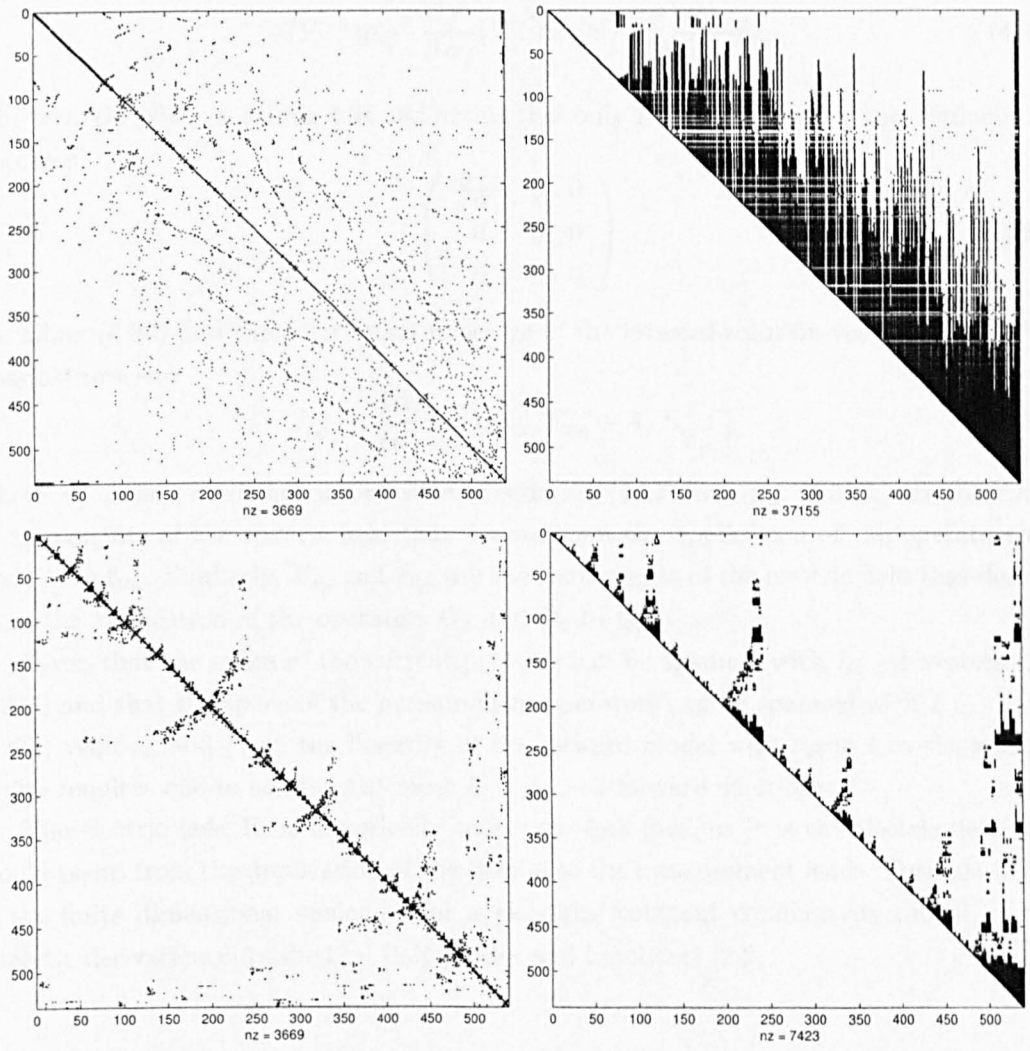


Figure 4.3: Sparsity plots of a system matrix $Y \in \mathbb{R}^{539\times539}$. On the top left, sparsity of the unordered matrix, and on the top right sparsity of the factorisation. On the bottom left, sparsity of the symmmd reordered Y and on the bottom right sparsity of its Cholesky factorisation.

the vectors $(Y^{-1}\,\mathbf{c}_p)$ and $(Y^{-1}\,\mathbf{m}_l)$ can be recognised to be forward solutions vectors corresponding respectively to the actual current pattern and to a current pattern numerically equal to the measurement operator. We will label them respectively $\mathbf{f}_{\mathbf{c}_p}$ and $\mathbf{f}_{\mathbf{m}_l}$, thus

$$-(Y^{-1}\,\mathbf{m}_l)^T\,\frac{\partial Y}{\partial \boldsymbol{\sigma}_j}(Y^{-1}\mathbf{c}_p) = -\mathbf{f}_{\mathbf{m}_l}^T\,\frac{\partial Y}{\partial \boldsymbol{\sigma}_j}\mathbf{f}_{\mathbf{c}_p} \tag{4.49}$$

The term $\partial Y/\partial \boldsymbol{\sigma}_j$, recalling 4.34 and noting that only Y' is dependent on the conductivity becomes

$$\begin{pmatrix} \frac{\partial Y'}{\partial \boldsymbol{\sigma}_j} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.50}$$

Recalling (4.37) and the partitioned structure of the forward solution vectors (4.31), it is possible to write

$$J_{i,j} = \frac{\partial \mathbf{d}_i}{\partial \boldsymbol{\sigma}_j} = A_j\,E_{x_l}^T\,E_{x_p} + A_j\,E_{y_l}^T\,E_{y_p} \tag{4.51}$$

where A_j is the j-th diagonal term of A of equation (4.37), and E_{x_l} and E_{y_l} are the x and y components of the electric field that derives from the application of the operators G_x and G_y to $\mathbf{f}_{\mathbf{m}_l}$. Similarly, E_{x_p} and E_{y_p} are the components of the electric field that derives from the application of the operators G_x and G_y to $\mathbf{f}_{\mathbf{c}_p}$.

Given that the space of the current patterns can be spanned with $L_c - 1$ vectors (see 4.20g) and that the space of the measurement operators can be spanned with $L_v - 1$ (see 4.20h) vectors, and given the linearity of the forward model with respect to the stimuli, (4.51) requires one to compute at most $L_c + L_v - 2$ forward solutions.

The electric field $\mathbf{E}_l$ is historically called the *lead field*, as it is the electric field that would result from the application of the stimuli to the measurement leads. Formula (4.51) is the finite dimensional analogue, for a piecewise constant conductivity model, of the analytic derivation published by Polydorides and Lionheart [22].

4.4 Simulations

In this paragraph are presented some forward simulations over a mesh resembling the Oxford Brookes experimental tank. It is highlighted, in particular, the ability of the Complete Electrode Model to describe correctly the phenomena at electrode interfaces. The simulations show the great complexity of the current density profiles, indicating that simpler choices of electrode models (i.e. flat current density profile under the electrode) are not justified and could significantly reduce the accuracy of the forward solution.

4.4.1 Meshing

Two meshes are required by our implementation of the forward solver: a coarser one for the representation of the conductivity, and a finer one for the calculation of the forward solution. This is a common arrangement, which allows one to have a sufficiently small number of model parameters, while attaining the necessary accuracy by having a refined forward mesh. A free program called EMC2 (Édition de Maillages et Contours en 2 dimensions), which has been developed at the French institute INRIA, was used for meshing. EMC2 is a 2D mesh generator; the program works similarly to a CAD application, allowing the user to define objects through a graphic interface (see Figure 4.4). The procedure we adopt to create the two meshes is: first, generate the coarser mesh in EMC2; secondly, import the mesh into MATLAB and, lastly, use adaptive refinement procedures to generate a suitable finer mesh. The adaptation procedure is intended to refine the mesh where the forward solution is most likely to suffer from discretisation errors. The procedure works in the following manner: at the first iteration a forward solution is computed over the coarse mesh. Error estimates are then computed with MATLAB's function pdejmps, which calculates the posteriori error as a weighted sum of the discontinuities of the current density's normal component across each element (as a reference on adaptation techniques the reader is referred to [107]). Elements with the largest error are then refined with a procedure based on MATLAB's refinemesh. The routine takes care to update structures associated with the mesh during its operation, as for example, the lists of nodes under each electrode which can vary if elements in contact with electrodes are refined. If necessary, the forward solution can be recomputed on the newly generated finer mesh, and the whole procedure repeated until the desired refinement level is obtained. Figure 4.5 shows on the left the coarse mesh obtained from EMC2 for the Brookes tank, and on the right the corresponding refinement. The coarse mesh has 510 elements, while the fine has 12444 elements. Figure 4.6 shows a magnification of the refined mesh. As the electric field is more intense close to the electrodes, a finer discretisation occurs there. On the right of

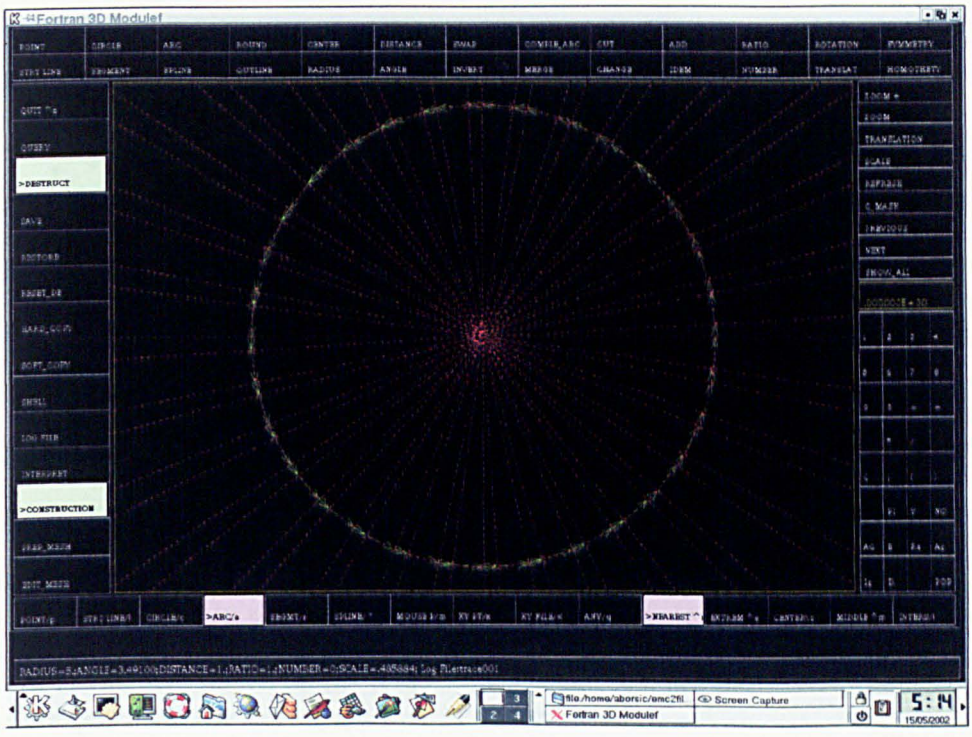


Figure 4.4: INRIA's EMC2 mesh generator running under Linux. Creation of the Brookes tank mesh.

the figure is shown a detail of the mesh; the refinement is the finest at the edges of the electrode, as in these regions the electric field presents peaks as illustrated by Figures 4.8 and 4.9.

4.4.2 Forward Solutions

Illustrative simulations were performed over the Brookes tank mesh. Stimuli were applied through 32 current electrodes, equispaced on the boundary of the disk, and covering 30% of the lateral surface. An homogeneous resistivity distribution was set, with a value of 60 $\Omega \cdot \mathrm{cm}$, and contact impedances were set to 15 $\Omega \cdot \mathrm{cm}^2$. The first two simulations, in Figure 4.7, show the electric potential and the current density on the boundary of the tank for the first trigonometric current pattern. As shown on the left part of the figure, the current density is discontinuous at electrode edges, and non uniform on the electrodes themselves. The resulting potential presents ripples corresponding to the variability of the injected current density. Figure 4.8 shows in detail the current density for two different electrodes. On the left of the figure is illustrated the current density for an injecting electrode. The

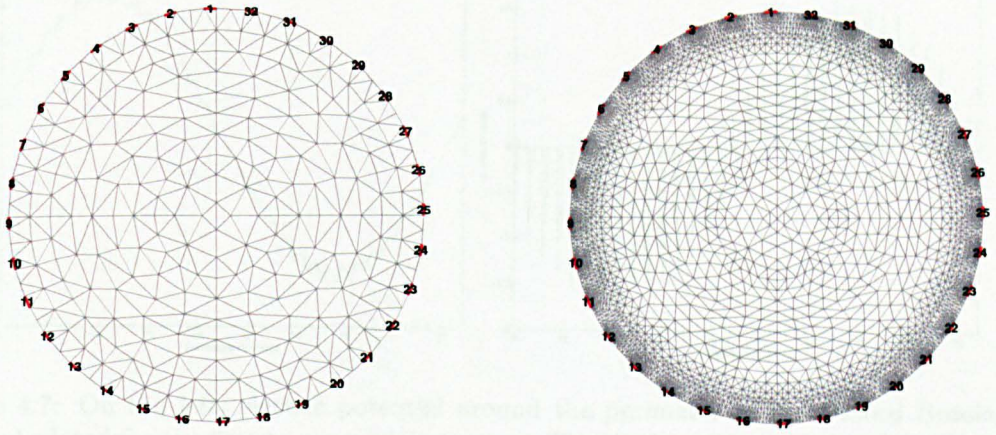


Figure 4.5: On the left is the coarser mesh, created with EMC2, used for the representation of conductivity. On the right, the finer mesh, used for the calculation of the forward solution. The finer mesh is obtained by an adaptive refinement of the coarser mesh. Initially electrodes are defined on the coarser mesh (red thick lines), the adaptive refinement procedure generates the finer mesh from it.

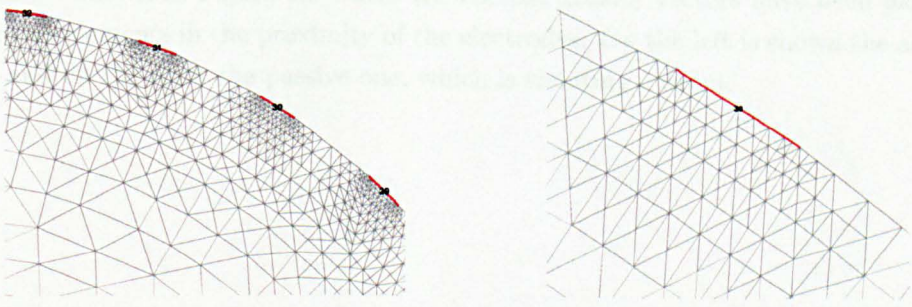


Figure 4.6: Details of the refined mesh. On the left it is possible to appreciate how the mesh is finer in the proximity of the electrodes, where the electric field is more intense. On the right, detail of a single electrode. The refinement of the mesh is finer near the edges, where the current density peaks cause an intense electric field (see Figures 4.8 and 4.9).

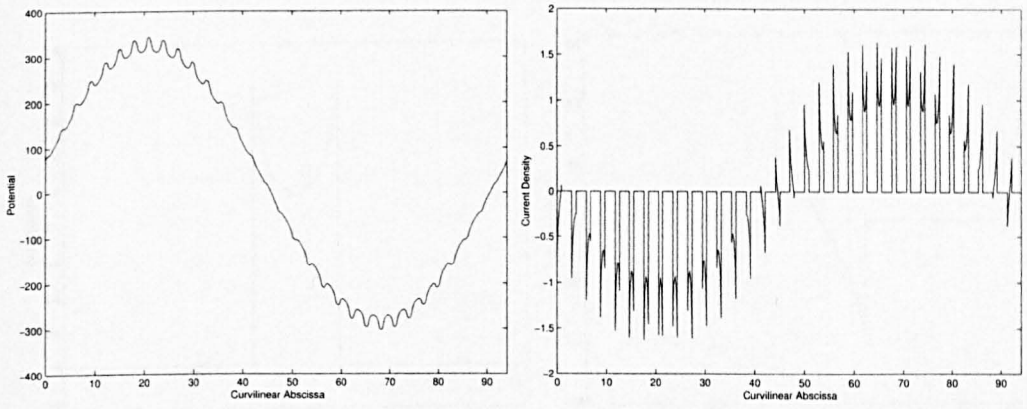


Figure 4.7: On the left, electric potential around the perimeter of the Oxford Brookes tank calculated for the first trigonometric pattern. The ripples are caused by the presence of the electrodes. On the right, current density on the perimeter of the tank, for the same stimulus. The current density is null in between electrodes, and non constant over electrodes.

applied current is positive, resulting in a net flux of charges across the electrode. On the right, is illustrated the current density for a passive electrode. The electrode is subject to a difference of potential, created by stimuli applied to other electrodes. As the electrode is a conductor, a flux of charges is crossing it. The phenomenon is called the *shunting effect*. The net flux across the electrode is null, since no current is applied to it. Similar phenomena are shown in Figure 4.9 where the current density vectors have been plotted for a few mesh elements in the proximity of the electrodes. On the left is shown the active electrode, and on the right the passive one, which is shunting current.

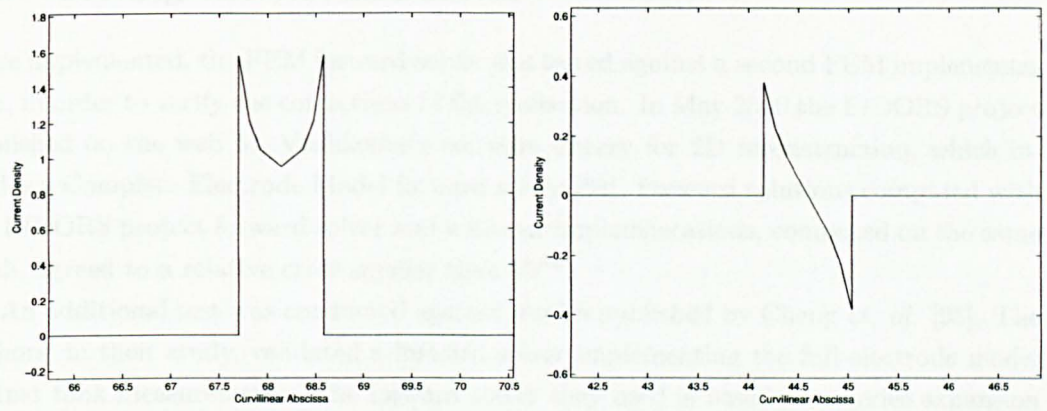


Figure 4.8: Details of the current density at two electrode interfaces. On the left is the plot of the current density of an active electrode: the net current is positive. On the right a passive electrode, subject to a difference of potential produced by other electrodes, is shunting current. Charges are entering the electrode and leaving it, the net flux is zero.

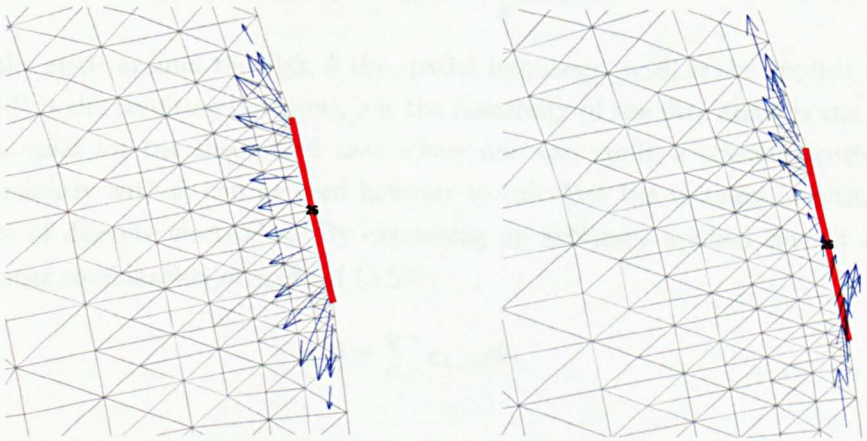


Figure 4.9: Details of the current density in the proximity of the two electrodes of Fig. 4.8. The current density vectors (blue arrows) are shown for few elements close to the electrode (red thick line). On the left, plot for the active electrode. On the right, plot for the passive electrode subject to shunting.

4.5 Testing the forward solver

Once implemented, the FEM forward solver was tested against a second FEM implementation, in order to verify the correctness of the realisation. In May 2000 the EIDORS project published on the web M. Vauhkonen's software library for 2D reconstruction, which includes a Complete Electrode Model forward solver [56]. Forward solutions computed with the EIDORS project forward solver and with our implementations, computed on the same mesh, agreed to a relative error smaller than 10^{-8}.

An additional test was conducted against results published by Cheng *et. al.* [95]. The authors, in their study, validated a forward solver implementing the full electrode model against tank measurements. The forward solver they used is based on a series expansion of the PDE [1] [96].

For a cylindrical homogeneous disk (such as the saline solution in a circular tank) it can be shown that when spatial sine or cosine current densities are applied to the boundary of the disk, the resulting potential on that surface is proportional to the applied current density [57]:

$$\begin{aligned} j_k(\theta) = \cos(k\,\theta) &\rightarrow u_k(\theta) = \frac{\rho\, b}{k}\cos(k\,\theta) \\ j_k(\theta) = \sin(k\,\theta) &\rightarrow u_k(\theta) = \frac{\rho\, b}{k}\sin(k\,\theta) \end{aligned} \tag{4.52}$$

where θ is the angle around the disk, k the spatial frequency, $j_k(\theta)$ is the applied current density, $u_k(\theta)$ is the resulting potential, ρ is the resistivity of the disk and b is the radius. Such result, valid for the theoretical case where one can apply a current density that varies continuously with θ, can be used however to calculate the voltages resulting from the presence of discrete electrodes. By expressing an arbitrary applied current density $J(\theta)$ as a linear combination of $j_k(\theta)$ of (4.52)

$$J(\theta) = \sum_k a_k\, j_k(\theta) \tag{4.53}$$

the resulting boundary potential $u(\theta)$ is

$$u(\theta) = \sum_k a_k \frac{\rho\, b}{k}\, j_k(\theta) \tag{4.54}$$

where a_k are the coefficients of the linear combination (Fourier coefficients).

Recalling (4.18), for the complete electrode model we have

$$z\,J(\theta) = V_l - u(\theta) \tag{4.55}$$

where z is the contact impedance (assumed constant across electrodes), V_l is the voltage on each l–th electrode, and $u(\theta)$ the potential on the boundary of the disk. For an even number L of symmetrically placed electrodes, one can restrict himself to even functions of θ. By using (4.53) and (4.54), by multiplying each side of (4.55) by $\cos(\frac{L}{2}\theta)$ and by integrating over θ, (4.55) becomes

$$z\,\pi\,a_{(\frac{L}{2})} = \sum_{l=1}^{L}\left(V_l \int_{\mathcal{E}_l} \cos(\frac{L}{2}\theta) - \sum_k a_k \frac{\rho\, b}{k} \int_{\mathcal{E}_l} \cos(\frac{L}{2}\theta)\cos(k\,\theta)\right) \tag{4.56}$$

where $\mathcal{E}_l$ are the segments of the boundary underneath each electrode. The unknowns of the linear system of equations (4.56) are the coefficients a_k and the electrode voltages V. To make the system complete, conditions (4.17) can be applied

$$\sum_k a_k \int_{\mathcal{E}_l} \cos(k\,\theta) = I_l \tag{4.57}$$

where I_l is the current injected into the l–th electrode. In such a way, (4.56) and (4.57) can be solved to yield the coefficients a_k and the electrode potentials V_l.

In their experiment, Cheng *et. al.* measured the tank's *characteristic resistances*. Such resistances are the eigenvalues of the impendence matrix $R(\sigma)$ of the object (see Section 3.3.3), and can be measured as

$$\rho_i = \frac{V(i)^T I(i)}{I(i)^T I(i)} \tag{4.58}$$

where $I(i)$ is the current pattern that corresponds to the i -th eigenvector of $R(\sigma)$ and $V(i)$ is the resulting voltage vector.

The experiment consisted in measuring the characteristic resistances for four different bath resistivities: 284.0 $\Omega \cdot$ cm, 139.7 $\Omega \cdot$ cm, 62.3 $\Omega \cdot$ cm and 29.5 $\Omega \cdot$ cm. Measurements were then compared to forward simulations. The simulations were computed by feeding the forward model with the geometry of the tank and with the bath resistivity values, both known. The contact impedances of the electrodes, which depend on the bath concentration and on the AC stimuli's frequency [108], were not known. The authors found these values by data fitting procedures. They showed that, when their forward solver was fed with the proper contact impedance value, the forward solutions accurately matched the measured data.

We reproduced similar results with the FEM forward solver.

For us, it was possible to gather all the necessary information to reproduce the test: a circular tank with a diameter of 30 cm was used, the tank had 32 current electrodes, each with an area of 12 cm^2, the tank was filled with a saline solution to a depth of 4.24 cm. A FE mesh of the described tank was generated with EMC2 and imported into MATLAB. Forward solutions were computed setting the simulated bath conductivity to the four different values and using the fitted contact impedance values reported by the authors. Tables 4.1 and 4.2 list the results. Each of the tables reports the experimental characteristic resistances, the simulated values by Cheng and FEM calculated ones. The first table contains the results for the 284.0 $\Omega \cdot cm$ and 139.7 $\Omega \cdot cm$ baths, to which contact impedances of 58.0 $\Omega \cdot cm^2$ 30.5 $\Omega \cdot cm^2$correspond. The second for 62.3 $\Omega \cdot cm$ and 29.5 $\Omega \cdot cm$ baths, to which contact impedances of 15.0 $\Omega \cdot cm^2$ and 7.5 $\Omega \cdot cm^2$correspond. Simulated values were close to measured characteristic resistances. The relative error for each bath, averaged on all the spatial frequencies, resulted in the best case in a value of 0.3%, which is comparable to the typical accuracy specification of an EIT data acquisition system. For a particular experiment, the 284.0 $\Omega \cdot cm$ bath, the error was 2%. The discrepancy is particularly significant for the first spatial frequency (319.45 Ω against 312.0 Ω), while all the remaining values are consistent. As a test, the original mesh used for the simulations was refined several times, and experiments repeated. The predictions from the FEM solver remained unchanged (to the significant digits), indicating that the level of refinement was sufficient. Given the very good consistency of the characteristic resistances for the same bath and contact impedances, for spatial frequencies from 2 to 16, we believe that it is most likely that the values published by Cheng could be affected by a typographic error, and that the first spatial frequency should read 320 or 321 instead of 312.

Table 4.1: Comparison of experimental and calculated characteristic resistances.

ρ	284.0 $\Omega \cdot$ cm			139.7 $\Omega \cdot$ cm		
z_c	58.0 $\Omega \cdot cm^2$			30.5 $\Omega \cdot cm^2$		
Spat. Freq.	Exp	Cheng Sim.	FEM Sim.	Exp	Cheng Sim.	FEM Sim.
1	312.00	312.00	319.45	156.80	156.80	157.66
2	150.00	148.40	150.90	75.00	74.40	74.69
3	96.00	94.80	96.02	47.90	47.50	47.64
4	69.50	68.70	69.36	34.60	34.40	34.49
5	54.00	53.50	53.92	27.00	26.80	26.86
6	44.10	43.70	44.02	22.00	21.90	21.97
7	37.30	37.00	37.25	18.60	18.60	18.62
8	32.40	32.30	32.44	16.20	16.20	16.23
9	28.90	28.80	28.91	14.40	14.40	14.48
10	26.20	26.20	26.28	13.10	13.10	13.18
11	24.30	24.20	24.33	12.20	12.10	12.20
12	22.80	22.80	22.87	11.40	11.40	11.48
13	21.80	21.70	21.82	10.90	10.90	10.96
14	21.00	21.00	21.11	10.60	10.60	10.61
15	20.60	20.60	20.70	10.30	10.40	10.40
16	20.50	20.50	20.57	10.30	10.30	10.34

Table 4.2: Comparison of experimental and calculated characteristic resistances.

ρ	62.3 $\Omega \cdot$ cm			29.5 $\Omega \cdot$ cm		
z_c	15.0 $\Omega \cdot cm^2$			7.5 $\Omega \cdot cm^2$		
Spat. Freq.	Exp.	Cheng Sim.	FEM Sim.	Exp.	Cheng Sim.	FEM Sim.
1	70.70	70.70	70.66	33.90	33.90	33.55
2	33.80	33.70	33.62	16.30	16.10	16.00
3	21.70	21.50	21.53	10.10	10.30	10.27
4	15.70	15.60	15.64	7.60	7.50	7.47
5	12.20	12.20	12.21	5.90	5.80	5.85
6	10.00	10.00	10.01	4.80	4.80	4.80
7	8.50	8.50	8.50	4.10	4.10	4.08
8	7.40	7.40	7.43	3.60	3.60	3.57
9	6.60	6.60	6.64	3.20	3.20	3.19
10	6.00	6.00	6.05	2.90	2.90	2.91
11	5.60	5.60	5.61	2.70	2.70	2.70
12	5.20	5.20	5.28	2.60	2.50	2.54
13	5.00	5.00	5.04	2.50	2.40	2.43
14	4.90	4.90	4.88	2.40	2.30	2.35
15	4.80	4.80	4.79	2.30	2.30	2.31
16	4.70	4.70	4.76	2.30	2.30	2.30

Chapter 5

Anisotropic Regularisation

5.1 Introduction

In the preceding Chapters standard reconstruction methods for EIT have been introduced. The present chapter is concerned with the introduction of prior information into reconstruction. The aim in introducing prior knowledge is to enhance the inversion of conductivity profiles in terms of achieving more accurate quantitative estimates, in terms of partially preventing the smoothing effect introduced by the regularisation and in terms of reducing the sensitivity to measurement errors. Different methods can pursue all these aims, or part of them.

The issues of enhancing quantitative estimation and of preventing excessive smoothness are related: smooth solutions often overestimate the area of a contrasting inclusion and as a consequence the value of the contrast is often underestimated.

Our particular interest is to treat the problem of reconstructing conductivity profiles with steep conductivity variations, as often encountered in medical imaging. A first approach is to assume that the conductivity profile presents a "blocky" nature: the image consists mainly of a few uniform regions, separated possibly by steep variations that occur on a domain of small size. Methods that lead to ℓ_1–norm regularisation functionals are best suited for these applications, as discussed in detail in Chapters 6 and 7. A second possible approach, viable when specific structural information (known as *anatomical priors*) is known a priori, is to use ℓ_2–norm regularisation and to relax the smoothing effect on the basis of such information. By these means it is possible to reconstruct conductivity profiles with sharp features and to estimate their values with greater accuracy than with traditional ℓ_2 norm regularisation techniques.

The present chapter discusses the realisation of such regularisation functionals in the

form of anisotropic Gaussian filters. An anisotropic Gaussian filter is a regularisation functional that weights conductivity variations differently depending on their direction. The aim is to use such filters to depenalise conductivity variations that occur at locations and in directions that match the prior information, allowing sharper reconstruction of the expected features.

5.2 Incorporating Prior Information in EIT

All reconstruction methods in EIT incorporate some form of prior information, via Bayesian approaches or via Tikhonov regularisation. Often the prior information is of generic nature. As seen in Chapter 2, in conjunction with Tikhonov regularisation, differential operators are often used as penalty functionals. Nearly constant inversions are obtained for first order differential operators, and smooth inversions are obtained for second order differential operators. Clearly many situations of practical interest do not conform to these assumptions. The choice of such regularisation functionals is justified by the fact that they guarantee the stability of the inversion by filtering out components in the higher singular spectrum.

When prior information of the object properties is known, it is desirable to incorporate it into the reconstruction. The Bayesian approach offers a natural framework for doing so, by coding the prior information into the distribution $\Pi_M(m)$ (2.12). In practice the construction of the prior information distribution is a difficult task, especially if the model conductivities are assumed to be correlated. Common approaches tend to simplify somehow the construction of such a state of information. In the following sections some methods that have been used in the medical imaging context are reviewed. A clear classification of the methods as deterministic or probabilistic is not possible. All of them can be justified with considerations of probabilistic nature, and can be thought, to some extent, to build a sate of prior information.

5.3 Incorporating Prior Information Directly in the Model

Grouping Constraints

Grouping constraints methods (GCM) are used in situations when regions of the object are known to present the same conductivity. Image elements belonging to the same grouping region are constrained to have the same value. This situation is typical in medical imaging, where elements belonging to the same organ exhibit the same electrical properties (if the organ is uniform or assumed so). Anatomical information can be exploited to designate the

grouping regions appropriately. In [109] a FE mesh of the human thorax was created using known structural information. Five different areas were identified to represent different organs. The images reconstructed with this anatomically constrained mesh exhibited changes of lung resistivity related to the breathing cycle. A similar approach was followed in [110].

With the GCM the reconstruction is formulated as an unregularised LS problem, where the resistivity values of the few anatomical regions are fitted to the data. Regularisation is not needed as the constrained regions are usually sufficiently large to guarantee sufficient sensitivity (large singular values in the linearised forward operator). The inverse problem is therefore relatively well-posed.

In [111] mean values and covariance of the conductivity values of the constrained regions were assumed to be known, as well as the statistics of the measurement errors. An approach called *statistically constrained minimum mean square error estimator* (MiMSEE) was adopted, with improved results over the traditional LS approach.

The introduction of grouping constraints and of prior statistics reduces the ill-posedness of the inverse problems and achieves better quantitative results over traditional methods [109][110][111]. The major disadvantages in grouping image elements on the basis of known anatomy are that the shape of the organs is not allowed to change during their physiological activity and, secondly, that the reconstruction is completely biased to produce results conforming to the prior information.

Basis Constraints

Vauhkonen [112] proposed the basis constraints method (BCM) in order to overcome partially the limitations of the GCM method. The BCM works by estimating the conductivity distribution as a linear combination of specially constructed basis functions

$$\boldsymbol{\sigma} = \sum_{i=1}^{S} \mu_i \mathbf{w}_i \tag{5.1}$$

where μ_i are appropriate weights and $\mathbf{w}_1, \ldots, \mathbf{w}_S$ are orthonormal basis functions. These basis functions are built in such a way to be the principal eigenvectors of a representative ensemble of possible conductivity distributions. Vauhkonen built an ensemble of 81 possible thoracic conductivity distributions, using anatomical information from MRI scans, and conductivity values from standard medical tables. The ensemble distributions arose from the combination of nine different states of the heart and lungs. Each state represented, respectively, a different stage during the cardiac and breathing cycle, and was

characterised by a particular value of resistivity and shape of the organ. In such a way the basis functions that he used span both changes in conductivity values and in shape that occur during the physiological activity. Vauhkonen was able to show that changes in the reconstructed conductivity values were closely related to air and blood volume changes occurring during inspiration and breath holding of a subject.

The BCM method overcomes the problem associated with shape change during a physiological cycle, which was not described by the GCM. The BCM method suffers however, as the GCM does, from a "hard" bias toward a possible set of reconstructed images.

5.4 Prior Information via Tikhonov Regularisation

A major problem of methods that act directly on model parameters is that they enforce "hard" priors, in the sense that the inverse solution is constrained to belong necessarily to a subspace of the model space $\mathcal{M}$. If the imaged object does not conform to the prior information misleading results are obtained. Recent methods for incorporating prior information [83] [84] [85] [35] do so via Tikhonov regularisation. Recalling (3.4) the EIT inverse solution is formulated as

$$\boldsymbol{\sigma}_{rec} = \operatorname{argmin} \|h(\boldsymbol{\sigma}) - \mathbf{d}\|^2 + \alpha F(\boldsymbol{\sigma}) \tag{5.2}$$

The role of the regularisation functional is to penalise solutions that according to some prior knowledge are unlikely. The solution is "pushed" towards $\mathcal{N}(F)$ by the penalising effect of the regularisation functional. The approach is to build F is such a way that $\mathcal{N}(F)$ spans a particular subspace of $\mathcal{M}$ in which the solution is believed to lie according to the prior information. The enforcement of such prior information is "soft" as the conformity to the prior knowledge can be controlled via the parameter α and via the structure of $F(\boldsymbol{\sigma})$.

As mentioned, the regularisation functional in (5.2) assumes usually the form $F(\boldsymbol{\sigma}) = \|L\boldsymbol{\sigma}\|^2$. Methods for incorporating prior information into regularisation therefore build appropriate matrices L according to the priors.

Subspace Regularisation Method

The Subspace Regularisation Method [112] [83] (SSRM) was introduced by Vauhkonen to overcome the "hard" prior bias of BCM. Suppose that some basis vectors $\mathbf{w}_1, \ldots, \mathbf{w}_S$ have been built, for example, by computing the principal vectors of a representative ensemble as for BCM. These vectors will span the subspace of the model parameters $S_w \in \mathcal{M}$ in

which the solution is expected to lie. Vauhkonen suggests calculating a regularisation matrix whose null space is S_w in the following way

$$L = I - WW^T \tag{5.3}$$

where W is the matrix holding the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_S$ as columns and I is the identity matrix. In [112] [83] it was shown that SSRM reconstructions are more robust in the presence of noise than reconstructions using traditional regularisation matrices L. Secondly, it was shown that the SSRM method is capable of reconstructing conductivity profiles that do not conform to the prior information, in other words that the prior information is "soft".

Anisotropic regularisation

Anisotropic regularisation is a technique that can be used to incorporate prior information in situations where steep variations in conductivity are expected, but the localisation of them is only roughly known. Typical examples of this situation arise in medical imaging, where the shape of each organ is know from the anatomy, but it will vary to some extent from patient to patient. The non perfect knowledge of an organ's shape can arise also from physiological activity: the shape of the organ will change during time as it will contract or expand. The anisotropic regularisation works by formulating the reconstruction as a traditional Tikhonov regularisation and by using a regularisation matrix L that penalises conductivity variations according to their direction. The smoothness constraints are relaxed along the the direction of the expected changes, allowing faster transitions in this direction while preserving the necessary smoothness tangentially. Kaipio [84] has shown that when localisation of the structures is not exactly known such relaxation can be performed on a whole region of the image, maintaining a sufficient regularisation effect but allowing rapid variations in the anticipated directions. In his work Kaipio proposes a way of constructing anisotropic regularisation functionals for a piece-wise linear 2D FEM forward model. As the conductivity is linear on each element, the partial derivatives of the conductivity are expressed as functions of the nodal values of the variable. The reconstruction then penalizes the conductivity variations differently, according to the local direction of the expected changes. Kaipio showed, through numerical simulations, that by using anisotropic regularisation techniques it is possible to reconstruct with good accuracy the high impedance step presented by the skull with respect to the surrounding matter.

In the next sections anisotropic regularisation is discussed in more detail and a method for construction Gaussian anisotropic regularisation filters for a piece-wise constant 2D

FEM forward model is presented. The method is introduced by firstly considering the isotropic Gaussian regularisation and its relation to statistical inversion, and then by introducing the anisotropic Gaussian regularisation on the basis of statistical considerations.

5.5 Isotropic Gaussian Regularisation

As introduced in Chapter 2 the ℓ_2-norm Tikhonov regularised inversion is intimately related to Gaussian hypothesis on the measurement errors and prior information distributions. Recalling (2.32), the MAP inverse solution with such assumptions is

$$\boldsymbol{\sigma}_{MAP} = \text{argmin}\,[(h(\boldsymbol{\sigma}) - \mathbf{d}))^T C_{\boldsymbol{\nu}}^{-1}(h(\boldsymbol{\sigma}) - \mathbf{d}) + \\ + (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)^T C_{\boldsymbol{\sigma}}^{-1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)] \tag{5.4}$$

where $C_{\boldsymbol{\nu}}$ and $C_{\boldsymbol{\sigma}}$ are the covariances of the measurement noise and of the conductivity. Formally, the Tikhonov solution of the inverse problem (5.2) is identical to the MAP approach when $C_{\boldsymbol{\nu}}^{-1} = I$, $L^{\mathrm{T}}L \propto C_{\boldsymbol{\sigma}}^{-1}$ and $\boldsymbol{\sigma}_0$ is a prior estimate of $\boldsymbol{\sigma}$. As already mentioned the choice of L is in such a way connected to the statistical information on $\boldsymbol{\sigma}$.

The properties of $C_{\boldsymbol{\sigma}}$ reflect the prior knowledge of the system under measurement. The diagonal elements of the matrix represent the variance of each element in the image, the off-diagonal elements of the matrix are a function of the correlation r between elements of the image $C_{\boldsymbol{\sigma}_{ij}} = r\sqrt{C_{\boldsymbol{\sigma}_{ii}} C_{\boldsymbol{\sigma}_{jj}}}$. It is therefore possible to envisage the possibility of constructing $C_{\boldsymbol{\sigma}}^{-1}$ on the basis of the structural knowledge of the system under measurement, and of using a regularisation matrix L such that $L^T L \propto C_{\boldsymbol{\sigma}}^{-1}$.

The latter approach, $L^T L \propto C_{\boldsymbol{\sigma}}^{-1}$, has been considered by Adler *et al.* [88], even though the authors finally assumed just the implicit information that a limited resolution is achieved by EIT. They therefore used a covariance matrix that allowed some correlation between neighbouring image elements. The authors observed that the inversion of such covariance matrix was ill-posed and thus numerically unstable. Noting that the covariance matrix has the structure of a low-pass filter, the solution they proposed is to use a Gaussian high-pass filter to represent its inverse. For the 2D case, the frequency response of the filter with a spatial frequency of ω_o is

$$G(\tau, \upsilon) = 1 - \exp(-\omega_o(\tau^2 + \upsilon^2)) \tag{5.5}$$

with the following convolution kernel

$$g(x,y) = \delta(x,y) - \frac{\pi}{\omega_o^2}\exp(-\frac{\pi^2}{\omega_o^2}(x^2+y^2)) \tag{5.6}$$

A regularisation matrix L that approximates the filter can be found by expressing the filtered value $\hat{\sigma}$ of the continuous conductivity σ at the mid-point coordinates x_i, y_i of the *i-th* element of the image

$$\hat{\sigma}(x_i,y_i) = g(x,y) * \sigma(x,y)|_{x=x_i,y=y_i} \tag{5.7}$$

and assigning the value at the mid-point to the discrete conductivity

$$\hat{\boldsymbol{\sigma}}_i = \int g(x_i - x, y_i - y)\sigma(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{5.8}$$

The integration can be carried out on the single elements E_j of a mesh, after bringing the piece-wise constant conductivity (such is the representation used by the forward model) out of the integration

$$\hat{\boldsymbol{\sigma}}_i = \sum_j \boldsymbol{\sigma}_j \int_{E_j} g(x_i - x, y_i - y)\,\mathrm{d}x\,\mathrm{d}y \tag{5.9}$$

the filtered conductivity can be expressed as $\hat{\boldsymbol{\sigma}} = L\,\boldsymbol{\sigma}$ with the following definition for L

$$L_{ij} = \int_{E_j} g(x_i - x, y_i - y)\,\mathrm{d}x\,\mathrm{d}y \tag{5.10}$$

5.6 Anisotropic Gaussian Smoothing

The statistical interpretation of the matrix L sets the basis for incorporation of prior structural information into regularisation. In their study, Adler *et. al.*, however did not assume this information and used an isotropic Gaussian filter in order to mimic $C_{\boldsymbol{\sigma}}^{-1}$. Nevertheless their proposed method inspired the use of an anisotropic Gaussian kernel that we adopt in this study for exploiting the anatomical priors, which are, in the medical imaging context, the equivalent of the structural information.

We now consider the problem for which we propose a method by examining the situation presented in Figure 5.1. Assume that a body Ω has an inclusion Ω_{inc} that presents a different conductivity from the surrounding body. The shape of the inclusion is not precisely known, but it is bounded by β_2 and β_1. The boundary $\partial\Omega_{\text{inc}}$ of the object is as-

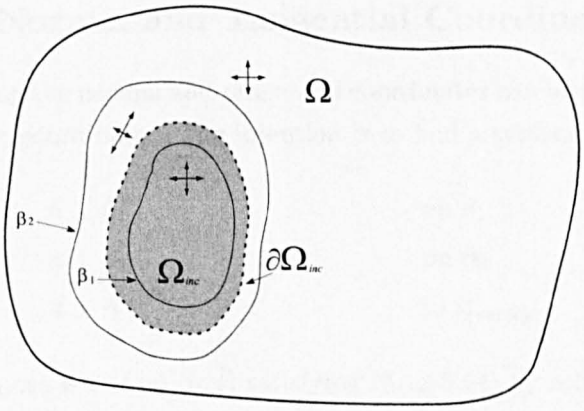


Figure 5.1: An object Ω has an inclusion Ω_{inc} with a different conductivity (gray area). The contour of the inclusion is bounded by β_1 and β_2 and is expected to follow approximately the bounding profiles. In order to exploit the structural information, the smoothing filter weights (represented by crosses) should be anisotropic in the region bounded by the two β curves, allowing for a faster variation of the conductivity in the direction of the expected changes.

sumed to follow approximately the bounding curves. An organ expanding and contracting during its physiological activity could be an example of the depicted situation.

The region enclosed between the two bounding curves β, which we label Ω_{change}, is where the expected conductivity discontinuity will occur. The objective is to relax the smoothing constraints in the region Ω_{change} along the direction normal to the line of changes, that has yet to be defined. The tangential smoothing can be maintained, expecting the changes to be orthogonal to that direction.

We therefore propose to use the anisotropic Gaussian kernel that is obtained by transforming (5.6)

$$g(n,t) = \delta(n,t) - \frac{\pi}{\omega_n \omega_t} \exp(-\pi^2(\frac{n^2}{\omega_n^2} + \frac{t^2}{\omega_t^2})) \tag{5.11}$$

where n and t are the tangential and normal directions of the expected changes in the conductivity. Such a kernel would separate the control of the smoothing along n and t by varying the parameters ω_n and ω_t. From a statistical point of view this is equivalent to assuming that the image elements are less correlated in the direction of the expected changes.

Theoretically the use of such a filter is straightforward; in practice the problem is to find a way of calculating the system of coordinates (n,t) given the geometry of the domains. The normal and tangential directions need to be defined somehow in the region Ω_{change}, in order to make the use of the Gaussian anisotropic kernel possible.

5.7 Finding Normal and Tangential Coordinates

The problem of finding the normal and tangential coordinates can be solved naturally with a system of harmonic coordinates. The intention is to find a system of coordinates where

$$\hat{n} \perp \beta_1 \qquad \text{on } \beta_1 \tag{5.12}$$

$$\hat{n} \perp \beta_2 \qquad \text{on } \beta_2 \tag{5.13}$$

$$\hat{t} \perp \hat{n} \qquad \text{in } \Omega_{\text{change}} \tag{5.14}$$

The solution we propose is to find $(\hat{n}, \hat{t})$ satisfying (5.12-5.14) by solving a PDE.

Solving $\nabla^2 n = 0$ over Ω_{change} with the following Dirichlet boundary conditions

$$\begin{aligned} n &= 1 \qquad \text{on } \beta_1 \\ n &= 0 \qquad \text{on } \beta_2 \end{aligned} \tag{5.15}$$

gives a solution n defined over all the domain Ω_{change} that can be regarded as one of the coordinates of the system (n, t). Points with $n = 0$ will lie on β_2, points with $n = 1$ will lie on β_1. The vector $\hat{n} \equiv \nabla n$ is orthogonal to β_1 for $n = 1$ and orthogonal to β_2 for $n = 0$. Therefore $\hat{n}$ satisfies (5.12) and (5.13). Now assume a crack in the domain Ω_{change} as illustrated in Figure 5.2, and that $\nabla^2 t = 0$ is solved over Ω_{change} with Dirichlet boundary conditions

$$\begin{aligned} t &= 0 \qquad \text{on one side of the crack.} \\ t &= 1 \qquad \text{on the other side of the crack.} \end{aligned} \tag{5.16}$$

and Neumann conditions $(\partial t / \partial \hat{n} = 0)$ on β_1 and β_2. The solutions n and t will form a system of harmonic coordinates. The level sets of t are distributed radially on the domain Ω_{change}, giving a vector $\hat{t} \equiv \nabla t$ that is orthogonal to $\hat{n}$ on it, satisfying (5.14). The coordinate change $(x, y) \rightarrow (n, t)$ maps the cracked domain Ω_{change} to the rectangle $[0\ 1] \times [0\ 1]$ in $\mathbb{R}^2$. The coordinates (n, t) can therefore be used to carrying out the integration of the kernel (5.11) producing an anisotropic weighting matrix L.

5.8 Comments on Calculation of L

The idea of solving a PDE for the calculation of (n, t) was based on the opportunity of relying mostly on the forward solver for the task. In electrical impedance tomography the

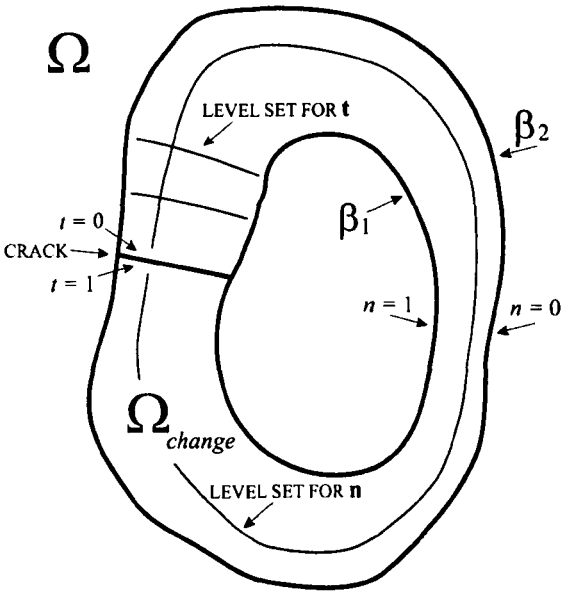


Figure 5.2: Calculation of the normal and tangential coordinates solving the Laplace equation over the domain Ω_{change} with opportune boundary conditions. The solutions n and t form a system of harmonic coordinates (n, t) that maps the cracked domain Ω_{change} to a rectangle.

forward algorithm solves

$$\nabla \cdot \sigma \nabla u = 0 \tag{5.17}$$

where u is the electric potential. It is sufficient to set σ to a constant value in (5.17) order to enable the forward solver to solve the Laplace equation. The calculation of (n, t) requires to solve the Laplace equation twice, with two different sets of boundary conditions. In the following the details of the numerical procedure are briefly explained.

5.8.1 Calculation of n

The calculation of n requires the solution of the Laplace equation over Ω_{change} with Dirichlet boundary conditions on β_1 and β_2. In order to achieve this, the numerical procedure uses the following steps:

- The FEM system matrix for the whole mesh of the domain Ω is assembled, setting the conductivity of each element to a constant value (eg. 1). The calculation of the FEM matrix is achieved making use of the forward solver. As the forward solver normally implements boundary conditions that simulate the presence of electrodes,

slight modifications are needed. Referring to Chapter 4, with the notation introduced for (4.32), the system matrix Y is calculated as $Y = Y'$, with Y' of Equation (4.34a)

- The system matrix Y is modified in order to apply Dirichlet boundary conditions: for each node belonging to β_1 or β_2 the corresponding row and column of Y are set to 0, the diagonal elements of Y corresponding to those nodes are set to 1.

- The right hand side vector, that with the notation of Chapter 4 we label $\mathbf{c}$, is created with all zero elements except for those elements corresponding to the nodes on β_1 , which are set to 1 (Dirichlet condition $n = 1$).

- The linear system $Y\mathbf{n} = \mathbf{c}$ is solved, and the unknowns vector $\mathbf{n}$, restricted to the domain Ω_{change}, represents the value of n on such domain.

5.8.2 Calculation of t

The calculation of t is slightly more complicated than that of n as Dirichlet boundary conditions need to be applied on the crack of the domain Ω_{change} and Neumann boundary conditions need to be applied to the boundaries β_1 and β_2. In order to achieve this, the domain Ω_{change} needs to be "cut out" from the whole domain Ω, in such a way that the condition $\partial t/\partial \hat{n} = 0$ is naturally applied on β_1 and β_2. Then the Dirichlet conditions can be applied to the crack. The numerical procedure uses the following steps:

- In order to "cut out" the subdomain Ω_{change} from Ω, the submesh of elements belonging to Ω_{change} need to be identified, and a reduced system matrix calculated just for the elements of the submesh. In order to identify a list of the elements belonging to Ω_{change} a numerical trick is used, based on the following consideration:

  The previously calculated solution $\mathbf{n}$ is constant on the inside of β_1 and on the outside of β_2, the nodal values are in fact identically 1 in the inside of β_1 and identically 0 on the outside of β_2. Elements belonging to Ω_{change} can be identified therefore by the fact that $\nabla n \neq 0$. This is achieved numerically by searching the vector $G_x\mathbf{n} + G_y\mathbf{n}$ for elements that are greater than a small positive constant (eg. 1×10^{-6}), where G_x and G_y are the gradient operators defined in (4.36) and already available to the forward solver. The output of this procedure is therefore the list of mesh elements belonging to Ω_{change}.

- A submesh for the domain Ω_{change} is built from the list of elements identified with the previous procedure. The mesh represents the domain Ω_{change}, "cut out" from Ω.

- The crack is practiced to the submesh.

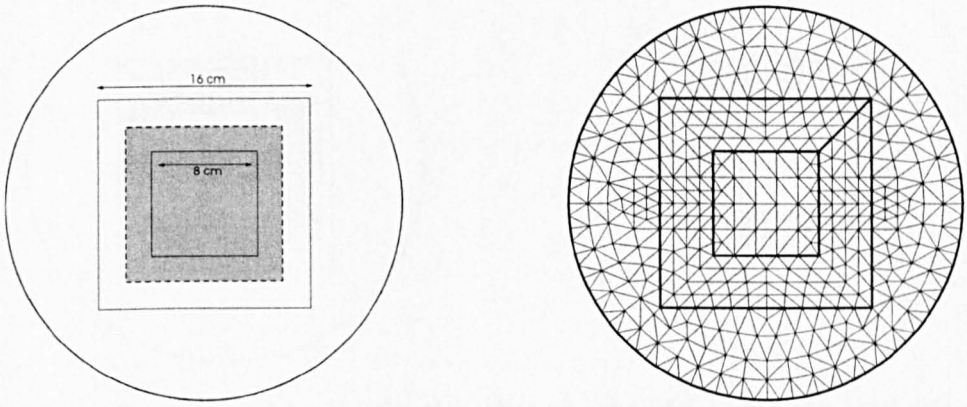
- The routines of the forward solver are used to calculate the reduced system matrix Y_{change} for the submesh. This task is achieved in the same manner as in 5.8.1 (i.e. $Y_{\text{change}} = Y'_{\text{change}}$, where Y'_{change} is computed with an arbitrary constant value of conductivity).
- The system matrix Y_{change} is modified in order to apply Dirichlet boundary conditions: for each node belonging crack the corresponding row and column of Y_{change} are set to 0, the diagonal elements of Y_{change} corresponding to those nodes are set to 1.
- The right hand side vector, that with the notation of Chapter 4 we label $\mathbf{c}$, is created with all zero elements except for those elements corresponding to the nodes on one side of the crack, which are set to 1 (Dirichlet condition $t = 1$).
- The linear system $Y\mathbf{t} = \mathbf{c}$ is solved, and the unknowns vector $\mathbf{t}$, represents the value of t on Ω_{change}.

An algorithm for the calculation of (n, t) has been developed in the MATLAB environment and integrated with the routines currently in use for the forward solution. The algorithm presupposes that a FEM mesh matching the external and internal boundaries of the object to be imaged has been produced. The user describes the structural information by selecting on the screen the nodes on β_1, β_2 and on the crack. The algorithm computes (n, t) with the aforementioned procedures. Given the two spatial frequencies ω_n and ω_t, it then calculates the anisotropic regularisation matrix L.

For each finite element E_i belonging to the domain Ω_{change} the corresponding row of L_i is calculated by integrating (5.11), while for the rest of the elements the isotropic kernel (5.6) is used. As it happens that the convolution kernels do not decay completely to zero on the integration domain, the sum over rows of L is not equal to zero. Uniform conductivity regions would therefore raise the response of the filter. In order to prevent this, the diagonal elements of L have been set to be equal to the negative sum of the remaining elements on the same row. In such a way a uniform conductivity distribution is orthogonal to L. The alteration is equivalent to assuming a flat prolongation of the image outside the domain of integration.

5.9 Simulations

In this section we compare reconstructions using prior information, in the form of anisotropic smoothing, and traditional reconstructions using Gaussian isotropic smoothing. A numer-



a) b)

Figure 5.3: Setup of the simulated experiment: round object with a square inclusion, definition of the domains and dimensions in cm. Generated mesh, boundaries and crack are shown in thick lines.

ical experiment involving a square inclusion embedded in a round object was set up. The experiment has no physiological meaning but allows easy visual comparison of the reconstructed profiles with respect to the original conductivity distribution, secondly the square object was chosen because of its sharp corners, that are a challenge for traditional regularised solutions. The numerical simulations allowed us to compare anisotropic and isotropic regularisation techniques, and to find some interesting results on the effect of the prior information on the regularisation.

5.9.1 Setup of the Experiment

The numerical experiment was set up as illustrated in Figure 5.3a. An outer round object with a diameter of 30 cm is expected to contain a square shaped inclusion with a different conductivity. The dimensions of the inclusion can vary, the side of the anomaly can range from 8 cm to 16 cm.

A mesh of 798 triangular elements, shown in Figure 5.3b, was used for the inverse computations. The mesh matches the internal boundaries β_1 and β_2 in order to allow the calculation of (n, t) with the PDE method. The forward solver uses a finer mesh for calculating the electric potential, attaining higher accuracy in the forward solutions. The finer mesh was obtained from an adaptive refinement of the first mesh, resulting in 6346 elements. The mesh was used also for the generation of the test conductivity profiles of the simulations. The inclusions of the test profiles were generated not coincide with the discretisation of the coarse mesh, as this would be not representative of a real situation.

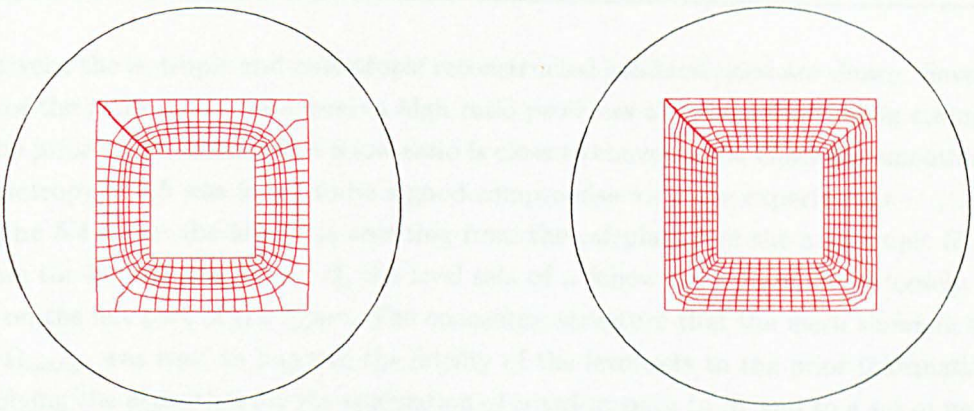


Figure 5.4: Level sets of (n,t). On the left part, the plot shows the level sets of (n,t) calculated by means of unsign the list of nodes on β_1 and β_2. In this case the level sets of n are rounded, far from the two boundaries they follow loosely the anticipated shape. On the right part, the plot shows the level sets of (n,t) calculated using an additional set of nodes, selected halfway between β_1 and β_2 (as the mesh is concentric in the region Ω_{change}). The algorithm for the calculation of n has been applied twice: once between β_1 and the halfway set of nodes, and then between the halfway set of nodes and β_2. The arrangement allows the level sets of n to more closely describe the prior information.

The disposition of the electrodes resembles the hybrid arrangement of the OXBACT III tomograph, with 32 electrodes for voltage measurements, and 32 for current injection. The simulations were performed applying the first 31 trigonometric patterns to the object, as this resembles the typical OXBACT III experiment, and enhances the sensitivity of the measurements to conductivity changes in the inner part of the object.

5.9.2 Reconstructions – Correct Prior Information

A first test conductivity profile matching the expected structure was used to compare the reconstructions with and without prior information. In the simulations, the conductivity of the surrounding circular object was set to 1 $\Omega \cdot$cm, and an 11 cm square inclusion, with conductivity 0.7 $\Omega \cdot$cm, was generated, as shown in Figure 5.5a.

For the reconstructions, trigonometric current patterns were used and the resulting measurements were calculated with a forward solver implementing the complete electrode model. A Gaussian white noise with zero mean and a standard deviation of 0.1% of the voltage range was added to the measurements to simulate instrument noise [25]. The reconstruction is solved iteratively, starting from a homogeneous conductivity that best fits the data. The first four steps of the algorithm were performed, as no further improvement in the reconstructed profile was noticed prolonging the iterations. In Figure 5.5b and 5.5c

respectively, the isotropic and anisotropic reconstructed conductivities are shown. Several values of the ratio ω_n/ω_t were used; a high ratio produces an image conforming strongly with the prior information, while a low ratio is close to conventional Gaussian smoothing. An anisotropy of 2.5 was found to be a good compromise for these experiments.

Figure 5.4 shows the level sets resulting from the calculation of the anisotropic filter. Far from the boundaries β_1 and β_2 the level sets of n follow the square shape loosely, as shown on the left part of the figure. The concentric structure that the mesh shows in the region Ω_{change} was used to improve the fidelity of the level sets to the prior information. By applying the algorithm for the calculation of n twice: once to β_1 and to a set of mesh nodes halfway between β_1 and β_2, and the then to the latter set of nodes and to β_2, the level sets can be forced to describe the prior information more closely, as shown in the right part of the figure. This arrangement was used in the simulations, as this improves the description of the corners of the square inclusion. The overall effect however depends both from the shape of the level sets, and from the discrete nature of the filter, which is calculated by integration on the mesh.

As expected, in the reconstructions, the isotropic solution smooths the lateral discontinuities of the square inclusion and rounds off the corners. The anisotropic solution, incorporating the prior information, estimates the square shape of the detected object more accurately. In Figure 5.6a and 5.6b cross sections of the true and reconstructed conductivities are shown for an easier quantitative comparison. Figure 5.6a illustrates a cross section along the x axis. The anisotropic solution follows the sharp transitions and settles closer to the correct value in the centre of the object than the isotropic one. The effect is more evident in Figure 5.6b. where the cross section is cut on the 45^o diagonal, crossing the corners of the inclusion. The discontinuity of the corners is even more difficult to describe for isotropic smoothing, resulting in a larger difference in the two reconstructions.

5.9.3 Reconstructions – Incorrect Prior Information

The risk in using prior information in the reconstruction process is to bias the solution toward the assumed distribution, and to miss inclusions that do not respect the prior assumptions.

In this section are presented reconstructions comparing the two methods in the case where the priors are incorrect. Figure 5.7a shows the test conductivity to be reconstructed. The inclusion is a rectangular object with the lateral edges orthogonal to the direction assumed for the conductivity changes in Ω_{change}. The object therefore violates the prior information assumed by the regularisation. The conductivity values for this test are again $1\ \Omega \cdot \text{cm}$ for the embedding object and $0.7\ \Omega \cdot \text{cm}$ for the inclusion. In Figure 5.7b and

5.7c, the isotropic and anisotropic reconstructions are shown. Both reconstructions were performed with the same parameters ($\alpha, \omega_n, \omega_t$) as the ones used for the results of Figure 5.5, and the same level of noise. The anisotropic solution detects the anomaly and locates it correctly. Some fake responses are however triggered in the Ω_{change} region, resulting in a poorer performance when compared to the isotropic reconstruction.

5.10 GSVD Analysis

As mentioned in Chapter 2, the linearised inverse problem

$$\delta\boldsymbol{\sigma}_{\text{rec}} = \text{argmin } \|J\,\delta\boldsymbol{\sigma} - \delta\mathbf{d}\|^2 + \alpha\|L\delta\boldsymbol{\sigma}\|^2 \tag{5.18}$$

is said to be in standard form when the regularisation matrix L is equal to the identity. For such a case, or when L is a diagonal matrix, the inverse solution can be studied by means of SVD analysis. For the general case, when L is a full matrix, there are two possible methods that allow similar characterisation of the inverse solution. The first method is to use transformations to standard form, and the second is to use GSVD analysis [43].

For the simple case where L is square and invertible the transformation to standard form is straightforward: by choosing $\overline{J} = JL^{-1}$ and $\overline{\boldsymbol{\sigma}} = L\boldsymbol{\sigma}$ and $\overline{\mathbf{d}} = \mathbf{d}$, (5.18) is transformed to

$$\delta\overline{\boldsymbol{\sigma}}_{\text{rec}} = \text{argmin } \|\overline{J}\,\delta\overline{\boldsymbol{\sigma}} - \delta\overline{\mathbf{d}}\|^2 + \alpha\|\delta\overline{\boldsymbol{\sigma}}\|^2 \tag{5.19}$$

which can be analysed via SVD, and the results transformed back as $\boldsymbol{\sigma} = L^{-1}\overline{\boldsymbol{\sigma}}$. The matrix L can be understood to define simply a different norm for weighting the conductivity vector $\boldsymbol{\sigma}$, and the transformation to standard form to be a method of transforming a problem regularised in such norm to a problem regularised in the traditional ℓ_2 norm.

The GSVD decomposition allows study of the problem (5.18) in the generic case where L is rectangular (although transformation methods exist also for this case [43]). The generalised decomposition can be applied therefore to the Gaussian filter. The matrices $J \in \mathbb{R}^{m\times n}$ and $L \in \mathbb{R}^{p\times n}$ (where p is the number of "regularisation constraints", $p = n$ in the Gaussian case) are then decomposed as

$$\begin{pmatrix} J \\ L \end{pmatrix} = \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & I_{n-p} \\ M & 0 \end{pmatrix} X^{-1} \tag{5.20}$$

where $r = \text{rank}(L)$, $U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{p\times p}$ and $X \in \mathbb{R}^{n\times n}$. The matrices U and V have

orthonormal columns and X is non singular, Λ and M are diagonal matrices $\in \mathbb{R}^{p\times p}$, $\Lambda = \text{diag}(\lambda_1, ..., \lambda_p)$ and $M = \text{diag}(\mu_1, ..., \mu_p)$ with

$$1 \geq \lambda_1 \geq ... \geq \lambda_r > \lambda_{r+1} = \cdots = 0\,, \quad 0 \leq \mu_1 \leq ... \leq \mu_p \leq 1 \tag{5.21}$$

and

$$\lambda_i^2 + \mu_i^2 = 1\,, \quad i = 1, ..., p \tag{5.22}$$

The generalized singular values are defined as $\gamma_i = \lambda_i/\mu_i$ for $i = 1 \ldots r$, they appear in non-increasing order:

$$\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_{r-1} \geq \gamma_r > 0 \tag{5.23}$$

Using this decomposition the regularised solutions to the linearised problem (5.18) can be expressed as

$$\delta\boldsymbol{\sigma}_{\text{rec}} = \sum_{i=1}^{r} \frac{\gamma_i^2}{\gamma_i^2 + \alpha^2} \frac{(\mathbf{u}_i^T \delta\mathbf{d})}{\lambda_i} \mathbf{x}_i + \sum_{i=r+1}^{n} (\mathbf{u}_i^T \delta\mathbf{d})\, \mathbf{x}_i \tag{5.24}$$

Similar conclusions can be drawn from (5.24) to those from the SVD analysis. The generalized singular values λ_i show the ill-posedness of the problem by rapidly decaying to zero for increasing i. The term $\gamma_i^2/(\gamma_i^2 + \alpha^2)$ should therefore tend to zero with sufficient rapidity to prevent the first term at the right hand side of (5.24) from diverging. The matrix pair (J, L) is therefore understood to regularise the inversion in a similar fashion to (J, I): by damping the content of the singular vectors (SV) for which the corresponding singular values are too small.

When anisotropic filtering is adopted, we have found interesting changes in the structure that the singular vectors assume. Typically the singular vectors associated with bigger singular values are smooth; they have components only in the lower part of the spectrum of spatial frequencies and do not present discontinuities. The GSVD analysis of the pair (J, L) revealed that with the anisotropic Gaussian filter the structure of the singular vectors changes.

Figure 5.8a illustrates the second SV for the isotropic Gaussian filter, which is a vertical gradient as usually happens for traditional choices of L. The corresponding vector for the anisotropic Gaussian filter is shown if Figure 5.8b. The plot of the singular vector presents null region corresponding to the area delimited by β_2. The singular vector appears to span only changes in the outer region of the image, to carry information decorrelated from the area of prior information. Figure 5.9 shows the $14^{th}, 42^{nd}, 108^{th}$ and 131^{st} singular vectors for the anisotropic case. The structural prior information appears to having been embedded in the SVs, modifying the smooth structure that one would expect otherwise. Some

of the vectors $\mathbf{x}_i$ decouple the information relative to the inclusion from the background by spanning only particular regions of the image. The non-smooth properties of the lower SVs allow the reconstruction to describe inclusions matching the prior information, as explained in more detail via spectral analysis in the next section.

5.11 Spectral Analysis

As with the SVD decomposition, it is possible to express a given conductivity distribution $\boldsymbol{\sigma}$ as a linear combination of the singular vectors $\mathbf{x}_i$

$$\boldsymbol{\sigma} = \sum_{i=1}^{n} w_i \mathbf{x}_i \tag{5.25}$$

The w_i are said to be the spectral coefficients of $\boldsymbol{\sigma}$. The actual calculation of the coefficients differs from the standard SVD case in that $\mathbf{x}_i$ are not orthogonal but just linearly independent. The first r spectral coefficients w_i can be calculated as

$$w_i = \langle \boldsymbol{\sigma}, \mathbf{x}_i \rangle_L \quad i = 1, ..., r \tag{5.26}$$

since the vectors x_i are $L^{\mathrm{T}}L$ orthogonal for $i = 1, ..., r$, and $\langle \mathbf{x}_i, \mathbf{x}_j \rangle_L = 0$ for $i = 1 \ldots r$, $j = r + 1 \ldots n$.

The remaining coefficients can be calculated as

$$w_i = \langle \boldsymbol{\sigma}, \mathbf{x}_i \rangle_J \quad i = r + 1, ..., n \tag{5.27}$$

since the last $n - r$ columns of X are $J^{\mathrm{T}}J$ orthogonal and $\langle \mathbf{x}_i, \mathbf{x}_j \rangle_J = 0$ for $i = 1 \ldots r$, $j = r + 1 \ldots n$.

Traditionally, the lower singular vectors are smooth and the higher ones are oscillatory. Smooth conductivity distributions will therefore have components in the lower part of the spectrum and vice versa. In this sense the decomposition (5.25) is similar to a Fourier analysis of $\boldsymbol{\sigma}$.

As introduced earlier, the regularisation dampens the contribution of the higher singular vectors in the image. For a given conductivity distribution $\boldsymbol{\sigma}$, the higher the spectral content, the more the reconstructed image will suffer from the dampening. Thus, an image with sharp changes, having significant components in the higher part of the spectrum, will be heavily smoothed. However, when anisotropic filtering is used, the lower singular vectors are able to span certain sharp transitions, shifting the corresponding compo-

	iso x_i	aniso x_i
σ_1	273	11
σ_2	38	64

Table 5.1: Number of singular vectors required to describe a test conductivity within 0.04 error.

nents downwards in the spectrum. These conductivity patterns will therefore survive the smoothing effect of the regularisation.

The spectral shifting property was verified by expanding the test conductivity of Figure 5.5a in its spectral coefficients, using the singular vectors derived both from the isotropic and anisotropic filters. The singular vectors span the conductivity space of the mesh used for the inverse calculations. The test conductivity, defined on a finer mesh, was therefore projected onto this mesh. The distribution of the spectral energy was compared by truncating the two expansions at an index k for which

$$\frac{\| \sum_{i=1}^{k} w_i \mathbf{x}_i - \boldsymbol{\sigma} \|}{\|\boldsymbol{\sigma}\|} < 0.04 \tag{5.28}$$

the value of 0.04 was chosen by visual inspection as a threshold that would guarantee a good approximation to the original image. The results are shown in the first row of Table 5.1 labelled $\boldsymbol{\sigma}_1$. The lower 273 singular vectors are needed to describe the conductivity $\boldsymbol{\sigma}_1$ within 4% error, if isotropic filtering is used. The same conductivity is spanned within the same error by the lower 11 singular vectors in the anisotropic case. The non-smooth nature of the SVs deriving from the anisotropic filter can thus describe a sharp conductivity profile with a smaller spectral content, resulting in less smoothing from the regularisation as shown by the reconstructions. The conductivity $\boldsymbol{\sigma}_1$ matches the prior information used for setting up the matrix L and for calculating the corresponding singular vectors. For this reason the sharp changes in $\boldsymbol{\sigma}_1$ are spanned by the lower SVs.

The same experiment was repeated for the test conductivity of Figure 5.7a; the results are reported in the second row of Table 1, labelled $\boldsymbol{\sigma}_2$. In this case the situation differs, 64 singular vectors are needed to span the image within the 4% error in the anisotropic case, versus 38 for the isotropic regularisation. The conductivity $\boldsymbol{\sigma}_2$ does not match the prior information, it isn't spanned by the lower singular vectors, resulting in a slightly worse spectral distribution.

5.12 Discussion

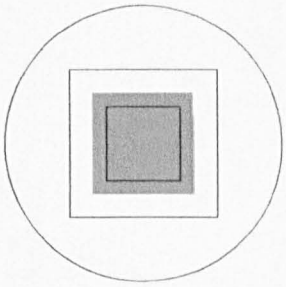
In the present chapter a novel method for using structural prior information in EIT image reconstruction has been proposed. The method is based on considerations of statistic nature as an extension of the traditional Gaussian regularisation techniques. The approach is to enhance the reconstructions of sharp conductivity variations by incorporating the structural information into the Tikhonov regularisation functional. The effectiveness of the approach has been positively compared to the use of isotropic functionals by means of simulations. The simulations show that the sharpness and quantitative estimation of the conductivity are enhanced when the experiment matches the prior information. On the other hand, with a careful selection of the regularisation parameters, the algorithm was able to detect a contrast that violated the prior assumptions. In the simulations the regularisation parameter was chosen by visual inspection. If needed, common methods for the choice of the parameter, as the L–curve method, the Discrepancy Principle or the Generalised Cross Validation could be used. Alternatively, from the strict Bayesian point of view, the regularisation parameter should be chosen according to the prior statistics, by imposing that $\alpha L^T L = C_{\boldsymbol{\sigma}}^{-1}$.

In our view the present work contributes in three different respects to the field of anisotropic regularisation: the use of anisotropic Gaussian filters is proposed, a method for the calculation of the regularisation matrix L is proposed, the regularised problem is analysed via GSVD decomposition and some interesting aspects are highlighted. The use anisotropic Gaussian smoothing is a contribution of general validity, and a natural extension of the isotropic equivalent. The use of such regularisation filters is believed to be useful for forward models that assume the conductivity to be piece-wise linear. As with such a representation of the conductivity the partial derivatives of are not readily available, the Gaussian filter achieves directional properties by weighting opportunely neighbouring elements, allowing the use of structural prior information. Regarding the method we propose for the calculation of L, we believe it to be best suited for objects with a relatively simple contour, and of reasonably large dimensions compared to the mesh size. Both the calculation of the normal and tangential coordinates and the effect of the filter are independent from discretisation if the mesh elements are small compared to the details of the prior structures. Regarding the analysis of the effect of anisotropic filtering on the regularisation, some interesting modifications to the structure of the generalised singular vectors were found. The capability of the reconstruction method of describing steep conductivity variations is explained in terms of such analysis.

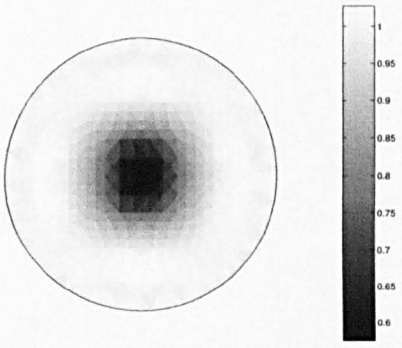
In this work only the 2D dimensional case was studied. As a concluding remark, the

extension to 3D is briefly discussed here. The extension of the Gaussian filter to the three dimensional case is straightforward, the integration of the kernel (5.11) should be carried out on the volume elements. The calculation of a system of coordinates for such integration (one normal and two tangentials to the surfaces) is instead complicated by the fact that a system of harmonic coordinates cannot be found with the method we propose for 2D. In the 3D setting the curves β_1 and β_2 would be two surfaces. One possibility could be to find the normal coordinate by solving again the Laplace equation with Dirichlet boundary conditions on the two surfaces. The other two coordinates, both orthogonal to the first one, could be found by choosing an arbitrary system of coordinates on, for example, β_2 and mapping it onto β_1 through Ω_{change} following the streamlines of the field used for the calculation of t. On all the region Ω_{change} and on β_1 and β_2 those two new coordinates would be orthogonal to the tangential coordinate as required.

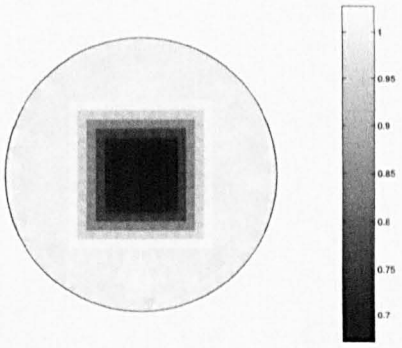
The choice of the initial system of coordinates on β_2 is not straightforward. For simple topologies polar coordinates could be projected from a sphere onto β_2.



a)

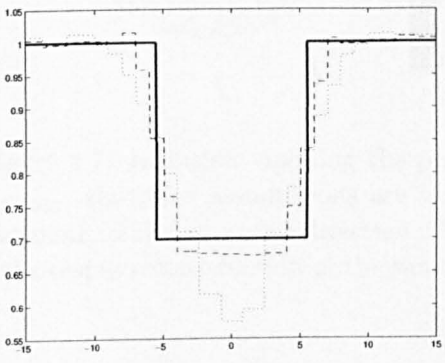


b)

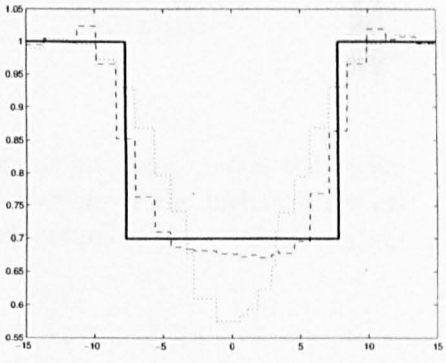


c)

Figure 5.5: Comparisons of isotropic and anisotropic priors. a) Test conductivity profile. b) Reconstruction with Gaussian isotropic smoothing of the 10 cm square inclusion. c) Reconstruction using a Gaussian anisotropic filter



a)



b)

Figure 5.6: Cross sections of the reconstructions. The thick line represents the true conductivity, the dashed line the anisotropic reconstruction and the dotted line the isotropic one. a) Cross section along the horizontal axis. b) Cross section along the 45^o diagonal.

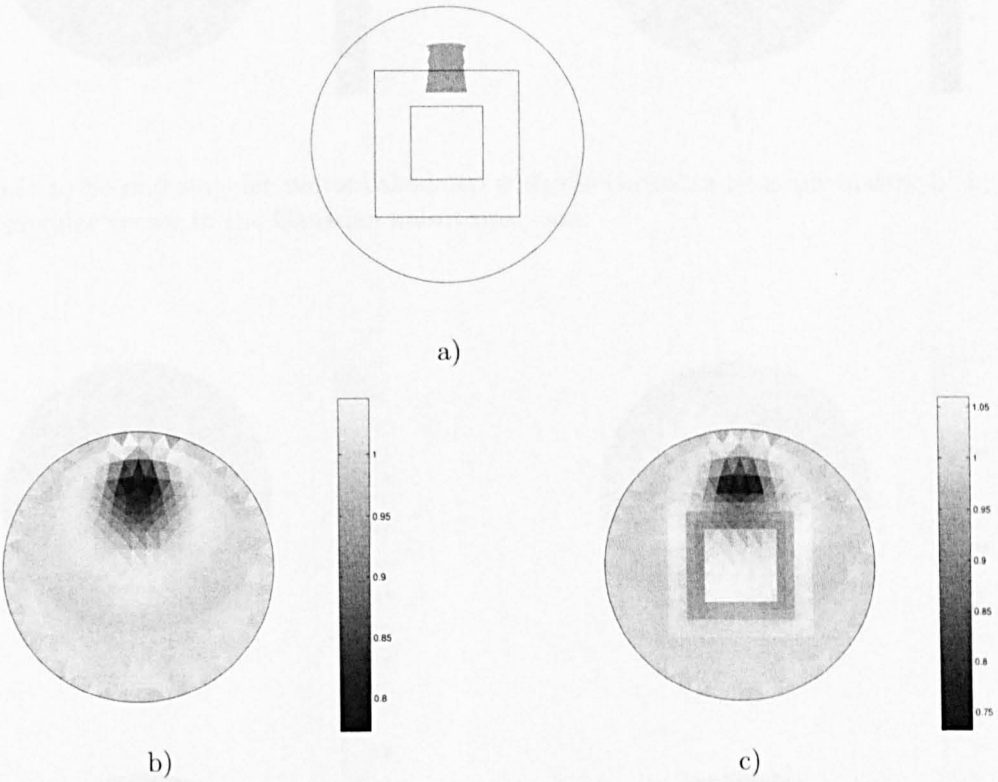
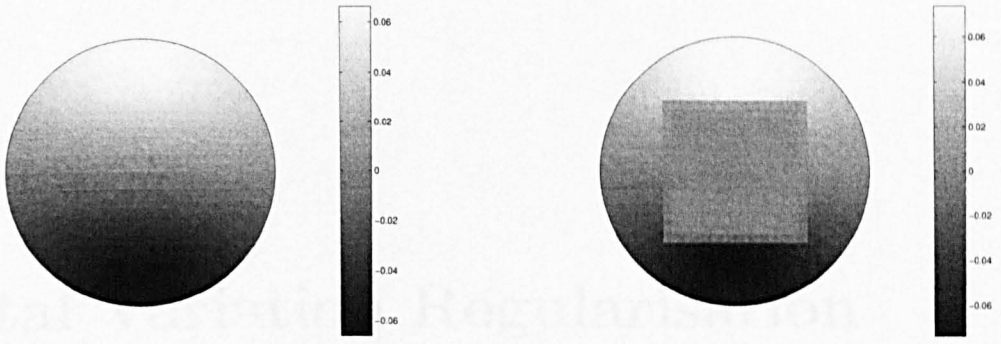
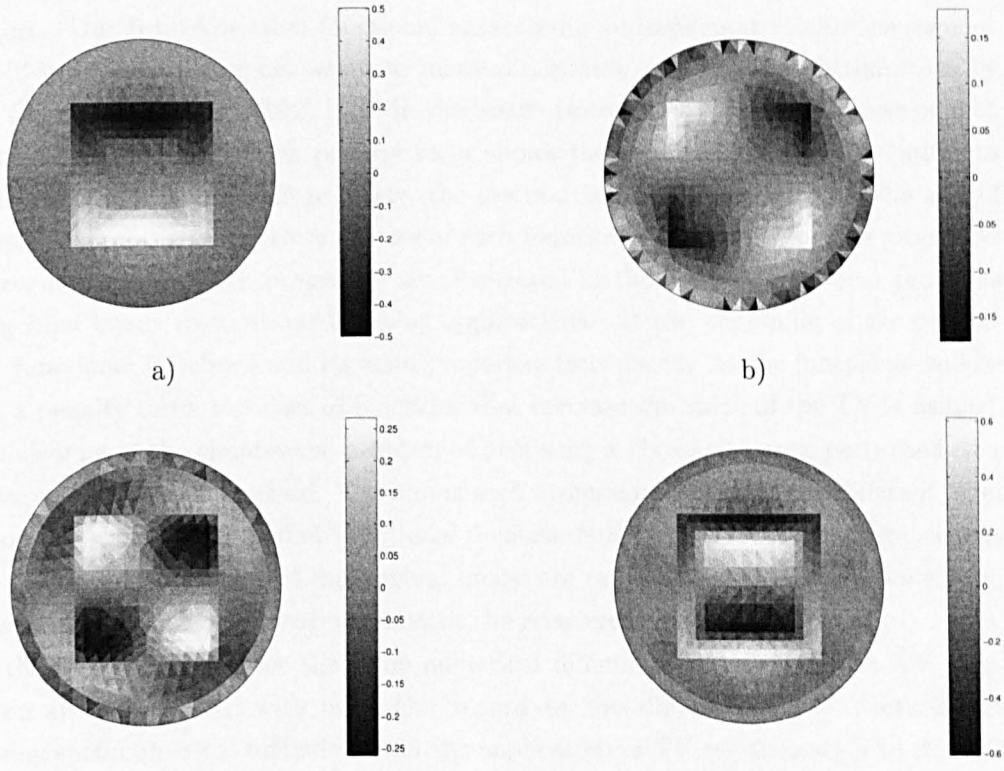


Figure 5.7: Inclusion violating the priors. a) A rectangular inclusion crosses the region Ω_{change}, the prior assumptions are violated, the lateral borders of the inclusion are orthogonal to the expected direction. b) Isotropic reconstruction of the conductivity. c) Anisotropic reconstruction of the same conductivity.



a) b)

Figure 5.8: a) Second singular vector calculated with the Gaussian isotropic matrix L. b) Second singular vector in the Gaussian anisotropic case.



a) b)

c) d)

Figure 5.9: Generalized singular vectors (14^{th}, 42^{nd}, 108^{th}, 131^{st}) calculated with the Gaussian anisotropic matrix L.

Chapter 6

Total Variation Regularisation

6.1 Introduction

The present chapter is dedicated to introduction of the Total Variation (TV) regularisation techniques. The Total Variation functional is assuming an important role in the regularisation of inverse problems belonging to many disciplines, after its first introduction by Rudin, Osher and Fatemi (1992) [113] in the image restoration context. The use of such a functional as a regularisation penalty term allows the reconstruction of discontinuous profiles. As this is a desirable property, the method is gaining popularity. The aim of the present chapter is to motivate the use of such functional by illustrating the properties of TV regularisation. Such properties are illustrated in the context of inverse problems deriving from image restoration/denoising applications. At the beginning of the chapter the TV functional is defined and its main properties introduced. As the functional will be used as a penalty term, the class of functions that increase the value of the TV is defined. In the following of the chapter the problem of denoising a "blocky" image perturbed by a high-frequency noise is discussed. The aim of such discussion is to show the different regularisation properties of TV and of traditional ℓ_2-norm functionals. Numerical experiments show how the sharp features of the original image are retained by the TV regularisation, and how the ℓ_2-norm regularisation smooths the recovered image.

At the end of the chapter the some numerical difficulties associated with TV regularisation are highlighted, with particular regard to non-differentiability. Methods for addressing such numerical difficulties and the application of TV regularisation to the EIT inverse problem are discussed in the next chapter.

6.2 Total Variation Functional

Total variation measures the total amplitude of the oscillations of a function. For a differentiable function $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$, with $d = 1, 2, 3$ the total variation is [114]

$$TV(f) = \int_\Omega |\nabla f| \tag{6.1}$$

The definition can be extended to non-differentiable functions [115] as:

$$TV(f) = \sup_{\vec{v} \in \mathcal{V}} \int_\Omega f \operatorname{div} \vec{v} \tag{6.2}$$

where $\mathcal{V}$ is the space of $C^1(\Omega; \mathbb{R}^d)$ functions that vanish on $\partial\Omega$ and $\|\vec{v}\|_\Omega \leq 1$.

As the TV functional measures the variations of a function over its domain, it can be understood to be effective at reducing oscillations in the inverted profile, if used as a penalty term. The same properties apply however to ℓ_2 regularisation functionals. The important difference is that the class of functions with bounded total variation also includes discontinuous functions, which makes the TV particularly attractive for the regularisation of non-smooth profiles. The following one-dimensional example illustrates the advantage of using the TV against a quadratic functional in non-smooth contexts

Let $F = \{f : [0, 1] \rightarrow \mathbb{R}, | f(0) = a, f(1) = b\}$, we have

- $\min_{f \in F} \int_0^1 |f'(x)| \mathrm{d}x$ is achieved by any monotonic function, including discontinuous ones.
- $\min_{f \in F} \int_0^1 (f'(x))^2 \mathrm{d}x$ is achieved only by the straight line connecting the points $(0, a)$ $(1, b)$.

Figure 6.1 shows three possible functions f_1, f_2, f_3 in F. All of them have the same total variation, including f_3 which is discontinuous. Only f_2 however minimises the H^1 semi-norm

$$|f|_{H_1} = \left(\int_0^1 \left(\frac{\partial f}{\partial x} \right)^2 \mathrm{d}x \right)^{1/2} \tag{6.3}$$

which is called semi-norm, as $|f|_{H_1} = 0$ does not imply that $f = 0$. The quadratic functional, if used as penalty, would therefore bias the inversion toward the linear solution and the function f_3 would not be admitted in the solution set as its H^1 semi-norm is infinite.

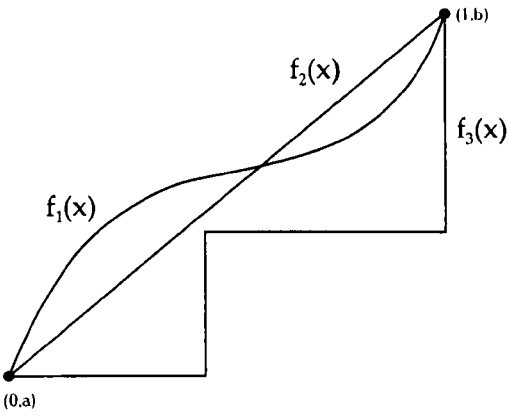


Figure 6.1: Three possible functions: $f_1, f_2, f_3 \in F$. All of them have the same TV, but only f_2 minimises the H^1 semi-norm.

6.3 Perturbations That Increase the Total Variation

In order to motivate the use of the TV functional as a regularisation penalty term and to understand the benefits associated with it, it is helpful to characterise the perturbations that will increase the TV of a function. We report in this section some results published by Dobson and Santosa [116]. As most of the results arise from inverse problems in image processing applications, the model parameter to be recovered, the image intensity, will be indicated by m, and it is defined on $\Omega = (0\ 1) \times (0\ 1)$ in $\mathbb{R}^2$.

Lets consider the following noise removal problem: an image m_o is observed in the presence of a perturbation δm_o, the original image m_o is to be restored from the observations $d = m_o + \delta m_o$. We consider restoring m_o by minimising the total variation of m, subject to $\|d - m_o\|_2 \leq \eta$, or similarly adopting the Tikhonov unconstrained formulation. To establish that the process could recover m_o, it is necessary (but not sufficient) to show that the perturbation δm_o will increase the total variation of the image m_o. We study therefore the class of functions that will result in such an increase.

6.3.1 Nearly Piecewise Constant Images

Assume that $m_o \in C^1(\Omega)$, and that it is a "nearly piecewise constant" function. That is: the set upon which m_o is not constant:

$$B = \{x \in \Omega : \nabla m_o(x) \neq 0\} \tag{6.4}$$

has some "small" extent with respect to Ω.

Lets represent the signal perturbation δm_o with it's Fourier transform, assuming periodic boundary conditions

$$\delta m_o = \sum_{k \in Z^2} \check{\delta} m_{o_k} e^{2\pi i x \cdot k} \tag{6.5}$$

where $Z = \{0, \pm 1, \pm 2, \ldots\}$ and

$$\check{\delta} m_{o_k} = \int_\Omega \delta m_o(x) e^{-2\pi i x \cdot k} \mathrm{d}x \tag{6.6}$$

then

$$\nabla \delta m_o = -2\pi i \sum_{k \in Z^2} k \check{\delta} m_{o_k} e^{2\pi i x \cdot k} \tag{6.7}$$

and provided $\nabla \delta m_o \in L^1(\Omega)$

$$\int_B |\nabla \delta m_o| \leq 2\pi \int_B \sum_{k \in Z^2} |k\, \check{\delta} m_{o_k}| = 2\pi |B|\, \|k\, \check{\delta} m_{o_k}\|_1 \tag{6.8}$$

where $|B|$ is the measure of B. It follows that for any $k \in Z^2$

$$2\pi |k\, \check{\delta} u_{o_k}| \leq \int_{\Omega - B} |\nabla \delta m_o| + \int_B |\nabla \delta m_o| \leq \int_{\Omega - B} |\nabla \delta m_o| + 2\pi |B|\, \|k\, \check{\delta} m_{o_k}\|_1 \tag{6.9}$$

taking the supremum of the left-hand side

$$2\pi (\|k\, \check{\delta} m_{o_k}\|_\infty - |B|\, \|k\, \check{\delta} m_{o_k}\|_1) \leq \int_{\Omega - B} |\nabla (m_o + \delta m_o)| \tag{6.10}$$

Combining (6.8) and (6.10)

$$\begin{aligned} \int_\Omega |\nabla (m_o + \delta m_o)| &\geq \int_B |\nabla m_o| - \int_B |\nabla \delta m_o| + \int_{\Omega - B} |\nabla (m_o + \delta m_o)| \\ &\geq \int_\Omega |\nabla m_o| + 2\pi (\|k\, \check{\delta} m_{o_k}\|_\infty - 2|B|\, \|k\, \check{\delta} m_{o_k}\|_1) \end{aligned} \tag{6.11}$$

Thus the total variation of $m_o + \delta m_o$ is greater than the total variation of m_o if

$$2|B|\, \|k\, \check{\delta} m_{o_k}\|_1 < \|k\, \check{\delta} m_{o_k}\|_\infty \tag{6.12}$$

Two important conclusions can be drawn from the results

- Given $\|k\, \delta m_{o_k}\|_1$ is finite, m_o could be recovered if $|B|$ is sufficiently small.
- Given $|B|$ is finite, m_o could be recovered as long as the spread of the perturbation's

spectrum is sufficiently small.

So for example, a "nearly piecewise constant" function polluted by a few neighbouring pure frequencies would be easily recovered. On the other hand, a smoothly varying function polluted by a spatially localised perturbation (which implies a large spectral spread) may be impossible to recover.

6.3.2 Discrete Images

Results similar to the continuous case exist for discrete images. Assume that m is piecewise constant on a uniform square grid over Ω, with $N \times N$ pixels, and that $m_{z,h}$ denotes the value of m on the pixel (z,h). In this case $|\nabla m_o|$ is a measure supported on the lines of the grid. In order to calculate it we define the horizontal and vertical difference operators D_H and D_V as

$$\begin{aligned}(D_H\, m)_{z,h} &= m_{z+1,h} - m_{z,h} \\ (D_V\, m)_{z,h} &= m_{z,h+1} - m_{z,h}\end{aligned} \tag{6.13}$$

where $m_{z,N+1} = m_{z,1}$ and $m_{N+1,h} = m_{1,h}$, due to the periodicity of m. From definition (6.13) we see that

$$\int_\Omega |\nabla m| = \|D_H\, m\|_1 + \|D_V\, m\|_1 \tag{6.14}$$

Again assume that $m = m_o + \delta m_o$, and define the set $\mathcal{S}$ upon which m_o is not constant

$$\mathcal{S} = \{(z,h) : (D_H\, m_o)_{z,h} \neq 0 \text{ or } (D_V\, m_o)_{z,h} \neq 0\} \tag{6.15}$$

Lets denote with S the number of elements in $\mathcal{S}$, and express the perturbation δm_o as

$$\delta m_{o_{z,h}} = \frac{1}{N^2} \sum_{k,l} \check{\delta} m_{o_{k,l}} \exp\left(\frac{-2\pi i(z-1)(k-1)}{N}\right) \exp\left(\frac{-2\pi i(h-1)(l-1)}{N}\right) \tag{6.16}$$

Assuming that δm_o is band–limited, i. e. $(k,l) \in \mathcal{K}$, where $\mathcal{K}$ is some finite set with K elements, by application of the triangle inequality one finds that

$$\sum_{(z,h)\in\mathcal{S}} |D_H\,(m_o + \delta m_o)_{z,h}| \geq \sum_{(z,h)\in\mathcal{S}} |D_H\,(m_o)_{z,h}| - \frac{K\,S}{N^2}\, \|\check{\delta} m_o\|_\infty \tag{6.17}$$

With a similar calculation it can be shown that

$$\left(1 - \frac{K\,S}{N^2}\right) \|\check{\delta} m_o\|_\infty \leq \sum_{(z,h)\notin\mathcal{S}} |D_H\,(m_o + \delta m_o)_{z,h}| \tag{6.18}$$

Combining the previous two inequalities, it follows that

$$\|D_H\,(m_o + \delta m_o)\|_1 \geq \|D_H\,m_o\|_1 + \left(1 - \frac{2K\,S}{N^2}\right) \|\check{\delta} m_o\|_\infty \tag{6.19}$$

The same derivation yields an analogous inequality for the vertical difference operator. Adding the horizontal and vertical contributions, one finds that

$$\int_\Omega |\nabla m_o + \delta m_o| \geq \int_\Omega |\nabla m_o| + 2\left(1 - \frac{2K\,S}{N^2}\right) \|\check{\delta} m_o\|_\infty \tag{6.20}$$

Therefore $TV(m_o + \delta m_o) \geq TV(m_o)$ if

$$KS \leq \frac{N^2}{2} \tag{6.21}$$

In the discrete case no restrictions are imposed on the spread of spectrum of the perturbation, simply the number of spectral components of δm_o must be small compared to the proportion of edges over which m_o is not constant (S/N^2).

6.3.3 Application: perfect recovery from band–limited perturbations

In the preceding sections we have shown under which conditions the total variation of an image is increased by a perturbation. However it is not sufficient to show that the unperturbed image is the infimum in a set of images to prove that it can be recovered by minimisation of the TV functional, as the recovery process depends also on the effect of the perturbation on the observations. An image can be restored either with a noise constrained minimisation

$$\begin{aligned} m_{rec} &= \text{argmin} \int_\Omega |\nabla m| \\ s.t. \quad & \|m - d\| \leq \eta \end{aligned} \tag{6.22}$$

where η qualifies the error level on the observations, either with a Tikhonov approach

$$m_{rec} = \text{argmin}\ \|m - d\| + \alpha \int_\Omega |\nabla m| \tag{6.23}$$

where α is the Tikhonov factor. In general, the effect of the perturbation would alter d in such a way that $\|m_o - d\| \neq 0$, consequently the original unperturbed image m_o might not be the infimum of the objective function, and perfect recovery would not be possible.

An interesting case to study is the one where δm_o belongs to the class of perturbations that increase the total variation for the given image, and δm_o does not alter the observations. In this case the original image can be recovered in it's original form, since it is the extremum of the objective function. A practical example of this case is the one of band-limited perturbations. Suppose that the perturbation δm_o belongs to a space of functions Q with a limited spectral support, the reconstruction problem can be formulated as

$$\begin{aligned} &\min_m \int_\Omega |\nabla m| \\ &\text{s.t.} \quad (m - d) \in Q \end{aligned} \tag{6.24}$$

making the constraints insensitive to the effect of the perturbations. Lets assume that Q is a class of perturbations supported in the discrete frequencies $|k| > R$ with R a known constant. The set K being finite, the total variation of any image m_o with $S \leq \frac{N^2}{2R}$ will be increased by a perturbation belonging to Q, guaranteeing that m_o can be recovered exactly by (6.24). By building a convolution operator A, whose kernel is an ideal low pass filter with a cut off frequency $f_c \leq k$, it is possible to recover the unperturbed image, reformulating (6.24) as

$$m_{rec} = \operatorname{argmin} \|A(m - d)\| + \alpha \int_\Omega |\nabla m| \tag{6.25}$$

The Tikhonov regularised formulation (6.25) is equivalent to the constrained formulation (6.24) in that only the components of $(m - d)$ orthogonal to Q are considered.

6.3.4 Numerical Experiments

This section reports the results of numerical experiments that we conducted to support the general findings of Section 6.3, and to show an application of the findings of Section 6.3.3. A noise-free checkerboard image m_o, 32×32 pixels wide, was generated as shown on the left of Figure 6.2. The image is not constant over 180 edges. On the right of Figure 6.2 a perturbation δm_o is shown. The perturbation has support limited to four discrete frequencies in the higher part of the spectrum. With the introduced notation $N = 32$, $S = 180$, and $K = 4$, the perturbation therefore satisfies (6.21) and increases the total variation of m_o. The perturbed image $m_o + \delta m_o$ shown in Figure 6.3 has in fact a TV of

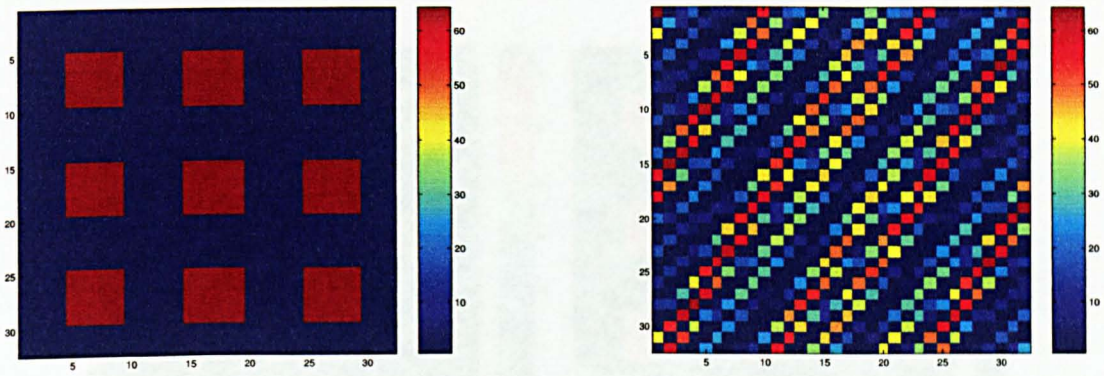


Figure 6.2: Two images, each 32×32 pixel wide, represented in an arbitrary color scale. On the left, an unperturbed checker board image m_o with $S = 180$ and $TV(m_o) = 5.6250$. On the right, a high frequency perturbation δm_o with $K = 4$.

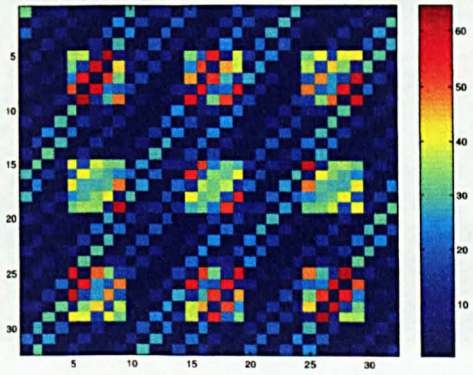


Figure 6.3: Pertubed image $m_o + \delta m_o$ in arbitrary colour scale, $TV(m_o + \delta m_o) = 51.4813$

.

51.4813, against 5.6250 for the unperturbed image.

In order to recover m_o with (6.25) we utilised an operator A that corresponds to a Gaussian low pass convolution kernel with a standard deviation of 3 pixels. The filter's response gives a good attenuation at the frequencies of the perturbation, being $\|A\,\delta m_o\|/\|\delta m_o\| = 0.031$, making the first term of (6.25) almost insensitive to δm_o. The PD-IPM algorithm (see Section 7.3) was used for the actual minimisation of (6.25). Figure 6.4 shows several steps of the restoration process ($1^{st}, 5^{ft}, 10^{th}, 15^{th}$). The original image can be restored almost perfectly, small differences remain due to the fact that A is not completely orthogonal to the perturbation. The restored image presents a total variation of 5.6256, close to the original image's TV of 5.6250.

For comparison with ℓ_2–norm regularisation methods, a Pseudo–TV (PSTV) restora-

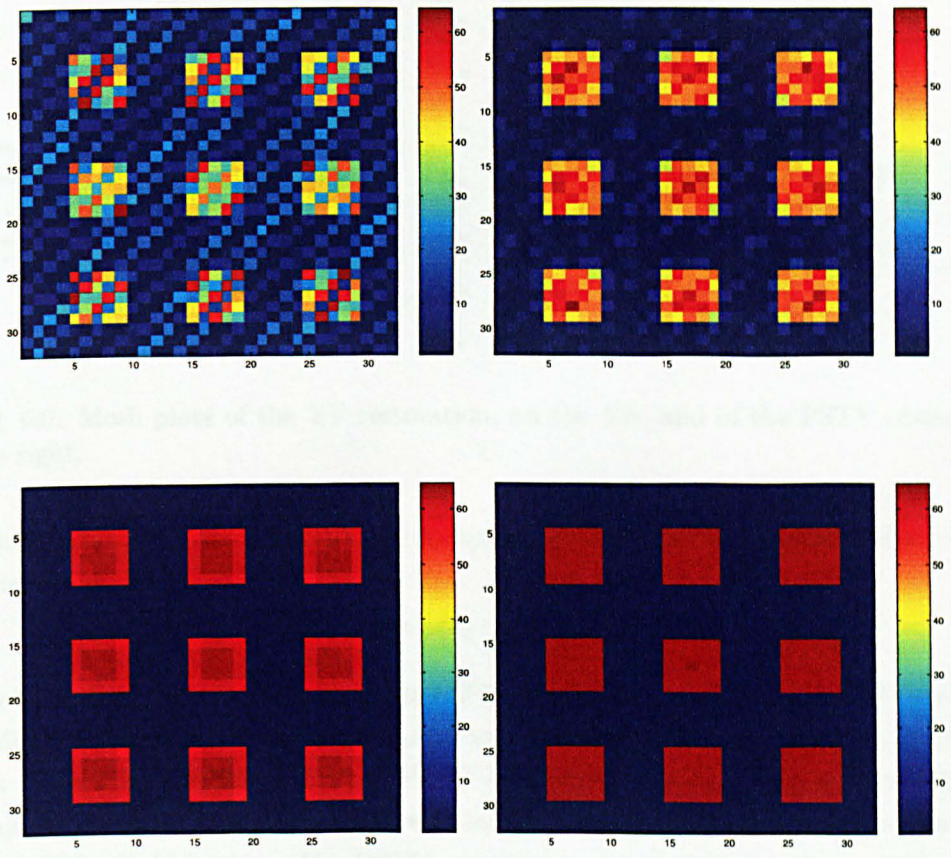


Figure 6.4: Several steps ($1^{st}, 5^{ft}, 10^{th}, 15^{th}$) of the TV restoration of the perturbed image.

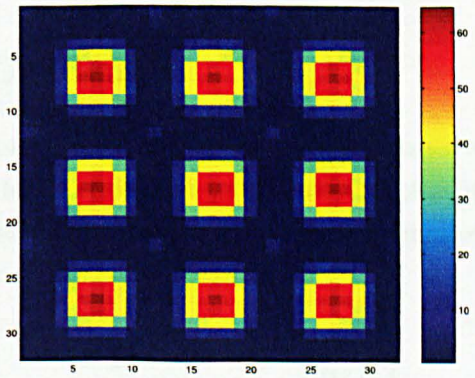


Figure 6.5: PSTV restoration.

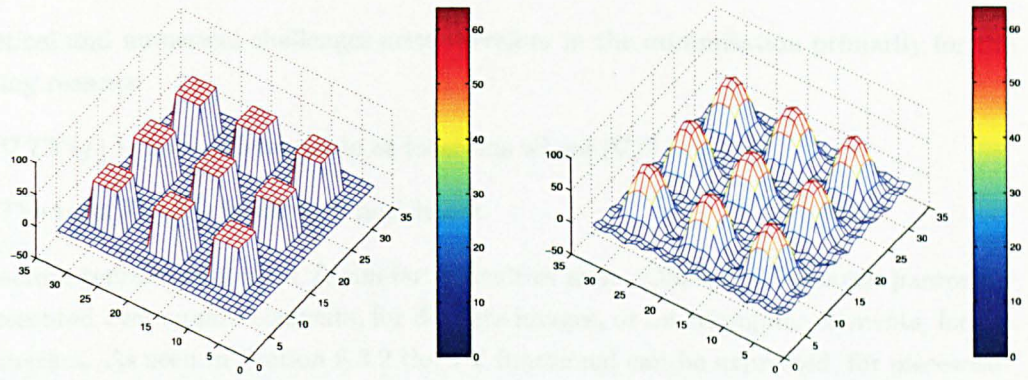


Figure 6.6: Mesh plots of the TV restoration, on the left, and of the PSTV restoration, on the right.

tion was performed. We call PSTV the ℓ_2 equivalent of the TV functional, defined for 2D discrete images as

$$\int_\Omega |\nabla m| = \|D_H\, m\|_2 + \|D_V\, m\|_2 \tag{6.26}$$

The results of the restoration are shown in Figure 6.5. For the TV and the PSTV restorations different values of the Tikhonov factor were used, as the penalty terms are in different norms. The Tikhonov factor for the PSTV experiment was chosen in such a way as to offer in the restored image a similar rejection of the high-frequency perturbation as offered by TV regularisation. The PSTV experiment shows the smoothing characteristic of ℓ_2 regularisation: the restoration process alters significantly the edgy characteristics of the original image. The effect is evident in the mesh plots of Figure 6.6 where the two restorations are compared.

6.4 Numerical Challenges

Total variation based regularisation methods necessarily involve the minimisation of the TV functional. In order to minimise the total variation of a function f, optimality conditions must be met. That is: partial derivatives of the functional should be nulled. The gradient of $TV(f)$ can be shown to be [117]

$$\nabla TV(f) = -\nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right) \tag{6.27}$$

theoretical and numerical challenges arise therefore in the minimisation primarily for the following reasons:

- $\nabla TV(f)$ is non–differentiable at locations where $|\nabla f| = 0$
- The term $\nabla \cdot \frac{\nabla f}{|\nabla f|}$ is strongly non–linear

For discrete representations of f, similar difficulties arise. Often the unknown parameter is represented over square elements, for discrete images, or on triangular elements, for 2D FEM meshes. As seen in Section 6.3.2 the TV functional can be expressed, for piecewise constant functions, as:

$$TV(\mathbf{f}) = \sum_{i=1}^{N} l_i \, |\mathbf{f}_{m(i)} - \mathbf{f}_{n(i)}| \tag{6.28}$$

where N is the number of edges over the domain of $\mathbf{f}$ and $m(i)$ and $n(i)$ are the indices of the elements adjacent the i–th edge and $\mathbf{f}_{m(i)}, \mathbf{f}_{n(i)}$ are the values assumed by $\mathbf{f}$ on those elements. A similar representation exists for piecewise–linear parameters, which involves sums of the absolute values of each element's nodal values. Whichever form the discrete parameter $\mathbf{f}$ assumes, $TV(\mathbf{f})$ can be expressed in matrix form as:

$$TV(\mathbf{f}) = \|L\,\mathbf{f}\|_1 \tag{6.29}$$

where L is an opportune sparse matrix. The discrete functional presents, like its continuous counterpart, points of non–differentiability for any value of $\mathbf{f}$ for which $|L_i\,\mathbf{f}| = 0$.

6.4.1 Discussion

The present chapter shows by theoretical considerations and by numerical experiments the different regularisation properties of the TV functional and of typical ℓ_2–norm regularisation functionals. In particular, the capability of TV regularisation of retaining sharp features in the sought parameters, is shown. Theoretical results indicate that it is possible to reconstruct blocky images from partial knowledge of their spectrum. In this sense the TV regularisation is attractive for use in EIT, as observations of the higher spectral components of the conductivity are poor. The use of TV regularisation poses computational difficulties. The aim of the next chapter is therefore to address the difficulties associated with the technique and to apply this type of regularisation to EIT inversion.

Chapter 7

Deterministic methods for TV Regularised EIT Inversion

The aim of this chapter is to overview the numerical methods that have been applied in TV regularised inverse problems, and to examine their suitability for application in EIT. Two different approaches have been already proposed for application in electrical imaging, the first by Dobson *et. al.* [86] and the second by Somersalo *et. al.* and Kolehmainen *et. al.* [41] [11]. The approach proposed by Dobson and Santosa is suitable for the linearised problem and suffers from poor numerical efficiency. Somersalo and Kolehmainen successfully applied MCMC methods (see Section 2.2) to solve the TV regularised inverse problem. The advantage in applying MCMC methods over deterministic methods is that they do not suffer from the numerical problems involved with non-differentiability of the TV functional. They do not require *ad hoc* techniques. Probabilistic methods, such as MCMC, offer central estimates and errors bars by sampling the posterior probability density of the sought parameters. The sampling process involves a substantial computational effort, often the inverse problem is linearised in order to speed up the sampling. Our aim is to find efficient methods for deterministic Tikhonov style regularisation, to offer a non-linear TV regularised inversion in a short time.

A second aspect, which adds importance to the study of efficient MAP (Tikhonov) methods, is that the linearisation in MCMC methods is usually performed after an initial MAP guess. Kolehmainen [11] reports calculating several iterations of a Newton method before starting the burn-in phase of his algorithm. A good initial deterministic TV inversion could therefore bring benefit to these approaches.

Examining the relevant literature, a variety of deterministic numerical methods have been used for the regularisation of image denoising and restoration problems with the TV

functional (a good review is offered by Vogel in [50]). The numerical efficiency and stability are the main issues to be addressed. Use of *ad–hoc* techniques is common, given the poor performance of traditional algorithms. Most of the deterministic methods draw from ongoing research in optimisation, as TV minimisation belongs to the important classes of problems known as "Minimisation of sum of norms" [118] [119] [120] and "Linear ℓ_1 problems" [121] [122].

The present chapter is concerned with the applicability of such techniques to the EIT inverse problem. The first part of the chapter analyses by theoretical considerations and by numerical experiments of the application of well known optimisation methods (Steepest Descent and Newton Methods) to the TV regularised EIT inverse problem. Such methods are shown to be inefficient and to present numerical difficulties. The second part of the chapter discusses the novel use of optimisation techniques called Primal Dual Interior–Point Methods in EIT. Such new techniques allow the formulation of a numerical framework for efficient and stable TV regularisation. The properties of the PD–IPM methods are analysed by theoretical considerations and by numerical experiments.

7.1 Approximations of the TV Functional

The TV functional is an element of a class F of regularisation functionals that can be represented as

$$F(f) = \int_{\Omega} \phi(|\nabla f|) \tag{7.1}$$

where $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ can assume different forms. The choice $\phi(t) = t^2$ yields the H^1 seminorm, and $\phi(t) = |t|$ the TV functional. A number of modifications to $\phi(t) = |t|$ have been proposed to deal with the non–differentiability of the absolute value for $t = 0$. One of the most recurring examples is [123] [124] [114] [125] [117]

$$\phi_\beta(t) = \sqrt{t^2 + \beta} \tag{7.2}$$

where β is a small positive parameter. A second example is the approximation used by Chambolle and Lions [126]

$$\phi_\epsilon(t) = \begin{cases} \frac{t^2}{2\epsilon} & t \leq \epsilon \\ t - \frac{\epsilon}{2} & \epsilon \leq t \leq \frac{1}{\epsilon} \\ \frac{\epsilon t^2}{2} \frac{1-\epsilon^2}{2\epsilon} & t \geq \frac{1}{\epsilon} \end{cases} \tag{7.3}$$

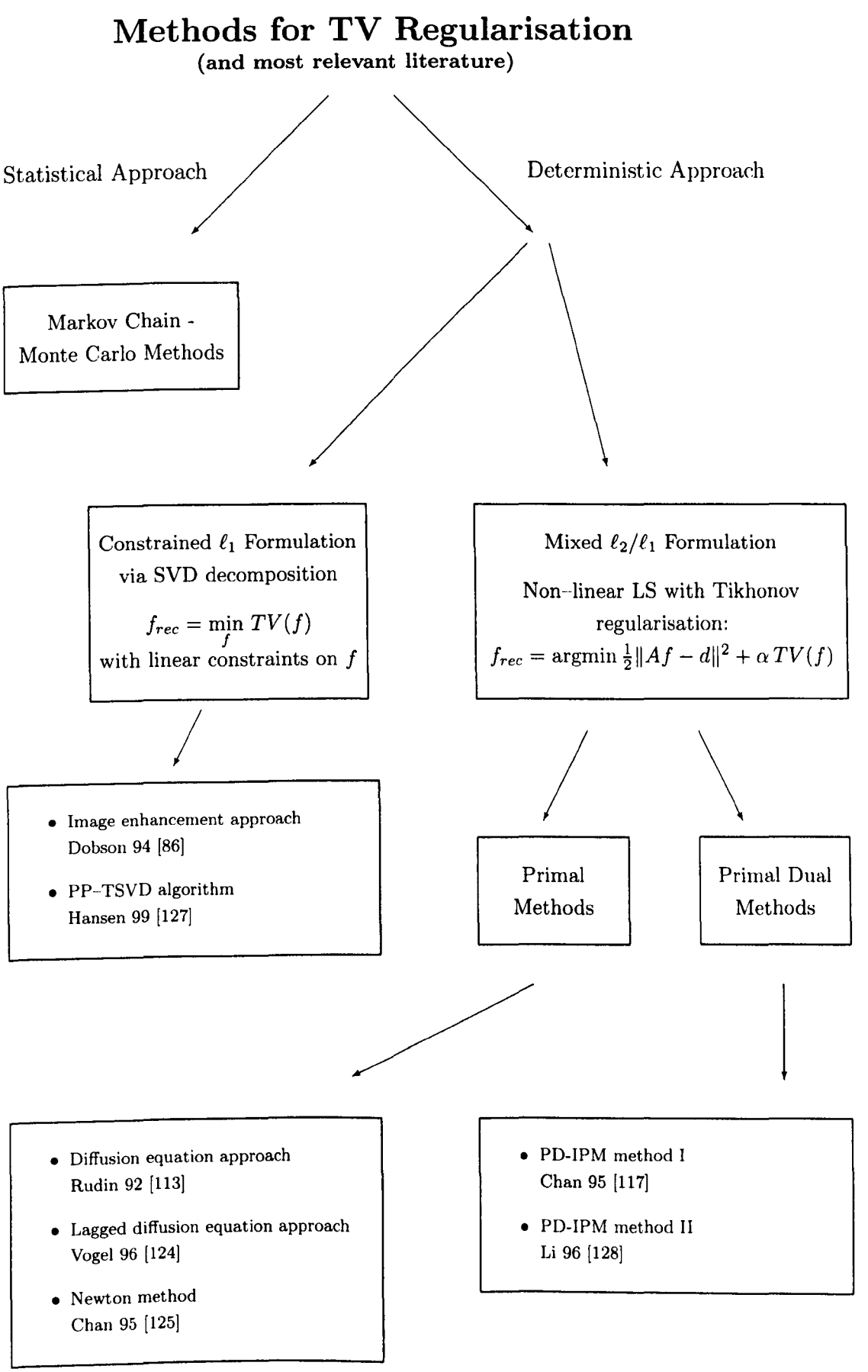
where ϵ is a small positive parameter. Dobson and Santosa adopt a very similar modification for small values of t in [86], while Kolehmainen [11] adopts a different modification

$$\phi(t) = \frac{1}{\gamma} \log(\cosh(\gamma t)) \tag{7.4}$$

where γ is a small positive parameter. The rationale behind all these modifications to the choice $\phi(t) = |t|$ is that the resulting functionals are differentiable and they tend to the TV functional for the parameters β, ϵ, γ tending to zero. In the following we will adopt $\phi_\beta(t) = \sqrt{t^2 + \beta}$ defining the TV_β regularisation functional

$$TV_\beta(f) = \int_\Omega \sqrt{|\nabla f|^2 + \beta} \tag{7.5}$$

We refer to Acar and Vogel [123] for an existence, uniqueness and stability analysis for TV_β regularised problems with $\beta > 0, \beta \to 0$.


Methods for TV Regularisation
(and most relevant literature)
Statistical Approach
Deterministic Approach
Markov Chain - Monte Carlo Methods
Constrained ℓ_1 Formulation via SVD decomposition
$f_{rec} = \min_f TV(f)$
with linear constraints on f
Mixed ℓ_2/ℓ_1 Formulation
Non–linear LS with Tikhonov regularisation:
$f_{rec} = \text{argmin } \frac{1}{2}\|Af - d\|^2 + \alpha TV(f)$
• Image enhancement approach Dobson 94 [86]
• PP–TSVD algorithm Hansen 99 [127]
Primal Methods
Primal Dual Methods
• Diffusion equation approach Rudin 92 [113]
• Lagged diffusion equation approach Vogel 96 [124]
• Newton method Chan 95 [125]
• PD-IPM method I Chan 95 [117]
• PD-IPM method II Li 96 [128]

7.2 Primal Methods for the Mixed ℓ_1/ℓ_2 Problem

Primal methods solve the following convex [123] optimisation problem

$$\begin{aligned} f_{rec} &= \operatorname{argmin} \psi(f) \\ \psi(f) &= \tfrac{1}{2}\|Af - d\|^2 + \alpha TV_\beta(f) \end{aligned} \tag{7.6}$$

where (7.6) is called the primal problem (P), as a dual formulation exists for which f_{rec} is a supremum of an optimisation problem called dual (see Paragraph 7.3). Primal methods were the first to be used in TV regularisation, as they are commonly used in optimisation, and widely known.

7.2.1 Artificial Time Evolution (or Steepest Descent Method)

Rudin Osher and Fatemi [113] have introduced the use of the TV functional for the regularisation of inverse problems in image restoration. Their approach, called artificial time evolution, is a time marching technique based on a diffusion equation. Originally the technique was applied to the noise-level constrained inversion. We will show its application to (7.6), as the two forms are equivalent [129].

The Euler Lagrange equation for $\psi(f)$, obtained by imposing the condition that the gradient $\psi'(f)$ should be zero at an optimal solution, is:

$$\psi'(f) = A^T(Af - d) - \alpha \nabla \cdot \left(\frac{1}{\sqrt{|\nabla f|^2 + \beta}} \nabla f \right) = 0 \tag{7.7}$$

or equivalently

$$A^T(Af - d) + \alpha \mathcal{L}_\beta(f) f = 0 \tag{7.8}$$

where $\mathcal{L}_\beta(f)$ is the diffusion operator that applied to v gives

$$\mathcal{L}_\beta(f) v = -\nabla \cdot \left(\frac{1}{\sqrt{|\nabla f|^2 + \beta}} \nabla v \right) \tag{7.9}$$

Rudin *et. al.* then assume that f is a function of time (as well as space) and to compute the steady state of the following non-linear diffusion equation for $f(x, y, t)$

$$\frac{\partial f}{\partial t} = A^T(Af - d) + \alpha \mathcal{L}_\beta(f) f \tag{7.10}$$

with $f(x, y, 0)$ given as initial guess. After spatial discretisation, with finite difference schemes, the authors time-integrated (7.10) to the steady-state, applying the recurring

equation:

$$f^{(k+1)} = f^{(k)} - \Delta t\, \psi'(f^{(k)}), \quad k = 0, 1, \ldots \tag{7.11}$$

with a constant step size Δt (and some restrictions on the size for stability). The approach is therefore a steepest descent method, with fixed step size.

The approach has a few limitations, mainly due to stability and efficiency. Rudin *et. al.* did not perform a line search along $\psi'(f^{(k)})$, as this would involve some processing time, and opted for a fixed step Δt. Issues of stability arise from choosing the step size a priori, as $f^{(k+1)}$ should fall within the *trust region* around $f^{(k)}$. In other words, as a linear approximation to ψ is used in (7.11), not every step size Δt is guaranteed to produce a reduction of ψ (the approximation has local validity). The step size Δt must be therefore restricted in order to guarantee that

$$\psi(f^{(k)} - \Delta t\, \psi'(f^{(k)})) < \psi(f^{(k)}) \tag{7.12}$$

is satisfied. The neighbourhood around $f^{(k)}$, identified by all the possible step sizes Δt and search directions $\psi'(f^{(k)})$ for which (7.12) holds, is called trust region of $f^{(k)}$, and it is commonly used to restrict the possible step sizes.

As Rudin *et. al.* in their algorithm have chosen a fixed the step size, the step size had to be small in order to result a priori within the trust region of $f^{(k)}$ for each k. As a result the convergence of their implementation was slow. The steepest descent algorithm is however slow even with a line search procedure, as its convergence is linear [130]. Secondly, the method proves to be particularly inefficient when applied to (7.6) as the behavior of $\psi(f)$ results in small trust regions, and therefore small step sizes as noted by several authors [86] [117] [124] [128].

Numerical Experiment

For the purpose of comparison with more efficient methods we have applied the steepest descent method 2D EIT reconstruction. The Tikhonov formulation for the EIT problem, substituting the proper variables in (7.6) and writing it as a non-linear inversion, is

$$\begin{aligned} \boldsymbol{\sigma}_{rec} &= \operatorname{argmin} \psi(\boldsymbol{\sigma}) \\ \psi(\boldsymbol{\sigma}) &= \tfrac{1}{2}\|h(\boldsymbol{\sigma}) - \mathbf{d}\|^2 + \alpha\, TV_\beta(\boldsymbol{\sigma}) \end{aligned} \tag{7.13}$$

Steepest Descent Algorithm

find homogeneous $\boldsymbol{\sigma}_0 : \boldsymbol{\sigma}_0 = \text{argmin } \|h(\boldsymbol{\sigma}_0) - \mathbf{d}\|$;
set β
k=0;
while k<kmax
  $\delta \mathbf{d}_k = (h(\boldsymbol{\sigma}_k) - \mathbf{d})$;
  $J_k = J(\boldsymbol{\sigma}_k)$;
  $E_k = \text{diag } (\sqrt{|L_i \boldsymbol{\sigma}_k|^2 + \beta})$;
  $\psi'_k = J_k^T (h(\boldsymbol{\sigma}_k) - \mathbf{d}_k) + \alpha L^T E_k^{-1} L \boldsymbol{\sigma}_k$;
  $\psi'_k = \psi'_k / \|\psi'_k\|$;
  $\lambda = \text{argmin } \psi(\boldsymbol{\sigma}_k - \lambda \psi'_k)$;
  $\boldsymbol{\sigma}_{k+1} = \boldsymbol{\sigma}_k - \lambda \psi'_k$;
  k=k+1;
end while

Figure 7.1: Pseudo code for the Steepest Descent algorithm.

As the conductivity is discretised on a mesh using piecewise constant representation, and with the notation introduced for (6.29), the TV_β functional becomes

$$TV_\beta(\boldsymbol{\sigma}) = \sum_{i=1}^{N} \sqrt{|L_i \boldsymbol{\sigma}|^2 + \beta} \tag{7.14}$$

The discretised diffusion operator $\mathcal{L}_\beta(\boldsymbol{\sigma})$ can be expressed as

$$\mathcal{L}_\beta(\boldsymbol{\sigma}) = L^T E^{-1} L \tag{7.15}$$

where $E = \text{diag}(\eta_i)$ and $\eta_i = \sqrt{|L_i \boldsymbol{\sigma}|^2 + \beta}$. The gradient $\psi'(\boldsymbol{\sigma})$ becomes

$$\psi'(\boldsymbol{\sigma}) = J^T (h(\boldsymbol{\sigma}) - \mathbf{d}) + \alpha L^T E^{-1} L \boldsymbol{\sigma} \tag{7.16}$$

The steepest descent algorithm was implemented using the following procedure:

In order to test the algorithm a resistivity profile with a background value of 1 $\Omega \cdot \text{m} \cdot 10^{-2}$ containing three inclusions with respective values of 0.5, 1.3 and 2 $\Omega \cdot \text{m} \cdot 10^{-2}$ was generated, as shown in the left part of Figure 7.2. Reconstructions were performed solving the non linear problem by updating the matrix J at each iteration, and performing a simple non-linear line search procedure on the objective function $\psi(\boldsymbol{\sigma})$ with the following fixed step lengths: $\lambda = [1 \cdot 10^{-4}, 1 \cdot 10^{-3}, 1 \cdot 10^{-2}, 1 \cdot 10^{-1}, 0.2, 0.5, 0.8, 1]$.

Different values of β were used in the reconstructions, precisely $1 \cdot 10^{-2}, 1 \cdot 10^{-4}, 1 \cdot 10^{-6}$.

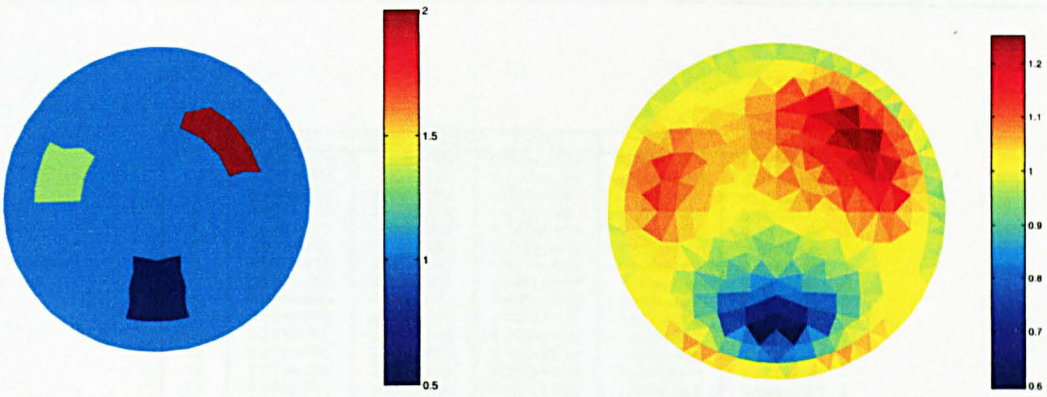


Figure 7.2: On the left, test resistivity profile. On the right, 20^{th} step of the steepest descent reconstruction with $\beta = 1 \cdot 10^{-6}$. (Conductivity scale in arbitrary units)

The value of $\beta = 1 \cdot 10^{-2}$ produces a good smoothing of $\psi(\boldsymbol{\sigma})$, but results in a poor approximation of the TV functional as, with the chosen mesh and conductivity values, the mean value of $|L_i\,\boldsymbol{\sigma}|$ is $8 \cdot 10^{-3}$. The two smaller values of β: $1 \cdot 10^{-4}$ and $1 \cdot 10^{-6}$ produce less smoothing but are better approximation of the ℓ_1 behavior of the regularisation functional.

Figure 7.2 shows the reconstructed conductivity at the 20^{th} step of the steepest descent algorithm, for a value of β of $1 \cdot 10^{-6}$. Table 7.1 reports the main numerical parameters during the iterations of the algorithm, for the same choice of β. Different values of β did not have an influence on the convergence of the algorithm, and on the final reconstructed image, as the algorithm is slow to converge for all β's values. Figure 7.3 displays the behavior of the gradient's norm $\|\psi'_k\|$, showing a very slow reduction with iterations. From the numerical experiments, the steepest descent algorithm becomes inefficient as the step size after the first few iterations becomes small, preventing the algorithm from achieving a significant reduction of the objective function, and resulting in a very slow convergence rate. The algorithm has shown the same slow behavior even for the extreme value of $\beta = 10$ for which a reduction of $\|\psi'_k\|$ to the order of $1 \cdot 10^{-2}$ is obtained at the 20^{th} iteration (as obtained for smaller values of β).

7.2.2 Newton's Method

Newton's Method has been considered by Vogel and Oman [124], Chan, Zhou and Chan [125], and by Chan, Golub and Mulet [117] as an obvious choice over the poor effectiveness of the steepest descent algorithm. Newton's Method exhibits in fact quadratic performance in a region around the optimal point [130]. The method works by locally approximating

Iter.	$\psi(\sigma)$	$\|h(\sigma) - \mathrm{d}\|$	$TV(\sigma)$	λ	$\|\psi'(\sigma)\|$
0	2.430e+000	1.559e+000	0.000e+000	0.000e+000	1.557e+000
1	5.131e-001	7.162e-001	1.596e+001	1.000e+000	4.079e-001
2	3.368e-001	5.801e-001	2.164e+001	5.000e-001	3.532e-001
3	2.614e-001	5.111e-001	2.070e+001	2.000e-001	2.090e-001
4	2.158e-001	4.643e-001	2.182e+001	2.000e-001	2.284e-001
5	1.861e-001	4.312e-001	2.232e+001	1.000e-001	1.397e-001
6	1.074e-001	3.274e-001	2.817e+001	5.000e-001	2.556e-001
7	8.103e-002	2.842e-001	2.840e+001	1.000e-001	7.212e-002
8	6.027e-002	2.449e-001	3.082e+001	2.000e-001	9.387e-002
9	5.858e-002	2.414e-001	3.090e+001	1.000e-002	7.595e-002
10	5.721e-002	2.385e-001	3.100e+001	1.000e-002	6.205e-002
11	5.491e-002	2.336e-001	3.223e+001	1.000e-001	1.238e-001
12	5.266e-002	2.288e-001	3.227e+001	1.000e-002	1.011e-001
13	5.085e-002	2.248e-001	3.232e+001	1.000e-002	7.998e-002
14	4.945e-002	2.216e-001	3.238e+001	1.000e-002	6.167e-002
15	4.836e-002	2.192e-001	3.246e+001	1.000e-002	4.826e-002
16	4.748e-002	2.171e-001	3.364e+001	1.000e-001	1.236e-001
17	4.526e-002	2.119e-001	3.367e+001	1.000e-002	9.908e-002
18	4.351e-002	2.078e-001	3.372e+001	1.000e-002	7.592e-002
19	4.221e-002	2.046e-001	3.379e+001	1.000e-002	5.541e-002
20	4.127e-002	2.023e-001	3.388e+001	1.000e-002	4.015e-002

Table 7.1: Main numerical parameters for the first 20 iterations of the Steepest Descent algorithm with $\beta = 1 \cdot 10^{-6}$.

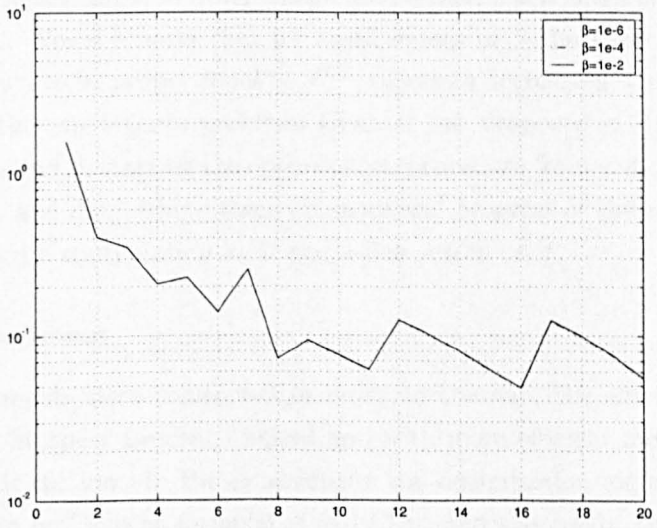


Figure 7.3: Gradient's norm during the first 20 steps of the steepest descent algorithm for the three values of β. The graphical results are superimposed as the numerical values are close.

the objective function with second order terms (see Section 2.3). For the Tikhonov TV_β regularised inverse problem (7.6) the gradient has been shown to be

$$\psi'(f) = A^T(Af - d) + \alpha \mathcal{L}_\beta(f)f \tag{7.17}$$

consequently the Hessian is

$$\psi''(f) = A^T A + \alpha \mathcal{L}_\beta(f) + \mathcal{L}'_\beta(f)f \tag{7.18}$$

with the operator $\mathcal{L}'_\beta(f)$ equal to [125]

$$\mathcal{L}'_\beta(f)v = \nabla \cdot \left(\frac{\nabla v \cdot \nabla f}{\sqrt{(|\nabla f|^2 + \beta)^3}} \nabla f \right) \tag{7.19}$$

and the iteration of Newton's Method being

$$f^{(k+1)} = f^{(k)} - [\psi''(f^{(k)})]^{-1}\psi'(f^{(k)}), \quad k = 0, 1, \ldots \tag{7.20}$$

The convergence properties of the method were studied Vogel *et. al.* in [131] and by Chan *et. al.* in [125]. Both authors report that the method results convergent only for values of β which are sufficiently large; in other words the method has a small convergence region, with respect to β. Vogel reports that for small values of β the size of the convergence region can be shown to be proportional to $\beta^{3/2}$, therefore decreasing with beta decreasing.

To overcome the convergence problems Chan *et. al.* proposed in [125] a continuation method on both α and β, starting the optimisation procedure from sufficiently large values of the parameters and diminishing them on successful progress of the algorithm, while in [117] they proposed a continuation on β and a line search on f.

Numerical Experiment

Numerical experiments were conducted in order to compare the efficiency of Newton's Method with the Steepest Descent method and with more efficient methods to be introduced later in this section. In the experiments the optimisation algorithm was applied to the EIT inverse problem as formulated in (7.13), and the conductivity test profile of Figure 7.2 was used.

The method requires the calculation of gradient $\psi'(\boldsymbol{\sigma})$ and of the Hessian $\psi''(\boldsymbol{\sigma})$ at each step. The gradient can be calculated as in (7.16). With the notation introduced for

(7.15) the discretised operator $\mathcal{L}'_\beta(\boldsymbol{\sigma})\boldsymbol{\sigma}$ of equation (7.18) becomes

$$\mathcal{L}'_\beta(\boldsymbol{\sigma})\boldsymbol{\sigma} = L^T E^{-1} Z L \tag{7.21}$$

with

$$Z = \mathrm{diag}(\frac{|L_i\boldsymbol{\sigma}|^2}{\eta_i^2}) \tag{7.22}$$

and the Hessian of the objective function

$$\psi''(\boldsymbol{\sigma}) = J^T J + \alpha L^T E^{-1} F L \tag{7.23}$$

with

$$F = \mathrm{diag}(1 - \frac{|L_i\boldsymbol{\sigma}|^2}{\eta_i^2}) \tag{7.24}$$

The Newton's Method algorithm can be therefore implemented with the procedure of Figure 7.4. Numerical experiments were conducted both with and without continuation techniques on β and line searches on $\boldsymbol{\sigma}$. When applied to the test EIT inverse problem Newton's Method appeared to be particularly unstable even for large values of β if no continuation and line search procedures were adopted. The algorithm converged in fact only for values of β greater then 0.5. The left part of Figure 7.5 shows the plot of the norm of gradient ψ'_k during iterations for values of β equal to $10, 1, 0.5$. The convergence of the method in this case is quicker than that observed with the Steepest Descent method, as in 20 steps a reduction of 7 decades is obtained for the gradient's norm, as opposed to 1.5 decades for the second method. Our interest however is to use smaller values for the parameter β, to have a ℓ_1 behaviour for TV_β.

To achieve the convergence of the algorithm for smaller values of beta, both a continuation technique technique on β and a line search on $\boldsymbol{\sigma}$ were adopted. The continuation procedure on β decreases the parameter by 20% at each iteration if a reduction of the objective function $\psi(\boldsymbol{\sigma}_k)$ was achieved. The reduction rate was chosen empirically, as it resulted in a steady reduction of the objective function. The line search procedure that was adopted is the same as used for the Steepest Descent algorithm, where $\psi(\boldsymbol{\sigma}_k + \lambda\,\delta\boldsymbol{\sigma}_k)$ is evaluated for the fixed values of $\lambda = [1 \cdot 10^{-4}, 1 \cdot 10^{-3}, 1 \cdot 10^{-2}, 1 \cdot 10^{-1}, 0.2, 0.5, 0.8, 1]$.

The right part of Figure 7.5 shows the plot of the norm of gradient ψ'_k for the two different initial values $\beta = 10$ and $\beta = 1 \cdot 10^{-1}$. In the first case, for the first 8 iterations the reduction rate of $\|\psi'_k\|$ is higher than with no continuation and line search techniques. The progress of the algorithm comes however to a halt after the 11^{th} iteration, when the gradient's norm keeps oscillating in between the values $1 \cdot 10^{-4}$ and $1 \cdot 10^{-5}$ (if iterations

are prolonged). With the smaller initial value of β the algorithm progress is slower, and the gradient of the objective function is not reduced beyond $1 \cdot 10^{-2}$. Continuation and line search procedures seem therefore to bring stability to the algorithm, but performance remains compromised by choices of small betas.

The main numerical parameters for both the initial choices of β are reported in Table 7.1. In both cases a similar final value of the objective function is reached, $2.6 \cdot 10^{-5}$ in the first case and $7.0 \cdot 10^{-5}$ in the second. The final value total variation of the reconstructed conductivity is smaller when the algorithm is started with $\beta = 0.1$, as the TV functional is approximated better. The step size for the smaller initial value of β deteriorates during iterations becoming quickly very small and preventing the algorithm from making further significant progress. Table 7.1 reports also the angle in degrees between the Newton direction and the Steepest Descent direction. The angle always appears to be high, reaching almost orthogonality at the last iterations. At first we tried to relate the behaviour of Newton's algorithm to the fact that the curvature of $\psi(\boldsymbol{\sigma})$ could change abruptly in correspondence with a sign change in $|L_i \boldsymbol{\sigma}|$. We thought that Newton's algorithm, using curvature information, could avoid sign changes by following directions almost perpendicular to $-\psi'(\boldsymbol{\sigma})$, still yielding a significant reduction on ψ having possibly a larger trust region then the Steepest Descent method. In practice we found that the step directions and step lengths are not directly linked to the "fine" behaviour of ψ as for both the algorithms each step corresponds to tenths or hundredths of sign changes in $|L_i \boldsymbol{\sigma}|$. The Newton's algorithm seems to lead to step directions and lengths that results in more sign changes in $|L_i \boldsymbol{\sigma}|$ at each iteration, and to be able in this way to identify sooner the right relative conductivity values for the elements.

7.2.3 Iterative Reweighting Properties

An interesting insight on the regularisation properties of the TV_β functional is given by inspection of the Newton's iteration 7.20 for the discrete case:

$$[J^T J + \alpha L^T E^{-1} F L]\delta\boldsymbol{\sigma} = -[J^T(h(\boldsymbol{\sigma}) - \mathbf{d}) + \alpha\, L^T E^{-1} L\boldsymbol{\sigma}] \tag{7.25}$$

Apart from the presence of diagonal matrices $E^{-1}F$ at the LHS and E^{-1} at the RHS (7.25) is similar to the normal equations of a ℓ_2 regularised problem, with a penalty term equal to $\|L\boldsymbol{\sigma}\|^2$. Equation (7.25) is however non–linear as both E and F depend on $\boldsymbol{\sigma}$. From this point of view, the Newton's iteration can be thought to be an IRLS (Iterative Reweighted Least Squares)[132] [133] [134] method applied to the penalty term TV_β. The

Newton's Method Algorithm

find homogeneous $\boldsymbol{\sigma}_0 : \boldsymbol{\sigma}_0 = \text{argmin } \|h(\boldsymbol{\sigma}_0) - \mathrm{d}\|$;
set initial β
k=0;
while k<kmax
  $\delta \mathrm{d}_k = (h(\boldsymbol{\sigma}_k) - \mathrm{d})$;
  $J_k = J(\boldsymbol{\sigma}_k)$;
  $E_k = \text{diag } (\sqrt{|L_i\,\boldsymbol{\sigma}_k|^2 + \beta})$;
  $F_k = \text{diag } (1 - \frac{|L_i \boldsymbol{\sigma}|^2}{\eta_i^2})$;
  $\psi'_k = J_k^T(h(\boldsymbol{\sigma}_k) - \mathrm{d}_k) + \alpha\, L^T E_k^{-1} L\, \boldsymbol{\sigma}_k$;
  $\psi''_k = J^T J + \alpha L^T E_k^{-1} F_k L$;
  $\delta\boldsymbol{\sigma}_k = -[\psi''_k]^{-1}\psi'_k$;
  $\lambda = \text{argmin } \psi(\boldsymbol{\sigma}_k + \lambda\,\delta\boldsymbol{\sigma}_k)$;
  if sufficient reduction of $\psi(\boldsymbol{\sigma}_k)$ achieved
    $\boldsymbol{\sigma}_{k+1} = \boldsymbol{\sigma}_k + \lambda\,\delta\boldsymbol{\sigma}_k$;
    decrease β
    k=k+1;
  else
    increase β
  end if else
end while

Figure 7.4: Pseudo code for the Newton's Method algorithm with continuation on β and line search on $\boldsymbol{\sigma}$.

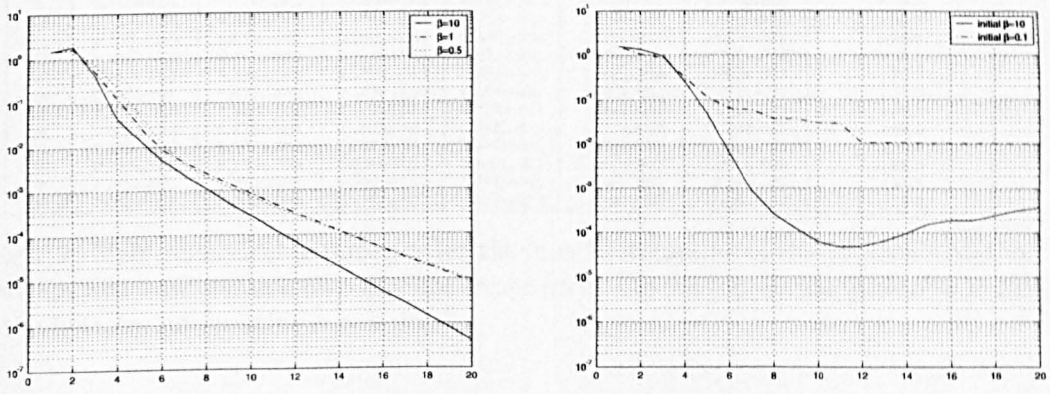


Figure 7.5: On the left, plot of $\|\psi'_k\|$ for the Newton's Method algorithm with no continuation on β and no line search on $\boldsymbol{\sigma}$, for two different constant values of β. On the right, plot of $\|\psi'_k\|$ for the same algorithm and with continuation on β and line search on $\boldsymbol{\sigma}$, for two different initial values of β.

Iter.	$\psi(\sigma)$	$\|h(\sigma)-\mathbf{d}\|$	$TV(\sigma)$	β	λ	$\|\psi'(\sigma)\|$	Angle$(\delta\sigma, -\psi'(\sigma))$
0	2.431e+000	1.559e+000	0.000e+000	1.000e+001	0.000e+000	1.557e+000	+81.7
1	1.060e+000	1.029e+000	2.911e+002	8.000e+000	5.000e-001	1.341e+000	+82.7
2	5.001e-001	7.071e-001	1.687e+002	6.400e+000	5.000e-001	9.594e-001	+77.8
3	2.599e-002	1.607e-001	9.236e+001	5.120e+000	1.000e+000	2.465e-001	+84.4
4	1.256e-003	3.331e-002	7.418e+001	4.096e+000	1.000e+000	5.044e-002	+86.6
5	1.508e-004	4.505e-003	7.074e+001	3.277e+000	1.000e+000	6.541e-003	+88.8
6	1.175e-004	7.770e-004	6.996e+001	2.621e+000	1.000e+000	9.564e-004	+89.6
7	1.048e-004	3.530e-004	6.947e+001	2.097e+000	1.000e+000	2.732e-004	+89.8
8	9.394e-005	3.041e-004	6.899e+001	1.678e+000	1.000e+000	1.245e-004	+89.9
9	8.422e-005	2.998e-004	6.832e+001	1.342e+000	1.000e+000	6.162e-005	+89.9
10	7.556e-005	3.099e-004	6.767e+001	1.074e+000	1.000e+000	4.780e-005	+89.9
11	6.782e-005	3.211e-004	6.690e+001	8.590e-001	1.000e+000	4.853e-005	+89.9
12	6.093e-005	3.338e-004	6.608e+001	6.872e-001	1.000e+000	6.402e-005	+89.9
13	5.477e-005	3.498e-004	6.515e+001	5.498e-001	1.000e+000	9.637e-005	+89.9
14	4.929e-005	3.757e-004	6.415e+001	4.398e-001	1.000e+000	1.533e-004	+89.9
15	4.441e-005	3.965e-004	6.326e+001	3.518e-001	8.000e-001	1.867e-004	+89.9
16	4.004e-005	4.033e-004	6.255e+001	2.815e-001	5.000e-001	1.914e-004	+89.9
17	3.618e-005	4.369e-004	6.162e+001	2.252e-001	5.000e-001	2.513e-004	+89.9
18	3.274e-005	4.827e-004	6.053e+001	1.801e-001	5.000e-001	3.241e-004	+89.9
19	2.968e-005	5.152e-004	5.936e+001	1.441e-001	5.000e-001	3.708e-004	+89.8
20	2.692e-005	5.186e-004	5.819e+001	1.153e-001	5.000e-001	3.737e-004	+89.8

Iter.	$\psi(\sigma)$	$\|h(\sigma)-\mathbf{d}\|$	$TV(\sigma)$	β	λ	$\|\psi'(\sigma)\|$	Angle$(\delta\sigma, -\psi'(\sigma))$
0	2.430e+000	1.559e+000	0.000e+000	1.000e-001	0.000e+000	1.557e+000	+72.5
1	4.373e-001	6.612e-001	1.761e+002	8.000e-002	8.000e-001	1.053e+000	+86.3
2	3.209e-001	5.665e-001	1.238e+002	6.400e-002	2.000e-001	8.831e-001	+82.5
3	4.912e-002	2.216e-001	9.003e+001	5.120e-002	1.000e+000	3.064e-001	+84.9
4	7.716e-003	8.774e-002	6.859e+001	4.096e-002	1.000e+000	1.181e-001	+87.5
5	2.510e-003	4.993e-002	5.896e+001	3.277e-002	8.000e-001	6.420e-002	+89.3
6	2.210e-003	4.686e-002	5.651e+001	2.621e-002	1.000e-001	6.007e-002	+88.6
7	1.038e-003	3.200e-002	5.796e+001	2.097e-002	8.000e-001	3.746e-002	+89.7
8	9.712e-004	3.096e-002	5.408e+001	1.678e-002	1.000e-001	3.675e-002	+89.1
9	6.756e-004	2.576e-002	5.343e+001	1.342e-002	5.000e-001	2.990e-002	+89.6
10	6.324e-004	2.493e-002	5.171e+001	1.074e-002	1.000e-001	2.878e-002	+88.5
11	8.474e-005	8.624e-003	5.242e+001	8.590e-003	1.000e+000	1.074e-002	+89.9
12	8.288e-005	8.554e-003	5.210e+001	6.872e-003	1.000e-002	1.065e-002	+89.9
13	8.111e-005	8.484e-003	5.178e+001	5.498e-003	1.000e-002	1.057e-002	+89.9
14	7.941e-005	8.415e-003	5.148e+001	4.398e-003	1.000e-002	1.048e-002	+89.9
15	7.779e-005	8.345e-003	5.117e+001	3.518e-003	1.000e-002	1.040e-002	+89.9
16	7.623e-005	8.276e-003	5.087e+001	2.815e-003	1.000e-002	1.031e-002	+89.9
17	7.473e-005	8.207e-003	5.057e+001	2.252e-003	1.000e-002	1.022e-002	+89.9
18	7.329e-005	8.138e-003	5.028e+001	1.801e-003	1.000e-002	1.014e-002	+89.9
19	7.188e-005	8.070e-003	4.999e+001	1.441e-003	1.000e-002	1.005e-002	+89.9
20	7.051e-005	8.000e-003	4.972e+001	1.153e-003	1.000e-002	9.966e-003	+89.9

Table 7.2: Main numerical parameters for the first 20 iterations of the Newton's Method algorithm with continuation on β and line search on $\boldsymbol{\sigma}$. On the top, initial value $\beta = 1 \cdot 10^{1}$, On the bottom, initial value $\beta = 1 \cdot 10^{-1}$.

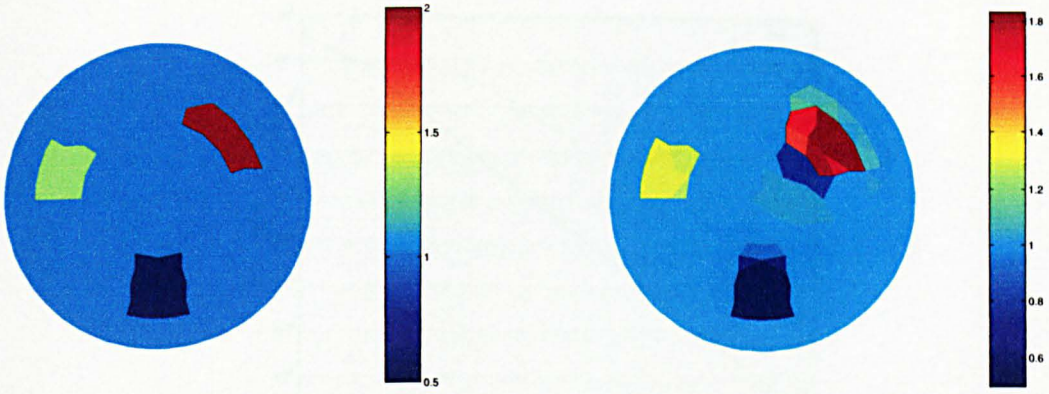


Figure 7.6: On the left, test resistivity profile. On the right, 40^{th} step of the Newton's Method reconstruction with continuation on β, initial value $\beta = 0.1$, and line search on $\boldsymbol{\sigma}$. (Conductivity scale in arbitrary units)

IRLS method has been widely used for the solution of the ℓ_p problems of the kind

$$\min_{\mathbf{x}} \frac{1}{p}\|A\mathbf{x} - \mathbf{y}\|_p^p \tag{7.26}$$

By defining the residuals $\mathbf{r} = A\mathbf{x} - \mathbf{y}$, the gradient of the objective function in (7.26) can be expressed as

$$A^T W(A\mathbf{x} - \mathbf{y}) \tag{7.27}$$

where the elements of diagonal the weighting matrix W are $W_i = |\mathbf{r}_i|^{p-2}$, and the so–called *non–linear normal equations* can be expressed as

$$A^T W A\mathbf{x} = A^T W\mathbf{y} \tag{7.28}$$

Equations (7.28) are called non–linear normal equations as the weights W depend on $\mathbf{x}$. The method works by iterating (7.28) and updating W using the current residuals values until convergence is reached. The base method can be generalised [42], to allow solution of the following problems

$$\min_{\mathbf{x}} \phi(A\mathbf{x} - \mathbf{y}) \tag{7.29}$$

This formulation is used to handle non-differentiability occurring for $p = 1$, by choosing ϕ similarly as done in 7.1. The reweighting technique allows thus to solve ℓ_p–norm problems by repeated application of traditional least squares algorithms. The IRLS algorithm successfully achieves convergence, the convergence is however linear [135] [136]; usually more efficient algorithms are preferred [137] [138]. Taking the view that the Newton's algorithm

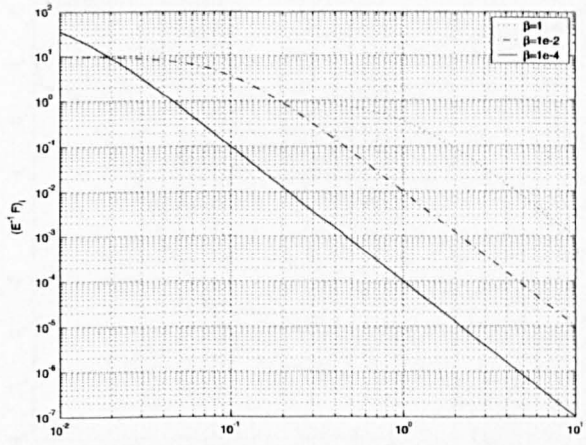


Figure 7.7: Logarithmic plot of the weights $(E^{-1}F)_i$ against $|L_i\boldsymbol{\sigma}|$ for several values of β.

is an IRLS method applied to the TV_β functional, this can explain why the convergence of that algorithm is not rapid as expected. However, these considerations do not apply fully to our case, as the situation is complicated by the fact that the reweighting procedure is actually dynamically varying the regularisation of an ill-posed inverse problem. In the product $L^T E^{-1} L$ at the RHS of (7.25), the diagonal elements of $E^{-1} = \mathrm{diag}(\eta_i^{-1})$, with

$$\eta_i^{-1} = \frac{1}{\sqrt{|L_i\boldsymbol{\sigma}|^2 + \beta}} \tag{7.30}$$

are a monotonically decreasing function of $|L_i\boldsymbol{\sigma}|$. As EIT algorithms are usually started from homogeneous initial guesses, for which $L_i\boldsymbol{\sigma}$ is null on the whole domain, the regularisation weights η_i^{-1} are initially dominated by β: the first step corresponds therefore to a traditional ℓ_2 regularised inversion. As areas of contrast arise, regularisation weights behave like $1/\sqrt{|L_i\boldsymbol{\sigma}|}$ for large gradients in $\boldsymbol{\sigma}$. The regularisation therefore penalises large variations less and less as they appear in the reconstructed profile. Analysis of the TV_β regularised Newton's Method step depends therefore on the dynamic behaviour of such weights. Besides, the numerical well-posedness of (7.25) depends on the properties of the matrix pencil $(J^T J + \alpha L^T E^{-1} F L)$. The diagonal elements of the Hessian's weighting matrix $E^{-1}F$ are

$$(E^{-1}F)_i = \frac{1}{\sqrt{|L_i\boldsymbol{\sigma}|^2 + \beta}} \left(1 - \frac{|L_i\boldsymbol{\sigma}|^2}{|L_i\boldsymbol{\sigma}|^2 + \beta}\right) \tag{7.31}$$

and they behave like $1/\sqrt{\beta}$ for small gradients and like $1/|L_i\boldsymbol{\sigma}|^3$ for large gradients. Figure 7.7 shows the plot of $(E^{-1}F)_i$ against $|L_i\boldsymbol{\sigma}|$ for the three values of β 1, $1 \cdot 10^{-2}$ and $1 \cdot 10^{-4}$.

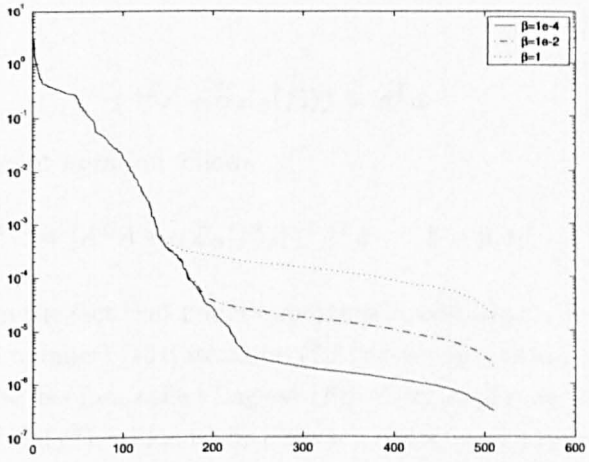


Figure 7.8: Logarithmic plot of the singular values of the matrix $(J^TJ + \alpha L^T E^{-1} FL)$ for several values of β.

For larger values of β the dynamic range of the weights is smaller covering approximately 3 decades for $\beta = 1$ and covering almost 9 decades for $\beta = 1 \cdot 10^{-4}$. During iterations, as the reconstruction approaches the true conductivity distribution (which could present discontinuities) $|L_i\boldsymbol{\sigma}|$ will assume a large range of values and consequently the regularisation weights will too. As some weights can become small during the reconstruction progress, the regularisation is likely to suffer. This effect is supported by the SVD analysis of the matrix $(J^TJ + \alpha L^T E^{-1} FL)$ calculated using the true conductivity profile (to which the algorithm tends as it progresses). Figure 7.8 shows a logarithmic plot of the singular values of the Hessian matrix for the three different values of β. The conditioning of the matrix becomes poorer and poorer as β is reduced, due to the fact that the weights values assume a larger range. The sensitivity of Newton's Method to the value of β, and it's potential instability [124] [125] [117] can therefore be explained in terms of the dynamic regularisation properties that the terms E^{-1} and $E^{-1}F$ bring to (7.25).

7.2.4 Lagged Diffusivity Method (and Fixed Point iteration)

The Legged Diffusivity method was introduced by Vogel and Oman [124] and others [131] [114] as an alternative to the Steepest Descent and Newton's Methods, given the poor performance of the first, and stability issues of the second. The method was originally formulated as a fixed point iteration, obtained by setting the gradient of the objective function in (7.6) to zero

$$A^TAf - A^Td + \alpha \mathcal{L}_\beta(f)f = 0 \tag{7.32}$$

or

$$(A^T A + \alpha \mathcal{L}_\beta(f))f = A^T d \tag{7.33}$$

from which the fixed point iteration follows

$$f^{k+1} = [A^T A + \alpha \mathcal{L}_\beta(f^k)]^{-1} A^T d \qquad k = 0, 1, \ldots \tag{7.34}$$

The method is based on the fact that under opportune conditions on A (linearity, injectivity, bounded condition number) [131] iteration (7.34) converges to the minimiser of (7.6). In literature the method is often called Lagged Diffusivity approach, as each f^{k+1} is calculated on the basis of $\mathcal{L}_\beta(f^k)$, which is the diffusion operator at the preceding step. The equivalent Quasi–Newton's Method can be derived from the Newton's Method dropping the term $\mathcal{L}'_\beta(f)$ from the Hessian matrix, with the following iteration

$$f^{(k+1)} = f^{(k)} - [A^T A + \alpha \mathcal{L}_\beta(f)^{(k)}]^{-1} \psi'(f^{(k)}), \quad k = 0, 1, \ldots \tag{7.35}$$

The method is believed to have better convergence properties with respect to Newton Method thanks to dropping the negative semidefinite matrix $\mathcal{L}'_\beta(f)$ from the Hessian. The method is globally convergent, though local convergence is only linear [114].

Numerical Experiment

With the notation introduced in the preceding paragraphs, the Lagged Diffusivity iteration for the discretised inverse problem is

$$[J^T J + \alpha L^T E^{-1} L]\delta\boldsymbol{\sigma} = -[J^T(h(\boldsymbol{\sigma}) - \mathbf{d}) + \alpha L^T E^{-1} L\,\boldsymbol{\sigma}] \tag{7.36}$$

Numerical experiments were conducted on the EIT inverse problem (7.13) using the same test conductivity profile as for the Steepest Descent and Newton's Methods tests, using the formulation of Figure 7.9. The algorithm appeared stable for smaller values of β in comparison to Newton's algorithm, being convergent for values of the parameter as small as $1 \cdot 10^{-4}$ requiring no continuation. Such techniques were however adopted in the algorithm to reduce the smoothing parameter even further during iterations. The continuation policy reduced β by 50% on each successful step, starting from initial values of β equal to $1, 1 \cdot 10^{-2}, 1 \cdot 10^{-4}$. The method showed good progress, being able to reduce $\|\psi'(\boldsymbol{\sigma})\|$ even at the last iterations, when β reached values smaller than $1 \cdot 10^{-10}$. Figure 7.10 shows the norm gradient of the objective function during the progress of the algorithm, Table 7.3 reports the main numerical parameters for the initial value of $\beta = 1 \cdot 10^{-2}$. In

Lagged Diffusivity Algorithm

find homogeneous $\boldsymbol{\sigma}_0 : \boldsymbol{\sigma}_0 = \text{argmin } \|h(\boldsymbol{\sigma}_0) - \mathrm{d}\|$;
set β
k=0;
while k<kmax
  $\delta \mathrm{d}_k = (h(\boldsymbol{\sigma}_k) - \mathrm{d})$;
  $J_k = J(\boldsymbol{\sigma}_k)$;
  $E_k = \text{diag } (\sqrt{\|L_i \boldsymbol{\sigma}_k\|^2 + \beta})$;
  $\psi'_k = J_k^T(h(\boldsymbol{\sigma}_k) - \mathrm{d}_k) + \alpha L^T E_k^{-1} L \boldsymbol{\sigma}_k$;
  $\psi''_k = J^T J + \alpha L^T E_k^{-1} L$;
  $\delta\boldsymbol{\sigma}_k = -[\psi''_k]^{-1}\psi'_k$;
  $\lambda = \text{argmin } \psi(\boldsymbol{\sigma}_k + \lambda\,\delta\boldsymbol{\sigma}_k)$;
  if sufficient reduction of $\psi(\boldsymbol{\sigma}_k)$ achieved
    $\boldsymbol{\sigma}_{k+1} = \boldsymbol{\sigma}_k + \lambda\,\delta\boldsymbol{\sigma}_k$;
    decrease β
    k=k+1;
  else
    increase β
  end if else
end while

Figure 7.9: Pseudo code for the Lagged Diffusivity algorithm with line search on $\boldsymbol{\sigma}$.

comparison with Newton's Method the algorithm is capable of achieving a better reduction in the objective function's gradient, approximately one order of magnitude higher, and to do that for much smaller values of the smoothing parameter. The Lagged Diffusivity algorithm is therefore an improvement over Newton's Method, especially from the stability point of view. The method is however less efficient with respect to Primal Dual methods, which are introduced in the next section.

7.3 Primal Dual Methods for the Mixed ℓ_1/ℓ_2 Problem

As seen in Paragraph 7.2, traditional methods for TV_β regularised inversion solve the following minimisation problem, called primal (P)

$$\begin{aligned} f_{rec} &= \text{argmin}\, \psi(f) \\ \psi(f) &= \tfrac{1}{2}\|Af - d\|^2 + \alpha\, TV_\beta(f) \end{aligned} \tag{7.37}$$

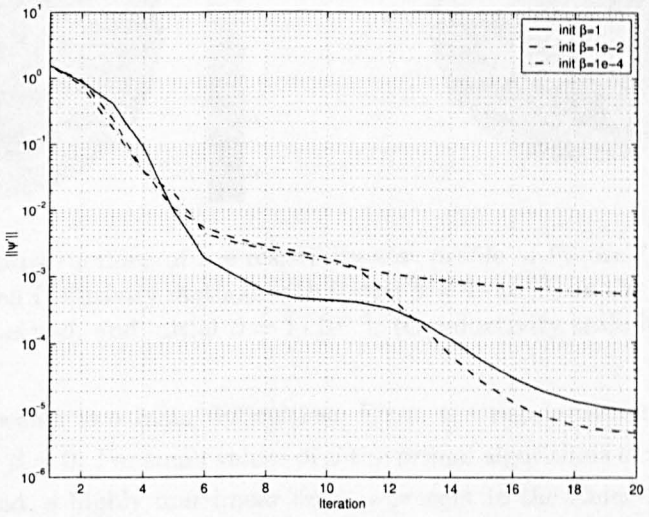


Figure 7.10: Gradient's norm during the first 20 steps of the Lagged Diffusivity algorithm for three initial values of β.

Iter.	$\psi(\boldsymbol{\sigma})$	$\|h(\boldsymbol{\sigma}) - \mathbf{d}\|$	$TV(\boldsymbol{\sigma})$	β	λ	$\|\psi'(\boldsymbol{\sigma})\|$	Angle$(\delta\boldsymbol{\sigma}, -\psi'(\boldsymbol{\sigma}))$
0	1.215e+000	1.559e+000	0.000e+000	1.000e-002	0.000e+000	1.557e+000	+70.3
1	1.739e-001	5.897e-001	1.326e+002	5.000e-003	8.000e-001	9.241e-001	+81.2
2	2.115e-002	2.056e-001	8.460e+001	2.500e-003	1.000e+000	2.439e-001	+80.5
3	5.403e-004	3.262e-002	6.985e+001	1.250e-003	1.000e+000	4.010e-002	+87.2
4	3.987e-005	8.137e-003	5.853e+001	6.250e-004	1.000e+000	1.039e-002	+89.1
5	1.663e-005	4.606e-003	5.373e+001	3.125e-004	8.000e-001	5.309e-003	+89.3
6	1.084e-005	3.242e-003	5.154e+001	1.563e-004	5.000e-001	3.652e-003	+89.5
7	8.638e-006	2.599e-003	4.970e+001	7.813e-005	5.000e-001	2.843e-003	+89.5
8	7.456e-006	2.212e-003	4.816e+001	3.906e-005	5.000e-001	2.353e-003	+89.5
9	6.368e-006	1.759e-003	4.692e+001	1.953e-005	5.000e-001	1.832e-003	+89.5
10	5.428e-006	1.258e-003	4.548e+001	9.766e-006	8.000e-001	1.296e-003	+89.4
11	4.706e-006	6.171e-004	4.454e+001	4.883e-006	1.000e+000	4.880e-004	+89.7
12	4.583e-006	5.209e-004	4.407e+001	2.441e-006	1.000e+000	1.865e-004	+89.7
13	4.536e-006	5.081e-004	4.380e+001	1.221e-006	1.000e+000	6.545e-005	+89.5
14	4.510e-006	5.066e-004	4.364e+001	6.104e-007	1.000e+000	2.656e-005	+89.1
15	4.494e-006	5.054e-004	4.354e+001	3.052e-007	1.000e+000	1.343e-005	+88.5
16	4.483e-006	5.040e-004	4.348e+001	1.526e-007	1.000e+000	8.104e-006	+88.0
17	4.475e-006	5.026e-004	4.344e+001	7.629e-008	1.000e+000	5.927e-006	+87.7
18	4.469e-006	5.011e-004	4.340e+001	3.815e-008	1.000e+000	5.038e-006	+87.6
19	4.465e-006	4.993e-004	4.338e+001	1.907e-008	1.000e+000	4.623e-006	+87.7
20	4.462e-006	4.974e-004	4.336e+001	9.537e-009	1.000e+000	4.502e-006	+87.9

Table 7.3: Main numerical parameters for the first 20 iterations of the Lagged Diffusivity algorithm with continuation on β and line search on $\boldsymbol{\sigma}$, for the initial value $\beta = 1 \cdot 10^{-2}$.

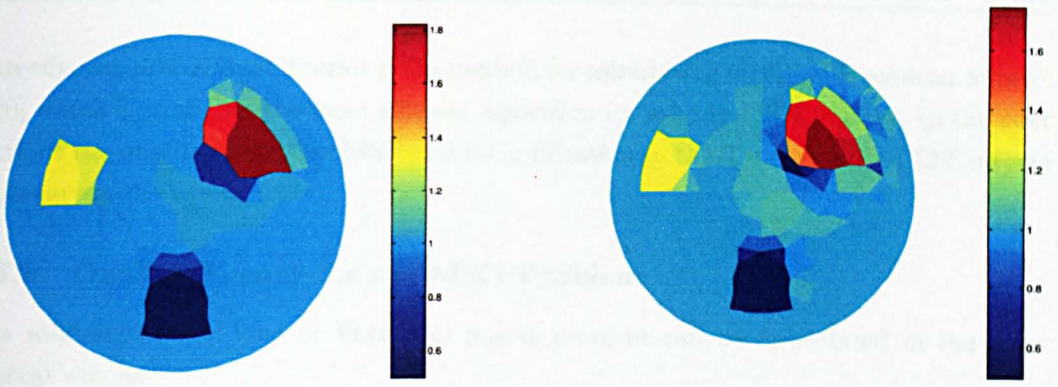


Figure 7.11: Reconstructions of the resistivity test profile of Figure 7.2. On the left, 20^{th} step of the Lagged Diffusivity method with initial $\beta = 1 \cdot 10^{-2}$, on the right reconstruction with the same method, and initial $\beta = 1 \cdot 10^{-4}$. (Conductivity scale in arbitrary units)

Solving 7.37 presents two main difficulties: First, the regularisation functional is non–differentiable for $\beta = 0$. For small values of β the primal algorithms can prove unstable and inefficient. Second, a highly non–linear term is present in the Euler–Lagrange equations, as we will discuss shortly. Apposite techniques are therefore needed to solve the primal problem. Similar conclusions arise from analysing the discretised form of 7.37 that, with the notation introduced for EIT and with $\beta \to 0$, can be written as

$$\begin{aligned} \boldsymbol{\sigma}_{rec} &= \operatorname{argmin} \psi(\boldsymbol{\sigma}) \\ \psi(\boldsymbol{\sigma}) &= \tfrac{1}{2}\|h(\boldsymbol{\sigma}) - \mathbf{d}\|^2 + \alpha \|L\,\boldsymbol{\sigma}\|_1 \end{aligned} \tag{7.38}$$

or equivalently as

$$\begin{aligned} \boldsymbol{\sigma}_{rec} &= \operatorname{argmin} \psi(\boldsymbol{\sigma}) \\ \psi(\boldsymbol{\sigma}) &= \tfrac{1}{2}\|h(\boldsymbol{\sigma}) - \mathbf{d}\|^2 + \alpha \sum_i \|L_i\,\boldsymbol{\sigma}\| \end{aligned} \tag{7.39}$$

The minimisation of $\|L\,\boldsymbol{\sigma}\|_1$ with constraints is classified in Optimisation as the ℓ_1 *linear approximation problem*. Several authors have proposed *ad hoc* techniques to overcome the associated difficulties [121] [139], [140], [118]. Similarly, when (7.38) is expressed as (7.39), the minimisation of $\sum_i \|L_i\,\boldsymbol{\sigma}\|$ with constraints is known in Optimisation as the *minimisation of sum of norms problem* (MSN) and has been studied by several authors [141], [142], [134], [119], [137], [119], [122] [120]. Andersen [120] reports that he was the first to solve the MSN problem in an efficient and accurate manner for large problems even when many norms $\|L_i\,\boldsymbol{\sigma}\|$ are zero at a solution point. Andersen [137] developed the duality theory for the MSN problem, which led to the publication with other coauthors of

"An efficient primal dual interior point method for minimising a sum of Euclidean norms" [120] which is probably the most efficient algorithm for solving MSN to date. In the next sections the duality theory for MSN and its application to the TV regularised EIT inverse problem are discussed.

7.3.1 Duality Theory for the MSN Problem

The minimisation of sum of Euclidean norms problem can be formulated in the most general way as

$$\min_{y} \sum_{i=1}^{n} \|A_i y - c_i\| \tag{7.40}$$

with $y \in \mathbb{R}^m; c_i \in \mathbb{R}^d$ and $A_i \in \mathbb{R}^{d\times m}$, which is equivalent to

$$(P) \quad \min_{y} \left\{ \sum_{i=1}^{n} \|z_i\| : A_i y + z_i = c_i, \quad i = 1, \ldots, n \right\} \tag{7.41}$$

with $z_i \in \mathbb{R}^d$. We call (7.40) primal problem. The duality of (P) can be described in the following way

$$\begin{aligned} \min_{y:A_iy+z_i=c_i} \sum_{i=1}^{n} \|z_i\| &= \min_{y:A_iy+z_i=c_i} \max_{x_i:\|x_i\|\le 1} \sum_{i=1}^{n} x_i^T z_i \\ &= \max_{x_i:\|x_i\|\le 1} \min_{y:A_iy+z_i=c_i} \sum_{i=1}^{n} x_i^T z_i \\ &= \max_{x_i:\|x_i\|\le 1} \min_{y\in\mathbb{R}^m} \left(\sum_{i=1}^{n} c_i^T x_i - y^T \sum_{i=1}^{n} A_i^T x_i \right) \\ &= \max_{x_i} \left\{ \sum_{i=1}^{n} c_i^T x_i : \|x_i\| \le 1; \sum_{i=1}^{n} A_i^T x_i = 0 \right\} \end{aligned} \tag{7.42}$$

where the first equality follows from Cauchy–Schwartz, the second from min-max theory [120] [143], the third trivially, and the fourth because if $\sum_{i=1}^{n} A_i^T x_i$ is not zero, the minimised value would be $-\infty$. The dual problem of (P) is therefore

$$(D) \quad \max_{x_i} \left\{ \sum_{i=1}^{n} c_i^T x_i : \|x_i\| \le 1; \sum_{i=1}^{n} A_i^T x_i = 0, \quad i = 1, \ldots, n \right\} \tag{7.43}$$

and the variables y are called primal variables and the variables $x_i \in \mathbb{R}^d$ dual variables. The problems (P) and (D) are therefore equivalent. The concept of duality and the relation

between primal and dual optimal points can be formalised defining the primal feasible region as

$$\mathcal{Y} = \left\{(y, z) \in \mathbb{R}^m \times \mathbb{R}^{dn} : Ay + z = c\right\} \tag{7.44}$$

and the dual feasible region as

$$\mathcal{X} = \left\{x \in \mathbb{R}^{dn} : A^T x = 0; \|x_i\| \leq 1, \quad i = 1, \ldots, n\right\} \tag{7.45}$$

where x is obtained by stacking the vectors x_i. The following theorem was proved by Andersen [137]:

Theorem 7.1 *Let $(y, z) \in \mathcal{Y}$, $x \in \mathcal{X}$ be feasible and let $(y^*, z^*) \in \mathcal{Y}$, $x^* \in \mathcal{X}$ be optimal solutions to (P) and (D) respectively, then*

$$\begin{aligned} \sum_{i=1}^{n} \|z_i\| - \sum_{i=1}^{n} c_i^T x_i > 0 \\ \sum_{i=1}^{n} \|z_i^*\| - \sum_{i=1}^{n} c_i^T x_i^* = 0 \end{aligned} \tag{7.46}$$

In words: for feasible points the term $\sum_{i=1}^{n} \|z_i\|$ is an upper bound to $\sum_{i=1}^{n} c_i^T x_i$ and vice-versa. The difference $\sum_{i=1}^{n} \|z_i\| - \sum_{i=1}^{n} c_i^T x_i = \sum_{i=1}^{n} (\|z_i\| - x_i^T z_i)$ is called the *primal dual gap*; it is positive except at an optimal point where it vanishes. The primal-dual gap can be zero if and only if, for each $i = 1, \ldots, n$, either $\|z_i\|$ is zero or $x_i = z_i / \|z_i\|$. This can be expressed conveniently in a form called *complementarity condition*

$$z_i - \|z_i\| x_i = 0, \quad i = 1, \ldots, n \tag{7.47}$$

The complementarity condition encapsulates therefore the optimality of both (P) and (D). An important class of algorithms called Primal Dual Interior Point Methods (PD IPM) is based on the observation that (7.47) with the feasibility conditions (7.44) and (7.45) captures completely the optimality of both problems. The framework for a PD IPM algorithm for MSN problem works by enforcing the three following conditions (primal feasibility, dual feasibility, complementarity)

$$Ay + z = c \tag{7.48a}$$

$$A^T x = 0 \tag{7.48b}$$

$$z_i - \|z_i\| x_i = 0 \tag{7.48c}$$

The Newton Method cannot be applied in a straightforward manner to (7.48) as the complementarity condition is not differentiable for $\|z_i\| = 0$. Andersen *et. al.* [120] suggest replacing it with the so called *centering condition*

$$z_i - (\|z_i\|^2 + \beta^2)^{\frac{1}{2}} x_i = 0, \quad i = 1, \ldots, n \tag{7.49}$$

Even if at first sight the centring condition is very similar to the smooth approximations used for the TV functional in 7.1, it has different implications in this context. Particularly, it was shown in [137] the the centring condition is the complementarity condition of the following pair of smooth optimisation problems

$$\begin{aligned} (P_\beta) \quad & \min \left\{ \sum_{i=1}^{n} (\|z_i\|^2 + \beta^2)^{\frac{1}{2}} : (y, z) \in \mathcal{Y} \right\} \\ (D_\beta) \quad & \min \left\{ c^T x + \beta \sum_{i=1}^{n} (1 - \|x_i\|^2)^{\frac{1}{2}} : x \in \mathcal{X} \right\} \end{aligned} \tag{7.50}$$

and that the following theorem holds:

Theorem 7.2 *The problem P_β and D_β are a primal dual pair. Specifically, D_β has the solution $(y(\beta), z(\beta))$ and P_β has the solution $x(\beta)$, all satisfying* (7.48a), (7.48b), (7.49).

The theorem shows that introducing the perturbation β in the complementarity condition for the original pair of problems is equivalent to smoothing the norms in (P) and introducing a cost into (D). Particularly the cost function $\sum_{i=1}^{n}(1 - \|x_i\|^2)^{\frac{1}{2}}$ can be understood to keep the dual solution away from its boundary ($\|x_i\| = 1$), from which the name of centring condition for (7.49), and of interior point method for the algorithm. The concept of keeping iterates away from the boundary of feasible regions originates from interior point methods for linear programming (LP) [144]. In LP optimal points are known to lie on vertices of the feasible set; traditional algorithms, such as the simplex method, exploited this by working on the frontier of the feasible region and examining vertices to find the solution. This approach changed in the mid 80s with Karmarkar's [145] introduction of interior point methods, which work by following a smoother path inside the feasible region called a *central path* (identified by a centering condition), and possibly making larger steps at each iteration. In MSN the central path is defined by the solutions $(y(\beta), z(\beta), x(\beta))$ of P_β, D_β for $\beta > 0$, $\beta \to 0$. Using these results Andersen *et. al.* realised an efficient PD–IPM algorithm that works maintaining feasibility conditions (7.48a),

(7.48b) and applies the centering condition (7.49) with a centering parameter β which is reduced during iterations, following the central path to the optimal point. Chan *et. al.* [117] applied Andersen's PD-IPM algorithm for solving TV regularised inverse problems in image restoration. In the following we describe the application of the PD-IPM method to inversion with ℓ_1 functionals, similarly to Chan *et al.*

7.3.2 Duality for Tikhonov Regularised Inverse Problems

In inverse problems, with linear forward operators, the discretised TV regularised inverse problem, can be formulated as

$$(P) \quad \min_{\mathbf{f}} \frac{1}{2}\|A\mathbf{f} - \mathbf{d}\|^2 + \alpha \sum_{i=1}^{n} \|L_i \mathbf{f}\| \tag{7.51}$$

We will label it as the primal problem. With similar arguments as used in the preceding paragraph, the dual of (P) can be shown to be

$$(D) \quad \max_{\mathbf{x}:\|\mathbf{x}_i\|\leq 1} \min_{\mathbf{f}} \frac{1}{2}\|A\mathbf{f} - \mathbf{d}\|^2 + \alpha \mathbf{x}^T L\mathbf{f} \tag{7.52}$$

The optimisation problem

$$\min_{\mathbf{f}} \frac{1}{2}\|A\mathbf{f} - \mathbf{d}\|^2 + \alpha \mathbf{x}^T L\mathbf{f} \tag{7.53}$$

has an optimal point defined by the first order conditions

$$A^T(A\mathbf{f} - \mathbf{d}) + \alpha L^T\mathbf{x} = 0 \tag{7.54}$$

the dual problem can be written therefore as

$$(D) \quad \begin{array}{c} \max \\ \mathbf{x} : \|\mathbf{x}_i\| \leq 1 \\ A^T(A\mathbf{f} - \mathbf{d}) + \alpha L^T\mathbf{x} = 0 \end{array} \quad \frac{1}{2}\|A\mathbf{f} - \mathbf{d}\|^2 + \alpha \mathbf{x}^T L\mathbf{f} \tag{7.55}$$

The complementarity condition for (7.51) and (7.55) is set by nulling the primal dual gap

$$\frac{1}{2}\|A\mathbf{f} - \mathbf{d}\|^2 + \alpha \sum_{i=1}^{n} \|L_i \mathbf{f}\| - \frac{1}{2}\|A\mathbf{f} - \mathbf{d}\|^2 - \alpha \sum_{i=1}^{n} \mathbf{x}_i^T L_i \mathbf{f} = 0 \quad i = 1, \ldots, n \tag{7.56}$$

which with the dual feasibility $\|\mathbf{x}_i\| \leq 1$ is equivalent to requiring that

$$L_i\mathbf{f} - \|L_i\mathbf{f}\|\mathbf{x}_i = 0 \quad i = 1, \ldots, n \tag{7.57}$$

in analogy with the findings of the previous paragraph. The PD-IPM framework for the TV regularised inverse problem can thus be written as

$$\|\mathbf{x}_i\| \leq 1 \tag{7.58a}$$

$$A^T(A\mathbf{f} - \mathbf{d}) + \alpha L^T\mathbf{x} = 0 \tag{7.58b}$$

$$L_i\mathbf{f} - \|L_i\mathbf{f}\|\mathbf{x}_i = 0 \quad i = 1, \ldots, n \tag{7.58c}$$

Again it is not possible to apply the Newton Method directly to (7.58) as (7.58c) is not differentiable for $L_i\mathbf{f} = 0$. A centering condition has to be applied, obtaining a smooth pair of optimisation problems (P_β) and (D_β) and a central path parameterised by β. This is done by replacing $\|L_i\mathbf{f}\|$ by $(\|L_i\mathbf{f}\|^2 + \beta)^{\frac{1}{2}}$ in (7.58c).

7.3.3 Application to EIT

The PD-IPM algorithm in its original form [117] was developed for inverse problems with linear forward operators, as non–linear problems have not yet been addressed by research in this field at the time of this writing. The following section describes the numerical implementation for inversion in EIT tomography. The implementation is based on the results of the duality theory for inverse problems with linear forward operators. The extension of the theoretical results to non–linear inversion is out of the scope of the present work. It was possible to apply the original algorithm to the EIT inverse problem with minor modifications, and to obtain successful reconstructions. The formulation for the EIT inverse problem is

$$\begin{aligned} \boldsymbol{\sigma}_{rec} &= \operatorname{argmin} \psi(\boldsymbol{\sigma}) \\ \psi(\boldsymbol{\sigma}) &= \tfrac{1}{2}\|h(\boldsymbol{\sigma}) - \mathbf{d}\|^2 + \alpha TV(\boldsymbol{\sigma}) \end{aligned} \tag{7.59}$$

With a similar notation as used in Section 7.2, the system of non linear equations that defines the PD–IPM method for (7.59) can be written as

$$\begin{aligned} \|x_i\| &\leq 1 \\ J^T(h(\boldsymbol{\sigma}) - \mathbf{d}) + \alpha L^T\mathbf{x} &= 0 \\ L\boldsymbol{\sigma} - E\mathbf{x} &= 0 \end{aligned} \tag{7.60}$$

with $\eta_i = \sqrt{\|L_i\boldsymbol{\sigma}\|^2 + \beta}$ and $E = \operatorname{diag}(\eta_i)$, and J the Jacobian of the forward operator $h(\boldsymbol{\sigma})$. Newton's method can be applied to solve (7.60) obtaining the following system for

the updates $\delta\boldsymbol{\sigma}$ and $\delta\mathbf{x}$ of the primal and dual variables

$$\begin{bmatrix} J^TJ & \alpha L^T \\ \overline{F}L & -E \end{bmatrix} \begin{bmatrix} \delta\boldsymbol{\sigma} \\ \delta\mathbf{x} \end{bmatrix} = - \begin{bmatrix} J^T(h(\boldsymbol{\sigma}) - \mathbf{d}) + \alpha L^T\mathbf{x} \\ L\boldsymbol{\sigma} - E\mathbf{x} \end{bmatrix} \tag{7.61}$$

with

$$\overline{F} = \operatorname{diag}(1 - \frac{x_i L_i \boldsymbol{\sigma}}{\eta_i}) \tag{7.62}$$

which in turn can be solved as follows

$$[J^TJ + \alpha L^T E^{-1}\overline{F}L]\,\delta\boldsymbol{\sigma} = -[J^T(h(\boldsymbol{\sigma}) - \mathbf{d}) + \alpha L^T E^{-1} L\boldsymbol{\sigma}] \tag{7.63a}$$

$$\delta\mathbf{x} = -\mathbf{x} + E^{-1}L\boldsymbol{\sigma} + E^{-1}\overline{F}L\,\delta\boldsymbol{\sigma} \tag{7.63b}$$

Equations (7.63) can therefore be applied iteratively to solve the non linear inversion (7.59). Some care must be taken on the dual variable update, to maintain dual feasibility. A traditional line search procedure with feasibility checks is not suitable as the dual update direction is not guaranteed to be an ascent direction for the penalised dual objective function (D_β). The simplest way to compute the update is called the *scaling rule* [120] which is defined to work as follows

$$\mathbf{x}^{(k+1)} = \lambda(\mathbf{x}^{(k)} + \delta\mathbf{x}^{(k)}) \tag{7.64}$$

where

$$\lambda = \max\{\lambda :\ \lambda\|\mathbf{x}_i^{(k)} + \delta\mathbf{x}_i^{(k)}\| \leq 1, \quad i = 1, \ldots, n\} \tag{7.65}$$

An alternative way is to calculate the exact step length to the boundary, applying what is called the *steplength rule* [120]

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \min(1, \lambda)\,\delta\mathbf{x}^{(k)} \tag{7.66}$$

where

$$\lambda = \max\{\lambda :\ \|\mathbf{x}_i^{(k)} + \lambda\,\delta\mathbf{x}_i^{(k)}\| \leq 1, \quad i = 1, \ldots, n\} \tag{7.67}$$

In the context of EIT, and in tomography in general, the computation involved in calculating the exact step length to the boundary of the dual feasibility region is negligible compared to the whole algorithm iteration. It is convenient therefore to adopt the exact update, which in our experiments resulted in a better convergence. The scaling rule has the further disadvantage of always placing $\mathbf{x}$ on the boundary of the feasible region, which prevents the algorithm from following the central path. Concerning the updates on the

primal variable, the update direction $\delta\sigma$ is a descent direction for (P_β) therefore a line search procedure could be opportune. In our numerical experiments we have found that for relatively small contrasts (e.g. 3:1) the primal line search procedure is not needed, as the steps are unitary. For larger contrasts a line search on the primal variable guarantees the stability of the algorithm.

As a general comment, it is important to note that equation (7.63a) is identical to the update equation for the primal Newton's method (7.25) except for the presence of $\overline{F}$ in place of F; secondly $\overline{F} \to F$ as the solution is approached. Furthermore as the dual variable is usually started from the feasible initial point $\mathbf{x} = 0$, the matrix $\overline{F}$ is initially the identity matrix. The PD-IPM method can thus be considered as an interpolation between the Lagged diffusivity method and Newton's Method for the primal problem. Initially the algorithm behaves like the Lagged Diffusivity method, which is stable, and as the optimal point is approached the algorithm behaves like Newton's Method, which converges quadratically (at least for linear forward operators [117]) in a region around the solution. The method therefore shows stability and efficiency [120],[117],[36] (see also Section 7.5) at the small additional cost (over primal methods) of computing the updates of the dual variables.

7.3.4 Comments on the Duality for the Continuous Problem

Some insight on the greater efficiency of primal dual methods over primal methods can be gained from some observations reported by Chan and Mulet in [117] and [146]. One major difficulty associated with finding an optimal point for the TV regularised inverse problem (7.6) is that the associated Euler-Lagrange equation

$$A^T(Af - d) - \alpha \nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2 + \beta}} \right) = 0 \tag{7.68}$$

contains the highly non-linear term $\nabla \cdot \left(\frac{\nabla f}{\sqrt{|\nabla f|^2+\beta}} \right)$. The linearisation of equation (7.68) and its solution by Newton's method is therefore likely to have a very small convergence region. A better way to solve (7.68) is to note that $\frac{\nabla f}{|\nabla f|}$ is the unit normal vector to the level sets of f, and would, in general, be smoother than $\nabla \cdot \frac{\nabla f}{|\nabla f|}$. An auxiliary variable $w = \frac{\nabla f}{|\nabla f|}$ can thus be introduced and the Euler-Lagrange equation solved by means of this equivalent system of partial differential equations

$$A^T(Af - d) - \alpha \nabla \cdot w = 0 \tag{7.69a}$$

$$\nabla f - |\nabla f| w = 0 \tag{7.69b}$$

which is "more linear" than (7.68) and thus better suited to be solved via linearisation techniques. By inspection, equations (7.69) can be thought to be the continuous equivalent to the discrete system of equations (7.58). The primal dual method can therefore be expected to achieve better results thanks to a better linearisation of the problem.

7.3.5 Comments on an Alternative PD–IPM Algorithm

We would like to make some comments on an alternative primal–dual formulation found in literature [147] [128] [138] [148]. This implementation is based on the original algorithm for ℓ_1–norm minimisation proposed by Coleman [118]. The primal dual formulation of Coleman is equivalent to (7.48). While Andersen [120] suggests dealing with the non differentiability of such system of equations by using a centering condition (see (7.49)), Coleman proposes a different alternative. Observing that (7.48) is piecewise differentiable and that non–differentiability occurs only for $\|z_i\| = 0$, instead of smoothing the problem with a parameter β, he suggests keeping all the iterates inside the differentiability domains. Since the optimal point might lie on a differentiable region not connected to the one where the initial guess falls, the algorithm crosses the hypersurfaces of non–differentiability. To avoid numerical instability and to avoid falling on a non–differentiable point he uses an equivalent framework to (7.48), with a line search procedure that guarantees a "clearance" of τ from points with $\|z_i\| = 0$ (boundary of the differentiable region). The parameter τ is reduced during the progress of the algorithm, in such a way that, if for the optimal point $\|z_i\| = 0$, the numerical solution can approach such an optimal point. The algorithm can therefore be understood to be an interior point method and the "clearance" condition to act similarly to the centering condition.

As a last remark, we noted that it is not clear from the original publication [118] or from [128] [138] [148] how to find an initial guess that would satisfy such "clearance". The algorithm is in fact capable of keeping a "clearance" but must be started from a differentiable region. In EIT it is common practice to use an initial homogeneous guess. As the conductivity values are the same, $\|z_i\| = 0$ for all i; such a natural starting guess would be an infeasible point for the algorithm. Other initial guesses, like the first step of a Newton method, are not guaranteed to be feasible points either.

7.4 Comments on Methods That Use the SVD Decomposition

Two methods have been proposed that formulate the TV regularised solution using the SVD decomposition of the linearised forward operator. The aim of this chapter is to use the TV functional to regularise the non-linear least squares problem. Methods that make use of the SVD decomposition are not suitable for iterative reconstruction techniques required to solve non-linear problems, due to the high computational requirement of such decomposition. Given that the two methods are well known, we briefly comment on them.

The first method was proposed by Dobson [86] for the specific case of EIT. The forward operator was linearised and the inverse problem expressed as

$$\delta\boldsymbol{\sigma}_{rec} = \text{argmin} \int |\nabla \delta\boldsymbol{\sigma}| \tag{7.70}$$

with the following constraints

$$J\,\delta\boldsymbol{\sigma} = \delta\mathbf{d} \tag{7.71}$$

where J is the Jacobian of the forward operator. Given the ill-conditioning of the matrix J, Dobson suggests stabilising the constraints by replacing J with a matrix $\tilde{J}$ obtained from a truncated SVD decomposition of the linear operator. Considering that $\boldsymbol{\sigma}$ is discrete, the formulation becomes

$$\delta\boldsymbol{\sigma}_{rec} = \text{argmin} \sum_i |L_i\,\delta\boldsymbol{\sigma}| \tag{7.72}$$

with the following constraints

$$\tilde{J}\,\delta\boldsymbol{\sigma} = \delta\tilde{\mathbf{d}} \tag{7.73}$$

where $\tilde{\mathbf{d}}$ is the projection for the measurements on the truncated space. Dobson proposes to solve the inverse problem by smoothing (7.72) with a formula similar to (7.3), and using projected gradient methods for maintaining the feasibility of the constraints. As Dobson admits in his publication and as verified in [36], the numerical efficiency of the method is very poor for the same problems encountered by the Artificial Time Evolution method.

If the TV regularised inversion formulated by Dobson is used, the minimisation of (7.72) subject to (7.73) can be performed efficiently with PD-IPM schemes. The duality of such a problem is studied by Andersen in [119], where he then solves the problem with Newton's barrier methods. The results for the MSN problem that Andersen obtained later [120], can be applied to the primal-dual formulation in [119] and the problem can be solved with a similar framework as presented here.

A second method that employs the SVD decomposition in a TV regularised formulation is the PP-TSVD algorithm of Hansen [127]. The author considers first the linearised inverse problem, and approximates the Jacobian with its SVD truncated version of order k

$$\delta\boldsymbol{\sigma}_k = \text{argmin} \|J_k\delta\boldsymbol{\sigma} - \delta\mathbf{d}\|^2 \tag{7.74}$$

where $\delta\boldsymbol{\sigma}_k$ is the TSVD solution of (7.74). Suppose now that the matrix B holds in its columns the higher right singular vectors of the Jacobian $B = [\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots]$. Then $\delta\boldsymbol{\sigma}_k + B\mathbf{z}$, with $\mathbf{z}$ a generic vector, is a solution of (7.74), as B spans the null space of J_k. Hansen proposes therefore to formulate the TV regularised inversion as

$$\delta\boldsymbol{\sigma}_{rec} = \text{argmin} \|L(\delta\boldsymbol{\sigma}_k + B\mathbf{z})\|_1 \tag{7.75}$$

where L is the discrete TV operator, as in the preceding sections, and to solve it with the Barrodale algorithm [121]. As (7.75) is a MSN problem, it could be solved with the methods of Section 7.3.1.

7.5 Numerical experiments with the PD-IPM algorithm

Numerical Experiments Regarding the Efficiency of the Algorithm

The primal dual algorithm, which is illustrated in Figure 7.12, was tested against the same numerical experiment used for the primal methods. The PD-IPM method showed a greater efficiency and stability over the traditional methods. In particular, much more aggressive continuation procedures on the parameter β were possible, while retaining stability. The adopted continuation procedure reduces the $\beta_{reduction}$ factor at each successful step, resulting in a fast reduction rate of the parameter, as shown in Table 7.4. In these conditions the algorithm showed steady progress, even for very small βs. The value of β was limited to a lower bound of $1 \cdot 10^{-12}$ as the regularisation functional TV_β becomes a close approximation to TV, while retaining differentiability for the primal problem. In contrast to Newton's and Lagged Diffusivity algorithms, the PD-IPM algorithm is therefore capable of reducing the primal objective function's gradient for very small values of β. Similar gradient reduction rates were experienced with primal algorithms for values of β bigger than $1 \cdot 10^{-2}$, which result in a very poor approximation of the TV functional. Table 7.4 also reports in the first and second column the terms $\sum_{i=1}^{n} |L_i\boldsymbol{\sigma}|$, which is the TV of the reconstructed image, and $\sum_{i=1}^{n} \mathbf{x}_i L_i\boldsymbol{\sigma}$, which is the dual of the first term. The difference between the two terms is proportional to the primal dual gap: $\alpha\left(\sum_{i=1}^{n} |L_i\boldsymbol{\sigma}| - \sum_{i=1}^{n} \mathbf{x}_i L_i\boldsymbol{\sigma}\right)$, which is reported in the third column. The algorithm, with

PD-IPM Algorithm

find homogeneous $\boldsymbol{\sigma}_0 : \boldsymbol{\sigma}_0 = \text{argmin } \|h(\boldsymbol{\sigma}_0) - \mathrm{d}\|$;
initialise primal variable $\boldsymbol{\sigma}$ with one step of LS algorithm
initialise dual variable x to zero
set initial β
k=0;
while k<kmax
  $\delta \mathrm{d}_k = (h(\boldsymbol{\sigma}_k) - \mathrm{d})$;
  $J_k = J(\boldsymbol{\sigma}_k)$;
  $E_k = \text{diag } (\sqrt{\|L_i \boldsymbol{\sigma}_k\|^2 + \beta})$;
  $\overline{F}_k = \text{diag } (1 - \frac{x_i L_i \boldsymbol{\sigma}}{\eta_i})$;
  $\delta \boldsymbol{\sigma}_k = -[J^T J + \alpha L^T E_k^{-1} \overline{F}_k L]^{-1} \; J_k^T (h(\boldsymbol{\sigma}_k) - \mathrm{d}_k) + \alpha L^T E_k^{-1} L \boldsymbol{\sigma}_k$;
  $\delta \mathbf{x}_k = \mathbf{x}_k + E_k^{-1} L \boldsymbol{\sigma}_k + E_k^{-1} \overline{F}_k L \, \delta \boldsymbol{\sigma}_k$;
  $\lambda_{\boldsymbol{\sigma}} = \text{argmin } \psi(\boldsymbol{\sigma}_k + \lambda_{\boldsymbol{\sigma}} \, \delta \boldsymbol{\sigma}_k)$;
  $\lambda_{\mathbf{x}} = \max\{\lambda_{\mathbf{x}} : \|x_i + \lambda_{\mathbf{x}} \delta x_i\| \leq 1, \; i = 1, \ldots, n\}$;
  if a reduction of primal objective function has been achieved
    $\boldsymbol{\sigma}_{k+1} = \boldsymbol{\sigma}_k + \lambda_{\boldsymbol{\sigma}} \, \delta \boldsymbol{\sigma}_k$;
    $\mathbf{x}_{k+1} = \mathbf{x}_k + \min(1, \lambda_{\mathbf{x}}) \, \delta \mathbf{x}_k$;
    decrease β by a factor $\beta_{reduction}$
    decrease $\beta_{reduction}$
    k=k+1;
  else
    increase β
  end if else
end while

Figure 7.12: Pseudo code for the PD–IPM algorithm with continuation on β, line search on $\boldsymbol{\sigma}$ and dual steplength rule on $\mathbf{x}$.

the present test situation (no measurement noise and contrasts matching the mesh used for the inverse parameters) shows the capability of perfectly reconstructing the blocky objects of the test conductivity profile, as shown in Figure 7.14, thanks to the properties of the regularisation functional and of a very small Tikhonov factor. The simulated data does not however reflect any practical situation, where measurement noise is expected to affect the measurements and the conductivity profiles would not match the mesh used for the inversion.

Iter.	$\sum_i \lvert L_i\boldsymbol{\sigma}\rvert$	$\sum_i x_i\, L_i\boldsymbol{\sigma}$	PD-GAP/α	$\lVert h(\boldsymbol{\sigma})-\mathbf{d}\rVert$	β	$\lambda_{\boldsymbol{\sigma}}$	$\lVert\psi'(\boldsymbol{\sigma})\rVert$	Angle$(\delta\boldsymbol{\sigma}, -\psi'(\boldsymbol{\sigma}))$
1	1.034e+002	0.000e+000	1.034e+002	8.370e-001	1.000e-001	0.000e+000	1.547e+000	+83.1
2	1.574e+002	4.256e+001	1.148e+002	4.038e-001	7.000e-002	8.000e-001	6.322e-001	+83.5
3	9.455e+001	3.190e+001	6.264e+001	1.189e-001	3.920e-002	1.200e+000	1.607e-001	+85.8
4	7.042e+001	3.408e+001	3.634e+001	2.739e-002	1.756e-002	1.400e+000	2.615e-002	+88.2
5	5.372e+001	3.779e+001	1.593e+001	1.467e-002	6.294e-003	8.000e-001	1.851e-002	+89.1
6	4.971e+001	3.959e+001	1.012e+001	1.064e-002	1.805e-003	5.000e-001	1.326e-002	+89.0
7	4.405e+001	3.962e+001	4.425e+000	6.935e-003	4.139e-004	8.000e-001	8.052e-003	+87.9
8	4.122e+001	3.973e+001	1.490e+000	5.002e-004	7.596e-005	1.000e+000	6.650e-004	+89.6
9	4.058e+001	4.002e+001	5.633e-001	3.265e-005	1.115e-005	1.000e+000	4.513e-005	+89.9
10	4.020e+001	4.008e+001	1.118e-001	6.016e-006	1.310e-006	1.400e+000	1.952e-006	+89.8
11	4.018e+001	4.010e+001	7.704e-002	3.674e-006	1.230e-007	1.200e+000	5.883e-007	+89.5
12	4.013e+001	4.011e+001	1.924e-002	2.191e-006	9.248e-009	1.400e+000	7.428e-007	+89.6
13	4.012e+001	4.011e+001	8.245e-003	1.256e-006	5.561e-010	1.200e+000	9.279e-008	+88.2
14	4.011e+001	4.011e+001	1.803e-003	7.099e-007	2.675e-011	1.200e+000	6.210e-008	+88.3
15	4.011e+001	4.011e+001	4.619e-004	3.734e-007	1.029e-012	1.400e+000	3.387e-008	+83.3
16	4.011e+001	4.011e+001	1.415e-004	2.790e-007	1.000e-012	1.200e+000	1.582e-008	+65.5
17	4.011e+001	4.011e+001	1.103e-004	2.755e-007	1.000e-012	5.000e-001	9.131e-009	+71.2
18	4.011e+001	4.011e+001	1.087e-004	2.751e-007	1.000e-012	1.000e-001	8.285e-009	+72.0
19	4.011e+001	4.011e+001	1.096e-004	2.751e-007	1.000e-012	1.000e-002	8.207e-009	+71.6
20	4.011e+001	4.011e+001	1.088e-004	2.750e-007	1.000e-012	1.000e-002	8.130e-009	+72.0

Table 7.4: Main numerical parameters for the first 20 iterations of the PD–IPM algorithm with continuation on β and line search on $\boldsymbol{\sigma}$ and Steplength update on $\mathbf{x}$.

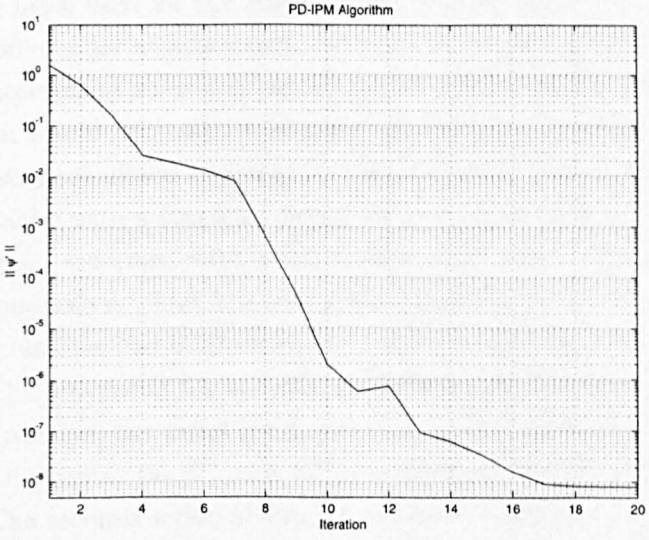


Figure 7.13: Gradient's norm during the first 20 steps of the PD-IPM algorithm.

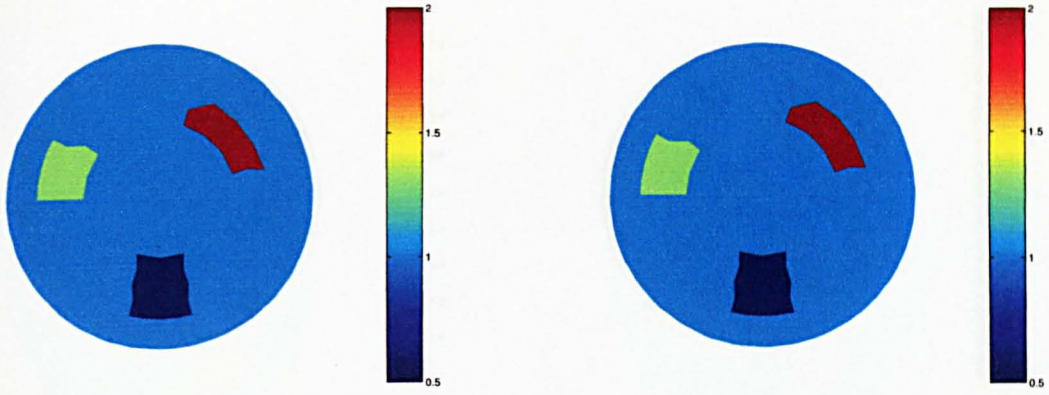


Figure 7.14: On the left, test resistivity profile. On the right, 20^{th} step of the PD–IPM reconstruction, for the noiseless test setup.

Further Numerical Experiments

The setup of the previous numerical test was intended to highlight the numerical efficiency of the different methods in relation to the differentiability issues involved with TV regularisation. Having proved first that the PD-IPM method is efficient, we conducted some further tests in order to evaluate its usefulness in practical situations. For this second group of experiments the conductivity test profiles were generated on one mesh, and then reconstructions were performed over a second mesh. It is not possible in fact to assume that the mesh used for the inverse computations will match in the real word the conductivity profile under measurement. Secondly noise was added to the measurements.

In the first experiment a test conductivity profile comprising two small circular anomalies was set up on a mesh with 6081 triangular elements, as shown in Figure 7.15. Forward measurements were calculated applying 31 trigonometric current patterns and adding to them Gaussian noise with a standard deviation of 0.1% of each measurement value. The reconstruction was performed with a coarse/fine mesh pair. The coarse mesh, used to represent the conductivity, had 510 triangular elements, while the finer mesh, used for the forward calculations had 6232 elements. On the right part of Figure 7.15 is shown a pseudo TV (PSTV) inversion (as in 6.26), computed with the same regularisation matrix as for the TV inversion, but using the ℓ_2–norm, for comparison purposes.

Figure 7.16 illustrates the 5^{th} and 10^{th} iterations of the TV regularised PD-IPM reconstruction. The reconstruction shows the characteristics of a ℓ_1 inversion, with sharp definition of the contrasts. The profiles of the two circular object of the test profiles have been superimposed to the plots. As the targets do not match the coarse mesh, it is difficult for the algorithm to estimate their shape and value correctly. The overestimation of

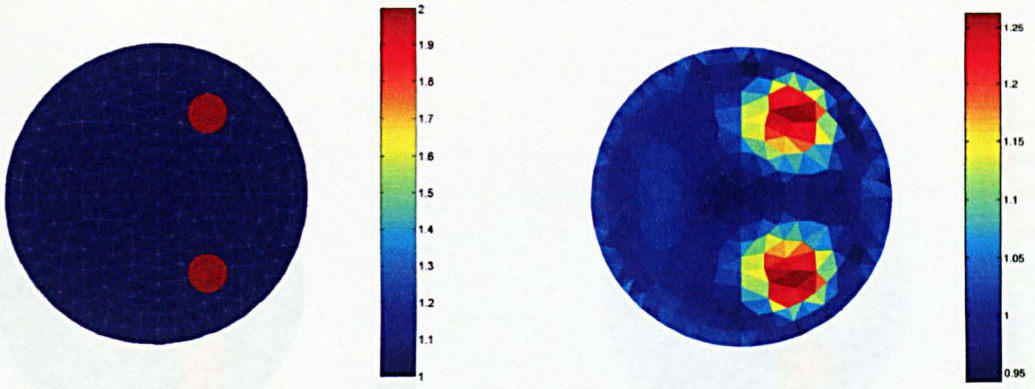


Figure 7.15: On the left, test resistivity profile defined on a 6081 elements mesh, with the discretisation used for the inversion superimposed. On the right, traditional reconstruction with ℓ_2–norm regularisation, plotted for comparison with the TV regularised inversions.

the contrasts area, due the the coarseness of the discretisation, results in a lower estimation of the conductivities. Given this limitation, the TV regularised inversion shows good localisation properties, and a relatively closer estimation of the contrasts' values when compared to the PSTV inversion.

In order to show the different properties of the TV and PSTV reconstructions, the values of $|L_i\boldsymbol{\sigma}|$, which is proportional to the conductivity difference between two neighbouring elements sharing the i–th mesh edge, have been plotted in Figure 7.17 for both methods. The values have been sorted in increasing order. The TV reconstruction presents values of $|L_i\boldsymbol{\sigma}|$ as high as 1.8, approximately 10 times as high as for the PSTV reconstructions. The plot of $|L_i\boldsymbol{\sigma}|$ for the TV reconstruction presents a slowly rising trend followed by a steep bend on the right side of the graph. By setting a threshold it is possible to identify a small number of edges as belonging to the boundary of the reconstructed targets, thanks to their large jump of conductivity. Adjusting the level of such a threshold for the PSTV reconstruction is not immediate, as for this method the graph of $|L_i\boldsymbol{\sigma}|$ has a very smooth trend. A classification of edges would be somehow more arbitrary.

Given this consideration, we decided to conduct some numerical experiments incorporating adaptive mesh refinement strategies into the reconstruction algorithm. The original PD–IPM reconstruction algorithm has been modified to perform a refinement of the mesh used for the inverse computations after the 5–th step. This arrangement allows the sharp characteristic of the TV inversion to be prominent enough at that stage of the reconstruction to allow an easy classification of the edges, as explained for the graph of Figure 7.17. The algorithm selects edges for which the value of $|L_i\boldsymbol{\sigma}|$ is greater than 0.3, and for each

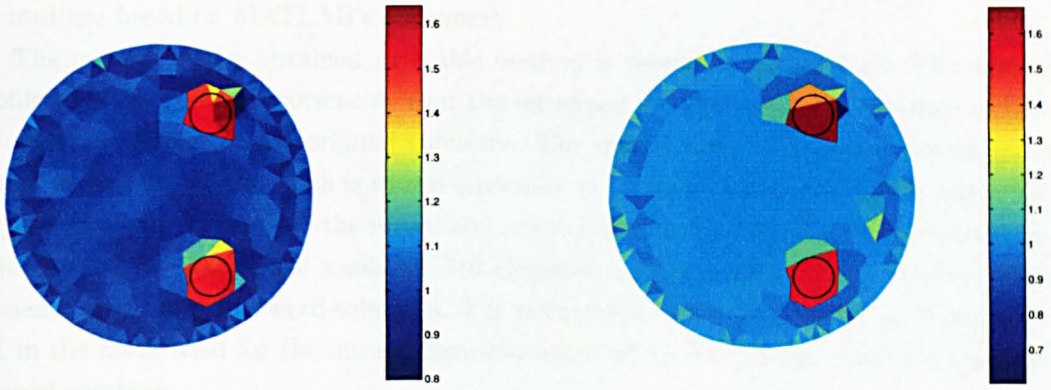


Figure 7.16: On the left, the 5^{th} step of the PD–IPM reconstruction, on the right, the 10^{th} step. The forward measurements were computed from the test profile of Figure 7.15 with trigonometric current patterns and 0.1% additive Gaussian noise. The contours of the two circular inclusions are superimposed to the graph.

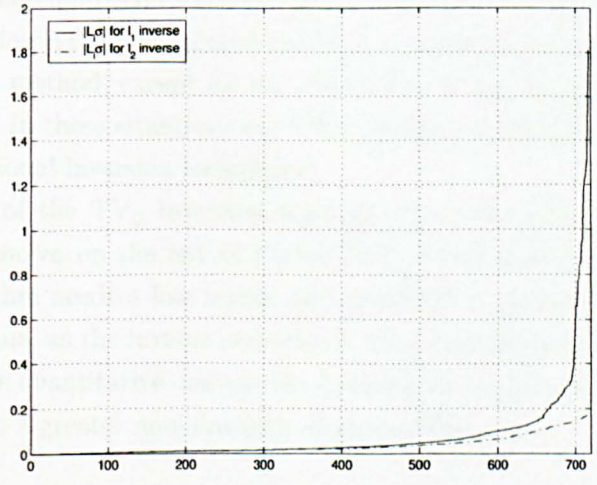


Figure 7.17: Plots of $|L_i\boldsymbol{\sigma}|$ for the TV (continuous line) and for the PSTV (dash–dotted line) reconstructions. The values have been sorted in increasing order.

edge elects the neighbouring triangles for refinement. The refinement is then carried out by routines based on MATLAB's refinemesh.

The reconstruction obtained with this method is shown in Figure 7.18. The inverse profile presents good improvements over the unrefined reconstructionism in terms of better shape matching to the original contours. The quantitative estimation showed some moderate improvement, which is shown especially in terms of better uniformity inside the targets' area. In Figure 7.19 the initial and refined meshes used for the inverse computations are shown. The initial mesh has 510 elements, and is coupled with a mesh of 6232 elements used for the forward solutions. The refinement raised the number of elements to 741 in the mesh used for the inverse computations and to 7005 in the mesh used for the forward solutions.

Experimenting with the technique, it was noted that further refinement of the mesh did not bring further improvements, and that the convergence of the algorithm slowed down. Clearly the refinement technique cannot be protracted indefinitely, as the the elements become too small, as does the sensitivity of the measurements to their conductivity values, worsening the conditioning of the problem. Dobson, in a theoretical analysis of methods for TV regularisation [114], reports similar results concerning the level of discretisation and the rate of convergence for the Lagged Diffusivity algorithm.

With mild contrasts such as the ones in the test profile of Figure 7.15, the algorithm does not require line search procedures on the primal variables $\boldsymbol{\sigma}$. The computational time involved in each iteration is therefore almost identical to traditional inversion methods such as the Gauss–Newton method, except for the calculation of the dual update $\delta\mathbf{x}$ which in practice is negligeble. In these situations the TV_β inversion is therefore very competitive with respect to traditional inversion techniques.

The performance of the TV_β inversion algorithm was also tested against the high contrast test profile shown on the left of Figure 7.20. From numerical experiments we noted that the algorithm needs a line search on the primal variables for contrasts larger that 2.5 to 1 or 3 to one, as the inverse problem shows a pronounced non linearity. With such high contrasts the quantitative discrepancy between the estimation the original value becomes larger, due to a greater non-linearity of the problem.

7.6 Discussion

The present chapter focuses on use of the TV regularisation as a technique for enhancing the reconstruction of sharp conductivity profiles. As introduced in Chapter 6, numerical difficulties arise in the solution of the inverse problem, and *ad hoc* methods are needed for

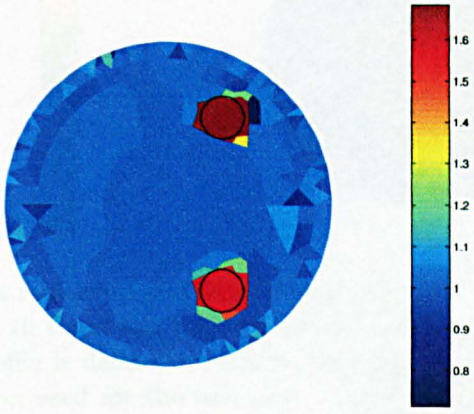


Figure 7.18: Reconstruction of the test profile of Figure 7.15 with adaptive mesh refinement. The contours of the two circular inclusions are superimposed to the graph.

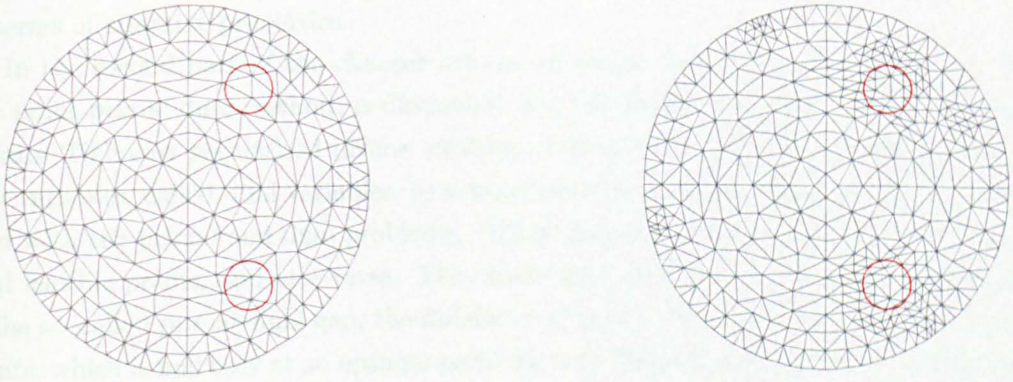


Figure 7.19: Mesh refinement during reconstruction. On the left, initial mesh used for inverse computations, the mesh has 510 elements. On the right, the refined mesh used for the inverse computations, for which the number of elements rose to 741. The contours of the two circular inclusions are superimposed to the graphs.

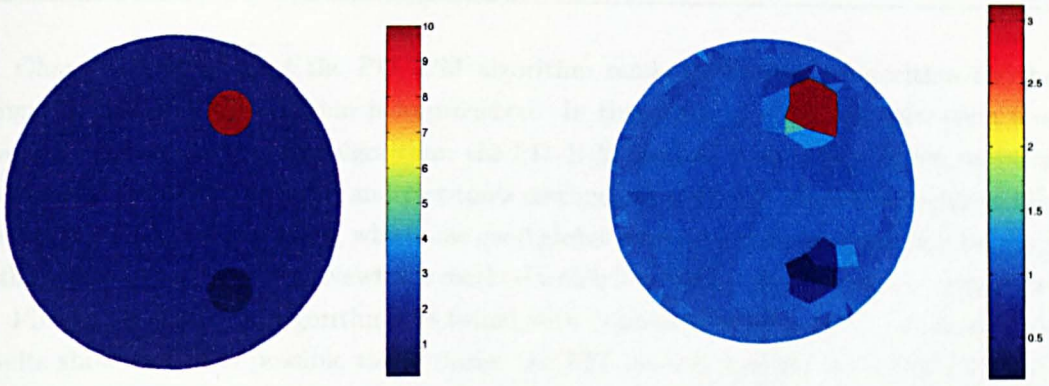


Figure 7.20: On the left, a test resistivity profile presenting two high contrast targets: the top target has a conductivity 10 times the background and the bottom one of 0.1 times the background. The test profile is defined on a 6081 elements mesh. Superimposed on the picture is the discretisation used for the inversion. On the right, the corresponding 10^{th} step of the PD–IPM reconstruction. The forward measurements were computed using trigonometric current patterns and 0.1% additive Gaussian noise.

it's solution. Such methods have been discussed and analysed by theoretical considerations and by numerical experiments. Traditional algorithms like the Steepest Descent, the Newton's method and the so–called Lagged Diffusivity Method have been shown to present numerical difficulties in conjunction with TV regularisation. The Newton's method formulation of the inverse problem gave rise to a novel interpretation of the TV regularisation, in terms of adaptive properties.

In the second part of the chapter the use of recent optimisation methods, the interior point primal dual method, is discussed. Via the duality theory, the traditional least squares Tikhonov regularised inverse problem, called primal problem, is reformulated in new variables, called dual variables, as a maximisation problem. Feasible sets are associated with the primal and dual problems. Primal feasible points are upper bounds to the dual feasible points, and vice versa. The observation of this fact allows the introduction of the so–called primal–dual gap, the difference of primal feasible points and dual feasible points, which is null only at an optimal point for both the problems. Instead of minimising the primal problem, or maximising the dual problem, the PD–IPM framework seeks the zero of the primal–dual gap.

The numerical efficiency and stability of this formulation are explained in terms of a better suitability to linearisation. The non–linear problem can therefore be solved with improved results by iterative methods that make successive linear approximations, such as Newton's method (applied to PD-IPM framework).

Chan [117] noted that the PD–IPM algorithm tends to Newton's algorithm for the primal problem as the solution is approached. In this work, a novel interpretation was given to the behaviour of the algorithm: the PD–IPM method is an interpolation between the Lagged Diffusivity method and Newton's method. Initially the algorithm behaves like the Lagged Diffusivity method, which has good global convergence properties, in proximity of the solution it behaves like Newton's method which has fast local convergence properties.

Finally the PD–IPM algorithm was tested with numerical experiments. The numerical results show that it is possible to regularise the EIT inverse problem with the TV functional, and calculate inverse solutions in a stable and efficient way. The computational effort required by the algorithm is similar to the one required by ℓ_2 norm regularisation as the computation of the dual variables update is trivial.

Apart from the considerations regarding the benefits of using the TV as a regularisation functional in conjunction with non–smooth conductivity profiles, from the analysis of the literature and from the numerical experiments we believe the PD–IPM algorithm to be the right framework to be used for ℓ_1–norm regularisation. The algorithm extends easily to the 3D case, where only the construction of the regularisation matrix L changes.

Chapter 8

Conclusions and Future Work

In this work a brief review of the theory of inverse problems, and of the EIT measurement and forward modelling techniques has been given in the first four chapters. The remaining part of the work deals with ways of reconstructing sharp conductivity profiles. As traditional ℓ_2–norm regularisation functionals produce smooth reconstructions, two novel methods to overcome the limitation are proposed. The first is the use of anisotropic Gaussian filters with the ℓ_2–norm, the second is to use the total variation functional that leads to ℓ_1–norm regularisation.

The anisotropic Gaussian filtering approach allows introduction of prior structural information into the regularisation functional, and therefore to reconstruct with better accuracy sharp conductivity profiles that match that information. On the other hand, with a careful selection of the regularisation parameter the method is capable of reconstructing conductivity profiles that violate the prior information, with slightly worse results in comparison to a similar method with no prior information. The capability of the method of describing profiles containing sharp features is explained in terms of generalised singular vectors. When the prior information is introduced, such vectors assume a different structure: some of the lowest SVs span the sharp conductivity profiles. Such SVs, not attenuated by the regularisation, contribute to forming a sharp image of the reconstructed object.

The Gaussian anisotropic smoothing has the advantage over TV regularisation techniques of not requiring particular numerical schemes for the reconstruction and to allow analysis with numerical tools such as GSVD. The complexity of the method shows in the calculation of the regularisation matrix and in the requirement of having *ad hoc* meshes. Some work is therefore needed on the side of the user to tailor the method for any specific application.

In contrast, the TV regularisation method does not require the user to tailor the reconstruction for the particular application. The method requires however particular numerical schemes, as it involves the minimisation of a non–differentiable functional. Total variation regularisation has been successfully used in EIT by Kaipio [87] with MCMC methods. Such statistical methods work by sampling posterior probability densities, therefore they are not affected by non–differentiability. The sampling of such space is a computationally intensive task; for this reason the forward problem is linearised, as this allows faster sampling.

In Chapter 7 we analysed deterministic methods for TV regularisation, with the aim of solving quickly the full non–linear problem. In the first part of the chapter traditional optimisation methods are analysed, and their inefficiency shown. In the second part of the chapter the novel application of primal–dual interior–point methods to EIT is discussed and successful reconstructions are shown. The PD-IPM method is very efficient; the computational effort required for a non–linear TV regularised inversion is similar to that of the same inversion regularised with a traditional ℓ_2–norm functional.

Concerning the analysis of the regularisation properties of TV, there are no standard tools. A novel interpretation was given to such functional, interpreting it as a non linear ℓ_2–norm functional that weights less and less large conductivity variations, allowing steep changes in the sought parameter.

Concerning future work, at present we foresee three possible extensions of the results of this thesis. The first is the extension of the TV regularisation to the 3D case. In the 2D case, for piece–wise constant elements, the TV functional is represented as

$$TV(\sigma) = \|L\sigma\|_1 \tag{8.1}$$

where L is a matrix that weights conductivity differences across neighbouring elements by the length of the shared edge. In 3D, for piece–wise constant conductivities, the expression of the functional is the same as (8.1), but the matrix L will weight differences of conductivities by the area of the shared triangular face (for FEM tetrahedral elements). Apart from this trivial change, the numerical scheme does not require any modification. Such extension is under way and numerical testing is currently performed with the help of the EIDORS 3D package [22].

A second possible extension of the present work is to experiment with anisotropic TV regularisation. With piece–wise constant representations of the conductivity, where the variations are associated with the edges, rows of the regularisation matrix could be weighted depending on the alinement of the corresponding edge to the direction of expected

changes. Otherwise piece–wise linear representations could be used, and the anisotropic TV functional be defined as

$$TV_{aniso}(\boldsymbol{\sigma}) = \int \sqrt{\xi \left(\frac{\partial \boldsymbol{\sigma}}{\partial x}\right)^2 + \eta \left(\frac{\partial \boldsymbol{\sigma}}{\partial y}\right)^2} \tag{8.2}$$

where ξ and η are the anisotropic weights and where the partial derivatives of $\boldsymbol{\sigma}$ are easily expressed as functions of the nodal values of the variable. Such an approach is interesting for the fact that traditional TV regularisation enforces only the prior information that a discontinuity might occur, but specific structural information is not enforced. By using the anisotropic TV functional, structural information would be incorporated in reconstructions, as with Gaussian anisotropic smoothing.

A third extension that we propose is to mix ℓ_2 and ℓ_1–norm regularisation functionals on the same image. There are possibly two aims in doing so, the first is to apply the two functionals to different areas of the image. This approach is equivalent to assuming that particular areas of the image are smooth and others are not. The aim is therefore to introduce prior structural information. A second approach is to apply the two functionals to the whole domain of the image, but to weight them differently, on a element by element basis, possibly in an adaptive way, as prosed by Chan in denoising applications [149]. The aim in such case is to mix the smooth properties of ℓ_2 –norm regularisation with the capability of describing sharp details of the TV regularisation in an automated way.

The inversion would be formulated as

$$\boldsymbol{\sigma}_{rec} = \text{argmin } \|h(\boldsymbol{\sigma}) - \mathbf{d}\| + \alpha \|L_2 \boldsymbol{\sigma}\| + \beta \|L_1 \boldsymbol{\sigma}\|_1 \tag{8.3}$$

where α and β are regularisation parameters and L_2 and L_1 the regularisation matrices respectively for the ℓ_2 and ℓ_1 norms.

We propose to use the PD-IPM method for these problems. In general the two matrices L_2 and L_1 will not have the same number of rows, as the two types of regularisation can be applied to subdomains of the image, or as in the case of piece wise constant representations, L_2 is associated with elements and L_1 with edges. In any case, the inversion would be formulated as

$$\boldsymbol{\sigma}_{rec} = \text{argmin } \|h(\boldsymbol{\sigma}) - \mathbf{d}\| + \alpha \|L_2 \boldsymbol{\sigma}\| + \beta \sum_i \|(L_1)_i \boldsymbol{\sigma}\| \tag{8.4}$$

which is a MSN problem; the PD-IPM method can be applied for the numerical solution of such a problem.

As a concluding remark, we believe the methods described in the thesis and the suggestions for future development to be useful in several applications. The present work deals mainly with structural prior information and with non–smooth reconstruction. In the three principal areas of application of the EIT technique, medical imaging, industrial process imaging and geophysical prospecting, it is common both for the imaged bodies to present sharp conductivity variations, and that some form of structural information is known a priori. We believe therefore that separate application of the proposed methods, or the combination of them as suggested earlier could improve current imaging techniques.

Bibliography

[1] E. Somersalo, M. Cheney, and D. Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J Appl Math*, 52:1023 1040, 1992.

[2] A. P. Calderón. On an inverse boudary problem. In *Seminar on Numerical Analysis and its Applications to Continuum Physiscs*, pages 65 73. W.H. Meyer and M.A. Raupp, Brasilian Math. Society, Rio de Janeiro, 1980.

[3] J. Sylvester and G. Uhlman. A global uniqueness theorem for an inverse boundary value problem. *Ann. Math.*, 125:153–169, 1987.

[4] A. G. Ramm. A simple proof of the uniqueness theorem in impedance thomography. *Appl. Math. Lett.*, 1:287–290, 1988.

[5] R. Kohn and M. Vogelius. Determining conductivity by boudary measurements. *Commun. Pure Appl. Math.*, 37:289–298, 1984.

[6] A. I. Nachman. Global uniqueness fro a two–dimensional inverse boundary value problem. *Annals of Math.*, 143:71–96, 1996.

[7] J. Sylvester. An anisotropic inverse baoundary value problem. *Commun. Pure Appl. Math.*, 43:201–232, 1990.

[8] W. R. B. Lionheart. Conformal uniqueness results in anisotropic electrical impedance imaging. *Inverse Problems*, 13:125 134, 1997.

[9] W. R. Breckon. *Image Reconstruction in Electrical Impedance Tomography.* PhD thesis, Oxford Brookes Polytechnic, 1990.

[10] M. Vauhkonen. *Electrical Impedance Tomography and Prior Information.* PhD thesis, Department of Applied Physics, Kuopio University, 1997.

[11] V. Kolehmainen. *Novel Approaches to Image Reconstruction in Diffusion Tomography.* PhD thesis, Department of Applied Physics - Kuopio University, 2001.

[12] J. G. Webster. *Electrical Impedance Tomography.* Adam Higler, 1990.

[13] D. C. Barber and B. H. Brown. Applied potential tomography. *J. Phys. E: Sci. Instrum.*, 17:723–733, 1984.

[14] N. D. Harris, B. H. Brown, and D. C. Barber. Continuous monitoring of lung ventilation with electrical impedance tomography. *Med. Biol. Soc.*, pages 1754 1755, 1992.

[15] I. Freichs, G. Hahn, and G. Hellige. Thoracic electrical impedance tomographic measurements during volume controlled ventilation-e ects of tidal volume and positive end-expiratory pressure. *IEEE Trans. Med. Imaging*, 18:763–773, 2000.

[16] P. W. A. Kunst, A. Vonk Noordegraaf, O. S. Hoekstra, P. E. Postmus, and P. M. J. M. de Vries. Ventilation and perfusion imaging by electrical impedance tomography: A comparison with radionuclide scanning. *Physiol. Meas.*, 19:481–490, 1998.

[17] B.M. Eyüboglu, B.H. Brown, and D.C. Barber. In vivo imaging of cardiac related impedance changes. *IEEE Eng. Med. Biol. Mag.*, 8:39–45, 1989.

[18] J.C. Newell, D. Isaacson, M. Cheney, G.J. Saulnier, D.G. Gisser, J.C. Goble R.D. Cook, and P.M. Edic. Impedance images of the chest. *Proc 14th Int. Conf. IEEE Eng. Med. Biol. Society*, pages 1752–1753, 1992.

[19] N. Kerrouche, C. N. McLeod, and W. R. B. Lionheart. Time series of EIT chest images using singular value decomposition and fourier transform. *Physiol Meas*, 22:147–157, 2001.

[20] D.S. Holder. Electrical impedance tomography of brain function. *Brain Topography*, 5:87–93, 1992.

[21] A. Gibson, R. H. Bayford, and D. S. Holder. Two-dimensional nite element modelling of neonatal head. *Physiol Meas.*, 21:45–52, 2000.

[22] N. Polydorides and W. R. B. Lionheart. A MATLAB based toolkit for three dimensional electrical impedance tomography: A contribution to the EIDORS project. *To appear in Meas. Sci. Technol.*, 2002.

[23] R. J. Sadleir, R. A. Fox, F. J. Van Kann, and Y. Attiliouzel. Estimating volumes of intra-abdominal blood using electrical impedance imaging. *In Proc. 14th Int. Conf. IEEE Eng. Med. Biol. Society*, pages 1750–1751, 1992.

[24] D. Murphy, P. Burton, R. Combs, L. Tarassenko, and P. Rolfe. Impedance imaging in the newborn. *Clin. Phys. Physiol. Meas.*, 8:A131–A140, 1987.

[25] B. H. Brown, D. C. Barber, and A. D. Seagar. Applied potential tomography: Possible clinical applications. *Clin. Phys. Physiol Meas.*, 6:109–121, 1985.

[26] R.H. Smallwood, S. Nour, Y. Mangnall, A. Smythe, and B.H. Brown. Impedance imaging and gastric motility. *In Proc. 14th Int. Conf. IEEE Eng. Med. Biol. Society*, pages 247–260, 1992.

[27] R. H. Smallwood, Y. F. Mangnall, and A.D. Leathard. Transport of gastric contents. *Physiol Meas*, 15:175–188, 1994.

[28] F. Dickin and M. Wang. Electrical resistance tomography for process tomography. *Meas. Sci. Tech.*, 7:247–260, 1996.

[29] R. Mann, F. J. Dickin, M. Wang, T. Dyakowski, R. A. Williams, R. B. Edwards, A. E. Forrest, and P. J. Holden. Application of electrical resistance tomography to interrogate mixing processes at plant scale. *Chem. Eng. Sci.*, 52:293–302, 1997.

[30] O. C. Jones, J.-T. Lin, L. Ovacik, and H. Shu. Impedance imaging relative to gas-liquid systems. *Nuclear Engineering and Design*, 141:159–176, 1993.

[31] G. Alessandrini and L. Rondi. Stable determination of a crack in a planar inhomogeneous conductor. *SIAM J. Math. Anal.*, 30:326–340, 1998.

[32] A. Friedman and M. Vogelius. Determining cracks by boundary measurements. *Indiana University Mathematics Journal*, 38:527–556, 1989.

[33] M. H. Loke. Electrical imaging surveys for environmental and engineering studies, 1997.

[34] A. Borsic, C. McLeod, W. Lionheart, and N. Kerrouche. Realistic 2d human thorax modelling for EIT. *Physiol Meas*, 22:77–83, 2001.

[35] A. Borsic, W. R. B. Lionheart, and C. N. McLeod. Generation of anisotropic-smoothness regularisation filters for EIT. *To appear in IEEE Trans. Med. Img.*, 2002.

[36] A. Borsic, C. N. McLeod, and W. R. B. Lionheart. Total variation regularisation in EIT reconstruction. *Proceedings of the 2nd World Congress on Industrial Process Tomography*, 2001.

[37] A. Tarantola. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation.* Elsevier, 1987.

[38] J. L. Melsa and D. L. Cohn. *Decision and Estimation Theory.* McGraw Hill, 1978.

[39] W. R. Gilks, R. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo Methods in Practice.* Chapmann & Hall, 1996.

[40] D. Gamerman. *Markov Chain Monte Carlo - Stocastic Simulation for Bayesian Inference.* Chapmann & Hall, 1997.

[41] E. Somersalo, J.P. Kaipio, M. Vauhkonen, and D. Baroudi. Impededance imaging and markov cahin monte carlo methods. *Proc. SPIE's 42nd Annual Meeting*, pages 175–185, 1997.

[42] K. P. Bube and R. T. Langan. Hybrid ℓ_1/ℓ_2 minimization with applications to tomography. *Geophysics*, 62:1183–1195, 1997.

[43] Per Christian Hansen. Regularisation tools. Technical report, Department of Mathematical Modelling, Technical University of Denmark - http://www.imm.dtu.dk/p̃ch, 1992.

[44] A. N. Tikhonov. Solution of incorrectly formulated problems and the regularisation method. *Dokl. Akad. Nauk SSSR*, 151:501–504, 1963.

[45] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems.* Winston & Sons, Washington D.C., 1977.

[46] V. A. Morozov. *Methods for Solving Icorrectly Posed Problems.* Springer, New York, Berlin, Heidelberg, 1984.

[47] M. Hanke and P.C. Hansen. Regularization methods for largescale problems. *Surv Math Ind*, 3:253–315, 1993.

[48] V. A. Morozov. On the solution of functional equations by the method of regularisation. *Soviet Mathematics - Doklady*, 7:414–417, 1966.

[49] G. H. Golub, M. Heath, and G. Wahba. Generalised cross validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.

[50] C. Vogel. *Computational Methods for Inverse Problems.* SIAM, 2002.

[51] M. Cheney, D. Isaacson, J. C. Newell, S. Simake, and J. Goble. NOSER: An algorithm for solving the inverse conductivity problem. *Int J Imag Sys Technol*, 2:66–75, 1990.

[52] J. Locker and P. M. Prenter. Regulariszation with differential operators. I. General theory. *J. Math. Analys. Applic.*, 74:504–529, 1980.

[53] R. H. Smallwood, A. R. Hampshire, B. H. Brown, R. A. Primhack, S. Marven, and P. Nopp. A comparison of neonatal and adult lung impedances derived from EIT images. *Physiol. Meas.*, 20:401–413, 1999.

[54] D. S. Holder, Y. Hanquan, and A. Rao. Some practical biological phantoms for calibrating multifrequency electrical impedance tomography. *Physiol. Meas.*, 17:A167–A177, 1996.

[55] P. J. Riu, J. Rosell, A. Lonzano, and R. Pallas-Areny. A broadband system for multifrequency static imaging in electrical impedance tomography. *Clin. Phys. Physiol. Meas.*, 13:A61–A65, 1992.

[56] 2nd EPSRC Engineering Network Meeting on Biomedical Applications of EIT., editor. *A Matlab Toolbox for the EIDORS Project to Reconstruct Two- and Three-Dimensional EIT Images*, London, 2000.

[57] D. G. Gisser, D. Isaacson, and J. C. Newell. Theory and performance of an adaptive tomography system. *Clin. Phys. Physiol. Meas.*, 9:A35–A41, 1988.

[58] B. Rigaud., Y. Shi, N. Chauveau, and J. P. Morucci. Experimental acquisition system for impedance tomography with active electrodes approach. *Med. Biol. Eng. Comput.*, 31:593–599, 1993.

[59] R. D. Cook, G. J. Salunier, D. G. Gisser, J. C. Goble, J. C. Newell, and D. Isaacson. ACT3: High–speed, high–precision electrical impedance tomograph. *IEEE Trans. Biomed. Imag.*, 41:713–722, 1994.

[60] Q. S. Zhu, C. N. McLeod, C. W. Denyer, F. J. Lidgey, and W. R. B. Lionheart. Development of a real-time adaptive current tomograph. *Physiol Meas*, 15:A37–A43, 1994.

[61] R. W. M. Smith, I. L. Freeston, and B. H. Brown. A real-time electrical impedance tomography system for clinical usee – design and preliminary results. *IEEE Trans. Biomed. Eng.*, 42:133–140, 1995.

[62] S. H. Mitra. *Analysis Ans Synthesis of Linear Active Networks.* John Wiley & Sons, New York, 1963.

[63] W. R. Breckon and M. K. Pidcock. Some mathematical aspects of impedance imaging. *Mathematics and Computer Science in Medical Imaging, Ed. Viergever and Todd-Pokropek, NATO ASI series F*, 39, 1988.

[64] D. C. Barber. A review of reconstruction techniques for electrical impedance tomography. *Med. Phys.*, 16:162–169, 1989.

[65] P. Matherall, D. C. Barber, R. H. Smallwood, and B. H. Brown. Three-dimensional electrical impedance tomography. *Nature*, 380:509–512, 1996.

[66] N. J. Avis and D. C. Barber. Image reconstruction using non-adjacent drive configuration. *Physiol. Meas.*, 16:A153–A160, 1994.

[67] B. M. Eyüboglu and T. C. Pilkington. Comments on distinguishability in electrical impedance tomography. *IEEE Trans. Biomed. Eng.*, 40:1328–1330, 1993.

[68] A. D. Seagar and R. H. T. Bates. Full-wave computed tomography part 4: Low frequency electric current CT. *IEE Proceedings*, 132:455–466, 1985.

[69] D. Isaacson. Distinguishability of conductivities by electric current computed tomography. *IEEE Trans. Med. Imag.*, 5:91–95, 1986.

[70] G. Gisser, D. Isaacson, and J. C. Newell. Current topics in impedance imaging. *Clin. Physiol. Meas.*, 8:38–46, 1987.

[71] G. Gisser, D. Isaacson, and J. C. Newell. Electric current computed tomography and eigenvalues. *SIAM J. Appl. Math.*, 50:1623–1634, 1990.

[72] W. R. B. Lionheart, J. Kaipio, and C. N. McLeod. Generalized optimal current patterns and electrical safety in EIT. *Physiol. Meas.*, 22:85–90, 2001.

[73] N. Polydorides and H. McCann. Electrode configurations for improved spatial resolution in electrical impedance tomography. *Meas. Sci. Technol.*, 13:1–9, 2001.

[74] J. Goble, M. Cheney, and D. Isaacson. Electrical impedance tomography in three dimensions. *Appl. Comput. Electromag. Soc. J.*, 7:128–147, 1992.

[75] P. J. Vauhkonen, M. Vauhkonen, T. Savolainen, and J. P. Kaipio. Three-dimensional electrical impedance tomography based on the complete electrode model. *IEEE Trans Biomed Eng*, 46:1150–60, 1999.

[76] K. Jerbi, W. R. B. Lionheart, P. J. Vauhkonen, and Vauhkonen M. Sensitivity matrix and reconstruction algorithm for EIT assuming axial uniformity. *Physiol. Meas.*, 21:61–66, 2000.

[77] B. H. Blott, G. J. Daniell, and S. Meeson. Nonlinear reconstruction constrained by image properties in electrical impedance tomography. *Phys. Med. Biol.*, 43:1215-1224, 1998.

[78] E. J. Woo, P. Hua, J. G. Webster, and W. J. Tompkins. A robust image reconstruction algoruthm and its prallel implementation in electrical impedance tomography. *IEEE Trans. Med. Imaging*, 12:137-146, 1993.

[79] P. Hua, J. G. Webster, and W. J. Tompkins. A regularised electrical impedance tomography reconstruction algorithm. *Clin Phys Physiol Meas*, 9:A137-141, 1988.

[80] M. Cheney D. Isaacson J. C. Newell. Electrical impedance tomography. *SIAM Rev.*, 41:85–101, 1999.

[81] T. J. Yorkey, J. G. Webster, and W. J. Tomkins. An optimal impedance tomographic reconstruction algorithm. *Proc. 8th Ann. Conf. IEEE Eng. Med. Biol. Soc.*, pages 339–342, 1986.

[82] P. Hua, E. J. Woo, and J. G. Webster. Iterative reconstruction methods using regularisation and optimal current patterns in elecrical impedance tomography. *IEEE Trans. Med. Imag.*, 10:621–628, 1991.

[83] M. Vauhkonen, D. Vadasz, J. P.Kaipio, E. Somersalo, and P. A. Karjalainen. Tikhonov regularization and prior information in electrical impedance tomography. *IEEE Trans Med Imaging*, 17:285–293, 1998.

[84] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, and E. Somersalo. Construction of nonstandard smothness priors. *Inverse Problems*, 15:713–729, 1999.

[85] L. M. Heikkinen, M. Vauhkonen, T. Savolainen, K. Leinonen, and J. P. Kaipio JP. Electrical process tomography with known internal structures and resistivities. *Inverse Probl. Eng.*, 9:431–454., 2001.

[86] D. C. Dobson and F. Santosa. An image enhancement technique for electrical impedance tomography. *Inverse Problems*, 10:317–334, 1994.

[87] J. P. Kaipio, V. Kolehmainen, E. Somersalo, and M. Vauhkonen. Statistical inversion and monte carlo sampling methods in electrical impedance tomography. *Inverse Problems*, 16:1487–1522, 2000.

[88] A. Adler and R. Guardo. Electrical impedance tomography: Regularized imaging and contrast detection. *IEEE Trans Med Imgaging*, 15:170–179, 1996.

[89] A. Adler, R. Guardo, and Y. Berthiaume. Impedance imaging of lung ventilation: Do we need to account for chest expansion? *IEEE Trans. Biomed. Eng.*, 43:414–420, 1996.

[90] F. Santosa and M. Vogelius. A backprojection algorithm for electrical impedance imaging. *SIAm J. Appl. Math.*, 50:216–243, 1990.

[91] D. C. Barber and B. H. Brown. Errors in reconstruction of resistivity images using a linear reconstruction technique. *Clin. Phys. Physiol. Meas.*, 9:A101–104, 1988.

[92] A. Seppänen, M. Vauhkonen, E. Somersalo, and J. P. Kaipio. State space models in process tomography – approximation of state noise covariance. *Inv. Prob. Eng.*, Accepted for publication 2000.

[93] A. Seppänen, M. Vauhkonen, P. J. Vauhkonen, E. Somersalo, and J. P. Kaipio. State estimation with fluid dynamical evolution models in process tomography – an application to impedance tomography. *Inv. Probl.*, 17:467–484, 2001.

[94] L. A. Geddes and L. E. Barker. The specific resistance of biological material - a compendium of data for the biomedical engineer and physiologist. *Medical and Biological Engineering*, 5:271–293, 1976.

[95] K. S. Cheng, D. Isaacson, J. C. Newell, and D. G. Gisser. Electrode model for electric current cmputer tomography. *IEEE Transations on Biomedical Engineering*, 36:918–924, 1989.

[96] K. Paulson, W. Breckon, and M. Pidcock. Electrode modelling in electrical impedance tomography. *SIAM Journal on Applied Mathematics*, 52(4):1012–1022, 1992.

[97] M. K. Pidcock, M. Kuzuoglu, and K. Leblebicioglu. Analytic and semi-analytic solutions in electrical impedance tomography: I. two-dimensional problems. *Physiol. Meas.*, 16:77–90, 1995.

[98] M. K. Pidcock, M. Kuzuoglu, and K. Leblebicioglu. Analytic and semi-analytic solutions in electrical impedance tomography: II. three-dimensional problems. *Physiol Meas*, 16:91–110, 1995.

[99] F. Kleinermann, N. J. Avis, S. K. Judah, and D. C. Barber. Three-dimensional image reconstruction for electrical impedance tomography. *Physiol. Meas.*, 17:A77 A83, 1996.

[100] D. Barber and B. Brown. *Recent Developments in Applied Potential Tomography APT.* S.L. Bacharach, M. Nijhoff, Amsterdam, 1986.

[101] C. R. Johnson. Computational and numerical methods for bioelectric field problems. *Critical Reviews in Biomedical Engineering*, 25, 1997.

[102] I. R. Mufti. A practical approach to finite-difference resistivity modelling. *Geophysics*, 43:930–942, 1978.

[103] T. Murai and Y. Kagawa. Electrical impedance computed tomography based on finite element model. *IEEE Transactions on Biomedical Engineering*, 32:177 184, 1985.

[104] E. J. Woo, P. Hua, W. J. Tompkins, and J. G. Webster. A finite element model with node renumbering for adaptive impedance imaging. *Proceedings Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, 10:277–278, 1988.

[105] K. H. Huebner. *The Finte Element Method for Engineers.* Jhon Wiley & Sons, New York, 1974.

[106] D. B. Geselowitz. An application of electrocardiographic lead theory to impedance plethysmography. *IEEE Transactions on Biomedical Engineering*, 18:38 41, 1971.

[107] M. Salazar-Palma, T. K. Sarkar, L. E. Garcia-Castillo, T. Roy, and A. Djorevic. *Iterative and Self-Adaptive Finite Elements in Electromagnetic Modelling.* Artech House, Norwood, 1998.

[108] L. A. Geddes, C. P. DaCosta, and G. Wise. The impedance of stainless-steel electrodes. *Med. Biol. Eng.*, 9:511–521, 1971.

[109] E.J. Woo, P. Hua, J.G. Webster, and W.J. Tompkins. Measuring lung resistivity using electrical impedance tomography. *IEEE Trans Biomed Eng*, 39:756–760, 1992.

[110] M. Glidewell and K.T. Ng. Anatomically constrained electrical impedance tomography for anisotropic bodies via a twostep approach. *IEEE Trans Med Imgaging*, 14:498–503, 1995.

[111] U. Baysal and B. M. Eyüboğlu. Use of a priori information in estimating tissue resitivities – a simulation study. *Phys. Med. Biol.*, 43:3589–3606, 1998.

[112] M. Vauhkonen, J.P. Kaipio, E. Somersalo, and P.A. Karjalainen. Electrical impedance tomography with basis constraints. *Inverse Problems*, 13(2):523–530, 1997.

[113] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[114] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. on Numerical Analysis*, 43:1779–1791, 1997.

[115] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation.* Birkhauser, 1984.

[116] D. C. Dobson and F. Santosa. Recovery of blocky images from noisy and blurred data. *SIAM Journal on Applied Mathematics*, 56(4):1181–1198, 1996.

[117] T. F. Chan, G. Golub, and P. Mulet. A nonlinear primal dual method for TV-based image restoration. *UCLA CAM Report 95-43.*

[118] T. F. Coleman and Y. Li. A globally and quadratically convergent affine scaling method for linear l_1 problems. *SIAM J. on Optimization*, 3(3):609–629, 1992.

[119] K. D. Andersen and E. Christiansen. A Newton barrier method for minimizing a sum of Euclidean norms subject to linear equality constraints. *Technical Report, Department of Mathematics and Computer Science, Odense University, Denmark*, 1995.

[120] K. D. Andersen, E. Christiansen, A. Conn, and M. L. Overton. An efficient primal-dual interior-point method for minimizing a sum of Euclidean norms. *SIAM J. on Scientific Computing*, 22:243–262, 2000.

[121] I. Barrodale and F. D. K. Roberts. An efficient algorithm for discrete l_1 linear approximation with linear constraints. *SIAM J. on Numerical Analysis*, 15:603–611, 1978.

[122] G. Xue and Y. Ye. An efficient algorithm for minimizing a sum of p–norms. *SIAM J. on Optimization*, 10:551–579, 2000.

[123] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10:1217–1229, 1994.

[124] C. R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Computing*, 17:227–238, 1996.

[125] T. F. Chan, H. M. Zhou, and R. H. Chan. A continuation method for total variation denoising problems. *UCLA CAM Report 95-18.*

[126] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Research Report N. 9509, CEREMADE, Université de Paris-Dauphine*, 1995.

[127] P. C. Hansen, M. Jacobsen, J. M. Rasmussen, and H. Söresen. *The PP-TSVD Algorithm for Image Restoration Problems. In P.C. Hansen, B.H. Jacobsen and K. Mosegaard. Inverse Methods II.* Springer, 1999.

[128] Y. Li and F. Santosa. A computational algorithm for minimizing total variation in image restoration. *IEEE Trans. Image Proc.*, 5:987–995, 1996.

[129] D. Luenberger. *Introduction to Linear and Nonlinear Programming.* Addison-Wesley, 1965.

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

[131] C. R. Vogel. *Nonsmooth Regularization, in H. W. Engl, A. K. Louis, and W. Rundell (Eds.), Inverse Problems in Geophysical Applications.* SIAM, 1995.

[132] U. Eckhardt. On a optimization problem related to minimal surfaces with obstacles. *AIIE Transactions*, 5:1–6, 1973.

[133] H. W. Kuhn. A note on fermat's problem. *Math. Programming*, 4:98–107, 1973.

[134] J. Rosen and G. Xue. On the convergence of miehle's algorithm for the Euclidean multifacility location problem. *Operations Research*, 40:188–191, 1992.

[135] M. R. Osborne. *Finite Algorithms in Optimization and Data Analusys.* Wiley Series in Probability and Mathematical Statistics, 1985.

[136] J. M. Wolfe. On the convergence of an algorithm for discrete ℓ_p approximation. *Numer. Math.*, 32:439–459, 1979.

[137] K. D. Andersen. An efficient Newton barrier method for minimizing a sum of Euclidean norms. *SIAM J. on Optimization*, 6:74–95, 1996.

[138] Y. Li. Piecewise differentiable minimization for ill-posed inverse problems. *Mathematical Programming*, 56:189–222, 1992.

[139] R.H. Bartels, A. R. Conn, and J.W. Sinclair. Minimisation tehniques for piecewise differentiable functions: The ℓ_1 solution to an overdetermined linear system. *SIAM J. Numer. Anal.*, 15:224–240, 1978.

[140] E. Seneta and W. L. Steinger. A new lad curve–fitting algorithm: Slightly overdetermined equations systems in ℓ_1. *Discrete Applied Mathematics*, 7:79–91, 1984.

[141] J. Eyster, J. White, and W. Wierwille. On solving multifacility location problems using a hyperboloid approximation procedure. *AIIE Trans.*, 5:1–6, 1973.

[142] M. Overton. A quadratically convergent method for minimizing a sum of Euclidean norms. *Math. programming*, 27:34–63, 1983.

[143] R. T. Rockefellar. *Convex Analysis.* Princeton University Press, Princeton University, 1970.

[144] S. J. Wright. *Primal Dual Interiror Point Methods.* SIAM, Philadelphia, 1997.

[145] N. Karmarkar. A new polynomiacl–time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[146] T. Chan and P. Mulet. Iterative methods for total variation restoration. *UCLA CAM Report 96-38*, 1996.

[147] Y. Li and F. Santosa. An affine scaling algorithm for minimizing total variation in image enhancement. *Manuscript*, 1994.

[148] T. F. Coleman, Y. Li, and A. Mariano. Segmentation of pulmonary nodule images using total variation minimization. *Technical Report*, Cornell University, USA, 1998.

[149] T. Chan, A. Marquina, and P. Mulet. High order total variation-based image restoration. *SIAM J. on Scientific Computing*, 22(2):503–516, 2000.

INSTITUTE OF PHYSICS PUBLISHING PHYSIOLOGICAL MEASUREMENT

Physiol. Meas. 22 (2001) 77–83 www.iop.org/Journals/pm PII: S0967-3334(01)19913-2

Realistic 2D human thorax modelling for EIT

Andrea Borsic[1,3], Chris McLeod[1], William Lionheart[2], Nacer Kerrouche[1]

[1] School Of Engineering, Oxford Brookes University, Oxford OX3 OBP, UK
[2] Department of Mathematics, UMIST, PO Box 88, Manchester M60 1QD, UK

E-mail: aborsic@brookes.ac.uk

Received 22 September 2000

Abstract

Electrical impedance tomography is a technique that permits estimation of resistivity within a subject by reconstructing from boundary measurements. Due to the ill conditioning of the problem, images are greatly affected by boundary errors. Reconstruction algorithms account for boundary shape and electrode position by using a forward solver. In this paper is presented the realization of a 2D realistic thoracic model as a better approach compared to circular meshing. Several meshes with different discretizations are compared in terms of accuracy with a complete electrode model forward solver. By comparing the different discretizations it is possible to choose an appropriate mesh density for the 2D EIT problem.

Keywords: Electrical impedance tomography (EIT), finite element modelling (FEM), mesh generation, complete electrode model, full electrode model, thoracic imaging, lungs

1. Introduction

Electrical impedance tomography (EIT) is an imaging technique that estimates the resistivity distribution inside the subject. Current patterns are applied to the subject through several electrodes and surface potentials are measured. The acquired boundary data are processed to reconstruct images. Reconstruction algorithms account for boundary shape and electrode position by making use of a forward model. Usually the electrode positions and boundary shape are not accurately known. Due to the ill posedness of the problem this greatly affects the image quality; strong regularization techniques are needed to obtain the convergence of the reconstruction process. Therefore, ideally, the actual shape should be measured for each experiment.

For practical reasons 2D EIT thoracic images have often been reconstructed over circular meshes. In our group progress in modelling has been represented by the use of a non-circular mesh (McLeod *et al* 1997). The contour was calculated by averaging shape measurements over several subjects. The mean contour was then expressed in polar coordinates—four components of the Fourier series of the radius against the circular angle were considered, obtaining a mean smoothed thoracic shape.

[3] Addressee for correspondence.

0967-3334/01/010077+07$30.00

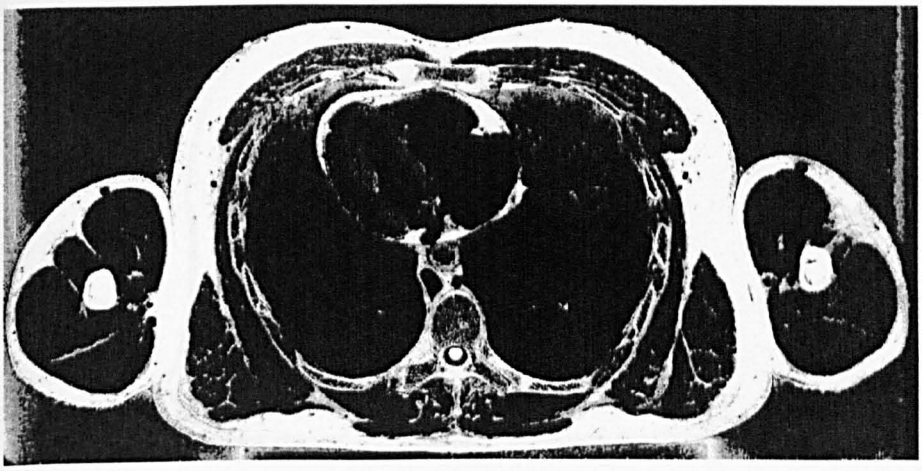


Figure 1. Cryostatic image of the human thorax.

As an alternative attempt to closely describe the human thoracic shape, the possibility of meshing a cryostatic image from the National Library of Medicine's Visible Human Project (figure 1) was considered.

In this paper are presented results obtained with this approach. Five meshes resembling the human thorax were generated. Different degrees of discretization were set for splines describing electrodes and organs in order to assess the accuracy of the forward solutions. A complete electrode model forward solver was employed in the simulations.

2. Mesh generation

In order to generate realistic meshes, pixel coordinates of the exterior boundary and main organs were imported to the INRIA's[4] EMC2[5] mesh generator. By connecting imported points, splines describing heart, left lung, right lung and spine were defined (figure 2).

The measurement protocol carried out with the OXBACT III data acquisition system (Zhu *et al* 1994, Denyer 1996) consists of stimulating the body with 32 electrodes carrying current and in measuring potentials over another 32 interleaved passive electrodes. For this purpose 64 electrodes have been defined on the exterior boundary. The ratio of the covered to free surface for these electrodes is 10:1.

Having set up all the necessary information, five different meshes were generated, with different degrees of coarseness.

The meshes are based on unstructured triangular elements, and as reported in table 1 the number of elements varies from 5390 to 35 314. The first three columns of the table report the number of nodes with which the splines defining electrodes, lungs, heart and spine were discretized. The last two columns report the number of nodes and triangles that resulted for each mesh corresponding to a particular discretization.

The discretization of inner splines was set in such a way to produce coarser meshes toward the centre, where the intensity of the electric field is weaker. The discretization of the electrodes was set to deliver a finer mesh toward the boundary where the intensity of the electric field is higher due to the presence of discontinuities (figure 3).

[4] Institut National de la Recherche en Informatique et en Automatique.

[5] Editor of meshes and contours in two dimensions.

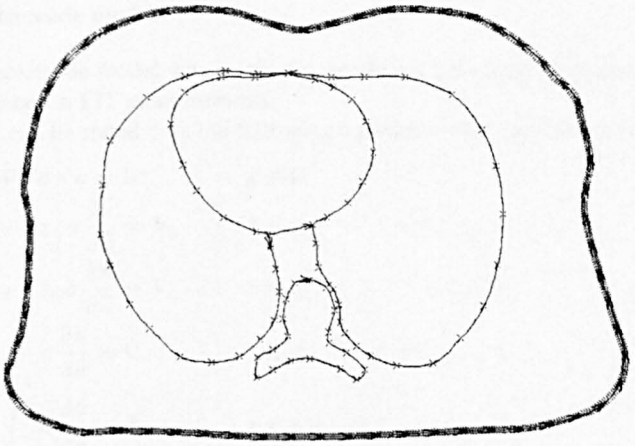
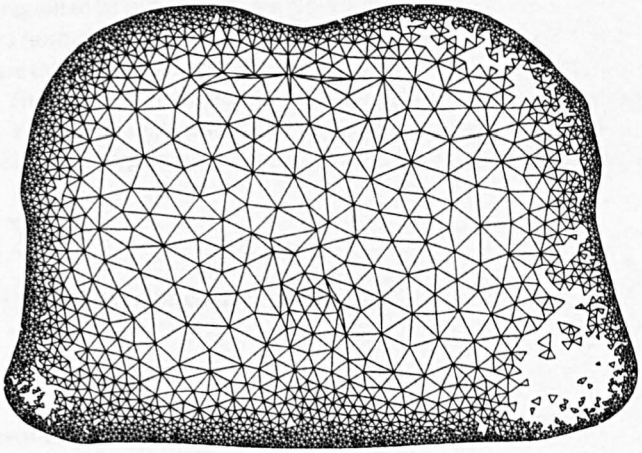


Figure 2. Splines defining organs and boundary.



3. Complete electrode model

The complete electrode model was originally developed for taking in account the effects of contact impedances in EIT measurements.

The model can be stated with the following equations, when applied to our system:

$$\nabla \cdot \sigma \nabla u = 0 \qquad x \in \Omega \tag{1}$$

$$u + z_{v_l} \sigma \frac{\partial u}{\partial \vec{n}} = V_{v_l} \qquad x \in e_{v_l} \qquad l = 1, \ldots, L_v \tag{2}$$

$$u + z_{c_l} \sigma \frac{\partial u}{\partial \vec{n}} = V_{c_l} \qquad x \in e_{c_l} \qquad l = 1, \ldots, L_c \tag{3}$$

$$\int_{e_{v_l}} \sigma \frac{\partial u}{\partial \vec{n}} = 0 \qquad x \in e_{v_l} \qquad l = 1, \ldots, L_v \tag{4}$$

$$\int_{e_{c_l}} \sigma \frac{\partial u}{\partial \vec{n}} = I_l \qquad x \in e_{c_l} \qquad l = 1, \ldots, L_c \tag{5}$$

$$\sigma \frac{\partial u}{\partial \vec{n}} = 0 \qquad x \in \partial\Omega \qquad x \notin e_{v_l} \qquad x \notin e_{c_l} \tag{6}$$

where Ω is the region to be imaged, u is the electric potential, σ is the conductivity distribution, $\vec{n}$ is the outward normal to the boundary $\partial\Omega$, e_{v_l} and e_{c_l} are the current and voltage electrodes and z_{v_l} and z_{c_l} are their respective contact impedances and L_v and L_c are the respective numbers of electrodes. The vector I represent the currents injected at the current electrodes and the vectors V_v and V_c the potentials assumed by voltage and current electrodes.

Since the electric charge must be conserved the following rule must be obeyed by the I vector:

$$\sum_l I_l = 0$$

and a reference potential can be chosen imposing that

$$\sum_l V_{v_l} = 0.$$

4. Finite element implementation

The FEM method allows discretization of the continuous problem by subdividing the region to be imaged into elements of simple shape, which are usually triangles or quadrangles for the 2D case. The electric potential is expressed as a linear combination of nodal basis functions. The nodal basis functions have support limited to each element—they assume the value of 1 at node i and 0 on each other node.

The approximated potential can then be expressed as

$$u' = \sum_{i=1}^{M} N_i \phi_i$$

where ϕ_i are the nodal basis functions, N_i are the nodal values assumed by the approximate solution and M is the number of nodes.

Wishing to calculate the nodal values N_i, the continuous problem is then turned into a linear system by applying the Galerkin method:

$$\int\int_{\Omega} \phi_i (\nabla \cdot \sigma \nabla u') = 0. \tag{7}$$

Using Green's theorem this can be rewritten as

$$\iint_{\Omega} \sigma \nabla \phi_i \cdot \nabla u' - \int_{\partial\Omega} \phi_i \sigma \frac{\partial u'}{\partial \vec{n}} = 0. \tag{8}$$

The boundary integral of equation (8) enables us to take into account the conditions expressed by equation (2) and (3). Additionally when we express the conditions stated by (4)–(6) we may write the system of linear equations

$$YN = C \tag{9}$$

where N is the vector of nodal values and Y and C are the system matrix and its right-hand side.

The Y matrix and the C vector have the following form:

$$Y = \begin{bmatrix} Y_y & Y_v & Y_c \\ Y_v^{\mathrm{T}} & D_v & 0 \\ Y_c^{\mathrm{T}} & 0 & D_c \end{bmatrix} \qquad C = \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}$$

where:

$$Y_y(i,j) = \iint_{\Omega} \sigma \nabla \phi_i \cdot \nabla \phi_j + \sum_{l=1}^{L_v} \frac{1}{z_{v_l}} \int_{e_{v_l}} \phi_i \phi_j + \sum_{l=1}^{L_c} \frac{1}{z_{c_l}} \int_{e_{c_l}} \phi_i \phi_j \qquad i,j = 1,\ldots,M$$

$$Y_v(i,j) = -\frac{1}{z_{v_j}} \int_{e_{v_j}} \phi_i \qquad i = 1,\ldots,M \qquad j = 1,\ldots,L_v$$

$$Y_c(i,j) = -\frac{1}{z_{c_j}} \int_{e_{c_j}} \phi_i \qquad i = 1,\ldots,M \qquad j = 1,\ldots,L_c$$

$$D_v(i,i) = \frac{\text{length}(e_{v_i})}{z_{v_i}} \qquad i = 1,\ldots,L_v$$

$$D_v(i,j) = 0 \qquad i \neq j$$

$$D_c(i,i) = \frac{\text{length}(e_{c_i})}{z_{c_i}} \qquad i = 1,\ldots,L_c$$

$$D_c(i,j) = 0 \qquad i \neq j.$$

The finite element version of the complete electrode model was implemented in MATLAB 5. The implementation was tested against the forward solver of the Kuopio MATLAB 2D EIDORS[6] toolbox (Vauhkonen *et al* 2001). Forward solutions computed with the two different forward solvers on the same mesh agreed with a relative error of the order of 1×10^{-8}.

Studies have also been conducted in order to validate the model against tank measurements. In 'Electrode models for electric current computed tomography' (Cheng *et al* 1989) the authors present a validation of the complete electrode model against measurements conducted on a tank. In the study characteristic impedances (p 921, table 1) measured on the tank are compared to simulations obtained with a boundary fourier series implementation of the electrode model. In order to validate the finite element method implementation of the model, the tank described in the paper by Cheng *et al* was meshed, and similar results reproduced; the measured characteristic impedances and the FEM simulated ones agreed to within 2% down to 0.3% depending on the bath concentration.

The realized forward solver has been used to compare different thoracic meshes as explained in the next section.

[6] EIT and diffuse optical tomography reconstruction software.

Table 2. Error and computational time of the forward solutions.

	Percentage error	Computation time (s)
Mesh_1	7.216%	23.2
Mesh_2	0.522%	75.9
Mesh_3	0.248%	146.6
Mesh_4	0.092%	350.2
Mesh_5	Not applicable	806.0

5. Forward solutions

Due to the discontinuities of the electric field near the electrodes, fine discretization is needed there. Discretization of electrodes greatly affects the total number of elements in the mesh. The reason for this is that the number of electrodes is high, and setting for each electrode a fine discretization will result in a very fine mesh near the boundary. By contrast the inner region of the mesh does not require the same accuracy, so usually the inner splines are discretized with a lower number of nodes, giving a locally coarser mesh. For these reasons the greatest number of elements is located near the boundary and it is highly influenced by the discretization of the electrodes.

Forward solutions[7] with the various meshes were performed and considering the solution over the mesh with the highest number of elements as the most accurate one, errors were computed with the following formula:

$$\text{Err} = \frac{\|V(\text{Mesh\_}i) - V(\text{Mesh\_5})\|}{\|V(\text{Mesh\_5})\|} \times 100. \tag{10}$$

Table 2 reports the errors and the forward solution computational time.

As is possible to see from table 1 the electrodes were discretized in seven, 11 or 15 nodes, giving six, ten and 14 edges underneath each electrode. For each of the two finer discretizations of electrodes two different levels of definition were adopted for the inner splines describing organs. This enables estimation of the sensitivity of the error to the coarseness of the inner portion of the mesh. Examining tables 1 and 2 together it is possible to make several observations. Mesh_1 was generated with a reasonably low electrode discretization resulting in a high error. Leaving unchanged the inner portion of the mesh, the discretization of electrodes was raised to 11 nodes (Mesh_2). This significantly reduced the error of the forward solution. To evaluate the sensitivity of the error to the coarseness of the central portion of the mesh, the splines describing heart, lungs and spine were discretized with a much higher number of nodes (Mesh_3). In this case the error decreased by a factor of two with respect to the previous mesh. Therefore the inaccuracy delivered by Mesh_1 should be attributed primarily to an insufficient discretization of electrodes. In Mesh_4 the number of nodes per electrode was increased to 15. The gain in the error is about a factor of two, not as significant as from Mesh_1 to Mesh_2 since the approximate solution is converging. It is impractical to raise the number of nodes per electrode any further, and having reached anyway a trend of convergence in the error, a reference mesh (Mesh_5) was generated with a finer central portion.

6. Choosing a mesh

In order to choose a mesh to use with a reconstruction algorithm one must consider that the goal of the forward solver is to provide an accuracy comparable to the one delivered by the data

[7] The solutions were computed with a homogeneous body with a resistivity of 5 Ω m and with contact impedances of 0.09 Ω m^2.

acquisition system. Any accuracy less than this will spoil the accuracy of the data acquisition system, since its measurements would be compared to an inaccurate set of data. Otherwise employing a much more accurate forward solver would result in a waste of computation time, since the overall accuracy will be set by the acquired data. Therefore the mesh should be chosen considering the accuracy delivered by the DAS in use.

In our case the data acquisition system used is the OXBACT III, which is designed to deliver an accuracy of 12 bits. Considering this accuracy as the one we can expect from the real measurements is too optimistic since many other factors other then the DAS accuracy affect the measurements themselves. In particular we can consider that even if the mesh shape is realistic, it will not match exactly the patient's body shape, the electrodes in practice will be not equispaced as assumed during the mesh generation, and the problem is 3D rather then 2D as assumed by the forward solver. For these reasons, most probably a sensible mesh to use in a forward solver is Mesh_2, which delivers a much higher accuracy then Mesh_1, and does not require the high number of elements of Mesh_4 to have an evident increase in accuracy.

7. Conclusions

Usually EIT chest images have being reconstructed on circular meshes. The ill posedness of the problem make this approach viable for differential imaging, where errors due to boundary shape and electrodes position are lessened by using previous acquired voltage measurements as a reference and reconstructing differences with respect to them. Wishing to reconstruct absolute images, a realistic model could deliver a better forward solution, even though this effort on its own is not sufficient to resolve all the challenges of the problem.

In this paper is presented the realization of a realistic human thorax mesh, along with a complete electrode model implementation of the forward problem as a step to increasing the accuracy delivered by the forward solver.

This approach, supported by the implementation of anatomical prior information in the reconstruction algorithm could enhance the quality of absolute imaging.

References

Cheng K S, Isaacson D, Newell J C and Gisser D G 1989 Electrode models for electric current computed tomography *IEEE Trans. Biomed. Eng.* **36** 918–23

Denyer C W 1996 Electronics for real-time and three-dimensional electrical impedance tomographs *PhD Thesis* Oxford Brookes University School of Engineering

McLeod C N, Shi Y, Lionheart B, Lottiaux J L and Paulson K 1997 Improvements in EIT image quality from the use of adaptive currents and accurate body shape *Proc. World Cong. on Medical Physics* vol 35, part 1

Paulson K, Breckon W and Pidcock M 1992 Electrode modelling in electrical impedance tomography *SIAM J. Appl. Math.* **52** 1012–22

Somersalo E, Cheney M and Isaacson D 1992 Existence and uniqueness for electrode models for electric current computed tomography *SIAM J. Appl. Math.* **52** 1023–40

Vauhkonen M 1997 Electrical impedance tomography and prior information *PhD Thesis* University of Kuopio Department of Applied Physics

Vauhkonen M, Lionheart W R B, Heikkinen L M, Vauhkonen P J and Kaipio J P 2001 A MATLAB package for the EIDORS project to reconstruct two-dimensional EIT images *Physiol. Meas.* **22** 107–11

Zhu Q S, McLeod C N, Denyer C W, Lidgey F J and Lionheart W R B 1994 Development of a real-time adaptive current tomograph *Physiol. Meas.* **15** (suppl) A37–A43

Generation of anisotropic-smoothness regularization filters for EIT

Andrea Borsic, William R.B. Lionheart, Christopher N. McLeod

Abstract— **In the inverse conductivity problem, as in any ill-posed inverse problem, regularization techniques are necessary in order to stabilize inversion. A common way to implement regularization in EIT is to use Tikhonov regularization. The inverse problem is formulated as a minimization of two terms: the mismatch of the measurements against the model, and the regularization functional. Most commonly, differential operators are used as regularization functionals, leading to smooth solutions. Whenever the imaged region presents discontinuities in the conductivity distribution, such as inter organ boundaries, the smoothness prior is not consistent with the actual situation. In these cases the reconstruction is enhanced by relaxing the smoothness constraints in the direction normal to the discontinuity. In this paper we derive a method for generating Gaussian anisotropic regularization filters. The filters are generated on the basis of the prior structural information, allowing a better reconstruction of conductivity profiles matching these priors. When incorporating prior information into a reconstruction algorithm, the risk is of biasing the inverse solutions toward the assumed distributions. Simulations show that, with a careful selection of the regularization parameters, the reconstruction algorithm is still able to detect conductivities patterns that violate the prior information. A GSVD analysis of the effects of the anisotropic filters on regularization is presented in the last sections of the paper.**

Keywords— **Electrical Impedance Tomography, Prior Information, Anisotropic Smoothing, Regularization, GSVD.**

I. Introduction

ELECTRICAL Impedance Tomography involves reconstructing the conductivity of an object from current and voltage measurements on the boundary. Usually electrodes are applied to the object and known currents are passed through some of them; the resulting voltages are measured on the electrodes. Reconstruction algorithms make use of a forward model: simulated measurements are matched to the real ones by acting on the discretized conductivity of the model; the reconstructed conductivity is the solution of the least-squares problem

$$\mathbf{s}_{\text{rec}} = \operatorname{argmin} \|\mathrm{h}(\mathbf{s}) - \mathbf{z}\|^2 \qquad (1)$$

where $\mathbf{z}$ is the vector of measured voltages, $\mathbf{s}$ is the discrete conductivity, h is the nonlinear forward operator from model space to measurements space.

The reconstruction problem is ill-conditioned and regularization techniques are necessary in order to stabilize the process. Commonly (1) is solved using the Tikhonov regularization, formulating the reconstruction as

$$\mathbf{s}_{\text{rec}} = \operatorname{argmin} \{\|\mathrm{h}(\mathbf{s}) - \mathbf{z}\|^2 + \alpha^2 \mathrm{F}(\mathbf{s})\} \qquad (2)$$

A. Borsic and C. N. McLeod are with the School Of Engineering, Oxford Brookes University, Oxford, UK.

W. R. B. Lionheart is with the Department of Mathematics, UMIST, Manchester, UK.

where $F(\mathbf{s}) \geq 0$ is the regularization functional, and α the regularization parameter. The conductivity being discrete, the regularization functional is usually expressed as $F(\mathbf{s}) = \|L\mathbf{s}\|^2$, where L is the regularization matrix. The reconstruction is therefore formulated as

$$\mathbf{s}_{\text{rec}} = \operatorname{argmin} \{\|\mathrm{h}(\mathbf{s}) - \mathbf{z}\|^2 + \alpha^2 \|\mathrm{Ls}\|^2\} \qquad (3)$$

The role of the regularization functional is to penalize solutions that according to some prior knowledge are unlikely. A classic choice for the matrix L is suggested by the identity matrix, a similar choice is made by the NOSER algorithm [1] which uses a positive diagonal matrix. Matrices that approximate first and second order differential operators have also been commonly used in EIT over the last decade [2]. All these regularization methods achieve the stability of the inversion by penalizing sudden variations in the conductivity; the cost is that the reconstruction is rendered incapable of describing sharp variations.

There are of course situations of practical interest where the actual conductivity presents sudden variations. In the literature two different approaches have been proposed for dealing with those situations. The first approach seeks solutions with the least total variation [3] or uses the total variation as a regularization functional [4], allowing the presence of step changes in the reconstructed images. This approach is particularly suitable for reconstructing piece-wise constant conductivities, although it might lead to staircase effects in the presence of conductivity gradients.

The second approach is that of incorporating structural prior information into the reconstruction process to estimate the unknown conductivity more closely [5],[6],[7],[8],[9]. By this means it is possible to enhance the sharpness of the images when the prior information is matched by the actual experiment, as shown by Kaipio *et. al.* [10].

The approach followed in [10] is to use anisotropic smoothness constraints in the regularization. The smoothness constraints are relaxed along the direction of the expected changes, allowing faster transitions in this direction while preserving the necessary smoothness tangentially. The study proposes a way of constructing the filters for a piece-wise linear 2D FEM forward model. The conductivity being linear, it is possible to express the first partial derivatives of each element as functions of the nodal values of $\mathbf{s}$. The reconstruction then penalizes them differently, according to the local direction of the expected changes.

In this paper we present a method for constructing Gaussian anisotropic filters for a piece-wise constant 2D FEM

forward model. As the spatial derivatives are not readily available, the filters express their directional properties by weighting appropriately the conductivity values of neighboring elements. The approach is introduced first by analyzing the traditional forms of regularization, then extending Gaussian isotropic filtering to the anisotropic case.

II. Standard Regularization

An insight on the ill-conditioning of the reconstruction problem and on traditional regularization techniques is gained by linearizing (1) about a point $\mathbf{s}_0$

$$\delta\mathbf{s}_{\mathrm{rec}} = \operatorname{argmin} \|J\,\delta\mathbf{s} - (\mathbf{z} - \mathrm{h}(\mathbf{s}_0))\|^2 \tag{4}$$

and by analyzing the singular value decomposition (SVD) of the Jacobian J of the forward operator. The matrix $J \in \mathbb{R}^{m\times n}$ (m is the number of observations, and n is the number of elements in which the conductivity has been discretized) is decomposed as

$$J = U\,\Sigma\,V^{\mathrm{T}} \tag{5}$$

where U and V are $m \times m$ and $n \times n$ orthonormal matrices and $\Sigma = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ with $\lambda_1 \geq ... \geq \lambda_n \geq 0$. Given the decomposition, the mapping $J\,\delta\mathbf{s}$ of a change in the conductivity $\delta\mathbf{s}$ can be expressed as

$$J\,\delta\mathbf{s} = \sum_{i=1}^{n} \mathbf{u}_i\,\lambda_i\,(\mathbf{v}_i^{\mathrm{T}}\delta\mathbf{s}) \tag{6}$$

The ill-posedness of the problem is shown by the singular values λ_i rapidly decaying to zero: there are certain changes in $\mathbf{s}$ parallel to the higher singular vectors $\mathbf{v}_i$ that are strongly attenuated and would become unobservable in the presence of noise.

When inverting the conductivity, assuming that J has full rank, solutions of (4) are expressed as

$$\delta\mathbf{s}_{\mathrm{rec}} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{(\mathbf{u}_{\mathrm{i}}^{\mathrm{T}}\delta\mathbf{z})}{\lambda_{\mathrm{i}}}\,\mathbf{v}_{\mathrm{i}} \tag{7}$$

Terms with higher values of i, for which observations are poor, are strongly amplified by the factor $1/\lambda_i$, leading to a solution that is dominated by noise.

The aim of regularization is therefore to dampen the contribution of the higher singular vectors to the reconstructed image. This is achieved explicitly by the truncated SVD technique stopping the summation of (7) at an index $k < n$ for which the observations are not exceeded by the noise level.

A similar filtering effect is achieved implicitly, at a minor computational expense, by using the identity matrix in the regularization term or by using the NOSER algorithm, for which the regularized solutions of (4) are expressed as

$$\delta\mathbf{s}_{\mathrm{rec}} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\lambda_{\mathrm{i}}}{\lambda_{\mathrm{i}}^2 + \alpha^2\,\mathrm{l}_{\mathrm{i}}}(\mathbf{u}_{\mathrm{i}}^{\mathrm{T}}\delta\mathbf{z})\mathbf{v}_{\mathrm{i}} \tag{8}$$

where l_i are the diagonal elements of L. The regularization prevents the multiplicative term from diverging for $\lambda_i \to 0$.

Similarly, the effect of regularization matrices that are discrete approximations of first order and second order differential operators can be understood in terms of SVD decomposition by considering the fact that higher singular vectors $\mathbf{v}_i$ tend to be more and more oscillatory, therefore constraining the image to be smooth rejects their contribution.

All the standard techniques therefore stabilize the inversion by limiting the unreliable contribution of the higher singular vectors to the reconstructed images.

III. Statistical Interpretation of L

A similar regularization approach is suggested by the statistical interpretation of the reconstruction [11]. In this case the discretized conductivity $\mathbf{s}$ is assumed to be a random variable and the observations $\mathbf{z}$ to be contaminated by the random noise $\mathbf{n}$

$$\mathbf{z} = \mathrm{h}(\mathbf{s}) + \mathbf{n} \tag{9}$$

Assuming also that $\mathbf{s}$ is a Gaussian variable with mean $\mathbf{m_s}$ and covariance matrix $C_{\mathbf{s}}$ and that $\mathbf{n}$ has zero mean and covariance matrix $C_{\mathbf{n}}$, the posterior probability density for the vector $\mathbf{s}$ given the observation $\mathbf{z}$ is

$$p_{\mathbf{s}}(\mathbf{s} \mid \mathbf{z}) = \frac{p_{\mathbf{z}}(\mathbf{z} \mid \mathbf{s})p_{\mathbf{s}}(\mathbf{s})}{p_{\mathbf{z}}(\mathbf{z})} \tag{10}$$

Where subscripts indicate which probability density p is used. Additionally recalling (9)

$$p_{\mathbf{z}}(\mathbf{z} \mid \mathbf{s}) = p_{\mathbf{n}}(\mathbf{z} - \mathrm{h}(\mathbf{s}) \mid \mathbf{s}) \tag{11}$$

If $\mathbf{s}$ and $\mathbf{n}$ are statistically independent

$$p_{\mathbf{z}}(\mathbf{z} \mid \mathbf{s}) = p_{\mathbf{n}}(\mathbf{z} - \mathrm{h}(\mathbf{s})) \tag{12}$$

Given that $p_{\mathbf{z}}(\mathbf{z})$ is a constant since $\mathbf{z}$ is fixed, and using (10) and (12)

$$p_{\mathbf{s}}(\mathbf{s} \mid \mathbf{z}) \propto \exp\left\{-\frac{1}{2}(\mathbf{z} - \mathrm{h}(\mathbf{s}))^T C_{\mathbf{n}}^{-1}(\mathbf{z} - \mathrm{h}(\mathbf{s}))\right\} \cdot \exp\left\{-\frac{1}{2}(\mathbf{s} - \mathbf{m_s})^T C_{\mathbf{s}}^{-1}(\mathbf{s} - \mathbf{m_s})\right\} \tag{13}$$

If $\mathbf{s}$ is now estimated with the maximum a posteriori (MAP) criterion, the maximum of $p_{\mathbf{s}}(\mathbf{s} \mid \mathbf{z})$ is sought, which is equivalent to minimizing the argument of the exponentials in (13)

$$\mathbf{s}_{MAP} = \operatorname{argmin} \{(\mathbf{z} - \mathrm{h}(\mathbf{s}))^T C_{\mathbf{n}}^{-1}(\mathbf{z} - \mathrm{h}(\mathbf{s})) + (\mathbf{s} - \mathbf{m_s})^T C_{\mathbf{s}}^{-1}(\mathbf{s} - \mathbf{m_s})\} \tag{14}$$

Formally, the Tikhonov solution of the inverse problem (3) is identical to the MAP approach when $C_{\mathbf{n}}^{-1} = I$, $L^{\mathrm{T}}L \propto C_{\mathbf{s}}^{-1}$ and $\mathbf{m_s}$ is a prior estimate of $\mathbf{s}$. This connects the choice of L to the statistical information on $\mathbf{s}$.

The properties of $C_{\mathbf{s}}$ reflect the prior knowledge of the system under measurement. The diagonal elements of the

matrix represent the variance of each element in the image, the off-diagonal elements of the matrix are a function of the correlation r between elements of the image $C_{\mathbf{s}_{ij}} = r\sqrt{C_{\mathbf{s}_{ii}} C_{\mathbf{s}_{jj}}}$. It is therefore possible to envisage the possibility of constructing $C_{\mathbf{s}}^{-1}$ on the basis of the structural knowledge of the system under measurement, and of using a regularization matrix L such that $L^T L \propto C_{\mathbf{s}}^{-1}$.

IV. Isotropic Gaussian Smoothing

The latter approach, $L^T L \propto C_{\mathbf{s}}^{-1}$, has been considered by Adler *et al.* [12], even though the authors finally assumed only the implicit information that a limited resolution is achieved by EIT. They therefore used a covariance matrix that allowed some correlation between neighbouring image elements. In this case the inversion of the covariance matrix is ill-posed and thus numerically unstable. Observing that the covariance matrix has the structure of a low-pass filter, the solution they proposed is to use a Gaussian high-pass filter to represent its inverse. For the 2D case, the frequency response of the filter with a spatial frequency of ω_o is

$$G(\tau, \upsilon) = 1 - \exp(-\omega_o(\tau^2 + \upsilon^2)) \tag{15}$$

with the following convolution kernel

$$g(x, y) = \delta(x, y) - \frac{\pi}{\omega_o^2} \exp(-\frac{\pi^2}{\omega_o^2}(x^2 + y^2)) \tag{16}$$

A regularization matrix L that approximates the filter can be found by expressing the filtered value $\hat{\sigma}$ of the continuous conductivity σ at the mid-point coordinates x_i, y_i of the i-*th* element of the image

$$\hat{\sigma}(x_i, y_i) = g(x, y) * \sigma(x, y)|_{x=x_i, y=y_i} \tag{17}$$

and assigning the value at the mid-point to the discrete conductivity

$$\hat{\mathbf{s}}_i = \int g(x_i - x, y_i - y)\sigma(x, y)\, \mathrm{d}x\, \mathrm{d}y \tag{18}$$

The integration can be carried out on the single elements E_j, after bringing the piece-wise constant conductivity out of the integration

$$\hat{\mathbf{s}}_i = \sum_j \mathbf{s}_j \int_{E_j} g(x_i - x, y_i - y)\, \mathrm{d}x\, \mathrm{d}y \tag{19}$$

the filtered conductivity can be expressed as $\hat{\mathbf{s}} = L\,\mathbf{s}$ with the following definition for L

$$L_{i\,j} = \int_{E_j} g(x_i - x, y_i - y)\, \mathrm{d}x\, \mathrm{d}y \tag{20}$$

V. Anisotropic Gaussian Smoothing

The statistical interpretation of the matrix L sets the basis for incorporation of prior structural information into regularization. In their study, Adler *et. al.*, however did not assume this information and used an isotropic Gaussian filter in order to mimic $C_{\mathbf{s}}^{-1}$. Nevertheless their proposed method inspired the use of an anisotropic Gaussian kernel that we adopt in this study for exploiting the anatomical priors, which are, in the medical imaging context, the equivalent of the structural information.

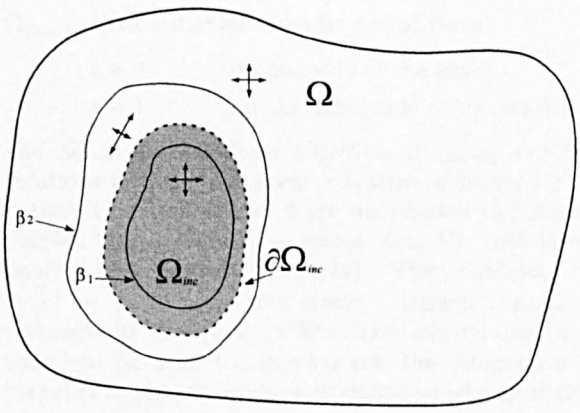


Fig. 1. An object Ω has an inclusion Ω_{inc} with a different conductivity (gray area). The contour of the inclusion is bounded by β_1 and β_2 and is expected to follow approximately the bounding profiles. In order to exploit the structural information, the smoothing filter weights (represented by crosses) should be anisotropic in the region bounded by the two β curves, allowing for a faster variation of the conductivity in the direction of the expected changes.

We now consider the problem for which we propose a method by examining the situation presented in Figure 1. Assume that a body Ω has an inclusion Ω_{inc} that presents a different conductivity from the surrounding body. The shape of the inclusion is not precisely known, but it is bounded by β_2 and β_1. The boundary $\partial\,\Omega_{\text{inc}}$ of the object is assumed to follow approximately the bounding curves. An organ expanding and contracting during its physiological activity could be an example of the depicted situation.

The region enclosed between the two bounding curves β, which we label Ω_{change}, is where the expected conductivity discontinuity will occur. The objective is to relax the smoothing constraints in the region Ω_{change} along the direction normal to the line of changes, that has yet to be defined. The tangential smoothing can be maintained, expecting the changes to be orthogonal to that direction.

We therefore propose to use the anisotropic Gaussian kernel that is obtained by transforming (16)

$$g(n, t) = \delta(n, t) - \frac{\pi}{\omega_n \omega_t} \exp(-\pi^2(\frac{n^2}{\omega_n^2} + \frac{t^2}{\omega_t^2})) \tag{21}$$

where n and t are the tangential and normal directions of the expected changes in the conductivity. Such a kernel would separate control of the smoothing along n and t by varying the parameters ω_n and ω_t. From a statistical point of view this is equivalent to assuming that the image elements are less correlated in the direction of the expected changes.

Theoretically the use of such a filter is straightforward; in practice the problem is to find a way of calculating the

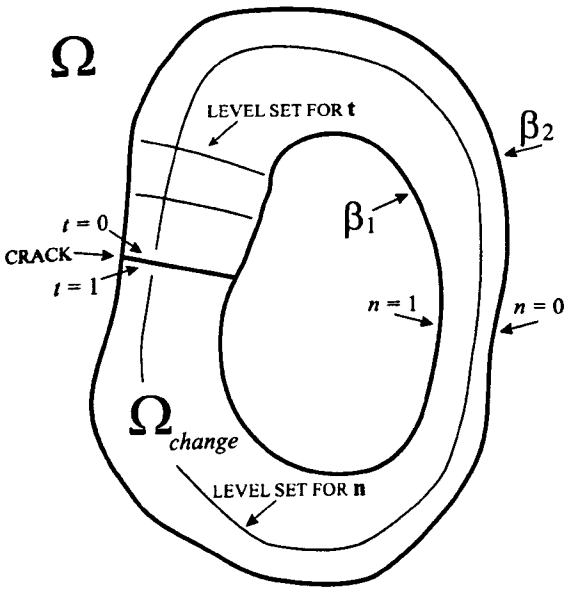


Fig. 2. Calculation of the normal and tangential coordinates solving the Laplace equation over the domain Ω_{change} with opportune boundary conditions. The solutions n and t form a system of harmonic coordinates (n,t) that maps the cracked domain Ω_{change} to a rectangle.

system of coordinates (n,t) given the geometry of the domains. The normal and tangential directions need to be defined somehow in the region Ω_{change}, in order to make the use of the Gaussian anisotropic kernel possible.

VI. Finding Normal and Tangential Coordinates

The problem of finding the normal and tangential coordinates can be solved naturally with a system of harmonic coordinates. The intention is to find a system of coordinates where

$$\hat{n} \perp \beta_1 \qquad \text{on } \beta_1 \tag{22}$$

$$\hat{n} \perp \beta_2 \qquad \text{on } \beta_2 \tag{23}$$

$$\hat{t} \perp \hat{n} \qquad \text{in } \Omega_{\text{change}} \tag{24}$$

The solution we propose is to find $(\hat{n},\hat{t})$ satisfying (22-24) by solving a PDE.

Solving $\nabla^2 n = 0$ over Ω_{change} with the following Dirichlet boundary conditions

$$n = 1 \qquad \text{on } \beta_1 \tag{25}$$

$$n = 0 \qquad \text{on } \beta_2$$

gives a solution n defined over all the domain Ω_{change} that can be regarded as one of the coordinates of the system (n,t). Points with $n = 0$ will lie on β_2, points with $n = 1$ will lie on β_1. The vector $\hat{n} \equiv \nabla n$ is orthogonal to β_1 for $n = 1$ and orthogonal to β_2 for $n = 0$. Therefore $\hat{n}$ satisfies (22) and (23). Now assume a crack in the domain Ω_{change} as illustrated in Figure 2, and that $\nabla^2 t = 0$ is solved over Ω_{change} with Dirichlet boundary conditions

$$t = 0 \qquad \text{on one side of the crack.} \tag{26}$$

$$t = 1 \qquad \text{on the other side of the crack.}$$

and Neumann conditions $(\partial t/\partial \hat{n} = 0)$ on β_1 and β_2. The solutions n and t will form a system of harmonic coordinates. The level sets of t are distributed radially on the domain Ω_{change}, giving a vector $\hat{t} \equiv \nabla t$ that is orthogonal to $\hat{n}$ on it, satisfying (24). The coordinate change $(x,y) \rightarrow (n,t)$ maps the cracked domain Ω_{change} to the rectangle $[0\ 1] \times [0\ 1]$ in $\mathbb{R}^2$. The coordinates (n,t) can therefore be used to carrying out the integration of the kernel (21) producing an anisotropic weighting matrix L.

VII. Comments on Calculation of L

The idea of solving a PDE for the calculation of (n,t) was based on the opportunity of relying mostly on the forward solver for the task. In electrical impedance tomography the forward algorithm solves

$$\nabla \cdot \sigma \nabla u = 0 \tag{27}$$

where u is the electric potential. It is sufficient to set σ to a constant value in (27) order to enable the forward solver to solve the Laplace equation. In order to apply Dirichlet boundary conditions however, the solver needs to be slightly modified; in fact Neumann or constrained Robin conditions are applied in EIT depending on which electrode model is implemented [13].

An algorithm for the calculation of (n,t) has been developed in the MATLAB environment and integrated with the routines currently in use for the forward solution. The algorithm presupposes that a FEM mesh matching the external and internal boundaries of the object to be imaged has been produced. The user describes the structural information by selecting on the screen the nodes on β_1, β_2. Given the two spatial frequencies ω_n and ω_t the algorithm computes the transformation of the domain Ω_{change} and calculates the anisotropic regularization matrix L.

For each finite element E_i belonging to the domain Ω_{change} the corresponding row of L_i is calculated by integrating (21), while for the rest of the elements the isotropic kernel (16) is used.

VIII. Simulations

In this section we compare reconstructions using prior information, in the form of anisotropic smoothing, and traditional reconstructions using Gaussian isotropic smoothing. A simple experiment involving a square inclusion embedded in a round object was set up. The experiment has no physiological meaning but its validity is general. The numerical simulations allowed us to compare the two methods, and to present a generalized singular value decomposition (GSVD) analysis of the effect of prior information on the regularization.

A. Setup of the experiment

The numerical experiment was set up as illustrated in Figure 3. An outer round object with a diameter of 30

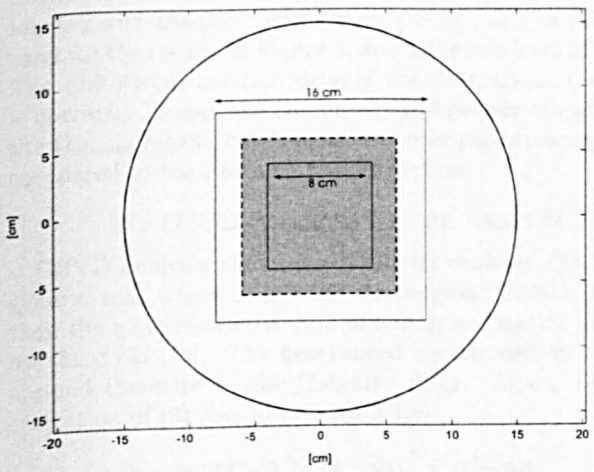


Fig. 3. Setup of the simulated experiment: round object with a square inclusion, definition of the domains and dimensions in cm.

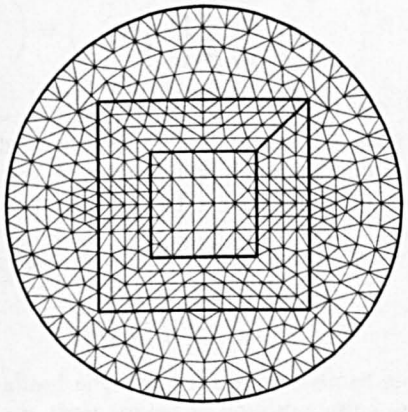


Fig. 4. Generated mesh, boundaries and crack are shown in thick line.

cm is expected to contain a square shaped inclusion with a different conductivity. The dimensions of the inclusion can vary, the side of the anomaly can range from 8 cm to 16 cm.

A mesh of 798 triangular elements, shown in Figure 4, was used for the inverse computations. The mesh matches the internal boundaries β_1 and β_2 in order to allow the calculation of (n, t) with the PDE method. The forward solver uses a finer mesh for calculating the electric potential, attaining higher accuracy in the forward solutions. The finer mesh was obtained from an adaptive refinement of the first mesh, resulting in 6346 elements. The mesh was used also for the generation of the test conductivity profiles of the simulations. The inclusions of the test profiles were generated not coincide with the discretization of the coarse mesh, as this would be not representative of a real situation. The disposition of the electrodes resembles the setup of the OXBACT III [14] adaptive current tomograph: 32 current electrodes, each one capable of injecting a current, are interleaved with another 32 electrodes, used to measure the electric potential, resulting in a total of 64 electrodes equispaced around the object.

B. Reconstructions

A first test conductivity profile matching the expected structure was used to compare the reconstructions with and without prior information. In the simulations, the conductivity of the surrounding circular object was set to 1 $\Omega^{-1}\,\mathrm{m}^{-1}$, and an 11 cm square inclusion, with conductivity 0.7 $\Omega^{-1}\,\mathrm{m}^{-1}$, was generated, as shown in Figure 5a.

For the reconstructions, trigonometric current patterns were used and the resulting measurements were calculated with a forward solver implementing the complete electrode model. A Gaussian white noise with zero mean and a standard deviation of 0.1% of the voltage range was added to the measurements to simulate instrument noise [15]. The reconstruction (3) is solved iteratively, starting from a homogeneous conductivity that best fits the data. The first four steps of the algorithm were performed. In Figure 5b and 5c respectively, the isotropic and anisotropic reconstructed conductivities are shown. Several values of the ratio ω_n/ω_t were used; a high ratio produces an image conforming strongly with the prior information, while a low ratio is close to conventional Gaussian smoothing. An anisotropy of 2.5 was found to be a good compromise for these experiments. As expected the isotropic solution smooths the lateral discontinuities of the square inclusion and rounds off the corners. The anisotropic solution, incorporating the prior information, estimates the square shape of the detected object more accurately.

In Figure 6a and 6b cross sections of the true and reconstructed conductivities are shown for an easier quantitative comparison. Figure 6a illustrates a cross section along the x axis. The anisotropic solution follows the sharp transitions and settles closer to the correct value in the centre of the object than the isotropic one. The effect is more evident in Figure 6b. where the cross section is cut on the 45° diagonal, crossing the corners of the inclusion. The discontinuity of the corners is even more difficult to describe for the isotropic smoothing, resulting in a larger difference in the two reconstructions.

C. Incorrect priors

The risk in using prior information in the reconstruction process is to bias the solution toward the assumed distribution, and to miss inclusions that do not respect the prior assumptions.

In this section are presented reconstructions comparing the two methods in the case where the priors are incorrect. Figure 7a shows the test conductivity to be reconstructed. The inclusion is a rectangular object with the lateral edges orthogonal to the direction assumed for the conductivity changes in Ω_{change}. The object therefore violates the prior information assumed by the regularization. The conductivity values for this test are again 1 $\Omega^{-1}\,\mathrm{m}^{-1}$ for the embedding object and 0.7 $\Omega^{-1}\,\mathrm{m}^{-1}$ for the inclusion. In Figure 7b and 7c, the isotropic and anisotropic

reconstructions are shown. Both reconstructions were performed with the same parameters ($\alpha, \omega_n, \omega_t$) as the ones used for the results of Figure 5, and the same level of noise. The anisotropic solution detects the anomaly and locates it correctly. Some fake responses are however triggered in the Ω_{change} region, resulting in a poorer performance when compared to the isotropic reconstruction.

IX. GSVD Analysis of the Results

GSVD analysis allows study of the problem (3) in the generic case where L is a full rectangular matrix, rather then the more restrictive case of a diagonal matrix allowed by the SVD [16]. The generalised decomposition can be applied therefore to the Gaussian filter. Again, the linearization of (3) should be considered

$$\delta \mathbf{s}_{\text{rec}} = \operatorname{argmin} \|\mathrm{J}\,\delta\mathbf{s} - \delta\mathbf{z}\|^2 + \alpha^2 \|\mathrm{L}\delta\mathbf{s}\|^2 \qquad (28)$$

The matrices $J \in \mathbb{R}^{m\times n}$ and $L \in \mathbb{R}^{p\times n}$ (where p is the number of "regularisation constraints", $p = n$ in the Gaussian case) are then decomposed as

$$\begin{pmatrix} J \\ L \end{pmatrix} = \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & I_{n-p} \\ M & 0 \end{pmatrix} X^{-1} \qquad (29)$$

where $r = \text{rank(L)}$, $U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{p\times p}$ and $X \in \mathbf{R}^{n\times n}$. The matrices U and V are orthonormal and X non singular, Λ and M are diagonal matrices $\in \mathbb{R}^{p\times p}$, $\mathbf{\Lambda} = \text{diag}(\lambda_1, ..., \lambda_p)$ and $M = \text{diag}(\mu_1, ..., \mu_p)$ with

$$1 \geq \lambda_1 \geq ... \geq \lambda_r \geq 0\,, \quad 0 \leq \mu_1 \leq ... \leq \mu_r \leq 1 \qquad (30)$$

and

$$\lambda_i^2 + \mu_i^2 = 1\,, \quad i = 1, ..., p \qquad (31)$$

The generalized singular values are defined as $\gamma_i = \lambda_i/\mu_i$ for $i = 1 \ldots r$, they appear in non-increasing order:

$$\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_{r-1} \geq \gamma_r > 0 \qquad (32)$$

Using this decomposition the regularized solutions to the linearized problem (28) can be expressed as

$$\delta \mathbf{s}_{\text{rec}} = \sum_{\mathrm{i}=1}^{\mathrm{r}} \frac{\gamma_\mathrm{i}^2}{\gamma_\mathrm{i}^2 + \alpha^2} \frac{(\mathbf{u}_\mathrm{i}^{\mathrm{T}} \delta\mathbf{z})}{\lambda_\mathrm{i}} \mathbf{x}_\mathrm{i} + \sum_{\mathrm{i}=\mathrm{r}+1}^{\mathrm{n}} (\mathbf{u}_\mathrm{i}^{\mathrm{T}} \delta\mathbf{z})\, \mathbf{x}_\mathrm{i} \qquad (33)$$

Similar conclusions can be drawn from (33) to those from the SVD analysis. The generalized singular values λ_i show the ill-posedness of the problem by rapidly decaying to zero for increasing i. The term $\gamma_i^2/(\gamma_i^2 + \alpha^2)$ should therefore tend to zero with sufficient rapidity to prevent the first term at the right hand side of (33) from diverging. The matrix pair (J, L) is therefore understood to regularize the inversion in a similar fashion to (J, I): by damping the content of the singular vectors (SV) for which the corresponding singular values are too small.

When anisotropic filtering is adopted, we have found interesting changes in the structure that the singular vectors assume. Typically the singular vectors associated with bigger singular values are smooth; they have components only in the lower part of the spectrum of spatial frequencies and do not present discontinuities. The GSVD analysis of the pair (J, L) revealed that with the anisotropic Gaussian filter the structure of the singular vectors changes.

Figure 8a illustrates the second SV for the isotropic Gaussian filter, which is a vertical gradient as usually happens. The corresponding vector for the anisotropic Gaussian filter is shown if Figure 8b. The central region of the image presents a sudden variation corresponding to the area delimited by β_2. Figure 9 shows the $14^{th}, 42^{nd}, 108^{th}$ and 131^{st} singular vectors for the anisotropic case. The structural prior information appears to having been embedded in the SVs, modifying the smooth structure that one would expect otherwise. Some of the vectors $\mathbf{x}_i$ decouple the information relative to the inclusion from the background by spanning only particular regions of the image. The non-smooth properties of the lower SVs allow the reconstruction to describe more easily inclusions matching the prior information, as the spectral analysis explains in more detail.

X. Spectral Analysis

As with the SVD decomposition, it is possible to express a given conductivity distribution $\mathbf{s}$ as a linear combination of the singular vectors $\mathbf{x}_i$

$$\mathbf{s} = \sum_{i=1}^{n} w_i\, \mathbf{x}_i \qquad (34)$$

The w_i are said to be the spectral coefficients of $\mathbf{s}$. The actual calculation of the coefficients differs from the standard SVD case in that $\mathbf{x}_i$ are not orthogonal but just linearly independent. The first r spectral coefficients w_i can be calculated as

$$w_i = \langle \mathbf{s}, \mathbf{x}_i \rangle_L \quad i = 1, ..., r \qquad (35)$$

since the vectors x_i are $L^{\mathrm{T}}L$ orthogonal for $i = 1, ..., r$, and $\langle \mathbf{x}_i, \mathbf{x}_j \rangle_L = 0$ for $i = 1 \ldots r$, $j = r+1 \ldots n$.

The remaining coefficients can be calculated as

$$w_i = \langle \mathbf{s}, \mathbf{x}_i \rangle_J \quad i = r+1, ..., n \qquad (36)$$

since the last $n - r$ columns of X are $J^{\mathrm{T}}J$ orthogonal and $\langle \mathbf{x}_i, \mathbf{x}_j \rangle_J = 0$ for $i = 1 \ldots r$, $j = r+1 \ldots n$.

Traditionally, the lower singular vectors are smooth and the higher ones are oscillatory. Smooth conductivity distributions will therefore have components in the lower part of the spectrum and viceversa. In this sense the decomposition (34) is similar to a Fourier analysis of $\mathbf{s}$.

As introduced earlier, the regularization dampens the contribution of the higher singular vectors in the image. For a given conductivity distribution $\mathbf{s}$, the higher the spectral content, the more the reconstructed image will suffer from the dampening. Thus, an image with sharp changes, having significant components in the higher part of the spectrum, will be heavily smoothed. However, when anisotropic filtering is used, the lower singular vectors are able to span certain sharp transitions, shifting the corresponding components downwards in the spectrum. These

TABLE I
NUMBER OF SINGULAR VECTORS REQUIRED TO DESCRIBE A TEST CONDUCTIVITY WITHIN 0.04 ERROR.

	iso x_i	aniso x_i
s_1	273	11
s_2	38	64

conductivity patterns will therefore survive the smoothing effect of the regularization.

The spectral shifting property was verified by expanding the test conductivity of Figure 5a in its spectral coefficients, using the singular vectors derived both from the isotropic and anisotropic filters. The singular vectors span the conductivity space of the mesh used for the inverse calculations. The test conductivity, defined on a finer mesh, was therefore projected onto this mesh. The distribution of the spectral energy was compared by truncating the two expansions at an index k for which

$$\frac{\| \sum_{i=1}^{k} w_i \mathbf{x}_i - \mathbf{s} \|}{\|\mathbf{s}\|} < 0.04 \qquad (37)$$

The results are shown in the first row of Table 1 labelled $\mathbf{s}_1$. The lower 273 singular vectors are needed to describe the conductivity $\mathbf{s}_1$ within 4% error, if isotropic filtering is used. The same conductivity is spanned within the same error by the lower 11 singular vectors in the anisotropic case. The non-smooth nature of the SVs deriving from the anisotropic filter can thus describe a sharp conductivity profile with a smaller spectral content, resulting in less smoothing from the regularization as shown by the reconstructions. The conductivity $\mathbf{s}_1$ matches the prior information used for setting up the matrix L and for calculating the corresponding singular vectors. For this reason the sharp changes in $\mathbf{s}_1$ are spanned by the lower SVs.

The same experiment was repeated for the test conductivity of Figure 7a; the results are reported in the second row of Table 1, labelled $\mathbf{s}_2$. In this case the situation differs, 64 singular vectors are needed to span the image within the 4% error in the anisotropic case, versus 38 for the isotropic regularization. The conductivity $\mathbf{s}_2$ does not match the prior information, it isn't spanned by the lower singular vectors, resulting in a slightly worse spectral distribution.

XI. CONCLUSIONS

In this paper a method is proposed for dealing with the reconstruction of conductivity images with sharp variations, as encountered in situations of medical interest. The approach is to enhance the reconstructions by incorporating the prior structural information into the regularization. This is achieved by using Gaussian anisotropic filters, which relax the smoothness in the direction of the expected changes. The effectiveness of the approach has been positively compared to the use of isotropic filters by means of simulations. The simulations show that the sharpness and quantitative estimation of the conductivity are enhanced when the experiment matches the prior information. On the other hand, with a careful selection of the regularization parameters, the algorithm was able to detect a contrast that violated the prior assumptions.

In our view the paper addresses three different aspects of anisotropic regularisation: the introduction of anisotropic Gaussian smoothing, a method for the calculation of the regularization filter and a GSVD analysis of the regularized problem. The first and the last contribution are of general validity. We believe the method we propose for the calculation of L to be best suited for objects with a relatively simple contour, and of reasonably large dimensions compared to the mesh size.

REFERENCES

[1] M. Cheney, D. Isaacson, J. C. Newell, S. Simake, and J. Goble, "NOSER: An algorithm for solving the inverse conductivity problem," *Int J Imag Sys Technol*, vol. 2, pp. 66–75, 1990.

[2] P. Hua, J. G. Webster, and W. J. Tompkins, "A regularised electrical impedance tomography reconstruction algorithm," *Clin Phys Physiol Meas*, vol. 9, pp. A137–141, 1988.

[3] D. C. Dobson and F. Santosa, "An image enhancement technique for electrical impedance tomography," *Inverse Problems*, vol. 10, pp. 317–334, 1994.

[4] A. Borsic, C. N. McLeod, and W. R. B. Lionheart, "Total variation regularisation in EIT reconstruction," *Proceedings of the 2nd World Congress on Industrial Process Tomography*, 2001.

[5] E.J. Woo, P. Hua, J.G. Webster, and W.J. Tompkins, "Measuring lung resistivity using electrical impedance tomography," *IEEE Trans Biomed Eng*, vol. 39, pp. 756–760, 1992.

[6] M. Glidewell and K.T. Ng, "Anatomically constrained electrical impedance tomography for anisotropic bodies via a twostep approach," *IEEE Trans Med Imgaging*, vol. 14, pp. 498–503, 1995.

[7] B.M. Eyuboglu, T.C. Pilkington, and P.D. Wolf, "Estimation of tissue resistivities from multipleelectrode impedance measurements," *Phys Med Biol*, vol. 39, pp. 1–17, 1994.

[8] M. Vauhkonen, J.P. Kaipio, E. Somersalo, and P.A. Karjalainen, "Electrical impedance tomography with basis constraints," *Inverse Problems*, vol. 13(2), pp. 523–530, 1997.

[9] M. Vauhkonen, D. Vadasz, J. P.Kaipio, E. Somersalo, and P. A. Karjalainen, "Tikhonov regularization and prior information in electrical impedance tomography," *IEEE Trans Med Imaging*, vol. 17, pp. 285–293, 1998.

[10] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, and E. Somersalo, "Construction of nonstandard smothness priors," *Inverse Problems*, vol. 15, pp. 713–729, 1999.

[11] A. Tarantola, *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, 1987.

[12] A. Adler and R. Guardo, "Electrical impedance tomography: Regularized imaging and contrast detection," *IEEE Trans Med Imgaging*, vol. 15, pp. 170–179, 1996.

[13] E. Somersalo, M. Cheney, and D. Isaacson, "Existence and uniqueness for electrode models for electric current computed tomography," *SIAM J Appl Math*, vol. 52, pp. 1023–1040, 1992.

[14] Q. S. Zhu, C. N. McLeod, C. W. Denyer, F. J. Lidgey, and W. R. B. Lionheart, "Development of a real-time adaptive current tomograph," *Physiol Meas*, vol. 15, pp. A37–A43, 1994.

[15] B.H. Brown and A.D. Seagar, "Applied potential tomography - data collection problems," *Proceedings of the IEE International Conference on Electric and Magnetic Fields in Medicine and Biology*, 1985.

[16] M. Hanke and P.C. Hansen, "Regularization methods for largescale problems," *Surv Math Ind*, vol. 3, pp. 253–315, 1993.

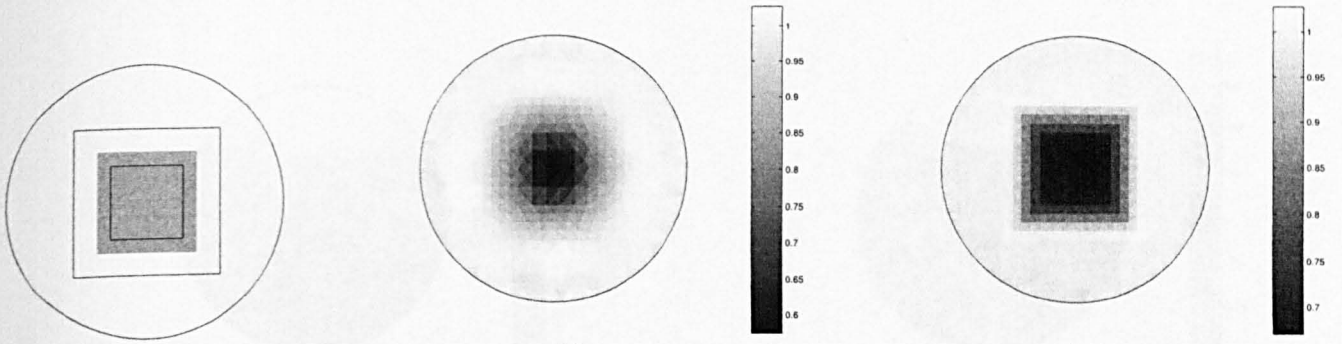


Fig. 5. Comparisons of isotropic and anisotropic priors. a) Test conductivity profile. b) Reconstruction with Gaussian isotropic smoothing of the 10 cm square inclusion. c) Reconstruction using a Gaussian anisotropic filter

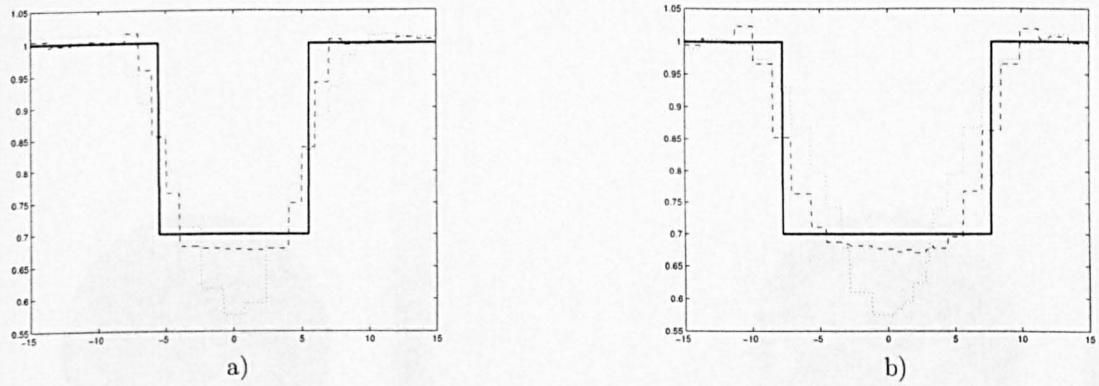


Fig. 6. Cross sections of the reconstructions. The thick line represents the true conductivity, the dashed line the anisotropic reconstruction and the dotted line the isotropic one. a) Cross section along the horizontal axis. b) Cross section along the 45° diagonal.

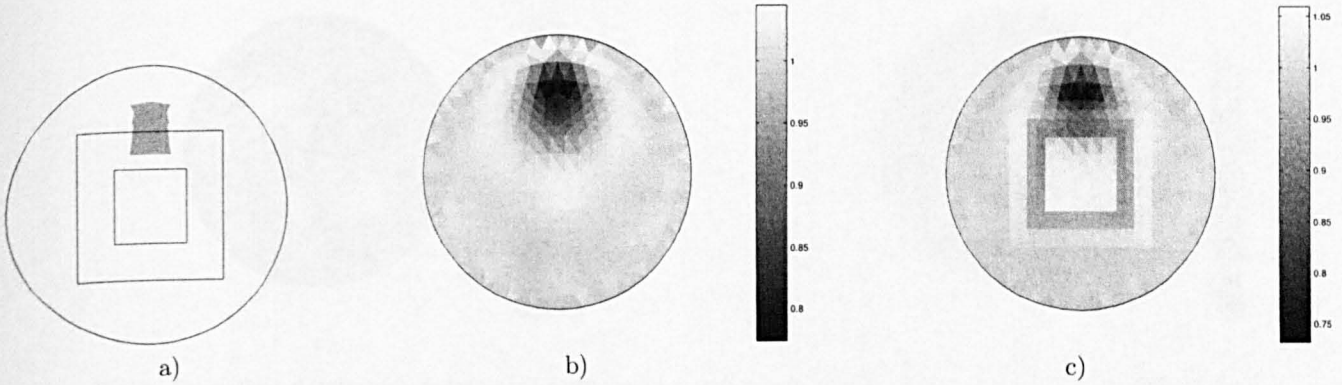


Fig. 7. Inclusion violating the priors. a) A rectangular inclusion crosses the region Ω_{change}, the prior assumptions are violated, the lateral borders of the inclusion are orthogonal to the expected direction. b) Isotropic reconstruction of the conductivity. c) Anisotropic reconstruction of the same conductivity.

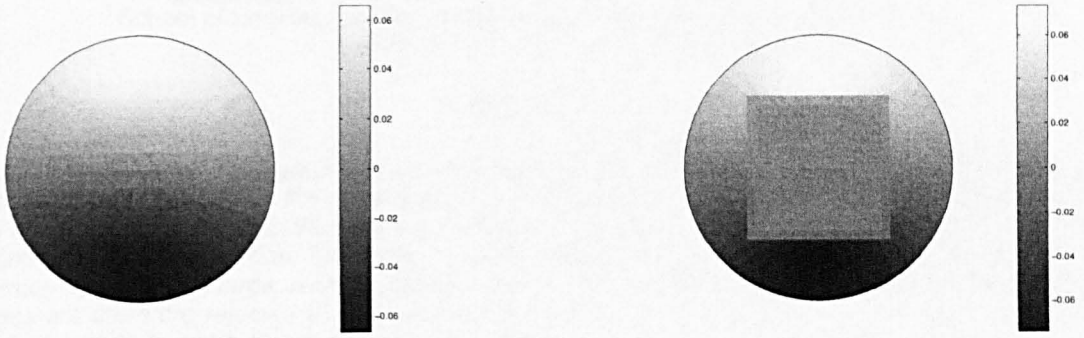


Fig. 8. a) Second singular vector calculated with the Gaussian isotropic matrix L. b) Second singular vector in the Gaussian anisotropic case.

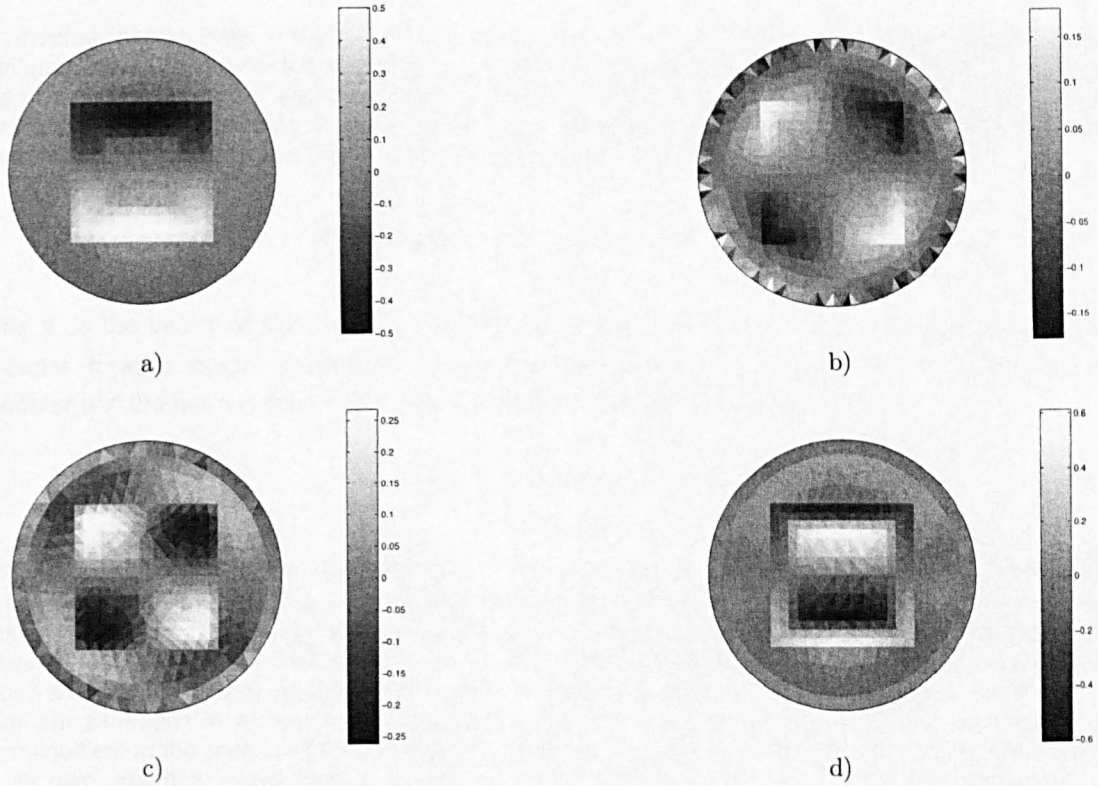


Fig. 9. Generalized singular vectors calculated with the Gaussian anisotropic matrix L. a) 14^{th}, b) 42^{nd}, c) 108^{th}, d) 131^{st}.

Total Variation regularisation in EIT Reconstruction

Andrea Borsic[1], Christopher N. McLeod[1], William R.B. Lionheart[2]
[1] School of Engineering, Oxford Brookes University, aborsic@brookes.ac.uk
[2] Department of Mathematics, UMIST

ABSTRACT

Electrical Impedance Tomography is an ill-posed inverse problem. Regularisation techniques are adopted in order to stabilise the reconstruction. The problem is formulated as the minimisation of the sum of two terms: the first is the mismatch between the real measurements and simulated ones, the second is the regularisation functional. The functional accomplishes the task of stabilising the inversion by assuming large values corresponding to conductivity distributions that are unlikely. Such images are therefore rejected by the minimisation process. The choice of a regularisation functional selects the class in which the inverse solutions will lie.
Commonly adopted regularisation functionals in EIT penalise non-smooth images, rendering the reconstruction algorithm incapable of describing step conductivity distributions. Situations where the imaged body has a non continuous conductivity distribution are of practical interest both in process and medical imaging. Total Variation regularisation is an emerging technique that we believe is particularly successful for solving such problems. The technique involves a regularisation functional which is not differentiable everywhere. An appropriate framework for solving the problem efficiently is presented in this paper.

Keywords Electrical Impedance Tomography, Total Variation, Primal Dual Interior Point Methods

1 INTRODUCTION

The inverse conductivity problem is well known to be ill-posed (Brekon 1990); regularisation techniques have been adopted in order to stabilise the inversion. Usually a FEM forward model is used in the reconstruction algorithm: the reconstruction process estimates the conductivity of each discrete element by matching the simulated measurements to the real ones. The reconstruction is commonly stabilised using the Tikhonov regularisation; the inversion is stated as:

$$\sigma_{rec} = \arg\min \frac{1}{2}\left\| v - h(\sigma) \right\|^2 + \alpha F(\sigma) \tag{1}$$

where v is the vector of the measured voltages, σ is the discretised conductivity vector, $h(\sigma)$ the non-linear forward model prediction, $F(\sigma)$ the regularisation functional, α the regularisation parameter and the norm is the 2-norm. The functional is very often expressed as:

$$F(\sigma) = \left\| L(\sigma - \sigma^*) \right\|^2 \tag{2}$$

where L is an appropriate regularisation matrix and σ^* a prior estimate of the conductivity distribution. In the literature there are several choices for the matrix L, for example the identity matrix (Yorkey 1986), a positive diagonal matrix (Cheney 1990) and approximations of first and second order differential operators (Hua 1990). All these choices tend to smooth the reconstructed images. They impose a certain degree of smoothness in order to obtain stability in the reconstruction process.
There are situations in almost every field of application of EIT where the imaged conductivity has discontinuities. In the medical field an example is that of the inter organ boundaries where each organ has its own electrical properties. In archaeology a buried wall will give rise to a sudden jump in conductivity, and in process tomography a multiphasic fluid will give rise to discontinuities at each phase interface.
It is therefore important to be able to reconstruct these situations correctly, even though such conductivities are difficult to deal with using a traditional algorithms. Several approaches have been investigated in order to overcome these limitations, often they can be categorised as a way to introduce prior information. An example is the anisotopic regularisation (Kaipio 1999; Borsic 2001a)

where the structure of the expected sudden changes is assumed to be roughly known. The smoothness constraints are relaxed therefore in the direction normal to the discontinuities; in this way the algorithm describes rapid variations in the object better.
Algorithms of this class use a particular matrix L which is calculated *ad hoc* given the specific prior information of the problem. They fall into general framework expressed by equations (1) and (2), that is:

$$\sigma_{rec} = \arg\min \frac{1}{2}\| v - h(\sigma) \|^2 + \alpha \sum_i \left[L_i (\sigma - \sigma^*) \right]^2 \tag{3}$$

where L_i is the ith line of the matrix L.

The framework expressed by eq. (3) can be called L^2 regularisation since the 2-norm is used. A norm guarantees that the functional is always positive, as a penalty term should be, and more important, the resulting L^2 functional is differentiable, leading to an easier solution of the minimisation problem.

L^2 regularisation, because of its simplicity, has been the common framework for solving several inverse problems, and particularly for EIT (Yorkey 1986; Cheney 1990; Hua 1990; Brekon 1990; Kaipio 1999; Borsic 2001a). The drawback is that the technique is bad at describing sudden changes in the solution, irrespective of the choice of L it tends to distribute the changes evenly across the image and therefore to favour smooth solutions.

2 TOTAL VARIATION FUNCTIONAL

Many regularisation matrices are discreet representations of differential operators and they are used in conjunction with the L^2 norm. A different approach is represented by the choice of the total variation functional, which is still a differential operator but leads to a L^1 regularisation.
The total variation (TV) of a conductivity image is defined as:

$$TV(\sigma) = \int_\Omega |\nabla \sigma| \, d\Omega \tag{4}$$

where Ω is the region to be imaged.
The TV functional was firstly employed by Rudin, Osher and Fatemi (1992) for regularising the restoration of noisy images. The technique is particularly effective for recovering 'blocky' images, and has become well known to the image restoration community (Chan 1998).
The effectiveness of the method in recovering discontinuous images can be understood by examining the following one dimensional situation.

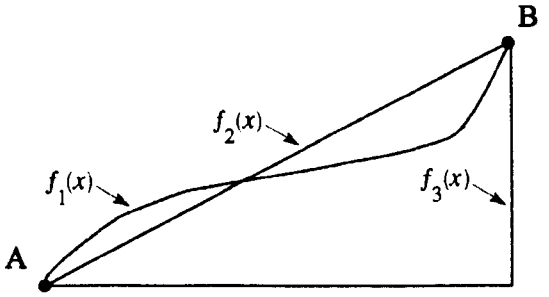


Figure 1. Two points A and B can be connected by several paths. All of them have the same TV

Suppose that the two points A and B of Figure 1 are to be connected by a path. Three possible functions $f(x)$ connecting them are shown. Each one has the same total variation as the others:

$$TV = \int_A^B f'(x)\,dx = f(B) - f(A) \tag{5}$$

The three possible functions would have the same penalty; TV regularisation treats f_1, f_2 and f_3 in the same way. A much broader class of functions are allowed therefore to be the solution of the inverse problem, including functions with discontinuities.
An other way to understand the differences with other techniques is to consider the discretised version of eq. (4). Suppose that the conductivity is described by piecewise constant elements, the TV of the image, in the sense of distributions, can be expressed as:

$$TV(\sigma) = \sum_k l_k \left| \sigma_{m(k)} - \sigma_{n(k)} \right| \tag{6}$$

where l_k is the length of the k th edge in the mesh, $m(k)$ and $n(k)$ are the indices of the two elements on opposite sides of the k th edge, and the index k ranges over all the edges. Equation (6) can be expressed in terms of matrices as:

$$TV(\sigma) = \sum_k \left| L_k \sigma \right| \tag{7}$$

where L is a sparse matrix, with one row per each edge in the mesh. Every row L_k has two non zero elements in the columns $m(k)$ and $n(k)$: $L_k = [0,\ldots,0,l_k,0,\ldots,0,-l_k,0,\ldots,0]$.

TV regularisation is therefore of the L^1 kind: it is a sum of absolute values. The absolute value guarantees the positivity of the penalty function but not the differentiability at the points where $\sigma_{m(k)} = \sigma_{n(k)}$. The numerical problem needs to be addressed properly. The important gain is that the L^1 regularisation doesn't distribute the jumps across the image evenly.

3 IMAGE ENHANCEMENT APPROACH

Early published work making use of the TV properties employed the technique as an image enhancement method. Dobson and Santosa (1994) published the first attempt to use the new technique in EIT. They do not use TV as a regularisation functional but as an approach to image enhancement. Traditional techniques were used in order to reconstruct a first image; the sharpness of the image was enhanced by minimising the TV in succeeding iterations.
Given a conductivity distribution that satisfies the measurements, an iterative constrained minimisation can be applied:

$$\sigma_{rec} = \arg\min \int_\Omega |\nabla \sigma| d\Omega \quad \text{subject to} \quad h(\sigma) = v \tag{8}$$

The TV of the conductivity is minimised, providing that the simulated measurements match the real ones. There are two difficulties associated with the formulation of eq. (8). The first is that noise on the measurements can bring the vector v outside the range of the forward operator, and therefore there would not exist any conductivity distribution satisfying the constraints. The second is that the inverse problem is ill-conditioned: small changes in v could lead to large changes in any σ satisfying the constraints.
The solution proposed by Dobson and Santosa is to stabilise the constraints by truncating the SVD decomposition of the linearised forward operator. The proposed method therefore addresses the linear step. The conductivity increment $\delta\sigma_{rec}$ is calculated as:

$$\delta\sigma_{rec} = \arg\min \int_\Omega |\nabla \delta\sigma| d\Omega \quad \text{subject to} \quad M\delta\sigma = v' \tag{9}$$

where M is the linearised and SVD truncated forward operator and v' is the projection of the real measurements onto the reduced measurements space.

3.1 Numerical solution

The problem of eq. (9) is numerically difficult to solve because of the non-differentiability of the TV functional. Dobson and Santosa define the following function:

$$w(x)=\begin{cases} x & if\ x>\varepsilon \\ \dfrac{x^2}{2\varepsilon}+\dfrac{\varepsilon}{2} & if\ x\le\varepsilon \end{cases} \tag{10}$$

and solve (9) redefining (6) as:

$$TV(\sigma)=\sum_i l_i w(\left|\sigma_{m(i)}-\sigma_{n(i)}\right|) \tag{10}$$

The function $w(x)$ is of class C^1 for $\varepsilon>0$ and rounds off the corners of the absolute value. The numerical scheme proposed for solving (9) consists in finding a feasible starting point for σ by applying the Moore-Penrose inverse of M to the vector v'. Successively, the conductivity is updated by projecting the gradient of (10) on the null space of M thus maintaining the feasibility during the iterations. The problem is solved for a sufficiently small value of ε.
The published results show the ability of the TV approach to reconstruct blocky images, but the numerical efficiency is limited and only the linearised problem is solved.

4 EFFICIENT METHODS FOR TV REGULARISATION

Recent developments in operations research (Andersen 2000) have provided new classes of methods to deal efficiently with the problems of minimising the sum of absolute values.
Chan, Golub and Mulet (1995) have drawn from these advances (already public as a technical report) and investigated the problem of restoring images with Primal Dual Interior Point Methods (PD-IPM). The formulation of the image restoration problem is very similar to the EIT reconstruction problem, and results can be easily exploited.
The reconstruction problem is Tikhonov regularised:

$$\sigma_{rec}=\arg\min\frac{1}{2}\left\|v-h(\sigma)\right\|^2+\alpha\sum_k l_k\left|\sigma_{m(k)}-\sigma_{n(k)}\right| \tag{11}$$

or more conveniently:

$$\sigma_{rec}=\arg\min\frac{1}{2}\left\|v-h(\sigma)\right\|^2+\alpha\sum_k\left\|L_k\sigma\right\| \tag{12}$$

The absolute value has been exchanged with the 2-norm since each argument is a scalar. The non smooth optimisation problem of eq. (12) is efficiently addressed by the primal dual interior point method. If we call problem (12) primal (P), there exists a second problem, called dual (D) that can be derived from P using the equality: $\left\|x\right\|=\max_{|y|\le 1} x^{\mathrm{T}}y$

$$\min_\sigma\frac{1}{2}\left\|v-h(\sigma)\right\|^2+\alpha\sum_k\left\|L_k\sigma\right\|=\min_\sigma\max_{|\eta_k|\le 1}\left(\frac{1}{2}\left\|v-h(\sigma)\right\|^2+\alpha\,\sigma^{\mathrm{T}}L^{\mathrm{T}}\eta\right) \tag{13}$$

in (13) the min and max can be exchanged (Rockafellar Cor. 37.3.2) leading to:

$$\min_\sigma\frac{1}{2}\left\|v-h(\sigma)\right\|^2+\alpha\sum_k\left\|L_k\sigma\right\|=\max_{|\eta_k|\le 1}\min_\sigma\left(\frac{1}{2}\left\|v-h(\sigma)\right\|^2+\alpha\,\sigma^{\mathrm{T}}L^{\mathrm{T}}\eta\right) \tag{14}$$

the problem $\min_{\sigma} \frac{1}{2}\|v-h(\sigma)\|^2 + \alpha\,\sigma^{\mathrm{T}} L^{\mathrm{T}}\eta$ can be solved by solving the linear equation:

$$\alpha\, L^{\mathrm{T}}\eta - J^{\mathrm{T}}(v-h(\sigma)) = 0 \tag{15}$$

where J is the Jacobian of the forward operator. The problem D can therefore be written as:

$$\max_{\substack{|\eta_k|\le 1 \\ \alpha L^{\mathrm{T}}\eta - J^{\mathrm{T}}(v-h(\sigma))=0}} \left(\frac{1}{2}\|v-h(\sigma)\|^2 + \alpha\,\sigma^{\mathrm{T}} L^{\mathrm{T}}\eta\right) \tag{16}$$

The problems P and D form what is called a primal dual pair. The variables σ are called primal variables and η dual variables. An optimal solution of P is also an optimal solution of D and under opportune conditions feasible points of P are upper bounds to feasible points of D, and feasible points of D are lower bounds to feasible points of P. The solution is therefore bounded above and below by feasible points of P and D.
The way to solve efficiently (12) or (16) is to reduce the gap between P and D rather then trying to minimise (12) or maximise (16) separately. The primal dual gap has to be zero at an optimal solution:

$$\frac{1}{2}\|v-h(\sigma)\|^2 + \alpha\sum_k \|L_k\sigma\| = \frac{1}{2}\|v-h(\sigma)\|^2 + \alpha\sum_k \eta_k{}^{\mathrm{T}} L_k\sigma \tag{17}$$

the condition $\|\eta_k\| \le 1$ together with condition (17) is equivalent to the condition (18):

$$\|L_k\sigma\| = \eta_k^{\mathrm{T}} L_k\sigma \quad \forall k \tag{18}$$

which in turn is equivalent to:

$$\|L_k\sigma\|\eta_k - L_k\sigma = 0 \quad \forall k \tag{19}$$

The solution of the P and D problems is therefore yielded by the system of equations:

$$\begin{aligned} \|L_k\sigma\|\eta_k - L_k\sigma &= 0 \\ \alpha\, L^{\mathrm{T}}\eta - J^{\mathrm{T}}(v-h(\sigma)) &= 0 \\ \|\eta_k\| &\le 1 \end{aligned} \tag{20}$$

The condition (19) is not-differentiable if $\|L_k\sigma\| = 0$. This is addressed by replacing the term by $\sqrt{\|L_k\sigma\|^2 + \beta}$. The system can then be linearised and properly solved iteratively with $\beta \to 0$.
The non-differentiability of (19) may leave one to wonder if there was any gain in solving a non-differentiable problem with a method that presents the same difficulty. The method however has proven to be extremely efficient, and has been applied to a number of applications (Alpert 1997; Alexander 1998)

5 SIMULATIONS

A reconstruction algorithm that formulates the inverse problem as in (11) and solves it as in (20) was developed in the MATLAB environment. The method proposed by Chan *et al.* to solve (20) assumes the forward operator to be linear. The reconstructions that we present in this section of the paper are fully non-linear, the algorithm has shown to work in this case, but we do not provide a proof of convergence at present.

The forward solver used for the reconstructions uses Finite Element Method and implements the complete electrode model; the practical implementation of the algorithm has been published by Borsic, McLeod and Lionheart (2001b).
Reconstructions were performed on a simulated round object and are presented in figures 2, 3 and 4. The conductivity of the object was discretised with a mesh of 510 elements. Measurements were performed by stimulating the object with a set of 32 current electrodes and by measuring the electric potential on a separate set of 32 voltage electrodes interleaved with the first ones. The first 10 trigonometric current patterns were applied, resulting in 320 measurements given the measurement configuration.
All the results presented in this paper are a preliminary study: the conductivity was reconstructed on the same mesh used for generating the forward solutions, and no noise was added to the measurements. However the simulations were useful for testing the efficiency of the algorithm and its ability to reconstruct step conductivity images.
On the left side of figures 2, 3 and 4 is shown the true conductivity and on the right side the reconstructed conductivity. Figure 2 and 3 show how the L^1 regularisation is capable of reconstructing particular situations that would be otherwise difficult to treat with the L^2 scheme. The TV regularisation is know to give good results in the case of 'blocky' objects, and to perform less well in case of 'strip' objects (objects that have a very large perimeter compared to the surface). The case simulated in figure 4 is therefore more difficult to be handled by the TV approach, even more because a circular annulus is in general a difficult situation to deal with in EIT.
The reconstruction method has been demonstrated to be very efficient compared to the solution proposed by Dobson and Santosa, which was also implemented in order to allow comparisons. The progress achieved by the first 6 steps of the PD-IPM is usually reached after more then a thousand iterations by the projected gradient method. The presented images are the results of the 15th iteration of the DP-IPM algorithm, the overall processing time is approximately 34 seconds on a Pentium II machine running at 400Mhz. Conversely the processing time for the first 1000 steps of the projected gradient method (PGM) is greater then 600 seconds. However it must be noted that the PGM solves the linear problem, its progress decreases thus during the iterations, and even if left working for hours it is not capable of reaching the progress of the first 15 steps of PD-IPM.

6 CONCLUSIONS AND FUTURE WORK

Practical results of the TV regularisation and the efficiency of PD-IPM method are of interest in process and medical imaging. More work has to be done in order to evaluate the robustness of the method in the presence of noisy measurements, we expect therefore to publish results from tank studies in future. It is important to note that the generality of the PD-IPM scheme allows its use for the 3D TV regularisation. The method is expected to work equally well in three dimensions, and to be easily extended to this case.

7 BIBLIOGRAPHY

Alexander J. C., Christiansen E., Overton M. L. *Computing limit loads by minimizing a sum of norms.* SIAM Journal on Scientific Computing, 19, 1046–1062, 1998.

Alpert J., Chan T.F., Huang D. J. H., Kahng A. B., Markov I. L., Moulet P., Yan K. *Faster minimization of linear wirelength for global placement.* Proc. International Symposium on Physical Design, 4–11, April 1997.

Andersen K. D., Christiansen E., Conn A. R., Overton M. L. *An efficient primal-dual interior-point method for minimizing a sum of Euclidean norms,* SIAM J. on scientific computing, 22, 243–262, 2000.

Borsic A., Lionheart W. R. B., McLeod C. N. *Introduction of a priori information in EIT regularisation.* Proceedings of 3rd EPSRC Engineering network meeting on biomedical applications of EIT, 2001–a.

Borsic A., McLeod C., Lionheart W. *Realistic 2D human thorax modelling for EIT.* Physiological Measurement, 22, 77–83, 2001–b.

Breckon W. R. *Image reconstruction in electrical impedance tomography.* PhD Thesis, Oxford Polytechnic, 1990.

Chan T. F., Golub G. H., Mulet P. *A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration.* UCLA Math Department CAM Report 95–43, 1995.

Chan T., Marquina A., Mulet P. *High-order total variation-based image restoration.* SIAM Journal on Scientific Computing, 19, 1046–1062, 1998.

Cheney M., Isaacson D., Newell J.C., Simake S., Goble J. *NOSER: An algorithm for solving the inverse conductivity problem.* Int. J. Imag. Sys. Technol. 2, 66–75, 1990.

Dobson D. C., Santosa F. *An image enhancement technique for electrical impedance tomography.* Inverse Problems, 10, 317–334, 1994.

Hua P., Webster J. G., Tompkins W. *A regularised electrical impedance tomography reconstruction algorithm.* Clin. Phys. Physiol. Meas. 2,66–75,1990.

Kaipio J. P., Kolehmainen V., Vauhkonen M., Somersalo E. *Construction of nonstandard smothness priors,* Inverse Problems, 15, 713–729, 1999.

Rockafellar T. *Convex Analysis,* SIAM.

Rudin L. I., Osher S., Fatemi E. *Nonlinear total variation based noise removal algorithms.* Physica D, 60, 259–268, 1992.

Yorkey T. J. *A quantitative comparison of the reconstruction algorithms used in impedance tomography.* PhD Thesis, University of Wisconsin, 1986.

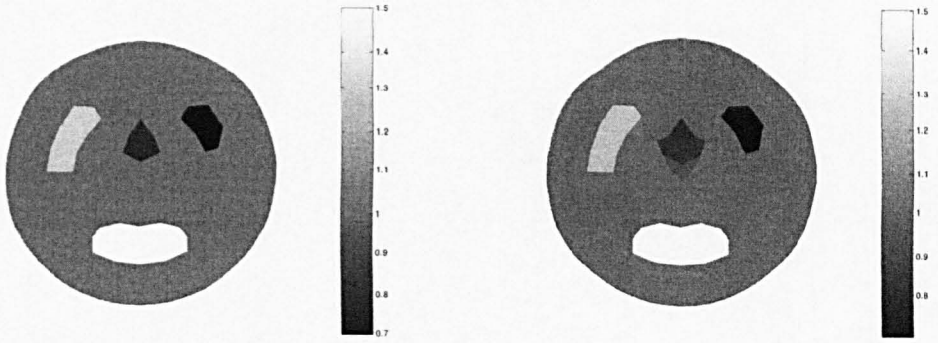


Figure 2: Reconstruction of and object with several inclusions, left true conductivity, right reconstruction.

Figure 3: Reconstruction of U shaped object, left true conductivity, right reconstruction.

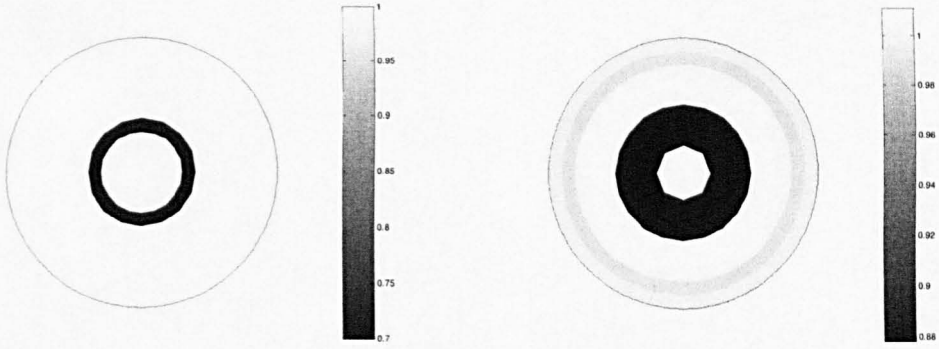


Figure 4: Reconstruction of a round annulus, left true conductivity, right reconstruction.