

***Total Variation Regularization in Electrical
Impedance Tomography***

Borsic, Andrea and Graham, Brad M. and Adler,
Andy and Lionheart, William R.B.

2007

MIMS EPrint: **2007.92**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Total Variation Regularization in Electrical Impedance Tomography

A. Borsic*, B. M. Graham†, A. Adler‡, W. R. B. Lionheart§

May 31, 2007

Abstract

This paper presents an evaluation of the use of Primal Dual Methods for efficiently regularizing the electric impedance tomography (EIT) problem with the Total Variation (TV) functional.

The Total Variation functional is assuming an important role in the regularization of inverse problems thanks to its ability to preserve discontinuities in reconstructed profiles. This property is desirable in many fields of application of EIT imaging, such as the medical and the industrial, where inter-organ boundaries, in the first case, and inter-phase boundaries, in the latter case, present step changes in electrical properties which are difficult to be reconstructed with traditional regularization methods, as they tend to smooth the reconstructed image. Though desirable, the TV functional leads to the formulation of the inverse problem as a minimization of a non-differentiable function which cannot be efficiently solved with traditional optimization techniques such as the Newton Method. In this paper we demonstrate the use of Primal Dual - Interior Point Methods (PD-IPM) as a framework for TV regularized inversion.

This paper introduces the smoothing properties of the traditional quadratic regularization algorithms, the discontinuity preserving properties of the TV functional are then outlined. The paper follows introducing the general PD-IPM framework and its application to inverse problems. Specifically 2D and 3D results from the application of TV regularization to the EIT inverse problem are presented and analyzed. Trough practical examples the discontinuity preserving capabilities of

*A. Borsic is with the Thayer School of Engineering, Dartmouth College, USA, email: andrea.borsic@dartmouth.edu

†B. M. Graham is with the School of Information Technology and Engineering (SITE), University of Ottawa, Canada, email: graham.bm@sympatico.ca

‡A. Adler is with the Department of Systems and Computer Engineering, Carleton University, Canada, email: adler@sce.carleton.ca

§W. R. B. Lionheart is with the School of Mathematics, University of Manchester, Manchester, UK, email: bill.lionheart@manchester.ac.uk

the method are shown, and compared to quadratic regularization results.

Keywords: Electrical Impedance Tomography, EIT, Regularization, Total Variation, TV, Primal Dual Interior Point Methods

1 Introduction

Electrical Impedance Tomography (EIT) uses surface electrodes to make measurements from which an image of the conductivity distribution within some medium is calculated. The inverse conductivity problem is ill-posed [2]; consequently regularization techniques have been adopted in order to stabilize the inversion. Most common regularization methods impose (explicitly or implicitly) a penalty on non-smooth regions in a reconstructed image. Such methods confer stability to the reconstruction process, but limit the capability of describing sharp variations in the sought parameter.

One technique to permit image regularization without imposing smoothing is the Total Variation (TV) formulation of regularization. The Total Variation functional is assuming an important role in the regularization of inverse problems belonging to many disciplines, thanks to its ability to preserve discontinuities in the reconstructed profiles. Application of non-smooth reconstruction techniques is important for medical and process imaging applications of EIT, as they involve discontinuous profiles. Qualitative and quantitative benefits can be expected in these fields.

We outline the properties of the TV functional in the next section, to motivate its use as a regularization penalty term and to understand the numerical difficulties associated with it. The use of the TV functional leads in fact to the formulation of the inverse problem as a minimization of a non-differentiable function. Application of traditional minimization techniques (Steepest Descent Method, Newton Method) has proven to be inefficient [1][23]. Recent developments in non-smooth optimization (Primal Dual-Interior Point Methods) have brought the means of dealing with the minimization problem efficiently. The performance of this algorithm with respect to traditional smooth algorithms is the subject of this paper.

2 Methods

This paper introduces the PD-IPM algorithm as follows. In the Methods section we describe the traditional family of EIT reconstruction algorithms used in our research, describe the TV functional and its PD-IPM implementation for EIT, and describe the evaluation procedure. In the Results section we describe the effectiveness of the TV functional compared to the quadratic regularized inverse. In the Discussion section we consider some additional observations of this work.

2.1 Static Image Reconstruction

We consider static EIT imaging where the goal of the algorithm is to recover the absolute conductivity of the medium under analysis. The technique requires a forward operator F on the conductivity vector, $\boldsymbol{\sigma}$, which calculates $\mathbf{V} = F(\boldsymbol{\sigma})$, the simulated voltages at the boundary. The reconstruction is commonly stabilized using regularization; the inversion is stated as:

$$\hat{\boldsymbol{\sigma}}_{rec} = \arg \min_{\boldsymbol{\sigma}} \frac{1}{2} \|F(\boldsymbol{\sigma}) - \mathbf{V}_{meas}\|^2 + \alpha G(\boldsymbol{\sigma}) \quad (1)$$

where \mathbf{V}_{meas} is the vector of the measured voltages $F(\boldsymbol{\sigma})$ the forward model prediction, $G(\boldsymbol{\sigma})$ the regularization functional, α is a hyperparameter controlling the level of applied regularization and the norm $\|\cdot\|$ is the 2-norm.

2.2 Quadratic Solution

The functional $G(\boldsymbol{\sigma})$ is often assumed to be of the form:

$$G(\boldsymbol{\sigma}) = \|\mathbf{L}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^*)\|^2 \quad (2)$$

where \mathbf{L} is an appropriate regularization matrix and $\boldsymbol{\sigma}^*$ a prior estimate of the conductivity distribution. In the literature there are several choices for the matrix \mathbf{L} , for example the identity matrix [3], a positive diagonal matrix [5], approximations of first and second order differential operators [4], and the inverse of a Gaussian matrix [6]. Algorithms of this class fall into general framework expressed by equations (1) and (2), that is:

$$\hat{\boldsymbol{\sigma}}_{rec} = \arg \min_{\boldsymbol{\sigma}} \frac{1}{2} \|F(\boldsymbol{\sigma}) - \mathbf{V}_{meas}\|^2 + \alpha \|\mathbf{L}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^*)\|^2 \quad (3)$$

The framework expressed by eq (3) can be called quadratic regularization since the 2-norm is used. A norm guarantees that the functional is always non-negative, as a penalty term should be, and more important, the resulting functional is differentiable, leading to an easier solution of the minimization problem. Quadratic regularization, because of its simple differentiability, has been the common framework for solving several inverse problems, and particularly for EIT [3],[5],[4],[2],[7],[8].

The optimization problem (3) can be solved by replacing $F(\boldsymbol{\sigma})$ with its linear approximation for a small change about an initial conductivity distribution $\boldsymbol{\sigma}_0$

$$F(\boldsymbol{\sigma}) \approx F(\boldsymbol{\sigma}_0) + \mathbf{J}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \quad (4)$$

where \mathbf{J} is the Jacobian matrix of $F(\boldsymbol{\sigma})$ calculated at the initial conductivity estimate $\boldsymbol{\sigma}_0$. The function to be minimized (1) with regularizing penalty term (2) becomes a quadratic function when F is replaced by its linear

approximation (4). Defining $\delta\boldsymbol{\sigma} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_0$ and $\delta\mathbf{V} = F(\boldsymbol{\sigma}_0) - \mathbf{V}_{\text{meas}}$, the solution to the linearized regularization problem is given by

$$\delta\boldsymbol{\sigma} = (\mathbf{J}^T \mathbf{J} + \alpha \mathbf{L}^T \mathbf{L})^{-1} \mathbf{J}^T \delta\mathbf{V} + \alpha \mathbf{L}^T \mathbf{L} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_{ref}) \quad (5)$$

Equation 5 is solved iteratively with $\boldsymbol{\sigma}_{i+1} = \boldsymbol{\sigma}_i + \delta\boldsymbol{\sigma}$

The drawback is that, regardless of the choice of \mathbf{L} , the technique cannot reconstruct step changes, smooth solutions are favoured.

2.3 Total Variation Functional

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 step 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 traditional algorithms. Several approaches have been investigated in order to overcome these limitations. Often they can be considered a way to introduce prior information. An example is anisotropic regularization [7][8] 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 better describes rapid variations in the object, however prior structural information needs to be known in order to adopt such methods.

Many regularization matrices are discrete representations of differential operators and are used in conjunction with the 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 regularization. The total variation (TV) of a conductivity image is defined as:

$$TV(\sigma) = \int_{\Omega} |\nabla\sigma| d\Omega \quad (6)$$

where Ω is the region to be imaged.

The TV functional was first employed by Rudin, Osher, and Fatemi [13] for regularizing the restoration of noisy images. The technique is particularly effective for recovering “blocky” images, and has become well known to the image restoration community [14]. The effectiveness of the method in recovering discontinuous images can be understood by examining the following one dimensional situation.

Suppose that the two points A and B of figure 1 are connected by a path. Three possible functions $f(x)$ connecting them are shown. As the functions

are monotonically increasing, the TV of each is:

$$TV(f) = \int_A^B f'(x) dx = f(B) - f(A) \quad (7)$$

which is the same value for each function. TV treats f_1 , f_2 and f_3 in the same way and when used as a penalty term in a Tikhonov regularized inverse problem, will not bias the result towards a smooth solution. On the other hand, the ℓ^2 norm assumes different values for f_1, f_2 and f_3 . When used as a penalty term the ℓ^2 norm will bias the solution towards smoother functions, for which the ℓ^2 norm assumes smaller values. In the cited example f_3 is inadmissible as a quadratic solution since its ℓ^2 norm is infinity. With the use of TV as a regularization penalty term a much broader class of functions are therefore allowed to be the solution of the inverse problem, including functions with discontinuities. Another way to understand the differences with other techniques is to consider the discretized version of equation (6). Suppose that the conductivity is described by piecewise constant elements, the TV of the a 2D image can be expressed as the sum of the TV of each of the k edges, with each edge weighted by its length:

$$TV(\sigma) = \sum_k l_k |\sigma_{m(k)} - \sigma_{n(k)}| \quad (8)$$

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 (8) can be expressed in terms of matrices as:

$$TV(\sigma) = \sum_k |\mathbf{L}_k \sigma| \quad (9)$$

where \mathbf{L} is a sparse matrix, with one row per each edge in the mesh. Every row \mathbf{L}_k has two non zero elements in the columns $m(k)$ and $n(k)$: $\mathbf{L}_k = [0, \dots, 0, l_k, 0, \dots, 0, -l_k, 0 \dots 0]$. TV regularization is therefore of the ℓ^1 kind: it is a sum of absolute values, in this case a sum of vector lengths. The absolute value guarantees the positivity of the penalty function but unfortunately results in non-differentiability in the points where $\sigma_{m(k)} = \sigma_{n(k)}$. The numerical problem thus needs to be addressed properly. However, the important gain is that the ℓ^1 regularization does not penalize discontinuities.

2.3.1 Solving TV - Early Approaches.

Two different approaches were proposed for application of TV to EIT, the first by Dobson and Santosa [1] and the second by Somersalo *et al.* [10] and Kolehmainen [11]. Dobson and Santosa replace the absolute value function in the neighbourhood of zero by a polynomial to obtain continuously

differentiable function upon which steepest descent is then used to perform the minimization. Their approach is suitable for the linearized problem but suffers from poor numerical efficiency. Somersalo and Kolehmainen successfully applied Markov Chain Monte Carlo (MCMC) methods to solve the TV regularized 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 error bars by sampling the posterior probability density of the sought parameters (therefore differentiability is not required). The sampling process involves a substantial computational effort, often the inverse problem is linearized in order to speed up the sampling. What is required is an efficient method for deterministic Tikhonov style regularization, to calculate a non-linear TV regularized inversion in a short time.

Examination of the literature shows that a variety of deterministic numerical methods have been used for the regularization of image de-noising and restoration problems with the TV functional (a good review is offered by Vogel in [15]). 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 optimization, as TV minimization belongs to the important classes of problems known as “Minimization of sum of norms” [16] [21] [18] and “Linear ℓ^1 problems” [19] [20].

Recent developments in operations research [21] have provided new classes of methods to deal efficiently with the problems of minimising the sum of absolute values. Chan, Golub and Mulet [22] have drawn from these advances 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. In the next section we summarize some results from Andersen, Christiansen, Conn and Overton [21] that are at the base of the method proposed by Chan [22] in image restoration applications, and of the method we propose for EIT.

2.4 Duality Theory for the Minimization of Sums of Norms Problem

The minimization of the term $TV(\boldsymbol{\sigma}) = \sum_k |\mathbf{L}_k \boldsymbol{\sigma}|$, can be thought to be a Minimization of Sum of Norms problem (MSN) as $\sum_k |\mathbf{L}_k \boldsymbol{\sigma}| = \sum_k \|\mathbf{L}_k \boldsymbol{\sigma}\|$, and in this case important results for MSN problems can be applied.

The most general way of expressing the MSN problem is

$$\min_{\mathbf{y}} \sum_{i=1}^n \|A_i \mathbf{y} - \mathbf{c}_i\| \quad (10)$$

with $\mathbf{y} \in \mathbb{R}^m$; $\mathbf{c}_i \in \mathbb{R}^d$ and $A_i \in \mathbb{R}^{d \times m}$, which is equivalent to

$$(P) \quad \min_{\mathbf{y}} \left\{ \sum_{i=1}^n \|\mathbf{z}_i\| : A_i \mathbf{y} + \mathbf{z}_i = \mathbf{c}_i, \quad i = 1, \dots, n \right\} \quad (11)$$

with $\mathbf{z}_i \in \mathbb{R}^d$. We call (11) primal problem, and we label it (P). An equivalent problem to (P), which is called dual, and which is a maximization problem, can be obtained in the following way

$$\begin{aligned} \min_{\mathbf{y}: A_i \mathbf{y} + \mathbf{z}_i = \mathbf{c}_i} \sum_{i=1}^n \|\mathbf{z}_i\| &= \min_{\mathbf{y}: A_i \mathbf{y} + \mathbf{z}_i = \mathbf{c}_i} \max_{\mathbf{x}_i: \|\mathbf{x}_i\| \leq 1} \sum_{i=1}^n \mathbf{x}_i^T \mathbf{z}_i \\ &= \max_{\mathbf{x}_i: \|\mathbf{x}_i\| \leq 1} \min_{\mathbf{y}: A_i \mathbf{y} + \mathbf{z}_i = \mathbf{c}_i} \sum_{i=1}^n \mathbf{x}_i^T \mathbf{z}_i \\ &= \max_{\mathbf{x}_i: \|\mathbf{x}_i\| \leq 1} \min_{\mathbf{y} \in \mathbb{R}^m} \left(\sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i - \mathbf{y}^T \sum_{i=1}^n A_i^T \mathbf{x}_i \right) \\ &= \max_{\mathbf{x}_i} \left\{ \sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i : \|\mathbf{x}_i\| \leq 1; \sum_{i=1}^n A_i^T \mathbf{x}_i = 0 \right\} \end{aligned} \quad (12)$$

where the first equality follows from Cauchy–Schwartz, the second from min–max theory [21] [26], the third trivially, and the fourth because if $\sum_{i=1}^n A_i^T \mathbf{x}_i$ is not zero, the minimised value would be $-\infty$. The dual problem of (P) is therefore

$$(D) \quad \max_{\mathbf{x}_i} \left\{ \sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i : \|\mathbf{x}_i\| \leq 1; \sum_{i=1}^n A_i^T \mathbf{x}_i = 0, \quad i = 1, \dots, n \right\} \quad (13)$$

and the variables \mathbf{y} are called primal variables and the variables $\mathbf{x}_i \in \mathbb{R}^d$ 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

$$\mathcal{Y} = \left\{ (\mathbf{y}, \mathbf{z}) \in \mathbb{R}^m \times \mathbb{R}^{dn} : A\mathbf{y} + \mathbf{z} = \mathbf{c} \right\} \quad (14)$$

and the dual feasible region as

$$\mathcal{X} = \left\{ \mathbf{x} \in \mathbb{R}^{dn} : A^T \mathbf{x} = 0; \|\mathbf{x}_i\| \leq 1, \quad i = 1, \dots, n \right\} \quad (15)$$

where \mathbf{x} is obtained by stacking the vectors \mathbf{x}_i . Andersen *et al.* [21] have shown that for feasible points $(\mathbf{y}, \mathbf{z}) \in \mathcal{Y}$, $\mathbf{x} \in \mathcal{X}$

$$\sum_{i=1}^n \|\mathbf{z}_i\| - \sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i > 0 \quad (16)$$

and that for optimal points $(\mathbf{y}^*, \mathbf{z}^*) \in \mathcal{Y}$, $\mathbf{x}^* \in \mathcal{X}$

$$\sum_{i=1}^n \|\mathbf{z}_i^*\| - \sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i^* = 0 \quad (17)$$

In words: for feasible points the term $\sum_{i=1}^n \|\mathbf{z}_i\|$ is an upper bound to $\sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i$ and vice-versa. The difference $\sum_{i=1}^n \|\mathbf{z}_i\| - \sum_{i=1}^n \mathbf{c}_i^T \mathbf{x}_i = \sum_{i=1}^n (\|\mathbf{z}_i\| - \mathbf{x}_i^T \mathbf{z}_i)$ 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 $i = 1, \dots, n$, either $\|\mathbf{z}_i\|$ is zero or $\mathbf{x}_i = \mathbf{z}_i / \|\mathbf{z}_i\|$. This can be expressed conveniently in a form called *complementary condition*

$$\mathbf{z}_i - \|\mathbf{z}_i\| \mathbf{x}_i = 0, \quad i = 1, \dots, n \quad (18)$$

The complementary 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 (18) with the feasibility conditions (14) and (15) 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, complementary)

$$A\mathbf{y} + \mathbf{z} = \mathbf{c} \quad (19a)$$

$$A^T \mathbf{x} = 0 \quad (19b)$$

$$\mathbf{z}_i - \|\mathbf{z}_i\| \mathbf{x}_i = 0 \quad (19c)$$

The Newton Method cannot be applied in a straightforward manner to (19) as the complementary condition is not differentiable for $\|\mathbf{z}_i\| = 0$. Andersen *et al.* [21] suggest replacing it with the so called *centering condition*

$$\mathbf{z}_i - (\|\mathbf{z}_i\|^2 + \beta^2)^{\frac{1}{2}} \mathbf{x}_i = 0, \quad i = 1, \dots, n \quad (20)$$

where β is a small positive scalar parameter. Even if at first sight the centring condition is very similar to the smooth approximations that are generally used, where $TV(\boldsymbol{\sigma})$ is approximated with $\sum_k \sqrt{\|\mathbf{L}_k \boldsymbol{\sigma}\|^2 + \beta}$, it has different implications in this context. Particularly, it was shown in [17] that

the centring condition is the complementary condition of the following pair of smooth optimization problems

$$\begin{aligned}
 (P_\beta) \quad & \min \left\{ \sum_{i=1}^n (\|\mathbf{z}_i\|^2 + \beta^2)^{\frac{1}{2}} : (\mathbf{y}, \mathbf{z}) \in \mathcal{Y} \right\} \\
 (D_\beta) \quad & \min \left\{ \mathbf{c}^T \mathbf{x} + \beta \sum_{i=1}^n (1 - \|\mathbf{x}_i\|^2)^{\frac{1}{2}} : \mathbf{x} \in \mathcal{X} \right\}
 \end{aligned} \tag{21}$$

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 (19a), (19b), (20).

Introducing the perturbation β in the complementary condition for the original pair of problems is therefore equivalent to smoothing the norms in (P) and introducing a cost into (D). Particularly the cost function $\sum_{i=1}^n (1 - \|\mathbf{x}_i\|^2)^{\frac{1}{2}}$ can be understood to keep the dual solution away from its boundary ($\|\mathbf{x}_i\| = 1$), from which the name of centring condition for (20), 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) [27]. 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 [28] 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 $\beta > 0, \beta \rightarrow 0$. Using these results Andersen *et al.* realised an efficient PD-IPM algorithm that works maintaining feasibility conditions (19a), (19b) and applies the centering condition (20) with a centering parameter β which is reduced during iterations, following the central path to the optimal point.

In the next Section we describe the application of the PD-IPM framework to TV regularized linear inverse problems.

2.5 Duality for Tikhonov Regularized Inverse Problems

In inverse problems, with *linear* forward operators, the discretized TV regularized inverse problem, can be formulated as

$$(P) \quad \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \alpha \sum_k |\mathbf{L}_k \mathbf{x}| \tag{22}$$

where \mathbf{L} , as in (9), is a discretization of the gradient operator. We will label it as the primal problem (P). The dual problem, can be derived noting, as

for the MSN problem, that

$$|\mathbf{L}_k \mathbf{x}| = \|\mathbf{L}_k \mathbf{x}\| = \max_{\mathbf{y}: \|\mathbf{y}\| \leq 1} \mathbf{y}^T \mathbf{L}_k \mathbf{x} \quad (23)$$

By applying (23) to (P), the dual problem (D) is obtained as follows [23]

$$(D) \quad \max_{\mathbf{y}: \|\mathbf{y}_i\| \leq 1} \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \alpha \mathbf{y}^T \mathbf{L}\mathbf{x} \quad (24)$$

The optimization problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \alpha \mathbf{y}^T \mathbf{L}\mathbf{x} \quad (25)$$

has an optimal point defined by the first order conditions

$$\mathbf{A}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) + \alpha \mathbf{L}^T \mathbf{x} = \mathbf{0} \quad (26)$$

the dual problem can be written therefore as

$$(D) \quad \max_{\mathbf{y}: \|\mathbf{y}_i\| \leq 1} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \alpha \mathbf{y}^T \mathbf{L}\mathbf{x} \quad (27)$$

$$\mathbf{A}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) + \alpha \mathbf{L}^T \mathbf{x} = \mathbf{0}$$

The primal–dual gap for (P) and (D) is therefore:

$$\frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \alpha \sum_k |\mathbf{L}_k \mathbf{x}| - \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 - \alpha \mathbf{y}^T \mathbf{L}\mathbf{x} =$$

$$\alpha \left(\sum_k |\mathbf{L}_k \mathbf{x}| - \mathbf{y}^T \mathbf{L}\mathbf{x} \right) \quad (28)$$

The complementary condition, which nulls the primal–dual gap, for (22) and (27) is therefore:

$$\sum_k |\mathbf{L}_k \mathbf{x}| - \mathbf{y}^T \mathbf{L}\mathbf{x} = \mathbf{0} \quad (29)$$

which with the dual feasibility $\|\mathbf{y}_i\| \leq 1$ is equivalent to requiring that

$$\mathbf{L}_i \mathbf{x} - \mathbf{y}_i \|\mathbf{L}_i \mathbf{x}\| = 0 \quad i = 1, \dots, n \quad (30)$$

The PD-IPM framework for the TV regularized inverse problem can thus be written as

$$\|\mathbf{y}_i\| \leq 1 \quad i = 1, \dots, n \quad (31)$$

$$\mathbf{A}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) + \alpha \mathbf{L}^T \mathbf{x} = \mathbf{0} \quad (32)$$

$$\mathbf{L}_i \mathbf{x} - \mathbf{y}_i \|\mathbf{L}_i \mathbf{x}\| = 0 \quad i = 1, \dots, n \quad (33)$$

It is not possible to apply the Newton method directly to (31),(32),(33) as (33) is not differentiable for $\mathbf{L}_i \mathbf{x} = 0$. A centering condition has to be applied [21] [23], obtaining a smooth pair of optimization problems (P_β) and (D_β) and a central path parameterised by β . This is done by replacing $\mathbf{L}_i \mathbf{x}$ by $(\|\mathbf{L}_i \mathbf{x}\|^2 + \beta)^{\frac{1}{2}}$ in (33).

2.6 PD-IPM for EIT

The PD-IPM algorithm described in section (2.5) was developed by Chan *et al.* [22] for inverse problems with *linear* forward operators. We next describe a numerical implementation of the PD-IPM algorithm, based on [23], to calculate the non-linear solution to

$$\boldsymbol{\sigma}_{rec} = \arg \min_{\boldsymbol{\sigma}} \frac{1}{2} \|F(\boldsymbol{\sigma}) - \mathbf{V}_{meas}\|^2 + \alpha TV(\boldsymbol{\sigma}) \quad (34)$$

With a similar notation as used in Section 2.1. This is recognized as equation (1) with $G(\boldsymbol{\sigma}) = TV(\boldsymbol{\sigma})$. The system of *non-linear* equations that defines the PD-IPM method for (34) can be written as

$$\begin{aligned} \|\mathbf{y}_i\| &\leq 1 \\ \mathbf{J}^T (F(\boldsymbol{\sigma}) - \mathbf{V}_{meas}) + \alpha \mathbf{L}^T \boldsymbol{\sigma} &= \mathbf{0} \\ \mathbf{L}\boldsymbol{\sigma} - \mathbf{E}\mathbf{y} &= \mathbf{0} \end{aligned} \quad (35)$$

with \mathbf{E} a diagonal matrix defined by $\mathbf{E} = \text{diag} \left(\sqrt{\|\mathbf{L}_i \boldsymbol{\sigma}\|^2 + \beta} \right)$ and \mathbf{J} the Jacobian of the forward operator $F(\boldsymbol{\sigma})$. Newton's method can be applied to solve (35) obtaining the following system for the updates $\delta\boldsymbol{\sigma}$ and $\delta\mathbf{y}$ of the primal and dual variables

$$\begin{bmatrix} \mathbf{J}^T \mathbf{J} & \alpha \mathbf{L}^T \\ \mathbf{K} \mathbf{L} & -\mathbf{E} \end{bmatrix} \begin{bmatrix} \delta\boldsymbol{\sigma} \\ \delta\mathbf{y} \end{bmatrix} = - \begin{bmatrix} \mathbf{J}^T (F(\boldsymbol{\sigma}) - \mathbf{V}_{meas}) + \alpha \mathbf{L}^T \mathbf{y} \\ \mathbf{L}\boldsymbol{\sigma} - \mathbf{E}\mathbf{y} \end{bmatrix} \quad (36)$$

with

$$\mathbf{K} = \text{diag} \left(1 - \frac{\mathbf{y}_i \mathbf{L}_i \boldsymbol{\sigma}}{\mathbf{E}(i, i)} \right) \quad (37)$$

equation (36) can be solved as follows

$$[\mathbf{J}^T \mathbf{J} + \alpha \mathbf{L}^T \mathbf{E}^{-1} \mathbf{L}] \delta\boldsymbol{\sigma} = - [\mathbf{J}^T (F(\boldsymbol{\sigma}) - \mathbf{V}_{meas}) + \alpha \mathbf{L}^T \mathbf{E}^{-1} \mathbf{L}\boldsymbol{\sigma}] \quad (38)$$

$$\delta\mathbf{y} = -\mathbf{y} + \mathbf{E}^{-1} \mathbf{L}\boldsymbol{\sigma} + \mathbf{E} \mathbf{K} \mathbf{L} \delta\boldsymbol{\sigma} \quad (39)$$

Equations (38) and (39) can therefore be applied iteratively to solve the non-linear inversion (34). The iterative procedure must be initialized which is done by setting $\mathbf{y}_0 = \mathbf{0}$. Thus in the first iteration (38) is solved as

$$\delta\boldsymbol{\sigma} = (\mathbf{J}^T \mathbf{J} + \alpha \mathbf{L}^T \mathbf{L})^{-1} (\mathbf{J}^T (F(\boldsymbol{\sigma}) - \mathbf{V}_{meas})) \quad (40)$$

and $\delta\mathbf{y} = \mathbf{E}^{-1} \mathbf{L}\boldsymbol{\sigma} + \mathbf{E} \mathbf{K} \mathbf{L} \delta\boldsymbol{\sigma}$. This is recognizable as the first step of the 2-norm regularized inverse of equation (3).

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 modified dual objective function (D_β).

The simplest way to compute the update is called the *scaling rule* [21] which is defined to work as follows

$$\mathbf{y}^{(k+1)} = \varphi^* \left(\mathbf{y}^{(k)} + \delta \mathbf{y}^{(k)} \right) \quad (41)$$

where φ^* is a scalar value such that

$$\varphi^* = \sup \left\{ \varphi : \varphi \left\| \mathbf{y}_i^{(k)} + \delta \mathbf{y}_i^{(k)} \right\| \leq 1, \quad i = 1, \dots, n \right\} \quad (42)$$

An alternative way is to calculate the exact step length to the boundary, applying what is called the *step length rule* [21]

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \min(1, \varphi^*) \delta \mathbf{y}^{(k)} \quad (43)$$

where φ^* is a scalar value such that

$$\varphi^* = \sup \left\{ \varphi : \left\| \mathbf{y}_i^{(k)} + \varphi \delta \mathbf{y}_i^{(k)} \right\| \leq 1, \quad i = 1, \dots, n \right\} \quad (44)$$

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{y} 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 appropriate. 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.

3 Evaluation Procedure

A reconstruction algorithm that formulates the inverse problem as in equation (34) and solves it as in equations (38) and (39) was developed in the MATLAB environment. The method proposed by Chan *et al.* [14] to solve equation (35) assumes the forward operator to be linear. The reconstructions that we present in this section of the paper are fully non-linear, the algorithm is shown to work on the cases we used as tests, but we do not provide a proof of convergence.

Evaluation was performed by comparing the performance of the TV PD-IPM algorithm with that of the quadratic algorithm equation (3). In equation (3) $\mathbf{L} = \mathbf{R}_{HPF}$ where \mathbf{R}_{HPF} is the Gaussian spatial high pass filter

originally described in [6]. 2D Simulated data were computed on a 1024 element circular mesh using the two phantoms shown in figures 3(a) and 3(b).

Phantom A is a single “blocky” contrast with a conductivity of 0.90, phantom B consists of 2 “blocky” contrasts with conductivities of 0.90 and 1.10. Background conductivity was set at 1.0. 15 sets of reconstructions were made for each phantom and for each algorithm (TV regularization and ℓ^2 Gaussian regularization) with increasing amounts of simulated noise added. The 16 electrode adjacent protocol was used [6]. 2D reconstructions were performed on a 576 element circular mesh, not matching the mesh used for forward computations, in order to avoid what is referred as an *inverse crime* [24].

4 Results

4.1 Phantom 1

Figure 4(a) and 4(b) shows that after the first iterative step the TV and the quadratic solutions are similar. The resolution, in terms of blur radius, is slightly better for the TV solution, however visual inspection of figures 4(a) and 4(b) shows that the TV solution has more noise. Blur Radius (BR) is defined as a measure of the resolution: $BR = \sqrt{A_z/A_0}$ where A_0 is the area of the entire 2D medium and A_z is the area of the reconstructed contrast containing half the magnitude of the reconstructed image [6]. BR calculates the area fraction of the elements that contain 50% of the total image amplitude. We call this the half amplitude (HA) set.

The convergence behaviour of the two algorithms is illustrated in Figure 5 in which Residual Error, Total Variation, and Resolution are plotted against iteration number. Both the L^2 and L^1 solutions show steady decrease in Residual Error over the first 4 iterations. By the 5th iteration both solutions have converged in this measure.

3D visualizations of selected TV solutions are shown in Figure 7. The characteristic blocky structures of a TV solution start to emerge by the 3rd iteration as shown in Figure 7 and by the profile plots of figure 6. Visibly detectable improvements in the TV solution are impossible to detect after the 8th iteration with no appreciable changes in the total variation or in the reconstructed images. The profile plots of figure (6) show that the TV algorithm is able to reconstruct the profile of phantom almost exactly in the noise free case. Although the residual error of the L^2 solution decreases over the first 4 iterations the Resolution, shown in Figure 5(c), has peaked by the 5th iteration. Although not shown, the resulting L^2 images are visually similar. Figure 5(c) indicates that the resolution measure of Blur Radius is not a good indicator of TV image quality since the TV and visual images steadily improve while the Blur Radius decrease for the first 3 iterations

then increase until it has stabilized by the 8th iteration.

4.2 Noise Effects

Noise was added to the simulated data in 15 increments from 0 to a maximum standard deviation of 3% of the signal. Good images, such as those in figures 8(a) and 8(b) were produced by both algorithms for noise levels smaller than 0.6%. AWGN up to 1.0% produced TV images that by the 7th iteration were recognizable but had large noise artefacts.

TV reconstructions of data with more than 1.5% noise, such as figure illustrated in 9(b), were dominated by noise artefacts. The quadratic algorithm was more robust to noise with the best reconstructions occurring with the first step of the algorithm. As more iterations were used the quadratic reconstructions became corrupted by noise. However, the first step of the quadratic algorithm produced a relative good image quality with noise as high as 2.5%, see Figure 9(a).

4.3 Phantom 2

With low noise, the TV algorithm is able to recover a single blocky contrast almost exactly. With two contrasts the TV algorithm provides a reasonable reconstruction however it is unable to recover the profile as accurately as it does in the phantom A case. Figure 10 shows the profiles for the TV and L^2 algorithms while figures 11(a) and 11(b) show reconstructions from both algorithms for the 8th iteration.

4.4 Parameters

The PD-IPM method has two tuneable parameters β and λ . The value of β has a large effect on convergence. Too large a value of β (greater than 10^{-6}) prevented convergence. to the desired “blocky” solution; the solution stabilized but showed smoothed features that were not consistent with the edges obtained with smaller values of β . Ultimately it was determined that the quickest convergence occurred when β was initialized to a small value (we used 10^{-12}) and left unchanged. This was the method used in the results shown in this paper.

With an iterative algorithm multiple values of the regularization hyperparameter, λ , could be used for each iteration. In this work, for the TV algorithm, a different value was used for λ_0 , in the initialization step (40) and for λ_i in the iterative steps (38). λ_0 was selected using the *BestRes* method described in [25]. *BestRes* is an algorithm for objectively calculating the hyperparameter for linearized one-step EIT image reconstruction algorithms. This method suggests selecting a hyperparameter that results in a reconstruction that has maximum resolution for an impulse contrast. The PD-IPM algorithm did not show to be strongly sensitive to the value of

λ_0 . We varied the value of λ_0 three orders of magnitude above and one order of magnitude below λ_{BR} without appreciably changing the TV solution at convergence or the rate of convergence.

Although the initial hyperparameter, λ_0 , was always selected using *Best Res* several numerical experiments were performed to determine the effect of the iterative hyperparameter, λ_i , on algorithm performance. Although λ_i could be changed at each iteration, in the reconstructions shown in this manuscript λ_i was maintained constant, thus $\lambda_i = \lambda_{i+1}$. Figure 12 shows the results of running the algorithm to convergence six times with a different value λ_i for each run. It is obvious from the figure that the algorithm is sensitive to the value of λ_i ; too small a value of λ_i prevents a “blocky” solution, too large a value of λ_i will allow blocky reconstructions but suppress the amplitude. The *BestRes* method was originally used to calculate λ_i however the method was unable to find a good value for λ_i . Best results were obtained by the *ad hoc* visual inspection of figures such as figure 12 for various values of λ_i . Further work is required to develop an objective method to select λ_i .

The original PD-IPM methods includes updating the Jacobian matrix at each iteration. In our work numerical experiments this did not result always in a significant improvement in reconstructed images. We adopted therefore the arrangement of not updating the Jacobian at each single iteration. This provides a reduction in the reconstruction computational time.

As an additional numerical experiment, we evaluated the use of the same regularization matrix L as for TV regularization, (equation 9), with the quadratic algorithm (3). Although reconstructions from the first step were identical to TV reconstructions, the quadratic solutions rapidly degraded, producing noisy reconstructions that were dominated by noise artefacts after the 10th iteration. The TV prior is not recommended for use with the quadratic algorithm.

4.5 Preliminary testing in 3D

The generality of the PD-IPM scheme allows its use for the 3D EIT reconstructions. The method was expected to work equally well in three dimensions, and to be easily extended to this case. To validate this a single experiment with the simulated tank of figure (13) was performed. The tank has 315 nodes, 1104 elements, 32 electrodes and is constructed of 4 identical layers of tetrahedrons. A single contrast in the shape of a crescent was used to generate simulated data. Although an *inverse crime* we reconstructed the data on the same mesh, as the aim was simply of demonstrating that the PD-IPM framework could be used for TV

The convergence of the PD-IPM algorithm is shown in figure 14. Convergence occurred rapidly with a reasonable image appearing in the first iteration and convergence being achieved by the 8th iteration - there was was

no appreciable improvement in the image or change in the error norm after the 8th iteration. Figure 15 shows slices taken at the five layer boundaries (including top and bottom tank surfaces) of the simulated tank. Figure 15 shows reconstructed conductivities after the first iteration, Figure 16 shows reconstructed conductivities after 8 iterations.

The results were not as good as the results obtained from the 2D numerical experiments. This may be attributable to poor quality of the 3D model in terms of number of mesh elements. More work is required in order to properly evaluate the performance of PD-IPM in 3D.

5 Discussion and Conclusion

Practical results of the TV regularization and the efficiency of PD-IPM method are of interest in process and medical imaging. In this work we have demonstrated a practical implementation of a TV regularized reconstruction algorithm for EIT, and compared its performance to a traditional L^2 regularized reconstruction algorithm. Currently, TV regularized reconstruction are considerably more expensive to calculate than quadratic reconstructions, however the TV PD-IMP algorithm is able to compute non-smooth reconstructions in the presence of moderate noise, and it is therefore of practical use.

The typical number of iterations required by the TV PD-IPM algorithm for convergence, and thus for being able to show sharp profile in the reconstructed images, is in the order of 10 iterations. The quadratic algorithm produces good, albeit smooth, solutions in 1 to 3 steps. Thus there is a clear speed advantage in using the quadratic regularization. On a 1.8GHz AMD Turion 64 with 1GB ram, one step of the quadratic algorithm took 0.78 seconds for the 576 element 2D mesh, while one step of the PD-IPM algorithm took 0.86 seconds. Thus the TV solution at convergence takes about 9 seconds to calculate compared to the 1 to 2 seconds needed by the quadratic solution. The absolute difference in the computational times is obviously much more significant in 3D.

In our experiments we have found that the quadratic algorithm is slightly more robust to noise however both algorithms produce useful reconstructions at realistic noise levels. We feel that we will need to carry further work in better understanding the convergence behaviour of the PD-IPM algorithm, as to possibly reduce the number of iterations that this algorithm typically requires in order to converge, and thus to make it more competitive in terms of computational requirements. Application of the proposed method in 3D settings must be studied in more detail, with dense meshes and a number of inverse parameters representative of real applications.

References

- [1] D. C. Dobson, F. Santosa, An image enhancement technique for electrical impedance tomography. *Inverse Problems*, 10, 317-334, 1994.
- [2] W. R. Breckon, Image reconstruction in electrical impedance tomography. PhD Thesis, Oxford Polytechnic, 1990.
- [3] T. J. Yorkey, A quantitative comparison of the reconstruction algorithms used in impedance tomography. PhD Thesis, University of Wisconsin, 1986.
- [4] P. Hua, J. G. Webster, W. Tompkins, A regularized electrical impedance tomography reconstruction algorithm. *Clin. Phys. Physiol. Meas.*, 9, 137-141, 1988.
- [5] M. Cheney, D. Isaacson, J.C. Newell, S. Simake, J. Goble, NOSER: An algorithm for solving the inverse conductivity problem. *Int. J. Imag. Sys. Technol.*, 2, 66-75, 1990.
- [6] A. Adler, R. Guardo, Electrical impedance tomography: regularized imaging and contrast detection. *IEEE Trans. Med. Imaging*, 15, 170-179, 1996.
- [7] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, E. Somersalo, Construction of nonstandard smoothness priors. *Inverse Problems*, 15, 713-729, 1999.
- [8] A. Borsic, W. R. B. Lionheart, C. N. McLeod, Generation of anisotropic regularisation filters for Electrical Impedance Tomography. *IEEE Trans. Med. Imaging*, 21, 579- 587, 2002.
- [9] E. Somersalo, D. Isaacson, M. Cheney, Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.*, 52, 1023-40, 1992.
- [10] E. Somersalo, J. Kaipio, M. Vauhkonen, D. Baroudi, S. Järvenp, Impedance imaging and Markov chain Monte Carlo methods. *SPIE Proceedings Series*, 3171, 175-185, 1997.
- [11] V. Kolehmainen, Novel approaches to image reconstruction in diffusion tomography. PhD Thesis, Kuopio University, 2001.
- [12] J. P. Kaipio, V. Kolehmainen, E. Somersalo, M. Vauhkonen, Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography. *Inverse Problems*, 16, 1487-1522, 2000.
- [13] L. Rudin, S. J. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms. *Physica D*, 60, 259-268, 1992.

- [14] T. Chan, A. Marquina, P. Mulet P, High-order total variation-based image restoration. *SIAM Journal on Scientific Computing*, 19, 1046-1062, 1998.
- [15] C. Vogel, *Computational methods for inverse problem*. Philadelphia, SIAM, 2001.
- [16] K. D. Andersen, E. Christiansen, A Newton Barrier method for Minimizing a sum of Euclidean norms subject to linear equality constraints. Technical Report, Dept of Mathematics and Computer Science, Odense University, 1995.
- [17] K. D. Andersen, An efficient Newton Barrier method for minimizing a sum of Euclidean norms. *SIAM J. on Optimization*, 6, 74-95, 1996.
- [18] T. F. Coleman, Y. Li, A globally and quadratically convergent affine scaling method for linear ℓ_1 problems. *Mathematical Programming*, 56, 189-222, 1992.
- [19] I. Barrodale, F. D. K. Roberts, An efficient algorithm for discrete ℓ_1 linear approximation with linear constraints. *SIAM J. Numer. Anal.*, 15, 603-611, 1978.
- [20] G. Xue, Y. Ye, An efficient algorithm for minimizing a sum of Euclidean norms with applications. *SIAM Journal on Optimization*, 10, 551-579, 2000.
- [21] K. D. Andersen, E. Christiansen, A. R. Conn, 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.
- [22] T. Chan, G. H. Golub, P. Mulet, A nonlinear Primal-Dual method for Total Variation-based image restoration. UCLA Math Department CAM Report, 95-43, 1995.
- [23] A. Borsic, *Regularisation methods for imaging from electrical measurements*. PhD Thesis, Oxford Brooks University, 2002.
- [24] A. Wirgin, The inverse crime. *ArXiv Mathematical Physics e-prints arXiv:math-ph/0401050*, 2004.
- [25] B. M. Graham, A. Adler, Objective selection of hyperparameter for EIT. *Physio. Meas.*, 27, S65-S79, 2006.
- [26] R. T. Rockefellar, *Convex Analysis*, Princeton, Princeton University Press, 1970.
- [27] S. J. Wright, *Primal Dual Interior Point Methods*. Philadelphia, SIAM, 1997.

- [28] N. Karmarkar, A New Polynomial-Time Algorithm for Linear Programming. *Combinatorica*, 4, 373-395, 1984.
- [29] W. R. B. Lionheart, N. Polydorides, A. Borsic, The reconstruction Problem. In “Electrical Impedance Tomography, Methods, History and Applications”, 3-64, Bristol, Institute of Physics Publishing, 2005.
- [30] J. Wade, K. Senior, S. Seubert, Convergence of derivative approximations in the inverse conductivity problem. Bowling Green State University, Tech Rep. 96-14, 1996.

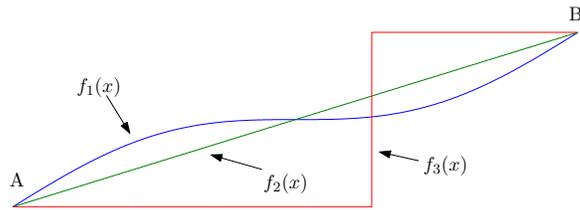


Figure 1: *Two points A and B can be connected by several paths. All of them have the same TV.*

PD-IPM Algorithm

$$\psi(\boldsymbol{\sigma}) = \frac{1}{2}\|F(\boldsymbol{\sigma}_0) - \mathbf{V}_{\text{meas}}\| + \alpha TV(\boldsymbol{\sigma})$$

```

find a homogeneous  $\boldsymbol{\sigma}_0$  to minimise  $\|F(\boldsymbol{\sigma}_0) - \mathbf{V}_{\text{meas}}\|$ ;
initialise dual variable  $\mathbf{x}$  to zero;
initialise primal variable  $\boldsymbol{\sigma}$  with one step of traditional quadratic
regularized inversion;
set initial  $\beta$ ;
k=0;
while (termination condition not met)
     $\delta\mathbf{V}_k = (F(\boldsymbol{\sigma}_k) - \mathbf{V}_{\text{meas}})$ ;
     $J_k = J(\boldsymbol{\sigma}_k)$ ;
     $E_k = \text{diag}(\sqrt{\|L_i \boldsymbol{\sigma}_k\|^2 + \beta})$ ;
     $K_k = \text{diag}(1 - \frac{\mathbf{y}_i L_i \boldsymbol{\sigma}}{E_k(i,i)})$ ;
     $\delta\boldsymbol{\sigma}_k = -[J^T J + \alpha L^T E_k^{-1} K_k L]^{-1} J_k^T \delta\mathbf{V}_k + \alpha L^T E_k^{-1} L \boldsymbol{\sigma}_k$ ;
     $\delta\mathbf{y}_k = \mathbf{y}_k + E_k^{-1} L \boldsymbol{\sigma}_k + E_k^{-1} K_k L \delta\boldsymbol{\sigma}_k$ ;
     $\lambda_{\boldsymbol{\sigma}} = \text{argmin} \psi(\boldsymbol{\sigma}_k + \lambda_{\boldsymbol{\sigma}} \delta\boldsymbol{\sigma}_k)$ ;
     $\lambda_{\mathbf{y}} = \max\{\lambda_{\mathbf{y}} : \|y_i + \lambda_{\mathbf{y}} \delta y_i\| \leq 1, i = 1, \dots, n\}$ ;
    if a reduction of primal objective function has been achieved
         $\boldsymbol{\sigma}_{k+1} = \boldsymbol{\sigma}_k + \lambda_{\boldsymbol{\sigma}} \delta\boldsymbol{\sigma}_k$ ;
         $\mathbf{y}_{k+1} = \mathbf{y}_k + \min(1, \lambda_{\mathbf{y}}) \delta\mathbf{y}_k$ ;
        decrease  $\beta$  by a factor  $\beta_{\text{reduction}}$ ;
        decrease  $\beta_{\text{reduction}}$ ;
    else
        increase  $\beta$ ;
    end if else
k=k+1; evaluate termination condition;
end while

```

Figure 2: Pseudo code for the PD-IPM algorithm with continuation on β , line search on $\boldsymbol{\sigma}$ and dual steplength rule on \mathbf{y} .

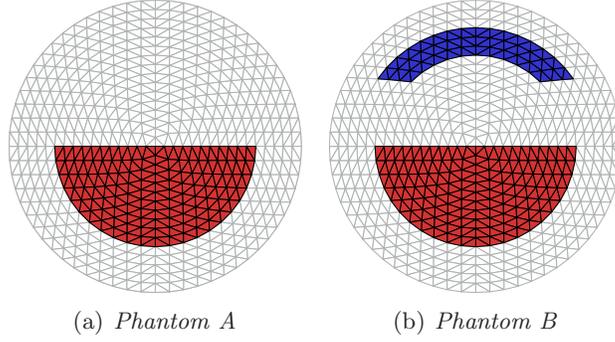


Figure 3: *2D Phantom contrasts on a 1024 element mesh, used to generate simulated data using 16 electrode adjacent current injection protocol.*

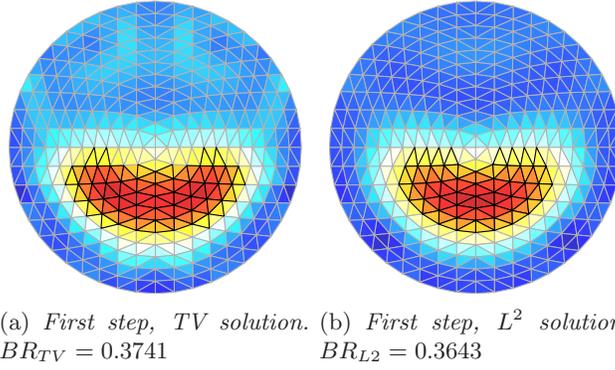


Figure 4: *Black bordered triangles are elements of the HA set. No noise added.*

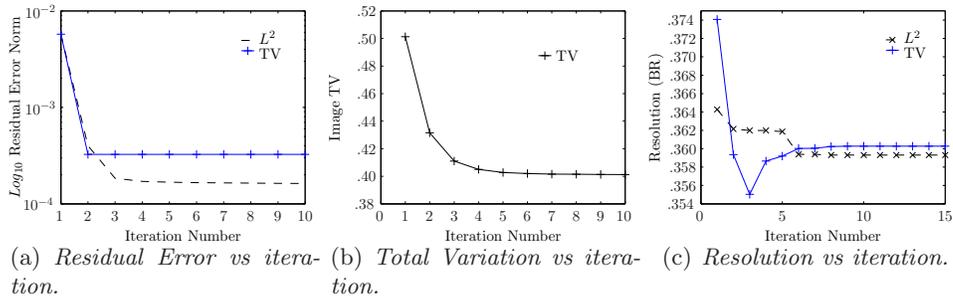


Figure 5: *Convergence Behaviour of Algorithms. No Noise added.*

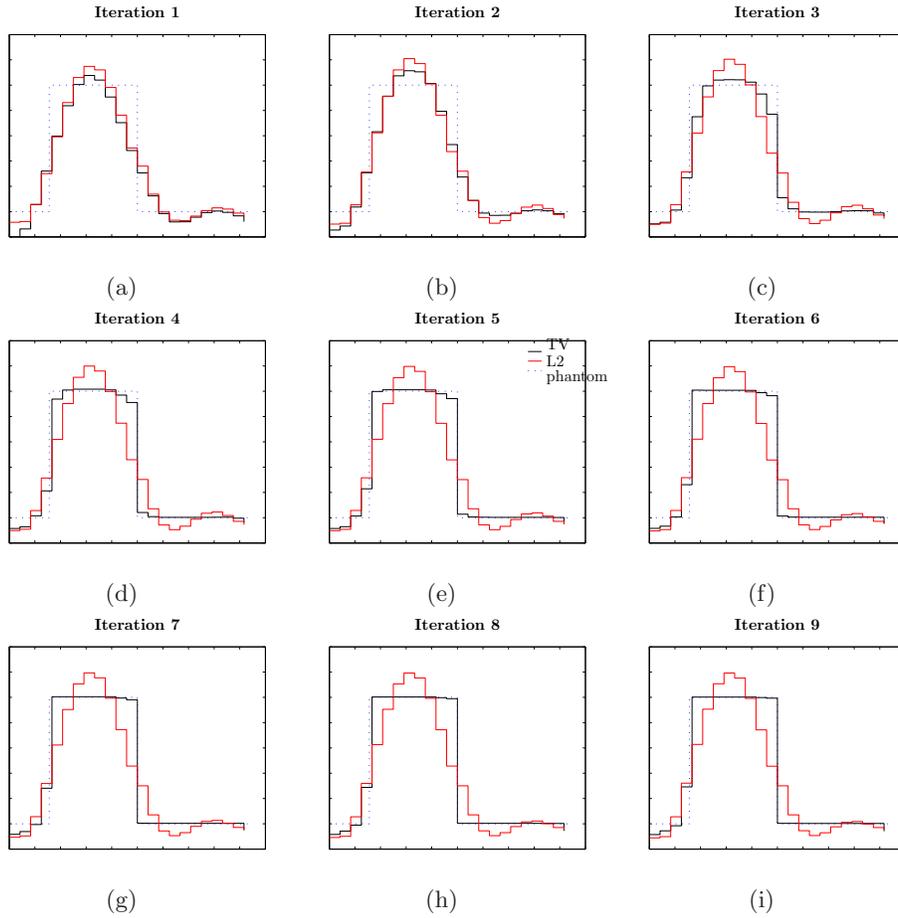


Figure 6: Profile plots of the originating contrast, TV , and L^2 reconstructions. No Noise added. Profiles are vertical slices through the middle of the reconstructed image.

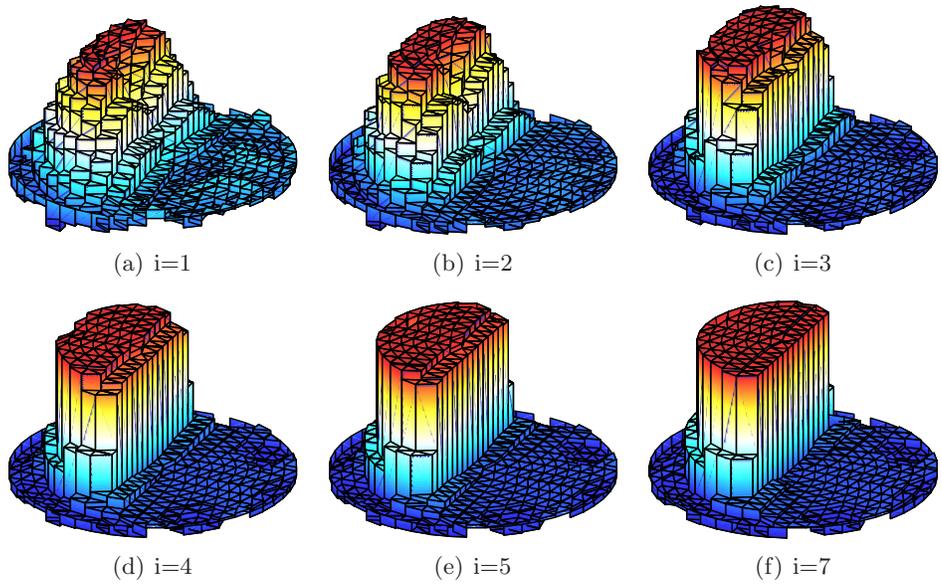


Figure 7: *TV reconstructions of Phantom A at increasing iterations. Vertical axis is absolute conductivity. Normalized to 0. No Noise added.*

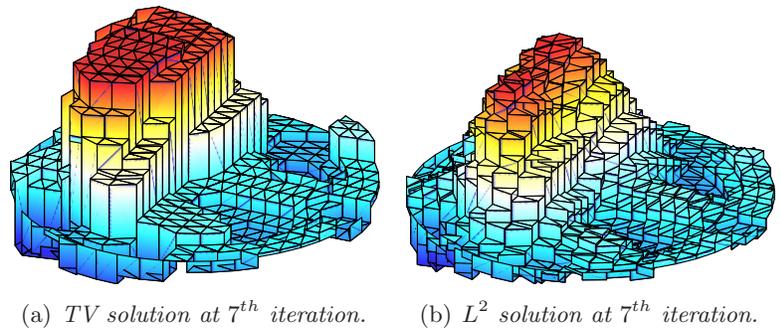
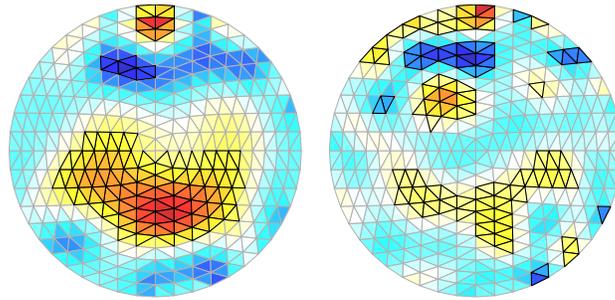


Figure 8: *Reconstructions of Phantom A with 0.6% AWGN.*



(a) L^2 solution with 2.5% AWGN, first step. Noisy but useful reconstruction. (b) TV solution with 1.5% AWGN, first step. Noise dominated solution.

Figure 9: Reconstructions of Phantom A.

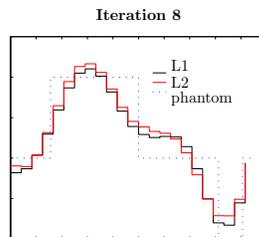
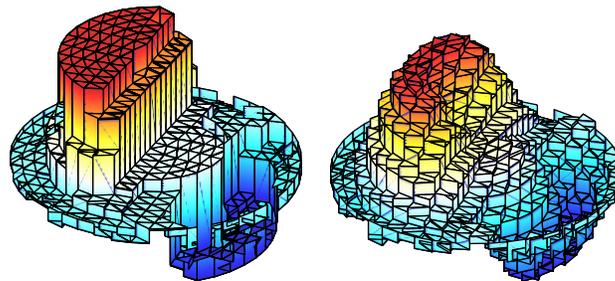


Figure 10: Phantom B profiles.



(a) TV solution at 8th iteration (b) L^2 solution at 8th iteration

Figure 11: Reconstructions of Phantom B with 2.5% AWGN.

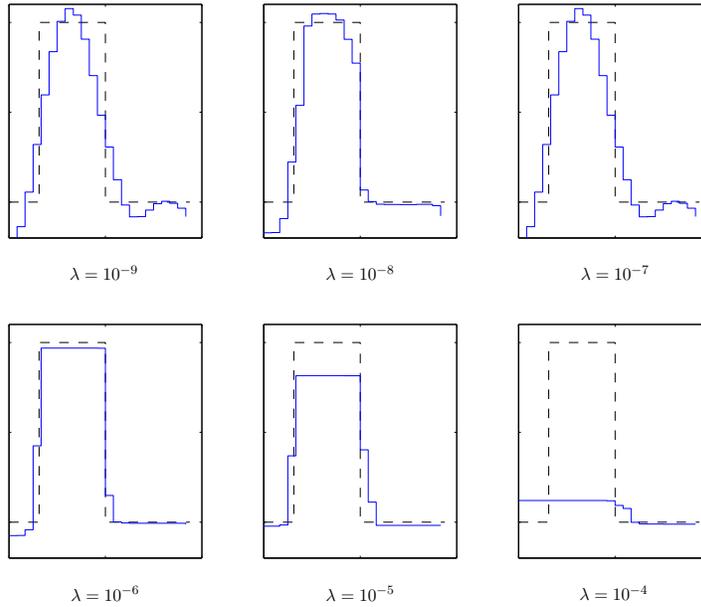


Figure 12: Profiles of TV solutions at the 7th iteration (convergence). Showing effect of using different λ_i values in equation (38). Dotted line is generating contrast, solid line is TV solution. $\lambda_i \in [10^{-9}, 10^{-4}]$

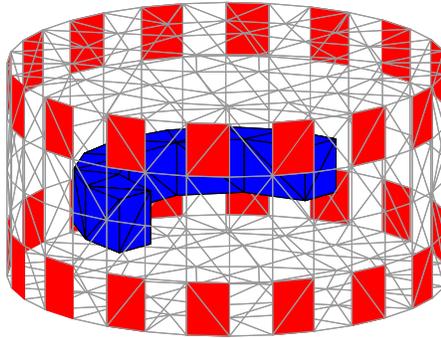


Figure 13: Four layer tank used for 3D reconstructions. Red patches are the 32 electrodes in 2 layers. Phantom contrast are the blue elements which are only in the second layer (between $z=1$ and $z=2$).

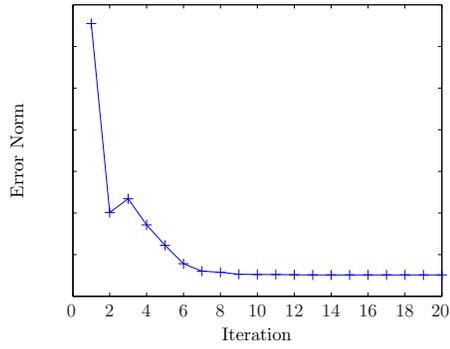


Figure 14: *Convergence of 3D PD-IPM algorithm.*

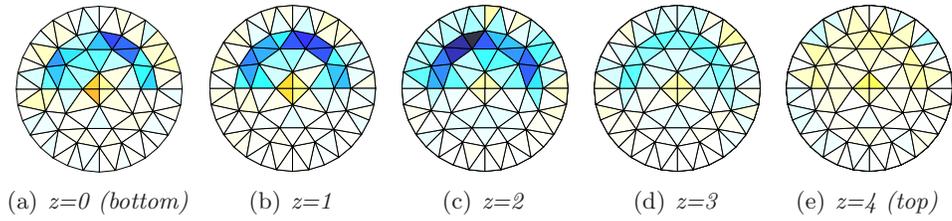


Figure 15: *Slices of 3D reconstructions for Iteration 1. No noise added.*

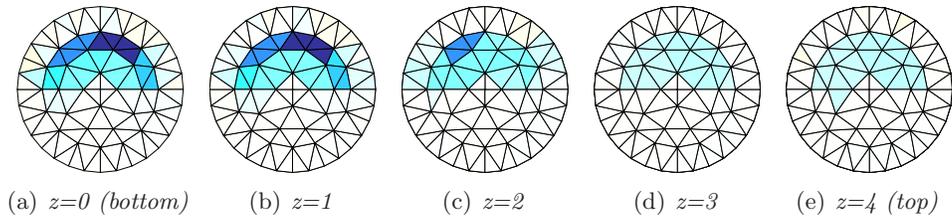


Figure 16: *Slices of 3D reconstructions for Iteration 8. No noise added.*