Geophysical Applications of Electrical Impedance Tomography

by

Alistair James Scott Boyle, PEng, MASc, Carleton University

A thesis submitted to the

Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Ph.D. Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario, Canada, K1S 5B6

The undersigned recommend to the

Faculty of Graduate Studies and Research

acceptance of the thesis

Geophysical Applications of Electrical Impedance Tomography

Submitted by Alistair James Scott Boyle, PEng, MASc, Carleton University

in partial fulfillment of the requirements for the degree of

Ph.D. Electrical and Computer Engineering

———————————————

Chair, Department of Systems and Computer Engineering

Dr. Yvan Labiche

———————————————

Thesis Supervisor, Carleton University

Dr. Andy Adler

———————————————

Thesis Co-Supervisor, British Geological Survey

Dr. Paul Wilkinson

Apr 2016, Carleton University

Abstract

Impedance imaging is a technique where stimulus currents are applied through electrodes to a body or the ground and measurements of the potential at other electrodes are collected. The data, along with any available prior information, are used to reconstruct an image of the conductivity distribution throughout the interior which provides diagnostic, cost effective information upon which decisions can be based for a broad array of geophysics, biomedical and industrial applications. The same technique is known as (biomedical) Electrical Impedance Tomography (EIT) and (geophysics) Electrical Resistivity Tomography (ERT). New geophysical applications have arisen for the automated monitoring of slope stability risks for natural landslides, transport embankments and cuttings, mine tailings dams and piles, and remote infrastructure in changing climatic environments. When impedance imaging is used in challenging scenarios, image quality can suffer unless sources of data error and instability can be addressed. This work develops computational techniques to address the issue of data set stability under adverse measurement conditions and builds practical implementations that demonstrate the effectiveness of our approach. We seek to achieve images with fewer artifacts and better detectability through improved methods for addressing boundary movement which permit the use of this technology on unstable surfaces where the positions of electrodes can change over time. Processes are developed for evaluating the correctness of an implementation and the overall validity of reconstructed images. Results demonstrated by adapting well understood strategies show improved reconstruction quality for simulated and measured geophysics data sets.

Acknowledgements

Thank you to the many people who made this work such a rewarding experience: Dr Andy Adler of Carleton University, my supervisor, for his friendship, guidance, inspiration and assistance; Dr Paul Wilkinson of the British Geological Survey, my co-supervisor, for his support, insightful comments and enlightening discussion; the geophysical tomography group at the British Geological Survey and Dr Jonathan Chambers in particular for their openness and willingness to share data and collaborate; my coworkers, for allowing me the flexibility in working hours to hold down both a professional and academic career; my peers at university, for their camaraderie, encouragement, and collaboration; our dog, for keeping my feet warm on those long nights; and finally, to my family and particularly my wife for their understanding, unflinching support of this endeavour, the many hours spent re-reading and correcting, and for the patient renegotiation of "working hours."

> *"[....] Unfortunately practice generally precedes theory, and it is the usual fate of mankind to get things done in some boggling way first, and find out afterward how they could have been done much more easily and perfectly."*
>
> — C. S. Peirce (1882)
>
> "Introductory Lecture on the Study of Logic"
>
> a lecture introducing experimental design

To my wife,

without whose love, support and encouragement

this work, as with the last, would not have been possible

To my children,

who bring joy, laughter, and love into every room

For my parents,

who are, no doubt, bursting with pride

Contents

II   Contributions   80

5   Problems with Inverse Problems   81

6   Data Quality and Model Mismatch   104

Appendices   191

List of Figures

List of Tables

List of Algorithms

List of Acronyms & Terminology

AGWN   Additive Gaussian White Noise

PCG   Preconditioned Conjugate Gradient

PCA   Principle Component Analysis

BEM   Boundary Element Method

CEM   Complete Electrode Model

EIDORS Electrical Impedance and Diffuse Optics Reconstruction Software

EIT   Electrical Impedance Tomography

ERT   Electrical Resistivity Tomography

FEM   Finite Element Method

GCV   Generalized Cross-Validation

PD-IPM Primal-Dual Interior Point Method

PDE   Partial Differential Equation

PEM   Point Electrode Model

RTK GPS Real-Time Kinematic Global Positioning System

SNR Signal-to-Noise Ratio

SVD Singular Value Decomposition

TSVD Truncated Singular Value Decomposition

TV Total Variation

in situ in place (latin), in its natural setting

ad hoc for this (latin), a solution for a specific task, non-generalizable, improvised

List of Symbols

Notational Convention

a scalar

$\mathbf{a}$ vector $[a_1\ a_2\ a_3]^{\mathsf{T}}$

$\mathbf{A}$ matrix $[\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3]$

$\mathbf{A}^{\mathsf{T}}$ matrix transpose

$\mathbf{A}^{*}$ matrix complex conjugate transpose $\mathbf{A}^{*} = \mathbf{A}^{\mathsf{T}}$ when $\mathbf{A} \in \mathbb{R}$

$\mathbf{A}^{\dagger}$ Moore-Penrose pseudoinverse $\mathbf{A}^{\dagger} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}$

$\mathbf{A} \cdot \mathbf{B}$ dot product

$\mathbf{A} \times \mathbf{B}$ cross product

$\mathbf{J}$ Jacobian, first partial derivatives

$\mathbf{H}$ Hessian, second partial derivatives

$\int_{\Omega} x$ integral of x over the volume Ω

$\oint_{\partial\Omega} x$ surface integral of x on the surface $\partial\Omega$

Notational Convention *(continued)*

$\frac{\partial}{\partial x}$ partial derivative with respect to x also ∂_x

i   imaginary number, $\sqrt{-1}$

∇u gradient operator $\frac{\partial u}{\partial x}$

$\nabla \cdot \mathbf{A}$ divergence operator

$\nabla \times \mathbf{A}$ curl operator

$\nabla^2 u$ Laplacian operator $\frac{\partial^2 u}{\partial x_i \partial x_j}$

$\mathcal{F}$ non-linear model or function

$\mathcal{F}(\square, \mathbf{x})$ parametrized about variable(s) $\square$ (box) with fixed variables $\mathbf{x}$

Statistics Symbols

μ mean

σ^2 variance

σ standard deviation

Electrical Symbols

σ conductivity, reciprocal of resistivity $1/\rho$ [S/m = $1/\Omega$·m]

ε permittivity [S·s/m·rad]

γ admittivity [S/m] ($\sigma + \mathrm{i}\omega\varepsilon$) at angular frequency ω

ϕ potential [V]

Electrical Symbols *(continued)*

ω angular frequency [rad/s], where frequency f is $\omega = 2\pi f$

$\mathbf{E}$ electric field [V/m = N/C]

$\mathbf{H}$ magnetic field [A/m]

$\mathbf{D}$ electric flux density [C/m^2 = N/V·m^2]

$\mathbf{B}$ magnetic flux density [T = V·s/m^2]

I current [A]

$\bar{\mathbf{J}}$ current density [A/m^2]

ρ charge density [C/m^3]

Ω domain, the region of interest

$\partial\Omega$ boundary of the domain Ω

$\hat{n}$ unit vector, outward normal from the boundary $\partial\Omega$

N FEM expansion (basis, shape) function

$\mathbf{x}$ voltages everywhere on the domain Ω

$\mathbf{b}$ boundary currents

$\mathbf{A}$ FEM system matrix

$\mathbf{R}$ regularization matrix

$\mathbf{W}$ inverse measurement covariance matrix $\sum_n^{-1} = \mathbf{W}$

$\mathcal{T}$ measurement selection matrix $\mathbf{v} = \mathcal{T}\mathbf{x}$

Geophysics Symbols

ϕ pore volume fraction (porosity) [%]

S saturation fraction [%]

n saturation exponent

a tortuosity factor

m cementation exponent

F formation factor

I resistivity index

B average ion mobility

Q_v cation concentration per unit pore volume

Constants

μ_0 magnetic constant, permeability of free space [H/m = N/A^2]

ϵ_0 electric constant, permittivity of free space [F/m]

Part I

Background

Chapter 1

Introduction

Electrical Impedance Tomography (EIT) is an imaging technique where stimulus currents are applied through electrodes to a body or the ground and measurements of the potential at other electrodes are collected. The data, along with any available prior information, are used to reconstruct an image of the conductivity distribution throughout the interior which provides diagnostic, cost-effective information upon which decisions can be based for a broad array of geophysics, biomedical and industrial applications.

An image is interpreted by the observer. Perceived objects may represent true or false features. Similarly, the background may represent a true absence of features or an insensitivity of processing method or measurement technique. These four observation outcomes may be arranged into a "confusion matrix" (Figure 1.1). An artifact is an apparent structural detail caused by the processing or capture of the data. This thesis focuses on converting sources of error, expressed as reconstruction artifacts, into useful information. By accounting for the source of error within the
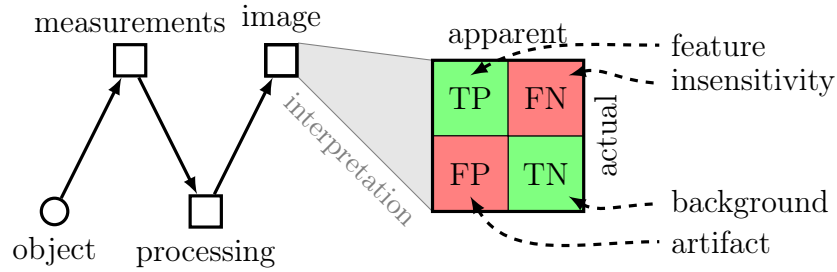
Figure 1.1: Confusion matrix; an object is measured, and the data is processed into an image; measuring and processing may introduce inaccuracies

reconstruction, the occurrence and magnitude of reconstruction artifacts caused by those former errors may be reduced.

1.1 Problem

Whatever the application, the capture and processing of EIT data with the goal of maximizing useful information content immediately encounters two issues: determining what portions of a reconstruction represent structure or artifact and then, having identified an artifact, how to minimize modelling errors which result in misleading artifacts of a particular type. Achieving a good outcome on these two issues helps in correctly interpreting the data (Kolehmainen et al., 1997; Hartinger et al., 2007). In biomedical applications, these issues have typically been addressed by using "complete" data measurements on a fixed number of symmetrically placed electrodes. A difference in conductivity is then reconstructed between sets of measurements; the difference reconstruction alleviates many of the immediate causes of reconstruction artifacts by reducing errors common to the two sets of measurements. Difference reconstruction generally achieves improved artifact performance at the cost of potentially useful constant components of the boundary or conductivity information.

When Electrical Resistivity Tomography (ERT)[1] is applied to the ground, absolute reconstructions from a single set of data are typically of keen interest because the constants discarded in the difference solution contain valuable information. The absolute reconstruction method exacerbates issues with data collection and modelling errors because errors common to the measurements must now be accounted for in the reconstruction.

As EIT/ERT is used in more complex scenarios than ones for which stimulus strategies and image reconstruction algorithms were originally intended, the quality and robustness of the images decreases. Techniques that are capable of adapting well understood strategies to achieve optimal robustness and quality may improve decision outcomes by supporting accurate image interpretation for the automated monitoring of slope stability risks for natural landslides, transport embankments and cuttings, mine tailings dams and piles, and remote infrastructure in a changing climatic environment. In an environment where climate change is affecting Canada's north, infrastructure built on permafrost is at risk. Technologies such as impedance imaging may provide cost-effective monitoring solutions for detecting and potentially predicting changes in the stability of the subsurface underlying such remote infrastructure.

1.2 Goal

Our goal is to develop computational techniques to address the issue of data set stability under adverse measurement conditions and build practical implementations

---

[1]In the biomedical literature, the term EIT is preferred, while in the geophysical literature, ERT denotes the same technique.

that demonstrate the effectiveness of our approach. We seek to achieve images with fewer artifacts and better detectability through improved methods for addressing boundary movement which permit the use of this technology in challenging new environments (Figure 1.2).
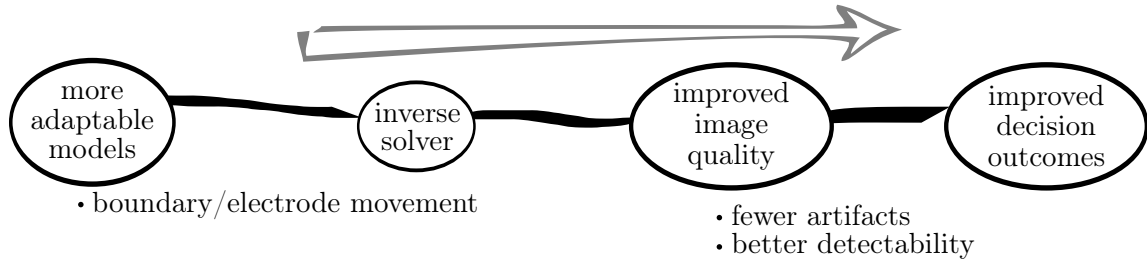
Figure 1.2: The Goal; improve decision outcomes

1.3 Contributions

The following contributions are applicable to EIT and similar inverse problems:

- a framework for checking inherent algorithm behaviours for inverse problems and for evaluating the quality of input models and data,

- a method for reducing electrode location modelling errors, or conversely, estimating electrode movement relative to a set of known locations,

- an application of our framework to a new absolute iterative Gauss-Newton algorithm and its implementation, demonstrated using data collected at a flooded mine site in Pont-Péan, France,

- an application of our framework to four nominally equivalent methods for calculating the electrode movement Jacobian, including a new Fréchet derivative method for electrode movement, and

- an application of the electrode movement estimation technique to simultaneously reconstruct electrode movement and conductivity for a slow-moving landslide in Yorkshire, UK.

1.3.1 Problems with Inverse Problems

Inverse problems can lead to inherently low resolution images which can be difficult to verify. Identifying incorrect reconstructions is difficult as they tend not to present clear indications of failure but rather introduce image artifacts which may lead to misinterpretation and incorrect decisions based on these interpretations.

We report our processes and the techniques used to find issues in the specific context of impedance imaging and more generally, for inverse problems (Figure 1.3). We focus on the algorithmic aspects: the challenges in validating inverse problem codes. This represents early work on a novel aspect of inverse problems and is related to the various theories of software and algorithm debugging, defect and root cause analysis, and business process analysis. We attempt to quantify and make repeatable processes which are typically heuristic and *ad-hoc*. Examples are presented from the author's implementation of a new absolute iterative Gauss-Newton solver for impedance imaging problems.

1.3.2 Data Quality and Model Mismatch

In inverse problems, procedures for evaluating data quality inherently involve some form of model correlation. We construe "data" to include not only the voltage measurements but also the stimulus currents, the stimulus and measurement sequence, the geometry near the region of interest, and the location of each electrode. Error
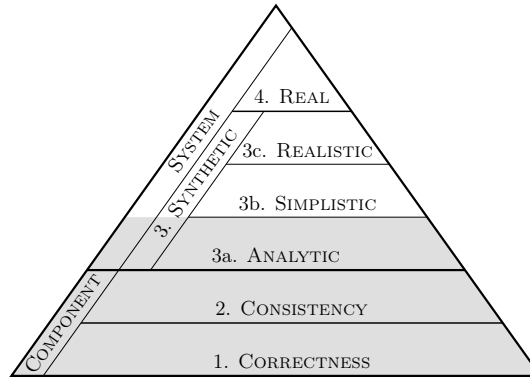
Figure 1.3: Proposed hierarchy of validation; proposed in this thesis, internal checks of components and system (grey background ▢) form a trusted platform upon which the input data may be compared to expectations (white background ▢)

estimates for all values are important but at times missing from the available data set. Identifying and then acting on the quality of relevant data is important for reducing reconstruction artifacts.

Overall, we are motivated by the desire to compare observed and expected information. First we advise comparing the measurement data to the output of an appropriate model. Then we outline a number of additional techniques for identifying problematic data. We illustrate these quality assessment techniques using data collected from a flooded historical mine site at Pont-Péan, France (Figure 1.4).

1.3.3 Electrode Movement and Modelling Errors

Prior work has shown that errors in boundary shape and electrode placement in the forward EIT problem can lead to significant artifacts (Adler et al., 1996; Zhou and Dhalin, 2003) when the errors are not conformal deformations of the correct shape and position (Boyle and Adler, 2011; Boyle et al., 2012a). To reconstruct electrode movement, a movement Jacobian must be constructed to calculate the first derivative

(a) artifacts dominate (b) corrected reconstruction
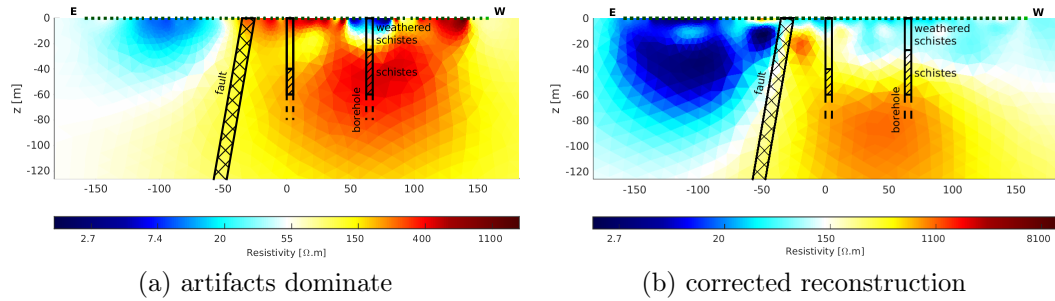
Figure 1.4: Pont-Péan reconstructed conductivity; (a) reconstructed conductivity shows little agreement with geological data due to incorrect assumed stimulus sequence, (b) reconstructed resistivity for correct stimulus sequence, electrode positions (green) on the top surface running West (W) to East (E), overlaid with known borehole and fault information

of electrode movement with respect to the measurements.

In biomedical two-dimensional reconstructions, a Finite Element Method (FEM) based electrode perturbation scheme has been used (Soleimani et al., 2006) with an optimization that performs a single-step matrix rank-one update (Gómez-Laberge and Adler, 2008) as part of a single-step Gauss-Newton solution. These single-step solutions are not appropriate for absolute reconstructions where the solution space is nonlinear, and a reasonable reconstruction requires an iterative solution to accurately report arbitrary conductivity discontinuities and electrode movement (Figure 1.5).

This work evaluates four methods for estimating the movement Jacobian using a simplified four electrode homogeneous half-space model. These methods are nominally equivalent, but we show that they have significant differences within the range of expected real-world parameter variation. We illustrate the computational complexity of the algorithms to understand which algorithm is the most appropriate and under what conditions each may be a "safe" choice.
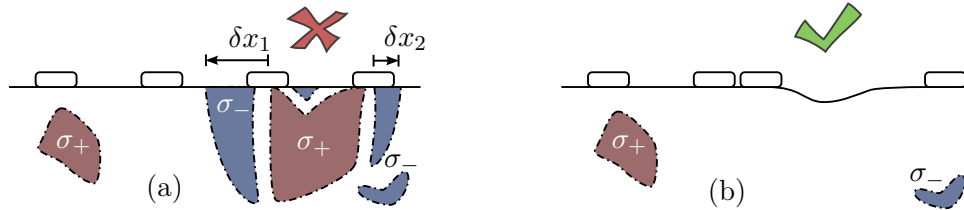
Figure 1.5: Movement artifacts in a conductivty σ reconstruction (a) may dominate an image, yet (b) contain useful information about, for example, electrode movement δx

1.3.4 Reconstructing Surface Movement

The absolute iterative electrode movement reconstruction method has been applied to a slow moving landslide located in Yorkshire, UK. The slope has moved by as much as two metres per year, typically during the spring wet season. The location is a test site for monitoring of long-term slope stability by a solar-powered remote EIT system (ALERT) (Kuras et al., 2009; Wilkinson et al., 2010). Data were recorded daily from 2008 and communicated to a central control site. The electrode movement and conductivity were jointly estimated throughout the seasons as the ground conductivity of the mudstone, clay and sandstone/ironstone formations varied over time due to water saturation changes and slope movement (Figure 1.6).

Electrode movement was reconstructed with reasonable accuracy when compared to Real-Time Kinematic Global Positioning System (RTK GPS) measurements of the initial and final electrode locations over the 2008–2009 season and on a nearby electrode array over the 2013–2014 season. Incorporating electrode movement significantly reduced artifacts observed in the image as identified by comparing conductivity-only reconstructions with electrode placement errors to reconstructions using the true electrode positions. Previous techniques for joint conductivity and movement reconstruction were limited to small electrode movements using a single-

step Gauss-Newton solution. This work extends this joint conductivity and movement technique to large electrode movements (our data exhibit movements of up to 32% of electrode spacing) and conductivity contrasts supporting applications where this joint inversion technique was not previously applicable. Correcting for electrode movement significantly reduced conductivity artifacts and achieved an acceptable degree of agreement with true electrode movements, on the order of 4% of electrode spacing (0.20 m).
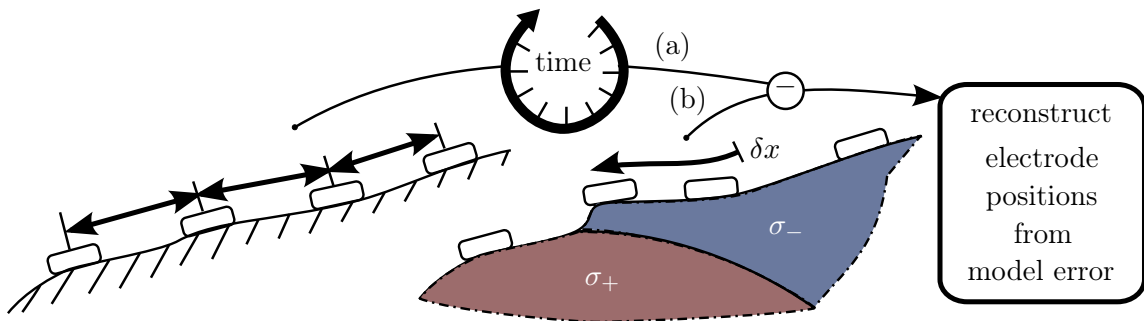
Figure 1.6: Data from a slow-moving landslide; (a) known but dated electrode location information is used to (b) reconstruct the absolute conductivity σ and an estimate of electrode ⌣ movement δx; a movement threshold could be established to alert operators to a potentially dangerous slope movement

1.4 Publications

Publications based on the work presented in this thesis are as follows.

- A. Boyle, M. Jehl, M. Crabb, and A. Adler. Estimating electrode movement in two dimension. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015a.

- A. Adler, A. Boyle, M. Crabb, H. Gagnon, B. Grychtol, N. Lesparre, and W. R. B. Lionheart. EIDORS Version 3.8. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015.

- A. Boyle, P. Wilkinson, J. Chambers, N. Lesparre, and A. Adler. Slope stability monitoring through impedance imaging. In *15th Conf. Electrical Impedance Tomography*, Gananoque, Canada, April 2014.

The current work builds on the author's prior contributions regarding conformal movements in two-dimensional EIT and the computational challenges in EIT.

- A. Boyle, A. Adler, and W. R. B. Lionheart. Shape deformation in two-dimensional electrical impedance tomography. *IEEE Trans. Med. Imag.*, 31(12):2185–2193, December 2012b.

- A. Boyle, A. Borsic, and A. Adler. Addressing the computational cost of large EIT solutions. *Physiol. Meas.*, 33(5):787–800, April 2012a.

- A. Boyle and A. Adler. Impact of electrode area, contact impedance and boundary shape on EIT images. *Physiol. Meas.*, 32(7):745–754, June 2011.

- A. Boyle, A. Adler, and A. Borsic. Scaling the EIT problem. In *12th Conf. Electrical Impedance Tomography*, Univ. of Bath, Bath, UK, May 2011.

- A. Boyle. The Effect of Boundary Shape Deformation on Two-Dimensional Electrical Impedance Tomography. Master's thesis, Carleton University, Ottawa, Canada, 2010.

While working on this thesis, the author was also involved in related work which resulted in the following publications.

- N. Lesparre, A. Boyle, B. Grychtol, J. Cabrerra, J. Marteau, and A. Adler. Electrical resistivity imaging in transmission between the surface and underground tunnel for fault characterization. *J. Applied Geophys.*, 2016.

- A. Boyle, Y. Mamatjan, and A. Adler. Selection of stimulus and measurement schemes. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015b.

- N. Lesparre, B. Grychtol, A. Boyle, and A. Adler. Consideration of inverse problem parametrization meshing for electrical resistivity tomography (submitted). *Sci. and Tech.* 2014.

- C. Jones, B. Grychtol, Herve Gagnon, A. Boyle, A. Adler C. He, and P. Gaggero. Open electrical impedance tomography (OEIT) file format. In *15th Conf. Electrical Impedance Tomography*, Gananoque, Canada, April 2014.

- A. Boyle, Y. Mamatjan, and A. Adler. Robust stimulation and measurement patterns in biomedical EIT. In *Workshop on 100 Years of Electrical Imaging*, Paris, France, July 2012c.

1.5 Conclusion

This work presents four contributions in the field of impedance imaging. First, methods for confirming the correct behaviour of an algorithm implementation. Second, methods for evaluating measurement quality and modelling errors. Third, an evaluation of four methods of estimating the effect of an electrode that was misplaced or moved. Lastly, a method of reconstructing large electrode movements and conductivity discontinuities. These contributions aim to improve the identification and reduce the occurance of image artifacts which will allow the application of impedance imaging to challenging new environments.

Chapter 2

Impedance Imaging

Impedance imaging, as it was originally conceived, was a technique for using electricity to see into the Earth with applications in mining and later, oil exploration (Allaud and Martin, 1977). Developments in math, computing and electronics have had a strong impact on the experimental success of impedance imaging techniques over the intervening years (Cheney et al., 1999). The mathematical model of a layered, isotropic earth (Slichter, 1933; Langer, 1933, 1936) has evolved into a flexible FEM-based scheme that can handle a broader range of geometries (Strang and Fix, 1973; Schöberl, 1997; Adler and Lionheart, 2006) (Figure 2.1). The availability of digital computers and mathematical tools for solving ill-posed nonlinear problems have enabled sophisticated techniques for inferring underlying conductivity distributions from limited information (Borcea, 2002; Mueller et al., 2002). At the same time, the collection of empirical evidence for measuring the performance of these tools remains little changed since the early experiments of Schlumberger in which a bathtub was filled with clay as a model for field electrical measurements. The techniques involved

(a) two-layer half-space (c.1936)   (b) FEM-based reconstruction (c.2016)
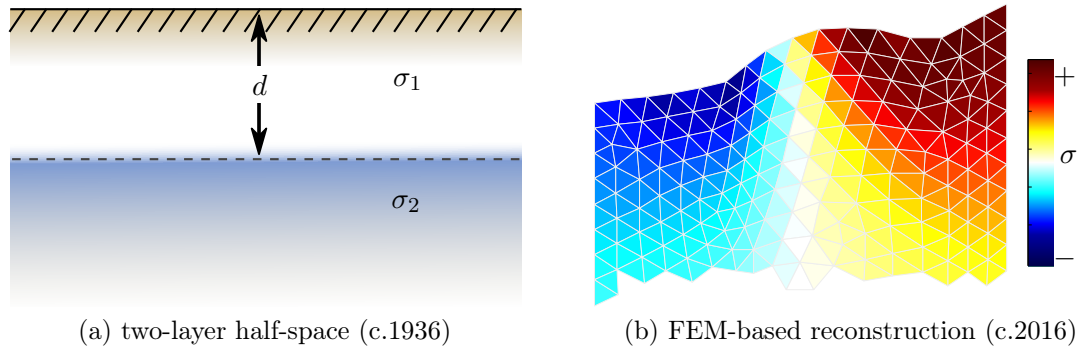
Figure 2.1: Impedance imaging: then and now; Mathematical and computational improvements have afforded the ability to create detailed conductivity estimates for arbitrarily shaped domains without prior knowledge of interface depth or structure.

in building electronics, constructing experimental simulacra, and searching for correlations with the "real world" remain important skills in the ongoing development of impedance imaging techniques. Ultimately, many of the basic impedance imaging techniques used today have been proven upon fundamental mathematical principles, but this does not exclude the use of heuristic, or even arbitrary, choices when faced with real world data sets (Graham and Adler, 2006).

Impedance imaging is widely used in commercial geophysics applications today for ground water monitoring (Slater et al., 1997), soil contamination and remediation monitoring (LaBrecque et al., 1996b), shallow gas prospecting (Ahmad et al., 2009), CO_2 sequestration monitoring (Christensen et al., 2006), and large-scale mineral prospecting (Daniels and Dyck, 1984). Impedance imaging has also been used to investigate nuclear waste storage sites (Lesparre et al., 2013) and behaviour of live volcanoes (Portal et al., 2012). A crude form of impedance imaging has been recognized as the means by which "weakly[1] electric fish," for example the Peters' ele-

---

[1]"Strongly electric fish" have a strong enough discharge to stun prey, for example the electric eel *Electrophorus electricus*.

phantnose fish *Gnathonemus petersii* and black ghost knifefish *Apteronotus albifrons*, find food, communicate and navigate in the muddy African and South American freshwater waterways (Caputi et al., 1998; Nelson and Maciver, 1999; Neveln et al., 2013). For robotics and human-machine interfaces, a skin-like flexible conductive rubber can localize the force applied to a material with a limited number of sensors along the boundary of the material through impedance imaging techniques (Tawil et al., 2011). Impedance imaging also finds applications in an industrial setting for monitoring multiphase fluid flow such as bubble mixing columns in industrial chemistry (George et al., 2000), nondestructive testing for crack, void and strain detection in solids such as cement or metal (Bryan and Vogelius, 1994; Karhunen et al., 2010), landmine detection on land or underwater (Church et al., 2006; Bouchette et al., 2014), management of mixing and heating processes for food, pharmaceuticals, and industrial chemistry (York, 2001; Guérin et al., 2007), monitoring of cement and epoxy curing (Chen et al., 2011; Tallman et al., 2014), and optimization of lumber milling in saw-mills (Steele et al., 2004; Cooper et al., 2008). In forestry management, it is used for monitoring tree health (fungal growth) without core sampling (Brazee et al., 2011). In the biomedical field, there have been promising developments in respiratory lung monitoring (Adler et al., 2009; Gómez-Laberge et al., 2012), non-invasive measurement of heart stroke volume (Vonk-Noordegraaf et al., 2000), brain imaging for management of stroke and epilepsy events (Packham et al., 2012), trauma care monitoring for internal bleeding (Tucker et al., 2011), studying stomach emptying (Soulsby et al., 2006; Nakae et al., 2000), screening for breast and prostate cancers (Kao et al., 2006; Wan et al., 2010), and the measurement of electrocautery and cryosurgery effectiveness/completion during surgery (Edd et al., 2008).

Impedance imaging techniques have a broad range of applications. Each of these applications represents unique challenges in their geometry, mathematics, computational algorithms, data collection and interpretation but are unified by their common mathematical and algorithmic basis. In this work, we focus on the geophysics applications with the understanding that methods developed for a particular application can provide valuable insight into new ways of tackling many of these challenging problems across a range of applications.

2.1 Fundamental Equations

The flow of a direct current (DC) or quasi-static electric current in the presence of a continuous impedance distribution can be derived from Maxwell's equations (see Appendix B).

The voltage (potential) u on the domain Ω can be computed, with the Complete Electrode Model (CEM) (Cheng et al., 1989; Somersalo et al., 1992; Hyvönen, 2004), as

$$\nabla \cdot (\sigma \nabla u) = 0 \quad \text{in} \quad \Omega, \quad \sigma > 0 \tag{2.1}$$

$$\int_{E_m} \sigma \frac{\partial u}{\partial \hat{n}} dE_m = I_m \quad m = 1..M \tag{2.2}$$

$$u + z_m \sigma \frac{\partial u}{\partial \hat{n}} = U_m \quad \text{on} \quad E_m, \quad m = 1..M \tag{2.3}$$

$$\frac{\partial u}{\partial \hat{n}} = 0 \quad \text{on} \quad \partial\Omega \setminus \bigcup_m E_m \tag{2.4}$$

for a domain Ω with a specified conductivity σ and $m = 1..M$ electrodes E_m on the boundary $\partial\Omega$. The divergence of the current density $\sigma \nabla u$ is zero within the domain

which implies that there are no current sources or sinks within (2.1). Over each electrode E_m, the total current leaving the electrode I_m is the sum of the current normal to the boundary $\sigma \, \partial u / \partial \hat{n}$ under the electrode (2.2). The sum of all currents at the boundary is zero $\sum_m I_m = 0$ due to (2.1). The voltage U_m measured by the electrode is a function of the voltage at the boundary, the contact impedance z_m and the current flowing through the electrode (2.3). No current flows across the boundary of the domain in areas without an electrode $\partial \Omega \setminus \cup_m E_m$ (2.4).

The simpler Point Electrode Model (PEM) models the electrode as a single impedance connected to a point on the boundary (Somersalo et al., 1992). The PEM replaces the CEM electrode behaviour (2.2) (2.3) with

$$\frac{U_m - u_m}{z_m} = I_m \tag{2.5}$$

The PEM fails to capture three important effects: the discrete region of the boundary over which stimulus currents flow, shunting of current across the face of the electrode, and the effect of contact impedance on current flow (Figure 2.2). The error is relatively small when many small electrodes are used; the FEM approximates the continuous boundary data scenario. Shunting effects are minimal when contact impedance is much greater than local conductivities under the electrode or electrode diameter (relative to surface area) is small.

(a) PEM	(b) CEM	(c) CEM
$z_c = 0.01\ \Omega \cdot \mathrm{m}$,	$z_c = 1\ \Omega \cdot \mathrm{m}$,	$z_c = 0.01\ \Omega \cdot \mathrm{m}$,
$\sigma = 1\ \mathrm{S/m}$	$\sigma = 1\ \mathrm{S/m}$	$\sigma = 1\ \mathrm{S/m}$
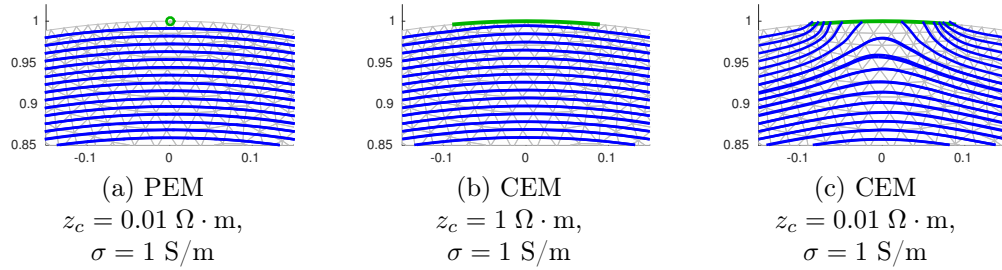
Figure 2.2: Complete and point electrode models; the effect of contact impedance on the two models of a measurement electrode (green, upper surface), current stream-lines near the electrode (blue) for a domain with conductivity 1 S/m and electrode with contact impedance 1 Ω·m (b) and 0.01 Ω·m (a,c) using PEM (a) and CEM (b,c)

2.2 Discrete Approximations

In general, an approximation of the Partial Differential Equation (PDE) is selected: a truncated Fourier series, finite difference, Boundary Element Method (BEM) or FEM. For simple geometric domains, Fourier transforms of (2.1) can be manipulated to provide direct insight into the effect of impedance inhomogeneities of specific shape (Isaacson, 1986). Examples of these simple geometries are the half-space (a flat world), sphere (the head) or circle (the chest). These insights can lead to a deeper understanding of the fundamental behaviour of impedance imaging with respect to sensitivity, resolution and the choice of electrode placement and usage.

For arbitrary geometries, an FEM discretization of the domain is commonly used.

$$u(x) = \sum_e \sum_n u_{e,n} N_{e,n}(x) \tag{2.6}$$

The discretization is a "weak form" that applies a weighted shape function N as a discrete approximation of an arbitrary function. The domain is discretized into adjacent regions or *elements e* forming a *mesh*. The shape function is one at a *node n*

of the element, has a value over the element, and is zero outside of the element. The sum of the shape functions gives a value for the function u over the element that is a linear weighting of the shape functions $N_{e,n}(x)$ by the node values $u_{e,n}$. When nodes are enumerated globally, then (2.6) simplifies to $u = \sum_i u_i N_i(x)$ for node i. By taking the sum of weighted shape functions, one can approximate arbitrary functions with minimal error over the domain (Strang and Fix, 1973). A simple choice of shape functions are the linear elements with nodes at the corners of the element. Piece-wise linearity is enforced across the domain by constraining collocated nodes of adjacent elements to the same value.

The FEM discretization of the EIT equations (2.1)–(2.4) forms the block-wise matrix

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{2.7}$$

$$\begin{bmatrix} \mathbf{G}+\mathbf{B} & \mathbf{C} \\ \mathbf{C}^{\mathsf{T}} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \vec{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{\mathbf{b}} \end{bmatrix} \tag{2.8}$$

where column vectors $\mathbf{u}$ are potential at each node, $\vec{\mathbf{u}}$ are the measurements at the electrodes, and $\vec{\mathbf{b}}$ are applied currents at the electrodes.

The remaining matrices **G B C D** are then defined in terms of

$$G_{ij} = \int_{\Omega} \sigma \nabla N_i \cdot \nabla N_j \tag{2.9}$$

$$B_{ij} = \sum_{\ell=1}^{L} \int_{E_\ell} N_i N_j \tag{2.10}$$

$$C_{\ell j} = -\frac{1}{z_\ell} \int_{E_\ell} N_j \tag{2.11}$$

$$D_{ij} = \frac{|E_\ell|}{z_\ell} \quad \text{for } i = j = \ell \text{ and is } 0 \text{ otherwise} \tag{2.12}$$

for shape function N, conductivity σ, contact impedance z_ℓ and electrode surface area $|E_\ell|$. If the conductivity on the element is considered constant, then (2.9) can be calculated as $G_{ij} = \sum_k \sigma_k \int_{\Omega_k} \nabla N_i \cdot \nabla N_j$, the sum of the element-wise shape functions weighted by conductivity. A ground node is typically restricted to zero potential, removing a column and row from **A** and a row from **b**. Difference measurements are selected as the difference between electrode potentials $\vec{\mathbf{u}}$.

Stimulus (current) I is typically applied across a pair of electrodes $(s+, s-)$ so that $(\vec{\mathbf{b}}_{s+} - \vec{\mathbf{b}}_{s-} = 0)$. Likewise, the i-th measurement $\mathbf{v}_i$ is typically a difference in electrical potential (voltage) across a pair of electrodes $(m+, m-)$ due to a particular stimulus $(\mathbf{v}_i = \vec{\mathbf{u}}_{m+} - \vec{\mathbf{u}}_{m-})$. A selection vector $\mathbf{t}$ may be used to calculate the difference voltages $\mathbf{v}_i$ from the complete voltage set $(\mathbf{v}_i = \mathbf{t}^\mathsf{T}\mathbf{x})$ by selecting the difference between the appropriate electrodes.

Pair-wise stimulus and measurement are used in most impedance systems. A pair-wise stimulus and measurement pattern **P** can be defined as a set of tetrapolar stimulus and measurement pairs in a $k \times 4$ array with rows $(s+, s-, m+, m-)$ for k difference measurements. Each row of the stimulus and measurement pattern

$$\mathbf{P}_i = [3, 2, 5, 4]$$
$$\mathbf{B}_i = I\,[\,0,\ -1,\ +1,\quad 0,\quad 0,\ 0,\ 0\,]^\mathsf{T}$$
$$\mathcal{T}_i = G\,[\,0,\quad 0,\quad 0,\ -1,\ +1,\ 0,\ 0\,]$$
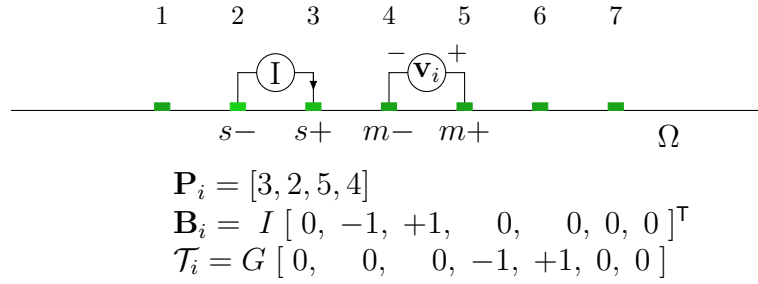
Figure 2.3: Stimulus and measurement electrode pairs; electrodes (green, numbered left-to-right) applied to a half-space domain Ω, a pair of stimulus electrodes $(s+, s-)$ has current I driven between them $\mathbf{B}_i$, a pair of measurement electrodes $(m+, m-)$ selected from the available electrodes $\mathcal{T}_i$ measures a differential voltage that has been scaled by gain G all due to the voltage distribution caused by the stimulus according to row i from the stimulus and measurement pattern $\mathbf{P}_i$, the result is a single entry in the measurement vector $\mathbf{v}_i$

corresponds to an entry in the measurement vector $\mathbf{v}$. A set of measurements may be constructed by a stimulus vector $\mathbf{b}$ for each $(s+, s-)$ pair and a selection vector $\mathbf{t}$ for each measurement pair $(m+, m-)$ in the stimulus and measurement pattern. The stimulus vectors $\mathbf{b}$ may then be concatenated into a stimulus matrix $\mathbf{B}$. Likewise, the measurement selection vectors $\mathbf{t}$ may be collected into a matrix $\mathcal{T}$ (Figure 2.3).

A naïve least squares solution to the square system of equations will calculate an LU-decomposition of the system $\mathbf{A}$ for each stimulus $\mathbf{b}$.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{2.13}$$

$$\mathbf{LU} = \mathbf{A} \qquad\qquad \text{LU-decomposition} \tag{2.14}$$

$$\mathbf{Ly} = \mathbf{b} \qquad\qquad \text{solved for } \mathbf{y}, \text{ forward substitution} \tag{2.15}$$

$$\mathbf{Ux} = \mathbf{y} \qquad\qquad \text{solved for } \mathbf{x}, \text{ backward substitution} \tag{2.16}$$

where $\mathbf{y}$ is a temporary variable. The LU-decomposition for a square matrix gives

lower $\mathbf{L}$ and upper $\mathbf{U}$ triangular matrices. The LU-decomposition, for an $n \times n$-matrix is $O(\frac{2}{3}n^3)$ and is generally twice as fast as QR-decomposition $O(\frac{4}{3}n^3)$ (Trefethen and Bau, 1997, §20). The solution will give the result at the electrodes and at all nodes $\mathbf{x}$. Iterative refinement is usually applied (with sufficient additional precision) to correct for accumulated numerical error by using the residual $\mathbf{r}$.

$$\mathbf{r} = \mathbf{b} - \mathbf{Ax} \tag{2.17}$$

$$\mathbf{x}' = \mathbf{x} + \mathbf{A}^{-1}\mathbf{r} \tag{2.18}$$

Multiple stimuli $\mathbf{b}$ and the iterative refinement steps can be calculated by reusing the $\mathbf{L}, \mathbf{U}$ matrices.

Alternatively, a pseudo-inverse $(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{Ab}$ can be computed using the Cholesky decomposition for a Hermitian positive definite matrix $\mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{LL}^\mathsf{T}$ in roughly half the time $O(\frac{1}{3}n^3)$ of an LU-decomposition (Trefethen and Bau, 1997, §23) at the cost of some lost precision. The Cholesky decomposition is not applicable when $\mathbf{A}$ is constructed from complex-valued conductivities; $\mathbf{A}$ is then symmetric but not Hermitian. For very large matrices, linear Conjugate Gradient methods may be more efficient.

Heuristic algorithms may be used to search for an efficient factorization of the sparse matrix $\mathbf{A}$ into an ordering that best utilizes the computer's resources; calculating a decomposition into dense matrices. These factorization algorithms are restricted in their search; spending too long looking for a better matrix arrangement wastes time that could have been spent solving the current "best" matrix (Duff et al., 2002; Davis et al., 2004).

In general, for the stimulus matrix $\mathbf{B}$ there are many more rows m (nodes and electrodes in the mesh) than columns n (stimulus electrode pairs). A QR-decomposition can be applied to the column-wise stimulus $\mathbf{B}$ to rearrange the stimulus so that

$$\mathbf{QR} = \mathbf{B} \qquad\qquad\qquad \text{QR-decomposition} \qquad (2.19)$$

$$\mathbf{QR} = \begin{bmatrix} \hat{\mathbf{Q}} \ \hat{\mathbf{Q}}_0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{R}} \\ 0 \end{bmatrix} = \hat{\mathbf{Q}}\hat{\mathbf{R}} \qquad\qquad (2.20)$$

$$\mathbf{X} = (\mathbf{A}^{-1}\hat{\mathbf{Q}})\hat{\mathbf{R}} \qquad\qquad (2.21)$$

The QR-decomposition for an $m \times n$ rectangular matrix $\mathbf{B}$ gives a unitary $m \times m$ matrix $\mathbf{Q}$ and upper triangular $m \times n$ matrix $\mathbf{R}$. The lower $(m-n)$ rows of $\mathbf{R}$ are zero, so the right-most columns of $\mathbf{Q}$ are all multiplied by zero. A "thin" QR-decomposition $\hat{\mathbf{Q}}\hat{\mathbf{R}}$ is the result of dropping these zero-multiplied rows and columns. The thin $\hat{\mathbf{Q}}\hat{\mathbf{R}}$ matrices reduce the quantity of multiply-adds required in the forward substitutions (2.15). The QR-decomposition will combine linearly dependent columns of stimulus $\mathbf{B}$, reducing the number of forward solutions required to find the measurements $\mathbf{v}$. These dependent columns may be dropped by removing the rows of $\mathbf{R}$ and columns of $\mathbf{Q}$ that are nearly zero ($\mathbf{R}_i < \epsilon$).

As an example, for four stimulus using three electrodes in a system with m electrodes and n stimulus pairs

$$\mathbf{B} = \begin{bmatrix} 0 & & & \cdots & & & 0 \\ & \ddots & & & & & \\ & & 1 & 1 & 0 & -1 & \\ & & 0 & -1 & 1 & 0 & \\ & & -1 & 0 & -1 & 1 & \\ & & & & & \ddots & \\ 0 & & & \cdots & & & 0 \end{bmatrix} \qquad (2.22)$$

the QR-decomposition will have two rows in $\mathbf{R}$ with values greater than machine precision ϵ so that $\hat{\mathbf{Q}}$ is $m \times 2$ and $\hat{\mathbf{R}}$ is $2 \times n$.

2.3 Hardware, Wires and Electrodes

The equations of the FEM provide a nice way of modelling the current flow through a domain, but how is that current actually delivered to the boundary? The equipment used for geophysical and biomedical applications is in some ways very similar and in others drastically different.

Biomedical equipment designs consider patient safety first; electrical isolation, current limiting and protection against electrical faults are of primary concern. Power is generally supplied from mains power and at a well controlled frequency and voltage. Repetitive events of interest occur over relatively short time spans: breathing and heart rate, for example. Interesting diagnostic information exists in the analysis of changes in this dynamic data.

Geophysical equipment is quite different; generally speaking, more power applied affords the opportunity for better signal-to-noise. Power is often only available through batteries or fuel cells charged by solar or wind power at remote locations.

Sufficient power to take a set of measurements may be accumulated over days. The geological events of interest occur over long time frames but can exhibit dramatic irreversible changes over the short term. Changes occur against a background of extreme geological conductivity contrast for which both absolute and time-lapse (difference) reconstructions may be employed. These conductivity reconstructions are used to quantify changes caused by geological processes and manmade interventions.

Ultimately, these two applications are tied together by their fundamental mathematical equivalence. The electronics, wiring, electrodes and algorithms are performing a similar function: to deliver current to the boundary of a domain and measure the resulting potential.

The electronics for an impedance measurement system are deceptively simple. A typical biomedical system performs its task with a system composed of digital and analog components as in Figure 2.4. Frequencies are digitally synthesized sinusoids. A modified Howland current source (Figure 2.5a) typically supplies low-frequency current (100 Hz – 10 MHz, 100 μA – 10 mA) to a switching array. A switching array selects electrodes to deliver and return current. While current is driven, difference voltage measurements are taken across a selection of non-driving electrodes. Differential measurements from an instrumentation amplifier (Figure 2.5b) are then converted to complex voltage measurements by accounting for phase offset between the stimulus current and measured voltage (quadrature measurements). An FPGA or DSP manages the electronics, synthesizes frequencies, and stores data for communication to a computer where the data can be analyzed and an image reconstructed (Boone and Holder, 1996; Ross et al., 2003).

For geophysics applications, the current source stimulus circuit has tended to
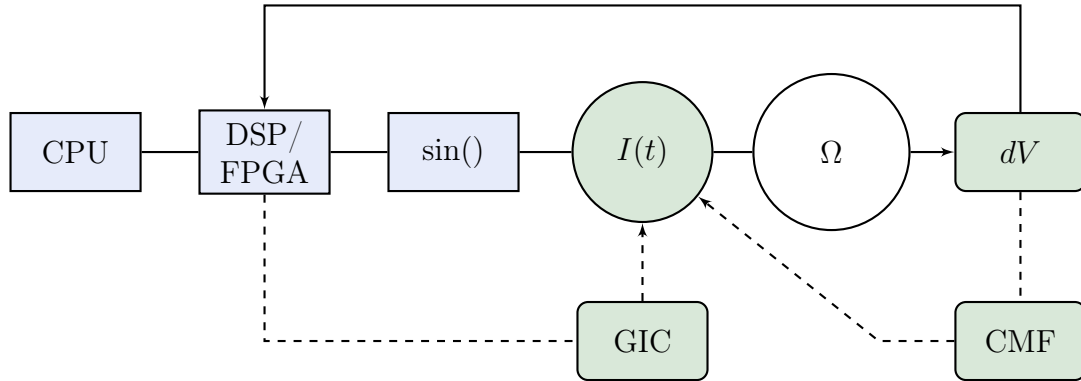
Figure 2.4: Block diagram of typical EIT hardware; A computer (CPU) configures the EIT system and receives measurements used for reconstructions of conductivity, a DSP or FPGA has local control of a digital frequency synthesis sin() which drives a modified Howland current source $I(t)$ across a body Ω, measurement of difference voltages dV due to stimulation currents are stored by the DSP/FPGA and reported back to the CPU; A digitally tuned GIC corrects for stray capacitance in the system; Common Mode Feedback (CMF) may be used to improve Common Mode Rejection Ratio (CMRR) on measurements.

take a simpler form given lower frequency requirements. An alternating square wave $(+1, 0, -1, 0, \ldots)$ at approximately one Hertz is driven onto a pair of stimulus electrodes. The applied current is measured and recorded. Wiring and differential measurement challenges remain largely the same between biomedical and geophysics applications (Ogilvy et al., 2009; Kuras et al., 2009).

These relatively loose constraints belie the challenge in building an accurate system. Common Mode Rejection Ratio (CMRR) of Operational and Instrumentation Amplifiers typically falls off rapidly for frequencies in the 100kHz range (McEwan et al., 2007; Langlois et al., 2015). CMRR is addressed through careful design of the circuits, matching of resistors, calibration and active Common Mode Feedback (CMF) circuits (Boone and Holder, 1996; Langlois et al., 2015).

Stray capacitance limits frequency range and leads to common mode voltages

from the current sources. Active shielding of tri-axial electrode cabling can reduce wiring capacitances (McEwan et al., 2007). A Generalized Impedance Converter (GIC) (Figure 2.5c) can be constructed to cancel any additional stray capacitance by using digital potentiometers to tune the circuit. The GIC can provide a significant improvement in current source output impedance (Ross et al., 2003). GICs require calibration for discrete frequencies and to be recalibrated at regular intervals as parts age and temperatures changes. Variable stray capacitances are very challenging to calibrate out.

The simplest systems minimize the number of active circuits. The number of circuits to be calibrated is reduced by using a single current source and differential measurement circuit. Stray capacitance is then introduced by the switching infrastructure needed to connect the circuits to the electrodes. Other systems construct many current sources and differential measurement circuits. Multiple current sources enable complex stimulus patterns (Boone and Holder, 1996). Multiple measurement circuits accelerate acquisition times. When each electrode has a dedicated current source and voltage measurement, the switching subsystem may be removed. Still other systems place the electronics at the electrode, so called "active electrodes," to remove wiring from the list of critical design issues (McEwan et al., 2007; Gaggero et al., 2012).

Electrodes and the stimulus and measurement patterns applied to them may be selected to minimize polarization effects (Dahlin, 2000; Wilkinson et al., 2012). The reciprocal error is used to estimate the error due to polarization and electrical noise. Reciprocal error is calculated as the difference in measurements for reversed stimulus and measurement electrode pairs (Parasnis, 1988; LaBrecque et al., 1996a).

(a) Modified Howland current source

(b) Instrumentation amplifier

(c) Generalized Impedance Converter
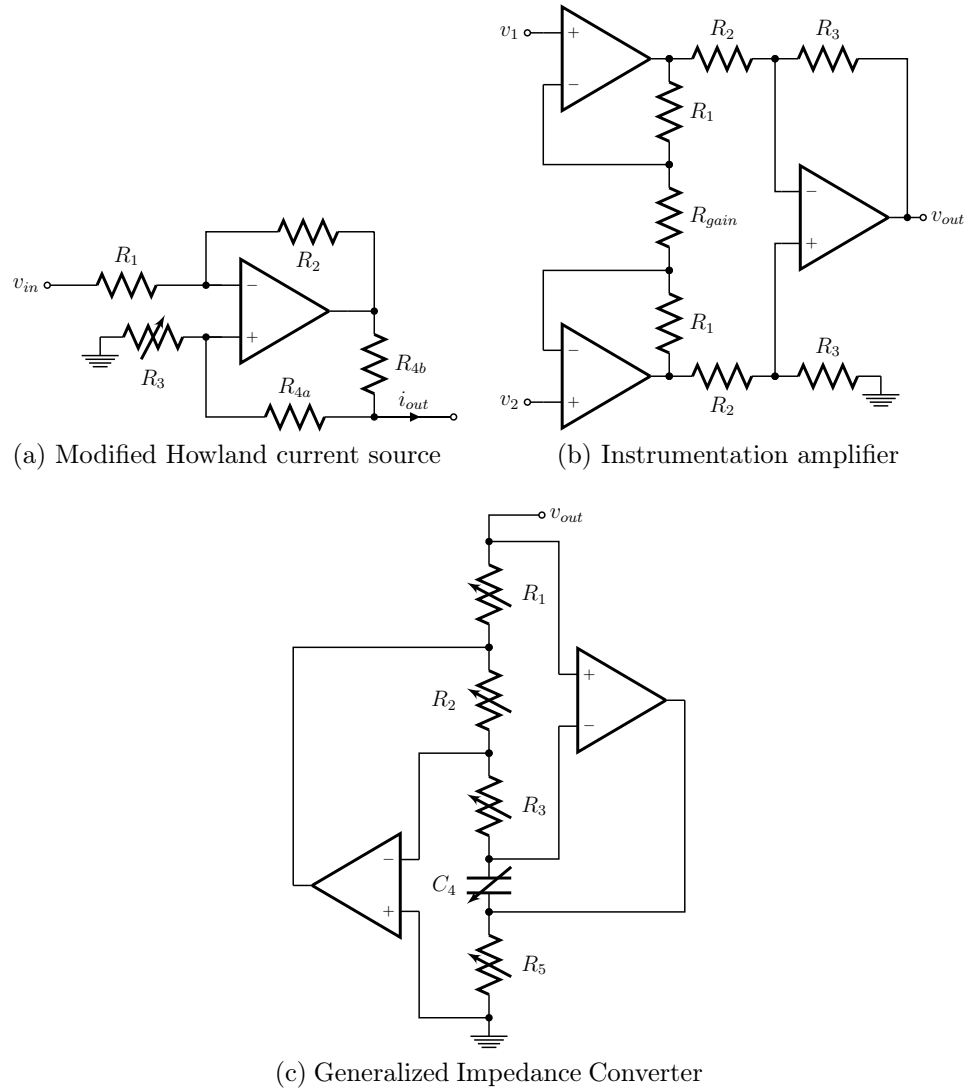
Figure 2.5: Common circuits found in EIT hardware; (a) a modified Howland current source is typically placed in parallel with (c) a Generalized Impedance Converter (GIC) tuned to cancel stray capacitance; difference voltage measurements are captured with modified versions of (b) the standard instrumentation amplifier

Contact impedance estimates can also be used to roughly determine the quality of electrode contact with the boundary. The contact impedance and reciprocal error estimates can be used to weight the data, eliminate bad measurements, detect equipment faults or correct electrode connectivity problems.

Various materials can be used for electrodes. In biomedical work, non-polarizing Ag-AgCl gel electrodes are in common use. For geophysics, spikes of stainless-steel or various brass composites are preferred (polarizing) electrode materials for their cost effectiveness and corrosion resistance. Polarization effects are more important to geophysical Self Potential (SP) and Induced Potential (IP) surveys where non-polarizing Pb-PbCl$_2$ electrodes are sometimes used. SP measurements are performed without artificial stimulus where spontaneously occuring voltages are due to natural processes such as geochemical and geobiological batteries, temperature gradients, salinity gradients and fluid flow (Jouniaux et al., 2009). IP measurements are the time domain potential response to a step input and historically, used non-polarizing electrodes. More recently, careful IP survey design have enabled data collection using stainless steel electrodes (Dahlin et al., 2002; Dahlin and Leroux, 2012; Gazoty et al., 2013). SP voltages would typically be calibrated out as a measurement offset at the EIT hardware. For ERT, IP measurements are avoided by waiting for the voltage response to reach an equilibrium. In any case, polarization effects do occur to a greater or lesser degree. When switching between electrodes, some EIT devices have the ability to "ground" the electrode for a period of time which allows accumulated charge to dissipate. An alternate scheme is to maximize the time between when electrodes are used for stimulus and when those electrodes (or nearby electrodes) are used for measurements (Wilkinson et al., 2012).

2.4 Discussion

Impedance imaging, as a technology will give a blurry picture which may still contain valuable information. If one considers the change in that information as interesting, then looking at ERT as a monitoring technology is a promising avenue. Changes in the medium indicate some change in the boundary or the impedance within. What does this indicate? One needs a deeper understanding of the underlying mechanisms of conductivity in rocks to "dig deeper."

Chapter 3

Rocks and Conductivity

Rocks can have conductivity and other properties that range over orders of magnitude. A reliable interpretation of a reconstruction of the conductivity underground depends on a sound understanding of the geology of the expected materials.

3.1 Rock Classifications

The classification of the world around us forms a basis for much of the natural sciences. Minerals form an important component of these classification schemes according to Linnaeus' broad taxonomic system (plant, animal, mineral) and the modern Nomenclature Codes for plants, animals, bacteria, viruses and minerals that have grown from Linnaeus' Latin binomial nomenclature (Linnaeus, c.1753). Solid homogeneous chemical compounds with a crystalline structure formed through geological processes are classified as minerals by the governing body of mineralogical naming, the International Mineralogical Association (Martin, 1998). Mineraloids are similar to minerals but do not exhibit their crystalline structure.

| (a) igneous | (b) sedimentary | (c) metamorphic |
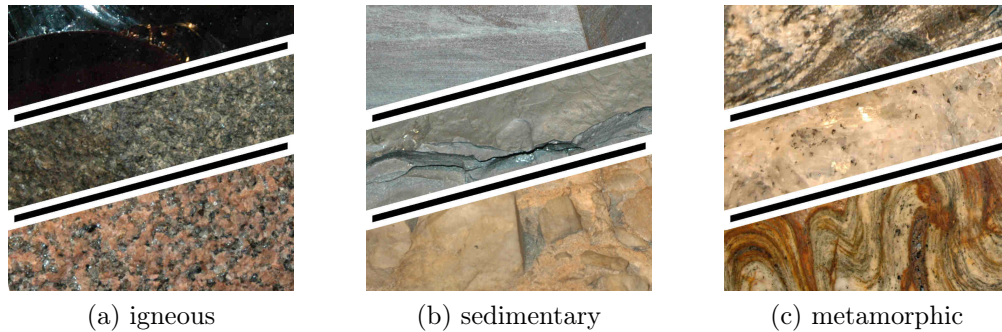
Figure 3.1: Examples of igneous, sedimentary and metamorphic rock; (a) igneous obsidian, gabbro and granite (top to bottom), (b) sedimentary sandstone, shale and brecia (top to bottom), (c) metamorphic gneiss, marble and folliated granitic gneiss (top to bottom)

The familiar rocks we see about us in our everyday lives are a solid conglomeration of a range of these minerals and mineraloids. The most significant differentiator between types of rock is their rate of formation. Igneous rocks are formed from magma flows; either molten lava cools underground, as with granite, or is forced to the surface and expelled through a volcanic eruption as in pumice and rhyolite. Sedimentary rocks, on the other hand, are formed near the surface from organic matter, chemical precipitates and fragments of older rocks. Sediment is compressed by layers above to form sandstone, shale and limestone amongst others. Metamorphic rocks are formed of other rock under much greater pressure and heat than sedimentary rocks. The heat and pressure recrystallizes existing minerals into new forms.

The possible combinations of minerals that form rocks exist in a broad spectrum. Combinations of minerals that can form a type of rock are dictated by their combined phase diagram. The Eutectic point in the phase diagram (a particular combination of mineral ratio and temperature) will identify when minerals solidify into a common material rather than cooling at different temperatures into separated layers (Klein

and Philpotts, 2013, §8.3). The first mineral to cool crystallizes and precipitates out of the combined liquefied mineral melt. The precipitation removes a portion of a mineral from the available combined mineral melt which changes the ratio of materials in liquid form. The operating point on the phase diagram (temperature versus mixture ratio) then moves, and the liquid's temperature cools. Eventually the mineral melt reaches a Eutectic point at which time the materials combine at the crystalline level and precipitate together, which holds temperature stable, until there is no liquid material remaining. The combined material can be homogeneous or exhibit particular bands, rods, or other crystalline structures of alternating material.

The naming and identification of a particular rock is primarily a communication tool. It selects a point from the spectrum to use as a reference for identifying rocks of a similar make-up. Rock names are used to classify and clearly communicate about the existing geology of a region. The art and science of identifying stone and inferring the geological story surrounding it forms an important part of the work of geology. The knowledge of how and when rocks were formed is used to understand where potential resources might exist and to predict the likelihood and outcome of natural events such as earthquakes and volcanic eruptions (Renton, 1994; Skinner and Porter, 1987).

The identification of rocks in the field and laboratory is performed with a range of tools from hammer and microscope to X-ray crystallography and chemical analysis (Figure 3.2). Direct observation remains a key tool in the field. Hardness, cleavage, specific gravity, luster, colour, streak, texture and crystalline structure are the common properties used to differentiate rocks (Coe, 2010). When the rock is not directly available for examination, indirect measures such as electric, magnetic, radioactive

Figure 3.2: Geologists hammers; Hardened steel head to prevent metal splinters when breaking rocks. Spike or chisel end for freeing specimens and clearing cover. (a) Estwing rock pick, (b) Wood-handled rock pick, (c) Estwing crack hammer, (d) Loupes (magnification) for examining rocks.

and acoustic properties can be acquired. Each of these properties may be used to aid in classifying a particular rock or group of rocks.

3.2 Conductivity Mechanisms

The conductivity of a particular rock specimen is mainly the combination of the conductive properties of the constituent minerals, the structure of the rock, and the degree of water saturation. The conductivity of rocks varies over nearly twenty-four orders of magnitude because rocks can consist of such a wide range of materials and those materials can be constituted in many forms (Touloukian et al., 1989; Keller and Frischknecht, 1966; Telford et al., 1990) (Figure 3.3).

Materials can be classified by their ability to provide charge transport mechanisms and their resulting conductivity. Charge as ions, electrons or protons are transported through the rock and across interfaces between media via diffusion, percolation, tunnelling and thermal-convection (Touloukian et al., 1989). A measurement of
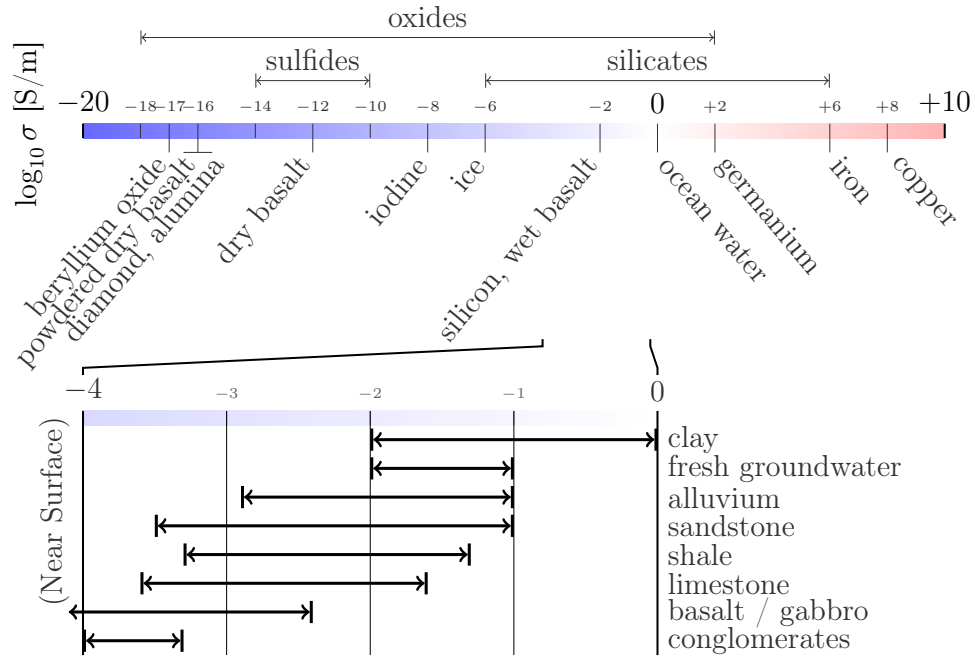
Figure 3.3: Typical rock and mineral conductivities; relevant conductivities for near surface rocks are identified in the lower portion of the chart

conductivity in a solution depends on the ionic mobility and concentration. Ions are formed from electrolytes that dissociate into charged particles when dissolved in an appropriate solution. The movement of those ions is hindered by fluid viscosity and particle interactions. An increase in temperature strongly affects ion mobility. At high temperatures, entire atoms move through ionic percolation; while at lower temperatures, tunnelling or defect diffusion allows charge transport. Most rocks tend to be insulators with large energy barriers between adjacent atoms so that electrons and protons are not the mechanism of charge transport. Conversely, metallic rock provides a uniform distribution of available valence electrons allowing conductivity through electron migration. Semiconductors fill the middle ground where an energy barrier just high enough to prevent electron transport without external energy exists.

Concentrations of naturally occurring semiconductors are rare.

When the charge distribution changes in a non-uniform manner, charge accumulates resulting in polarization. A force supplied through electric fields, mechanical deformation, material changes or chemical reaction can motivate charge separation and resulting polarization behaviour (Touloukian et al., 1989). In impedance imaging, polarization effects due to electric fields are usually dominant.

Charge separation mechanisms under an electric field at high frequency (10^{24} – 10^{12} Hz) vary in scale from distortions of electron clouds to preferential alignment of polarized molecules. For lower frequencies ($< 10^9$ Hz), interface polarization occurs due to charge accumulation (Touloukian et al., 1989). Bulk transport phenomena dominate when wave lengths are greater than the rock grain size. Interfaces may be abrupt, such as at material boundaries, or distributed due to regions of local variation in electrical properties of a material.

The electric fields that drive charge transport and charge separation are not exclusively supplied by external sources but can also be generated internally by a material (Touloukian et al., 1989). In rocks, the streaming, Nernst, Dorn (sedimentation), diffusion and redox potentials can play important roles. The Galvanic (contact) and Seebeck (thermoelectric) potentials tend to contribute little to charge transport within rocks. In particular, the streaming or electro-osmotic potential plays a key role in the behaviour of ionic flow through rock pore spaces by resisting the prevailing electric current.

3.3 Archie's Law for Sedimentary Rocks

The Schlumberger brothers first made electrical surface resistivity measurements in 1911 (Schlumberger, 1920; Allaud and Martin, 1977), but it wasn't until Archie of Shell Oil published (Archie, 1942) that the relationship between sedimentary rock porosity, resistivity and saturation became clear. There are many models for specific conductivity modes in rock, but these complex models generally can be reduced to a few commonly used empirical models.

Charge transport through the rock pores plays a primary role in many sedimentary rock types. Conductivity in sedimentary fluid filled rocks σ_r is commonly modelled using Archie's Law (Archie, 1942)

$$\sigma_r = \frac{1}{a}\sigma_w\phi^m S_w^n \tag{3.1}$$

where the water conductivity σ_w, pore volume fraction ϕ, and saturation S_w along with three empirical constants a, m and n can closely fit measured sedimentary rock conductivities. Archie's Law assumes that the main conductivity mechanism is charge transport through the pore structure rather than conduction through the rock fabric or due to water-rock chemical interaction. The pore volume fraction $0 \leq \phi < 1$ is the fraction of the rock that is empty space available to be filled with a fluid while the saturation fraction S_w indicates how much of that space is occupied by fluid. The empirical value m has been found to depend on pore and grain texture and is commonly referred to as the cementation exponent. The value of m would be 1.0 for parallel capillaries and is 1.3 for packed spheres (for example, uncompacted sand) and typically increases with the degree of cementation. For most sedimentary

rocks, m is in the range of 1.8–2.0. The empirical factor a corrects for variation in compaction, grain size and pore structure and depends on the length of the current flow paths. a is commonly referred to as the tortuosity because it potentially corrects for variation in the connectivity through the pore paths. The final empirical factor n reflects the "wettability" of rock where non-conductive fluids in the pore space such as hydrocarbons reduce the ability to maintain surface water layers in the pore as saturation is reduced. The saturation exponent n is commonly close to 2.0. At sea-level pressures, the pore volume fraction can be considered constant for a particular rock. At greater depths within the crust, pressures are high enough to compress the rock and reduce or close off pore spaces. Water conductivity σ_w is strongly affected by temperature, pressure and ion concentrations. For metamorphic and igneous rocks, the cementation exponent m can be much larger (Wright et al., 2009).

Two oft cited geophysics measures of rock formations use Archie's Law: the formation factor and resistivity index (Kemna, 2000). The formation factor F is the ratio of the water and rock conductivities at 100% water saturation ($S_w = 1.0$)

$$F = \frac{\sigma_w}{\sigma_{r|S_w=1.0}} = \frac{a}{\phi^m} \tag{3.2}$$

while the resistivity index I is the ratio of rock conductivity at 100% water saturation to the *in situ* rock conductivity

$$I = \frac{\sigma_{r|S_w=1.0}}{\sigma_r} = S_w^{-n} \tag{3.3}$$

The formation factor and resistivity index are notable because they cancel the effect of water conductivity σ_w which may not be known (Ucok et al., 1980).

3.4 Clay Minerals in the Near Surface

Clay minerals are the insoluble products of chemical weathering that result in fine particulate matter a thousand times smaller in diameter than sand. Clay can typically contain 10–60% water and exhibit grain arrangements that lead to variable conductivity depending on orientation.

Archie's Law assumes that the primary conduction mechanism is ionic transport through the pore spaces. In cases where this is not true, Archie's model fails to accurately estimate conductivity for those materials. In particular for clay, or rocks containing a significant ratio of clay, the clay itself provides the main conduction mechanism through cation/anion exchange.[1] Clay is frequently encountered, outside of the tropics, when acquiring near surface measurements (Renton, 1994).[2]

Clay develops an electrical double layer of ions on the fluid-solid interface which strongly affect its electrical and chemical behaviour. A layer of charged ions is trapped along the surface; ions are adsorbed due to chemical interactions with the solid. A second layer is attracted from adjoining fluids to the first through Coulomb forces. The combined layers are (approximately) electrically neutral, as the charges on the ions cancel. The "double layer" is not, in fact, a precisely layered sequence of ions but is commonly modelled as a diffuse region with exponential decay away from the interface (Stern, 1924).

An electrical double layer has two important properties: conduction and charge storage. Conduction can occur preferentially along the double layer; ion density is

---

[1]Materials such as metal, graphite or minerals with a "metallic luster" also break this assumption because they conduct mainly through electron transport.

[2]Clay is a dominant component of most soils. In the tropics, the extreme weathering decomposes clay into hydrated oxides of aluminum and silicon dioxide. The concentration of these oxides leads to reduced agricultural soil quality. (Renton, 1994)

higher than in the fluid. This conduction is important in materials with a large surface area to volume ratio, such as clay. The trapped ions are a stored charge; changes in the potential applied to the material exhibit capacitive effects. These capacitive effects are the same ones observed as polarization on electrodes.

A common adaptation of Archie's Law, the Waxman-Smits equation (Waxman and Smits, 1968), attempts to account for the effects of clay by adding elements that support variation in ion mobility and concentration

$$\sigma_r = \frac{S_w^n}{F}\left(\sigma_w + \frac{BQ_v}{S_w}\right) \tag{3.4}$$

where B is the average ion mobility and Q_v is the cation concentration per unit pore volume. The pore fraction, cementation exponent and tortuosity $a\phi^{-m}$ of Archie's law have been converted to a formation factor F using (3.2). These values must typically be fitted to careful lab measurements of core samples to correlate conductivity with saturation (Merrit et al., 2015).

For mixtures of clay in a rock matrix, the Maxwell-Garnett Effective Medium Approximation (EMA) is sometimes applied (Maxwell-Garnett, 1904). For an arrangement of spherical inclusions of conductivity σ_i in a matrix of conductivity σ_m, the effective conductivity σ_e is found to be

$$\sigma_e = \sigma_m \frac{2\delta_i(\sigma_i - \sigma_m) + \sigma_i + 2\sigma_m}{\delta_i(\sigma_m - \sigma_i) + \sigma_i + 2\sigma_m} \tag{3.5}$$

for inclusions comprising $0 \leq \delta_i \leq 1$ of the volume, so long as the denominator does not vanish. The mixing equation (3.5) provides a bound on expected conductivity for mixed rock materials (Robinson and Friedman, 2003).

3.5 Water Conductivity

Changes in the conductivity of water can confound the interpretation of a conductivity reconstruction. Water conductivity is influenced by ion concentration, pressure and temperature.

At temperatures between freezing and boiling, water can take a wide range of conductivities (Waxman and Thomas, 1974). At low temperatures, ionic mobility and the corresponding conductivity are reduced up to the point where water forms a solid. At high temperatures, the energetic movement of the water molecules tends to interfere with ionic conduction and reduce conductivity until the liquid turns to gas. As both solid and gas, the ionic conductivity mechanism is disabled and other, generally less conductive mechanisms, dominate. At typical groundwater temperatures, conductivity increases with increasing ionic mobility due to increasing temperature.

Pressure changes the boiling and melting points of water and simultaneously changes the conductivity. Under great pressure the maximum conductivity can be significantly different than at the surface (Wyble, 1958).

Periods of flooding and run-off can cause abrupt increases in water conductivity by dissolving dry minerals into solution at the surface. A dry period can result in an increase and then decrease in conductivity as ion concentration initially rises through evaporation and then falls as the wet material dries. Periods of prolonged heavy rain tend to reduce surface water conductivity through ionic dilution (Light et al., 2005).

A comprehensive model of water conductivity at the surface would account for pressure, temperature and ionic effects (ion types and their concentration). Modelling these effects is not trivial. Seasonal heating at depth is a damped sinusoid that is modified by layered materials with differing thermal conductivities. Inputs

for heating, pressure, rainfall and ion concentrations would need to be monitored to correlate measurements to a model. Alternatively, direct measurements (a thermistor array over depth) can be correlated to a model assuming homogeneous material

$$T(z,t) = \bar{T} + Ae^{-z/d}\sin(\omega t + \phi - z/d) \qquad (3.6)$$

for temperature T at time t and depth z. Mean air temperature $\bar{T}$, amplitude A, characteristic penetration depth d and phase offset ϕ are fitted to measured annual temperature variation (Chambers et al., 2014).

When pressure and ion concentrations do not vary significantly, conductivity variation is typically linearized within a temperature range. At the surface, water conductivity is commonly characterized as increasing by 2% per degree Celsius ($+2\%/°$C) over the range of $15 - 25°$ Celsius (Light et al., 2005). In combination with (3.6), the known dependence of conductivity with temperature can be used to estimate the expected seasonal conductivity changes in the ground. Knowing the expected conductivity due to seasonal variations more clearly reveals other changes in conductivity which are typically due to moisture content changes (Chambers et al., 2014).

3.6 Fractured and Porous Rock

Rocks with open interconnected fissures or breaks in the medium tend to be dominated by ion conduction through these gaps. The effect of the streaming potential in near surface rock is reduced when conduction occurs outside of the rock pore spaces. (The streaming potential retards ionic flow; ions trapped on the pore walls

resist movement of ions in the adjacent fluid.) Conduction of ions in water seeping through the rock fractures becomes the dominant conductive mechanism. The apparent macro scale conductivity is then the result of the connectivity and alignment of conduction paths. Fractures aligned along a plane can result in observations of strong anisotropic conductivity. Rock develops fractures perpendicular to the least principle stress. A vertically loaded shale will develop horizontal fractures that result in greater conductivity in the horizontal plane than vertically. The difference between unconnected and connected fractures with water seepage can be significant enough to be observed in fractured rocks (Hill, 1972; Scesi and Gattinoni, 2009). Changes in conductivity magnitude and anisotropy can be used to monitor the development of stress fractures over time (Nicollin et al., 2010).

Metamorphic and igneous rock exhibit different conductivity behaviours than sedimentary rock. Archie's model for conductivity does not reflect this difference. None the less, the model can be roughly fit to porous igneous pumice and similar materials (Wright et al., 2009). Metamorphic and compact igneous rock such as basalt tend to exhibit lower conductivity than sedimentary rocks since the pore space is much smaller. One notable feature of this tighter pore space is that hydration and surface temperature tend to have less affect on conductivity (Lee et al., 1983).

3.7 Impedance and Landslides

An image of impedance near the surface has a close relationship to the lithology of the structure being imaged. Electrical impedance contrasts between adjacent rock units can reveal variation in their physical characteristics (Jongmans and Garambois,

2007; Supper et al., 2014; Perrone et al., 2014).

On slopes, lithology can affect slope stability; moisture content changes the characteristics of rock and particularly clays. An increase in moisture content decreases shear strength due to raised pore pressures and increases shear stress due to increased mass. Through careful lab work and on-going monitoring, there is the opportunity to proactively monitor and predict slope failure using impedance imaging (Friedel et al., 2006; Lebourg et al., 2010).

3.8 Discussion

Knowledge of the relationship between *in situ* rock and its conductivity informs decisions about the validity of reconstructed conductivity images. The combination of images and geological observations can be used to make choices about where and when to gather further data. This data can take many forms such as where to drill or dig test holes, realignment of the conductivity survey, or the collection of different types of weather data. Additional data can also be used as prior information in the model for the inverse conductivity problem.

Chapter 4

Inverse Problems

The inverse problem is to take measurements and determine the model parameters that best fit these data. The definition is broad; the field of inverse problems has focused on functions that are mathematically hard to resolve. The definition of inverse problems encompasses all three types of PDE: elliptic, hyperbolic and parabolic. We focus here on the elliptic PDE inverse problem of impedance imaging. For impedance imaging, examples of interesting model parameters are interior conductivity, contact impedance, electrode position and boundary shape.

A forward problem translates model parameters and stimulus into measurements $\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) = \mathbf{v}$ (Figure 4.1). When a forward problem is linearized about a set of model parameters $\boldsymbol{\sigma}$, it is characterized by the sparse system matrix $\mathbf{A}$. The system matrix incorporates the implicit and explicit model parameters such as geometry, conductivity and electrode locations. The system matrix also embodies a particular choice of discretization. The measurements $\mathbf{v}$ are determined by applying stimulus
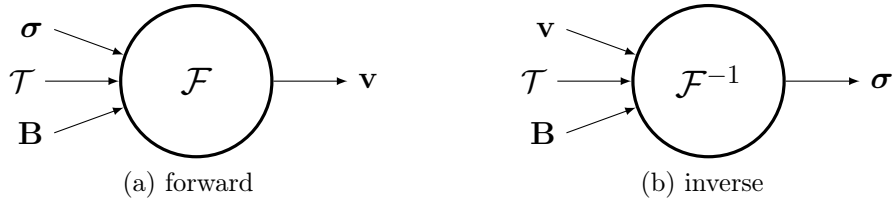
(a) forward   (b) inverse

Figure 4.1: Forward and inverse problems; the forward problem constructs a system based on model parameters $\boldsymbol{\sigma}$, and then estimates the measurements $\mathbf{v}$ that would result from some input stimulus $\mathbf{B}$ from selected electrodes $\mathcal{T}$; an inverse problem takes the input stimulus $\mathbf{B}$ and measurements $\mathbf{v}$ on selected electrodes $\mathcal{T}$ to reconstruct an estimate of the model properties $\boldsymbol{\sigma}$

$\mathbf{B}$ such that

$$\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) = \mathbf{v} \quad \rightsquigarrow \quad \mathbf{v}_i = \mathcal{T}_i \mathbf{A}^{-1} \mathbf{B}_i \tag{4.1}$$

The inverse problem is not the linear algebraic inverse of the forward problem $(\mathbf{A}^{-1})^{-1}\mathbf{x} = \mathbf{b}$. We are not interested in learning the stimulus $\mathbf{b}$ used to construct our measurements, nor do we know the voltages $\mathbf{x}$ throughout the interior. Instead, we are interested in the model parameters $\boldsymbol{\sigma}$. Therefore, while the forward model may be linearized so that $\mathcal{F} \simeq \mathbf{A}$, a linearization of the inverse model does not involve the same parameters $(\mathcal{F}^{-1} \neq \mathbf{A}^{-1})$. In the inverse problem, a procedure for learning the model parameters that best fit the measurements is required. The method needs to be robust in the face of adverse numerical conditions.

4.1 Ill-posed and Ill-conditioned

Forward problems for physical scenarios tend to be well-posed while their inverse may be ill-posed (Hadamard, 1902). A well-posed problem has a solution that exhibits

the conditions of existence, uniqueness and stability so that it changes continuously
and slowly with the initial conditions. A problem that does not meet the criteria for
a well-posed problem is said to be ill-posed. For many impedance imaging problems,
the condition of stability is violated such that a small perturbation of the input
measurements from a system may cause a large change in the estimate of output
model parameters. This unstable behaviour makes it difficult to find a solution
because most solution methods assume that as the correct solution is approached
through variation of parameters, the error in the solution is reduced.

When a discrete approximation to an ill-posed problem is formulated, such as with
the FEM, instability typically translates to an ill-conditioned discrete linearization.
Even if originally well-posed, when a problem is discretized and suffers from an
accumulation of numerical, measurement or model errors, the problem may become
ill-conditioned. Ill-conditioning of discrete problems can be estimated by calculating
the condition number $\kappa(\cdot)$ of a matrix $\mathbf{A}$

$$\kappa(\mathbf{A}) = ||\mathbf{A}^{-1}|| \cdot ||\mathbf{A}|| \ \geq \ 1 \tag{4.2}$$

Examining the relationship of the singular values from a Singular Value Decomposi-
tion (SVD) can provide insight into the source of instabilities and possible remedies.
The SVD is formed as

$$\mathbf{U\Sigma V} = \mathbf{A} \qquad\qquad \text{SVD} \tag{4.3}$$

where $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors and $\mathbf{\Sigma}$ is a diagonal matrix of
the singular values in decreasing order from top-to-bottom. The condition number is

then the ratio of the largest to the smallest singular values $\kappa(\mathbf{A}) = \max(\mathbf{\Sigma})/\min(\mathbf{\Sigma})$. In finite precision calculations, a large condition number can serve as an indication that calculations are likely to become numerically unstable due to loss of precision[1]. The successive additions and multiplications of the algorithms combine finite numbers with a wide range of exponents to reduce the number of digits containing useful information (Goldberg, 1991; Higham, 2002). A widely used example of finite precision variables are the IEEE-754 64-bit double precision floating point numbers (IEEE, 2008). A "double" has 15 significant digits: 1 sign bit, 11 bits of exponent, 52 bits of fraction and one implied fractional bit that gives a smallest possible fraction $\epsilon = 2^{-53} \simeq 1.11 \times 10^{-16}$.

For an overdetermined inverse system, there are fewer unknowns $\boldsymbol{\sigma}$ than data $\mathbf{v}$; an example being to calculate a best-fit homogeneous conductivity σ_1.

$$\hat{\sigma}_1 = \arg\min_{\sigma_1} ||\mathcal{F}(\sigma_1, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_2 \tag{4.4}$$

Overdetermined systems typically meet the criteria for well-posed problems.

Conversely, underdetermined systems have many unknown model parameters $\boldsymbol{\sigma}$ and few measurements $\mathbf{v}$. If the measurements are consistent enough a best-fit estimate will approximate the true model parameters. When the measurements are noisy or otherwise contradictory, the system may be regularized to facilitate convergence to a solution. Convergence is achieved by removing problematic small singular values as indicated by the condition number.

---

[1]In general, a condition number $\kappa(\mathbf{A}) = 10^k$ indicates an expected loss of k digits of precision.

4.2 Regularization

Regularization may be considered a form of additional prior information applied with the intent of achieving a better conditioned problem. Regularization provides a reduction in the variance (error sensitivity) of a reconstruction in exchange for bias on the solution space. A problem with additional regularization can be restated as an optimization with new penalty terms

$$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma}} ||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}|| + ||\lambda\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)|| \tag{4.5}$$

Linear regularization schemes can be formulated as a weighted penalty $\lambda\mathbf{R}$ on changes in the model parameters $\boldsymbol{\sigma}$ away from a prior assumed value $\boldsymbol{\sigma}_*$.

The simplest example is Tikhonov regularization which uses an identity matrix to regularize the problem $\mathbf{R} = \mathbf{I}$ (Tikhonov and Arsenin, 1977). Generalized Tikhonov regularization may be applied to all problems of the same form as (4.5). The Tikhonov regularization penalizes excessively large model parameters $\boldsymbol{\sigma}$. The hyperparameter λ adjusts the strength of the regularization relative to the measurement residual. Setting the hyperparameter to zero ($\lambda = 0$) removes the regularization. Note that removing the prior ($\boldsymbol{\sigma}_* = 0$) is the same as assuming the prior was zero everywhere.

A Truncated Singular Value Decomposition (TSVD) calculates the SVD of a matrix but drops all singular values less than a particular threshold. By examining the SVD singular values, it can be shown that using the TSVD is a form of Tikhonov regularization (Hansen, 1998). The truncation threshold for the TSVD determines the maximum value for which "small" singular values will be set to zero. The hy-

perparameter λ in Tikhonov regularization can be adjusted to achieve the same effect as a specific TSVD truncation threshold. The Tikhonov hyperparameter value to achieve an equivalent TSVD threshold will vary depending on the measurement residual $||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||$ and is problem dependent.

Laplace regularization is similar to Tikhonov regularization but applies a discrete derivative to the solution parameters through the regularization matrix $\mathbf{R}$ to enforce a non-smoothness penalty. The non-smoothness penalty is achieved by taking each parameter i as the diagonal parameters of the regularization matrix $\mathbf{R}_{i,i} = N$ and subtracting the N adjacent parameters j such that $\mathbf{R}_{i,j} = -1$. In the simple case of a one-dimensional problem with four parameters

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{4.6}$$

In two or more dimensions, the direction of the smoothing can be selected by careful application of these discrete derivatives.

The two terms,

$$||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}|| \qquad \text{measurement residual and} \qquad (4.7)$$

$$||\lambda \mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)|| \qquad \text{regularization,} \qquad (4.8)$$

may have different norms applied. For example, the measurement error can be treated as an L2-norm while the regularization is treated as an L1-norm. The L1-

norm sums two errors of half magnitude to equal a single error of full magnitude

$$||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_1 = \sum_i (\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}_i, \mathcal{T}_i) - \mathbf{v}_i) \tag{4.9}$$

while the L2-norm will penalize the single error more strongly since errors are sum-squared.

$$||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_2^2 = \sum_i (\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}_i, \mathcal{T}_i) - \mathbf{v}_i)^2 \tag{4.10}$$

An L2-norm $||\cdot||_2$ penalizes terms as though they were orthogonal and unrelated. An L1-norm $||\cdot||_1$ takes the sum of the errors, treating all differences as equivalent. The L2-norm has the effect of smoothing edges while the L1-norm tends to preserve those edges. The L1-norm preserves edges because a sharp transition between two values is not penalized as strongly as with an L2-norm. The L2-norm solution can be strongly skewed by outliers or erroneous imputed data due to the squaring of errors. The choice of which norm is appropriate is problem dependent; problems with strong discontinuities or numerous data outliers may be better served by the L1-norm.

In this work, we have focused on the L2-norm. For further information on the L1-norm and specific requirements around working with the non-differentiable regions of the Total Variation (TV) L1 solution through the Primal-Dual Interior Point Method (PD-IPM) see (Rudin et al., 1992; Borsic and Adler, 2012).

4.3 Hyperparameter Selection

Where a hyperparameter λ is present, determining its value becomes a secondary optimization problem. Two well understood methods that do not require the measurement noise to be known *a priori* are commonly employed: the L-curve and the Generalized Cross-Validation (GCV) methods. Other methods include the discrepancy principle, restricted maximum likelihood and the unbiased predictive risk estimator (Morozov, 1966; O'Sullivan, 1986). We focus here on the widely used L-curve and GCV methods to give some flavour for how these methods compare.

The L-curve method solves the problem many times in searching for an optimal hyperparameter value (Hansen, 1992). The procedure is to select a range of hyperparameters and plot a graph of measurement mismatch $||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||$ against regularization terms $||\lambda \mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)||$. For many problems, the curve will look somewhat like a curved L when plotted on a log-log scale. When the hyperparameter is very small, the measurement mismatch is large. When the hyperparameter is very large, the regularization term is large. Somewhere between these extremes a hyperparameter value may be found that minimizes both terms at the point of maximum curvature (Figure 4.2a). For measurements with large noise η or other sources of error, the "L-curve" can be stretched beyond recognition.

The GCV method is a leave-one-out method that solves the inverse problem while leaving out a single measurement and then calculates the residual for that measurement (Wahba and Wang, 1990). This procedure is repeated for each measurement, and the sum of the misfits gives the GCV misfit. These values are plotted against a range of hyperparameters, and the minimum GCV misfit gives the optimal hyperparameter value (Figure 4.2b).
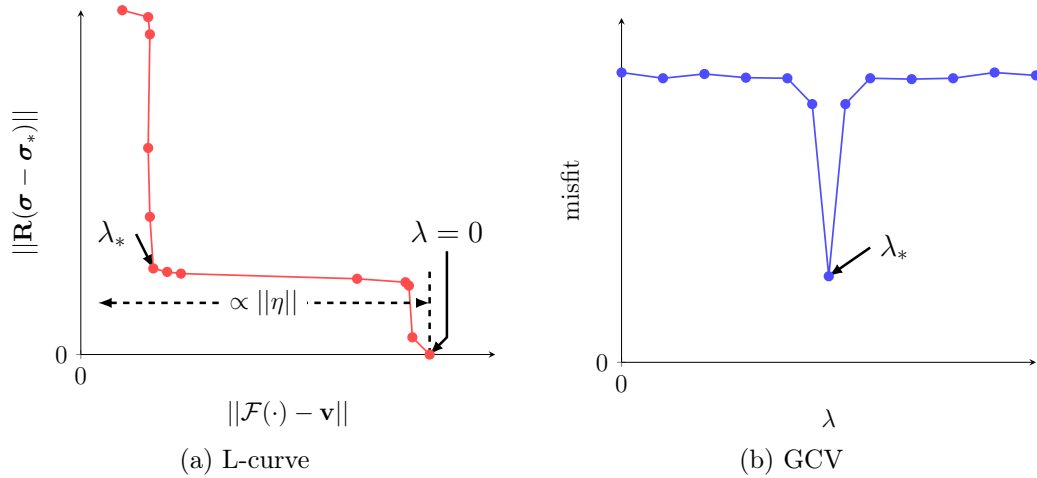
Figure 4.2: Idealized L-curve and Generalized Cross Validation (GCV) plots; (a) an L-curve plots measurement residual $||\mathcal{F}(\cdot) - \mathbf{v}||$ against regularization residual $||\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)||$ for a range of hyperparameters λ, the optimal value λ_* is at the inner-corner of the "L," maximum measurement error $\lambda = 0$ is proportional to measurement noise η, (b) GCV plots a range of λ against the sum of residuals for a single left-out measurement, the optimal value λ_* is the minimum misfit

Both methods are dependent on the specific problem, measurements and noise so that a hyperparameter value can not be chosen prior to evaluating the problem. The L-curve and GCV methods may fail on one problem and give very good results on another.

In general, these hyperparameter selection methods can be computationally expensive. For the L-curve, $n = 10$ or more inverse solutions must be found within a range of hyperparameters that brackets the optimal value with enough resolution to identify an appropriate point of curvature inflection. The GCV method expands the cost to nm forward solutions for m measurements. If single reconstructions take more than a short amount of time (short being subjective to the user), the additional cost of L-curve or GCV can quickly exceed the perceived value. Some reconstructions are compute limited and can already extend from hours to days to complete without

additional L-curve or GCV iterations. On the other hand, these methods provide an excellent retrospective check that an appropriate hyperparameter has been selected once other adjustments to the algorithm have stabilized.

In this work, initial hyperparameter values have been selected based on an SVD of the Jacobian $\mathbf{J}^\mathsf{T}\mathbf{J}$ to find the useful hyperparameter range (the singular value range for an initial homogeneous conductivity model) and an L-curve has then been plotted around a mid-point hyperparameter (singular value) to tune the selection. In an automated environment, a method would need to be selected that is tested against a range of problems likely to be seen in the field, but these issues are not addressed in this work.

4.4 Jacobian

A search direction is generally required to construct an inverse solution. This search direction may be calculated as a function of a linearized first derivative for changes in the measurements $\mathbf{v}$ with respect to the model parameters $\boldsymbol{\sigma}$. The first derivative is referred to as the Jacobian $\mathbf{J}$.

The Jacobian $\mathbf{J}$ is usually a dense matrix approximating the derivatives of a non-linear system $\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T})$ about a point $\boldsymbol{\sigma}$. For each model parameter perturbation $\partial\boldsymbol{\sigma}_j$ and measurement $\partial\mathbf{v}_i$, the Jacobian matrix is calculated as row i, column j such that

$$\mathbf{J}_{i,j} = \frac{\partial \mathbf{v}_i}{\partial \boldsymbol{\sigma}_j} = \frac{\partial \mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}_i, \mathcal{T}_i)}{\partial \boldsymbol{\sigma}_j} \tag{4.11}$$

There are two predominant methods for calculating an estimate of the Jacobian.

The simplest method is a direct numerical implementation of the Jacobian through the perturbation method. On the other hand, the adjoint method for the conductivity Jacobian is generally a more efficient technique.

In the perturbation method, each model parameter is perturbed independently and the forward problem is solved to estimate its effect on the measurements. Each element perturbation gives a column j of the Jacobian after applying each stimulus.

$$\mathbf{J}_j = \frac{\mathcal{F}(\boldsymbol{\sigma} + \delta\boldsymbol{\sigma}_j, \mathbf{B}, \mathcal{T})}{\delta\boldsymbol{\sigma}_j} \tag{4.12}$$

The key advantage of the perturbation method is its simplicity. There are two principle draw-backs: the computational cost of solving the forward problem once for every model parameter $\boldsymbol{\sigma}$, and the difficulty in choosing a "good" perturbation. A good perturbation is one that is sufficiently small so that it approximates the partial derivative, while being sufficiently large to avoid the accumulation of numerical errors.

The adjoint method for conductivity takes the dot product of the stimulus and measurement vector fields to calculate the Jacobian (Wang et al., 2001; Borsic et al., 2012). The stimulus field $\mathbf{E}_{i(stim)}$ is the current vector for each element N_j of the FEM mesh $-\nabla\mathcal{F}_\Omega(\boldsymbol{\sigma}, \mathbf{B}_k, \mathcal{T}_k)$ for a single stimulus k corresponding to measurement i. The measurement field $\mathbf{E}_{i(meas)}$ is found by swapping measurement and stimulus electrodes to create a new stimulus $\hat{\mathbf{B}}_i$ and solving the forward problem $-\nabla\mathcal{F}_\Omega(\boldsymbol{\sigma}, \hat{\mathbf{B}}_i, \mathcal{T}_i)$.

$$\mathbf{J}_{i,j} = \int_\Omega N_j \mathbf{E}_{i(stim)} \cdot \mathbf{E}_{i(meas)} \tag{4.13}$$

The forward solution typically solves the interior voltage distribution as a by-product

of finding the measurements at the electrodes. The solution for the interior potential would then be discarded, and a selection matrix would be used to construct the difference between potentials at the electrodes. In the adjoint method, these interior potentials are used to determine the fields: the negative gradient of the potential $\mathbf{E}_i = -\nabla\phi_i$. A measurement field is calculated as if it were a new stimulus using the measurement electrodes.

The adjoint method may reduce the number of calculations when the number of electrode pairs used for stimulus or measurements are fewer than the number of model parameters (for example, conductivity). An additional efficiency is gained by avoiding perturbations to the forward system matrix $\mathbf{A}$ which supports the reuse of the computationally expensive decomposition (QR, LU, SVD).

The adjoint method is nearly always a better choice when available; it is more efficient to calculate and numerically stable. Note that fields for each element need first be calculated for all electrode combinations (a forward solution without system matrix perturbations for each electrode combination followed by calculations of the field from the potential); then rows of the Jacobian can be calculated for each measurement i

$$\mathbf{J}_{i,j} = v_j(\mathbf{E}_{i,j:(stim)} \cdot \mathbf{E}_{i,j:(meas)}) \tag{4.14}$$

$$\mathbf{J}_i = \mathbf{V}(\mathbf{E}_{i:(stim)} \cdot \mathbf{E}_{i:(meas)}) \tag{4.15}$$

where $\mathbf{V}$ is a diagonal matrix of the m element volumes v_j and $\mathbf{E}_{i(stim)}$, $\mathbf{E}_{i(meas)}$ are $3 \times m$ matrices of the current vector dot-products.

The "standard method" takes this same adjoint approach (§C)

$$\mathbf{J}_\sigma = -\mathcal{T}\mathbf{A}^{-1}\mathbf{C}^{\mathsf{T}}\mathbf{S}\frac{\partial \mathbf{D}}{\partial \sigma_k}\mathbf{CX} \tag{4.16}$$

for measurement selection operator $\mathcal{T}$, system matrix $\mathbf{A}$, connectivity matrix $\mathbf{C}$ assigning local nodes to global mesh nodes, a shape function $\mathbf{S}$, and a conductivity $\mathbf{D}$ which is currently varied one element of the parametrization at a time.

4.5 Parameter Spaces

The parameter space in which an inverse problem is reconstructed need not be the same as the forward problem. An example is a three-dimensional forward problem which models the current flow from small electrodes on a large horizontal surface. The three-dimensional model of current flow is important; current flows throughout the medium and radiates in all directions from the stimulus electrode. In the inverse problem, the conductivity might only vary in a two-dimensional vertical plane under the electrodes. A mapping must project between the two parametrizations.

Selecting appropriate parametrizations can be challenging. In general, parametrizations that align with real properties of the physical system have the best opportunity to improve a reconstruction. The mapping from inverse to forward problem spaces is an important technique for controlling the ill-conditioning of a problem by reducing the degrees of freedom. Having fewer inverse parameters to search for while simultaneously having a highly accurate forward model may improve the conditioning of a problem.

Where a mapping $\mathbf{M}$ between the inverse and forward parameters is used, the
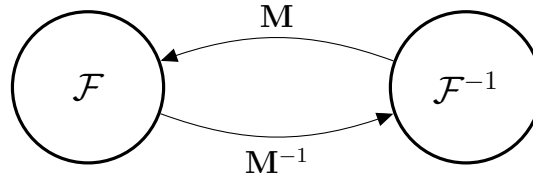
Figure 4.3: Forward and inverse mappings; a mapping $\mathbf{M}$ takes the inverse parameter space to the forward parameter space, a reverse mapping $\mathbf{M}^{-1}$ converts back to the inverse parameter space.

mapping $\mathbf{M}^{-1}$ translates the Jacobian from the forward space to the inverse space (Figure 4.3). The calculation of these mappings for irregular FEM meshes is complex; there are degenerate cases that must be safely detected. Overlap volume for tetrahedra in finite precision math depends on a few primitives that must be well implemented (Skiena, 2008, §13.9, §17.1). There are two common approaches to developing these primitives: estimation of tolerances and exact adaptive precision (Fortune, 1996; Shewchuk, 1997).

In simulation, selecting an inverse parametrization that exactly matches the forward problem is known as an "inverse crime." Such a choice ignores discretization errors and may lead to unreasonably good predictions of algorithm performance (Kaipio and Somersalo, 2007; Wirgin, 2004).

4.6 Measurement and Parameter Weighting

A measurement weighting term $\mathbf{W}$ can be applied based on noise estimates or other reliability criteria. The weighting matrix is commonly defined as the inverse noise covariance of uncorrelated Additive Gaussian White Noise (AGWN). For AGWN, the weighting matrix is a diagonal matrix with an entry for each measurement i

according to the estimate of its noise variance $\mathbf{W}_{i,i} = 1/\text{var}(\mathbf{v}_i)$. This term is applied by modifying the measurement residual L2-norm

$$||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_2^2 \quad \rightarrow \quad ||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_{\mathbf{W}}^2 \tag{4.17}$$

$$||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_{\mathbf{W}}^2 = \sum_i \mathbf{W}_{i,i}(\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}_i, \mathcal{T}_i) - \mathbf{v}_i)^2 \tag{4.18}$$

The result will be a set of weighted residuals, where noisy measurements are nearly zero and will no longer affect the choice of search direction. A new hyperparameter λ may be required to achieve similar levels of regularization when noise variance changes.

To avoid having to select a new hyperparameter when weights are adjusted, the measurement variance $\text{var}(\mathbf{v})$ or some equivalent measure of data quality may be scaled

$$\mathbf{W}_{i,i} = \exp\left(-\frac{\text{var}(\mathbf{v}_i)}{\max(\text{var}(\mathbf{v}))}\right) \; ; \qquad 0 < \text{var}(\mathbf{v}_i) < \max(\text{var}(\mathbf{v})) \tag{4.19}$$

so that a small quantity of noise will result in a nearly unit weighting $\mathbf{W}_{i,i} \simeq 1$ and the measurement with maximum noise will be de-weighted $\mathbf{W}_{i,i} \simeq 0$. The denominator $\max(\text{var}(\mathbf{v}))$ could be selected *a priori* based on system characterization so that a very low noise set of measurements would not be unduly penalized. For a nonlinear scaling of the measurements, the noise estimates should be similarly scaled.

The model parameters $\boldsymbol{\sigma}$ may also be weighted as a diagonal matrix $\mathbf{S}$ on the

L2-norm of the prior term

$$||\lambda\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)||_2^2 \quad \rightarrow \quad ||\lambda\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)||_{\mathbf{S}}^2 \tag{4.20}$$

$$||\lambda\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*)||_{\mathbf{S}}^2 = \lambda^2 \sum_j \mathbf{S}_{j,j}(\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_*))^2 \tag{4.21}$$

A common choice for parameter weighting is to normalize the parameters based on their sensitivity; insensitive elements far from the electrodes are de-weighted[2].

$$\mathbf{S}_{j,j} = \left(\sum_i \mathbf{J}_{i,j}{}^2\right)^k = ||\mathbf{J}_j||_2^{2k} \tag{4.22}$$

for measurement i and element j. A sensitivity-based weighting with Tikhonov regularization $\mathbf{R} = \mathbf{I}$ is commonly called the Noser prior (Cheney et al., 1990). Selecting $k = 1$ tends to push noise artifacts to the regions of the image with low sensitivity; away from the electrodes. Setting $k = 0$ (removing weighting matrix $\mathbf{S}$) tends to result in artifacts clustered near the electrodes and boundary. With $k = 0.5$, the calculation is the inverse 2-norm of a column of the Jacobian matrix $\mathbf{J}$. The intermediate $k = 0.5$ has been observed to distribute artifacts evenly throughout the image and gives updates that are almost independent of discretization geometry (Winkler and Rieder, 2015).

---

[2](Winkler and Rieder, 2015) have $\mathbf{S}_{\mathbf{J},\boldsymbol{\sigma}} = \mathbf{S}/\boldsymbol{\sigma}_0$: we note that for a homogeneous initial conductivity $\boldsymbol{\sigma}_0$, this amounts to a scaling of the hyperparameter $\lambda^2 \rightarrow \lambda^2/\boldsymbol{\sigma}_0$.

4.7 Measurement Scaling

A nonlinear scaling of the measurements is, at times, appropriate to achieve a more balanced weighting of measurements. For many EIT applications, electrodes are situated in a reasonably symmetric arrangement about the circumference of the object of interest. The average measurement between electrodes on these symmetric arrays, assuming a homogeneous conductivity and convex surface, tends to be close to other electrodes with similar spacing.

In arrangements where the electrodes are on an approximately planar surface, electrodes at the outer edges of the array can have significantly lower average measurements. The smaller edge measurements skew the reconstructed image because model parameter changes will not cause the same quantity of error in the cost function as more central measurements would.

Typically a measurement scaling using the geometric factor $\mathbf{G}$ is used to correct for this imbalance in measurement weightings. A diagonal matrix generated by computing the equivalent measurements on a homogeneous 1 S/m or, equivalently 1 Ohm·m, forward model $\mathcal{F}(1, \mathbf{B}, \mathcal{T})$ is constructed. The inverse of the resulting measurements becomes a geometric scaling factor where the measurements now have units of apparent resistivity ϱ

$$\mathbf{G}_\varrho(\mathbf{v}) = \frac{\mathbf{v}}{\mathcal{F}(1, \mathbf{B}, \mathcal{T})} \quad \rightarrow \quad \mathbf{G}_\varrho = \frac{1}{\mathcal{F}(1, \mathbf{B}, \mathcal{T})} \tag{4.23}$$

$$\mathbf{G}'_\varrho = \frac{\partial \mathbf{G}_\varrho(\mathbf{v})}{\partial \mathbf{v}} = \frac{1}{\mathcal{F}(1, \mathbf{B}, \mathcal{T})} = \mathbf{G}_\varrho \tag{4.24}$$

where $\mathbf{G}'$ is applied to the Jacobian $\mathbf{J}_\varrho = \mathbf{G}'\mathbf{J}_\sigma$ using the chain rule. In biomedical EIT, this scaling $\mathbf{G}$ has typically been referred to as "measurement normalization."

Geophysical ERT, where linear electrode arrays are wide-spread and absolute solutions are common, frequently uses apparent resistivity.

The log of apparent resistivity $\ln \varrho$ follows a similar mapping

$$\mathbf{G}_{\ln \varrho}(\mathbf{v}) = k \ln \mathbf{G}_{\varrho}(\mathbf{v}) = k \ln \frac{\mathbf{v}}{\mathcal{F}(1, \mathbf{B}, \mathcal{T})} \tag{4.25}$$

$$\mathbf{G}'_{\ln \varrho}(\mathbf{v}) = \frac{\partial \mathbf{G}_{\ln \varrho}(\mathbf{v})}{\partial \mathbf{v}} = \frac{\partial k \ln \mathbf{G}_{\varrho}(\mathbf{v})}{\partial \mathbf{G}_{\varrho}(\mathbf{v})} \frac{\partial \mathbf{G}_{\varrho}(\mathbf{v})}{\partial \mathbf{v}} = \frac{k}{\mathbf{v}} \frac{1}{\mathcal{F}(1, \mathbf{B}, \mathcal{T})} \tag{4.26}$$

Choosing $k = 1$ gives the natural log scaling $\ln(\cdot)$, while selecting $k = 1/\ln(10)$ gives the log base-10 scaling $\log_{10}(\cdot)$.

While uncommon, there are some situations where negative apparent resistivity does occur. When a measurement is small due to electrode geometry $\mathcal{F}(1, \mathbf{B}_i, \mathcal{T}_i) \to 0$ a small error in electrode placement can result in negative apparent resistivity. To manage electrode placement error, one may filter the negative apparent resistivity data based on a threshold geometric factor (Wilkinson et al., 2008). A region of high impedance contrast near electrodes may also give negative apparent resistivity. It may be better to avoid the use of apparent resistivity altogether when the source of the negative apparent resistivity is unclear, for example after attempting to remove negative values based on geometric considerations.

In some sense, a nonlinear scaling of the measurements is simple to implement because it is a one-way transformation; the measurements are rescaled and disappear into the solution without needing to compute the inverse mapping. This is not the case for model parameters in an iterative reconstruction; at each iteration the scaling is applied, and a new search vector in scaled-space is selected. The search vector then needs to be converted back to the original space to update the model parameters.

4.8 Model Parameter Scaling: Natural Limits

Inequality constraints may be implemented directly through specific algorithms such as quadratic programming (Bertsekas, 1982). On the other hand, unconstrained nonlinear solvers can be employed if an injective mapping is used to implement the constraints (Polydorides et al., 2014). The selected nonlinear mapping must be real analytic and monotonic. The mapping intentionally introduces nonlinearity in the solution space. The mappings appear in a similar form to the measurement scaling but (a) require an inverse mapping, and (b) are framed in the context of solution constraints. These constraints may be applied on a block-wise or per-parameter basis so that the constraints do not need to be uniform across the entire model parameter set.

To implement a positivity constraint on the model parameters ($\boldsymbol{\sigma} > 0$), a common choice for a mapping is into log units (Barber and Seagar, 1987). For each iteration, the model parameters are mapped to a new space $\boldsymbol{\sigma} \rightarrow \mathbf{m}$, a search direction is determined $\delta\mathbf{m}$, and the search direction is mapped back to the original parameter space $\delta\mathbf{m} \rightarrow \delta\boldsymbol{\sigma}$. The direction conversion is

$$\mathbf{N}_{\ln}(\boldsymbol{\sigma}) = \mathbf{m} = k \ln \boldsymbol{\sigma} \tag{4.27}$$

$$\mathbf{N}_{\ln}^{-1}(\mathbf{m}) = \boldsymbol{\sigma} = \exp\left(\frac{\mathbf{m}}{k}\right) \tag{4.28}$$

for conductivity restricted to $0 < \boldsymbol{\sigma} < \infty$ with $k > 0$ or for resistivity restricted to $0 < \boldsymbol{\rho} < \infty$ with $k < 0$ where $\boldsymbol{\rho} = 1/\boldsymbol{\sigma}$. For a small change in conductivity, the

chain rule for the Jacobian is calculated

$$\mathbf{J_m} = \frac{\partial \mathbf{v}}{\partial \mathbf{m}} = \underbrace{\frac{\partial \mathbf{v}}{\partial \boldsymbol{\sigma}}}_{\mathbf{J}_\sigma} \underbrace{\frac{\partial \boldsymbol{\sigma}}{\partial \ln \boldsymbol{\sigma}} \frac{\partial \ln \boldsymbol{\sigma}}{\partial k \ln \boldsymbol{\sigma}}}_{\mathbf{N}'_{\ln}(\boldsymbol{\sigma})} \tag{4.29}$$

$$\mathbf{N}'_{\ln}(\boldsymbol{\sigma}) = \frac{\boldsymbol{\sigma}}{k} \tag{4.30}$$

A change that exceeds the assumption of an infinitesimal change might be calculated when converting the line search result from log conductivity to conductivity units

$$\delta\boldsymbol{\sigma} = e^{\frac{\mathbf{m}+\delta\mathbf{m}}{k}} - e^{\frac{\mathbf{m}}{k}} = e^{\frac{\mathbf{m}}{k}}\left(e^{\frac{\delta\mathbf{m}}{k}} - 1\right) \simeq \boldsymbol{\sigma}\,\frac{\delta\mathbf{m}}{k} \tag{4.31}$$

where the approximation is a first-order Taylor series expansion of the natural exponent $\exp(x) = \sum_{n=0}^{\infty} x^n/n!$. As expected, the first-order Taylor series expansion is equal to the Jacobian chain rule (4.30). By rearranging (4.31) to find a multiplier to $\delta\mathbf{m}$, we get

$$\boldsymbol{\sigma}\,\frac{\delta\mathbf{m}}{k} = \delta\mathbf{m}\,\delta\mathbf{N}_{\ln}^{-1}(\boldsymbol{\sigma}) \tag{4.32}$$

$$\delta\mathbf{N}_{\ln}^{-1}(\boldsymbol{\sigma}) = \frac{\boldsymbol{\sigma}}{k} \tag{4.33}$$

Choosing $k = 1$ gives the natural log scaling $\ln(\cdot)$, while selecting $k = 1/\ln(10)$ gives the log base-10 scaling $\log_{10}(\cdot)$. Choosing $k = -1$ gives the natural log scaling in resistivity and $k = -1/\ln(10)$ gives log base-10 resistivity.

Typically, a new hyperparameter value λ is required when changing parameter spaces; the effect of the regularization will be similar but not identical. When sensitivity-based element weighting $\mathbf{N}$ is used, the new Jacobian $\mathbf{J_m}$ should be used

(a) logarithmic

(b) sigmoid; $p = 1, d = 3$
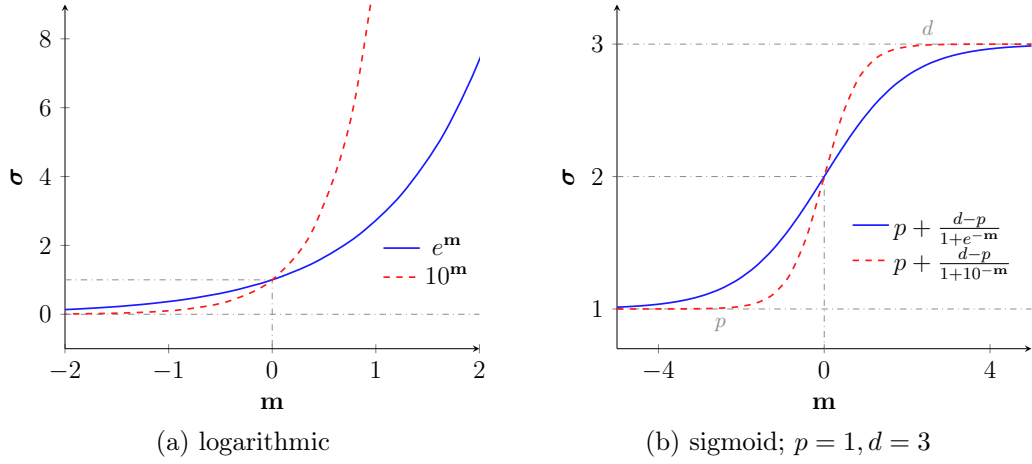
Figure 4.4: Natural limits: (a) logarithmic and (b) sigmoid mappings

to calculate element weighting in the modified parameter space $\mathbf{m}$.

Extending this method, a sigmoid function provides the ability to set upper d and lower p bounds on the model parameters in a similar fashion to the positivity constraint (Figure 4.4). The bounding is achieved using a sigmoid $\frac{1}{1+e^{-t}}$ or logistic regression function.

$$\mathbf{N}_{\curvearrowright}(\boldsymbol{\sigma}) = \mathbf{m} = -k \ln \left(\frac{d-p}{\boldsymbol{\sigma}-p} - 1 \right) \tag{4.34}$$

$$\mathbf{N}_{\curvearrowright}^{-1}(\mathbf{m}) = \boldsymbol{\sigma} = p + \frac{d-p}{1 + e^{-\mathbf{m}/k}} \tag{4.35}$$

for $0 < p < \boldsymbol{\sigma} < d < \infty$ with $k > 0$. The same procedure as for the log mapping is

applied again here. The Jacobian chain rule is calculated

$$\mathbf{J_m} = \frac{\partial \mathbf{v}}{\partial \mathbf{m}} = \underbrace{\frac{\partial \mathbf{v}}{\partial \boldsymbol{\sigma}}}_{\mathbf{J}_\sigma} \underbrace{\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{m}}}_{\mathbf{N}'_\smile(\boldsymbol{\sigma})} \tag{4.36}$$

$$\mathbf{N}'_\smile(\boldsymbol{\sigma}) = \frac{d-p}{k} \frac{e^{-\mathbf{m}/k}}{\left(1 + e^{-\mathbf{m}/k}\right)^2} \; ; \quad \text{for } \mathbf{m} = \mathbf{N}_\smile(\boldsymbol{\sigma}) \tag{4.37}$$

For a larger change in parameters, the change in conductivity $\delta\boldsymbol{\sigma}$ is calculated

$$\delta\boldsymbol{\sigma} = \mathbf{N}_\smile^{-1}(\mathbf{m} + \delta\mathbf{m}) - \mathbf{N}_\smile^{-1}(\mathbf{m}) \tag{4.38}$$

$$= \frac{(d-p)(1 - e^{-\delta\mathbf{m}/k})}{(1 + e^{-(\mathbf{m}+\delta\mathbf{m})/k})(1 + e^{\mathbf{m}/k})} \simeq \delta\mathbf{m}\, \delta\mathbf{N}_\smile^{-1}(\boldsymbol{\sigma}) \tag{4.39}$$

$$\delta\mathbf{N}_\smile^{-1}(\boldsymbol{\sigma}) = \frac{(\boldsymbol{\sigma}-p)}{k} \frac{1}{1 + e^{\mathbf{m}/k}} \; ; \quad \text{for } \mathbf{m} = \mathbf{N}_\smile(\boldsymbol{\sigma}) \tag{4.40}$$

For the sigmoid function, the Jacobian chain rule and the change in conductivity formulae are not the same. The sigmoid function may be used to replace the log conductivity constraint without modification to the rest of the algorithm.

In either the log or the sigmoid mapping, the direct mapping function $\mathbf{N}(\boldsymbol{\sigma})$, the Jacobian scaling $\mathbf{N}'(\boldsymbol{\sigma})$ and the updated inverse scaling $\delta\mathbf{N}^{-1}(\boldsymbol{\sigma})$ are used within the solver iterations. The prior $\boldsymbol{\sigma}_*$ and Jacobian $\mathbf{J}$ are scaled, and then, once the new search direction $\delta\mathbf{m}$ is determined, the scaling is reversed $\delta\mathbf{N}^{-1}(\boldsymbol{\sigma})$ to give the search direction in the original mapping $\delta\boldsymbol{\sigma}$.

| (a) model | (b) residual | (c) model | (d) residual |
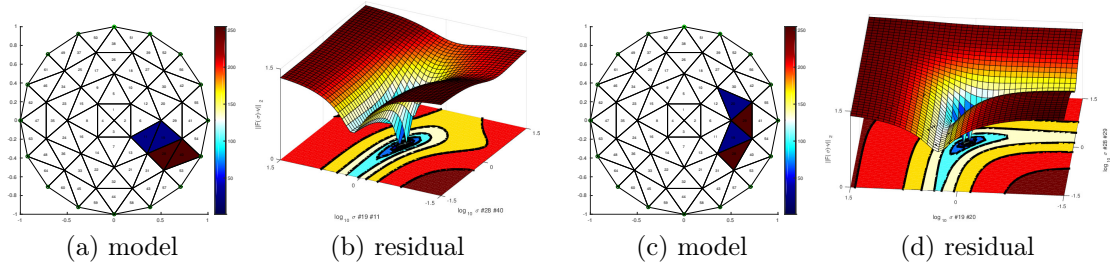
Figure 4.5: Residual plots; (a,c) a forward model with two regions, one of high conductivity ■ and one of low conductivity ■, (b,d) residual $\mathcal{R}$ as a function of the two sets of model parameters $\boldsymbol{\sigma}$, the plots exhibit "banana" shaped residuals: iterative (nonlinear) solvers are likely to handle these types of problems more successfully

4.9 Iterative Gauss-Newton

An iterative Gauss-Newton algorithm minimizes the L2-norm

$$\hat{\boldsymbol{\sigma}} = \arg \min_{\boldsymbol{\sigma}} ||\mathcal{F}(\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v})||_2^2 \tag{4.41}$$

The iterative nonlinear Gauss-Newton algorithm that solves (4.41) may be implemented as Algorithm 1. Reconstructions that exhibit asymmetric residual functions would benefit from such an iterative procedure. The algorithm can follow the nonlinear contours of the residual to find a minimum, rather than being restricted to minima that are linearly visible from the initial guess (Figure 4.5).

A homogeneous initial guess σ_0 can be estimated based on a scaling of the homogeneous solution $\mathcal{F}(1, \mathbf{B}, \mathcal{T})$. By applying Ohm's Law, one can find an L2-norm

---

Algorithm 1 Gauss-Newton
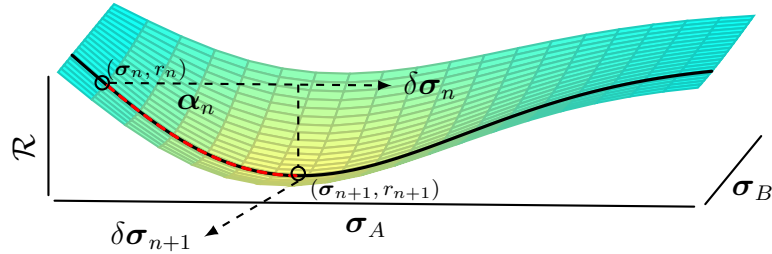
1: $\mathcal{F}_{\mathbf{b}}(\square) \leftarrow \mathcal{F}(\square, \mathbf{B}, \mathcal{T})$ ▷ for a fixed stim/meas pattern
2: **procedure** $\mathrm{GN}(\mathcal{F}_{\mathbf{b}}(\square), \mathbf{v})$ ▷ minimize (4.41)
3:     $\boldsymbol{\sigma}_0 \leftarrow (\mathbf{v}^\mathsf{T}\mathbf{v})^{-1}\mathbf{v}^\mathsf{T}\mathcal{F}_{\mathbf{b}}(1)$ ▷ initial guess as (4.44)
4:     $\mathcal{R}(\square) \leftarrow ||\mathcal{F}_{\mathbf{b}}(\square) - \mathbf{v}||_2$ ▷ residual func
5:     $r_0 \leftarrow \mathcal{R}(\boldsymbol{\sigma}_0)$
6:     $n \leftarrow 0$
7:     **repeat**
8:         $\mathbf{r}_n \leftarrow \mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n) - \mathbf{v}$
9:         $\mathbf{J}_n \leftarrow \partial\mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n)/\partial\boldsymbol{\sigma}_n$ ▷ Jacobian
10:         $\delta\boldsymbol{\sigma}_n \leftarrow -(\mathbf{J}_n^\mathsf{T}\mathbf{J}_n)^{-1}(\mathbf{J}_n^\mathsf{T}\mathbf{r}_n)$ ▷ $-\mathbf{J}^\dagger\mathbf{r}$
11:         $\alpha_n \leftarrow \mathrm{LINESEARCH}(\mathcal{R}(\square), \boldsymbol{\sigma}_n, \delta\boldsymbol{\sigma}_n)$ ▷ ---- for $0 < \alpha_n < 1$
12:         $\boldsymbol{\sigma}_{n+1} \leftarrow \boldsymbol{\sigma}_n + \alpha_n\,\delta\boldsymbol{\sigma}_n$
13:         $n \leftarrow n + 1$
14:         $r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma}_n)$
15:     **until** $(n > N)$ or $(r_n < \tau)$ or $(r_n - r_{n-1} < \kappa)$
16:     **return** $\boldsymbol{\sigma}_n$
17: **end procedure**

---

best-fit homogeneous solution $\hat{\sigma}_0$ using the pseudo-inverse of the measurements $\mathbf{v}$.

$$V_1 = IR_1 \; ; \;\; V_0 = IR_0 \qquad\qquad\qquad\qquad \text{where } R = 1/\sigma \qquad (4.42)$$

$$\frac{V_1}{R_1} = \frac{V_0}{R_0} \qquad\qquad \text{where } V_0 = \mathbf{v}, \, R_1 = 1, \, V_1 = \mathcal{F}(1, \mathbf{B}, \mathcal{T}) \qquad (4.43)$$

$$\hat{\sigma}_0 = \mathbf{v}^\dagger \mathcal{F}(1, \mathbf{B}, \mathcal{T}) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.44)$$

Iterations are halted when the solution converges to within an acceptable residual $(r_n < \tau)$, the iterations fail to improve the residual $(r_n - r_{n-1} < \kappa)$, or the iteration limit is exceeded $(n > N)$.

The line search to determine α_n can be computed using a range of algorithms. These algorithms are split into two classes: the exact and the inexact line searches. Exact line searches find the precisely optimal step size $(\partial \mathcal{R}/\partial \alpha = 0)$. An exact search is not generally applicable to nonlinear optimizations where the form of the residual function $\mathcal{R}$ must be tested rather than analytically derived. Inexact searches determine a sufficient step to ensure convergence without requiring the minimum to be precisely located (Nocedal and Wright, 1999). The Gauss-Newton method can make use of an inexact line search and still achieve a good convergence rate. Common inexact searches are iterative backtracking (Algorithm 2), golden section (Algorithm 3), and curve fitting searches (Algorithm 4) (Kiefer, 1953; Dennis and Schnabel, 1996). A careful combination of curve fitting and golden section search gives Brent's method (Brent, 1973). In general, a minimum is bracketed, and then, the range in which the minimum lies is localized sufficiently for the outer Gauss-Newton iterations to make further progress. In the limit, a random direction with a good line search can eventually find the minimum for many problems. The Wolfe or

the strong Wolfe conditions can be used to check if a step is "sufficient" by checking

the value and curvature of the cost function in the vicinity of the proposed step

(Armijo, 1966; Wolfe, 1969).

$$\mathcal{R}(\boldsymbol{\sigma} + \alpha\,\delta\boldsymbol{\sigma}) - \mathcal{R}(\boldsymbol{\sigma}) \leq \tau\alpha\,\delta\boldsymbol{\sigma}^{\mathsf{T}}\,\nabla\mathcal{R}(\boldsymbol{\sigma}) \quad \text{(Armijo rule: sufficient decrease)} \quad (4.45)$$

$$\delta\boldsymbol{\sigma}^{\mathsf{T}}\,\nabla\mathcal{R}(\boldsymbol{\sigma} + \alpha\,\delta\boldsymbol{\sigma}) > \kappa\delta\boldsymbol{\sigma}^{\mathsf{T}}\,\nabla\mathcal{R}(\boldsymbol{\sigma}) \quad\quad\quad\quad \text{(curvature condition)} \quad (4.46)$$

for $0 < \tau < \kappa < 1$. Typical values are $\tau = 1 \times 10^{-4}$ and $\kappa = 0.9$ or $\kappa = 0.1$. The

weak curvature condition (4.46) may be strengthened by replacing it with

$$|\delta\boldsymbol{\sigma}^{\mathsf{T}}\,\nabla\mathcal{R}(\boldsymbol{\sigma} + \alpha\,\delta\boldsymbol{\sigma})| \leq \kappa|\delta\boldsymbol{\sigma}^{\mathsf{T}}\,\nabla\mathcal{R}(\boldsymbol{\sigma})| \quad \text{(strong curvature condition)} \quad (4.47)$$

to arrive at the strong Wolfe conditions. The Wolfe conditions require the partial

derivatives of the residual function $\mathcal{R}$ which is unappealing when their calculation is

costly. The form of the residual $\mathcal{R}$ is generally not available for impedance imaging

reconstructions which makes the approximation of the partial derivative $\nabla\mathcal{R}$ through

perturbation time consuming and numerically imprecise. In this work, a fitted fifth-

order polynomial line search was employed ($N = 5$ maximum iterations) without

explicit calculation of the Wolfe conditions (stop if $r_0 - r_n > \sqrt{\epsilon}$, for $\epsilon = 2 \times 10^{-16}$).

Impedance imaging problems are typically ill-posed and ill-conditioned so that the

standard Gauss-Newton algorithm (Algorithm 1) will not converge. Regularization

may be applied to stabilize the solution. Restated as a regularized minimization of

---

Algorithm 2 Line Search: Backtracking

$\alpha = 0 \qquad\qquad \alpha \leftarrow b\,\alpha$
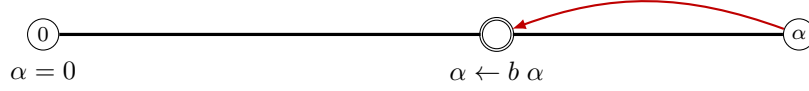
1: **procedure** BACKTRACKINGLINESEARCH$(\mathcal{R}(\square), \boldsymbol{\sigma}, \delta\boldsymbol{\sigma}, \tau, b)$ $\qquad \triangleright\ 0 < b < 1$
2: $\qquad \alpha \leftarrow 1$
3: $\qquad r_0 \leftarrow \mathcal{R}(\boldsymbol{\sigma}); \quad r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma} + \alpha\,\delta\boldsymbol{\sigma})$
4: $\qquad$ **while** $r_n - r_0 \leq \tau\alpha\,\delta\boldsymbol{\sigma}^\mathsf{T}\,\nabla\mathcal{R}(\boldsymbol{\sigma})$ **do** $\qquad \triangleright$ Armijo condition as (4.45)
5: $\qquad\qquad \alpha \leftarrow b\,\alpha$ $\qquad\qquad\qquad\qquad\qquad \triangleright$ backtrack $\longrightarrow$
6: $\qquad\qquad r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma} + \alpha\,\delta\boldsymbol{\sigma})$ $\qquad\qquad\qquad\qquad \triangleright$ for new ⓐ
7: $\qquad$ **end while**
8: $\qquad$ **return** α
9: **end procedure**

Algorithm 3 Line Search: Golden Section
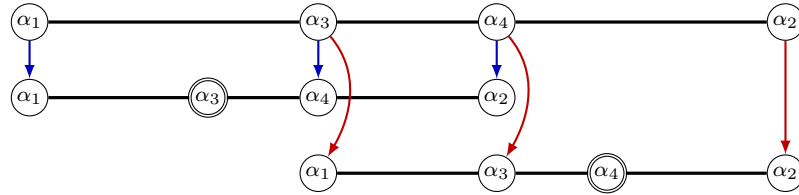
1: **procedure** GOLDENSECTIONLINESEARCH$(\mathcal{R}(\square), \boldsymbol{\sigma}, \delta\boldsymbol{\sigma}, \tau)$
2: $\qquad \phi \leftarrow \frac{1}{2}(1 + \sqrt{5}) \simeq 1.618$ $\qquad\qquad\qquad \triangleright$ Golden Section constant
3: $\qquad \alpha_1 \leftarrow 0; \quad \alpha_2 \leftarrow 1; \quad \alpha_3 \leftarrow \frac{1}{\phi^2}; \quad \alpha_4 \leftarrow \frac{1}{\phi}$ $\qquad \triangleright$ test points
4: $\qquad$ **repeat**
5: $\qquad\qquad r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma} + \alpha_n\,\delta\boldsymbol{\sigma})$ $\qquad\qquad\qquad \triangleright$ for new ⓐₙ
6: $\qquad\qquad$ **if** $r_1 > r_2$ **then** $\qquad\qquad\qquad\qquad \triangleright$ drop α_1 $\longrightarrow$
7: $\qquad\qquad\qquad \alpha_1 \leftarrow \alpha_3; \quad \alpha_3 \leftarrow \alpha_4; \quad \alpha_4 \leftarrow \frac{1}{\phi}\alpha_2$ $\qquad \triangleright$ new α_4
8: $\qquad\qquad\qquad r_1 \leftarrow r_3; \quad r_3 \leftarrow r_4$
9: $\qquad\qquad$ **else** $\qquad\qquad\qquad\qquad\qquad\qquad \triangleright$ drop α_2 $\longrightarrow$
10: $\qquad\qquad\qquad \alpha_2 \leftarrow \alpha_4; \quad \alpha_4 \leftarrow \alpha_3; \quad \alpha_3 \leftarrow \frac{1}{\phi^2}\alpha_2$ $\qquad \triangleright$ new α_3
11: $\qquad\qquad\qquad r_2 \leftarrow r_4; \quad r_4 \leftarrow r_3$
12: $\qquad\qquad$ **end if**
13: $\qquad$ **until** $|\alpha_2 - \alpha_1| < \tau(|\alpha_3| + |\alpha_4|)$ $\qquad \triangleright$ min spacing, typically $\tau = \sqrt{\epsilon}$
14: $\qquad$ **return** $(\alpha_1 + \alpha_2)/2$
15: **end procedure**

Algorithm 4 Line Search: Curve Fit
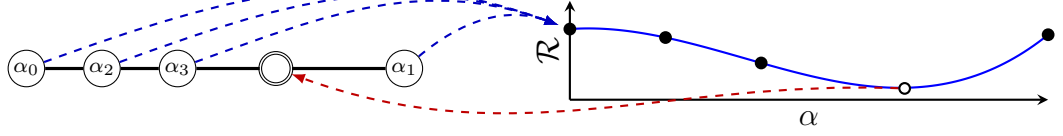
1: **procedure** CURVEFITLINESEARCH($\mathcal{R}(\square), \boldsymbol{\sigma}, \delta\boldsymbol{\sigma}, \tau$)
2:    $\alpha_1 \leftarrow 1; \quad n \leftarrow 1$
3:    $r_0 \leftarrow \mathcal{R}(\boldsymbol{\sigma}); \quad r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma} + \alpha \, \delta\boldsymbol{\sigma})$
4:    **while** $r_n - r_0 \leq \tau\alpha \, \delta\boldsymbol{\sigma}^\mathsf{T} \, \nabla\mathcal{R}(\boldsymbol{\sigma})$ **do**   ▷ Armijo condition as (4.45)
5:        $\alpha_n \leftarrow \text{FIT}(\alpha_{0..n-1}, r_{0..n-1})$ ▷ fit test points -▸, select minima ◂-·
6:        $r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma} + \alpha_n \, \delta\boldsymbol{\sigma})$ ▷ for new ⓐ
7:        $n \leftarrow n + 1$
8:    **end while**
9:    **return** α
10: **end procedure**

the L2-norm, additional regularization terms are added to (4.41),

$$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma}} ||\mathcal{F}(\mathbf{M}\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v}||_{\mathbf{W}}^2 + ||\lambda\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_\star)||_2^2 \qquad (4.48)$$

and iteratively updated by moving towards the inverse model parameters $\boldsymbol{\sigma}_n$ that minimize the residual (Algorithm 5). Additions to Algorithm 5 with respect to Algorithm 1 are highlighted.

The matrix $\mathbf{M}$ maps the inverse problem parametrization to the forward problem parametrization. When the forward and inverse parametrizations are FEM elements, the mapping is a geometric mapping between the two. Note that both the forward $\mathbf{M}$ and reverse $\mathbf{M}^{-1}$ mappings are needed in the Gauss-Newton iterations. The prior $\boldsymbol{\sigma}_*$ may be set to the initial guess $\hat{\boldsymbol{\sigma}}_0$ or some other value. A null prior estimate $(\boldsymbol{\sigma}_\star = 0)$ is a common choice.

Algorithm 5 Regularized Gauss-Newton

1: $\mathcal{F}_{\mathbf{b}}(\square) \leftarrow \mathcal{F}(\mathbf{M}\square, \mathbf{B}, \mathcal{T})$ ▷ fwd model
2: **procedure** REGGN($\mathcal{F}_{\mathbf{b}}(\square), \mathbf{v}, \mathbf{W}, \mathbf{R}, \mathbf{M}^{-1}$)   ▷ minimizes (4.48)
3:       $\boldsymbol{\sigma}_0 \leftarrow (\mathbf{v}^{\mathsf{T}}\mathbf{W}\mathbf{v})^{-1}\mathbf{v}^{\mathsf{T}}\mathbf{W}\mathcal{F}_{\mathbf{b}}(1)$ ▷ weighted initial guess as (4.44)
4:       $\boldsymbol{\sigma}_* \leftarrow \boldsymbol{\sigma}_0$ ▷ prior estimate
5:       $\mathcal{R}(\square) \leftarrow ||\mathcal{F}_{\mathbf{b}}(\square) - \mathbf{v}||_{\mathbf{W}} + ||\lambda\mathbf{R}(\square - \boldsymbol{\sigma}_*)||_2$ ▷ residual func
6:       $r_0 \leftarrow \mathcal{R}(\boldsymbol{\sigma}_0)$
7:       $n \leftarrow 0$
8:       **repeat**
9:            $\mathbf{r}_n \leftarrow \mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n) - \mathbf{v}$
10:           $\mathbf{J}_n \leftarrow \mathbf{M}^{-1}(\partial\mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n)/\partial\boldsymbol{\sigma}_n)$ ▷ Jacobian
11:           $\delta\boldsymbol{\sigma}_n \leftarrow -(\mathbf{J}_n^{\mathsf{T}}\mathbf{W}\mathbf{J}_n + \lambda^2\mathbf{R}^{\mathsf{T}}\mathbf{R})^{-1}(\mathbf{J}_n^{\mathsf{T}}\mathbf{W}\mathbf{r}_n + \lambda^2\mathbf{R}^{\mathsf{T}}\mathbf{R}(\boldsymbol{\sigma}_* - \boldsymbol{\sigma}_n))$
12:           $\alpha_n \leftarrow$ LINESEARCH($\mathcal{R}(\square), \boldsymbol{\sigma}_n, \delta\boldsymbol{\sigma}_n$)   ▷ for $0 < \alpha_n < 1$
13:           $\boldsymbol{\sigma}_{n+1} \leftarrow \boldsymbol{\sigma}_n + \alpha_n \delta\boldsymbol{\sigma}_n$
14:           $n \leftarrow n + 1$
15:           $r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma}_n)$
16:      **until** $(n > N)$ or $(r_n < \tau)$ or $(r_n - r_{n-1} < \kappa)$
17:      **return** $\boldsymbol{\sigma}_n$
18: **end procedure**

For difficult problems, where the regularization still results in iterations that fail to converge, constraints may be added in the form of weighting and scaling for both model parameters and measurements to minimize the functional

$$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma}}||\mathbf{G}(\mathcal{F}(\mathbf{M}\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - \mathbf{v})||_{\mathbf{W}}^2 + ||\lambda\mathbf{N}(\mathbf{R}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_\star))||_{\mathbf{S}}^2 \qquad (4.49)$$

which results in Algorithm 6. The constraints restrict the range of the algorithm's search to prevent finite precision overflow and restrict the solution space to physically realistic ranges. Additions to Algorithm 6 with respect to Algorithm 5 are highlighted.

For large inverse parametrizations, the dense matrix multiplication $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$ required to find $\delta\boldsymbol{\sigma}_n$ in (14) can be space prohibitive. An $m \times n$ Jacobian with $n =$

Algorithm 6 Scaled Regularized Gauss-Newton

1: $\mathcal{F}_{\mathbf{b}}(\square) \leftarrow \mathcal{F}(\mathbf{M}\square, \mathbf{B}, \mathcal{T})$ ▷ fwd model
2: **procedure** SCREGGN$(\mathcal{F}_{\mathbf{b}}(\square), \mathbf{v}, \mathbf{W}, \mathbf{R}, \mathbf{M}^{-1}, \mathbf{G}, \mathbf{G}', \mathbf{N}, \mathbf{N}', \delta\mathbf{N}^{-1})$ ▷ min (4.49)
3:      $\boldsymbol{\sigma}_0 \leftarrow (\mathbf{v}^{\mathsf{T}}\mathbf{W}\mathbf{v})^{-1}\mathbf{v}^{\mathsf{T}}\mathbf{W}\mathcal{F}_{\mathbf{b}}(1)$ ▷ weighted initial guess as (4.44)
4:      $\boldsymbol{\sigma}_* \leftarrow \boldsymbol{\sigma}_0$ ▷ prior estimate
5:      $n \leftarrow 0$
6:      **repeat**
7:          $\mathbf{r}_n \leftarrow \mathbf{G}(\mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n) - \mathbf{v})$
8:          $\mathbf{J}_n \leftarrow \mathbf{G}'\mathbf{M}^{-1}\left(\partial\mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n)/\partial\boldsymbol{\sigma}_n\right)\mathbf{N}'$ ▷ Jacobian
9:          **if** $n = 0$ **then**
10:            $\mathbf{S} \leftarrow$ SENSITIVITY$(\mathbf{J}_0, \mathbf{N}(\boldsymbol{\sigma}_0))$ ▷ from (4.22) (or $= \mathbf{1}$)
11:            $\mathcal{R}(\square) \leftarrow ||\mathbf{G}(\mathcal{F}_{\mathbf{b}}(\square) - \mathbf{v})||_{\mathbf{W}} + ||\lambda\mathbf{N}(\mathbf{R}(\square - \boldsymbol{\sigma}_*))||_{\mathbf{S}}$ ▷ residual func
12:            $r_0 \leftarrow \mathcal{R}(\boldsymbol{\sigma}_0)$
13:         **end if**
14:         $\delta\boldsymbol{\sigma}_n \leftarrow -(\mathbf{J}_n^{\mathsf{T}}\mathbf{W}\mathbf{J}_n + \lambda^2\mathbf{R}^{\mathsf{T}}\mathbf{S}\mathbf{R})^{-1}(\mathbf{J}_n^{\mathsf{T}}\mathbf{W}\mathbf{r}_n + \lambda^2\mathbf{R}^{\mathsf{T}}\mathbf{S}\mathbf{R}(\boldsymbol{\sigma}_* - \boldsymbol{\sigma}_n))\delta\mathbf{N}^{-1}$
15:         $\alpha_n \leftarrow$ LINESEARCH$(\mathcal{R}(\square), \boldsymbol{\sigma}_n, \delta\boldsymbol{\sigma}_n)$ ▷ for $0 < \alpha_n < 1$
16:         $\boldsymbol{\sigma}_{n+1} \leftarrow \boldsymbol{\sigma}_n + \alpha_n\,\delta\boldsymbol{\sigma}_n$
17:         $n \leftarrow n + 1$
18:         $r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma}_n)$
19:     **until** $(n > N)$ or $(r_n < \tau)$ or $(r_n - r_{n-1} < \kappa)$
20:     **return** $\boldsymbol{\sigma}_n$
21: **end procedure**

50,000 element inverse parametrization, when multiplied into an $m \times m$ dense matrix $\mathbf{J^TWJ}$ requires $50,000^2 \times 8$ Bytes of memory: 18 GB memory. Calculating (14) with a number of inner Preconditioned Conjugate Gradient (PCG) iterations is much more efficient. The PCG iterations do not require explicit construction of the $\mathbf{J^TWJ}$ matrix product (Hestenes and Stiefel, 1952).

To reconstruct a change in the parameters from a known baseline, one can substitute

$$\mathbf{v} \rightarrow \mathbf{v}_0 + \Delta\mathbf{v} \quad ; \qquad \boldsymbol{\sigma} \rightarrow \boldsymbol{\sigma}_0 + \Delta\boldsymbol{\sigma} \tag{4.50}$$

$$\mathbf{v}_0 = \mathcal{F}(\boldsymbol{\sigma}_0, \mathbf{B}, \mathcal{T}) \tag{4.51}$$

so that we now solve

$$\Delta\hat{\boldsymbol{\sigma}} = \arg\min_{\Delta\boldsymbol{\sigma}} ||\mathcal{F}(\boldsymbol{\sigma}_0 + \Delta\boldsymbol{\sigma}, \mathbf{B}, \mathcal{T}) - (\mathbf{v}_0 + \Delta\mathbf{v})||_\mathbf{W}^2 + ||\lambda\mathbf{R}(\Delta\boldsymbol{\sigma} - \Delta\boldsymbol{\sigma}_*)||_\mathbf{S}^2 \tag{4.52}$$

where $\Delta\hat{\boldsymbol{\sigma}}$ is the estimated change in the model parameters, $\Delta\boldsymbol{\sigma}_*$ is the prior estimate of the change in the expected solution ($\Delta\boldsymbol{\sigma}_* = \boldsymbol{\sigma}_* - \boldsymbol{\sigma}_0$), $\Delta\mathbf{v} = \mathbf{v}_B - \mathbf{v}_A$ is the difference between two sets of measurements, $\boldsymbol{\sigma}_0$ is a set of background model parameters, and $\mathbf{v}_0$ are the estimated measurements based on the forward model. The prior estimate is frequently assumed to be zero ($\Delta\boldsymbol{\sigma}_* = 0$). A unit conductivity homogeneous background is a common choice ($\boldsymbol{\sigma}_0 = 1$ S/m) or by homogeneous best-fit to measurements (4.44).

A single Gauss-Newton iteration is frequently sufficient when the parameter change is small $\Delta\boldsymbol{\sigma} \rightarrow 0$. The difference one-step Gauss-Newton solution can be

solved without a line search ($\alpha = 1$) as

$$\Delta\boldsymbol{\sigma} = -(\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J} + \lambda^2\mathbf{R}^\mathsf{T}\mathbf{S}\mathbf{R})^{-1}(\mathbf{J}^\mathsf{T}\mathbf{W}\Delta\mathbf{v} - \lambda^2\mathbf{R}^\mathsf{T}\mathbf{S}\mathbf{R}\Delta\boldsymbol{\sigma}_*) \qquad (4.53)$$

In solving the difference problem, any common errors between the two data sets tend to cancel. Errors in boundary geometry, electrode placement and contact impedance are not as important for difference reconstructions as when solving the absolute problem (4.49).

4.10 Conjugate Gradient

The nonlinear Conjugate Gradient method is algorithmically very similar to the Gauss-Newton method. The Conjugate Gradient method has the advantage of tracking the most recent search directions and being able to select an orthogonal search direction when it detects that the search direction is oscillating (Shewchuk, 1994). This conjugation behaviour allows the algorithm to quickly converge towards a minimum within a narrow "valley" in the solution space. The Gauss-Newton method, in this pathological case, will oscillate between the valley's sides and only converge slowly. The updates for a nonlinear Conjugate Gradient are solved as Algorithm 7 where additions with respect to the Gauss-Newton method (Algorithm 6) are highlighted.

The conjugate direction $\mathbf{s}_n$ is an adjusted search direction where the conjugate scalar β_n selects how much of the previous conjugate direction $\mathbf{s}_{n-1}$ to use. By setting ($\beta = 0$), one gets the Gauss-Newton iterations. The most common criteria for the

conjugate direction parameter β are (Nocedal and Wright, 1999)

$$\beta_{FR} = \frac{\delta\boldsymbol{\sigma}_n^{\mathsf{T}}\delta\boldsymbol{\sigma}_n}{\delta\boldsymbol{\sigma}_{n-1}^{\mathsf{T}}\delta\boldsymbol{\sigma}_{n-1}} \qquad\qquad \text{Fletcher-Reeves} \qquad (4.54)$$

$$\beta_{PR} = \frac{\delta\boldsymbol{\sigma}_n^{\mathsf{T}}(\delta\boldsymbol{\sigma}_n - \delta\boldsymbol{\sigma}_{n-1})}{\delta\boldsymbol{\sigma}_{n-1}^{\mathsf{T}}\delta\boldsymbol{\sigma}_{n-1}} \qquad\qquad \text{Polak-Ribière} \qquad (4.55)$$

$$\beta_{HS} = -\frac{\delta\boldsymbol{\sigma}_n^{\mathsf{T}}(\delta\boldsymbol{\sigma}_n - \delta\boldsymbol{\sigma}_{n-1})}{\mathbf{s}_{n-1}^{\mathsf{T}}(\delta\boldsymbol{\sigma}_n - \delta\boldsymbol{\sigma}_{n-1})} \qquad\qquad \text{Hestenes-Stiefel} \qquad (4.56)$$

$$\beta_{DY} = -\frac{\delta\boldsymbol{\sigma}_n^{\mathsf{T}}\delta\boldsymbol{\sigma}_n}{\mathbf{s}_{n-1}^{\mathsf{T}}(\delta\boldsymbol{\sigma}_n - \delta\boldsymbol{\sigma}_{n-1})} \qquad\qquad \text{Dai-Yuan} \qquad (4.57)$$

Only the Fletcher-Reeves criteria is guaranteed to converge, but in practice, using the Polak-Ribiére criteria with a negativity restriction $\max(0, \beta_{PR})$ works well for many problems by providing a reset to gradient descent (Dai and Yuan, 1999).

The Conjugate Gradient method requires the additional step of computing the conjugate direction $\mathbf{s}_n$ and conjugation parameter β_n, as well as storing the previous versions of the descent direction $\delta\boldsymbol{\sigma}_{n-1}$ and search direction $\mathbf{s}_{n-1}$, but the Conjugate Gradient algorithm typically converges more quickly than the Gauss-Newton algorithm.

As with the Gauss-Newton algorithm, iterations are halted when the solution converges to within an acceptable residual $(r_n < \epsilon)$, the iterations fail to improve the residual $(r_n - r_{n-1} < \tau)$, or the iteration limit is exceeded $(n > N)$.

Algorithm 7 Scaled Regularized Conjugate Gradient

1: $\mathcal{F}_{\mathbf{b}}(\square) \leftarrow \mathcal{F}(\mathbf{M}\square, \mathbf{B}, \mathcal{T})$ $\triangleright$ fwd model

2: **procedure** $\textsc{ScRegCG}(\mathcal{F}_{\mathbf{b}}(\square), \mathbf{v}, \mathbf{W}, \mathbf{R}, \mathbf{M}^{-1}, \mathbf{G}, \mathbf{G}', \mathbf{N}, \mathbf{N}', \delta\mathbf{N}^{-1})$ $\triangleright$ min (4.49)

3:     $\boldsymbol{\sigma}_0 \leftarrow (\mathbf{v}^{\mathsf{T}}\mathbf{W}\mathbf{v})^{-1}\mathbf{v}^{\mathsf{T}}\mathbf{W}\mathcal{F}_{\mathbf{b}}(1)$ $\triangleright$ weighted initial guess as (4.44)

4:     $\boldsymbol{\sigma}_* \leftarrow \boldsymbol{\sigma}_0$ $\triangleright$ prior estimate

5:     $n \leftarrow 0$

6:     **repeat**

7:         $\mathbf{r}_n \leftarrow \mathbf{G}(\mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n) - \mathbf{v})$

8:         $\mathbf{J}_n \leftarrow \mathbf{G}'\mathbf{M}^{-1}\left(\partial\mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n)/\partial\boldsymbol{\sigma}_n\right)\mathbf{N}'$ $\triangleright$ Jacobian

9:         **if** $n = 0$ **then**

10:             $\mathbf{S} \leftarrow \textsc{Sensitivity}(\mathbf{J}_0, \mathbf{N}(\boldsymbol{\sigma}_0))$ $\triangleright$ from (4.22) (or $= \mathbf{1}$)

11:             $\mathcal{R}(\square) \leftarrow ||\mathbf{G}(\mathcal{F}_{\mathbf{b}}(\square) - \mathbf{v})||_{\mathbf{W}} + ||\lambda\mathbf{N}(\mathbf{R}(\square - \boldsymbol{\sigma}_*))||_{\mathbf{S}}$ $\triangleright$ residual func

12:             $r_0 \leftarrow \mathcal{R}(\boldsymbol{\sigma}_0)$

13:         **end if**

14:         $\delta\boldsymbol{\sigma}_n \leftarrow -(\mathbf{J}_n^{\mathsf{T}}\mathbf{W}\mathbf{J}_n + \lambda^2\mathbf{R}^{\mathsf{T}}\mathbf{S}\mathbf{R})^{-1}(\mathbf{J}_n^{\mathsf{T}}\mathbf{W}\mathbf{r}_n + \lambda^2\mathbf{R}^{\mathsf{T}}\mathbf{S}\mathbf{R}(\boldsymbol{\sigma}_* - \boldsymbol{\sigma}_n))\delta\mathbf{N}^{-1}$

15:         $\beta_n \leftarrow \textsc{Conjugate}(\delta\boldsymbol{\sigma}_n, \delta\boldsymbol{\sigma}_{n-1}, \mathbf{s}_{n-1})$ $\triangleright$ $\beta_0 = 0$ then (4.54)–(4.57)

16:         $\mathbf{s}_n \leftarrow \delta\boldsymbol{\sigma}_n + \beta_n\mathbf{s}_n$

17:         $\alpha_n \leftarrow \textsc{LineSearch}(\mathcal{R}(\square), \boldsymbol{\sigma}_n, \mathbf{s}_n)$ $\triangleright$ for $0 < \alpha_n < 1$

18:         $\boldsymbol{\sigma}_{n+1} \leftarrow \boldsymbol{\sigma}_n + \alpha_n\,\mathbf{s}_n$

19:         $n \leftarrow n + 1$

20:         $r_n \leftarrow \mathcal{R}(\boldsymbol{\sigma}_n)$

21:     **until** $(n > N)$ or $(r_n < \tau)$ or $(r_n - r_{n-1} < \kappa)$

22:     **return** $\boldsymbol{\sigma}_n$

23: **end procedure**

4.11 Discussion

Regularization, scaling and weighting play an important role in achieving the best possible reconstruction. Collectively, these mathematical tools provide the machinery to reliably reconstruct challenging nonlinear inverse problems. Somewhat counter-intuitively, the complexity of these algorithms tends to mask subtle implementation errors. A robust algorithm can correct for these subtle errors by introducing artifacts so that detection and correction can be challenging.

Part II

Contributions

Chapter 5

Problems with Inverse Problems

Inverse problems lead to inherently low resolution images which can be difficult to verify. There are two primary reasons: ill-conditioning and ill-posed-ness. Sufficient regularization applied to an inverse problem will fix ill-conditioning by introducing bias in the reconstruction. Regularization of an ill-posed inverse problem selects one of many possible "right" solutions. On occasion, some part of a reconstruction will go poorly. Some failures are apparent: the hardware may not power on or an algorithm may produce Not a Number (NaN) floating point values. For the most part, impedance imaging hardware will provide data, and the algorithms will produce an image. The reconstructed impedance image may then lead to wrong interpretations. How should one go about determining whether a result is satisfactory? If the result is wrong, what caused the failure?

In this work, we report our processes, and the techniques used to find issues in the specific context of impedance imaging and, more generally, for inverse problems. We focus on the algorithmic aspects: the challenges in validating inverse problem

codes as well as their inputs and outputs. This represents early work on a novel aspect of inverse problems.

This work is naturally related to theories of software and algorithm debugging (Gill, 1951; Zeller, 2009), defect and root cause analysis (Ishikawa, 1982; Wilson et al., 1993) and business process analysis (Harry and Schroeder, 2000; Ohno, 1988). Methods for debugging software or hardware are very heuristic by nature; they tend to be problem dependent. The behaviour of experienced developers has been observed for generic software problems and various theories of hypothesis testing and defect finding proposed (Letovsky, 1987; Romero et al., 2007; Lawrance et al., 2013). The importance of a precise, concise description of a defect directly impacts the ability to reproduce and correct the root cause (Andersen and Fagerhaug, 2006). These observations are useful in giving some general advice on the process of finding defects: construct a clear problem description and then look to the design. Yet this advice is not directly useful in helping to diagnose a problem based on the available symptoms typical of an inverse problem. In the following, we propose to construct a process for evaluating the quality of a result. A systematic process such as this potentially allows the dissection of an algorithm to locate a specific cause within the code base or algorithm inputs.

5.1 Tests of Validity

In testing inverse problem codes, we draw a distinction between proofs of correctness[1] and the weaker claim of likely validity.[2] Correctness implies a pass or fail criteria

---

[1]"correct **a.** True, accurate; [...] *Hence* [...] correctness **n.**" (Fowler and Fowler, 1964)

[2]"valid **a.** [...] sound, defensible, well-grounded [...] *Hence* validity **n.**" (Fowler and Fowler, 1964)

which can be precisely defined. Validity, on the other hand, implies a statistical view about the likelihood of a solution. This distinction is particularly apparent for inverse problems where even a claim of validity can be difficult to construct because the distribution of solutions is generally not known *a priori* for measured data. Our inability to define an end-to-end outcome that is specific and correct drives our choice of testing strategy which is to break the problem into smaller pieces and apply the most appropriate method.

5.2 Challenges

Many computer programs may be decomposed into a sequence of boolean decisions. Exhaustively proving a correct implementation for such codes can be combinatorially overwhelming but is, ultimately, deterministic and lends itself to automation. Exploration of key points in the boolean decision space are possible through carefully crafted input sequences or by providing directed random stimulus and comparing the outcome to a simplified model. An automated and regularly executed test suite is desirable because it may capture the unintended side-effects of fixes or new features. An automated suite of tests is the standard software engineering solution for maintaining the quality and correctness of a code base (Rothermel et al., 2004).

Unconstrained optimization problems have "typical" test functions which may be applied to confirm that an algorithm can find the expected minima or compare performance (Moré et al., 1981). These test inputs generally represent low dimensional smooth cost functions so do not closely emulate expected inputs in an inverse problem. For inverse problems, it is inherently difficult to test a system which produces

(a) model   (b) no noise   (c) 12 dB noise   (d) elec. movement
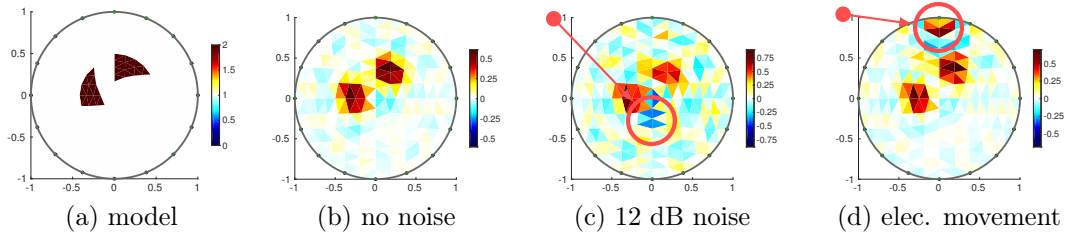
Figure 5.1: Sample reconstructions; (a) a forward model, (b) ideal reconstruction without noise, (c) reconstruction with 12 dB noise (artifact circled ●→○), (d) electrode movement (artifact circled ●→○)

images with significant element-to-element variance. Even with a gold standard image for comparison, two reconstructions with different noise samples can produce images with quite different artifacts. Methods such as Markov Chain Monte Carlo (MCMC) allow an empirical estimate of the source distribution per element (Kaipio et al., 2000), but such a sampling strategy requires a correct implementation to enable a correct estimate. Without knowing the true distribution, the image itself must be interpreted and judged equivalent: an inherently manual and human driven process.

A poor reconstruction with artifacts of a particular magnitude and distribution may mask the changes of interest: a false negative. Alternatively, artifacts may be incorrectly interpreted and assigned some meaning: a false positive. Experienced users can typically pick out particular patterns of artifacts to identify specific issues in the data. Examples are noisy measurements due to poor electrode connections or movement of the subject (Figure 5.1).

When debugging any problem, a clear idea of the expected outcome is important. When a result is outside of the expected range or somehow "unusual," a deeper investigation can be initiated. The ability to set an expectation is deeply connected

to an understanding of the system generating the result (Brooks, 1983). This expectation may be based on knowledge of the underlying mechanisms, physiology or geology and an understanding of the inner workings of the hardware and algorithms. Determining a "ground truth" for a site can be challenging for geological impedance reconstructions; sites are unique and cannot generally be disassembled to determine if the reconstruction matches the reality. A ground truth for geological applications is typically based on historical and current topological maps and site plans, geological maps and models, aerial photographs, intrusive sampling, a variety of sensor modalities (for example, nuclear or acoustic), borehole logs, trial pits, or other geophysical and geochemical investigations. Less reliably, an expectation may be based on a large number of previous reconstructions forming some empirical distribution. The ability to quickly set an expectation is then tied to a users experience with the system.

5.3 Proposed Process

Given the challenges inherent in demonstrating a correct or valid reconstruction, we propose that, for inverse problems in general, trust in the overall design may be constructed by testing units of functionality and then building upon that foundation. Individual components may (#1) be validated to perform as expected in isolation, or (#2) compared against some gold standard, possibly after applying an appropriate transformation. Next, (#3) the components can be reassembled and simulated data followed through the algorithmic machinery. To confirm overall correct operation, the outcome can be compared to a synthetic model which was used to construct

the input data. Finally, (#4) the real data can be compared to simulated data. A model can be manually refined to bring it into closer agreement with the real measurements. Model refinements may adapt the geometry, conductivity or other parameters to more closely match external data or hypotheses. If the two data sets appear reasonably aligned, the flow of the algorithm can be monitored between the two data sets to determine where processing diverges. The reconstructed image of the real data may be contrasted with other images, estimates of sensitivity, external information or extracted features (#5) to develop a meaningful interpretation of the image in context. These ideas are illustrated in Figure 5.2, inspired by Maslow's Hierarchy of Needs (Maslow, 1943). Ultimately, the desired outcome is to either (a) explain the unusual reconstruction as valid with a specific cause, (b) determine a mismatch in data quality, or (c) identify some flaw in the algorithm or its implementation. One would expect that, after some period of time with a stable code base, outcomes (a) and (b) would be predominant.

This philosophy of building up trust through proofs of correctness and validity is applied here to the absolute iterative Gauss-Newton/Conjugate Gradient solver, its input data and the resulting impedance reconstructions. To achieve these goals, two approaches present themselves. First, to identify units of the algorithm and independently confirm their correct operation. Second, to construct a means of monitoring the operation of the algorithm at various points in the iterations and in some sense develop a time series view of the algorithm behaviour around these points.
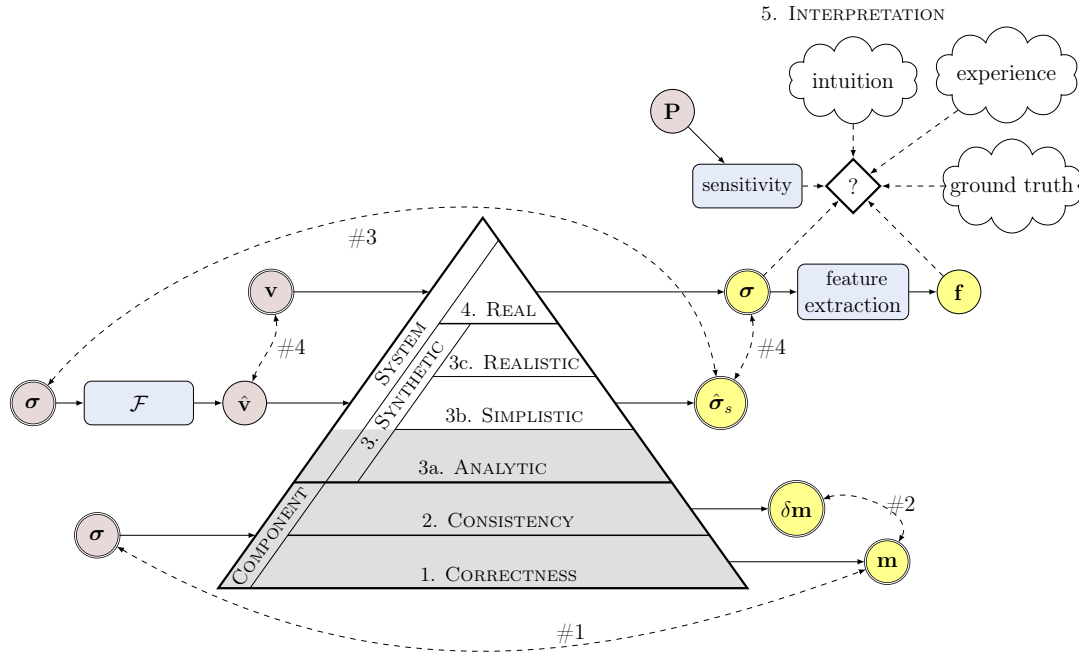
Figure 5.2: Proposed hierarchy of validation; proposed in this thesis, internal checks of components and system (grey background ▦, this chapter's focus) form a trusted platform upon which the input data may be compared to expectations (white background □, the following chapter's focus, Chapter 6), an interpretation (#5) decides whether a result is reasonable, (#4) input ◯ and output ◯ may be compared, (#3) synthetic data may be used to confirm correctness of the outcome for (a) analytic, (b) simplistic and (c) realistic data, analytic solutions provide a mechanism for independently confirming forward model $\boxed{\mathcal{F}}$ correctness, (#2) consistency between related groups of components may be validated and (#1) from each set of components checked for consistency at least one needs to be confirmed to provide correct results, symbols are listed in Table 5.1, points of comparison are identified with dashed lines ◀--▶

5.4 The Flow of Data

Tracing the data flow provides a visual roadmap. We use this roadmap for understanding how the components of the Gauss-Newton (Algorithm 6) and Conjugate Gradient (Algorithm 7) algorithms interact with the data set. Interactive algorithm visualization has had mixed success as a teaching aid in the classroom setting and debugging tasks are generally accelerated with textual logs rather than visual aids (Hundhausen et al., 2002). The flow chart of data movement (Figure 5.3), on the other hand, provides a quick reference to a detailed algorithm description. The data flow of the iterative solver is mapped out in detail in Figure 5.4 to aid in identifying potential targets for validation and monitoring. These targets would be points in the algorithm's flow where a sensible, low-dimensional interpretation of the current state can be made or where a set of these targets can be identified for comparison across or within iterations.

In the Gauss-Newton and Conjugate Gradient algorithms shown in Figure 5.3 and Figure 5.4, data flows from the inputs ◯ to the outputs ◯ over the course of a number of iterations. The main flow of the iterations is identified with heavier ⟶ arrows. Optional methods for constructing some inputs based on others are identified with dashed ⇢ arrows. Matrices multiplied with the data are identified with ▢ while more complex processing are shown with ▭. The symbols in the figure correspond to Algorithm 7 of Chapter 4 and are summarized in Table 5.1.
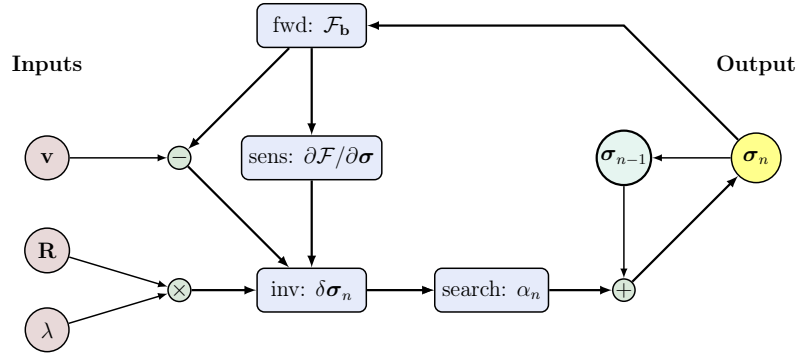
Figure 5.3: Representative iterative solver for inverse problems; an iterative Gauss-Newton solver (Algorithm 1) uses a forward model $\mathcal{F}_{\mathbf{b}}$ to estimate the sensitivity of a measurement to model parameter change σ, a misfit is calculated based on input data $\mathbf{v}$ and is used to estimate the next update direction $\delta\sigma$, a line search determines α how far in that direction to proceed, the process repeats until a proposed best fit for the data is determined
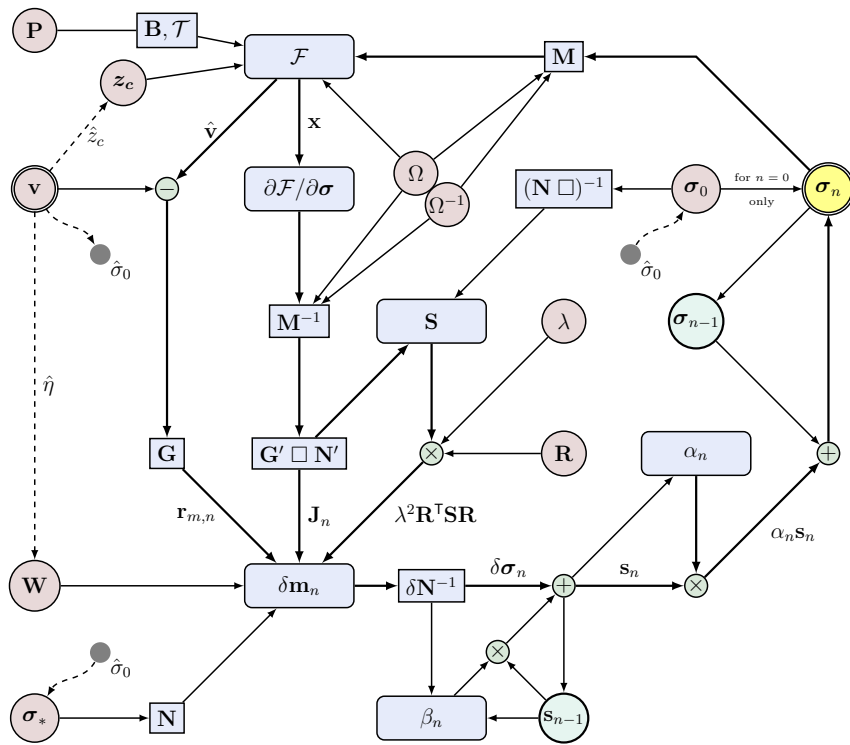
Figure 5.4: Gauss-Newton and Conjugate Gradient iterations; Algorithm 7 inputs ⬤, outputs ⬤, core data flow ⟶, optional input construction ---►, matrix multiplications ▢, equations and complex algorithms ⬭, symbols are listed in Table 5.1

Table 5.1: Gauss-Newton and Conjugate Gradient symbols

Symbol	Description
$\mathbf{N}, \mathbf{N}^{-1}, \mathbf{N}', \delta\mathbf{N}^{-1}$	model parameter scaling
$\mathbf{S}$	model parameter weighting
$\mathbf{G}, \mathbf{G}^{-1}, \mathbf{G}'$	measurement scaling
$\mathbf{W}$	measurement weighting
$\mathbf{M}, \mathbf{M}^{-1}$	forward $\leftrightarrow$ inverse parameter mapping
$\delta\mathbf{m}, \delta\boldsymbol{\sigma}$	descent direction; mapped, conductivity
$\mathbf{s}_n, \mathbf{s}_{n-1}$	search direction at iteration n and $n-1$
β_n	conjugate gradient update, $\beta = 0$ for GN
α_n	line search $(0 < \alpha_n \leq 1)$
$\partial\mathcal{F}/\partial\boldsymbol{\sigma}$	Jacobian $\mathbf{J}$
$\mathcal{F}$	forward model
Ω, Ω^{-1}	forward and inverse domains
$\boldsymbol{\sigma}_n, \boldsymbol{\sigma}_{n-1}$	model parameters at iteration n and $n-1$
$\boldsymbol{\sigma}_0, (\hat{\sigma}_0)$	(homogeneous) model parameter initial guess
$\boldsymbol{\sigma}_*$	model parameter prior
λ	hyperparameter
$\mathbf{R}$	regularization matrix
$\mathbf{v}$	measurements
$\mathbf{P}, \mathbf{B}, \mathcal{T}$	stimulus and measurement pattern
$\mathbf{z_c}(\hat{z}_c)$	contact impedance (estimate)
$\hat{\eta}$	measurement noise estimate

5.5 Consistency and Correctness

Figure 5.4 may be examined to identify matched sets of transformations; in a sense, we are checking for consistency of units. A scaling on one branch should be matched by a scaling on another, so that when two branches merge, data in the same units are combined. Some suggested consistency and correctness tests are listed in Table 5.2.

Injective mappings, a one-to-one mapping, may be checked for consistency directly ($\mathbf{M}^{-1} \circ \mathbf{M}(\boldsymbol{\sigma}) = \boldsymbol{\sigma}$). Derivatives may be checked through numerical differentiation across a range of perturbation lengths and test points ($\partial_{\mathbf{v}} \mathbf{G}(\mathbf{v}) = \mathbf{G}'(\mathbf{v})$). Ultimately, the correctness of implementation is most easily confirmed by plotting and visually checking for an appropriate range of values and expected slope. Visual checking has the added advantage of picking up on some types of unexpected outcomes.

Visual checks do not lend themselves to automated regression testing approaches (Memon, 2008). Reliably automating testing of numerical software that applies nonlinear functions and the matrix inverse to an ill-conditioned, ill-posed problem are inherently difficult; the range of "correct" answers is at times hard to specify exactly.

We propose that an appropriate approach is to confirm the correctness of detailed behaviour visually. Once visually checked, automated tests check the asymptotes or other stable properties of functions. In this way, the initial implementation is carefully checked, and an automated mechanism is constructed which would trigger this same detailed examination if any gross behaviours change within the function. This mixed approach offers a compromise between "one-off" testing and the extreme efforts required to achieve complete and automated test coverage.

Table 5.2: Examples of Consistency and Correctness

Symbol	Test
$\mathbf{N}, \mathbf{N}^{-1}, \mathbf{N}', \delta\mathbf{N}^{-1}$	plot $\mathbf{N}(\boldsymbol{\sigma}) = \mathbf{m}$
	$\mathbf{N}^{-1} \circ \mathbf{N}(\boldsymbol{\sigma}) = \boldsymbol{\sigma}$
	$\mathbf{N}' \circ \mathbf{N}^{-1}(\mathbf{m}) = \frac{\mathbf{N}^{-1}(\mathbf{m}+\delta\mathbf{m})-\mathbf{N}^{-1}(\mathbf{m})}{\delta\mathbf{m}}$
	$\delta\mathbf{N}^{-1} \circ \mathbf{N}^{-1}(\mathbf{m}) = \frac{\mathbf{N}^{-1}(\mathbf{m}+\delta\mathbf{m})-\mathbf{N}^{-1}(\mathbf{m})}{\mathbf{N}^{-1}(\delta\mathbf{m})}$
$\mathbf{G}, \mathbf{G}^{-1}, \mathbf{G}'$	plot $\mathbf{G}(\mathbf{v}) = \varrho$
	$\mathbf{G}^{-1} \circ \mathbf{G}(\mathbf{v}) = \mathbf{v}$
	$\mathbf{G}' \circ \mathbf{G}^{-1}(\varrho) = \frac{\mathbf{G}^{-1}(\varrho+\delta\varrho)-\mathbf{G}^{-1}(\varrho)}{\mathbf{G}^{-1}(\delta\varrho)}$
$\mathbf{S}$	image of diag($\mathbf{S}$)
$\mathbf{W}$	plot diag($\mathbf{W}$)
$\mathbf{M}, \mathbf{M}^{-1}$	image of $\boldsymbol{\sigma}$ in Ω and Ω^{-1} using $\mathbf{M}$
	$\mathbf{M}^{-1} \circ \mathbf{M}(\boldsymbol{\sigma}) = \boldsymbol{\sigma}$
α_n	plot for $\mathcal{F}$ analytic (low-dimensional) quadratic problems
$\mathcal{F}$	compare with analytic solutions
Ω, Ω^{-1}	obs. geometry matches expected
	obs. electrodes are correctly numbered and placed
$\mathbf{P}, \mathbf{B}, \mathcal{T}$	obs. correct assignment, magnitudes

An example of asymptotes that could be tested safely and consistently are the upper and lower bounds of an implementation of the sigmoid function. Another more subtle example is a consistency check $\mathbf{N}' \circ \mathbf{N}^{-1}(\mathbf{m}) = \partial_{\mathbf{m}} \mathbf{N}^{-1}(\mathbf{m})$ for log conductivity and sigmoid mappings which show the error introduced by using a first order approximation as well as finite precision numbers. The mappings are non-linear so that their error bounds also vary as a function of the non-linearity.

For most cases, it is sufficient to construct a worst case estimate of the error bounds, but this bound can at times be excessive to the point of defeating the construction of any useful test. The most common example of this phenomenon for inverse problems is in comparing noisy reconstructions where artifacts can be of the same magnitude as the true image: a maximum error bound of 100% of the true image. Some subset of checks will typically require on-going human intervention in the testing process.

We recommend automated checks of correctness should be run frequently: either at each code change committed to the repository or at a scheduled interval. A hybrid approach is sometimes applied where a quick subset of checks is run at every code change as a "sanity check" and the complete test suite is run at regular intervals. Reporting of the results should be automated so that all interested parties will be made aware of a failure. Independent changes may interact; considerable time can be saved when authors are made aware promptly so that they may collaborate. A report provides a record of what checks need maintenance and which checks are newly failing; this report may help isolate changes affecting the check. Failure of a check for correctness should trigger a visual comparison between old and new validation tests associated with that check to confirm that plots or figures remain consistent.

Regenerating or storing previous images used in a visual validation provide a record of "acceptable" examples should the original verifier become unavailable.

5.6 Analytic Models

Analytic models are geometrically simple domains for which mathematically tractable solutions are known. Analytic models are an important mechanism for proving a correct and sufficiently accurate FEM implementation of the forward model. Using a proven construction method for simple domains (an FEM assembly procedure), more complicated forward problems can be built. A proven forward model can then be used to build higher level checks on the overall inverse problem behaviour.

There are three main groups of analytic solutions for impedance imaging: models based on a homogeneous resistor, models incorporating a conductivity discontinuity, and models with continuous resistivity variation. The models may be applied in one, two or three dimensions. Resistor models are summarized in Table 5.3 for a homogeneous conductivity $\boldsymbol{\sigma}_0$ and contact impedance z_c. The three-dimensional models are drawn with a rectangular cross-section, but any fixed cross-section extruded in a straight line will give the same model for a given cross-sectional area. Conductivity $\boldsymbol{\sigma}_0$ has units of siemens per meter (S/m) while contact impedance z_c is generally defined as conductivity per unit surface area of the electrode ($\Omega \cdot \mathrm{m}^2$) in three dimensions. For contact impedance in simulation, the dimensionality of the surface area term is reduced to match the dimensions; the contact impedance units become $\Omega \cdot \mathrm{m}$ for two-dimensional models and Ω for one-dimensional models.

Contact impedance is sometimes referred to as "specific contact resistivity" ρ_c in

(a) four electrode measurements   (b) two electrode measurements
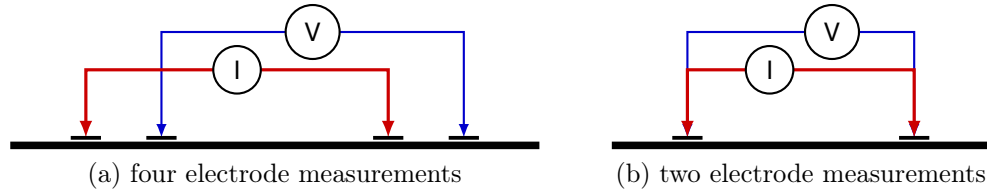
Figure 5.5: Contact impedance z_c measurement: the difference in (a) four electrode measurements V_4, and (b) two electrode measurements V_2 gives an estimate of the contact impedance $|V_4 - V_2|/I = z_c$

the context of metal-semiconductor contacts (Reeves and Harrison, 1982).

$$\rho_c = \left.\frac{\partial V}{\partial \bar{\mathbf{J}}}\right|_{V=0} \qquad \text{specific contact resistivity } [\Omega \cdot m^2] \qquad (5.1)$$
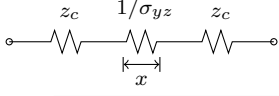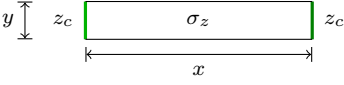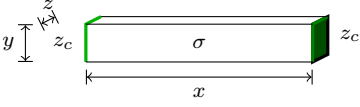
$$R_c = \frac{\rho_c}{A} \qquad \text{contact resistance } [\Omega] \qquad (5.2)$$

The specific contact resistivity is defined as the slope of the current-voltage I-V curve at $V = 0$. A linear I-V curve is referred to as an "Ohmic contact." Contact impedance is typically measured by taking the difference between a two electrode measurement and a four electrode measurement (Figure 5.5). Experimental measurement of contact impedance then incorporates wiring resistance, electrode properties and any transport phenomena across the electrodes. Impedance imaging hardware can collect data for an estimate of contact impedance by measuring at stimulus electrodes (two electrode measurements) and combining this with the normally collected four electrode measurements to calculate an estimate of contact impedance (Vilhunen et al., 2002; Winkler and Rieder, 2015).

Analytic models may be calculated by applying boundary conditions to the homogeneous solution. In one dimension, the analytic model is trivial: a simple resistor

Table 5.3: Resistor models

	σ Units	z_c Units	Model	Equation
1D	S·m	Ω	z_c $\quad$ $1/\sigma_{yz}$ $\quad$ z_c ; x	$R = 2z_c + \frac{x}{\sigma_{yz}}$
2D	S/□	$\Omega \cdot \mathrm{m}$	y ; z_c ; σ_z ; z_c ; x	$R = 2\frac{z_c}{y} + \frac{x}{y\sigma_z}$
3D	S/m	$\Omega \cdot \mathrm{m}^2$	z ; y ; z_c ; σ ; z_c ; x	$R = 2\frac{z_c}{yz} + \frac{x}{yz\sigma}$

model. In two dimensions, the Laplace equation

$$\nabla \cdot (\sigma \nabla u) = 0 \tag{5.3}$$

for conductivity σ and potential u may be written in polar coordinates for a homogeneous circular domain

$$\frac{1}{r} \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} = 0 \tag{5.4}$$

for $0 \leq r \leq R$ radial distance from the centre and angle $0 \leq \theta < 2\pi$. The current flow normal to the boundary at $r = R$ for real valued conductivity may be calculated through separation of variables which gives the Fourier series

$$u(r, \theta) = u_0 + \sum_{n=1}^{\infty} \left(\frac{r}{R}\right)^n a^n \cos(n\theta) + b_n \sin(n\theta) \tag{5.5}$$

for a_n, b_n restricted by the boundary conditions (Isaacson, 1986). The outer boundary conditions are that no current flow across the boundary except at electrode locations.

The outer boundary condition is then

$$u(R,\theta) + z_c \frac{\partial u}{\partial r}|_{r=R} = U_i \tag{5.6}$$

for non-overlapping electrodes of width w and radial location θ_i. One or more concentric circular inhomogeneities may be added within the unit circle (Figure 5.6a). To solve, continuous boundary conditions at the interfaces are matched and the Fourier series system is solved simultaneously for all regions. A detailed derivation and description of the Fourier series assembly procedure can be found in Demidenko (2011). The Fourier series is truncated when satisfactory convergence is achieved.

Continuous one-dimensional resistivity models on a half-space provide further variety in available analytic solutions which may match particular scenarios (Wait, 1982). An exponential function of depth or other arbitrary smooth functions are possible. Surface discontinuities such as wedges, spherical inclusions, or vertical discontinuities at the surface of the half-space allow modelling of dikes, surface anomalies, or fluid filled boreholes.

The two-dimensional analytic model may be extended by applying conformal deformations. Conformal deformations can move (and deform) the inhomogeneity away from the origin while leaving exterior boundary measurements unchanged. Any deformation of a domain with an isotropic conductivity can be modelled as the original shape with some anisotropic conductivity (Lionheart, 1998). A subset of these distortions can be modelled as isotropic conductivity changes with the original shape. This subset is exactly the distortions that are conformal maps. A conformal deformation preserves infinitesimal angles but not lengths (Hazewinkel, 2002). For two dimensions, the Cauchy-Riemann equations describe the conformal deformation

condition between original coordinates (x, y) and deformed coordinates (u, v)

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad ; \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{5.7}$$

Conformal deformations in three-dimensions are no longer interesting for constructing analytic models; they are restricted to rotation, reflection, scaling and translation. Restricting the model to the unit sphere at the origin, possibly containing many other spheres at the origin of differing conductivity, leaves no degrees of freedom to deform the model. The conformal inversion is available in two-dimensions and enables various useful deformations. Useful examples include movement of the inhomogeneity away from the origin, deformation of the boundary or transformation from the unit circle to the half-space (Figure 5.6). Measurements at the exterior boundary remain the same so the form of the solution does not change.

A subtlety of the two-dimensional conformal deformations lies in their implementation with a discrete FEM model (Boyle and Adler, 2011). Conformal deformations modelled by translating nodes on an FEM mesh with a piecewise linear discretization of the domain fail to locally preserve angles in the mesh. These changes in angle introduce voltage changes at interior mesh nodes where none would be otherwise expected (Boyle et al., 2012a). CEM electrodes will no longer have the same current density across their surface and interior elements will experience slightly different current flow. The effect at the electrodes may be corrected by applying a nonlinear contact impedance z_c along the length of the CEM electrode. A conformal deformation always transforms an isotropic conductivity into a new isotropic conductivity due to the Cauchy-Riemann equations (5.7). The conformal deformation assumes that the conductivity is constant meaning that the material maintains its conduc-
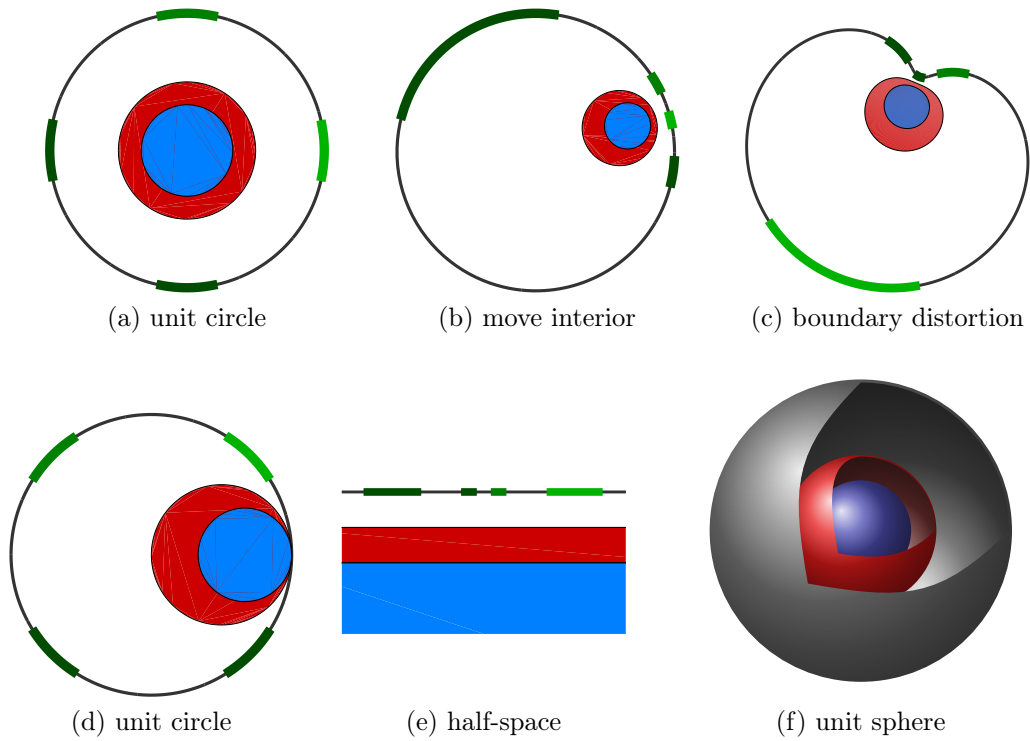
Figure 5.6: Conformal deformations; (x,y) coordinates are converted to the complex plane $z = x+y\mathrm{i}$, electrodes in green, (a) a unit circle $|z| = 1$ with inhomogeneities of diameter $z/2$ and $z/3$, (b) a conformal deformation $C = 7/10-\mathrm{i}/5$; $z \to (z+C)/(1+zC^*)$ applied to (a), (c) a conformal deformation $C = 7/10-\mathrm{i}/5$; $z \to -\mathrm{i}\exp(2z+2C)$ applied to (a), (d) a unit circle $|z| = 1$ with inhomogeneities $z/2+1/2$ and $z/3+2/3$, (e) a conformal deformation $z \to -\mathrm{i}(1+z)/(1-z)$ applied to (d) gives the half-space, (f) the unit sphere for which conformal deformations do not cause any interesting deformations, note that electrodes (green) are also deformed

tivity per unit volume under deformation. To capture the effect of a deformation on a compressible material requires an appropriately scaled conductivity. Practically speaking, the effect of compression and discretization errors are generally small and do not affect the solution except during extreme deformations (Figure 5.7). Conformal deformations that do not modify the boundary shape still result in modified electrode size and position. The deformation still may suffer from a failure to preserve angles in the FEM mesh deformation.

In summary, conformal deformations are useful for constructing analytic two-dimensional models for comparison with approximate methods such as the FEM. Conformal deformations introduce subtle errors if the geometric deformation is directly applied to a linear FEM forward model. The quantity of error is related to the magnitude of the deformation and how close the deformation is to any electrode.

(a) undeformed model

(b) conformal, fixed σ (c) conformal, modified σ
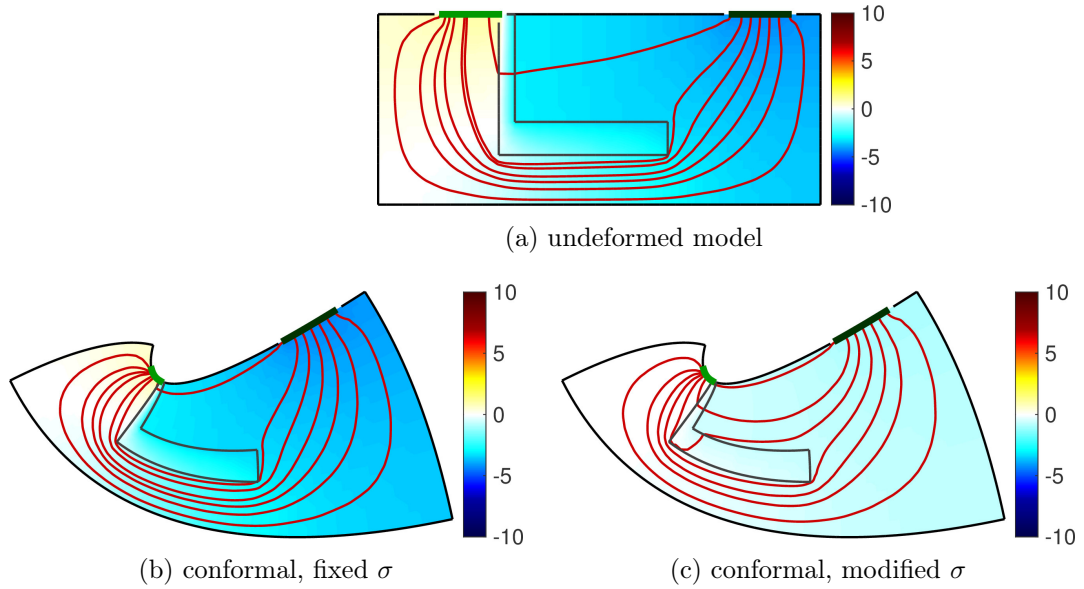
Figure 5.7: Conformal deformation examples; (x,y) coordinates are converted to the complex plane $z = x + y\mathrm{i}$, (a) half-space ($\sigma = 1$ S/m^2) with a grey outlined L-shaped inhomogeneity ($\sigma = 0.01$ S/m^2), current stream lines drawn in red, CEM electrodes ($z_c = 1 \times 10^{-5}$ $\Omega\cdot$m) on the top surface in green, voltage distribution across the domain as background colours with range shown in the colour bar, (b) conformal deformation applied to (a) with $z \to \frac{12\mathrm{i}}{240} \frac{\exp(-z-20-80\mathrm{i})}{100}(20\mathrm{i} - z)(-10\mathrm{i} - z)$ and unmodified conductivity modelling an incompressible fluid of fixed conductivity, (c) a modified conductivity no longer gives an identical voltage distribution after updating conductivity based on area change as a compressible conductive media, conformal deformations will always take an isotropic conductivity to another isotropic conductivity (5.7), for conformal deformations the current stream lines generally follow the same (deformed) paths to (deformed) electrodes

5.7 Contributions and Discussion

In this work, a framework for checking algorithm behaviours for inverse problems has been proposed. A distinction was drawn between a provably correct outcome and one that is judged to be valid based on an assumed distribution. We advocate a hybrid approach where automated tests for correctness are regularly executed while the more ambiguous tests of validity are moderated by a developer. Those tests that require human intervention should then be tied to an automated test so that a failure of the automated test indicates the need for an inspection of all related tests for validity.

Comparing the results of different implementations may be useful to show agreement, but black-box testing immediately hits a road-block for closed-source software when solutions differ: which implementation is to be trusted and how is a bug to be proven? Ultimately, the choice is one of either "believing" in an implementation or having a mechanism for testing that implementation in depth and at various levels of detail. Testing in this sort of detail is generally not possible without access to source code which leaves closed-source implementations in a challenging position. Access to commercial implementation's source code is typically under restrictive legal terms that endanger competing implementations or the code is altogether unavailable.

Comparative testing against open source software such as BERT or E4D may be practical (Günther et al., 2006; Johnson et al., 2010). A similar effort would be necessary to first show that the alternate implementation was sound. Such an effort has been applied to the E4D software; though, results appear to be internal to the originating organization. Ultimately, there is a choice to be made in determining how much effort to put into confirming an implementation's behaviour. Open source

software affords the opportunity to distribute this effort and cost. If the testing itself is open source, the results become reproducible and verifiable. Given a satisfactory level of confidence in these alternate implementations, a comparison of the final solutions may provide some form of confirmation in the event of similar outcomes.

The author implemented the Gauss-Newton/Conjugate Gradient solver as described in this and previous chapters. The implementation has been contributed to the Electrical Impedance and Diffuse Optics Reconstruction Software (EIDORS) project under the GNU GPL version 2 or version 3 open-source license. Tests for correctness and validity were implemented where time permitted. In the following chapter, we examine the inputs to the algorithm as a means of investigating the quality of the data.

Chapter 6

Data Quality and Model Mismatch

In the previous chapter, we have illustrated a process for evaluating whether an implementation is correct. In this chapter, we focus on the algorithm inputs as a possible cause of a bad reconstruction (Figure 5.2 #3, #4 and #5). In examining the data, we wish to avoid conflating the reconstruction algorithm with the quality of the data. To achieve some degree of independence between the two, we initially rely solely on the forward model's correct implementation to examine the data (Figure 6.1). Data that agrees with the model to a certain degree will result in a "reasonable" reconstruction while a significant data-model mismatch may cause a "suspicious" artifact. A suspicious artifact may either be ignored, by discarding or de-weighting measurements, or the cause of the problem can be resolved by adjusting the model to make the data useful in the reconstruction.

Data in impedance imaging are superficially considered to be the voltage measurements. A more complete definition includes voltage measurements, stimulus currents, the stimulus and measurement sequence, the geometry near the region of
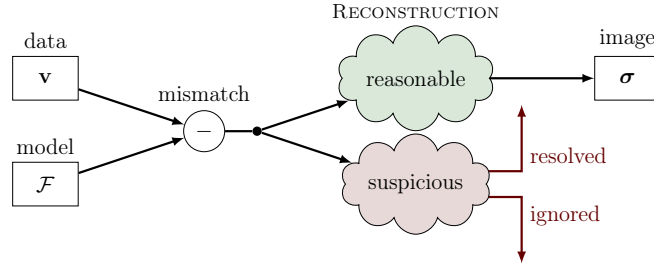
Figure 6.1: Explanations for data quality and model mismatch; data and model may have a mismatch which leads to a reasonable (green) or suspicious (red) reconstruction; a suspicious mismatch may be *resolved* and incorporated into the reconstruction or *ignored* by reducing its affect on the reconstruction

interest, and the location of each electrode. Error estimates for all values are important but at times, missing from the available data set. Typically, a measurement is the average of repeated measurements to minimize the effect of uncorrelated random noise. Measurement error estimates are typically approximated using the reciprocal measurement error[1] rather than measurement variance.

We contrast two sets of data to look for a broad range of issues in data acquisition. Simulated data with no noise provides a baseline expectation. The simulated data may be compared to the input data (Figure 6.2). If the input data are also simulated, care should be taken to avoid the inverse crime; the two data sets should not be constructed based on the same discretization (Kaipio and Somersalo, 2007).

Overall, we are motivated by the desire to contrast observed and expected information throughout the reconstruction process. As a first step, we advise directly comparing the measurement data to the output of an appropriate forward model. A direct comparison removes the opportunity for any intervening errors in reconstruction to disrupt the analysis. Common wisdom is to plot the simulated and

---

[1]Reciprocal measurement error is the change in measurement after reversing stimulus and measurement electrodes and for a linear model will be zero (Parasnis, 1988).

real data side-by-side or against each other and decide if they are "similar enough." This advice is appropriate for expert users but is hard to execute in practice for an in-experienced user. Having plotted the data, what is one expecting to see? How closely should the data sets be aligned, and to what degree can offsets be explained by common issues: electrode contact quality, broad differences in assumed impedance, and model boundary mismatches?

6.1 Representations of Quality and Mismatch

At a general level, the data can be examined for any extreme outliers or obvious inconsistencies. Two similar approaches are applied in the geophysics and biomedical fields which amount to plotting the measurements in two or three dimensions. Biomedical researchers commonly examine the data directly as two-dimensional plots of measurement magnitude y against measurement number x. With experience, it is possible to identify the rough location of a simple conductivity inhomogeneity or discontinuity near the boundary through approximate pattern matching between plots. Unexpected patterns in the raw data may indicate measurement protocol mismatches due to incorrect electrode numbering or repeating anomalous data that might indicate a bad electrode (Figure 6.2). Varying model configurations can help identify trends that might explain discrepancies between modelled and real data. Sorting the measurements x based on electrode number, spacing, voltage difference (real versus simulated) or other criteria can help to crystallize inconsistencies in the data set into a pattern of correlations.

In Figure 6.2, a range of models are shown which compare measurement plots
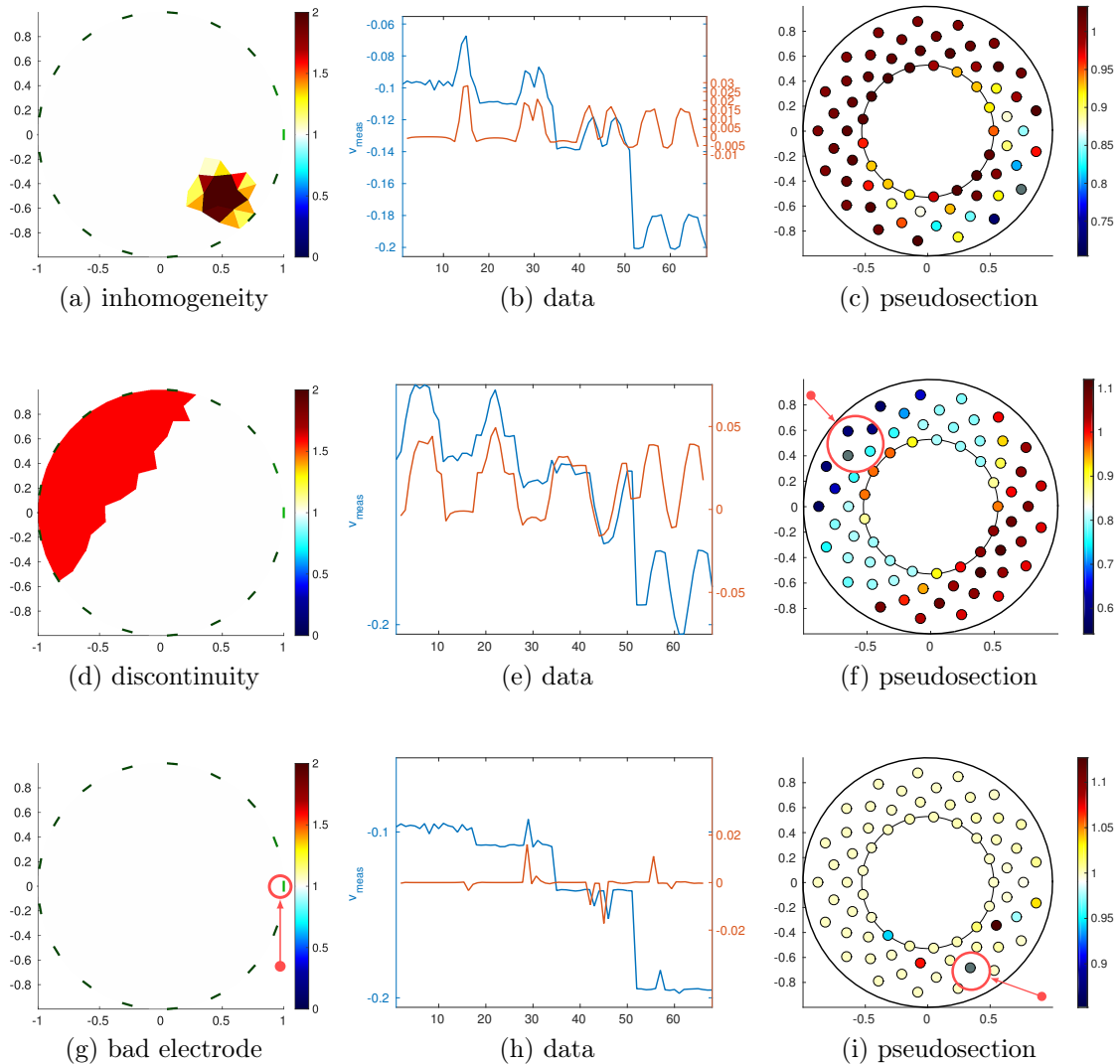
Figure 6.2: Data comparison for circular models; a circular domain with 17 electrodes, dipole-dipole stimulus and measurements separated by 1 to 4 electrodes $n = 1, a = 1..4$ (§A.3), for (a,d) inhomogeneous models, and (g) an electrode with high contact impedance circled (red), (b,e,h) measurements (blue) and difference from the homogeneous measurements (orange), (c,f,i) pseudosections for apparent resistivity of the same data with anomalies circled (red), (c) anomalies that do not touch the boundary are hard to localize in the pseudosection, pseudosection or raw measurements are equally useful for identifying an anomaly but then require careful examination to find their cause

(biomedical) and pseudosection (geophysics) representations of the measurement data for a circular domain. The stimulus and measurement sequence was a dipole-dipole pattern with electrode spacings of 1 to 4 electrodes $n = 1, a = 1..4$ (§A.3). Measurement plots have data plotted in the stimulus and measurement sequence order. For each pseudosection, a circular representation was used with coordinates (θ, r). The angle θ was the mean of all electrode locations for that measurement. The depth $r = R(1 - a/\pi)$ was calculated for measurement electrode separation a and surface radius R. Colour represents the measurement magnitude. Models of an inhomogeneity, discontinuity and bad electrode were selected for their unique signatures.

In geophysics, it is common to examine the pseudosection of the data. A pseudosection is a plot of the measurement data: apparent resistivity[2] or voltage measurements are colour coded and sorted onto a two-dimensional grid. The grid locations are motivated by the concept of a Depth of Investigation calculated from the positions of electrodes used in each measurement (Figure 6.3). Ultimately, the locations are somewhat arbitrary (a pseudosection is not a conductivity reconstruction) but the arrangement does provide a consistent representation of the input data independent of measurement order. Pseudosections can help to quickly identify problematic electrodes or strong contrasts in the ground conductivity (Edwards, 1977). The pseudosection can be helpful in identifying conductivity discontinuities or anomalies near electrodes.

A pseudosection is a flattened two-dimensional plot of a multidimensional space describing the distance between the electrodes of an impedance imaging system. Ex-

---

[2]Apparent resistivity is a measurement normalized by the expected measurement for a homogeneous conductivity model (§4.7).

(a) inhomogeneity

(b) data

(c) pseudosection

(d) discontinuity

(e) data

(f) pseudosection
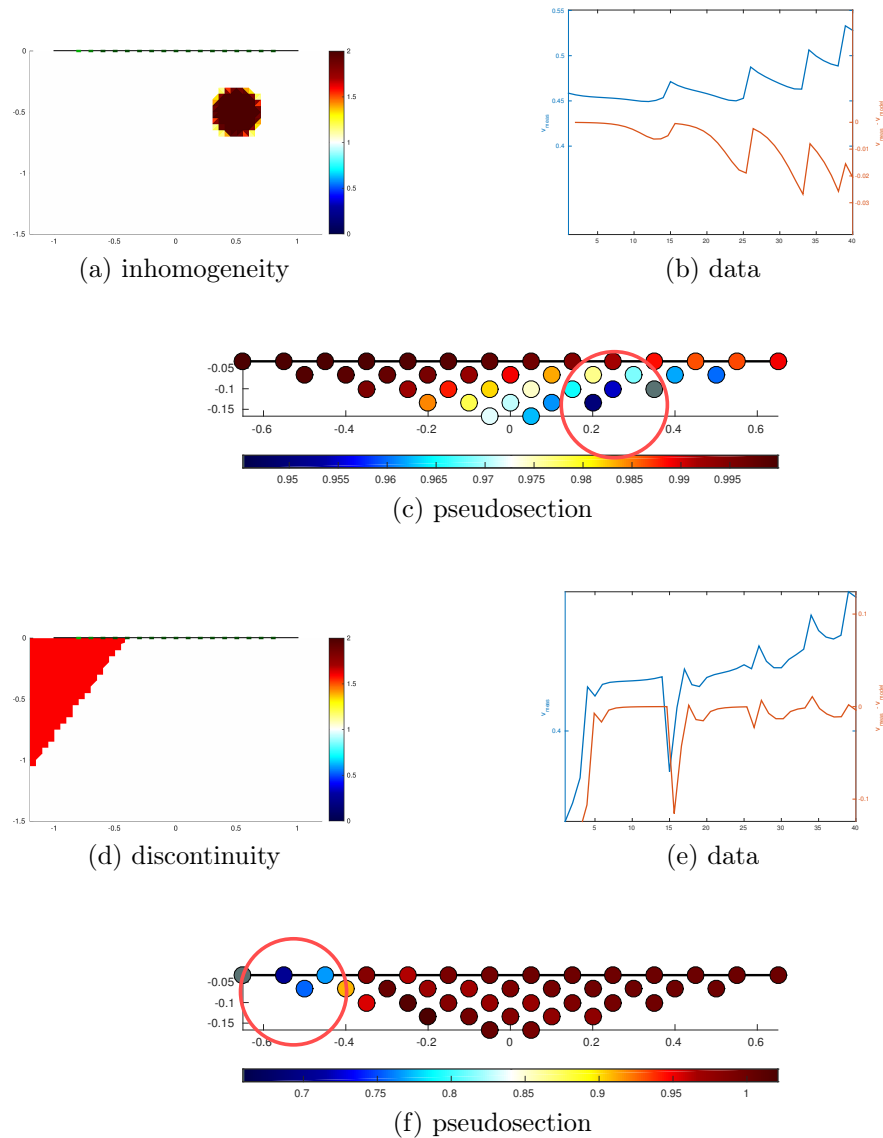
Figure 6.3: Data comparison for half-space models; a half-space domain with 17 electrodes, Wenner measurements separated by 1 to 5 electrodes $a = 1..5$ (§A.1), for (a,d) inhomogeneous models, (b,e) measurements (blue) and difference from the homogeneous measurements (orange), (c,f) pseudosections for apparent resistivity of the same data with anomalies circled (red), pseudosection or raw measurements are equally useful for identifying an anomaly but then require careful examination to find their cause

amples for a system using stimulus and measurement electrode pairs (dipoles) include the measurement dipole length, stimulus to measurement dipole spacing, and the average of active electrode positions. The plot is constructed by assigning, from the various measures of electrode spacing, a two-dimensional position (x, z). There is a considerable amount written about the preferred coordinates at which to plot points on the pseudosection. A specific depth may be calculated, if one assumes a homogeneous model, as the depth z at the median of the sensitivity curve on a homogeneous half-space (Edwards, 1977; Barker, 1989; Oldenburg and Li, 1999). The horizontal position x is the mean of the stimulus and measurement electrode locations. In this work, the depth z is set dependant on the separation of the measurement electrodes a so that $z = -a/3$, though more complicated choices are available. Arranging the data in this fashion allows quick and relatively standardized correlation between measurement electrode separation and electrode spacing while removing dependencies on data set ordering. This is beneficial for comparison purposes but does not give a good tool for direct interpretation. Reciprocal measurements are merged by taking their mean value. For circular domains, the pseudosection may be plotted in radial coordinates.

In Figure 6.3, two models are shown comparing measurement plots (biomedical) and pseudosection (geophysics) representations of the measurement data for a half-space model. The stimulus and measurement sequence was a Wenner pattern with electrode spacings of 1 to 5 electrodes $a = 1..5$ (§A.1). Colour represents the measurement magnitude. Models of an inhomogeneity and a discontinuity were selected for their unique signatures.

Identifying noisy measurements or those with high contact impedance can be

automated (Meeson et al., 1996; LaBrecque et al., 1996a; Vilhunen et al., 2002; Deceuster et al., 2013) but each cause of faulty data typically requires a new test. Symptoms are not all unique so that identifying the root cause from a selection of indicators can be complicated: we know something is suspiciously bad with our data but not the cause. An indication may validate that a reconstruction is suspect but does not help recover a good reconstruction. To recover a good reconstruction from measurements with troublesome data, the "bad" data may be removed through the de-weighting matrix $\mathbf{W}$ (§4.6). Better yet, if the cause can be identified, the forward model might be modified to account for the source of the error. If the source of the error is dynamic, the model can either be updated once based on some estimation framework or added as a parameter of the inverse solver.

6.2 Data-Model Mismatch and Reconstruction

Observing a sequential transformation may lead to insights about the cause of anomalies in the final product. Here, we apply this observational technique to impedance imaging by instrumenting the iterative Gauss-Newton solver to provide greater detail in reporting progress.

A data set, the impedance imaging measurements, undergoes a transformation into an estimate of some set of model parameters. A linear algorithm, for example a difference imaging single-step Gauss-Newton algorithm, has little internal state so that observing all inputs and outputs gives a relatively complete picture of behaviour. Nonlinear iterative algorithms may transition through many solutions: observing these may lead to insights as to why a reconstruction has failed. Monitoring the

evolution of two similar sets of data allows one to determine where that evolution diverges. The point of divergence can be useful in identifying unique features in the data or (in)sensitivities in the algorithm.

The iterative Gauss-Newton update $\delta\boldsymbol{\sigma}_n$

$$\delta\boldsymbol{\sigma}_n = (\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J} + \lambda^2\mathbf{R}^\mathsf{T}\mathbf{S}\mathbf{R})^{-1}(\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{r}_n + \lambda^2\mathbf{R}^\mathsf{T}\mathbf{S}\mathbf{R}(\boldsymbol{\sigma}_* - \boldsymbol{\sigma}_n)) \tag{6.1}$$

$$\boldsymbol{\sigma}_{n+1} = \boldsymbol{\sigma}_n + \alpha_n\,\delta\boldsymbol{\sigma}_n \tag{6.2}$$

depends on the previous iteration's estimate of model parameters $\boldsymbol{\sigma}_n$, measurement misfit $\mathbf{r}_n = \mathcal{F}(\boldsymbol{\sigma}_n) - \mathbf{v}$ based on the forward model $\mathcal{F}(\cdot)$, line search α_n and search direction $\delta\boldsymbol{\sigma}_n$. The search direction, in turn, depends on the first-order estimate of slope (Jacobian) $\mathbf{J}_n$ at the previous estimate of model parameters, measurement weighting $\mathbf{W}$, the regularization scheme $\mathbf{R}^\mathsf{T}\mathbf{S}\mathbf{R}$ and strength λ.

For a single-step reconstruction with no line search ($\alpha = 1$), the initial guess is generally homogeneous. For single-step difference imaging ($\Delta\mathbf{v} = \mathbf{v}_B - \mathbf{v}_A$), the initial guess is typically assumed to be the prior ($\hat{\boldsymbol{\sigma}}_0 = \boldsymbol{\sigma}_*$) with a uniform measurement weighting ($\mathbf{W} = 1$) giving a Gauss-Newton step

$$\Delta\boldsymbol{\sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J} + \lambda\mathbf{R}^\mathsf{T}\mathbf{S}\mathbf{R})^{-1}\mathbf{J}^\mathsf{T}\,\Delta\mathbf{v} \tag{6.3}$$

with the difference solution $\Delta\boldsymbol{\sigma}$ simplified considerably from (6.1) and (6.2) to a linear combination of Jacobian $\mathbf{J}$ and regularization λ, $\mathbf{R}$. An iterative algorithm clearly has considerably more internal behaviour to monitor.

Forming a mental picture of a multidimensional data set and the way an algorithm is traversing that space can be difficult, particularly for math intensive techniques

Listing 6.1: Sample log: EIDORS Gauss-Newton solver

```
1    iteration 2
2      calc Jacobian J(x) = (jacobian_adjoint, 3168 DoF, 349 meas, dx_dlogx)
3      calc hp^2 R^t R
4        {1,1} regularization RtR (prior_laplace), ne=3168x3168, hp=3.162
5      calc step size dx
6        ||dx||=41.6
7        ||dx_1||=41.6 (log_conductivity)
8      update step dx, beta=0, ||dx||=41.6
9        acceleration     d||dx||=-40
10       direction change ||ddx||=1.28 (79.3°)
11     line search, alpha = line_search_onm2
12       line search (finite precision) limits
13         log_conductivity: alpha range = 3.03e-16 -- 581
14         alpha (unchanged) = 0 0.062 0.126 0.252 0.504 1
15       i     =      [1]       [2]       [3]       [4]       [5]       [6]
16       alpha =        0     0.062     0.126     0.252     0.504         1
17               2.17e+04  8.97e+03  7.35e+03  4.73e+03  1.87e+03  8.02e+03
18       step size = 0.448, misfit = 2.21e+03, expected = 2.26e+03
19       did not like our step selection - choose one of perturb values
20       step size = 0.504, misfit = 1.87e+03 selected
21       update perturbations around step = 0.504 (limit alpha=1.0)
22       selected alpha=0.504
23     show_fem_pontpean()
24     calc residual
25       r = 2.21e+03 (3.05%)
26       dr=-1.95e+04 (-26.9%)
```

(Shewchuk, 1994). Constructing visualizations through dimension reduction techniques such as Principle Component Analysis (PCA) may be helpful but their useful application to inverse problems debugging is not well explored.

A textual log of the algorithm is an effective aid for debugging: it describes the procedure, critical choices and progress (Dillon, 2003; Zeller, 2009). An example[3] from a single nonlinear Gauss-Newton iteration is shown in Listing 6.1. The listing describes the sequence of execution. Indentation is loosely used to track the depth of calls to functions and detailed information about the state of the algorithm. The key ingredient for the log is a readable and concise description of the current activity.

The log starts with the iteration number, having skipped the start-up and first iteration for this example. Line 2 describes the configuration for the Jacobian es-

---

[3]The example log uses the Pont-Péan data set as described in the following section §6.3.

timate using log conductivity and the adjoint method for 349 measurements and 3168 parameters in the inverse model. Lines 3–4 describe the regularization matrix: Laplace regularization with $\lambda = 3.162$. Lines 5–7 record the calculation of the descent direction $\delta\sigma$ and calculate the 2-norm of the vector $||\delta\sigma||$ as a measure of magnitude. Line 8 mentions that $\beta = 0$: the iteration is performing a Gauss-Newton update, not a Conjugate Gradient iteration. Lines 9 and 10 attempt to describe the acceleration and direction change of the multidimensional search. The acceleration is defined as

$$\texttt{d||dx||} = ||\delta\sigma_n|| - ||\delta\sigma_{n-1}|| \tag{6.4}$$

the change in the 2-norm magnitude of the search vector from the previous iteration (Figure 6.4a). The direction change is defined

$$\texttt{||ddx||} = \left|\left|\frac{\delta\sigma_n}{||\delta\sigma_n||} - \frac{\delta\sigma_{n-1}}{||\delta\sigma_{n-1}||}\right|\right| \tag{6.5}$$

as the change in the direction of the search vector normalized to the unit hypersphere. Equivalently, (6.5) can be expressed as a dot product of two normalized vectors

$$\texttt{||ddx||} = 1 - \frac{\delta\sigma_n}{||\delta\sigma_n||} \cdot \frac{\delta\sigma_{n-1}}{||\delta\sigma_{n-1}||} \tag{6.6}$$

giving the same measure of chord length on the unit hypersphere. A value of zero indicates the same direction as the previous iteration, while a value of two indicates a complete reversal and likely oscillation in the search path (Figure 6.4b).

Lines 11–21 describe the line search. 5 points are tested along the search direction,

(a) magnitude   (b) direction
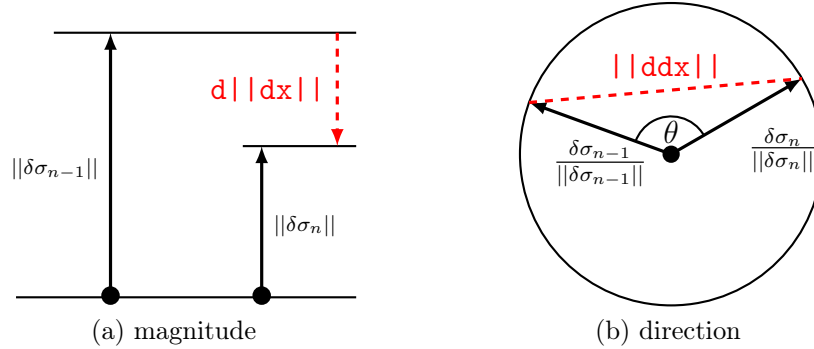
Figure 6.4: Search direction and magnitude change; (a) current $\delta\sigma_n$ and previous $\delta\sigma_{n-1}$ search direction magnitudes are compared `d||dx||`, (b) a hypersphere flattened to a circle representation, search directions are normalized to the unit circle and the chord length is used to describe the quantity of direction change `||ddx||` $= 2\sin(\theta/2)$

a line is fitted and an update length ($\alpha = 0.448$) is selected. The selected length is checked for an improvement to the residual. One of the test points is found to have a lower residual and is selected instead ($\alpha = 0.504$).

The final lines summarize the iteration status. Lines 23–24 indicate a new figure showing the current conductivity σ_n. Lines 24–26 show the absolute value of the residual and its improvement from the previous iteration. These values are of interest when determining why iterations terminated early or if the iteration count should be stopped earlier on future runs. From these data, secondary time series plots of iteration progress may be plotted.

6.3   Pont-Péan Background

We illustrate the behaviour of our visualization techniques by observing their behaviour on real data. Data collected at Pont-Péan, France were used to debug an

implementation of the iterative absolute solver (Algorithm 7). The resulting reconstruction exhibits a similar apparent resistivity pseudosection and reconstructed conductivity profile to published reconstructions (Pessel and Gibert, 2003). The reconstructed images match the available ground truth borehole data and known geology based on the historical findings at the mine site (Figure 6.6d). Data were then intentionally rearranged to model mis-wiring of electrodes. The error was modelled by renumbering electrodes in the forward model so that they no longer match the physical electrode arrangement. Erroneous and correct images were then contrasted using the methods proposed in this work to illustrate debugging the (known) issue.

Data were collected by researchers at Université Rennes 1, Rennes, France from 2000 to 2011 (Pessel and Gibert, 2003). Measurements were collected using an ABEM Terrameter SAS 4000 resistivity meter (ABEM Instrument AB, Sweden), ABEM ES464 electrode selector and 64 stainless steel (20cm long, 1cm diameter) electrodes placed alongside the road *La Grande Betuaudais* which is located at (48°00'29.5"N, 1°42'36.3"W) in a linear array with 5m spacing between electrodes. Stimulus currents were 100–200 mA alternating square wave $(0, +, 0, -, 0)$.

The Pont-Péan mines (Pont-Péan, Brétagne, France) were operational through nearly two centuries, producing silver, lead, zinc and sulphur. The site was initially identified in 1628 by Jean du Châtelet, Baron de Beausoleil et d'Auffembach and his wife, professional prospectors who travelled throughout Europe (Chauris, 1989). They were accused of being charlatans but overcame these charges and left France for a time. In 1632 they returned and the wife, Martine de Bertereau wrote a number of books describing their methods, tools and discoveries as well as declaiming their earlier treatment. Among them "La Restitution de Pluton" (1640) was addressed to

Cardinal de Richelieu. The couple were arrested on unknown charges and died in the Bastille (Jean in 1645) and Château de Vincennes (Martine in 1642), respectively (Kölbl-Ebert, 2009).

Works were eventually established in 1730 after a major investment by Compaigne des Mines de Brétagne and eventually abandoned in 1794 during the French Revolution after 65 years in operation (Fayn, 1863). Mining resumed in 1844 at the surface, and at depth in 1852. The mine became the primary producer of silver in France, providing 80% of domestic supply. Pumps to remove water from the mine were installed in 1880 but rising energy costs diminished the mine's financial viability. Finally, in April 1904, 60 years after reopening, a strong inflow of water (5 million L/min) overcame the 595m deep mine putting nearly a thousand miners out of work. Throughout its existence the mine is estimated to have produced 220 kilotons of silver bearing lead and 20 kilotons of zinc. The mine office buildings remain as a national historic monument (Besson, 1982; Lodin, 1908, 1911).

The Pont-Péan mines are situated on a mineralized dioritic dike running roughly North-South for 4 km at 15 m in width. The dike lies at an 80° eastward inclination. The dike splits a fault plane separating Precambrian schist (West) from Tertiary sediments (East) (Lodin, 1895). The ERT survey line runs perpendicular to this dike. The Tertiary sediments are roughly 80 m deep in the area of the ERT survey data.

In 1958, two boreholes were drilled 400 m southward of the resistivity survey line at 40m and 100m westward of the fault. These boreholes indicate weathered schist to a depth of 35 m and 25 m respectively. Both the dioritic dike and schist are very fractured and may be a hydrological sink draining water (Pessel and Gibert, 2003).

6.4 Conductivity Reconstruction

Following the procedure outlined so far, we first examine the measurement data and associated error estimates. The data is then reconstructed and the behaviour of the algorithm is observed. Data is intentionally corrupted to model electrode mis-wiring and the analysis is compared with good reconstructions to illustrate the utility of the outlined debugging process.

The specific Pont-Péan ERT survey data used in the following work was collected in February 2004. The data were collected at least a year after Pessel and Gibert (2003) (published January 2003), at the same site and with the same electrode positions, but with a different stimulus and measurement protocol. It is not clear if the data was collected in the same seasonal conditions. The geological data and reconstructions presented in Pessel and Gibert (2003) (Figure 6.5) are relevant comparison points but cannot be expected to exactly match reconstructions presented here.

The 519 voltage measurements were normalized against stimulus current. As in Pessel and Gibert (2003), the 170 adjacent measurements where measurement electrodes M-N were separated by the minimum electrode spacing of 5 meters ($a = 5$ m, $n = 1$) were found to be inconsistent with other measurements and were dropped from the data set. The 349 remaining measurements were plotted as a pseudosection (Figure 6.6a). Comparison of the pseudosection between Figure 6.5d and Figure 6.6a would suggest that the more recent measurements were obtained with an electrode array placed approximately 20 meters to the East of that used in Pessel and Gibert (2003), based on the location of the strong contrast at $x = $ -50 m and $x = $ -30 m, z = 0 m of the apparent resistivity pseudosections.
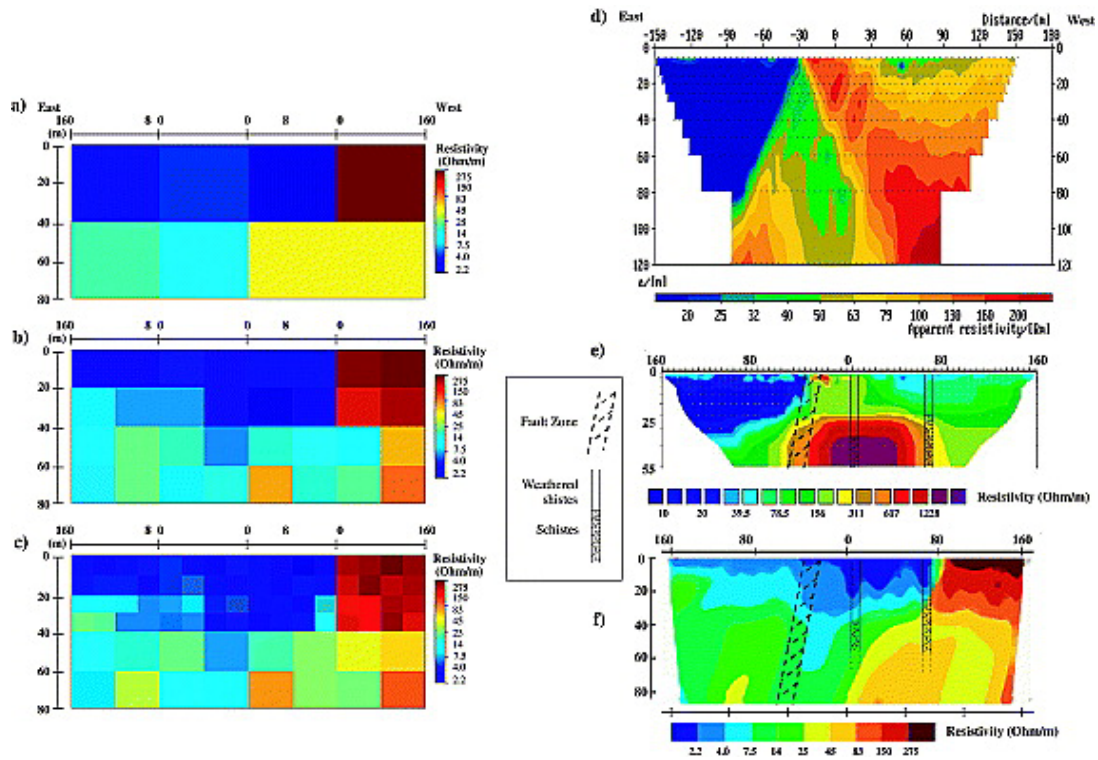
Figure 6.5: *[From Pessel and Gibert (2003), Figure 9:]*
Resistivity models obtained at the end of each stage of the inversion of the pole-dipole data for the (a) 4×2 coarse final model, (b) 8×4 resampled final model, and (c) final 80-block model obtained from a partial resampling of the top part of the 8×4 model. This allowed a refinement of the conductivity distribution in the shallowest part of the model and was accompanied by an additional slight decrease of the misfit function (see Figure 10). The geological information available is shown: the fault location coincides with a low-resistivity dipping zone and the thickness of the low-resistivity zone located immediately westward of the fault agrees with the weathered schists found in two boreholes (B1 and B2). (d) Data obtained for the pole-dipole array and arranged in a pseudo-section of the logarithm of the apparent resistivity. The remote electrode is 1000 m westward, i.e., on the right of the profile. The dike is approximately located at x = -30 m. Since the remote electrode is on the resistive right part of the profile, the apparent resistivities are biased toward the high values. (e) Resistivity model obtained by inverting the data with the RES2DINV software by Loke and Barker (1996b) (f) Smoothed version of the resistivity section shown in Figure 9c.
[Reproduced from Pessel and Gibert (2003), Figure 9, for comparison.]

A normalized exponential scaling of reciprocal measurement error was used to construct a diagonal measurement weighting matrix $\mathbf{W}$ (§4.6)

$$\mathbf{W}_{i,i} = \exp\left(-\left(\frac{\mathbf{e}_i}{0.01}\right)^2\right) \qquad \text{for } 0 < \mathbf{W}_{i,i} \leq 1 \qquad (6.7)$$

where the i-th measurement has reciprocal error $\mathbf{e}_i$ as a fraction of the measurement magnitude. In this data set, most measurements have less than 1% reciprocal error. Apparent resistivities with error much greater than the 1% cutoff ($\mathbf{e}_i/0.01$) were effectively ignored by applying this $\mathbf{W}$ weighting matrix.

A weighted best fit homogeneous resistivity estimate $\hat{\rho}_x$ for the measurements $\mathbf{v}_x$ was calculated

$$\hat{\rho}_x = \arg\min_{\rho_x} \left\|\frac{\mathbf{v}_x}{\rho_x} - \frac{\mathbf{v}_1}{\rho_1}\right\|_{\mathbf{W}} \qquad (6.8)$$

$$\hat{\rho}_x = (\mathbf{v}_1^\mathsf{T}\mathbf{W}\mathbf{v}_1)^{-1}\mathbf{v}_1^\mathsf{T}\mathbf{W}\mathbf{v}_x \qquad (6.9)$$

using the weighting matrix $\mathbf{W}$ (6.7) and the estimated homogeneous measurements $\mathbf{v}_1$ from a forward model ($\rho_1 = 1/\sigma = 1\ \Omega \cdot \text{m}$). Reciprocal error estimates for the 170 dropped measurements were generally low, so that all but 29 would not have been de-weighted by $\mathbf{W}$ in the reconstruction. The ($a = 5$ m, $n = 1$) measurements skew the weighted homogeneous resistivity estimate: $\hat{\rho}_x = 55.6\ \Omega \cdot \text{m}$ when they were included and $\hat{\rho}_x = 77.5\ \Omega \cdot \text{m}$ when they were dropped. We have observed the same data misfit (nearly an order of magnitude larger apparent resistivity) and have also dropped the 170 inconsistent measurements.

A two and a half dimensional ($2^1/_2$-D) forward model (Appendix D) of the linear 64 electrode array was constructed with 0.01 meter diameter CEM electrodes at 5

meter intervals. The forward model was used to calculate apparent resistivities (§4.7) based on a homogeneous conductivity ($\rho_1 = 1\ \Omega \cdot \text{m}$). The apparent resistivities were plotted against the homogeneous 77.5 $\Omega \cdot$ m forward model estimates and also presented as an apparent resistivity pseudosection (Figure 6.6b). The sub-vertical fault due to the 15 meter dike is apparent in the contrast between measurements near $x = -40, z = 0$ meters in the apparent resistivity pseudosection, as it was for Pessel and Gibert (2003). The measurement weighting matrix $\mathbf{W}$ was also presented as a pseudosection (Figure 6.6c). Of 349 measurements, 38 had reciprocal measurement errors greater than 1% and were strongly de-weighted in the reconstruction by applying the measurement weighting matrix. Finally, the resistivity section was reconstructed (Figure 6.6d).

A two-dimensional (x-z) conductivity parametrization was constructed of $25,046$ triangular mesh elements. A $2^1/_2$-D forward solver was implemented and a range of domain truncations were tested under homogeneous conditions, to confirm that the $2^1/_2$-D forward solution converged to an equivalently sized three-dimensional model and the analytic three-dimensional half-space model with no subsurface boundaries (Appendix A, (A.4)). These checks confirmed that systematic forward modelling errors due to FEM boundary conditions imposed on the finite domain and the new $2^1/_2$-D forward solver implementation were not skewing the reconstructed conductivity image. The Gauss-Newton absolute inverse solver (Algorithm 7, $\beta = 0$) was used to reconstruct a conductivity profile. A Laplace smoothing prior and L2-norm on measurements and regularization were applied. The measurements were converted to apparent resistivity. The reconstruction was performed in log conductivity parameter space (min $\sigma = 0$, max $\sigma = \infty$) with a limit of ten iterations or a residual

(a) raw measurements pseudosection

(b) apparent resistivity pseudosection

(c) measurement weighting $\mathbf{W}$ presented as a pseudosection

(d) reconstructed conductivity, $\lambda = 3.2$
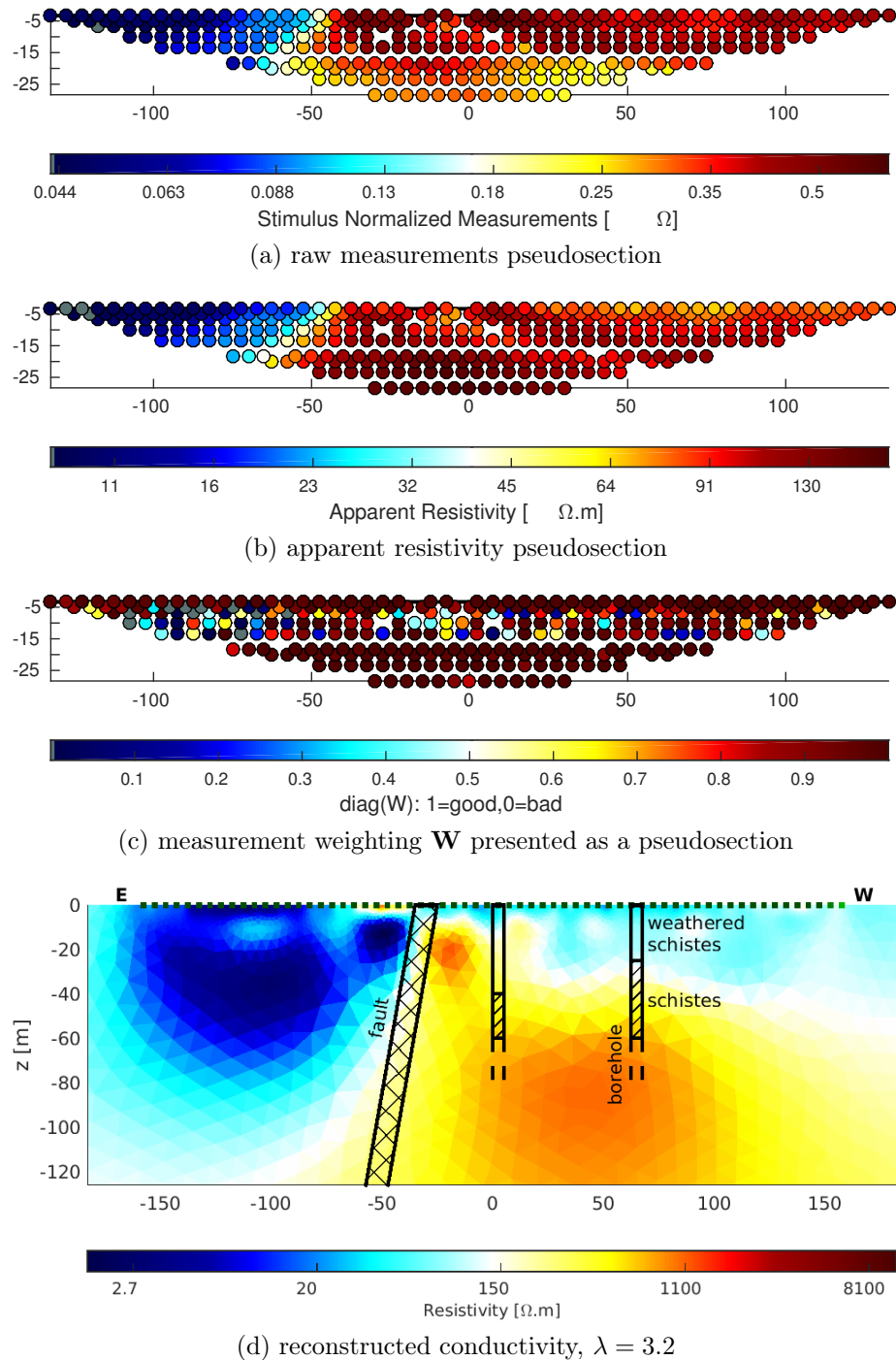
Figure 6.6: Pont-Péan data and reconstructed conductivity; after dropping adjacent measurements from the data set: (a) raw data normalized by the stimulus amplitude, (b) apparent resistivity based on a 3D forward model, (c) the diagonal of the weighting matrix $\mathbf{W}$ constructed from reciprocal error estimates, (d) the reconstructed resistivity ($\lambda = 3.2, \sigma_0 = 77.5 \,\Omega \cdot \mathrm{m}$), electrode positions in green on the top surface, overlaid with known borehole and fault information

reduction of at least 0.1%. An L-curve was calculated for the hyperparameter and the value $\lambda = 3.2$ was selected based on the curve and by looking for the location of variations in the reconstructed resistivity image at each hyperparameter value and selecting a hyperparameter that smoothed the smallest of these variations. After 5 Gauss-Newton iterations, the weighted norm of the apparent resistivity misfit was reduced from $||(\hat{\mathbf{v}}_0 - \mathbf{v})/\hat{\mathbf{v}}_1||_{\mathbf{W}} = 653.1$ to $||(\hat{\mathbf{v}}_5 - \mathbf{v})/\hat{\mathbf{v}}_1||_{\mathbf{W}} = 62.3$, an improvement of $10.5\times$ to achieve 8.6% of the original weighted apparent resistivity misfit. The weighted norm of reciprocal error as apparent resistivity was $||\epsilon/\hat{\mathbf{v}}_1||_{\mathbf{W}} = 3.5$. The reciprocal error was an order of magnitude smaller than the reconstruction misfit indicating that the data was not over-fit. The reconstructed conductivity was consistent with the available geological data and previously published RES2DINV reconstruction.

The reconstruction does not agree particularly well with the block-wise relaxation model (Figure 6.5a–c) or its post-processed, smoothed version (Figure 6.5f). This mismatch may be due to the selected discretization in Pessel and Gibert (2003) which may strongly affect the end solution by directing the initial search in a particular direction. Some tendency towards higher resistivity regions at the western edge of the electrode array (as in Figure 6.5a–c) were observed with Tikhonov ($\mathbf{R}^{\mathsf{T}}\mathbf{R} = \mathbf{I}$) and Noser ($\mathbf{R}^{\mathsf{T}}\mathbf{R} = \mathrm{diag}(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{0.5}$) (Cheney et al., 1990) regularization priors but these generally disappeared in subsequent iterations as the reconstruction was refined.

For the Gauss-Newton iterations, plots may be used to summarize key points of the algorithm state. These points include the Jacobian sensitivity, search direction, line search and residual (Figure 6.7). The Jacobian sensitivity plot serves to indicate where reconstructed conductivity regions might be considered significant. Regions

beyond $z = -50$ meters depth have rapidly declining significance. The search direction at each new iteration serves to indicate what regions did not fit the previous iteration's data and may also serve to indicate regions where convergence was slow due to oscillations in the search direction. For this data set, the conductivity discontinuity around the fault was a source of ongoing refinement. Finally, the line search and residual plots serve to indicate how well the algorithm was advancing. The line search plot indicates that a search direction was selected which enabled the residual to be reduced. At times, the line search may indicate multiple minima or a search direction which has a magnitude that is too large. Also, note that the line search fitted curve is not necessarily consistent with the final test point selected based on the predicted minima from the curve. This is not particularly surprising, considering the nature of non-linear inverse problems. A line search that checks the selected search magnitude α and selects a better test point, when a lesser residual has already been calculated at that other test point, tends to be more successful on difficult data sets. Ideally, as in these plots, the residual experiences exponential convergence. The residual plot also helps to identify when changes are being driven by the data misfit or by regularization terms. At the third and fourth iterations, the iterations were largely driven by the regularization terms which were smoothing conductivity variation (Laplace regularization).

Using the same observational methodology, we can now check how the algorithm handles erroneous data. There are an almost limitless range of possible errors. We chose to model a relatively common usage error: electrode mis-wiring. An error in electrode wiring was modelled by renumbering the electrodes in the forward model. Electrodes were nominally numbered from West to East, starting at electrode #1

(a) sensitivity

(b) search direction

(c) line search

(d) iteration progress
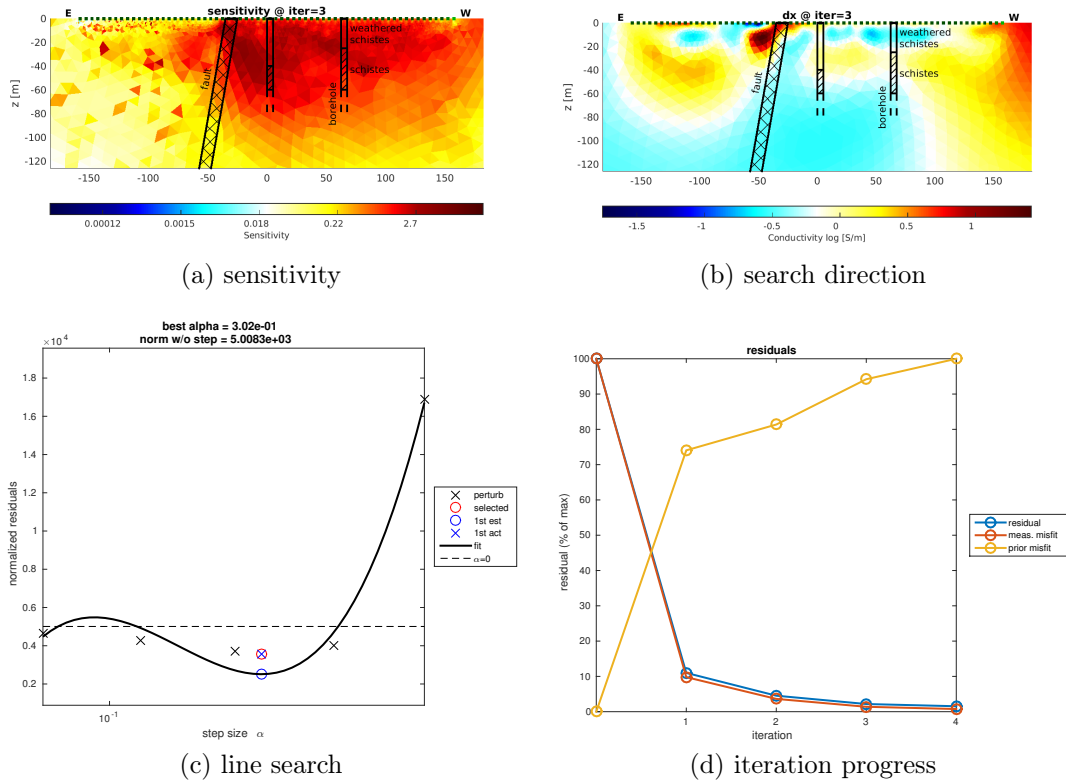
Figure 6.7: Pont-Péan, Gauss-Newton reconstruction at iteration 3; (a) sensitivity at iteration 3 has changed due to the conductivity contrast at the fault plane, (b) search direction at iteration 3 indicates continued refinement on either side of the fault plane and in regions with low sensitivity, (c) a curve fitted to the line search test points (×) reflects a relatively linear step with a smooth curve and minimum ($\alpha = 1.1$) near one (circles), (d) the reconstruction progresses over the course of 4 iterations with exponential convergence of the residual (blue) as the measurement misfit was reduced (red) at the cost of the prior misfit (yellow).

through #32. Electrodes #4 and #23 were swapped with #28 and #16 respectively. For ERT systems such as the ABEM hardware, these types of electrode mis-wiring errors are minimized because correct electrode numbering is enforced through the use of a single common cable on a linear array of electrodes. The opportunity for user error is greater for configurations with two-dimensional grids of electrodes or where electrodes are connected by independent wires to the base. An electrode mis-wiring error is much more common in experimental environments where the electrode placement strategy and wiring are (intentionally) left unrestricted through equipment design choices. In contrast, commercial equipment suppliers are interested in minimizing in-the-field user errors and design their equipment accordingly.

Plots of apparent resistivity pseudosection (Figure 6.8a) show some strong outliers around $x = 80$ m. In all likelihood, these would be candidates for de-weighting given further evidence to support a noise hypothesis. A hyperparameter was selected using the same procedure as before $\lambda = 32.0$. The search direction at iteration 3 (Figure 6.8d) and final conductivity reconstruction after 5 iterations (Figure 6.8e) exhibit some clear inconsistencies with the geological ground truth data. The line search (Figure 6.8b) at iteration 3 shows a step where little improvement occurred. The iteration progress over 5 iterations (Figure 6.8c) and converges to a significantly larger residual than the correct solution. The prior's contribution to the residual rose steadily showing that the Laplace regularization is limiting later iterations. Together these pieces of evidence tend to suggest that the reconstruction was suspect. The patterns of artifact in the conductivity reconstruction around the electrodes would be consistent with artifacts due to large electrode movements, electrode mis-wiring or widespread poor electrode contact.

(a) apparent resistivity pseudosection

(b) line search, iter=3

(c) iteration progress

(d) search direction, iter=3

(e) final reconstruction, iter=5
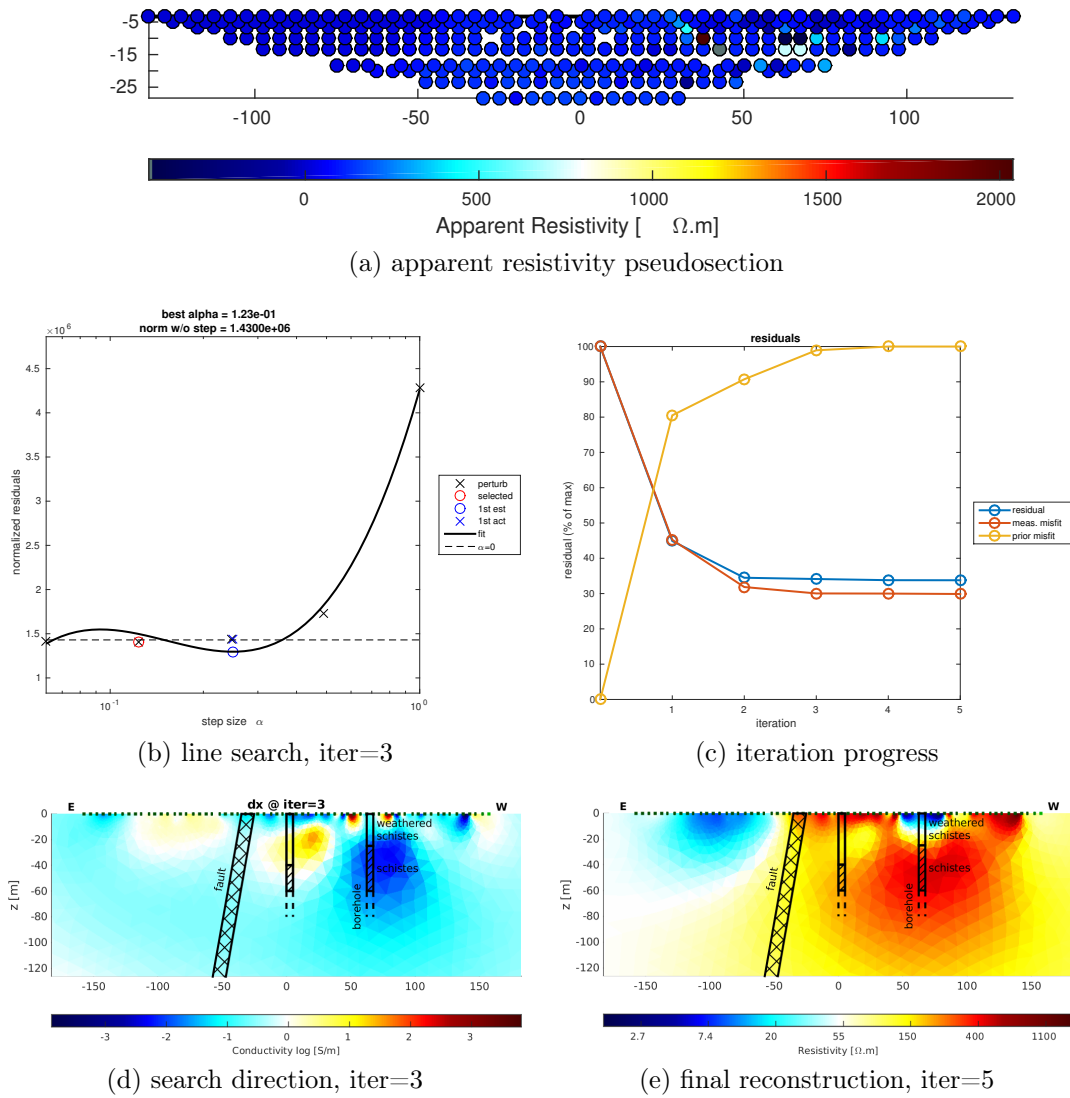
Figure 6.8: Example of incorrectly wired electrodes, Pont-Péan; (a) apparent resistivity with notable outliers, (b) a curve fitted to the line search test points (×) reflects a poor choice of direction (circles) at the third iteration, (c) the reconstruction progresses over the course of 5 iterations where the residual (blue) fails to converge sufficiently as the measurement misfit was reduced (red) while regularization misfit (yellow) was grew in suggesting little progress in fitting the data was being made after the second iteration, (d) search direction at iteration 3 and (e) final reconstructed conductivity ($\lambda = 32.0, \sigma_0 = 21.1 \ \Omega \cdot m$) at 5 iterations shows no agreement with geological data

6.5 Contributions and Discussion

In this work, a framework for evaluating the quality of input models and data on a trusted algorithm for inverse problems has been proposed. An emphasis was placed on the ability to observe the dynamics of the inverse solver behaviour. Increased internal visibility allows observation of the self-correcting behaviour inherent in iterative solutions to regularized inverse problems and iterative solvers in general. The increased observability facilitates narrowing of the root cause to a specific measurement's input data quality or model mismatches when compared to examination of solely the final result.

The framework outlined in this and the previous chapter support a comprehensive method for debugging complicated inverse problem algorithms. In the field, these methods allow the source of a bad reconstruction to be located and corrected. In after-the-fact reconstructions, the methods outlined here can confidently direct the choice to exclude data or adjust models.

The author implemented the Gauss-Newton/Conjugate Gradient solver as described in this and previous chapters. Integrated into this implementation is the ability to textually log algorithm progress and generate representative plots of iteration progress. Together, the logs and plots build a picture of algorithm behaviour that can be used to evaluate the quality of a reconstruction. The key contributions include (a) the absolute iterative inverse solver design and implementation, (b) a $2^{1/2}$-D forward solver implementation (Appendix D), (c) an analytic PEM half-space forward solver implementation (Appendix A, (A.4)), and (d) the debug methods outlined in this and the previous chapter. All implementations were contributed to the open source EIDORS project and licensed GNU GPL version 2 or version 3.

The Jacobian used in the Gauss-Newton iterations plays an important role in determining the search direction. When an electrode moves but the model does not allow for it, artifacts that obscure any useful reconstruction may result. A Jacobian for electrode movement, in combination with a forward model update method, may be used to adapt the model to electrode movements. We next take a closer look at the available Jacobian techniques for estimating electrode movement.

Chapter 7

Electrode Movement and Modelling Errors

[This work was, in part, presented at the International Conference on Biomedical Applications of Electrical Impedance Tomography, Neuchâtel, Switzerland, June 2– 5, 2015 (Boyle et al., 2015a).]

In both geophysics and biomedical impedance imaging, electrode movement and boundary modelling errors can be the cause of significant image reconstruction artifacts. Some boundary movements may be modelled as movement of the attached electrodes. Electrode movement may, in turn, be modelled as additional parameters in the inverse problem. To reconstruct electrode movement, a movement Jacobian must be constructed to estimate the first derivatives of the measurements with respect to the electrode movements.

In this work, we compare four methods of estimating the movement Jacobian in the context of a simplified four electrode homogeneous half-space. The four Jacobian
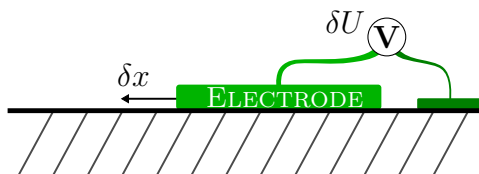
Figure 7.1: Electrode movement; an electrode ▮ with a measured potential difference **V** relative to a reference electrode ▮ is moved δx resulting in a change in the measured voltage δU

methods, detailed in the following sections, are: the naïve perturbation method, a minimal mesh perturbation method, a rank-one matrix update (Gómez-Laberge and Adler, 2008), and the recently developed Fréchet derivative (Dardé et al., 2012). By comparing these methods we mean to explore the sensitivity of the perturbation methods, understand and implement the Fréchet derivative, and ask the Engineering question: *Do these methods give the same answer for a simplified model?*

7.1 Movement Jacobian

A "movement Jacobian" is a first-order estimate of the effect of an electrode movement on the impedance imaging measurements. In impedance imaging, the movement Jacobian has a prominent role in some techniques for addressing the issue of boundary errors. Mismatches between the physical and model boundary, boundary errors, may result in incorrect interpretation of the image due to misleading reconstruction artifacts. Movement of the model boundary may be extrapolated from electrode displacements by interpolating surface movement between the electrodes. Some artifacts may be addressed by adapting the model to allow for electrode movement. Electrode displacements may be reconstructed by setting electrode positions as model parameters and estimating an electrode movement Jacobian. A movement

Jacobian matrix $\mathbf{J}_{m,x}$, captures how a small electrode movement δx affects each measurement δU (Figure 7.1)

$$J_{m,x}^{i,j} = \frac{\partial U_i}{\partial x_j} \tag{7.1}$$

where matrix row i, column j is the i-th measurement's change in voltage given a small movement on electrode j. The movement Jacobian estimates the effect of an electrode's movement in the same way that a conductivity Jacobian (§4.4) reflects how a regional conductivity change affects each measurement.

For the PEM, the Jacobian captures movement effects, but real electrodes modelled using the CEM have non-infinitesimal geometry and an associated contact impedance which affects current flow in the vicinity of the electrode. In general, the size, shape, rotation, and contact impedance are assumed to be constants in calculating the Jacobian. Electrodes moved in such a way are treated as inflexible and non-rotating electrodes. This has some clear limitations in the context of curved surfaces where the electrode can only move a certain distance without some sort of rotation to maintain contact with the boundary. In practise, the components of a movement vector are typically applied as tangential and normal movements rather than translation in absolute coordinates. Contact impedance and full movement vectors could, in principle, be estimated as columns in the Jacobian. Contact impedance and the conductivity distribution have been simultaneously reconstructed for impedance images (Vilhunen et al., 2002; Heikkinen et al., 2002; Winkler and Rieder, 2015). Sensitivity to normal movements are usually such that their effects are difficult to estimate accurately.

For Gauss-Newton updates, the Jacobian $\mathbf{J}$, regularization $\mathbf{R}$, hyperparameter λ

and data misfit $\mathbf{r}$, determine the search direction $\delta\mathbf{x}$. A line search then determines a distance α along that search direction. If the system is sufficiently linear, a line search is unnecessary and $\alpha = 1$ can work well.

$$\mathbf{r}_n = \mathcal{F}_{\mathbf{b}}(\boldsymbol{\sigma}_n, \mathbf{x}_n) - \mathbf{U} \tag{7.2}$$

$$\delta\mathbf{x}_n = (\mathbf{J}^\mathsf{T}\mathbf{J} + \lambda^2 \mathbf{R}^\mathsf{T}\mathbf{R})^{-1} \, \mathbf{J}^\mathsf{T}\mathbf{r}_n \tag{7.3}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha \, \delta\mathbf{x}_n \tag{7.4}$$

Errors of magnitude in the Jacobian may be circumvented by a reasonably accurate line search. On the other hand, any systematic errors in the movement Jacobian's direction estimate may entirely confound an attempt to address the electrode movement issue. Therefore, the movement Jacobian plays a key role in estimating any electrode placement error in the model or accounting for electrode movements.

7.2 Movement Artifacts

Surprisingly small amounts of electrode or surface movement can be a source of significant conductivity reconstruction artifacts when electrodes are incorrectly positioned. Electrode movements of as little as 5.7% of electrode spacing cause significant artifacts for elliptical 16 electrode (lung EIT) configurations (Adler et al., 1996). Electrode movements of 10% have been observed to give 20% resistivity artifacts far collinear electrode array (2D geophysics ERT) configurations (Zhou and Dhalin, 2003). To an experienced observer, an individual small electrode movement can be identified as conductivity image artifacts surrounding the erroneously placed electrode.

(a) Forward model

(b) No electrode movement

(c) 1% electrode movement

(d) 5% electrode movement

(e) 10% electrode movement

(f) 25% electrode movement

(g) 50% electrode movement

(h) −50% electrode movement

(i) Resistivity $[\Omega \cdot m]$
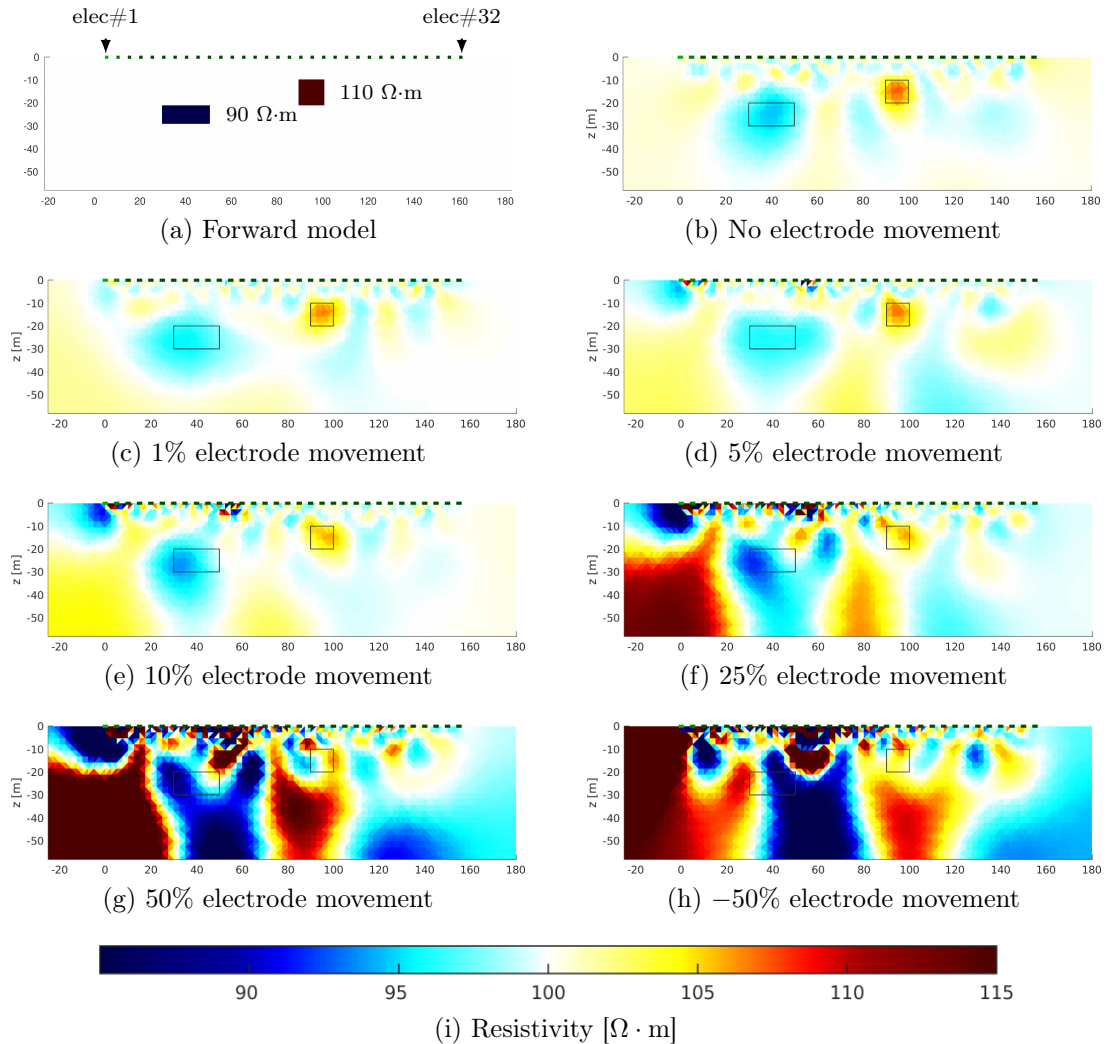
Figure 7.2: Electrode movement artifacts; simulated reconstructions based on forward model (a), each with two electrodes (electrode #2 and #12 of 32 electrodes numbered left-to-right at 5 m intervals) having electrode displacements of (b) 0% (c) 1%, (d) 5%, (e) 10%, (f) 25% and (g) 50% of electrode spacing on a 2-dimensional half-space reconstruction (40 dB SNR, $\lambda = 0.01$, Laplace regularization, Wenner stimulus pattern). Note that when electrode movement is reversed (h) -50%, the conductivity artifacts near the moved electrodes are, for the most part, reversed. Single or well separated electrode location errors introduce characteristic "ringing" artifacts that can overwhelm conductivity-based information.

In Figure 7.2, a two-dimensional half-space model with background resistivity of 100 $\Omega\cdot$m, a conductive target (90 $\Omega\cdot$m), and a resistive target (110 $\Omega\cdot$m) were simulated to generate measurements on a 32 electrode linear array (5 m electrode spacing) with the Wenner stimulus pattern. Conductivity images were reconstructed after adding noise (40 dB SNR AGWN) and moving two electrodes (electrode #2 and #12) in opposite directions by up to 50% (2.5 m) of electrode spacing (5 m). Artifacts initially appear near the moved electrodes for small movements and, for larger movements, spread to contaminate the entire image by masking the true targets (Figure 7.2b–g). When movements are reversed the artifacts near the electrodes are generally reversed: conductive artifacts become resistive and vise-versa (compare Figure 7.2g to Figure 7.2h). Larger movements or movements of multiple electrodes, where the effects interact, can lead to images that are very difficult to interpret. More extreme electrode movements may result in a reconstruction that fails to converge such that there is no useful image to interpret.

The electrode movement issue is, in fact, more tractable than it first appears. Impedance imaging methods are quite good at detecting small electrode movements. Using the same forward model (Figure 7.2a), the change in all measurements was plotted against a single electrode's movement over the range of -50% to +50% of electrode spacing (Figure 7.3). The measurement changes were nearly linear for a single electrode's movement: generally within 1% of measurements. The Jacobian, therefore, was nearly linear for a relatively broad range of electrode movement. We note that the assumption of orthogonality in the Jacobian will be approximately correct for "small" electrode movements or an individual large movement.

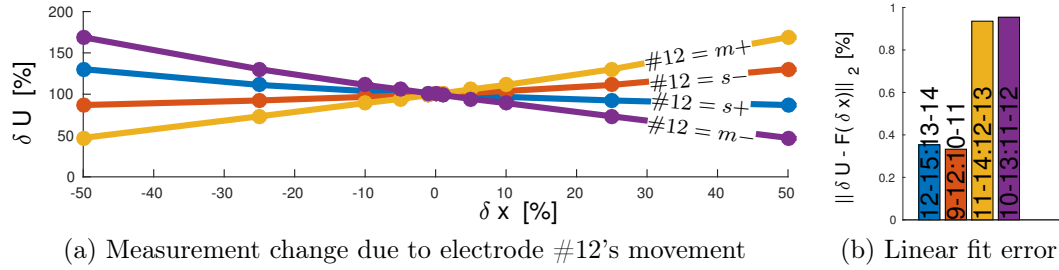(a) Measurement change due to electrode #12's movement

(b) Linear fit error

Figure 7.3: Change in measurement versus electrode movement; (a) showing selected measurements using electrode #12 (the movement electrode) when moved as both a stimulus and measurement electrode, measurement voltages are shown as a percentage of the zero movement measurement over the range of $\pm 50\%$ of electrode spacing (11 movement samples for each of 4 measurements, no noise, forward modelling only, Wenner 32 electrode stimulus pattern), colours between (a) and (b) are consistent, measurements used electrodes $[s+\ s-{:}m+\ m-]$ as indicated in (b), changes in measured voltages are nearly linear with electrode movement (b) as calculated by the 2-norm error on a linear line fit of measurement against electrode movement ($< 1\%$ measurement misfit over $\pm 50\%$ movement)

7.3 Movement Perturbation Jacobian

Calculations to estimate the movement Jacobian by direct perturbation of an FEM mesh requires multiple forward solutions: one per electrode and movement dimension. (In three dimensions, a 32 electrode array would require 96 additional forward solutions per Gauss-Newton iteration.) The method is a direct implementation of equation (7.1): the partial derivatives are replaced with a small perturbation δx_j of electrode j's location

$$J_{m,x}^{i,j} \simeq \frac{\delta U_i}{\delta x_j} = \frac{\mathcal{F}(\delta x_j) - \mathcal{F}(0)}{\delta x_j} \tag{7.5}$$

where δU_i is the difference between two forward model $\mathcal{F}$ solutions where the only change between forward models is the location of electrode j. The method inherently

captures all relevant movement effects in the model because the model is completely rebuilt to incorporate the updated electrode position.

Figure 7.4a illustrates a naïve perturbation implementation for movement of a single CEM electrode's nodes in an FEM mesh. Perturbations such as the one illustrated here must select a small enough perturbation to approximate the partial derivative while using a large enough perturbation to remain numerically stable. The perturbation method is fundamentally a calculation of the difference between two inverted FEM system matrices given the same stimulus.

$$\Delta\mathbf{u} = \mathbf{u}_b - \mathbf{u}_a = \mathcal{T}\mathbf{X}_b - \mathcal{T}\mathbf{X}_a = \mathcal{T}(\mathbf{A}_b^{-1} - \mathbf{A}_a^{-1})\mathbf{B} \qquad (7.6)$$

Numerical stability issues due to a perturbation are typically observed when the difference in floating point errors between two system matrices overwhelm the change in measured voltage at the electrodes. The movement Jacobian is a function of the system matrix which is dependent on the conductivity distribution, boundary, electrode placement, electrode size and electrode contact impedance. The selection of a stable perturbation magnitude is then dependent on variables that are typically unknown or uncertain to a significant degree.

We observe that a difference solution will cancel common errors in two inverted system matrices. Movement of all of an electrode's nodes affects the shape of all elements directly connected to the electrode. A piece-wise linear forward model of potential within the domain usually requires smaller elements near the electrode boundary to accurately approximate the electrical discontinuity at the edge of an electrode. Therefore, a small perturbation to the electrode location will affect a large number of mesh elements exactly where the mesh is most sensitive to numerical noise.
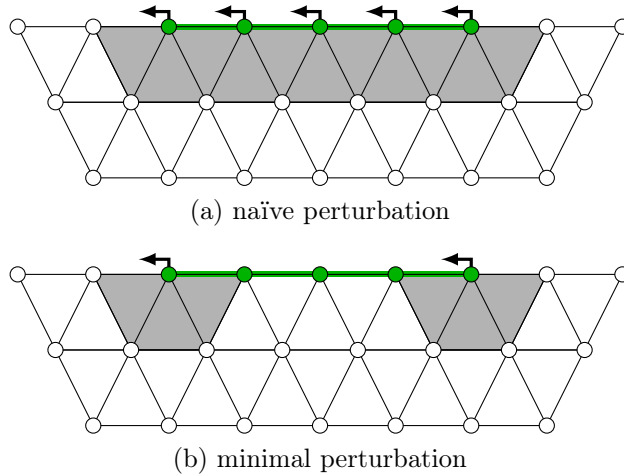
(a) naïve perturbation

(b) minimal perturbation

Figure 7.4: Movement Jacobian via perturbation; (a) all nodes ● of the FEM mesh belonging to a particular electrode are uniformly shifted by a small perturbation ↰, (b) only nodes along the edge of the electrode are perturbed giving a more stable Jacobian estimate and faster computations, grey elements ▽ are those elements connected to the perturbed nodes which must be updated in the FEM system matrix

7.4 Improved Perturbations

To improve compute time for the perturbation method, we introduced an alternate "minimal" perturbation strategy (Figure 7.4b). Rather than moving all nodes associated with an electrode, only nodes at the boundary of the electrode were shifted. The two solutions, naïve and minimal perturbation, are analytically equivalent because the boundary conditions are consistent. None the less, numerical differences exist due to the order of operations: at which point the problem is converted from a continuous domain to the discrete one. The choice of when and how to discretize implicitly selects a subset of the possible solutions. The minimal perturbation method should suppress instabilities in the naïve perturbation method because fewer elements of the difference in system matrices are subject to non-cancelling floating point errors.

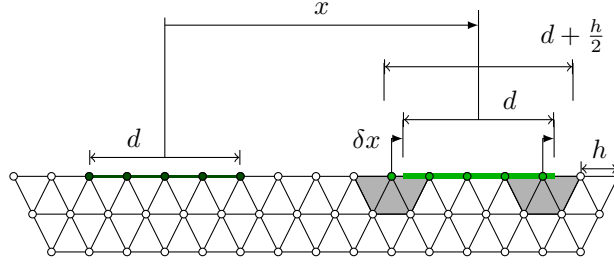The minimal perturbation method may be extended to handle large electrode

Figure 7.5: Large movement perturbations; an electrode ● is relocated by a distance x to a new site ●, (a) nodes in the vicinity of the new location $d + h/2$ replace the old electrode nodes, and (b) the new electrode boundary nodes are perturbed δx to exactly match the correct electrode boundary ↱, only nodes along the edge of the electrode are perturbed, grey elements ▽ are those elements connected to the perturbed nodes which must be updated in the FEM system matrix along with CEM connectivity and contact impedance distribution

movements (Figure 7.5). Assuming sufficient element refinement over the surface, the method may be applied by selecting nodes on the mesh for a new electrode location. The new electrode's boundary nodes are then perturbed to match the exact electrode location and diameter. First, a new electrode location x is identified. The local node spacing nearby is measured h and all nodes within the electrode diameter plus a margin $d_{e\ell} + h/2$ are assigned to replace the former CEM nodes. The new electrode's boundary is identified and nodes on that boundary are perturbed radially with respect to the electrode centre so that the electrode boundaries exactly match the prescribed shape and location. For iterative solutions, perturbations start with a common mesh so that mesh quality is not successively degraded by perturbations.

Finally, by identifying the elements affected by nodal perturbations and CEM connectivity changes, the system matrix $\mathbf{A}$ may be updated by only recalculating modified elements in $\mathbf{S}^{(e)}$ and reusing unmodified elements. The system matrix may then be reassembled at little relative cost. Typically, we observed a speed-up of between two and six times the naïve perturbation approach.

7.5 Rank-one Update Perturbation

One technique for reducing the quantity of calculations required for a mesh pertur-
bation is to develop a rank-one matrix update to the FEM system matrix. The
rank-one matrix update for electrode movement (Gómez-Laberge and Adler, 2008)
has generally been used for single-step Gauss-Newton solutions where the electrode
movement is small; typically less than 1% of electrode spacing.

The rank-one update is an application of the Sherman-Morrison formula (Sher-
man and Morrison, 1950)

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^\mathsf{T})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^\mathsf{T}\mathbf{A}^{-1}}{1 + \mathbf{v}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u}} \tag{7.7}$$

which computes the inverse sum of an invertible matrix $\mathbf{A}$ (our FEM system matrix)
and an outer product $\mathbf{u}\mathbf{v}^\mathsf{T}$: the movement perturbation[1]. The particular node to
perturb is selected by setting $\mathbf{u}, \mathbf{v}$ as zero column vectors and non-zero at the row $\mathbf{u}_i$
and column $\mathbf{v}_j$ of the matrix entry (i, j) that is to be perturbed. The determinant
of the perturbed matrix $\mathbf{A}$ is given by the well known matrix determinant lemma
(Ding, 2007)

$$\det(\mathbf{A} + \mathbf{u}\mathbf{v}^\mathsf{T}) = (1 + \mathbf{v}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u}) \det(\mathbf{A}) \tag{7.8}$$

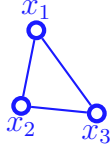for a perturbation $\mathbf{u}\mathbf{v}^\mathsf{T}$.

---

[1]The rank-one update is not invertible if $\mathbf{u}\mathbf{v}^\mathsf{T} = -\mathbf{A}$. This should not occur for any of our
applications: the perturbation would have to match, or match to within numerical precision after
inversion, the FEM system matrix.

The FEM matrices are calculated (Appendix C)

$$\mathbf{A} = \mathbf{C}^{\mathsf{T}}\mathbf{S}\mathbf{D}\mathbf{C} \tag{7.9}$$

$$\mathbf{J}_{\sigma_n} = \frac{\partial \mathbf{v}}{\partial \sigma_n} = -\mathcal{T}\mathbf{A}^{-1}\mathbf{C}^{\mathsf{T}}\mathbf{S}\frac{\partial \mathbf{D}}{\partial \sigma_n}\mathbf{C}\mathbf{X} \tag{7.10}$$

for a system matrix $\mathbf{A}$ decomposed into a connectivity matrix $\mathbf{C}$ which associates local and global node numbering in the element mesh, the element shape functions $\mathbf{S}$, and the conductivity per element $\mathbf{D}$. The conductivity Jacobian $\mathbf{J}_\sigma$ for element n is calculated as a difference in potentials $\mathcal{T}$, given potential distributions $\mathbf{X}$. (Yorkey et al., 1987; Adler and Guardo, 1996). We note that $\mathcal{T}$ is an operator which selects difference measurements from all calculated forward solutions and returns them as a vector. For a single stimulus, $\mathcal{T}$ can be constructed as a linear matrix followed by a vectorization, but this does not generalize unless all stimulus patterns use the same sequence of measurements. When the shape functions are first-order linear functions of potential with piece-wise linear conductivity, they can be calculated in two dimensions $(n_D = 2)$ as

$$\mathbf{S}^{(e)} = \frac{1}{n_D!}\,\frac{1}{|\det\mathbf{E}|}\,\mathbf{E}_{\backslash 1}^{\mathsf{T}}\mathbf{E}_{\backslash 1} \qquad \mathbf{E}_{\mathrm{2D}} = \begin{bmatrix} 1 & x_{1x} & x_{1y} \\ 1 & x_{2x} & x_{2y} \\ 1 & x_{3x} & x_{3y} \end{bmatrix}^{-1} \tag{7.11}$$

where $\mathbf{S}$ is a block diagonal matrix with each element e defined by its node locations $\mathbf{S}^{(e)}$, $\mathbf{E}_{\backslash 1}$ is the matrix $\mathbf{E}$ with the first row removed, and $|\det\mathbf{E}|$ gives the area or volume of an element. The variables in the matrix $\mathbf{E}$ are illustrated by an example

of a triangular element (blue) with nodes located at $x_1 = (x_{1x}, x_{1y})$, $x_2 = (x_{2x}, x_{2y})$, and $x_3 = (x_{3x}, x_{3y})$.

Applied to electrode movement in impedance imaging, the process of Gómez-Laberge and Adler (2008) is to use a decomposition of the FEM system matrix $\mathbf{A}$ to determine the impact of perturbing a node. The movement Jacobian requires the application of the product of derivatives rule

$$\mathbf{J}_{x,n} = \frac{\partial \mathbf{v}}{\partial x_n} = -\mathcal{T} \mathbf{A}^{-1} \mathbf{C}^\mathsf{T} \frac{\partial \mathbf{S}}{\partial x_n} \mathbf{DCX} \tag{7.12}$$

$$\frac{\partial \mathbf{S}}{\partial x_n} = \frac{1}{n_D!} \left(\frac{\partial |\det \mathbf{E}|^{-1}}{\partial x_n} \mathbf{E}_{\backslash 1}^\mathsf{T} \mathbf{E}_{\backslash 1} + \frac{1}{|\det \mathbf{E}|} \left(\frac{\partial \mathbf{E}_{\backslash 1}^\mathsf{T}}{\partial x_n} \mathbf{E}_{\backslash 1} + \mathbf{E}_{\backslash 1}^\mathsf{T} \frac{\partial \mathbf{E}_{\backslash 1}}{\partial x_n} \right) \right) \tag{7.13}$$

where x_n refers to a global node numbered n and affects all element shape functions $\mathbf{S}^{(e)}$ connected to that node. The partial derivatives of the first-order interpolatory shape function may then be approximated

$$\partial_x |\det \mathbf{E}|^{-1} = \frac{\mathbf{v}^\mathsf{T} \mathbf{E} \mathbf{u}}{|\det \mathbf{E}|} \tag{7.14}$$

$$\partial_x \mathbf{E}_{\backslash 1} = -(\mathbf{E} \mathbf{u} \mathbf{v}^\mathsf{T} \mathbf{E})_{\backslash 1} \tag{7.15}$$

using the rank-one perturbations vectors $\mathbf{u}$ and $\mathbf{v}$ to select the row and column to manipulate by a small perturbation (Gómez-Laberge and Adler, 2008).

The rank-one update avoids the need to reassemble an entire FEM system matrix and recalculate the matrix inverse in the forward solution for each electrode movement. The previously described "minimal perturbation" technique might be applied to reduce the number of nodes moved in the rank-one update. In practise, the benefit of a minimal perturbation is small because the rank-one update is already

computationally inexpensive and numerically stable for perturbations involving a small number of nodes.

7.6 Fréchet Derivative for Tangential Movement

The Fréchet derivative for tangential electrode movement (Dardé et al., 2012) accounts for contact impedance z_c effects when using the CEM for calculating the electrode movement Jacobian in a manner similar to the adjoint method for calculating the conductivity Jacobian (§4.4). The Fréchet derivative is computationally efficient because it can take advantage of the forward solutions already calculated for the conductivity Jacobian: no new forward solutions are required when using a "complete" stimulus sequence. A complete pair-wise stimulus sequence is one that uses all possible electrode pairs.

Each measurement m is calculated as the difference between potentials at two electrodes. For the CEM, each electrode has a measured voltage U and a varying potential u along the contact surface with the domain (Figure 7.6a). Between these two potentials, lies a distributed contact impedance layer which allows current to
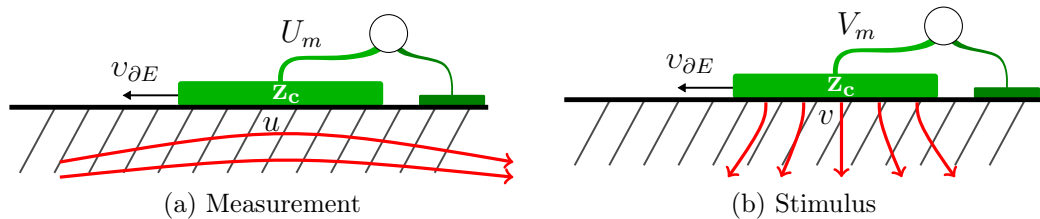

(a) Measurement   (b) Stimulus

Figure 7.6: Stimulus and measurement fields; under a particular electrode ▮ attached to a flat surface with contact impedance z_c, (a) the measurement field u, and (b) the stimulus field v are simulated to estimate the effect of the electrode movement $v_{\partial E}$ with respect to a possible measurement V_m, U_m.

flow along the electrode. In calculating the movement Jacobian, we are interested in the change in the measured electrode voltage δU assuming the reference electrode's measured voltage is unaffected by the primary electrode's movement.

For a small electrode movement h, the tangential component $\upsilon_{\partial E}$ of the movement Jacobian J_x may be calculated according to the main result derived in Dardé et al. (2012), equation (11)

$$\delta U_m = J_x h + O(h^2) \tag{7.16}$$

$$J_x\, h = \frac{1}{\ell z_c} \int_{\partial E} (h \cdot \upsilon_{\partial E})(U - u)(V - v)ds \tag{7.17}$$

where the measured voltage V and varying potential along the contact surface v are calculated when used as a stimulus electrode (Figure 7.6b). The 2D electrode is of length ℓ and contact impedance z_c, with contact impedance units of $[\Omega.\mathrm{m}]$. We note that our statement of this equation differs from that presented in Dardé et al. (2012) by adding the $1/\ell$ denominator term, where the units of contact impedance were not explicitly stated in the prior work. We find that this restatement gives consistent results for large contact impedances between the various methods while, at the same time, conforming to our definition of the contact impedance units used elsewhere in this work. Integrating over the surface of the electrode ∂E gives the movement Jacobian.

The implementation of the Fréchet derivative for electrode movement in two dimensions is particularly straightforward. For two-dimensional models, integration

over the surface of the electrode gives

$$\mathbf{J}_{x,e} = \frac{1}{z_c}\Big((U - u_1)(V - v_1) - (U - u_2)(V - v_2)\Big) \tag{7.18}$$

where electrode potential when used as a measurement electrode U and stimulus electrode V are combined with the potential at each edge of the electrode. The potential under the electrode when used as a measurement electrode, at both left u_1 and right u_2 edges of the electrode are combined with those as if used as a stimulus electrode, at both left v_1 and right v_2 edges of the electrode. The potentials were calculated by reversing the stimulus and measurement electrodes in the stimulus and measurement pattern and solving the forward problem again. These forward solutions would be available without additional calculations for a "complete" stimulus sequence where all combinations of stimulus electrodes were already calculated. We note that when the difference in potential across the electrode approaches zero due to low contact impedance, small electrodes, or weak stimulus, the Jacobian tends to zero.

The Fréchet derivative approaching zero for small contact impedances is inaccurate: the other methods presented here can correctly estimate a movement Jacobian under these conditions while the Fréchet derivative fails. This is somewhat disappointing, as the original proposition for the Fréchet derivative, beyond its computational efficiency, is that it accurately captures the CEM effects which are much more apparent for low contact impedances. The other methods succeed under low contact impedance conditions by sampling the rate of change in the electric fields beyond the electrode's boundary. This suggests a relatively simple idea for "fixing" the Fréchet derivative: to sample just beyond the electrode's boundary. Clearly,

this would have deep implications on the mathematical derivation of the Fréchet derivative: modifications are not considered further in this work.

We implement the Fréchet derivative for tangential movement in two dimensions and compare it to the three perturbation methods: naïve perturbation, minimal perturbation, and the matrix rank-one update.

7.7 Methods

A rectangular two-dimensional model with four CEM electrodes on its upper surface was constructed in EIDORS using NetGen (Figure 7.7). Many models were generated that conformed to the same geometry but with variations in the electrode diameter, mesh density and contact impedance. Mesh density, electrode diameter and contact impedance were varied over orders of magnitude. Simulations were run by generating a model for each parameter in the range: mesh density, contact impedance, and electrode diameter. A single parameter was varied over its range while the other two parameters were held fixed. On each model, the movement Jacobian was calculated for each electrode using our implementation of the naïve perturbation, minimal perturbation, rank-one update, and Fréchet derivative. The results were plotted to demonstrate the points of agreement and illustrate numerical instabilities between methods.

The model's conductivity was set to unity (1 Ω·m). The size of the model was set large enough so that expanding the sides or bottom did not significantly affect the measurements. An alternate approach would have been to use a mixed boundary condition that approximates the homogeneous conductivity in the $\pm x$ and $-z$
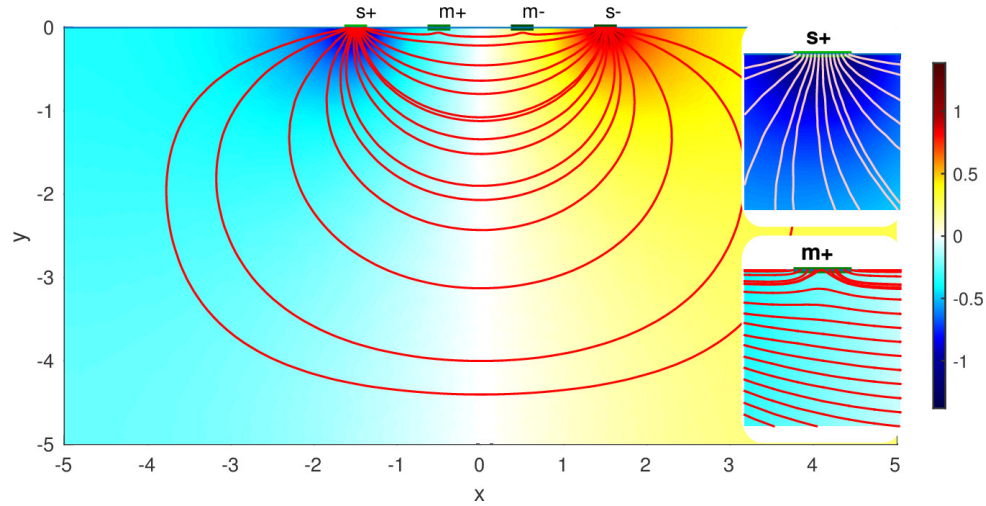
Figure 7.7: A two-dimensional four electrode half-space model; outer electrodes $s+$, $s-$ were used for stimulus and inner electrodes $m+$, $m-$ were used for measurements ($h_{max} = 0.1$ m, $d_{e\ell} = 0.2$ m, $z_c = 0.02$ Ω·m, $\sigma = 1$ S/m), streamlines (red) in the inset figures show current density near the electrodes. Image background colour shows the voltage distribution (voltage scale to the right of the figure),

directions to infinity. The expanded domain is substantially easier to implement and achieves the same result.

The four electrodes were spaced equidistant at 1 m intervals centre-to-centre. The electrode diameter $d_{e\ell}$ was varied from 0.9 m (nearly touching) down to 2 mm. Contact impedance z_c was varied between 10^{-14} Ω·m to 10^{+14} Ω·m. Mesh density, as measured by the maximum element height h_{max}, was varied between 0.10 m and 1 m.

The geometry and electrode parameters were selected so that the results should be general to any equally spaced linear electrode array. For example, contact impedance effects are a function of the ratio of background resistivity to contact impedance.

In Figure 7.7 ($z_c = 0.1$ Ω·m, $h_{max} = 0.1$ m, $d_{e\ell} = 0.1$ m), current flow density is indicated by stream lines (red) and voltage potential by the background colour map

(blue–white–yellow). Inset in Figure 7.7, are a close up of the positive stimulus electrode $s+$ (upper inset), and measurement electrode $m+$ (lower inset). The diversion of current flow through the measurement electrode can be observed for low electrode contact impedances in the lower inset.

## 7.8 Simulations

Plots illustrating variations in the Jacobian with respect to contact impedance (Figure 7.8a), mesh density (Figure 7.8b) and electrode diameter (Figure 7.8c) follow. Plots show the naïve perturbation, minimal perturbation, matrix rank-one update, and Fréchet derivative for the positive and negative measurement electrodes $(m+, m-)$. The dashed vertical line (red) indicates the configuration when a variable was fixed for other plots in Figure 7.8.

Similar plots for all electrodes were observed. We observed that the measurement electrode $m-$ plot's Jacobian were flipped vertically ($J_{m+} \simeq -J_{m-}$). The stimulus electrodes showed the same trends but with a reduced magnitude ($J_{m+} \simeq 3J_{s-} \simeq -3J_{s+}$). Looking to our model (Figure 7.7), we can see that this is the expected behaviour. Given a difference measurement between two electrodes $(m+, m-)$, in a fixed and smoothly varying field caused by stimulus electrodes $(s+, s-)$, we can consider movement of a single electrode. Moving a measurement electrode $m+$ in the $+x$ direction reduces the distance between the measurement electrodes, leading to a reduced difference measurement. Conversely, moving the other measurement electrode $m-$ in the $+x$ direction increases the distance between measurement electrodes, increasing the difference in measured potential. A similar thought experiment
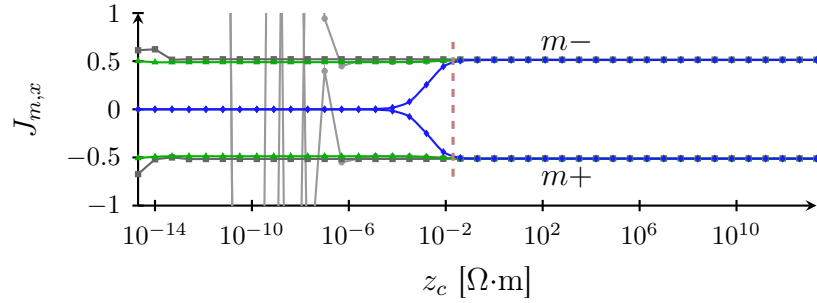
(a) Contact impedance z_c ($h_{max} = 0.1$ m, $d_{e\ell} = 0.2$ m)
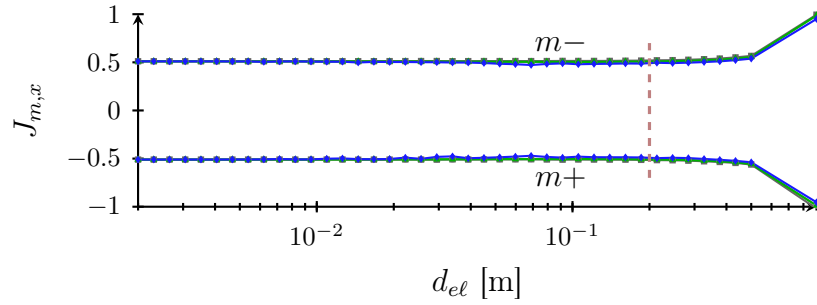
(b) Electrode diameter $d_{e\ell}$ ($h_{max} = 0.1$ m, $z_c = 0.02$ $\Omega \cdot$m)
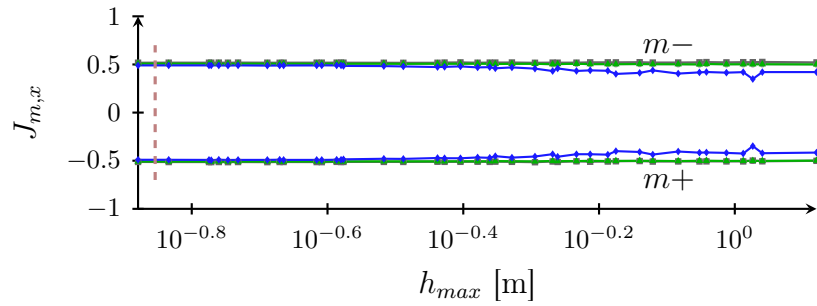
(c) Mesh density h_{max} ($d_{e\ell} = 0.2$ m , $z_c = 0.02$ $\Omega \cdot$m)

naïve perturb   min. perturb   rank-one update   Fréchet deriv.

Figure 7.8: Four Jacobian estimation methods for the $m+$ and $m-$ electrodes; (light grey) naïve perturbation, (dark grey) minimal perturbation, (green) matrix rank-one update (Gómez-Laberge and Adler, 2008), and (blue) Fréchet derivative (Dardé et al., 2012); showing variations in the Jacobian of the $m+$ measurement electrode for tangential surface movements $J_{m,x}$
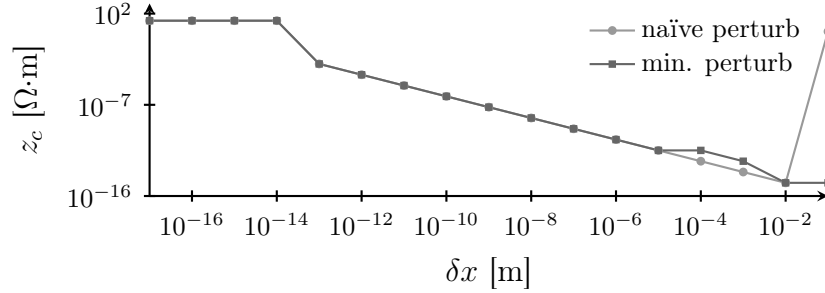
Figure 7.9: Perturbation instability; the point at which the perturbation Jacobian error became large relative to the rank-one update ($||\mathbf{J}_\Delta - \mathbf{J}_{R1}|| > 0.05$, $h_{max} = 0.1$ m, $d_{e\ell} = 0.2$ m), variation of mesh density h_{max} and electrode diameter $d_{e\ell}$ showed nearly identical plots, perturbation instability appeared to depended strongly on contact impedance z_c and step size δx, on the left perturbations fail when the perturbation approaches numeric precision limits $\delta x \to \epsilon = 1 \times 10^{-16}$, on the right of the plot the naïve perturbations fail for $\delta x > h_{min} = 0.017$ m the minimum element size

for the stimulus electrodes leads to the conclusion that these plots show the generally expected trends for electrode movement. The measurement electrodes in this model are separated by a third of the distance of the stimulus electrodes, so that the scaling between calculated stimulus and measurement Jacobians is approximately correct as well.

Our two independently developed implementations of the Fréchet derivative, based on Dardé et al. (2012), were found to give the same solution.

For large contact impedances, all methods performed well and gave similar results (Figure 7.8a). For small contact impedances, the naïve perturbation method became unstable for our selected perturbation step ($\Delta = 10^{-6}$ m). Changing the perturbation step size modified the threshold at which the naïve perturbation became unstable but never completely removed the instability (Figure 7.9). This unavoidable instability was expected because there are practical upper and lower limits on the perturbation step size. The mesh density and electrode diameter were varied

but these did not noticeably affect the plot of contact impedance versus perturbation size. For large step sizes, the "minimal perturbation for large steps" method gave "true" results to the limit where adjacent electrodes collided because the technique limits mesh element deformations to half an element's area. The rank-one update and naïve methods failed when nodal perturbations moved nodes beyond the enclosing elements leading to incorrect FEM meshes and an erroneous Jacobian.

For small contact impedances, we observed that the Fréchet derivative tends to zero. In the context of (7.17), this behaviour makes sense. The method uses the difference in potential across the electrode to calculate the effect of electrode movement. A contact impedance that tends towards zero will have a nearly zero potential difference across the electrode. We note that this behaviour is at odds with the perturbation method results and is a result of using finite precision calculations.

All methods exhibit the same stable and uniform behaviour for the majority of variations in electrode diameter (Figure 7.8b). For electrode diameters where the electrodes nearly touch, the absolute magnitude of the Jacobian grows. The case where electrodes nearly cover the boundary was not explored further, as all methods seemed to be in strong agreement.

For very coarse meshes, the Fréchet derivative was found to underestimate the Jacobian by some form of systematic noisy offset. This noise was possibly due to the sensitivity of the Fréchet derivative to errors in the FEM estimated voltages under the electrodes. The other perturbation-based methods rely solely on the change in measured electrode voltage, rather than a difference between measured voltage and voltage under the electrode. The naïve perturbation, minimal perturbation and rank-one matrix update gave consistent results for all tested mesh densities. We

note that the sign was correct in all cases: a line search could correct for any error introduced by the Fréchet derivative method.

Fast calculation of the movement Jacobian can improve productivity in algorithm development and throughput for Gauss-Newton iterations where the Jacobian is updated at each iteration. For a particular forward model ($h_{max} = 0.1$ m, $z_c = 10^{14}\Omega\cdot$m, $d_{e\ell} = 0.2$ m), the average calculation times for each movement Jacobian implementation were measured in Matlab[2] (Intel Core i5-2500K, 3.30 GHz, 32 GB mem) (Figure 7.10). Results are shown with and without EIDORS caching enabled. The EIDORS caching feature was not disabled because functions were called multiple times for certain operations, knowing that the cache will filter out most of the computational cost of these calls. EIDORS uses a function memoization implementation where a cached function result is returned when function inputs match a cache entry (Michie, 1968). To control the cache hit-rate, the ground node in the forward model was perturbed slightly so that the top-level function memoization would initially be defeated. Assembly and a single forward solution were not included in the timings, as these would already have been calculated in a Gauss-Newton iteration, and were considered a sunk cost. Subsequent calls to memoized results within a method benefited from the efficiency of a cached result if the forward model was not modified. When the cache was enabled or disabled, results showed that the Fréchet derivative method was fastest, while the rank-one update method was a close second.

Observing the profiling results, it was apparent that the vast majority of the computational time for the perturbation Jacobian (up to 96%) was consumed in recomputing the element shape functions $\mathbf{E}$ and then reassembling the system ma-

---

[2]Measurements performed with the `timeit` Matlab function.

trix **A**. With EIDORS' memoized caching mechanism, the entire system matrix was recomputed when a nodal perturbation was detected. For the specific case of computing the movement Jacobian, a finer grained caching of the inverted element matrices **E**, prior to being combined into the system matrix, considerably sped up the nodal perturbation methods because a handful of element matrices were recomputed. These improvements were implemented for the "min. perturb" method.

We note that the rank-one update, essentially, performs this optimization by only modifying elements of the system matrix that are directly connected to the perturbed node. The Fréchet derivative method avoids the system matrix computation costs altogether by not modifying the system matrices. It is possible, for a "complete" stimulus set where all possible combinations of stimulus electrodes are used, that no further forward solutions are required than have already been computed to that point in the Gauss-Newton algorithm. Such a computational saving is hard to quantify in a limited test bench such as the one presented here, but could reduce the incremental cost of calculating a movement Jacobian to a few multiply-and-accumulate operations if the forward solution nodal voltages are already available from the conductivity Jacobian.

3.25 2.76 3.48 cache disabled cache enabled 4 3 2 1 0

run time [sec]

0.52 1.51 0.23 0.93 0.13

naïve perturb min. perturb rank-one update Fréchet deriv.

(a) Total run time

fwd solve (0.096 sec)
system matrix (2.640 sec)

naïve perturb

min. perturb

update system matrix (0.221 sec)

rank-one update

specific calc (0.223 sec)

Fréchet deriv.

specific calc (0.053 sec)

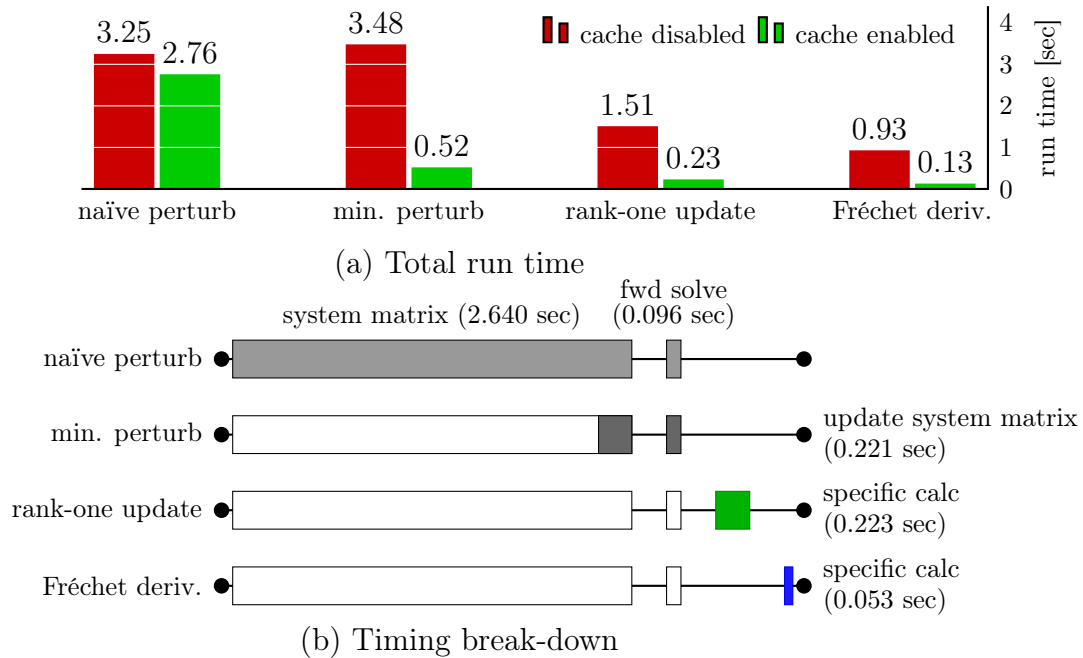(b) Timing break-down

Figure 7.10: Compute time for tangential movement Jacobian; (a) total run times compared between the different methods with and without function memoization (caching), (b) break down of timing where, for the rank-one update and Fréchet derivative, the clear boxes indicate a "sunk cost" that is already available if caching is enabled since the system matrix has not been perturbed

7.9 Discussion and Contributions

Boundary movement or imprecise electrode placement results in modelling errors that can cause significant reconstructed conductivity artifacts. Modelling boundary and electrode movement to address modelling errors and dynamic movement can correct some artifacts by incorporating electrode movement as inverse problem parameters to be resolved. The Jacobian calculation contributes to determining what parameter changes may be beneficial in minimizing a cost function.

In this work, four methods for calculating the electrode movement Jacobian were compared through simulation on a simplified two-dimensional homogeneous half-space. These methods were: the perturbation method, the rank-one matrix update (Gómez-Laberge and Adler, 2008), and the Fréchet derivative (Dardé et al., 2012). Simulations demonstrated that:

- Variations in electrode diameter gave stable results for all methods up to diameters where adjacent electrodes nearly touched.

- The perturbation method became unstable for small contact impedances depending on perturbation magnitude.

- The Fréchet derivative method was sensitive to coarse meshes and was not appropriate for small contact impedances.

- The Fréchet derivative and rank-one update methods were very fast, especially when accounting for previously solved forward problems in an iterative solution.

Based on these findings we recommend the Fréchet derivative for calculating the movement Jacobian when contact impedances are greater than the surrounding re-

sistivity distribution and the rank-one matrix update otherwise. When using the Fréchet derivative we recommend evaluation at two mesh densities to confirm that the Jacobian solution has converged.

We observed that from a computational point of view, the Fréchet derivative is most efficient, particularly when forward solutions are already available from calculating the conductivity Jacobian via the adjoint method or due to forward solutions previously calculated in an iterative reconstruction. The rank-one update is more efficient than the perturbation methods, but slower than the Fréchet derivative.

If a large change in electrode displacement is required to explain the data, a perturbation-based update will fail: overlapping element shape functions may be calculated. Overlapping element shape functions may lead to invalid FEM solutions. A superficial "fix" is to apply greater regularization to restrict the step size but this does not help achieve an accurate answer if larger movements are required to explain the measurements. In short, Gauss-Newton iterative solutions may make use of the rank-one update in calculating the Jacobian but will need an alternate strategy to safely update the FEM system matrix when movements exceed a fraction of the minimum element height of the mesh. Gauss-Newton updates and line search test points require an updated forward model where the electrode locations are updated. For tangential movements, we recommend a method where the new electrode location would be used to identify candidate nodes on the FEM mesh. Nodes from this set that are on the boundary of the electrode would then be perturbed to align them exactly with the new electrode location. This method is preferable to a mechanism where the whole mesh is distorted to push an electrode to its new location because it maintains the quality of the mesh elements, as measured by in-circle to

out-circle ratio or other mesh quality criteria (Shewchuk, 1996; Knupp, 2003; Pébay and Baker, 2003), which is important to the final FEM solution error (Berzins, 1999). Potentially, the system matrix may be updated more efficiently than a complete re-calculation by only updating shape sub-matrices $\mathbf{S}^{(e)}$ containing nodes that were perturbed and any electrode-to-node connectivity.

Unlike the conductivity Jacobian, the movement Jacobian may be nearly sparse for many common electrode configurations. Even for electrodes with very low contact impedance, effectively shunting current across their surface, the gaps between inactive electrodes must be small to have a measurable effect on the movement Jacobian. The sparsity of the movement Jacobian can be exploited to: reduce the number of calculations to build the Jacobian, reduce storage (for example, Compressed Sparse Row or Compressed Sparse Column storage), and expedite calculations using the Jacobian (for example, sparse matrix algebra).

Two independent implementations of the Fréchet derivative based on the derivation in Dardé et al. (2012) were built for this work. One implementation was written by the author, while another was contributed by a collaborator (M. Crabb). A third implementation (M. Jehl) was examined for consistency but was not integrated into the simulations presented here. This third implementation was difficult to integrate into our simulation framework without heavy modifications.

Further contributions in this work include: the simulation and evaluation framework, the minimal perturbation method, the extension of the minimal perturbation method for large electrode movements, and the system matrix update technique. These codes make use of EIDORS and NetGen for meshing, system matrix assembly, and forward simulations. An implementation of the system matrix update technique

and numerous performance optimizations were contributed to EIDORS.

Movement estimates tangential to the surface were faithfully calculated for many scenarios in this work. Using the same framework, movement normal to the surface was simulated with a naïve perturbation implementation, but appears to suffer from severe numerical instabilities that render the resulting Jacobian unusable in its current form. A Fréchet derivative for the normal component has also been developed (Dardé et al., 2013) and offers the possibility of a stable Jacobian for normal boundary movement. The methods in this work can be extended to normal and tangential movement in three dimensions; preliminary results for three-dimensional tangential perturbation methods show similar outcomes.

Chapter 8

Reconstructing Surface Movement

[This work was, in part, presentated at the 15th Conference on Electrical Impedance Tomography, Gananoque, Canada, April 2014 (Boyle et al., 2014).]

To demonstrate the practical application of both the electrode movement and iterative Gauss-Newton algorithms, conductivity and electrode movement were simultaneously reconstructed from data collected at an active landslide site. To our knowledge, this is the first simultaneous reconstruction of electrode movement and conductivity in a geophysics context. Electrode movement and conductivity have been reconstructed for biomedical problems where changes have been relatively mild, enabling single-step Gauss-Newton difference solutions (Soleimani et al., 2006). For geophysics, the conductivity contrasts are generally strong and electrode movements observed in the data presented here were as much as a third of the electrode spacing. These strong contrasts and large electrode movements necessitated an iterative solution because the combined effects are highly nonlinear. Two data sets from different locations were reconstructed using the same procedures to show the general

applicability of the method.

8.1 Background

Impedance imaging has been used in geophysical investigations of the behaviour and precursors of landslides and failure surfaces (Jongmans and Garambois, 2007; Perrone et al., 2004; Lapenna et al., 2005; Lebourg et al., 2005a; Naudetb et al., 2008; Sass et al., 2008). The technique is attractive because resistivity is strongly dependent on water saturation, fracturing, clay content and weathering which are all key factors in slope stability (Piegaria et al., 2009). Slopes may be monitored over time to observe changes in these key parameters using automated systems to collect and analyze data on a daily basis (Kuras et al., 2009; Lebourg et al., 2005b; Supper et al., 2014). Difference images may show immediate changes in water saturation (Suzuki and Higashi, 2001; Friedel et al., 2006; Jomard et al., 2007) but are limited in their ability to perform long-term monitoring due to background impedance changes and electrode movements. As illustrated in Figure 7.2, small amounts of electrode movement may introduce significant artifacts (Zhou and Dhalin, 2003; Oldenborger et al., 2005; Wilkinson et al., 2008). These artifacts may be reduced by accounting for the electrode movement.

An active landslide was identified in North Yorkshire[1], UK and has been monitored since 2008 (Wilkinson et al., 2010; Chambers et al., 2011). The landslide continues to move slowly, where a central portion of the slope has moved downhill by nearly two metres a year in some instances (Figure 8.1). At the top of the 40 metre hill, the landscape is lightly wooded and relatively flat. The top of the hill

---

[1]The hill is located at 54°06'39.2"N 0°57'34.9"W.

(a) slope rupture at main scarp, hilltop

(b) accumulated debris at
landslide toe, mid-slope

(c) satellite image 2016, overlaid electrode locations Aug 2014
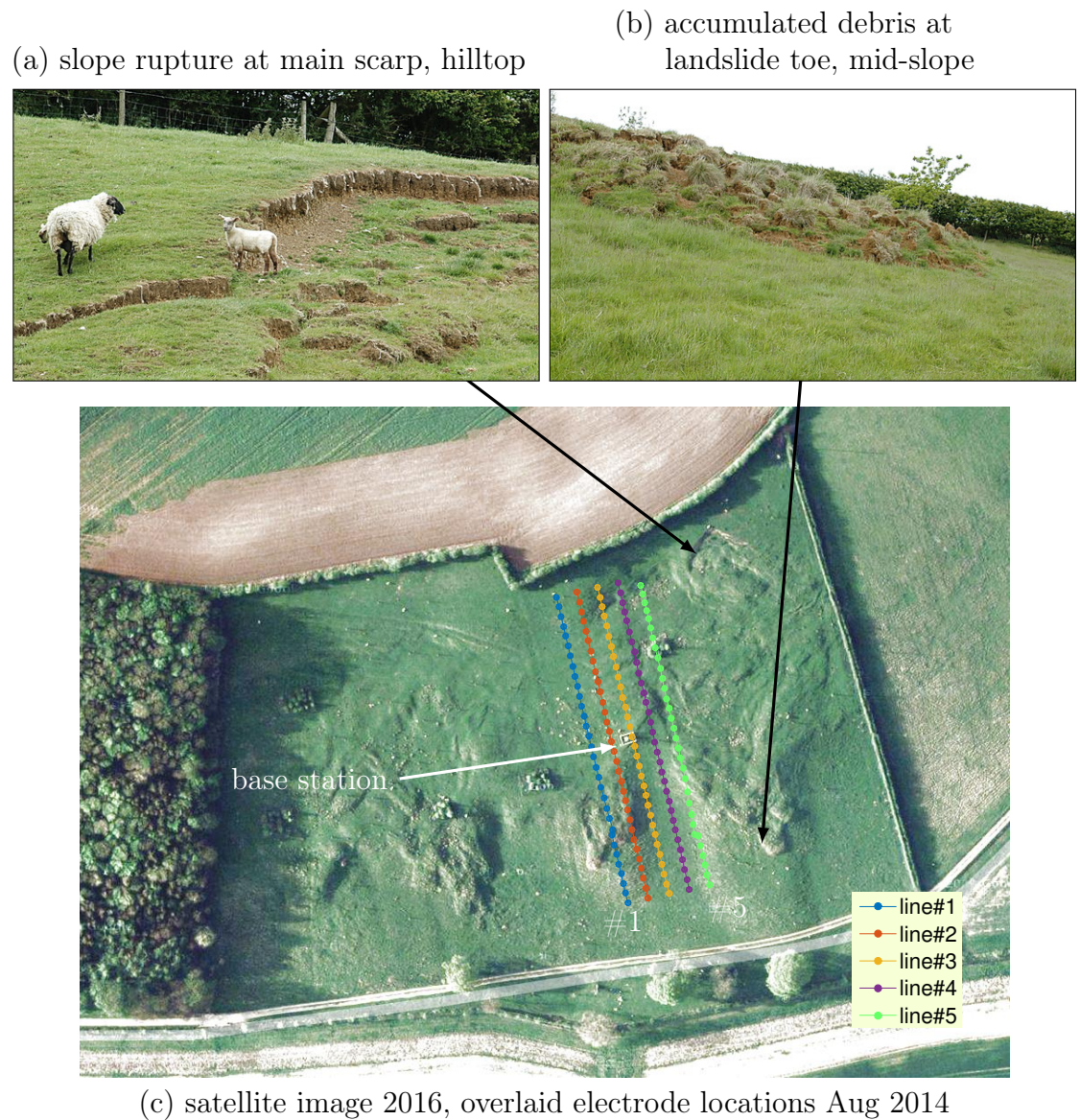
Figure 8.1: Slope failures at Hollin Hill; (a) rotational failures near the top of the slope above line#5, June 2015, (b) soil debris piled up at the toe of a landslide where line#5 runs through mid-slope with electrodes throughout the debris, June 2015, (c) satellite image of the hillside (2016), showing four landslide "lobes", 5 lines of 32 electrodes as of Aug 2014, and ALERT base station location

[(c) Satellite imagery ⓒ2016 DigitalGlobe, Getmapping plc, Infoterra Ltd Bluesky]

is relatively well anchored by the vegetation. The slope itself exposes four forma-tions: the Dogger Formation (DF), Whitby Mudstone Formation (WMF), Staithes Sandstone Formation (SSF) and Redcar Mudstone Formation (RMF), from top to bottom. The interfaces between sedimentary layers lie horizontally, with a gentle 5° dip to the North, determined through comparison of material interfaces at sur-rounding exposed slopes in the region (Chambers et al., 2011). The WMF, as the name implies, is a mudstone clay-based rock that highly weathered and is prone to movement during peak water saturation periods at Hollin Hill, occurring annually in the winter through early spring wet-season. The underlying SSF and unweathered RMF are structurally more competent, and landsliding is postulated to occur along the WMF/SSF interface (Uhlemann et al., 2015). The slope lies at an average angle of 12° over a change in elevation of 40 metres. Due to the low angle of repose and the geological structure, combined with the annual pattern of winter-spring precipi-tation, the portions of the slope near the WMF-SSF interface have tended to move annually, in a limited, slow moving landslide.

Five rows of thirty-two permanently installed electrodes travelled along with these movements, shifting their positions relative to each other (Figure 8.2). An automated impedance imaging survey was executed daily and data were transmitted to a remote office for storage and analysis. In 2008–2009, a middle section of line#1 exhibited a translational failure with movements of up to 1.6 metres. In 2013–2014, upper and middle segments of line#5 had rotational and translational failures of similar magnitudes.

(a) electrode locations

(b) stainless steel (polarizing) electrode
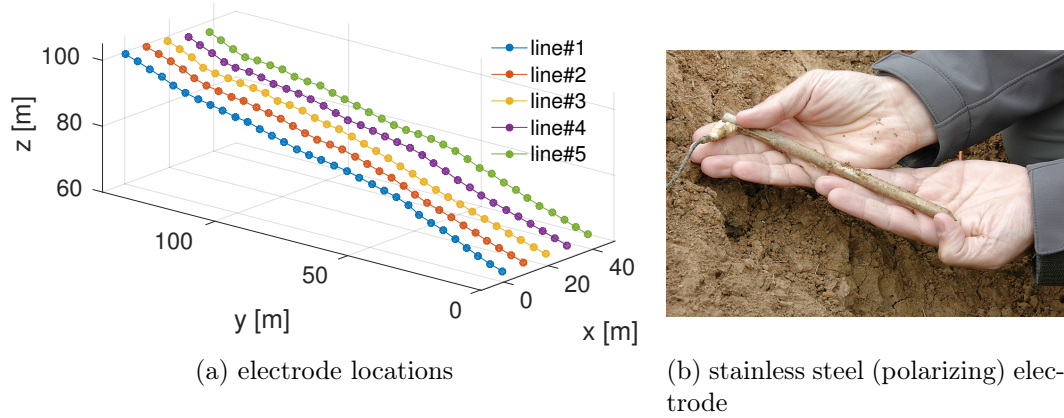
Figure 8.2: Electrode locations; (a) electrode locations for 5 lines of 32 electrodes each, line#1 (blue) to line#5 (green) as of August 2014, (b) where electrodes were 10 mm x 170 mm spikes of stainless steel (polarizing) selected for its conductivity, low cost, and corrosion resistance

8.2 Method

In this work, we reconstruct both conductivity and electrode movement simultaneously in a Gauss-Newton iterative framework. This joint inversion was achieved by balancing the sensitivity of the conductivity parameters against electrode movement

$$\hat{\sigma} = \arg\min ||\mathcal{F}(\sigma) - \mathbf{v}||_{\mathbf{W}} + ||\lambda\mathbf{R}(\sigma - \sigma_\star)||_2 \tag{8.1}$$

$$\delta\hat{\sigma} = -(\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J} + \lambda^2\mathbf{R}^\mathsf{T}\mathbf{R})^{-1}\left(\mathbf{J}^\mathsf{T}\mathbf{W}(\mathcal{F}(\sigma) - \mathbf{v}) + \lambda^2\mathbf{R}^\mathsf{T}\mathbf{R}(\sigma_\star - \sigma)\right) \tag{8.2}$$

for a forward model $\mathcal{F}$ taking model parameters σ, measurements $\mathbf{v}$, regularization $\mathbf{R}$, and a hyperparameter controlling regularization strength λ. The inversion parameters σ and regularization $\mathbf{R}$ have been extended to incorporate log conductivity

and longitudinal[2] electrode displacement.

$$\sigma \rightarrow [\mathbf{x} \ \log_{10} \sigma]^{\mathsf{T}} \tag{8.3}$$

$$\mathbf{J} \rightarrow [\mathbf{J_x} \ \mathbf{J}_\sigma \sigma \ln(10)] \tag{8.4}$$

$$\mathbf{R} \rightarrow \begin{bmatrix} \mathbf{R_x} & 0 \\ 0 & \beta \mathbf{R}_\sigma \end{bmatrix} \tag{8.5}$$

for inverse parameters as longitudinal electrode movement $\mathbf{x}$ and log conductivity $\log \sigma$. The Jacobian had additional columns added to estimate the first-order sensitivity of the measurements to longitudinal electrode movement. The conductivity Jacobian was corrected for a log conductivity parameter scaling. The regularization matrices were a Laplacian smoothing prior over log conductivity and Tikhonov prior $\mathbf{R_x} = \mathbf{I}$ over movement. The new regularization included a scaling parameter β which is a weighting term that sets the ratio of regularization applied to electrode movement and log conductivity. In principle, there are correlated changes between conductivity and electrode movement which may be partially accounted for by setting the off-diagonal blocks of the regularization matrix to non-zero values, but in practice these were not characterized and in the absence of a better guess were set to zero.

The forward model was constructed as a two-dimensional cross-section based on the original electrode locations and the mesh was then perturbed by PCHIP[3] interpolation (Carlson and Fritsch, 1989) for electrode displacements. Forward modelled measurements and the conductivity Jacobian were calculated using the $2^{1/2}$-D

---

[2]Longitudinal movement being movement inline with the electrodes and along the surface.

[3]Matlab `interp1(X,Y,Xq,'pchip')` one-dimensional interpolant based on Hermite derivatives.

method (Appendix D).

In this work, we have made use of the log conductivity to restrict the conductivity reconstruction to physically meaningful positive values. Experiments with the similar log movement constraint, to restrict electrodes to downslope movement, resulted in similar reconstructions to the ones presented here which used unscaled electrode movement. For the log movement parametrization, the behaviour at each iteration was different than an unscaled movement due to the structure of the movements in this data set. Because each electrode that moved had a different magnitude of movement, the log movement reconstruction tended to solve for each electrode's reconstructed displacement separately: one electrode per iteration. The apparent single electrode updates were actually an artifact of the log scaling, where smaller movements were reduced by orders of magnitude, so as to be inconsequential. Once the largest electrode placement error had been corrected, the next largest error would be addressed. This one electrode per iteration behaviour tends to favour accurate movements because it aligns with the construction of the electrode movement Jacobian. We feel that this highlights the importance of careful selection of the reconstruction parametrization. It is possible that some Fourier decomposition of electrode movement with appropriate regularization weighting per term might achieve similar reconstruction accuracy without artificially fixing any single electrode's location or degrees of freedom for movement.

8.3 Movement Jacobian

The movement Jacobian was initially calculated using the perturbation method with underlying $2^1/_2$-D forward simulations of the full FEM conductivity model. These movement perturbation calculations proved to be prohibitively slow. An alternate solution was implemented by adapting the half-space analytic PEM forward model (A.6) (Appendix A) and applying the electrode movement perturbation technique to calculate the Jacobian. Electrode positions were captured from the FEM model. A homogeneous conductivity was assigned based on the average conductivity of the current FEM model. The electrode movement Jacobian produced by the half-space analytic perturbation method was compared to the $2^1/_2$-D perturbation Jacobian under homogeneous conditions. The result was much faster to calculate and reasonably accurate: the effect of topography was found to be reasonably accounted for. The loss of accuracy due to changing electrode models (CEM to PEM) and using a homogeneous conductivity were not so disruptive as to change signs on the movement Jacobian. In the reconstructions presented here, this analytic perturbation Jacobian was used, though we believe that better results should be possible with a more accurate movement Jacobian.

Motivated by the efficiency of the 2D movement Jacobian of Gómez-Laberge and Adler (2008) (Chapter 7, Figure 7.10), the $2^1/_2$-D movement Jacobian was developed as the corollary of the $2^1/_2$-D conductivity Jacobian. Due to time constraints, the formulation described here was not fully evaluated in the following resistivity and movement reconstructions.

In general, the $2^1/_2$-D forward solver is a well known technique and is commonly used in geophysics ERT applications. An approximately half-space geometry, and

a conductivity that is nearly uniform along one axis, fit well with a 2¹/₂-D model, and occur naturally in many geological settings. By adapting the "standard method" of calculating the conductivity Jacobian to the 2¹/₂-D technique, conductivity may be efficiently reconstructed (Appendix D). To reconstruct electrode movement, we desire a similar 2¹/₂-D implementation for the electrode movement Jacobian. In the following, we outline the 2¹/₂-D conductivity Jacobian, referring to Appendices C and D for details of the FEM matrices, forward measurement calculations and conductivity Jacobian derivation. We present a new derivation for the 2¹/₂-D electrode movement Jacobian in this somewhat distributed context.

The 2D and 2¹/₂-D conductivity Jacobians were calculated

$$\mathbf{J}_{\sigma,\text{2D}} = -\mathcal{T}\mathbf{A}^{-1}\mathbf{C}^{\mathsf{T}}\mathbf{S}\frac{\partial \mathbf{D}}{\partial \sigma_n}\mathbf{C}\mathbf{X} \tag{8.6}$$

$$\mathbf{J}_{\sigma,\text{2}\frac{1}{2}\text{-D}} = -\frac{2}{\pi}\int_k \mathcal{T}\mathbf{A}_k^{-1}\mathbf{C}^{\mathsf{T}}(\mathbf{S}+k^2\mathbf{T})\frac{\partial \mathbf{D}}{\partial \sigma_n}\mathbf{C}\frac{\mathbf{X}}{2A} \tag{8.7}$$

for measurement selection $\mathcal{T}$, system matrix $\mathbf{A}$, mesh connectivity matrix $\mathbf{C}$, mesh shape functions $\mathbf{S}$, a conductivity change $\frac{\partial \mathbf{D}}{\partial \sigma_n}$, and the nodal voltages $\mathbf{X}$ over an electrode of width A modelled as a shunt in the y-direction.

Similarly, the 2¹/₂-D movement Jacobian may be calculated as an extension of the 2D Jacobian

$$\mathbf{J}_{x,\text{2D}} = -\mathcal{T}\mathbf{A}^{-1}\mathbf{C}^{\mathsf{T}}\frac{\partial \mathbf{S}}{\partial x_n}\mathbf{D}\mathbf{C}\mathbf{X} \tag{8.8}$$

$$\mathbf{J}_{x,\text{2}\frac{1}{2}\text{-D}} = -\frac{2}{\pi}\int_k \mathcal{T}\mathbf{A}_k^{-1}\mathbf{C}^{\mathsf{T}}\frac{\partial(\mathbf{S}+k^2\mathbf{T})}{\partial x_n}\mathbf{D}\mathbf{C}\frac{\mathbf{X}}{2A} \tag{8.9}$$

where the 2D movement Jacobian may be efficiently calculated using the rank-one

update (Gómez-Laberge and Adler, 2008). The rank-one update gives

$$\frac{\partial \mathbf{S}}{\partial x_n} = \frac{1}{2}\left(\frac{\partial |\det \mathbf{E}|^{-1}}{\partial x_n}\mathbf{E}_{\backslash 1}^{\mathsf{T}}\mathbf{E}_{\backslash 1} + \frac{1}{|\det \mathbf{E}|}\left(\frac{\partial \mathbf{E}_{\backslash 1}^{\mathsf{T}}}{\partial x_n}\mathbf{E}_{\backslash 1} + \mathbf{E}_{\backslash 1}^{\mathsf{T}}\frac{\partial \mathbf{E}_{\backslash 1}}{\partial x_n}\right)\right) \tag{8.10}$$

where x_n refers to a global node numbered n and affects all element shape functions $\mathbf{S}^{(e)}$ connected to that node. The partial derivatives of the first-order interpolatory shape function may then be approximated

$$\partial_x |\det \mathbf{E}|^{-1} = \frac{\mathbf{v}^{\mathsf{T}}\mathbf{E}\mathbf{u}}{|\det \mathbf{E}|} \qquad \text{and} \qquad \partial_x \mathbf{E}_{\backslash 1} = -(\mathbf{E}\mathbf{u}\mathbf{v}^{\mathsf{T}}\mathbf{E})_{\backslash 1} \tag{8.11}$$

using the rank-one perturbations vectors $\mathbf{u}$ and $\mathbf{v}$ to select the row and column the, specific element's node, to manipulate by a small perturbation (Gómez-Laberge and Adler, 2008). The $2^1/_2$-D movement Jacobian extends the 2D Jacobian

$$\frac{\partial \mathbf{S} + k^2 \mathbf{T}}{\partial x_n} = \frac{\partial \mathbf{S}}{\partial x_n} + k^2 \frac{\partial \mathbf{T}}{\partial x_n} \tag{8.12}$$

where $\partial \mathbf{S}/\partial x_n$ is already available, and the additional term $\partial \mathbf{T}/\partial x_n$ may be derived

$$\mathbf{T} = |\det \mathbf{E}|\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\frac{1}{12} \quad \text{so that} \quad \frac{\partial \mathbf{T}}{\partial x_n} = \frac{\partial |\det \mathbf{E}|}{\partial x_n}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\frac{1}{12} \tag{8.13}$$

for a 2D mesh. We then make use of the rule for determinants and their inverses

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det \mathbf{A}} \tag{8.14}$$

to reuse the previous calculation (8.11) for the partial derivative of the determinant of the shape matrix

$$\partial_x |\det \mathbf{E}|^{-1} = \frac{\mathbf{v}^\mathsf{T} \mathbf{E} \mathbf{u}}{|\det \mathbf{E}|} \qquad \rightarrow \qquad \partial_x |\det \mathbf{E}| = \frac{|\det \mathbf{E}|}{\mathbf{v}^\mathsf{T} \mathbf{E} \mathbf{u}} \tag{8.15}$$

which gives the additional partial derivative of the $2^1/_2$-D equation in terms of the same rank-one update perturbation applied to the 2D movement Jacobian

$$\frac{\partial \mathbf{T}}{\partial x_n} = \frac{|\det \mathbf{E}|}{\mathbf{v}^\mathsf{T} \mathbf{E} \mathbf{u}} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \frac{1}{12} \tag{8.16}$$

for a perturbation of a node selected by $\mathbf{u}, \mathbf{v}$ affecting a linear interpolatory shape function $\mathbf{E}$.

8.4 Prior Reconstructions

Electrode movement has been previously reconstructed for line#1 data (2008/2009) along with separate resistivity sections (Wilkinson et al., 2010). The algorithm used in that instance achieved estimates of within 0.2 metres (4% of electrode spacing) of the electrode's true positions as measured by RTK GPS[4]. An initial conductivity-only reconstruction with known electrode positions gave a plausible distribution that was in good agreement with available geological evidence: borehole and auger data, evaluation of local geology, aerial LiDAR, differential GPS, lab correlation of rep-

---

[4]Leica System 1200 RTK GPS in "kinematic mode" real-time correction achieves as much as 10 mm (RMS) horizontal and 20 mm vertical accuracies (Merritt, 2014).
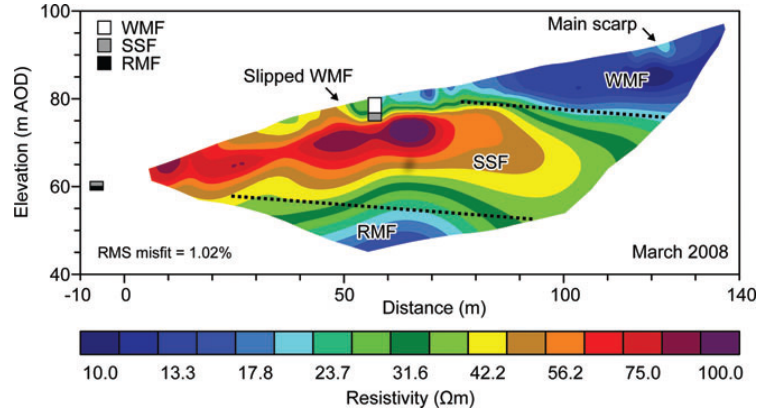
Figure 8.3: *[From Wilkinson et al. (2010), Figure 2:]*
2-D resistivity image inverted from the baseline data set (2008 March). The inferred boundaries between the Whitby (WMF), Staithes (SSF) and Redcar (RMF) formations are shown by dotted black lines. Stratigraphic logs of boreholes are shown in grey scale. The main scarp and slipped WMF material are indicated by the black arrows. *[Reproduced from Wilkinson et al. (2010), Figure 2, for comparison.]*

resentative samples to measured conductivities, piezometric pore pressure measurements, and *in-situ* rainfall and temperature records (Figure 8.3) (Chambers et al., 2011; Merritt et al., 2014). Resistivity reconstructions of data collected after movement exhibited artifacts. These artifacts were reduced when reconstructed movements were incorporated. The electrode movement was restricted to the downslope direction. The movement Jacobian was estimated based on an analytic half-space model with a homogeneous conductivity assigned to each group of measurements based on electrode separation. The electrode movement was then reconstructed by minimizing

$$\arg\min \sqrt{\sum_i |e_i|^2 + \alpha \sum_j |m_j| + \beta \sum_j \theta(m_j)|m_j|} \qquad (8.17)$$

for weighting factors $\alpha = 0.06$ m^{-1}, $\beta = 0.32$ m^{-1}, Heaviside step function θ, movement m_j, and misfit $e_i = r_b - r_a$ as the difference between predicted r_b and observed r_a apparent resistivity ratios. The apparent resistivity ratio was calculated as the ratio of the analytic half-space models (A.6) before and after movement

$$r = \frac{\rho'}{\rho} \left(\frac{\frac{1}{AM'} - \frac{1}{BM'} - \frac{1}{AN'} - \frac{1}{BN'}}{\frac{1}{AM} - \frac{1}{BM} - \frac{1}{AN} - \frac{1}{BN}} \right) \tag{8.18}$$

for homogeneous resistivity ρ, electrodes spaced AM, BM, AN, BN, and the updated locations and resistivity after movement indicated with a tick ρ'. Dipole-dipole data for measurements $n = 1$ were discarded, as they were judged to be too dependent on transverse movements which were not reconstructed.

A similar approach was applied (Wilkinson et al., 2016), to reconstruct two-dimensional xy-movements for the whole electrode grid by allowing for transverse movements through an additional weighting term

$$\arg\min \sum_i e_i^2 + \alpha \sum_j ||m_j|| + \beta \sum_j \theta(m_j^{(y)})|m_j^{(y)}| + \gamma \sum_j \theta(m_j^{(x)})|m_j^{(x)}| \tag{8.19}$$

facilitating balancing the sensitivities of transverse $m_j^{(y)}$ and longitudinal $m_j^{(x)}$ movements by adjusting the weighting ratio β/γ.

8.5 Results

Regions of high sensitivity were initially noted at depth where little sensitivity was expected. The model boundaries were extended approximately one electrode array length in each of the $+x, -x$ and $-z$ directions. To determine how far the FEM

model boundaries needed to be extended, an analytic PEM half-space model was compared to CEM homogeneous resistivity FEM simulations. The boundary extensions reduced boundary condition related errors in simulated measurements to within measured noise levels and removed the artifacts from the sensitivity plots when using the measured data.

The dipole-dipole stimulus and measurement sequence (Appendix A) used for line#1 was visualized in Figure 8.4, showing the sequence of tetra-polar measurements with stimulus in red and measurements in blue. A single difference measurement was captured as one row of the stimulus and measurement protocol in the figure. In the adjacent vertical strip chart, the horizontally aligned measurements (blue) were contrasted with the homogeneous resistivity estimate (green) of what those measurements would be, as well as showing their difference (red). The rightmost strip chart shows the estimated error for each measurement as estimated by reciprocal measurements.

The initial resistivity was reconstructed for line#1 (March 2008) and line#5 (February 2013) using the surveyed locations (Figure 8.5). The region near the electrodes was presented with annotations matching Figure 8.3, as well as an image of the complete model.

The final resistivity was also, independently, reconstructed using the final surveyed locations and used to create the expected resistivity change (Figure 8.7c,d). The initial resistivity for line#1 loosely matched those published in Wilkinson et al. (2010) (Figure 8.3) and achieved a similar <1% RMS measurement misfit relative to a homogeneous resistivity model. Superficially, the two reconstructions (Figure 8.3 and Figure 8.5) look very similar.
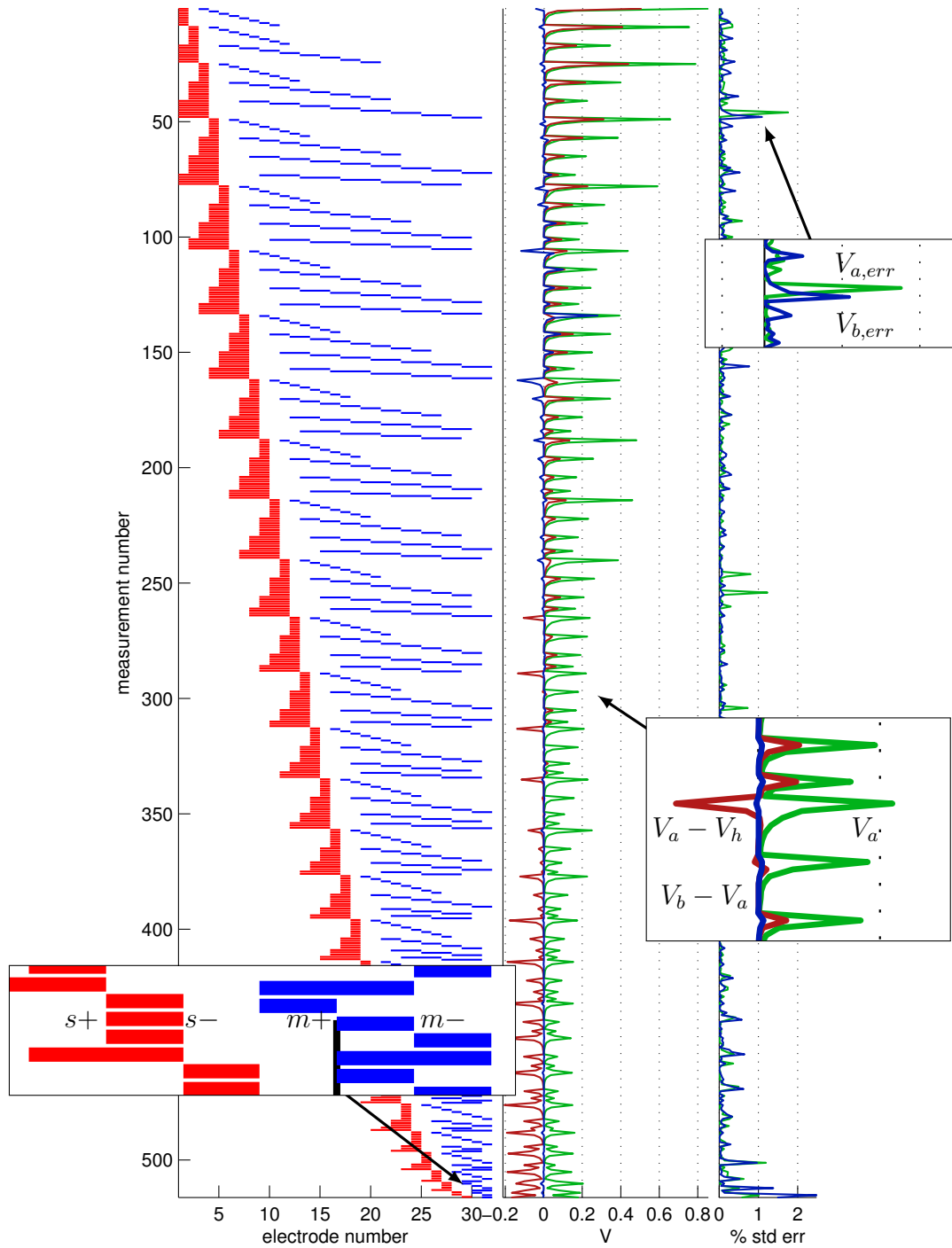
Figure 8.4: Dipole-dipole stimulus and measurement protocol for line#1; March 2008 measurements,(left) stimulus in red and measurements in blue, one row per difference measurement, (middle) initial difference measurements V_a (green) compared to homogeneous resistivity at $32.1\Omega \cdot$ m shown as $V_a - V_h$ (red), and the change from initial to final measurements $V_b - V_a$ (blue), and (right) the reciprocal standard error as a percentage of the measurements estimated by comparing reciprocal stimulus measurements for initial (green) and final (blue) measurements
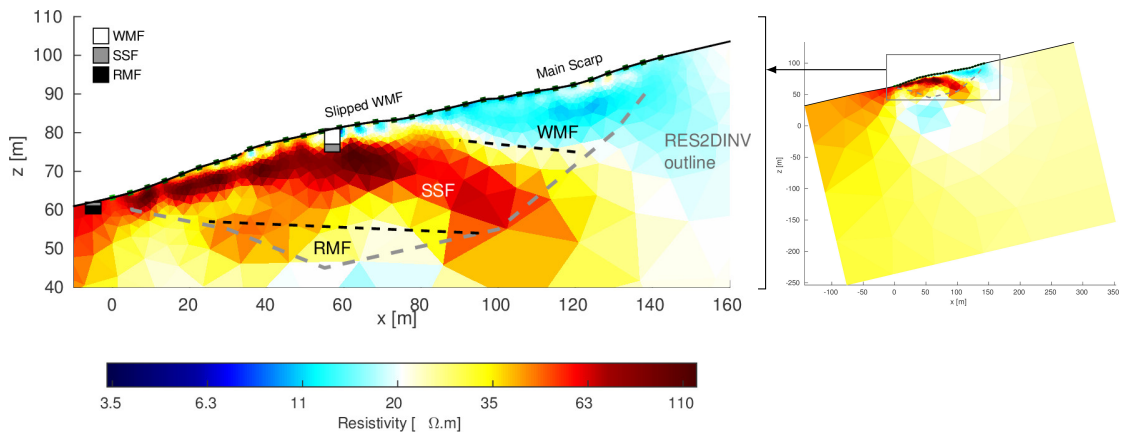
(a) line#1, March 2008
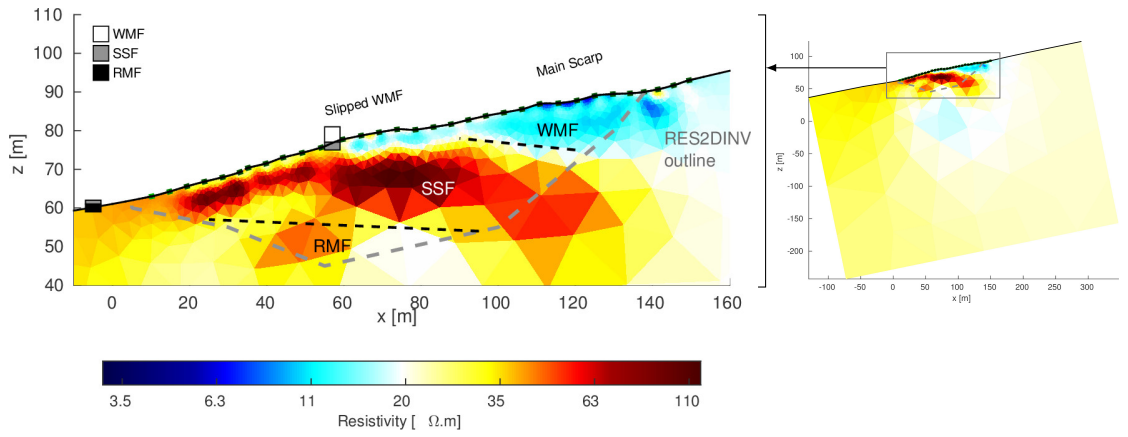
(b) line#5, February 2013

Figure 8.5: Initial resistivity-only reconstructions, (a) line#1 March 2008 ($\lambda = 130.4, \sigma_0 = 31.3 \ \Omega \cdot m$), and (b) line#5 February 2013 ($\lambda = 106.8, \sigma_0 = 25.6 \ \Omega \cdot m$)
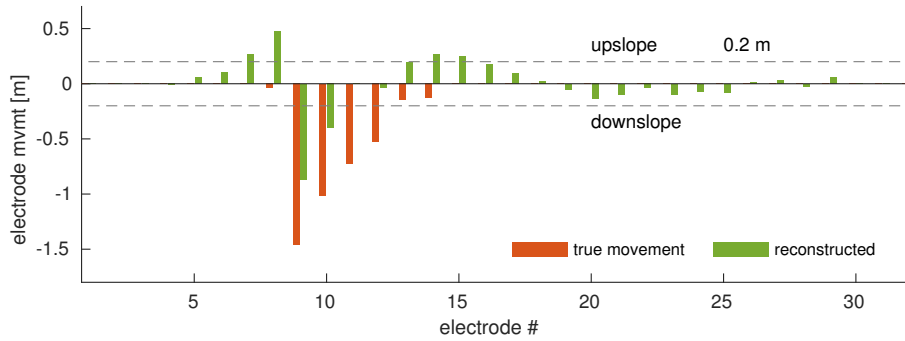
(a) reconstructed longitudinal movments for line#1
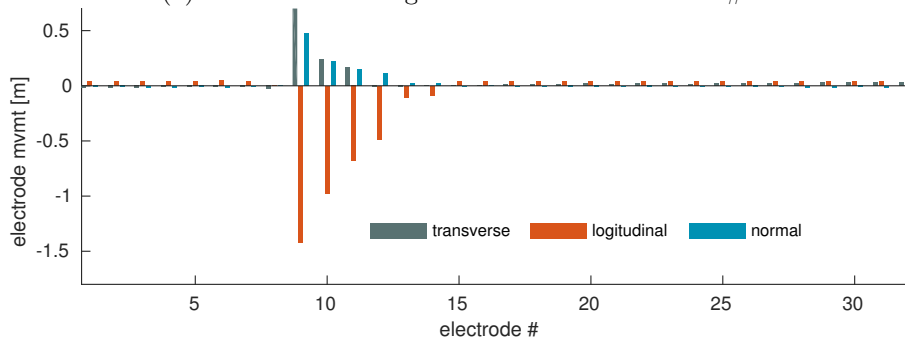
(b) true movements for line#1
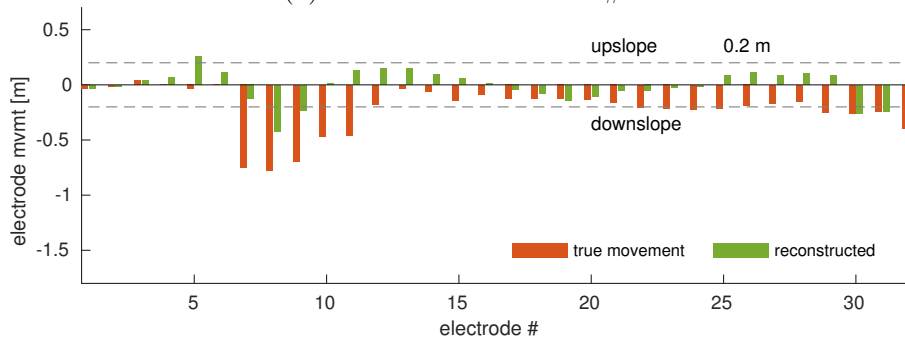
(c) reconstructed longitudinal movments for line#5

(d) true movements for line#5

Figure 8.6: Reconstructed electrode movement as joint resistivity-movement reconstruction, (a,b) true and reconstructed line#1 movement, March 2008 to April 2009, and (c,d) true and reconstructed line#5 movement, February 2013 to February 2014

Electrode movements were initially reconstructed by themselves using an independent implementation. The joint inversion code was checked against this result by setting the movement-resistivity balance parameter β to strongly favour electrode movement. Reconstructions showed no resistivity change and movements that were in close agreement with the movement-only reconstruction code. Small variations still existed between the two results due to differences in the implementation and resulting accuracy of the inexact line search. These variations were small with respect to the overall electrode movement solution. When electrode movements were reconstructed with resistivity changes (Figure 8.6), some portion of the reconstructed electrode movement was lost in favour of reconstructed resistivity change (Figure 8.7).

Due to the large electrode movements (Figure 8.8), it was found helpful to perform a crude version of successive relaxation. The first three iterations of the Gauss-Newton reconstruction were performed with an electrode movement hyperparameter that was an order of magnitude larger than following iterations. Without this adaptation, the reconstructed electrode movements showed poor agreement with measured locations, presumably because the Gauss-Newton iterations were trapped in a local minimum which favoured constructing resistivity artifacts near the electrodes. Exploring the space of hyperparameters near the selected hyperparameter did not reveal one which achieved better electrode movement reconstruction.

(a) line#1, reconstructed resistivity change

(b) line#1, resistivity change using true electrode positions

(c) line#5, reconstructed resistivity change

(d) line#5, resistivity change using true electrode positions
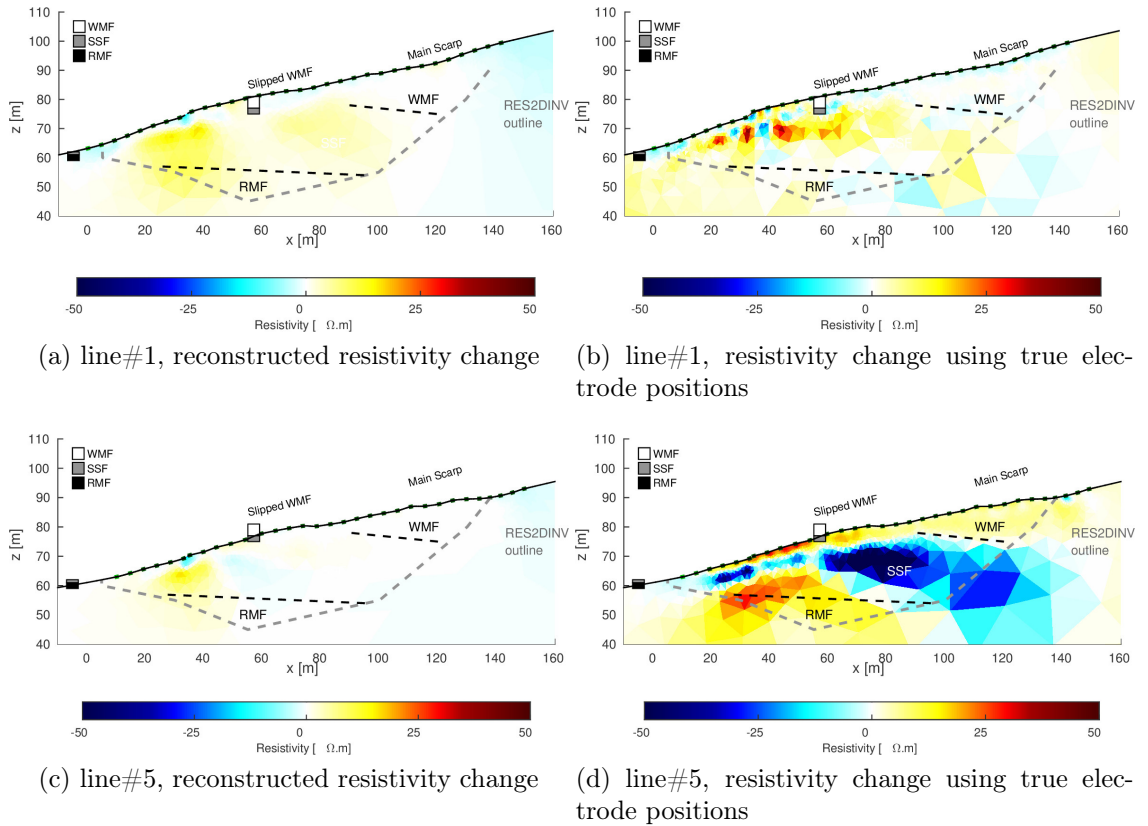
Figure 8.7: Change in resistivity for joint movement reconstructions ($\lambda_\sigma = 0.1, \lambda_x = 0.07$), (a,b) line#1, March 2008 to April 2009, and (c,d) line#5, February 2013 to February 2014
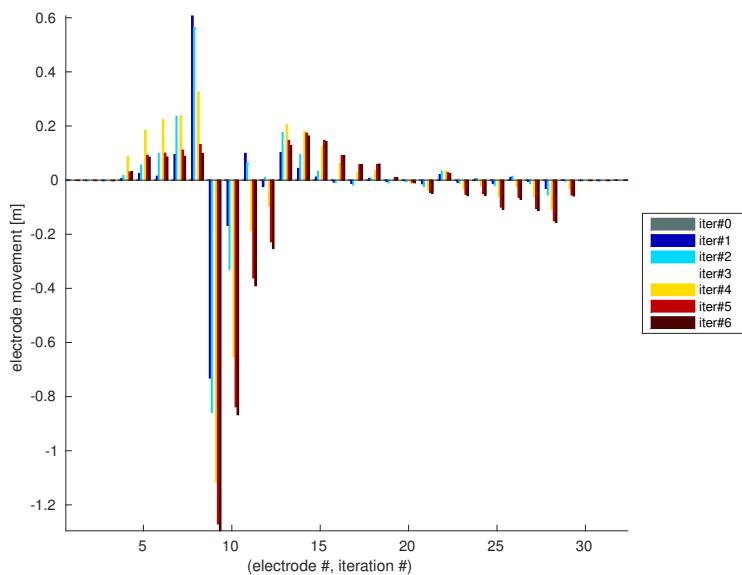
(a) reconstructed longitudinal movmenets for line#1
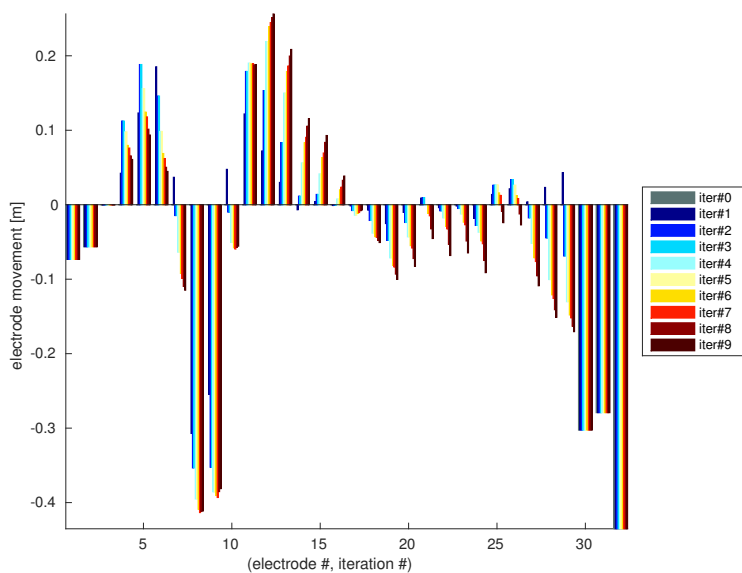
(b) reconstructed longitudinal movmenets for line#5

Figure 8.8: Electrode movement, iterations for (a) line#1, March 2008 to April 2009, and (b) line#5 movement, February 2013 to February 2014

8.6 Discussion

In this chapter, resistivity and electrode displacement were simultaneously reconstructed for a survey located on a slowly moving landslide. Results exhibit some measure of oscillatory artifacts in the reconstructed movement.

For a half-space model with a homogeneous resistivity, electrode positions are not unique. A translation of the entire set of electrodes will give identical measurements. Likewise, a scaling of all electrode positions is equivalent to a change in the homogeneous resistivity. When conductivities are inhomogeneous the electrode locations are somewhat fixed by the locations of the inhomogeneities. Examples of electrode position non-uniqueness manifested itself in this data as large oscillations in the reconstructed electrode movement when no measures were taken to address the issue.

Fixing the location of three electrodes at the upslope and downslope ends of the electrode array nearly eliminated these oscillations. As seen in the line#5 data, this is not necessarily a correct assumption, as both the top and bottom of a landslide may move, leading to resistivity artifacts. We infer that fixing these electrode locations was sufficient to damp the reconstructed movement's oscillations because it fixes the relationship between a stimulus current, a measured voltage, and electrode separation (distance). The physical scale near those electrodes has been fixed. Smoothing-type regularization of conductivity, used in this work, then controls how strictly the selected scaling of electrode separation is enforced. For example, electrodes that have contracted together in a region might cause a conductive artifact to be reconstructed under those electrodes. Increasing the conductivity regularization may suppress these artifacts and cause movement to be reconstructed by contracting the

local electrode spacing to account for smaller than expected voltage measurements in the region.

Both unscaled and log scaled electrode movement were tested and found to give results with similar absolute positional error. We have elected to present the unscaled electrode movement in our reconstructions as it is a less restrictive choice. One could imagine low angle slopes or floodplains where the expected direction of movement would not be known a priori. A further reason to avoid dependence on the log movement scaling is the extension of this work to transverse movements where the restriction to movements only to one side of the array seems inappropriate. In the data sets examined here, there are transverse movements which were caused by material accumulating at the toe of the landslide and towards the edges of the earth flow. These transverse movements can be predicted for this particular data set based on the pre-existing topology: line#1 electrodes moved east, downhill into a gully, while the line#5 electrodes moved west, downhill into the same gully.

Reconstructions generally matched the true electrode locations within 0.20 metres for movements of up to 1.46 metres. Exceptions were line#5 electrodes 12 and 13 where errors of nearly 0.40 m occurred. Adjusting the relationship between resistivity and movement regularization β caused more electrode placement error as resistivity regularization was reduced. These movement magnitudes represented movement of up to 32% of the average 4.73 metre electrode spacing, exceeding the prior joint resistivity-movement method's limitation to movements of approximately 1% of electrode spacing.

8.7 Conclusion and Contributions

This work demonstrated the practical application of a joint movement-resistivity reconstruction using an iterative Gauss-Newton regularized solver. Reconstructions show reasonable agreement with RTK GPS measured electrode locations, available resistivity estimates and geological structure.

The initial reconstructed resistivity, used as a starting point for the electrode movement and resistivity reconstruction, were in relatively close agreement (Figure 8.3 and Figure 8.5). Reconstructed changes in resistivity (Figure 8.7) showed considerable variation, particularly around the region at the toe of the landslide. These changes in resistivity could be indicative of water saturation changes due to water seepage at the toe of the landslide or other geological causes. Another possibility is that the resistivity changes represent artifacts due to transverse and normal components of electrode movement which were not accounted for in these reconstructions.

We note that, in general, even when electrode displacements were not entirely accurate, the spacing between electrodes was generally reduced first in the Gauss-Newton iterations. Errors in electrode spacing were distributed fairly evenly across the electrode array, after which the displacements shifted towards their true positions in most cases. This suggests that a parametrization for electrode movement that encompasses electrode spacing may lead to improved outcomes.

Identification of the Hollin Hill site, geology, field work and electrical measurements were carried out by the British Geological Survey. The model construction, algorithm development, inverse problem parametrization, hyperparameter selection, electrode movement and analysis were the key contributions of this work. Code

for the reconstruction was developed in the EIDORS Matlab environment. Models were meshed using NetGen and then massaged to construct an accurate model with the best possible mesh quality[5]. The absolute iterative Gauss-Newton solver was written from scratch by the author, based on a prior conductivity-only implementation, to support: solution limits through parametrization scaling (§4.8), voltage measurements to apparent resistivity conversion (§4.7), working units conversions (resistivity–conductivity, log scaling), heterogeneous initial conductivities, support for Tikhonov, Laplace and Noser regularization, implementations of a number of line search algorithms, expansion wherever required to add new joint inverse parametrizations (in this case, movement), unit tests and documentation. All simulations and analysis in this chapter were performed by the author except where explicitly noted (§8.4 and Figure 8.3).

Thanks go out to the researchers at the British Geological Survey, Geophysical Tomography Team for generously sharing their time and data.

---

[5]This is not an inverse crime: only electrode positions were matched accurately. Discretizations were selected by the meshing process and clearly did not agree with the continuous domain from which real measurement data were collected. Models used to compare initial and final electrode movement were constructed from different meshes, so did not exhibit overly optimistic discretization error.

Chapter 9

Discussion and Conclusions

This thesis developed new computational techniques to address the issue of maximizing useful information under adverse measurement conditions and built practical implementations that demonstrate the effectiveness of our approach. We have developed methods that reduce artifacts and improve detectability through improved methods for addressing boundary movement. The focus of this work has been on applications to geophysical problems that require iterative non-linear solutions, though the methods are generally applicable to the broader set of impedance imaging techniques.

In this work, we have developed a number of key contributions. We have developed a framework for checking inherent algorithm behaviours for inverse problems and evaluating input model and data quality. This evaluation framework has been applied to a new absolute iterative Gauss-Newton algorithm and its implementation. Methods for evaluating input data errors were demonstrated using data collected at a flooded mine site in Pont-Péan, France. The evaluation framework was applied

to four electrode movement Jacobian techniques, including the implementation of a new Fréchet derivative method. Finally, an impedance imaging technique for simultaneously addressing electrode movement and conductivity was evaluated on a slow-moving landslide in Yorkshire, UK using the infrastructure developed throughout this thesis. These contributions aim to improve the identification and reduce the occurrence of image artifacts which will allow the application of impedance imaging in challenging new environments.

These results have been published in

- A. Boyle, M. Jehl, M. Crabb, and A. Adler. Estimating electrode movement in two dimension. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015a.

- A. Adler, A. Boyle, M. Crabb, H. Gagnon, B. Grychtol, N. Lesparre, and W. R. B. Lionheart. EIDORS Version 3.8. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015.

- A. Boyle, P. Wilkinson, J. Chambers, N. Lesparre, and A. Adler. Slope stability monitoring through impedance imaging. In *15th Conf. Electrical Impedance Tomography*, Gananoque, Canada, April 2014.

Software implementations of the algorithms described in this work have been contributed to the EIDORS project under the GNU General Public License (GPL). Additional software not directly part of the EIDORS project has also been made available[1] and licensed under the GNU GPL, including code for performing the joint conductivity and movement reconstruction.

9.1 Recommendations

Some recommendations coming from this body of work are as follows.

---

[1]https://sourceforge.net/p/eidors3d/code/HEAD/tree/trunk/dev/a_boyle/thesis

- Determining if an algorithm is appropriate and whether an implementation of that algorithm is correct are challenging tasks. We recommend that these be decomposed into smaller units and that a comprehensive testing of these units be undertaken using direct numerical comparisons, analytic models, simplified models, and ultimately real data with a known ground truth.

- Evaluating whether a data set contains the expected data, whether error estimates are "reasonable," and if the data is of sufficient quality are critical to robust, reliable reconstructions. We observe that many data sets lack error estimates on key information. For example, electrodes are frequently assumed to be exactly placed and aligned. We recommend a number of techniques for evaluating data quality: direct examination of the data, comparison to simulated models, and graphical plots that enable an overview of the data. Analysis of sensitivity may be used to evaluate what levels of data error are manageable within a reconstruction framework.

- Four electrode movement Jacobian calculation methods were evaluated in this work. We recommend the rank-one update and Fréchet derivative for their computational efficiency. The Fréchet derivative method comes with caveats regarding contact impedance. Recent work suggests that the Fréchet derivative may enable successful reconstruction of normal movements for which alternative movement Jacobians (including the rank-one update) have not been stable. There are clearly some trade-offs between the two recommended techniques.

- We recommend joint resistivity and electrode movement reconstructions as a promising technique for dealing with electrode movement in ERT data sets.

The $2^1/_2$-D Fourier method was successful in reducing compute effort to solvable levels and we highly recommend it wherever it is geometrically applicable.

9.2 Limitations

Some specific limitations observed within this body of work are as follows.

- The time and cost of evaluating an algorithm, implementation, or data set must be carefully evaluated against the cost of interpretation errors. A thorough evaluation of the methods described in this work is not trivial to undertake. Shortcuts may be appropriate where the return on investment for a thorough validation effort can not be justified. In these cases, some checking is better than none.

- The evaluation framework presented in this work represents a new area of investigation with respect to inverse problems. While this work represents a wealth of "common" and not so common knowledge about best-practices for building robust and reliable reconstructions, the advice presented here is subject to experiential limitations. New data sets, algorithms and implementations may present new challenges not addressed by the recommendations contained in this work. None the less, the underlying philosophy of the presented techniques should be applicable to develop new methods.

- The rank-one update movement Jacobian does not capture the full CEM effects: care must be taken when used with large electrodes. It is not stable under normal movements and should not be used to estimate these types of electrode movement.

- The Fréchet derivative movement Jacobian technique is sensitive to low contact impedances. The correctness of the movement Jacobian should be checked when the contact impedance is of the same order or less than the surrounding conductivity.

- For joint resistivity and electrode movement reconstructions: care must be taken in selecting hyperparameters, and the scale of electrodes needs to be fixed somehow or unstable/oscillatory movement estimates result. Alternative solutions to these two issues may result in more robust reconstructions.

- The Conjugate Gradient and Gauss-Newton solutions were compared throughout this work. In general, the Gauss-Newton solution was found to be roughly equivalent. It is not certain that this is generally true.

9.3 Future Work

There are a number of directions in which the main themes of this work might be further explored:

- 3D reconstructions where electrodes are no longer collinear (grid electrode arrays, for example) require efficient mechanisms to construct models.

  - meshing: The definition and construction of these 3D meshes, including topology and arbitrary surface electrode geometry and numbering could be constructed as an extension of the `mk_geophysics_model` interface, already exhibited in the appendices of this work.

– compute: 3D meshes lead to computational challenges (big meshes, big
matrices) which may in part be addressed with algorithms compatible
with massively parallel compute resources: High Performance Computing
(HPC) and cloud-based scalable infrastructure.

- "Two-and-a-half dimensional" ($2^{1}/_{2}$-D) techniques provide an efficient technique
for addressing many common geological scenarios when conductivity is uniform
and parallel to the surface in a particular direction. Further $2^{1}/_{2}$-D techniques
might be developed:

  – An implementation of the $2^{1}/_{2}$-D movement Jacobian. The technique has
  been developed mathematically in this work but not yet fully implemented
  and tested. Results for the electrode movement reconstruction should be
  improved by its use.

  – A fast matrix inverse update method for calculating the inverse Fourier
  integral $\frac{2}{\pi} \int_{k} (\mathbf{A} + k^2 \mathbf{B})^{-1}$ possibly making use of the SVD since the matrix
  structure is not changing as k varies.

  – Enabling off-plane electrodes, so that an electrode may be placed at any
  point on the surface defined by the projection of the 2D surface into three
  dimensions.

  – Development of a CEM $2^{1}/_{2}$-D formulation for the Fourier dimension.

- The Conjugate Gradient technique should be algorithmically superior: further
evaluation should be undertaken to understand whether it might provide im-
proved solutions for some types of data set and reconstruction. Variations of
the algorithm, for example performing Conjugate Gradient on two disjoint sets

of parameters, may be particularly effective for joint resistivity and electrode movement reconstructions where residuals over the two parameter spaces may exhibit different characteristics.

- Improvements in our ability to take advantage of massively parallel computing resources, where the problem space which we can explore is limited by compute and storage.

- An open hardware platform for EIT which showcases design trade-offs for both biomedical and geophysics applications.

- Incorporating models of real hardware electronics into the inversion process to account for sources of error, possibly by coupling a circuit simulator and FEM/inverse problems implementation.

- The reliability and testing framework suggested in this work would be directly usable in biomedical applications. Better data quality and algorithm implementations could only improve the diagnostic quality of reconstructions.

- The application of joint electrode movement and absolute resistivity reconstructions may have application in biomedical domains with similar geometry, electrode movement issues, strong conductivity contrasts, or a lack of reference measurements. Such applications may be in the areas of breast and prostate cancer screening and brain imaging.

- Further evaluation of the effectiveness of various debug methodologies for inverse problems.

- Development of automated tests to catch the most common survey errors: imprecise placement of electrodes, poor electrode to ground contact, and poor vertical alignment of electrodes relative to the surface. Immediate feedback during surveying would provide an opportunity to correct problematic installations. Corresponding estimates of error might be recorded alongside collected data.

- Development of an automated slope movement detection hardware that would satisfy industrial cost and installation constraints. Evaluation of the reliability of joint conductivity and electrode movement techniques for a large number of (low cost) deployments under varying conditions.

- Addressing issues in the automation and continuous monitoring of remote infrastructure: semi-autonomous and remote systems, data quality management, and robust reconstruction techniques which can reliably detect certain conditions requiring intervention or remediation.

Appendices

Appendix A

Stimulus and Measurement Patterns

In this appendix, we follow the descriptions given in Barker (1979) and Loke and Barker (1996a) for common stimulus and measurement patterns found in ERT surveys. Stimulus and measurement patterns are used to form a set of measurements for impedance imaging data. A stimulus pattern is a sequence of electrodes where current or voltage are applied. A current return path always exists, but it is possible that this path is not through the electrodes due to buried conductors or other unanticipated paths. For ERT and EIT, it is generally assumed that the return path is always through the electrodes so that applied current in equals the return current out. Measurements are taken as the difference between potential across electrodes.

Current I from a single point electrode (a PEM) into a homogeneous half-space causes a radial potential ϕ

$$\phi = \frac{\rho I}{2\pi r} \tag{A.1}$$

for a current I applied to a medium of resistivity ρ, at radial distance r from the
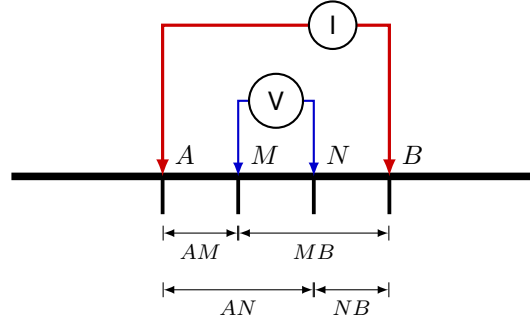
Figure A.1: Generalized pair-wise electrode configuration; for stimulus I and measurement V for four electrodes A M N B

electrode. There is a singularity at each stimulus electrode ($r = 0$) with the PEM.

A dipole AB, as in Figure A.1, causes a potential ϕ measured at M and N

$$\phi_M = \frac{\rho I}{2\pi} \left(\frac{1}{AM} - \frac{1}{MB} \right) \tag{A.2}$$

$$\phi_N = \frac{\rho I}{2\pi} \left(\frac{1}{AN} - \frac{1}{NB} \right) \tag{A.3}$$

at a particular radial distance from the two stimulus electrodes A and B. A potential difference V measured between MN, is then calculated as

$$\Delta\phi = V = \frac{\rho I}{2\pi} \left(\frac{1}{AM} - \frac{1}{MB} - \frac{1}{AN} + \frac{1}{NB} \right) \tag{A.4}$$

where each distance AM, MB, AN, NB is between a stimulus electrode and a measurement electrode. The model may be applied for any arbitrary pair-wise electrode placement.

Equation (A.4) may be rearranged, to find the "apparent resistivity" ρ_a

$$\rho_a = Rk \tag{A.5}$$

$$k = \frac{2\pi}{\left(\frac{1}{AM} - \frac{1}{MB} - \frac{1}{AN} + \frac{1}{NB}\right)} \tag{A.6}$$

where $R = V/I$ is the measured resistivity as the ratio of a measured difference in potentials V resulting from an applied current I, and a geometric factor k which accounts for the electrode configuration. The resistivity is "apparent" because inhomogeneity in the medium, electrodes with finite size, or a surface that is not flat introduce errors with respect to the assumptions implied by the model. None the less, the apparent resistivity is a useful tool for adapting to large variations in measurement magnitudes observed in the field on long collinear electrode arrays. The apparent resistivity normalizes measurements so that they are more equally weighted in the reconstruction under the aforementioned assumptions. The data may still be skewed by strong resistivity contrasts which are often encountered in geoelectrical applications. Reconstructions using log scaled data are sometimes used to rebalance the data weighting for these strong resistivity contrasts. Apparent resistivity is widely used in geophysical applications. In the biomedical community, a similar technique is called "measurement normalization."

In geophysics, standard pair-wise stimulus and measurement patterns are typically used for linear arrays of electrodes in a 32-, 64- or 128-electrode configuration. These patterns are named, based on historical usage, the: Wenner, Schlumberger, and dipole-dipole patterns. Many other variations and alternatives to these patterns have been proposed: we focus here on the most common definitions. Various

monopole configurations may be constructed based on the assumption of a second electrode at infinity. Some biomedical systems employ many electrodes simultaneously (Cook et al., 1994).

A.1 Wenner

A Wenner array has a stimulus dipole straddling a measurement dipole with equidistant electrodes (Figure A.2). For a homogeneous half-space, the apparent resistivity ρ_a for a Wenner array may be calculated using (A.4)

$$\rho_a = 2\pi a R \qquad\qquad (A.7)$$

for electrode spacing a. The geometric factor simplifies to $k = 2\pi a$. As before, the ratio of stimulus current applied I and the measured difference voltage V give the measured resistance R.

An example stimulus and measurement pattern was generated using the Wenner array by increasing a and shifting the starting electrode offset. The multiple n is always one: see the Sclumberger pattern for variations in stimulus versus measurement electrode spacing. The pattern is generally limited to discrete sequences because it
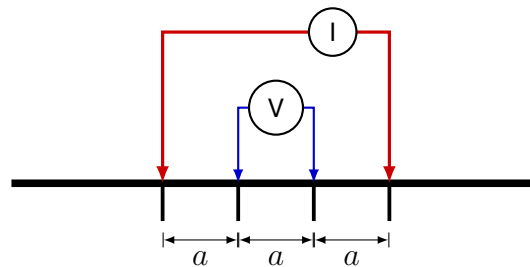
Figure A.2: Wenner electrode configuration

Table A.1: Example Wenner pattern

#	n	a	A	B	M	N	Electrode Placement stim A–B ▬▬ meas M–N
1	1	1	1	4	2	3	
2	1	1	2	5	3	4	
3	1	1	3	6	4	5	
4	1	1	4	7	5	6	
5	1	1	5	8	6	7	
6	1	1	6	9	7	8	
7	1	1	7	10	8	9	
8	1	2	1	7	3	5	
9	1	2	2	8	4	6	
10	1	2	3	9	5	7	
11	1	2	4	10	6	8	
12	1	3	1	10	4	7	

is time consuming and error prone to repeatedly relocate electrodes. An example Wenner stimulus and measurement pattern for ten equally spaced electrodes in a linear array is shown in Table A.1. The electrodes used for each measurement are shown with one measurement per row in the table. The spacing multiple a is shown in the table along with an illustration of the relative placement of the stimulus AB (red) and measurement MN (blue) electrode pairs with electrodes at the ends of the coloured lines.

Variations of the Wenner stimulus and measurement pattern alternate the placement of the stimulus and measurement electrodes. One arrangement matches a dipole-dipole configuration with $n = 1$ (§A.3).
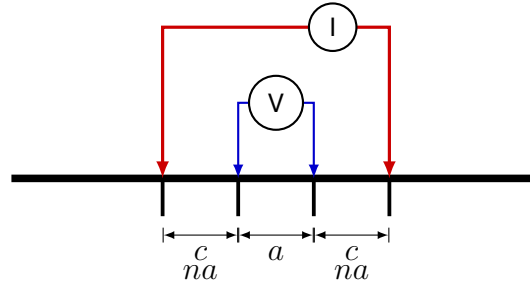
Figure A.3: Schlumberger electrode configuration

A.2 Schlumberger

Like the Wenner array, a Schlumberger array has a stimulus dipole straddling a measurement dipole with electrodes that are symmetric but not necessarily equidistant (Figure A.2). For the Schlumberger electrode configuration, the apparent resistivity ρ_a reduces to

$$\rho_a = \pi \frac{c(c+a)}{a} R \tag{A.8}$$

for measurement electrode spacing a and stimulus electrode separation $a + 2c$ (Figure A.3). The equation (A.8) may be simplified when $c = na$ for linear arrays of equally spaced electrodes

$$\rho_a = \pi n(n+1)aR \tag{A.9}$$

When electrode spacings are equidistant $(n = 1)$, the Schlumberger and Wenner electrode configurations are equivalent.

An example stimulus and measurement pattern was generated using the Schlumberger array by increasing a and n while shifting the starting electrode offset. As

with the Wenner array, the pattern is generally limited to discrete locations because it is time consuming and error prone to repeatedly relocate electrodes. An example Schlumberger stimulus and measurement pattern for ten equally spaced electrodes in a linear array is shown in Table A.2. The electrodes used for each measurement are shown with one measurement per row in the table. The spacing multiple n is shown in the table along with an illustration of the relative placement of the stimulus AB (red) and measurement MN (blue) electrode pairs with electrodes at the ends of the coloured lines.

Table A.2: Example Schlumberger pattern

#	n	a	A	B	M	N	stim A–B ——— meas M–N
			Electrode No.				Electrode Placement
1*	1	1	1	4	2	3	
2*	1	1	2	5	3	4	
3*	1	1	3	6	4	5	
4*	1	1	4	7	5	6	
5*	1	1	5	8	6	7	
6*	1	1	6	9	7	8	
7*	1	1	7	10	8	9	
8*	1	2	1	7	3	5	
9*	1	2	2	8	4	6	
10*	1	2	3	9	5	7	
11*	1	2	4	10	6	8	
12*	1	3	1	10	4	7	
13	2	1	1	6	3	4	
14	2	1	2	7	4	5	
15	2	1	3	8	5	6	
16	2	1	4	9	6	7	
17	2	1	5	10	7	8	
18	3	1	1	8	4	5	
19	3	1	2	9	5	6	
20	3	1	3	10	6	7	
21	4	1	1	10	5	6	

* Measurements #1–#12 are the same as the Wenner pattern, Table A.1.

A.3 Dipole-dipole

The dipole-dipole configuration places the stimulus and measurement electrode pairs adjacent to each other, rather than have the stimulus electrodes straddle the measurement electrodes Figure A.4.

The apparent resistivity calculated based on the general point electrode homogeneous half-space equation gives

$$\rho_a = n(n+1)(n+2)aR \tag{A.10}$$

for measurement electrode spacing a and measurement to stimulus electrode spacing na.

An example stimulus and measurement pattern was generated using the Dipole-Dipole array by increasing a and n while shifting the starting electrode offset. As with the Wenner and Schlumberger patterns, the dipole-dipole pattern is generally limited to discrete positions because it is time consuming and error prone to repeatedly relocate electrodes. An example dipole-dipole stimulus and measurement pattern for ten equally spaced electrodes in a linear array is shown in Table A.3. The electrodes used for each measurement are shown with one measurement per row in the table. The spacing multiple n is shown in the table along with an illustration of the relative
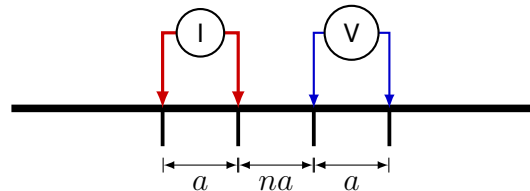
Figure A.4: Dipole-dipole electrode configuration

Table A.3: Example dipole-dipole pattern

#	n	a	A	B	M	N	stim A–B ▬▬ meas M–N
			Electrode No.				Electrode Placement
1	1	1	1	2	3	4	
2	1	1	2	3	4	5	
3	1	1	3	4	5	6	
4	1	1	4	5	6	7	
5	1	1	5	6	7	8	
6	1	1	6	7	8	9	
7	1	1	7	8	9	10	
8	1	2	1	3	5	7	
9	1	2	2	4	6	8	
10	1	2	3	5	7	9	
11	1	2	4	6	8	10	
12	1	3	1	4	7	10	
13	2	1	1	2	4	5	
14	2	1	2	3	5	6	
15	2	1	3	4	6	7	
16	2	1	4	5	7	8	
17	2	1	5	6	8	9	
18	2	1	6	7	9	10	
19	2	2	1	3	7	9	
20	2	2	2	4	8	10	

placement of the stimulus AB (red) and measurement MN (blue) electrode pairs with electrodes at the ends of the coloured lines.

Appendix B

From Maxwell's Equations to EIT

The equation $\nabla \cdot \sigma \nabla \phi = 0$ for conductivity σ and voltage potential ϕ defining the impedance imaging problem may be derived from Maxwell's equations (Somersalo et al., 1992; Holder, 2005). The problem of determining the interior conductivity based on boundary information is commonly called "Calderon's Problem" in the Inverse Problems literature (Calderón, 1980).

Starting from Maxwell's equations

$$\nabla \cdot \mathbf{D} = \rho \qquad \text{Gauss' Law} \qquad (\text{B.1})$$

$$\nabla \cdot \mathbf{B} = 0 \qquad \text{Gauss' Law for Electromagnetism} \qquad (\text{B.2})$$

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B} \qquad \text{Faraday's Law} \qquad (\text{B.3})$$

$$\nabla \times \mathbf{H} = \bar{\mathbf{J}} + \partial_t \mathbf{D} \qquad \text{Ampere's Law} \qquad (\text{B.4})$$

for charge density ρ, and using the notation for partial derivatives with respect to time $\partial/\partial t = \partial_t$, the vector cross product $\nabla \times \square$, and the vector divergence $\nabla \cdot \square$.

Maxwell's equations are coupled by constitutive relationships: electric field $\mathbf{E}$ and current density $\bar{\mathbf{J}}$ are related by conductivity σ, electric field $\mathbf{E}$ and electric flux density $\mathbf{D}$ are related by permittivity ϵ, and magnetic field $\mathbf{H}$ and magnetic flux density $\mathbf{B}$ are related by permeability μ

$$\bar{\mathbf{J}} = \sigma \mathbf{E} \qquad \text{Ohm's Law} \tag{B.5}$$

$$\mathbf{D} = \epsilon \mathbf{E} \qquad\qquad \mathbf{H} = \frac{1}{\mu}\mathbf{B} \tag{B.6}$$

for a linear, iostropic, nondispersive body. For a DC or quasi-static system, the electric and magnetic flux density does not vary significantly over time

$$\partial_t \mathbf{D} = 0 \qquad\qquad \partial_t \mathbf{B} = 0 \tag{B.7}$$

so that (B.3) and (B.4) become

$$\nabla \times \mathbf{E} = 0 \tag{B.8}$$

$$\nabla \times \mathbf{H} = \bar{\mathbf{J}} \tag{B.9}$$

From the Helmholtz decomposition, we know that a sufficiently smooth field $\mathbf{F}$ that decays towards infinity is fully specified by an irrotational (curl-free) vector field $\nabla \mathbf{u}$ and a solenoidal (divergence-free) vector field $\nabla \times \mathbf{A}$. An irrotational vector field has a scalar potential $\mathbf{u}$, while a solenoidal vector field has a vector potential $\mathbf{A}$.

$$\mathbf{F} = -\nabla \mathbf{u} + \nabla \times \mathbf{A} \tag{B.10}$$

We restrict ourselves to a scenario where transient signals are negligible by waiting for a "settling" time before making measurements and by applying low frequencies where magnetic effects are also negligible. We define low frequencies based on the size of the domain and the wavelength of propagating signals so that errors introduced by waves propagating to distant measurement electrodes are negligible. We are also restricted to frequencies below where "skin depth" does not affect the current flow in the medium. These two restrictions, introduced by our low frequency assumption, allow us to assume that the electric field is quasi-static: a low frequency system (Adler et al., 2011, §14.1.1).

For an electric field that is quasi-static and therefore irrotational (B.8), the electric field is fully specified by the scalar potential ϕ so that

$$\mathbf{E} = -\nabla\phi \tag{B.11}$$

The divergence of curl identity $\nabla \cdot (\nabla \times \square) = 0$ for magnetic fields $\mathbf{H}$ combined with the quasi-static assumption (B.9) and Ohm's Law gives

$$\nabla \cdot (\nabla \times \mathbf{H}) = 0 \tag{B.12}$$

$$\nabla \cdot \bar{\mathbf{J}} = 0 \tag{B.13}$$

$$\nabla \cdot (\sigma \mathbf{E}) = 0 \tag{B.14}$$

$$-\nabla \cdot \sigma \nabla\phi = 0 \tag{B.15}$$

which defines the DC behaviour of current flow within an impedance distribution.

An alternate solution can be found (Somersalo et al., 1992) by looking for time harmonic solutions and assuming very small magnetic permeability μ. These so-

lutions allow conductivty to be complex valued such that $\sigma \rightarrow \sigma + i\omega\epsilon$. Boundary conditions are typically specified using the CEM (Section §2.1) to account for contact impedance, shunting effects and descretization of the boundary.

Appendix C

The FEM Approximation

The FEM is a technique for numerically approximating the value of a function over a domain (Hughes, 2000). It is used in structural mechanics, electromagnetics, fluid dynamics and many other fields (Zienkiewicz et al., 2005). Following (Rücker and Günther, 2011), a PDE boundary value problem $\mathcal{F}u = f$, with boundary conditions

$$\mathcal{F}u = f \qquad \text{in } \Omega \quad \text{PDE} \tag{C.1}$$

$$u = g_0 \qquad \text{on } \partial\Omega_0 \quad \text{Dirichlet boundary condition} \tag{C.2}$$

$$\nabla u \cdot \hat{n} = g_1 \qquad \text{on } \partial\Omega_1 \quad \text{Neumann boundary condition} \tag{C.3}$$

$$\nabla u \cdot \hat{n} + \alpha u = g_0 \qquad \text{on } \partial\Omega_2 \quad \text{Robin boundary condition} \tag{C.4}$$

is converted to a weighted residual formulation by the following steps.

1. Multiply the residual of the PDE by a weighting function w that vanishes at the Dirichlet boundary $\partial\Omega_0$, where the combined residual and weighting function is to approach zero.

2. Integrate by parts (Green's first identity) and apply Neumann and Robin boundary conditions to arrive at the so-called "weak form."

3. Represent the approximate solution $u_h \simeq u$ as a linear combination of polynomial basis functions N_i defined on a particular mesh triangulation.

4. Select basis functions defined over individual elements of the mesh N_i and zero elsewhere with nodal values u_i defining their "control points." Choose to use these same basis functions for the discretized weighting function w_i, the so-called "Galerkin approach."

5. Solve the linear algebraic system for the nodal values u_i.

Applying the FEM to EIT we have a boundary value problem with CEM boundary conditions for voltage U_ℓ and current I_ℓ at the ℓ-th electrode E_ℓ with contact impedance z_ℓ. There is no current flow normal $\hat{n}$ to the boundary $\partial\Omega$ except at the electrodes

$$\nabla \cdot (\sigma \nabla \phi) = 0 \quad \text{in} \quad \Omega, \quad \sigma > 0 \tag{C.5}$$

$$\int_{E_\ell} \sigma \frac{\partial \phi}{\partial \hat{n}} dE_\ell = I_\ell \quad \ell = 1..L \tag{C.6}$$

$$\phi + z_\ell \sigma \frac{\partial \phi}{\partial \hat{n}} = U_\ell \quad \text{on} \quad E_\ell, \quad \ell = 1..M \tag{C.7}$$

$$\frac{\partial \phi}{\partial \hat{n}} = 0 \quad \text{on} \quad \partial\Omega \setminus \bigcup_\ell E_\ell \tag{C.8}$$

where the divergence $\nabla \cdot \square$ of the gradient of potential $\nabla \phi$ scaled by an isotropic conductivity σ is zero in the interior of a domain.

We make use of Green's first identity by rearranging

$$\int_\Omega (u\nabla^2 v + \nabla u \cdot \nabla v) = \oint_{\partial\Omega} u(\nabla v \cdot \hat{n}) \tag{C.9}$$

$$\int_\Omega u\nabla^2 v = -\int_\Omega \nabla u \cdot \nabla v + \oint_{\partial\Omega} u(\nabla v \cdot \hat{n}) \tag{C.10}$$

to agree with our Laplace equation on the interior, assuming a constant conductivity

$$\nabla \cdot (\sigma\nabla\phi) = \sigma\nabla^2\phi \tag{C.11}$$

We then multiply by the weighting function and integrate by parts using Green's first identity to get

$$-\int_\Omega \sigma\nabla w \cdot \nabla\phi + \oint_{\partial\Omega} \sigma w(\nabla\phi \cdot \hat{n}) = 0 \tag{C.12}$$

Discretizing over an approximation of the potential distribution, defined at nodes of the mesh ϕ_i and between by the linear combinations of basis functions N_i, with weighting function selected to be the same basis functions N_j gives a linear system of equations

$$\sum_i \int_\Omega \sigma\phi_i \nabla N_i \cdot \nabla N_j = \sum_i \oint_{\partial\Omega} \sigma\phi_i(\nabla N_i \cdot \hat{n}) \tag{C.13}$$

for columns i and rows j.

These terms may be concisely described for the system matrix $\mathbf{A}$ with block-wise matrix notation where additional boundary condition terms introduced by the CEM

are incorporated

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{C.14}$$

$$\begin{bmatrix} \mathbf{G} + \mathbf{B} & \mathbf{C} \\ \mathbf{C}^{\mathsf{T}} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \phi \\ \vec{\phi} \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{\mathbf{b}} \end{bmatrix} \tag{C.15}$$

where column vectors ϕ are potential at each node, $\vec{\phi}$ are the measurements at the electrodes, and $\vec{\mathbf{b}}$ are applied currents at the electrodes.

The remaining matrices $\mathbf{G}$ $\mathbf{B}$ $\mathbf{C}$ $\mathbf{D}$ are then defined in terms of

$$G_{ij} = \int_{\Omega} \sigma \nabla N_i \cdot \nabla N_j \tag{C.16}$$

$$B_{ij} = \sum_{\ell=1}^{L} \int_{E_\ell} N_i N_j \tag{C.17}$$

$$C_{\ell j} = -\frac{1}{z_\ell} \int_{E_\ell} N_j \tag{C.18}$$

$$D_{ij} = \frac{|E_\ell|}{z_\ell} \quad \text{for } i = j = \ell \text{ and is } 0 \text{ otherwise} \tag{C.19}$$

$$b_j = I_\ell \tag{C.20}$$

for shape function N, element conductivity σ, electrode contact impedance z_ℓ, and electrode surface area $|E_\ell|$. The b_j terms cancel everywhere except at the domain's edges with boundary currents $\mathbf{\bar{J}}$ normal $\hat{n}$ to the boundary $\partial\Omega$.

In practice, many stimulus patterns are concatenated

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \vec{\mathbf{b}}_1 & \vec{\mathbf{b}}_2 & \vec{\mathbf{b}}_3 & & \vec{\mathbf{b}}_n \end{bmatrix} \tag{C.21}$$

to construct an intermediate set of voltage distributions

$$\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \ldots \ \mathbf{x}_n] \tag{C.22}$$

$$\mathbf{AX} = \mathbf{B} \tag{C.23}$$

and eventually a set of measurements $\mathbf{u}$ are calculated as a difference in electrode potentials

$$\mathbf{u} = \mathcal{T}\mathbf{X} = \mathcal{T}\mathbf{A}^{-1}\mathbf{B} \tag{C.24}$$

using the system matrix $\mathbf{A}$. The difference function $\mathcal{T}$ typically differs for each stimulus (the measurements differ) but the matrix inverse of the system $\mathbf{A}^{-1}$ may be reused repeatedly to efficiently calculate new nodal voltages. The measurements $\mathbf{u}$ are typically stored as a single concatenated vector, though they relate to many different stimulus patterns.

For e elements and n nodes, the n row $\times$ n column system matrix $\mathbf{G}$, ignoring the electrode model for the moment, may be constructed

$$\mathbf{G} = \mathbf{C}^{\mathsf{T}}\mathbf{SDC} \qquad \text{for } \mathbf{S}^{(e)} = \nabla N_i^{(e)} \cdot \nabla N_j^{(e)} \text{ and } \mathbf{D}^{(e)} = \mathrm{diag}(\sigma_e) \tag{C.25}$$

from a connectivity matrix $\mathbf{C}$ with element shape functions $\mathbf{S}$ and element conductivities $\mathbf{D}$. The connectivity matrix converts local element nodes to global nodes

$$\mathbf{C} = \begin{cases} 1 & \text{column } j\text{'s global node number is} \\ & \quad \text{assigned to row } i\text{'s element local node} \\ 0 & \text{otherwise} \end{cases} \qquad (\text{C}.26)$$

Taking the matrix square root of the shape matrix allows the connectivity and shape matrices to be combined for efficient multiplications by the conductivity because the left and right multiplication are symmetric

$$\mathbf{G} = \mathbf{C}^{\mathsf{T}}\mathbf{S}^{1/2\mathsf{T}}\mathbf{S}^{1/2\mathsf{T}}\mathbf{D}\mathbf{C} = \mathbf{C}^{\mathsf{T}}\mathbf{S}^{1/2\mathsf{T}}\mathbf{D}\mathbf{S}^{1/2}\mathbf{C} \qquad \text{for } \mathbf{S} = \mathbf{S}^{1/2\mathsf{T}}\mathbf{S}^{1/2\mathsf{T}} \qquad (\text{C}.27)$$

$$\mathbf{G} = \mathbf{F}^{\mathsf{T}}\mathbf{D}\mathbf{F} \qquad \text{with } \mathbf{F} = \mathbf{S}^{1/2}\mathbf{C} \qquad (\text{C}.28)$$

where the shape functions are originally constructed from matrix squares so that the matrix square root is trivially known.

The $em \times n$ connectivity matrix $\mathbf{C}$, for elements with m local nodes, is a sparse sign matrix where all entries are 0 or ± 1. The connectivity matrix merges the e element-wise block matrices by constraining collocated nodes of adjacent elements to the same voltage. The matrix is applied both row-wise $\mathbf{C}$ and column-wise $\mathbf{C}^{\mathsf{T}}$ to collapse an $em \times em$ sized shape and conductivity matrix $\mathbf{SD}$ in local nodal coordinates down to an $n \times n$ $\mathbf{G}$ matrix in global nodal coordinates.

For an FEM mesh using a consistent element type throughout, the shape functions are formed of a common $m \times m$ shape function $\mathbf{S}^{(e)}$ and a deformation and scaling that reshapes the canonical FEM element into the shape required by a particular arrangement of nodes for each element. The completed $em \times em$ shape matrix

S forms a block-diagonal matrix

$$
\mathbf{S} = \begin{bmatrix} \mathbf{S}^{(1)} & 0 & \cdots & & 0 \\ 0 & \mathbf{S}^{(2)} & & & \vdots \\ \vdots & & \ddots & & \\ 0 & \cdots & 0 & \mathbf{S}^{(e)} \end{bmatrix}
\tag{C.29}
$$

with bandwidth matching the nodes per element m. The $em \times em$ conductivity matrix **D** is a diagonal matrix where the conductivity is repeated m times on the diagonal, once for each node of an element,

$$
\mathbf{D} = \mathrm{diag}(\boldsymbol{\sigma}) \otimes \mathbf{I}_m
\tag{C.30}
$$

where the Kronecker product $\otimes$ takes the diagonalized matrix of e element conductivities vector $\boldsymbol{\sigma}$ and combines with the $m \times m$ identity matrix $\mathbf{I}_m$. When the shape function is factored into its matrix square root $\mathbf{S}^{1/2}$, the size of the conductivity matrix is reduced $\mathbf{D}_m \rightarrow \mathbf{D}_{m-1}$ from $em \times em$ to $e(m-1) \times e(m-1)$. The shape functions $\mathbf{S}^{(e)}$ capture the structure of the elements and are of primary interest in evaluating the effect of an electrode movement, while the other **C** and **D** matrices remain constant.
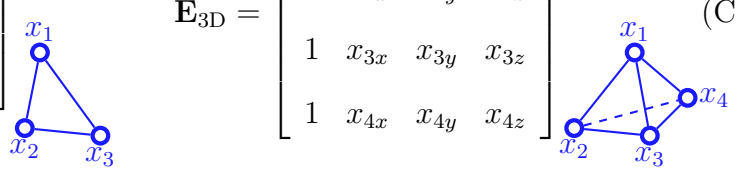
In general, the element integrals are numerically integrated to approximate their sums on arbitrary shape functions. For first-order interpolatory nodal (P1) shape functions (Arnold and Logg, 2014) with piecewise constant conductivity, an analytic solution exists. The gradients of the shape functions are constrained by the geometry

of the element, so that the "stiffness matrix" $\mathbf{S}^{(e)}$ is defined as

$$\mathbf{S}^{(e)} = \frac{1}{n_D!} \frac{1}{|\det \mathbf{E}|} \mathbf{E}_{\backslash 1}^{\mathsf{T}} \mathbf{E}_{\backslash 1} \tag{C.31}$$

$$\mathbf{S}^{(e)1/2} = \mathbf{E}_{\backslash 1} \sqrt{\frac{1}{n_D!} \frac{1}{|\det \mathbf{E}|}} \tag{C.32}$$

for an n_D-dimensional element, where $\mathbf{E}$ works out to be the matrix inverse of the node positions and $\mathbf{E}_{\backslash 1}$ is the same $\mathbf{E}$ matrix but with the top row removed, giving an $m - 1 \times m$ matrix. The inverse of the determinant $1/\det(\mathbf{E})$ gives twice the area of the element but may be positive or negative depending on the order of the nodes. In two and three dimensions, the $\mathbf{E}$ matrix is

$$\mathbf{E}_{2D} = \begin{bmatrix} 1 & x_{1x} & x_{1y} \\ 1 & x_{2x} & x_{2y} \\ 1 & x_{3x} & x_{3y} \end{bmatrix}^{-1} \qquad \mathbf{E}_{3D} = \begin{bmatrix} 1 & x_{1x} & x_{1y} & x_{1z} \\ 1 & x_{2x} & x_{2y} & x_{2z} \\ 1 & x_{3x} & x_{3y} & x_{3z} \\ 1 & x_{4x} & x_{4y} & x_{4z} \end{bmatrix}^{-1} \tag{C.33}$$

where the subscripts of each row denote the n-th node location x_n of a specific triangular or tetrahedral P1 element. Illustrations of a triangular and tetrahedral P1 element (blue) are adjacent to their respective equations, showing node locations x_n with respect to an example element's geometry.

The conductivity Jacobian for element k may be calculated using the identity for the partial derivative of a matrix inverse

$$\frac{\partial \mathbf{X}^{-1}}{\partial u} = -\mathbf{X}^{-1} \frac{\partial \mathbf{X}}{\partial u} \mathbf{X}^{-1} \tag{C.34}$$

so that our FEM mesh node voltages vary with the k-th element's conductivity σ as

$$\mathbf{J}_\sigma = \frac{\partial \mathbf{v}}{\partial \sigma_k} = \mathcal{T}\frac{\partial \mathbf{A}^{-1}}{\partial \sigma_k}\mathbf{B} \tag{C.35}$$

$$= -\mathcal{T}\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial \sigma_k}\underbrace{\mathbf{A}^{-1}\mathbf{B}}_{\mathbf{X}} \tag{C.36}$$

$$\mathbf{J}_\sigma = -\mathcal{T}\mathbf{A}^{-1}\mathbf{F}^{\mathsf{T}}\frac{\partial \mathbf{D}}{\partial \sigma_k}\mathbf{F}\mathbf{X} \tag{C.37}$$

where a difference $\mathcal{T}$ between nodes in a column of the voltage matrix $\mathbf{X}$ between a pair of electrodes gives the change in a measurement when conductivity k is perturbed (Yorkey et al., 1987; Adler and Guardo, 1996). Note that the matrix partial derivative of conductivity $\partial \mathbf{D}/\partial \sigma_k$ is zero everywhere but within the element k.

Appendix D

2 1/2-Dimensional Solutions

Two and a half-dimensional solutions are forward problems that are solved in three dimensions under the assumption that one dimension has uniform conductivity from positive to negative infinity. In real impedance imaging systems the electrodes are not of infinite length in a particular direction, even when the conductivity distribution is uniform in that direction. The finite size of the electrodes immediately invalidates attempts to correctly approximate three-dimensional solutions with a two-dimensional model. The two and a half-dimensional technique applies a Fourier transform in one-dimension and then integrates over the spatial frequencies to obtain a solution, while simultaneously solving with an FEM for the other two dimensions.

Following Dey and Morrison (1979), we show the derivation of the 2$1/2$-D method. By subscripts, we denote the dimensionality of our variables and partial derivatives, so that a three-dimensional potential ϕ_{xyz} is caused by current at the boundary applied to a three-dimensional conductivity distribution σ_{xyz}. If the conductivity is

constant in one dimension z, the partial derivative is zero

$$\frac{\partial \sigma_{xyz}}{\partial z} = 0 \qquad \text{for constant } \sigma_z \qquad \text{(D.1)}$$

and σ_{xyz} may now be denoted σ_{xy}. For such a conductivity, the potential may still vary in the z-dimension, depending on whether current density is applied uniformly in z. Current density is not generally uniform, given the finite size of electrodes used to apply current.

Given a constant conductivity in the z-dimension, the z-dependence of the potential ϕ_{xyz} may be Fourier transformed into the spatial frequency domain $\tilde{\phi}_{xy\tilde{k}}$ using the cosine transform and its inverse

$$\tilde{\phi}_{xy\tilde{k}} = \int_0^\infty \phi_{xyz} \cos(\tilde{k}z) dz \qquad \text{(D.2)}$$

$$\phi_{xyz} = \frac{2}{\pi} \int_0^\infty \tilde{\phi}_{xy\tilde{k}} \cos(\tilde{k}z) d\tilde{k} \qquad \text{(D.3)}$$

for a potential that is reflected across the xy-plane so that the potential is an even function $\phi(z) = \phi(-z)$.

Returning to the conductivity-potential relationship, for uniform conductivity in the z-dimension and assuming a PEM,

$$-\nabla \cdot (\sigma_{xy} \nabla \phi_{xyz}) = \frac{\partial \rho}{\partial t} \delta_{xyz} \qquad \text{(D.4)}$$

Applying the Fourier transform leads to

$$-\nabla \cdot (\sigma_{xy} \nabla \tilde{\phi}_{xy\tilde{k}}) + \tilde{k}^2 \sigma_{xy} \tilde{\phi}_{xy\tilde{k}} = \tilde{Q} \delta_{xy} \qquad \text{(D.5)}$$

for a fixed wave number $\tilde{k}$ and steady state current density in the spatial frequency domain

$$\tilde{Q}\delta_{xy} = \frac{I\partial\rho}{2\partial t}\delta_{xy} \simeq \frac{I}{2A} \tag{D.6}$$

where Dey and Morrison (1979) have approximated the electrode as a constant current patch, a shunt electrode model conducting current I over an electrode area A. We would prefer the more accurate boundary normal current over the electrode modelled by the CEM. In the construction presented here, we get a shunt model in the z-dimension and CEM in the xy-dimensions.

We note that (D.5) takes the same general form as the original equation (D.4), albeit in the spatial frequency domain in place of the z-dimension, with an additional dissipation term dependent on the square of the spatial wave number $\tilde{k}^2$.

The outcome is that the discretized FEM solution has an additional term applied to the FEM stiffness matrix.

$$\mathbf{G} = \mathbf{C}^\mathsf{T}\mathbf{S}\mathbf{D}\mathbf{C} \tag{D.7}$$

$$\mathbf{S} \quad \rightarrow \quad \mathbf{S} + \tilde{k}^2\mathbf{T} \qquad \text{such that} \tag{D.8}$$

$$\mathbf{G}_{\tilde{k}} = \mathbf{C}^\mathsf{T}(\mathbf{S} + \tilde{k}^2\mathbf{T})\mathbf{D}\mathbf{C} = \mathbf{C}^\mathsf{T}\mathbf{S}\mathbf{D}\mathbf{C} + \tilde{k}^2\mathbf{C}^\mathsf{T}\mathbf{T}\mathbf{D}\mathbf{C} \qquad \text{where} \tag{D.9}$$

$$\mathbf{S}^{(e)} = \frac{1}{2|\det\mathbf{E}|}\mathbf{E}_{\backslash 1}^\mathsf{T}\mathbf{E}_{\backslash 1} \qquad \text{and} \tag{D.10}$$

$$\mathbf{T}^{(e)} = \frac{1}{2|\det\mathbf{E}|}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\frac{1}{12} \tag{D.11}$$

where the core of the system matrix $\mathbf{G}$ is constructed from a connectivity matrix $\mathbf{C}$, a conductivity matrix $\mathbf{D}$ and an element shape matrix $\mathbf{S}$. In the $2\frac{1}{2}$-D formulation, the shape matrix gains an additional term $\mathbf{T}$ multiplied by the square of the spatial wave number $\tilde{k}$. The matrix $\mathbf{T}$ is an integral over the shape function of the form $\int_\Omega N_i N_j$ and depends solely on the area of the triangle. As with the shape matrix $\mathbf{S}$ the additional matrix $\mathbf{T}$ can be split into its matrix square root

$$\mathbf{G}_{\tilde{k}} = \mathbf{F_S^T D F_S} + \tilde{k}^2 \mathbf{F_T^T D F_T} \tag{D.12}$$

and converted into a symmetric multiplier of $\mathbf{D}$ to improve matrix assembly efficiency. To assemble matrices for many $\tilde{k}$ only the value of $\tilde{k}$ changes, so the matrix multiplications are a one-time cost.

The potentials found for the forward solution must have the inverse Fourier transform applied after inverting the system matrix.

$$\phi = \frac{2}{\pi} \int_{\tilde{k}} \mathcal{T} \tilde{A}_{\tilde{k}}^{-1} \mathbf{B} \tag{D.13}$$

An adaptive quadrature numerical integration[1] is then typically used to accumulate the inverted solutions at appropriate $\tilde{k}$. The first solution $\tilde{k} = 0$, is the two-dimensional solution. For increasing $\tilde{k}$, the potential is monotonically decreasing as $1/\tilde{k}^2$ for $\tilde{k} > 1$. For a sufficient summation of $\tilde{k}$, the solution will converge to the three-dimensional solution.

The difference between the two-dimensional and the three-dimensional or $2\frac{1}{2}$-D solutions can be significant as illustrated in Figure D.1 where widely separated

---

[1] `trapz`, `quadv`, or `integral` in Matlab

electrodes exhibit simulated measurements in error by as much as 71.5 times their true value. The 2D and $2\frac{1}{2}$-D simulations use the same 2D model (Figure E.3a), while the 3D simulations use a 3D model (Figure E.3b). Both models are 16 electrode CEM half-space models, sharing the same linear electrode arrangement (5 m spacing, 0.1 m diameter), with homogeneous conductivity ($\sigma = 1$). The analytic solution uses the 2D model geometry and estimates PEM electrodes at the centre of the CEM electrode positions and a homogeneous conductivity at the mean value of the original model which, in this case, is already homogeneous.

The agreement between the analytic, $2\frac{1}{2}$-D and 3D solutions illustrates that the FEM models extend far enough to approximate a half-space without introducing significant truncation errors. The difference between the analytic, $2\frac{1}{2}$-D, and the 3D solutions are likely due to differences in the modelled shape of the electrodes or insufficient mesh refinement. We do not explore the source of the analytic versus 3D error further here, though in principle it is straightforward to eliminate the possibilities such as PEM versus CEM, mesh density or electrode shape.
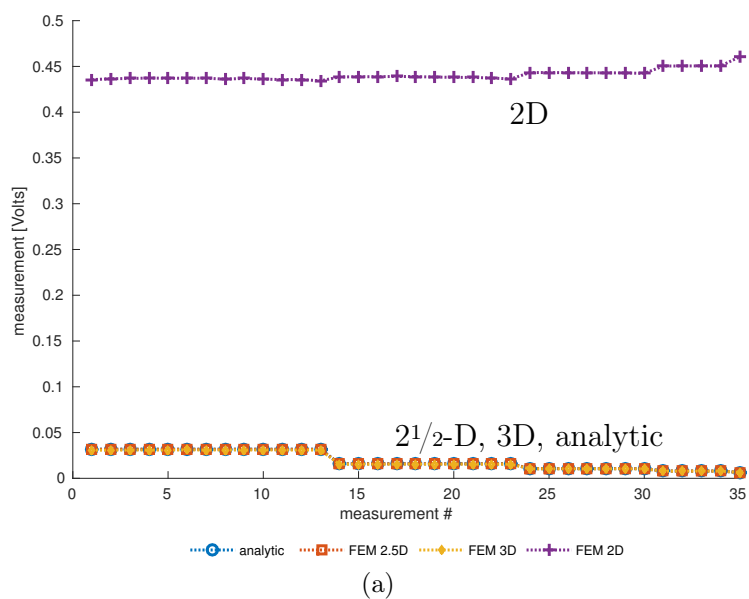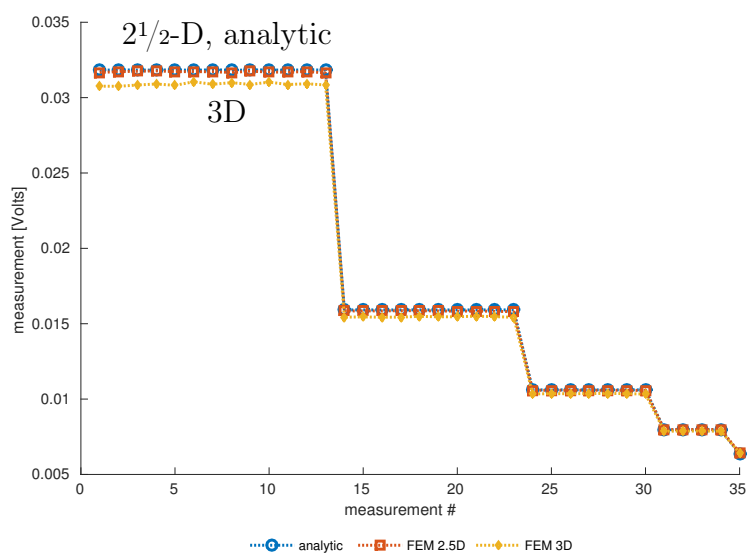
Figure D.1: 2D versus 3D measurements; a 16 CEM electrode half-space model with Wenner stimulus pattern, models for 2D and 2^1/$_2$-D used the model illustrated in Figure E.3a, 3D measurements used the model shown in Figure E.3b, and the analytic model used geometry from the 2D model, (a) 2D measurements are significantly different than 3D measurements and half-space (PEM) analytic model, while (b) 3D, 2^1/$_2$-D and analytic model are in close agreement

Appendix E

Inverse Solver in EIDORS

An implementation of the iterative Gauss-Newton algorithm described in this work has been contributed to the upcoming EIDORS release 3.9 as `inv_solve_gn` and `inv_solve_cg`. The implementation supports:

- Parameter units of conductivity, log conductivity, $\log_{10}$ conductivity, resistivity, log conductivity, $\log_{10}$ conductivity, as input, output and working units with automatic conversion between units.

- Joint inversion of conductivity/resistivity and electrode movement. In principle, joint inversion of multiple parameter types if supporting Jacobian and forward models are available.

- Input units of voltage, apparent resistivity, log apparent resistivity and $\log_{10}$ apparent resistivity.

- Support for dual models, for example a three-dimensional forward model and two-dimensional inverse model using a provided mapping function.

- The ability to solve absolute (single set of measurements) and difference (two sets of measurements) data iteratively.

A function to generate models, `mk_geophysics_model`, has also been built for building models in half-space and nearly half-space geometries, as found in many common geophysics problems. The function supports generating two-dimensional and three-dimensional models with control of electrode type (CEM or PEM), mesh density, and number of electrodes, and optional electrode placement. If electrode locations are provided, a mesh conforming to surface topology based on those electrode's elevations is constructed. Dual meshes (2D/3D models) and $2\frac{1}{2}$-D models are also supported. Fine grained control of all model parameters is possible by additional configuration options. Most meshes were generated with NetGen.

Detailed help is available for each function:

```
help inv_solve_gn; % wrapper for GN, uses inv_solve_core
help inv_solve_cg; % wrapper for CG, uses inv_solve_core
help inv_solve_core;
help mk_geophysics_model;
```

A minimal working example for a linear 32 CEM electrode array, with a simulated inhomogeneity and solved using the Gauss-Newton iterative solver in $2\frac{1}{2}$-D is shown in the following Matlab code. By default, EIDORS operates in units of conductivity. In the interests of brevity, some lines of code have been concatenated. The simulated data and the reconstruction were performed on different meshes and 40dB SNR was simulated to avoid the inverse crime where unreasonably optimistic results may be achieved through "cheating" in simulation studies (Kaipio and Somersalo, 2007). The forward model and the reconstruction show reasonable agreement after six iterations (Figure E.1).

```matlab
% --- simulated data ---
imdl = mk_geophysics_model('h2p5b', 32); % 2.5D half-space model
imdl.fwd_model.stimulation = stim_pattern_geophys(32, 'Wenner'); % Wenner stimulus
imgf = mk_image(imdl.fwd_model, 30); % homogeneous model, 30 Ohm.m
ctrs = interp_mesh(imdl.fwd_model); x = ctrs(:,1); y= ctrs(:,2);
r1 = sqrt((x-40).^2 + (y+20).^2); r2 = sqrt((x-120).^2 + (y+45).^2);
imgf.elem_data(r1<10) = 15; % 30 Ohm.m inhomogeneity @ (10,-15)
imgf.elem_data(r2<20) = 60; % 60 Ohm.m inhomogeneity @ (45,-25)
clf; show_fem(imgf,1); % show forward model
imgf.elem_data = 1./imgf.elem_data; % resistivity -> conductivity
vi = fwd_solve(imgf); vi = vi.meas; % simulated measurements
imgf.elem_data(:) = 1; % homogeneous model, 1 Ohm.m
vh = fwd_solve(imgf); vh = vh.meas; % homogeneous measurements
SNRdB = 40; % model 30dB SNR measurements (1% noise)
vi = add_noise(10^(SNRdB/20),vi); vi = vi.meas;
% --- reconstruction ---
imdl = mk_geophysics_model('h2p5a', 32); % different discretization
imgh = mk_image(imdl.fwd_model, 1); % homogeneous model, 1 Ohm.m
vh = fwd_solve(imgh); vh = vh.meas; % homogeneous measurements
R = 1./((vi'*vi)\(vi'*vh)); % initial estimate (homogeneous)
imdl.fwd_model.stimulation = stim_pattern_geophys(32, 'Wenner'); % Wenner stimulus
imdl.solve = 'inv_solve_gn'; % choose the new iterative Gauss-Newton solver
imdl.reconst_type = 'absolute'; % single set of measurements
imdl.hyperparameter.value = 1;
imdl.inv_solve_gn.prior_data = R; % initial guess
imdl.inv_solve_gn.elem_prior = 'resistivity'; % initial guess units
imdl.inv_solve_gn.elem_working = 'log_conductivity'; % working units
imdl.inv_solve_gn.elem_output = 'resistivity'; % return a resistivity image
imdl.inv_solve_gn.meas_working = 'apparent_resistivity'; % normalize measurements
imgi = inv_solve(imdl, vi); % reconstruct the image from the noisy data
clf; show_fem(imgi,1); % show the reconstruction
```

The non-linear Conjugate Gradient version of the inverse solver can be called by changing the requested solver and moving the iterative solver configuration.

```matlab
% choose the new iterative Conjugate Gradient solver
imdl.solve = 'inv_solve_cg';
```
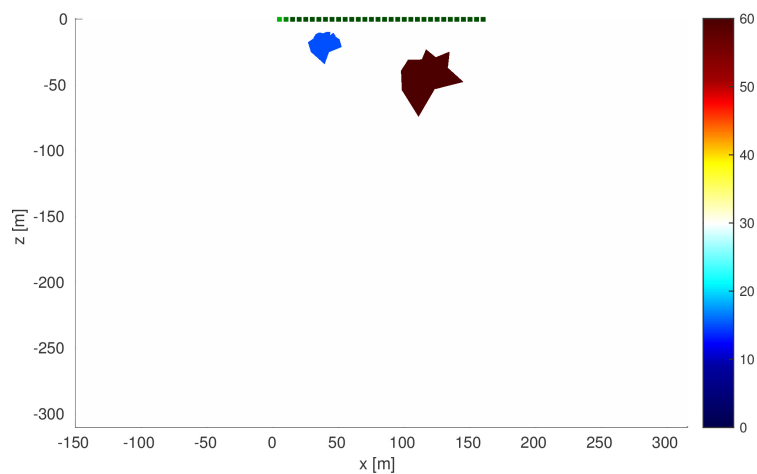
(a) forward model
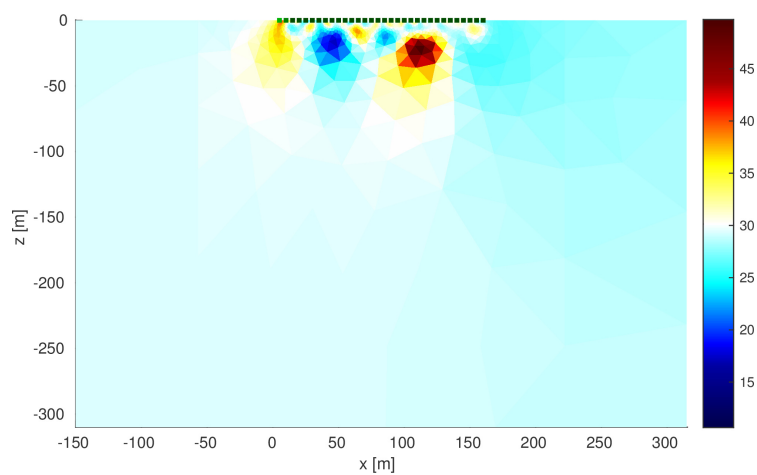
(b) reconstructed image

Figure E.1: Sample reconstruction; (a) $2^{1}/_{2}$-D forward model (32 CEM electrodes (green), half-space, Wenner stimulus, 40dB noise) was used to construct simulated data, (b) a reconstruction with a new discretization (32 CEM electrodes (green), $2^{1}/_{2}$-D, $\lambda = 1.0$) shows reasonable agreement with the forward model after six iterations of the Gauss-Newton iterative inverse solver
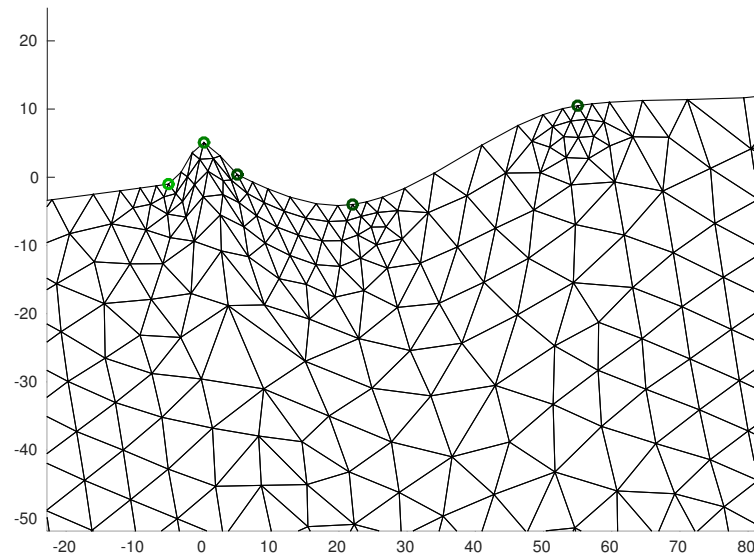
Figure E.2: Model surface topology; the surface topology of the mesh conforms to the electrode locations, PEM electrode nodes circled (green)

```
imdl.inv_solve_cg.prior_data = R; % initial guess
imdl.inv_solve_cg.elem_prior = 'resistivity'; % initial guess units
imdl.inv_solve_cg.elem_working = 'log_conductivity'; % working units
imdl.inv_solve_cg.elem_output = 'resistivity'; % return a resistivity image
imdl.inv_solve_cg.meas_working = 'apparent_resistivity'; % normalize measurements
imgi = inv_solve(imdl, vi); % reconstruct the image from the noisy data
clf; show_fem(imgi,1); % show the reconstruction
```

A model with topology can be constructed by assigning electrode locations, instead of just specifying the number of electrodes (Figure E.2). Electrode arrays that are not approximately collinear are not yet supported. Point electrode models `H` and a denser mesh `c` have been selected in this example. The model is configured for a two-dimensional solution.

```
% (x,z) coordinates by 5 electrodes
elec = [ -5 -1; 0.2 5.1; 5.1 0.4; 22 -4; 55 10.5];
imdl = mk_geophysics_model('H2c', elec); % 2D PEM model
```

Dual mesh models with matching geometry can be trivially constructed (Fig-

ure E.3), where the boundary, electrode geometry, and electrode positioning are aligned between a two- and three-dimensional model. The two models may have different mesh densities. A geometric mapping between the two models is calculated by projecting the two-dimensional model through the three-dimensional model. A two-dimensional topology may also be applied to both models, though not shown in this example.

```
imdl2 = mk_geophysics_model('h2a', 16); % 2D CEM model
imdl3 = mk_geophysics_model('h3a', 16); % 3D CEM model
imdl32 = mk_geophysics_model('h32a', 16); % dual mesh: 3D fwd and 2D rec model
```

A line search (`line_search_onm2`) that performs well on the impedance imaging problems presented in this work has been used by default in `inv_solve_gn` and `inv_solve_cg`.

The combination of simplified model generation and a highly configurable iterative inverse solver considerably eases the exploration of model dependencies and parameter choices. All the functions discussed here:

- `inv_solve_gn`

- `inv_solve_cg`

- `mk_geophysics_model`, and

- `line_search_onm2`

developed by the author, have been contributed to the EIDORS project and are freely available under a GNU GPL version 2 or later licence.
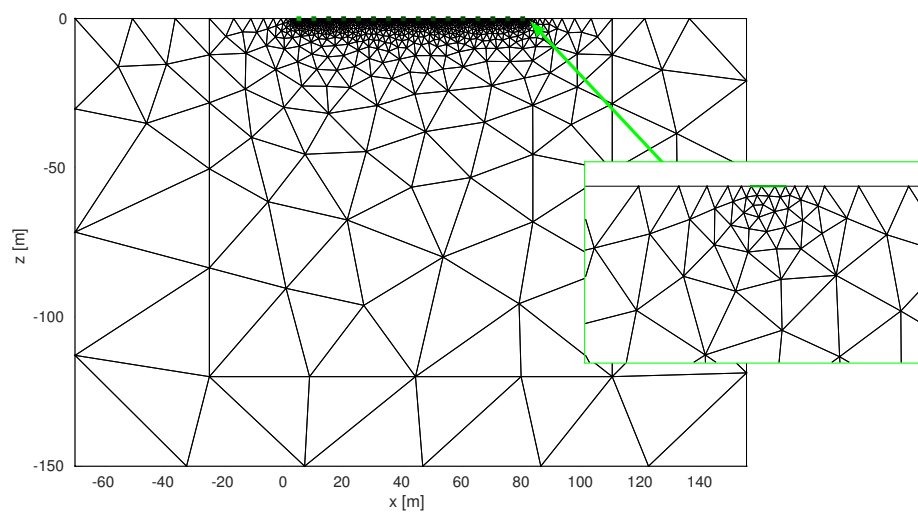
(a) 2D CEM half-space model
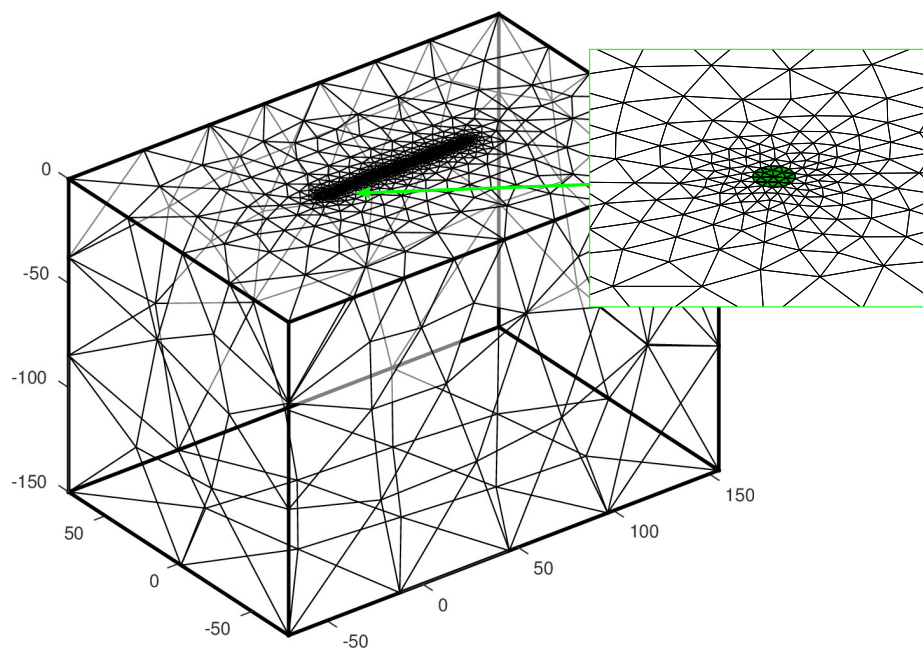
(b) 3D CEM half-space model

Figure E.3: 2D and 3D models; matching domain truncation, electrode dimension and placement for 16 electrode CEM (a) two-dimensional model showing element refinement around the electrodes (inset,green), and (b) three-dimensional model showing same (inset,green)

References

A. Adler and R. Guardo. Electrical impedance tomography: Regularised imaging and contrast detection. *IEEE Trans. Med. Imag.*, 15(2):170–179, April 1996. [cited at page 141, 214]

A. Adler and W. R. B. Lionheart. Uses and abuses of EIDORS: An extensible software base for EIT. *Physiol. Meas.*, 27(5):S25–S42, May 2006. [cited at page 13]

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(4): 414–420, April 1996. [cited at page 7, 133]

A. Adler, J. Arnold, R. Bayford, A. Borsic, B. Brown, P. Dixon, T. Faes, I. Frerichs, H. Gagnon, Y. Gärber, B. Grychtol, G. Hahn, W. R. B. Lionheart, A. Malik, R. Patterson, J. Stocks, A. Tizzard, N. Weiler, and G. Wolf. GREIT: a unified approach to 2D linear EIT reconstruction of lung images. *Physiol. Meas.*, 30(6): S35–S55, June 2009. [cited at page 15]

A. Adler, R. Gaburro, and W. R. B. Lionheart. Electrical impedance tomography. In Otmar Scherzer, editor, *Mathematical Methods in Imaging*, chapter 14, pages 601–654. Springer Science+Business Media, 2011. [cited at page 204]

A. Adler, A. Boyle, M. Crabb, H. Gagnon, B. Grychtol, N. Lesparre, and W. R. B. Lionheart. Eidors version 3.8. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015. [cited at page 10, 184]

J. Ahmad, D. R. Schmitt, C. D. Rokosh, and J. G. Pawlowicz. High-resolution seismic and resistivity profiling of a buried quaternary subglacial valley: Northern alberta, canada. *Geological Society of America Bulletin*, 121(11-12):1570–1583, November 2009. [cited at page 14]

L. Allaud and M. Martin. *Schlumberger: The History of a Technique.* John Wiley & Sons, 1977. [cited at page 13, 37]

B. Andersen and T. Fagerhaug. *Root Cause Analysis: Simplified Tools and Techniques.* ASQ Quality Press, 2nd edition, 2006. [cited at page 82]

G. E. Archie. The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the American Institue of Mining, Metallurgical, and Patroleum Engineers*, 146(1):54–62, 1942. [cited at page 37]

L. Armijo. Minimization of functions having lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics*, 16(1):1–3, 1966. [cited at page 70]

D. Arnold and A. Logg. Periodic table of the finite elements. *SIAM News*, 47(9), November 2014. [cited at page 212]

D. Barber and A. Seagar. Fast reconstruction of resistance images. *Clin. Phys. Physiol. Meas.*, 8(A):47–54, 1987. [cited at page 63]

R. Barker. Signal contribution sections and their use in resistivity studies. *Geophys. J. Int.*, 59(1):123–129, March 1979. [cited at page 192]

R. Barker. Depth of investigation of collinear symmetrical four-electrode arrays. *Geophysics*, 54(8):1031–1037, August 1989. [cited at page 110]

D. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982. [cited at page 63]

M. Berzins. Mesh quality: A function of geometry, error estimates or both? *Engineering with Computers*, 15(3):236–247, September 1999. [cited at page 157]

M. Besson. The metallogeny of france. *Mining Geology*, 32(3):273–280, 1982. [cited at page 117]

K. Boone and D. Holder. Current approaches to analogue instrumentation design in electrical impedance tomography. *Physiol. Meas.*, 17(4):229–247, 1996. [cited at page 25, 26, 27]

L. Borcea. Electrical impedance tomography. *Inv. Prob.*, 18(6):R99–R136, October 2002. [cited at page 13]

A. Borsic and A. Adler. A primal-dual interior-point framework for using the L1 or L2 norm on the data and regularization terms of inverse problems. *Inv. Prob.*, 28(9):1–27, August 2012. [cited at page 51]

A. Borsic, E. Attardo, and R. Halter. Multi-GPU Jacobian accelerated computing for soft-field tomography. *Physiol. Meas.*, 33(10):1703–1715, September 2012. [cited at page 55]

G. Bouchette, P. Church, J. E. McFee, and A. Adler. Imaging of compact objects buried in underwater sediments using electrical impedance tomography. *IEEE Trans. Geosci. Remote Sens.*, 52(2):1407–1417, February 2014. [cited at page 15]

A. Boyle. The Effect of Boundary Shape Deformation on Two-Dimensional Electrical Impedance Tomography. Master's thesis, Carleton University, Ottawa, Canada, 2010. [cited at page 11]

A. Boyle and A. Adler. Impact of electrode area, contact impedance and boundary shape on EIT images. *Physiol. Meas.*, 32(7):745–754, June 2011. [cited at page 7, 11, 98]

A. Boyle, A. Adler, and A. Borsic. Scaling the EIT problem. In *12th Conf. Electrical Impedance Tomography*, Univ. of Bath, Bath, UK, May 2011. [cited at page 11]

A. Boyle, A. Adler, and W. R. B. Lionheart. Shape deformation in two-dimensional electrical impedance tomography. *IEEE Trans. Med. Imag.*, 31(12):2185–2193, December 2012a. [cited at page 7, 11, 98]

A. Boyle, A. Borsic, and A. Adler. Addressing the computational cost of large EIT solutions. *Physiol. Meas.*, 33(5):787–800, April 2012b. [cited at page 11]

A. Boyle, Y. Mamatjan, and A. Adler. Robust stimulation and measurement patterns in biomedical EIT. In *Workshop on 100 Years of Electrical Imaging, Paris, France*, July 2012c. [cited at page 12]

A. Boyle, P. Wilkinson, J. Chambers, N. Lesparre, and A. Adler. Slope stability monitoring through impedance imaging. In *15th Conf. Electrical Impedance Tomography*, Gananoque, Canada, April 2014. [cited at page 11, 159, 184]

A. Boyle, M. Jehl, M. Crabb, and A. Adler. Estimating electrode movement in two dimension. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015a. [cited at page 10, 130, 184]

A. Boyle, Y. Mamatjan, and A. Adler. Selection of stimulus and measurement schemes. In *International Conference on Biomedical Applications of Electrical Impedance Tomography*, Neuchâtel, Switzerland, June 2015b. [cited at page 11]

N. J. Brazee, R. E. Marra, L. Göcke, and P. Van Wassenaer. Non-destructive assessment of internal decay in three hardwood species of northeastern north america using sonic and electrical impedance tomography. *Forestry*, 84(1):33–39, January 2011. [cited at page 15]

R. Brent. *Algorithms for minimization without derivatives*, chapter An Algorithm with Guaranteed Convergence for Finding a Zero of a Function. Prentice-Hall, 1973. [cited at page 69]

R. Brooks. Towards a theory of the comprehension of computer programs. *International Journal of Man-Machine Studies*, 18(6):543–554, June 1983. [cited at page 85]

K. Bryan and M. Vogelius. A computational algorithm to determine crack locations from electrostatic boundary measurements. the case of multiple cracks. *International Journal of Engineering Science*, 32(4):579–603, April 1994. [cited at page 15]

A. P. Calderón. On an inverse boundary value problem. In W. H. Meyer and M. A. Raupp, editors, *Seminar on Numerical Analysis and its Applications to Continuum Physics*, pages 65–73, Rio de Janeiro, 1980. Sociedade Brasileira de Matematica. [cited at page 202]

A. A. Caputi, R. Budelli, K. Grant, and C. C. Bell. The electric image in weakly electric fish: physical images of resistive objects in gnathonemus petersii. *The Journal of Experimental Biology*, 201:2115–2128, July 1998. [cited at page 15]

R. Carlson and F. Fritsch. An algorithm for monotone piecewise bicubic interpolation. *SIAM J. Numer. Anal.*, 26(1):230–238, February 1989. [cited at page 164]

J. Chambers, P. Wilkinson, O. Kuras, J. Ford, D. Gunn, P. Meldrum, C. Pennington, A. Weller, P. Hobbs, and R. Ogilvy. Three-dimensional geophysical anatomy of an active landslide in Lias Group mudrocks, Cleveland Basin, UK. *Geomorphology*, 125(4):472–484, February 2011. [cited at page 160, 162, 170]

J. Chambers, D. Gunn, P. Wilkinson, P. Meldrum, E. Haslam, S. Holyoake, M. Kirkham, O. Kuras, A. Merritt, and J. Wragg. 4D electrical resistivity tomography monitoring of soil moisture dynamics in an operational railway embankment. *Near Surface Geophysics*, 12(1):61–72, February 2014. [cited at page 42]

L. Chauris. Les exploitations minières dans le Massif Armoricain. Déclin ou progrès. *Norois*, 141(1):5–32, January 1989. [cited at page 116]

L. Chen, Y.-J. Du, S.-Y. Liu, and F. Jin. Evaluation of cement hydration properties of cement-stabilized lead-contaminated soils using electrical resistivity measurement. *Journal of Hazardous, Toxic, and Radioactive Waste*, 15(4):312–320, October 2011. [cited at page 15]

M. Cheney, D. Isaacson, J. Newell, S. Simske, and J. Goble. NOSER: an algorithm for solving the inverse conductivity problem. *International Journal of Imaging Systems and Technology*, 2(2):66–75, August 1990. [cited at page 60, 123]

M. Cheney, D. Isaacson, and J. C. Newell. Electrical impedance tomography. *SIAM Review*, 41(1):85–101, January 1999. [cited at page 13]

K. S. Cheng, D. Isaacson, J. C. Newell, and D. G. Gisser. Electrode models for electric current computed tomography. *IEEE Trans. Biomed. Eng.*, 36(9):918–924, September 1989. [cited at page 16]

N. B. Christensen, D. Sherlock, and K. Dodds. Monitoring CO_2 injection with cross-hole electrical resistivity tomography. *Exploration Geophysics*, 37(1):44–49, March 2006. [cited at page 14]

P. Church, J. E. McFee, S. Gagnon, and P. Wort. Electrical impedance tomographic imaging of buried landmines. *IEEE Trans. Geosci. Remote Sens.*, 44(9):2407–2420, September 2006. [cited at page 15]

A. L. Coe, editor. *Geological Field Techniques*. Wiley-Blackwell, October 2010. [cited at page 33]

R. Cook, G. Saulnier, D. Gisser, J. Goble abd J. Newell, and D. Isaacson. ACT3: a high-speed, high-precision electrical impedance tomograph. *IEEE Trans. Biomed. Eng.*, 41(8):713–722, August 1994. [cited at page 195]

J. E. Cooper, P. H. Steele, B. K. Mitchell, C. Boden, and W. R. B. Lionheart. Detecting juvenile wood in southern pine logs with brush electrodes. In *9th Conf. Electrical Impedance Tomography*, Dartmouth College, Hannover, USA, June 2008. [cited at page 15]

T. Dahlin. Short note on electrode charge-up effects in DC resistivity data acquisition using multi-electrode arrays. *Geophysical Prospecting*, 48(1):181–187, January 2000. [cited at page 27]

T. Dahlin and V. Leroux. Improvement in time-domain induced polarization data quality with multi-electrode systems by separating current and potential cables. *Near Surface Geophysics*, 10(6):545–565, April 2012. [cited at page 29]

T. Dahlin, V. Leroux, and J. Nissen. Measuring techniques in induced polarisation imaging. *J. Applied Geophys.*, 50(3):279–298, June 2002. [cited at page 29]

Y. Dai and Y. Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.*, 10(1):177–182, 1999. [cited at page 77]

J. Daniels and A. Dyck. Borehole resistivity and electromagnetic methods applied to mineral exploration. *IEEE Trans. Geosci. Remote Sens.*, GE-22(1):80–87, January 1984. [cited at page 14]

J. Dardé, H. Hakula, N. Hyvönen, and S. Staboulis. Fine-tuning electrode information in electrical impedance tomography. *Inverse Problems and Imaging*, 6(3): 399–421, August 2012. [cited at page 131, 143, 144, 149, 150, 155, 157]

J. Dardé, N. Hyvönen, A. Seppänen, and S. Staboulis. Simultaneous reconstruction of outer boundary shape and admittivity distribution in electrical impedance tomography. *SIAM J. Imaging Sci.*, 6(1):176–198, February 2013. [cited at page 158]

T. Davis, J. Gilbert, S. Larimore, and E. Ng. A column approximate minimum degree ordering algorithm. *ACM Trans. Math. Software*, 30(3):353–376, September 2004. [cited at page 22]

J. Deceuster, O. Kaufmann, and M. Van Camp. Automated identification of changes in electrode contact properties for long-term permanent ERT monitoring experiments. *Geophysics*, 78(2), March 2013. [cited at page 111]

E. Demidenko. An analytic solution to the homogeneous EIT problem on the 2D disk and its application to estimation of electrode contact impedances. *Physiol. Meas.*, 32(9):1453–1471, July 2011. [cited at page 97]

J. Dennis and R. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* SIAM, 1996. [cited at page 69]

A. Dey and H. Morrison. Resistivity modelling for arbitrarily shaped two-dimensional structures. *Geophysical Prospecting*, 27:106–136, March 1979. [cited at page 215, 217]

D. Dillon. *Debugging Strategies for .NET Developers*, chapter Debugging with Logs, pages 87–111. Apress, 2003. [cited at page 113]

J. Ding. Eigenvalues of rank-one updated matrices with some applications. *Applied Mathematics Letters*, 20(12):1223–1226, December 2007. [cited at page 140]

I. Duff, M. Heroux, and R. Pozo. An overview of the sparse basic linear algebra subprograms: The new standard from the BLAS technical forum. *ACM Trans. Math. Software*, 28(2):239–267, June 2002. [cited at page 22]

J. Edd, A. Ivorra, L. Horowitz, and B. Rubinsky. Imaging cryosurgery with EIT: tracking the ice front and post-thaw tissue viability. *Physiol. Meas.*, 29(8):889–912, July 2008. [cited at page 15]

L. Edwards. A modified pseudosection for resistivity and IP. *Geophysics*, 42(5): 1020–1036, August 1977. [cited at page 108, 110]

J. Fayn. Mémoire sur la mine de plomb et de zinc argentifères de Pontpéan (Ille-et-Vilaine). *Revue universelle des mines, de la métallurgie, des travaux publics, des sciences et des arts*, pages 279–349, 1863. [cited at page 117]

S. Fortune. Robustness issues in geometric algorithms. *Applied Computational Geometry Towards Geometric Engineering Lecture Notes in Computer Science*, 1148: 9–14, 1996. [cited at page 58]

H. W. Fowler and F. G. Fowler, editors. *The concise Oxford dictionary of current english.* Oxford University Press, 5th edition, 1964. [cited at page 82]

S. Friedel, A. Thielen, and S. Springman. Investigation of a slope endangered by rainfall-induced landslides using 3d resistivity tomography and geotechnical testing. *J. Applied Geophys.*, 60(2):100–114, October 2006. [cited at page 44, 160]

P. Gaggero, A. Adler, J. Brunner, and P. Seitz. Electrical impedance tomography system based on active electrodes. *Physiol. Meas.*, 33(5):831–847, April 2012. [cited at page 27]

A. Gazoty, G. Fiandaca, J. Pedersen, E. Auken, and A. Christiansen. Data repeatability and acquisition techniques for time-domain spectral induced polarization. *Near Surface Geophysics*, 11(4):391–406, August 2013. [cited at page 29]

D. L. George, J. R. Torczynski, K. A. Shollenberger, T. J. O'Hern, and S. L. Ceccio. Validation of electrical-impedance tomography for measurements of material distribution in two-phase flows. *International Journal of Multiphase Flow*, 26(4): 549–581, April 2000. [cited at page 15]

S. Gill. The diagnosis of mistakes in programmes on the edsac. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 206 (1087):538–554, May 1951. [cited at page 82]

D. Goldberg. What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys (CSUR)*, 23(1):5–48, March 1991. [cited at page 48]

C. Gómez-Laberge and A. Adler. Direct EIT jacobian calculations for conductivity change and electrode movement. *Physiol. Meas.*, 29(6):S89–S99, June 2008. [cited at page 8, 131, 140, 142, 149, 155, 166, 168]

C. Gómez-Laberge, J. H. Arnold, and G. K. Wolf. A unified approach for EIT imaging of regional overdistension and atelectasis in acute lung injury. *IEEE Trans. Med. Imag.*, 31(3):834–842, March 2012. [cited at page 15]

B. M. Graham and A. Adler. Objective selection of hyperparameter for EIT. *Physiol. Meas.*, 27(5):S65–S79, April 2006. [cited at page 14]

R. Guérin, G. Delaplace, L. Bouvier, P. Debreyne, O. Furling, D. Beudon, and J.C. Leuliet. Thermal and structural heterogeneities during cooking of a starch-based product by a non-intrusive ring sensor. *Chemical Engineering and Processing: Process Intensification*, 46(4):343–350, April 2007. [cited at page 15]

T. Günther, C. Rücker, and K. Spitzer. Three-dimensional modelling and inversion of dc resistivity data incorporating topography - II. Inversion. *Geophys. J. Int.*, 166(2):506–517, 2006. [cited at page 102]

J. Hadamard. Sur les problemes aux derivees partielles et leur signification physique. *Princeton University Bulletin*, 13(1):49–52, 1902. [cited at page 46]

P. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM review*, 34(4):561–580, May 1992. [cited at page 52]

Per Christian Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.* SIAM, Philadelphia, 1998. [cited at page 49]

M. Harry and R. Schroeder. *Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporations.* Random House, 2000. [cited at page 82]

A. Hartinger, H. Gagnon, and R. Guardo. Accounting for hardware imperfections in EIT image reconstruction algorithms. *Physiol. Meas.*, 28(7):S13–S27, June 2007. [cited at page 3]

M. Hazewinkel, editor. *Encyclopaedia of Mathematics*, volume 1, chapter Conformal mapping, pages 806–810. Springer-Verlag, 2002. [cited at page 97]

L. Heikkinen, T. Vilhunen, R. West, and M. Vauhkonen. Simultaneous reconstruction of electrode contact impedances and internal electrical properties: II. laboratory experiments. *Meas. Sci. and Tech.*, 13(12):1855–1861, November 2002. [cited at page 132]

M. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436, December 1952. [cited at page 75]

N. Higham. *Accuracy and Stability of Numerical Algorithms.* SIAM, 2nd edition, 2002. [cited at page 48]

D. G. Hill. A laboratory investigation of electrical anisotropy in precambrian rocks. *Geophysics*, 37(6):1022–1038, December 1972. [cited at page 43]

D. Holder, editor. *Electrical impedance tomography: Methods, history and applications.* IOP Publishing Ltd, 2005. [cited at page 202]

T. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Dover Publications, 2000. [cited at page 206]

C. Hundhausen, S. Douglas, and J. Stasko. A meta-study of algorithm visualization effectiveness. *Journal of Visual Languages & Computing*, 13(3):259–290, June 2002. [cited at page 88]

N. Hyvönen. Complete electrode model of electrical impedance tomography: Approximation properties and characterization of inclusions. *SIAM J. Appl. Math.*, 64(3):902–931, March 2004. [cited at page 16]

IEEE. Floating-Point Arithmetic. *IEEE Std. 754-2008*, pages 1–70, 2008. [cited at page 48]

D. Isaacson. Distinguishability of conductivities by electric current computed tomography. *IEEE Trans. Med. Imag.*, 5(2):91–95, June 1986. [cited at page 18, 96]

K. Ishikawa. *Guide to quality control.* Asian Productivity Organization Tokyo, 1982. [cited at page 82]

T. Johnson, R. Versteeg, A. Ward, F. Day-Lewis, and A. Revil. Improved hydrogeophysical characterization and monitoring through parallel modeling and inversion of time-domain resistivity andinduced-polarization data. *Geophysics*, 75(4): WA27–WA41, July 2010. [cited at page 102]

H. Jomard, T. Lebourg, S. Binet, E. Tric, and M. Hernandez. Characterization of an internal slope movement structure by hydrogeophysical surveying. *Terra Nova*, 19(1):48–57, February 2007. [cited at page 160]

C. Jones, B. Grychtol, H. Gagnon, A. Boyle, C. He, A. Adler, and P. Gaggero. Open electrical impedance tomography (OEIT) file format. In *15th Conf. Electrical Impedance Tomography*, Gananoque, Canada, April 2014. [cited at page 11]

D. Jongmans and S. Garambois. Geophysical investigation of landslides: a review. *Bulletin de la Société de France*, 178(2):101–112, March 2007. [cited at page 43, 160]

L. Jouniaux, A. Maineult, V. Naudet, M. Pessel, and P. Sailhac. Review of self-potential methods in hydrogeophysics. *Comptes Rendus Geoscience*, 341(10-11): 928–936, October 2009. [cited at page 29]

J. Kaipio and E. Somersalo. Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, 198(2):493–504, January 2007. [cited at page 58, 105, 222]

J. Kaipio, V. Kolehmainen, E. Somersalo, and M. Vauhkonen. Statistical inversion and monte carlo sampling methods in eit. *Inv. Prob.*, 16(5):1487–1522, June 2000. [cited at page 84]

T.-J. Kao, D. Isaacson, J. Newell, and G. Saulnier. A 3D reconstruction algorithm for EIT using a handheld probe for breast cancer detection. *Physiol. Meas.*, 27(5): S1–S11, April 2006. [cited at page 15]

K. Karhunen, A. Seppänen, A. Lehikoinen, P. J. M. Monteiro, and J. P. Kaipio. Electrical resistance tomography imaging of concrete. *Cement and Concrete Research*, 40(1):137–145, January 2010. [cited at page 15]

G. Keller and F. Frischknecht. *Electrical methods in geophysical prospecting*, vol-

ume 10 of *International series in electromagnetic waves*. Oxford: Pergamon Press, 1966. [cited at page 34]

A. Kemna. *Tomographic inversion of complex resistivity*. PhD thesis, Ruhr-Universität Bochum, Bochum, Germany, 2000. [cited at page 38]

J. Kiefer. Sequential minimax search for a maximum. *Proceedings of the American Mathematical Society*, 4(3):502–506, June 1953. [cited at page 69]

C. Klein and A. R. Philpotts. *Earth Materials: Introduction to Mineralogy and Petrology*. Cambridge University Press, 2013. [cited at page 32]

P. Knupp. Algebraic mesh quality metrics for unstructured initial meshes. *Finite Elements in Analysis and Design*, 39(3):217–241, January 2003. [cited at page 157]

M. Kölbl-Ebert. How to find water: The state of the art in the early seventeenth century, deduced from writings of martine de bertereau (1632 and 1640). *Earth Sciences History*, 28(2):204–218, 2009. [cited at page 117]

V. Kolehmainen, M. Vauhkonen, P. Karjalainen, and J. Kaipio. Assessment of errors in static electrical impedance tomography with adjacent and trigonometric current patterns. *Physiol. Meas.*, 18(4):289–303, August 1997. [cited at page 3]

O. Kuras, J. Pritchard, P. Meldrum, J. Chambers, P. Wilkinson, R. Ogilvy, and G. Wealthall. Monitoring hyrdaulic processes with automated time-lapse electrical resistivity tomography (ALERT). *Comptes Rendus Geoscience*, 341(10-11):868–885, November 2009. [cited at page 9, 26, 160]

D. LaBrecque, M. Miletto, W. Daily, A. Ramirez, and E. Owen. The effects of noise on Occam's inversion of resistivity tomography data. *Geophysics*, 62(2):538–548, March 1996a. [cited at page 27, 111]

D. LaBrecque, A. Ramirez, W. Daily, A. Binley, and S. Schima. ERT monitoring of environmental remediation processes. *Meas. Sci. and Tech.*, 7(3):375–383, March 1996b. [cited at page 14]

R. E. Langer. An inverse problem in differential equations. *Bulletin of the American Mathematical Society*, 39(10):681–824, October 1933. [cited at page 13]

R. E. Langer. On determination of earth conductivity from observed surface potentials. *Bulletin of the American Mathematical Society*, 42(10):747–754, October 1936. [cited at page 13]

P. Langlois, Y. Wu, R. Bayford, and A. Demosthenous. On the application of frequency selective common mode feedback for multifrequency EIT. *Physiol. Meas.*, 36(6):1337–1350, May 2015. [cited at page 26]

V. Lapenna, P. Lorenzo, A. Perrone, S. Piscitelli, E. Rizzo, and F. Sdao. 2d electrical resistivity imaging of some complex landslides in lucanian apennine chain, southern italy. *Geophysics*, 70(3):B11–B18, May 2005. [cited at page 160]

J. Lawrance, C. Bogart, M. Burnett, R. Bellamy, K. Rector, and S. Flemming. How programmers debug, revisited: An information foraging theory perspective. *IEEE Trans. Softw. Eng.*, 29(2):197–215, February 2013. [cited at page 82]

T. Lebourg, S. Binet, E. Tric, H. Jomard, and S. El Bedoui. Geophysical survey to estimate the 3D sliding surface and the 4D evolution of the water pressure

on part of a deep seated landslide. *Terra Nova*, 17(5):399–406, October 2005a. [cited at page 160]

T. Lebourg, S. Binet, E. Tric, H. Jomard, and S. El Bedoui. Geophysical survey to estimate the 3d sliding surface and the 4d evolution of the water pressure on part of a deep seated landslide. *Terra Nova*, 17(5):399–406, October 2005b. [cited at page 160]

T. Lebourg, M. Hernandez, S. Zerathe, S. El Bedoui, H. Jomard, and B. Fresia. Landslides triggered factors analysed by time lapse electrical survey and multi-dimensional statistical approach. *Engineering Geology*, 114(3-4):238–250, August 2010. [cited at page 44]

C. D. Lee, F. J. Vine, and R. G. Ross. Electrical conductivity models for the continental crust based on laboratory measurements on high-grade metamorphic rocks. *Geophysical Journal of the Royal Astronomical Society*, 72(2):353–371, February 1983. [cited at page 43]

N. Lesparre, D. Gibert, F. Nicollin, C. Nussbaum, and A. Adler. Monitoring the excavation damaged zone by three-dimensional reconstruction of electrical resistivity. *Geophys. J. Int.*, 195(2):972–984, November 2013. [cited at page 14]

N. Lesparre, B. Grychtol, A. Boyle, and A. Adler. Consideration of inverse problem parametrization meshing for electrical resistivity tomography (submitted). *Meas. Sci. and Tech.*, 2014. [cited at page 11]

N. Lesparre, A. Boyle, B. Grychtol, J. Cabrerra, J. Marteau, and A. Adler. Electrical resistivity imaging in transmission between the surface and underground

tunnel for fault characterization. *J. Applied Geophys.*, 128(1):163–178, May 2016. [cited at page 11]

S. Letovsky. Cognitive processes in program comprehension. *Journal of Systems and Software*, 7(4):325–339, December 1987. [cited at page 82]

T. S. Light, S. Licht, A. C. Bevilacqua, and K. R. Morash. The fundamental conductivity and resistivity of water. *Electrochem. Solid-State Lett.*, 8(1):E16–E19, 2005. [cited at page 41, 42]

Linnaeus. *Species plantarum.* L. Salvius, Stockholm, c.1753. [cited at page 31]

W. R. B. Lionheart. Boundary shape and electrical impedance tomography. *Inv. Prob.*, 14(1):139–147, February 1998. [cited at page 97]

A. Lodin. Note sur des dégagements de gaz inflammables survenus dans des mines métalliques, noamment dans celle de Pontpéan. *Annales des mines*, 9(8):40–105, 1895. [cited at page 117]

A. Lodin. Notice historique sur l'exploitation des mines de Pontpéan. *Annales des mines*, 10(14):5–72, 1908. [cited at page 117]

A. Lodin. Notice historique sur l'exploitation des mines de Pontpéan (suite et fin). *Annales des mines*, 10(20):457–553, 1911. [cited at page 117]

M. Loke and R. Barker. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-newton method. *Geophysical Prospecting*, 44(1):131–152, January 1996a. [cited at page 192]

M.H. Loke and R.D. Barker. Practical techniques for 3d resistivity surveys and data inversion. *Geophysical Prospecting*, 44(1):499–523, May 1996b. [cited at page 119]

R. F. Martin, editor. *The Nomenclature of Minerals: A Compilation of IMA Reports*, volume 0. Mineralogical Association of Canada, 1998. [cited at page 31]

A. Maslow. A theory of human motivation. *Psychological Review*, 50(4):370–396, July 1943. [cited at page 86]

J. Maxwell-Garnett. Colours in metal glasses and in metallic films. *Proceedings of the Royal Society of London*, 73:443–445, 1904. [cited at page 40]

A. McEwan, G. Cusick, and D. Holder. A review of errors in multi-frequency EIT instrumentation. *Physiol. Meas.*, 28(7):S197–S215, June 2007. [cited at page 26, 27]

S. Meeson, B. Blott, and A. Killingback. EIT data noise evaluation in the clinical environment. *Physiol. Meas.*, 17(4A):A33–A38, November 1996. [cited at page 111]

A. Memon. Automatically repairing event sequence-based gui test suites for regression testing. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 18(2), November 2008. [cited at page 91]

A. Merrit, J. Chambers, P. Wilkinson, L. West, W. Murphy, D. Gunn, and S. Uhlemann. Measurement and modelling of soil moisture – electrical resistivity relationship for the calibration of time-lapse electrical resistivity tomograms (ERT). *J. Applied Geophys.*, 2015. (*in review*). [cited at page 40]

A. Merritt. *4D Geophysical Monitoring of Hydrogeological Precursors to Landslide Activation*. PhD thesis, University of Leeds, Leeds, United Kingdom, April 2014. [cited at page 169]

A. Merritt, J. Chambers, W. Murphy, P. Wilkinson, L. West, D. Gunn, P. Meldrum, M. Kirkham, and N. Dixon. 3D ground model development for an active landslide in lias mudrocks using geophysical, remote sensing and geotechnical methods. *Landslides*, 11(4):537–550, August 2014. [cited at page 170]

D. Michie. "memo" functions and machine learning. *Nature*, 1(218):19–22, April 1968. [cited at page 152]

J. Moré, B. Garbow, and K. Hillstrom. Testing unconstrained optimization software. *ACM Transactions on Mathematical Software (TOMS)*, 7(1):17–41, March 1981. [cited at page 83]

V. Morozov. On the solution of functional equations by the method of regularization. In *Soviet Math. Dokl*, volume 7, pages 414–417, 1966. [cited at page 52]

J. Mueller, S. Siltanen, and D. Isaacson. A direct reconstruction algorithm for electrical impedance tomography. *IEEE Trans. Med. Imag.*, 21(6):555–559, June 2002. [cited at page 13]

Y. Nakae, M. Kagaya, R. Takagi, Y. Matsutani, H. Horibe, and T. Kondo. Cold pain prolongs gastric emptying of liquid but not solid meal: an electrical impedance tomography (EIT) study. *Journal of Gastroenterology*, 35(8):593–597, August 2000. [cited at page 15]

V. Naudetb, M. Lazzari, A. Perrone, A. Loperte, S. Piscitelli, and V. Lapenna. Integrated geophysical and geomorphological approach to investigate the snowmelt-triggered landslide of Bosco Piccolo village (Basilicata, southern Italy). *Engineering Geology*, 98(3-4):156–167, May 2008. [cited at page 160]

M. E. Nelson and M. A. Maciver. Prey capture in the weakly electric fish apteronotus albifrons: sensory acquisition strategies and electrosensory consequences. *The Journal of Experimental Biology*, 202:1195–1203, May 1999. [cited at page 15]

I. D. Neveln, Y. Bai, J. B. Snyder, J. R. Solberg, O. M. Curet, K. M. Lynch, and M. A. MacIver. Biomimetic and bio-inspired robotics in electric fish research. *The Journal of Experimental Biology*, 216:2501–2514, July 2013. [cited at page 15]

F. Nicollin, D. Gibert, N. Lesparre, and C. Nussbaum. Anisotropy of electrical conductivity of the excavation damaged zone in the Mont Terri Underground Rock Laboratory. *Geophys. J. Int.*, 181(1):303–320, February 2010. [cited at page 43]

J. Nocedal and S. Wright. *Numerical Optimization.* Springer Science Business Media, 1999. [cited at page 69, 77]

R. Ogilvy, P. Meldrum, O. Kuras, P. Wilkinson, J. Chambers, M. Sen, A. Pulido-Bosch, J. Gisbert, S. Jorreto, I. Frances, and P. Tsourlos. Automated monitoring of coastal aquifers with electrical resistivity tomograph. *Near Surface Geophysics*, 7(5-6):367–375, October 2009. [cited at page 26]

T. Ohno. *Toyota Production System: beyond large-scale production.* Productivity Press, 1988. [cited at page 82]

G. Oldenborger, P. Routh, and M. Knoll. Sensitivity of electrical resistivity tomography data to electrode position errors. *Geophys. J. Int.*, 163(1):1–9, January 2005. [cited at page 160]

D. Oldenburg and Y. Li. Estimating depth of investigation in dc resistivity and IP surveys. *Geophysics*, 64(2):403–416, March 1999. [cited at page 110]

F. O'Sullivan. A statistical perspective on ill-posed inverse problems. *Statistical Science*, 1(4):502–518, November 1986. [cited at page 52]

B. Packham, H. Koo, A. Romsauerova, S. Ahn, A. McEwan, S. Jun, and D. Holder. Comparison of frequency difference reconstruction algorithms for the detection of acute stroke using EIT in a realistic head-shaped tank. *Physiol. Meas.*, 33(5): 767–787, April 2012. [cited at page 15]

D. Parasnis. Reciprocity theorems in geoelectric and geoelectromagnetic work. *Geo-exploration*, 25(3):177–198, October 1988. [cited at page 27, 105]

P. Pébay and T. Baker. Analysis of triangle quality measures. *Mathematics of Computation*, 72(244):1817–1839, 2003. [cited at page 157]

A. Perrone, A. Iannuzzi, V. Lapenna, P. Lorenzo, S. Piscitelli, E. Rizzo, and F. Sdao. High-resolution electrical imaging of the Varco d'Izzo earthflow (southern Italy). *J. Applied Geophys.*, 56(1):17–29, May 2004. [cited at page 160]

A. Perrone, V. Lapenna, and S. Piscitelli. Electrical resistivity tomography technique for landslide investigation: A review. *Earth-Science Reviews*, 135:65–82, August 2014. [cited at page 44]

M. Pessel and D. Gibert. Multiscale electrical impedance tomography. *Journal of Geophysical Research*, 108(B1):EPM6–1 – EPM6–12, January 2003. [cited at page x, 116, 117, 118, 119, 121, 123]

E. Piegaria, V. Cataudella, R. Di Maio, L. Milano, M. Nicodemi, and M. Soldovieri. Electrical resistivity tomography and statistical analysis in landslide modelling: A conceptual approach. *J. Applied Geophys.*, 68(2):151–158, June 2009. [cited at page 160]

N. Polydorides, T. Ouypornkochagorn, and H. McCann. Inequality constrained EIT modelling and inversion. In *15th Conf. Electrical Impedance Tomography*, Gananoque, Canada, April 2014. [cited at page 63]

A. Portal, S. Béné, P. Boivin, E. Busato, C. Cârloganu, C. Combaret, P. Dupieux, F. Fehr, P. Gay, P. Labazuy, I. Laktineh, J.-F. Lénat, D. Miallier, L. Mirabito, V. Niess, , and B. Vulpescu. Inner structure of the puy de dôme volcano: cross-comparison of geophysical models (ert, gravimetry, muonic imagery). *Geosci. Instrum. Method. Data Syst.*, 2(2):703–717, September 2012. [cited at page 14]

G. Reeves and H. Harrison. Obtaining the specific contact resistance from transmission line model measurements. *IEEE Electron Device Lett.*, 3(5):111–113, May 1982. [cited at page 95]

J. J. Renton. *Physical Geology*. West Publishing Co., 1994. [cited at page 33, 39]

D. Robinson and S. Friedman. A method for measuring the solid particle permattivity or electrical conductivity of rocks, sediments, and granular materials. *Journal of Geophysical Research: Solid Earth*, 108(B2):ECV5–1–ECV5–9, February 2003. [cited at page 40]

P. Romero, B. du Boulay, R. Cox, R. Lutz, and S. Bryant. Debugging strategies and tactics in a multi-representation software environment. *International Journal of Human-Computer Studies*, 65(12):992–1009, December 2007. [cited at page 82]

A. Ross, G. Saulnier, J. Newell, and D. Isaacson. Current source design for electrical impedance tomography. *Physiol. Meas.*, 24(2):509–516, April 2003. [cited at page 25, 27]

G. Rothermel, S. Elbaum, A. Malishevsky, P. Kallakuri, and X. Qiu. On test suite composition and cost-effective regression testing. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 13(3):277–331, July 2004. [cited at page 83]

C. Rücker and T. Günther. The simulation of finite ERT electrodes using the complete electrode model. *Geophysics*, 76(4):F227–F238, July–August 2011. [cited at page 206]

L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, November 1992. [cited at page 51]

O. Sass, R. Bell, and T. Glade. Comparison of GPR, 2D-resistivity and traditional techniques for the subsurface exploration of the öschingen landslide, swabian alb (germany). *Geomorphology*, 93(1-2):89–103, January 2008. [cited at page 160]

L. Scesi and P. Gattinoni. *Water Circulation in Rocks*. Springer Science+Business Media, 2009. [cited at page 43]

C. Schlumberger. *Étude sur la Prospection Électrique du Sous-sol*. Gauthier-Villars et Cie, 1920. [cited at page 37]

J. Schöberl. NETGEN an advancing front 2D/3D-mesh generator based on abstract rules. *Computing and visualization in science*, 1(1):41–52, 1997. [cited at page 13]

J. Sherman and W. Morrison. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1):124–127, March 1950. [cited at page 140]

J. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, Carnegie Mellon University, 1994. [cited at page 76, 113]

J. Shewchuk. Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. In *Applied Computational Geometry Towards Geometric Engineering: FCRC'96 Workshop, WACG'96*, pages 203–222, Philadelphia, PA, May 1996. [cited at page 157]

J. Shewchuk. Adaptive precision floating-point arithmetic and fast robust geometric predicates. *Discrete & Computational Geometry*, 18(3):305–363, October 1997. [cited at page 58]

S. Skiena. *The Algorithm Design Manual, 2nd ed.* Springer, 2008. [cited at page 58]

B. J. Skinner and S. C. Porter. *Physical Geology.* John Wiley & Sons, 1987. [cited at page 33]

L. D. Slater, A. Binley, and D. Brown. Electrical imaging of fractures using groundwater salinity change. *Groundwater*, 35(3):436–442, May 1997. [cited at page 14]

L. B. Slichter. The interpretation of the resistivity prospecting method for horizontal structures. *Journal of Applied Physics*, 4(9):307–322, September 1933. [cited at page 13]

M. Soleimani, C. Gómez-Laberge, and A. Adler. Imaging of conductivity changes and electrode movement in EIT. *Physiol. Meas.*, 27(5):S103–S113, May 2006. [cited at page 8, 159]

E. Somersalo, M. Cheney, and D. Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.*, 52(4): 1023–1040, 1992. [cited at page 16, 17, 202, 204]

C. Soulsby, M. Khela, E. Yazaki, D. Evans, E. Hennessy, and J. Powell-Tuck. Measurements of gastric emptying during continuous nasogastric infusion of liquid feed: Electric impedance tomography versus gamma scintigraphy. *Clinical Nutrition*, 25 (4):671–680, August 2006. [cited at page 15]

P. Steele, J. Cooper, and W. R. B. Lionheart. Through-log density detector, August 2004. US Patent 6,784,672. [cited at page 15]

O. Stern. Zur theorie der elektrolytischen doppelschicht. *Zeitschrift für Elektro-chemie und angewandte physikalische Chemie*, 30(21-22):508–516, November 1924. [cited at page 39]

G. Strang and G. Fix. *An Analysis of the Finite Element Method.* Prentice Hall, September 1973. [cited at page 13, 19]

R. Supper, D. Ottowitz, B. Jochum, J. Kim, A. Römer, I. Baron, S. Pfeiler, M. Lovisolo, S. Gruber, and F. Vecchiotti. Geoelectrical monitoring: an innovative method to supplement landslide surveillance and early warning. *Near Surface Geophysics*, 12(1):133–150, February 2014. [cited at page 44, 160]

K. Suzuki and S. Higashi. Groundwater flow after heavy rain in landslide-slope area from 2-D inversion of resistivity monitoring data. *Geophysics*, 66(3):733–743, May 2001. [cited at page 160]

T. N. Tallman, S. Gungor, K. W. Wang, and C. E. Bakis. Damage detection and conductivity evolution in carbon nanofiber epoxy via electrical impedance tomography. *Smart Materials and Structures*, 23(4):1–9, April 2014. [cited at page 15]

D. S. Tawil, D. Rye, and M. Velonaki. Improved image reconstruction for an EIT-based sensitive skin with multiple internal electrodes. *IEEE Trans. Robot.*, 27(3): 425–435, June 2011. [cited at page 15]

W. Telford, L. Geldart, and R. Sheriff. *Applied Geophysics, 2nd ed.* Cambridge University Press, 1990. [cited at page 34]

A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. V. H. Winston and Sons (distributed by Wiley, New York), 1977. translation by F. John. [cited at page 49]

Y. S. Touloukian, W. R. Judd, and R. F. Roy, editors. *Physical Properties of Rocks and Minerals*, volume II-2 of *CINDAS data series on material properties*. Hemisphere Publishing Corp., 1989. [cited at page 34, 36]

L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM, 1997. [cited at page 22]

A. Tucker, E. Ross, J. Paugh-Miller, and R. Sadleir. In vivo quantification of accumulating abdominal fluid using an electrical impedance tomography hemiarray. *Physiol. Meas.*, 32(2):151–165, December 2011. [cited at page 15]

H. Ucok, I. Ershaghi, and G. R. Olhoeft. Electrical resistivity of geothermal brines. *Journal of Petroleum Technology*, 32(4):717–727, April 1980. [cited at page 38]

S. Uhlemann, P. Wilkinson, J. Chambers, H. Maurer, A. Merrit, D. Gunn, and P. Meldrum. Interpolation of landslide movements to improve the accuracy of

4D geoelectrical monitoring. *J. Applied Geophys.*, 121(1):93–105, October 2015. [cited at page 162]

T. Vilhunen, J. Kaipio, P. Vauhkonen, T. Savolainen, and M. Vauhkonen. Simultaneous reconstruction of electrode contact impedances and internal electrical properties: I. theory. *Meas. Sci. and Tech.*, 13(12):1848–1854, November 2002. [cited at page 95, 111, 132]

A. Vonk-Noordegraaf, A. Janse, J. Marcus, J. Bronzwaer, P. Postmus, T. Faes, and P. de Vries. Determination of stroke volume by means of electrical impedance tomography. *Physiol. Meas.*, 21(2):285–293, 2000. [cited at page 15]

G. Wahba and Y. Wang. When is the optimal regularization parameter insensitive to the choice of the loss function? *Communications in Statistics - Theory and Methods*, 19(5):1685–1700, 1990. [cited at page 52]

J. Wait. *Geo-Electromagnetism.* Academic Press, 1982. [cited at page 97]

Y. Wan, R. Halter, A. Borsic, P. Manwaring, A. Hartov, and K. Paulsen. Sensitivity study of an ultrasound coupled transrectal electrical impedance tomography system for prostate imaging. *Physiol. Meas.*, 31(8):S17–S29, July 2010. [cited at page 15]

K. Wang, D. Lary, D. Shallcross, S. Hall, and J. Pyle. A review on the use of the adjoint method in four dimensional atmospheric-chemistry data assimilation. *Quarterly Journal of the Royal Metrological Society, Part B*, 127(576):2181–2204, July 2001. [cited at page 55]

M. Waxman and L. Smits. Electrical conductivities in oil-bearing shaly sands. *SPE Journal*, 8(2):107–122, June 1968. [cited at page 40]

M. H. Waxman and E. C. Thomas. Electrical conductivities in shaly sands-i. the relation between hydrocarbon saturation and resistivity index; ii. the temperature coefficient of electrical conductivity. *Journal of Petroleum Technology*, 26(2), February 1974. [cited at page 41]

P. Wilkinson, J. Chambers, M. Lelliot, G. Wealthall, and R. Ogilvy. Extreme sensitivity of crosshole electrical resistivity tomography measurements to geometric errors. *Geophys. J. Int.*, 173(1):49–62, April 2008. [cited at page 62, 160]

P. Wilkinson, J. Chambers, P. Meldrum, D. Gunn, R. Ogilvy, and O. Kuras. Predicting the movements of permanently installed electrodes on an active landslide using time-lapse geoelectrical resistivity data only. *Geophys. J. Int.*, 183(2):543–556, August 2010. [cited at page xi, 9, 160, 169, 170, 172]

P. Wilkinson, M. Loke, p. Meldrum, J. Chambers, O. Kuras, D. Gunn, and R. Ogilvy. Practical aspects of applied optimized survey design for electrical resistivity tomography. *Geophysical Journal International*, 189(1):428–440, January 2012. [cited at page 27, 29]

P. Wilkinson, J. Chambers, S. Uhlemann, P. Meldrum, A. Smith, N. Dixon, and M. Loke. Reconstruction of landslide movements by inversion of 4-d electrical resistivity tomography monitoring data. *Geophysical Research Letters*, 43(3):1166–1174, February 2016. [cited at page 171]

P. Wilson, L. Dell, and G. Anderson. *Root cause analysis: A tool for total quality management.* ASQ Quality Press, 1993. [cited at page 82]

R. Winkler and A. Rieder. Model-aware Newton-type inversion scheme for electrical impedance tomography. *Inv. Prob.*, 31(4):1–27, March 2015. [cited at page 60, 95, 132]

A. Wirgin. The inverse crime. *ArXiv Mathematical Physics e-prints*, January 2004. [cited at page 58]

P. Wolfe. Convergence conditions for ascent methods. *SIAM Review*, 11(2):226–235, 1969. [cited at page 70]

H. M. N. Wright, K. V. Cashman, E. H. Gottesfeld, and J. J. Roberts. Pore structure of volcanic clasts: Measurements of permeability and electrical conductivity. *Earth and Planetary Science Letters*, 280(1-4):93–104, April 2009. [cited at page 38, 43]

D. O. Wyble. Effect of applied pressure on the conductivity, porosity and permeability of sandstones. *Journal of Petroleum Technology*, 10(11), November 1958. [cited at page 41]

T. York. Status of electrical tomography in industrial applications. *Journal of Electronic Imaging*, 10(3):608–619, July 2001. [cited at page 15]

T. Yorkey, J. Webster, and W. Tompkins. Comparing reconstruction algorithms for electrical impedance tomography. *IEEE Trans. Biomed. Eng.*, BME-34(11): 843–852, November 1987. [cited at page 141, 214]

A. Zeller. *Why programs fail: a guide to systematic debugging.* Morgan Kaufmann Publishers, Inc, 2nd edition, 2009. [cited at page 82, 113]

B. Zhou and T. Dhalin. Properties and effects of measurement errors on 2d resistivity imaging surveying. *Near Surface Geophysics*, 1(3):105–117, August 2003. [cited at page 7, 133, 160]

O. C. Zienkiewicz, R. L. Taylor, and J. Z. Zhu. *The Finite Element Method: Its Basis and Fundamentals.* Butterworth-Heinemann, sixth edition, 2005. [cited at page 206]

Vitae

Name:	Alistair James Scott Boyle, PEng.
Research Interests:	Electrical Impedance Tomography, the Finite Element Method, and Inverse Problems
Education:	Ph.D., Electrical and Computer Engineering, Carleton University, 2010–2016
	M.A.Sc., Biomedical Engineering, Carleton University, 2007–2010
	B.Sc., Electrical Engineering, University of Calgary, 1997–2002
Experience:	Senior ASIC Designer, Diablo Technologies, 2011–Present
	Hardware Design Engineer, Freescale Semiconductors, 2005–2011
	Design & Verification Engineer, Founder, Seaway Networks, 2001–2005
	Verification Engineer, Nortel, 2000–2001
Awards:	President's Doctoral Fellowship (Ph.D.), 2011
	Entry Scholarship (Masters), 2007
	EnCana Scholarship, 1997–2001