Electromagnetic Non-Destructive Testing of the Graphite Core of Advanced Gas-Cooled Reactors.

Joel Hampton

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

Department of Electrical and Electronic Engineering

University of Manchester

2022

Contents

List of Figures

List of Tables

Abbreviations

CO_2 Carbon Dioxide

ACG Augmented Conjugate Gradient

AE Auto-Encoder

AGR Advanced Gas-cooled Reactor

ALARP As Low As Reasonably Practicable

BFGS Broyden-Fletcher-Goldfarb-Shanno

CBMU Channel Bore Inspection Unit

CCC Committee on Climate Change

CDT Centre for Doctoral Training

CNN Convolutional Neural Network

DMI Differential Mutual Inductance

ECIT Eddy Current Inspection Tool

EDF Électricité de France

FCN Fully Connected Network

FEA Finite Element Analysis

FEM Finite Element Model

FNR Fast Neutron Radiation

GD Gradient Descent

GN Gauss-Newton

HC High Conductivity

ICA Independent Component Analysis

IS Inductance Spectroscopy

JFNL James Fisher Nuclear Laboratory

KNN k-nearest neighbours

KPCA Kernel Principal Component Analysis

LC Low Conductivity

LM Levenberg-Marquart

LR Logistic Regression

MC Medium Conductivity

MGN Modified Gauss-Newton

ML Machine Learning

MPE Mean Percentage Error

MPEN Mean Percentage Error Norm

MVPR Multi-Variable Polynomial Regression

NDE Non-Destructive Evaluation

NDT Non-Destructive Testing

NICIE New in-Core Inspection Equipment

ONR Office for Nuclear Regulation

PCA Principal component analysis

PD Powell's Dogleg

PECIT Prototype Eddy Current Inspection Tool

PEP Prototype Elliptical Probe

PGA Pile Grade-A

PoPECT Proof of Principle Eddy Current Tool

PWR Pressure Water Reactor

QN Quasi-Newton

RCNDE Research Centre in Non-Destructive Evaluation

RO Radiolytic Oxidation

SNR Signal-Noise Ratio

SVD Singular Value Decomposition

SVM Support Vector Machine

TSVD Truncated Singular Value Decomposition

UK United Kingdom

WENRA Western European Nuclear Regulators Association

Nomenclature

$\boldsymbol{\Gamma}$ The regularisation matrix

$\boldsymbol{D}$ The scaling matrix

$\boldsymbol{H}_f(\vec{\sigma}_k)$ The Hessian evaluated at some point $\vec{\sigma}_k$

$\boldsymbol{J}$ Jacobian

γ The damping parameter

λ The regularisation parameter

$\nabla_{\vec{\sigma}} f(\vec{\sigma}_k)$ The gradient evaluated at some point $\vec{\sigma}_k$

$\vec{d} = -\frac{\nabla_{\vec{\sigma}} f(\vec{\sigma}_k)}{\left\|\nabla_{\vec{\sigma}} f(\vec{\sigma}_k)\right\|}$ The gradient decent direction

$\vec{p} = -\boldsymbol{H}_f(\vec{\sigma}_k)^{-1} \nabla_{\vec{\sigma}} f(\vec{\sigma}_k)$ The Newton direction

$\vec{p} = -(\boldsymbol{H}_f(\vec{\sigma}_k) + \gamma \boldsymbol{D}^T \boldsymbol{D})^{-1} \nabla_{\vec{\sigma}} f(\vec{\sigma}_k)$ The re-scaled trust region direction

$\vec{p} = -(\boldsymbol{H}_f(\vec{\sigma}_k) + \gamma \boldsymbol{I})^{-1} \nabla_{\vec{\sigma}} f(\vec{\sigma}_k)$ The trust region direction

$f_\lambda(\vec{\sigma}) = \frac{1}{2}\left\|\vec{r}(\vec{\sigma})\right\|_2^2 + \frac{\lambda}{2}\left\|\boldsymbol{\Gamma}(\vec{\sigma} - \vec{\sigma}_0)\right\|_2^2$ The regularised least squares objective function

$q(\vec{p}) = f(\vec{\sigma}_k) + \nabla_{\vec{\sigma}} f(\vec{\sigma}_k)^T \vec{p} + \frac{1}{2}\vec{p}^T \boldsymbol{H}_f(\vec{\sigma}_k)\vec{p}$ The quadratic Taylor series expansions of an arbitrary function

Abstract

The UK's advanced gas cooled reactors use graphite as a moderator and structural component. The graphite core degrades over time, and this is a concern to both controlling reactivity and the safety of the core. To inspect the graphite core, inductance spectroscopy is used as the soft-field imaging modality. This thesis considers the reconstruction of the conductivity depth profile and the detection of subsurface cracks.

Within the thesis, multiple iterative inversion algorithms are compared for reconstructing the conductivity depth profile; of the iterative algorithms, second order methods are shown to be the best performing in different cases considering the prior estimate and noise level.

Two machine learning algorithms are compared for direct inversion: multivariable polynomial regression and a convolutional neural network. The results show that the two algorithms have a comparable performance.

A finite element model is used to generate data to train machine learning algorithms and in the iterative inversion of eddy current data. The model must accurately represent the physical experiment; therefore, there must be a rigorous calibration procedure. A constrained optimisation algorithm is presented for calibration and inversion of the eddy current data. The mean signal-noise ratio after calibration was 29.47 dB for a 3D coil model and 28.96 dB for the filament; prior to optimisation, the signal-noise ratio was 1.48 dB and 4.98 dB, respectively.

In the inversion of graphite data, it is shown that the iterative algorithm

generally does not improve on the prior estimate because the residual norm is below the discrepancy, causing steps to be rejected. The restrictions on the step could be relaxed but this increases susceptibility to modelling error. Instead, improvements can be made by increasing modelling accuracy.

Feature extraction and classification algorithms are investigated for detection of subsurface notches. The best balanced accuracy achieved on synthetic test data was 71.45% using principle component analysis and a neural network. It is not clear how this compares with existing techniques.

Declaration of Originality

I hereby confirm that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

Acknowledgements

I would like to give thanks to my supervisors and colleagues: Anthony Peyton, Matthew Brown, Henok Tesfalem and Adam Fletcher. Thank you for your patience and all your help during my time on the project.

Particular thanks are given to Adam Fletcher, for his help with characterising and combating resonance issues.

To Zoë, Niko and Renée. This wouldn't have been possible without you. To all my family and friends, thank you.

Joel Hampton.

Word count = 44771

Chapter 1

Introduction

1.1 Aims and Application Scope

The fleet of Advanced Gas-cooled Reactor (AGR) stations in the United Kingdom (UK) use graphite channels to: moderate neutrons, direct the coolant flow and provide a structural material in the core. Over time, the irradiation of the core degrades the graphite. Understanding the degradation processes is important for the safety case and for potentially controlling reactivity (due to loss of graphite); this requires measuring the density of the graphite.

The density of the bricks is currently determined by trepanning the graphite core and testing the sample. This is a destructive process which is limited to carefully selected sites. Furthermore, the trepanned samples do not span the full width of the brick, and therefore, the graphite behaviour at the periphery is generally unknown. This thesis presents the development of eddy current inspection of the graphite channels. The advantages of using eddy current inspection can be summarised as:

- Deployable in harsh environments, such as a reactor core. This is possible because the sensor is a coil of wire and has no active components

which would otherwise need shielding from the radiation.

- Non-contact and single sided inspection.
- Increased data acquisition. Trepanning is slow and has to be done in specific locations to minimise damage to the bricks.
- Reconstruction of the depth profile spanning the full width of graphite; the trepanned sample does not penetrate the full width of the graphite.
- A method of detecting subsurface cracking.
- Non-destructive, and thus, inspections can be repeated in the same locations for continued monitoring.

The aim is to recover the material properties from a set of measurements: this is an inverse problem. To date, the time taken to solve the inverse problem limits the reconstruction process to be offline. The objectives of this project are to speed up the reconstruction process without loss of accuracy (relative to the methodology used in previous iterations of the program), improve modelling methodology and to develop the crack detection capabilities. There is a balance to be found with speed and accuracy; tomographic images of a poor quality will be of little use for understanding the degradation of the core but slow algorithms limit the reconstructions to be offline and may take too long to be useful. Ideally, reconstructions could be performed whilst inspecting the core which would help identify locations which may require further testing.

1.2 Achievements

1.2.1 Publications

Primary Author

- J. Hampton, H. Tesfalem, A. Fletcher, A. Peyton, and M. Brown, "Reconstructing the conductivity profile of a graphite block using inductance spectroscopy with data driven techniques." Insight — Non Destructive Testing and Condition Monitoring, vol. 63, no. 2, pp. 82–87, 2021.

- J. Hampton, A. Fletcher, H. Tesfalem, A. Peyton, and M. Brown, "Comparison of Non-Linear Optimisation Algorithms for Recovering the Conductivity Depth Profile of an Electrically Conductive Block Using Eddy Current Inspection." NDT and E, vol. 125, 2022.

- J. Hampton, H. Tesfalem, O. Dorn, A. Fletcher, A. Peyton, and M. Brown, "Calibration of a Finite Element Forward Model in Eddy Current Inspection." IEEE Sensors Journal, vol. 22, no. 11, pp. 10699 - 10707, 2022.

Co-Author

- H. Tesfalem, J. Hampton, A. Fletcher, A. Peyton, and M. Brown, "Electrical Resistivity Reconstruction of Graphite Moderator Bricks From Multi-Frequency Measurements and Artificial Neural Networks." IEEE Sensors Journal, vol. 21, no. 15, pp. 17005 - 17016, 2021.

1.2.2 Conferences

- J. Hampton, H. Tesfalem, A. Fletcher, A. Peyton, and M. Brown, "Electromagnetic non-destructive testing of the graphite cores of advanced gas-cooled reactors (AGRs)." British Institute of Non-Destructive Testing Webinar Week, 2020 (Presented)
- J. Hampton, H. Tesfalem, O. Dorn, A. Fletcher, A. Peyton, and M. Brown, "Calibration of a Finite Element Forward Model in Eddy Current Inspection." Research Centre in Nondestructive Evaluation Conference, 2022 (Presented)
- J. Hampton, H. Tesfalem, A. Fletcher, A. Peyton, and M. Brown, "Using a Machine Learning Forward Model in an Iterative Depth Profile Reconstruction Algorithm." British Institute of Non-Destructive Testing/59th Annual British Conference on Non-Destructive Testing 2022 (Accepted)

1.3 Thesis Organisation

The thesis is written in standard format. There are 7 chapters in total, not including the introductory and closing conclusion chapters. Chapter 2 — The Graphite Core — aims to provide sufficient background detail in order to understand the use, material properties and degradation mechanisms of nuclear graphite. Chapter 3 — Inductance Spectroscopy and The Inverse Problem — is the most substantial component of the thesis and is concerned with the mathematical formulation of the problem, the background physics, the modelling of the graphite and inversion theory. Chapter 4 — Comparison of Iterative Algorithms — consolidates previous research with a comprehensive study on the performance of inversion algorithms in a number of different

cases studies. Chapter 5 — Inversion by Inference — reports on inversion using inferential techniques and compares two different algorithms. Chapter 6 — Sensor Characterisation — describes the calibration of the forward model using the laboratory and reactor probes. Moreover, the resonant issues with the reactor system are discussed. Chapter 7 — Depth Profiling the Graphite Core — introduces the algorithm used for inverting inductance data collected from a nuclear reactor, and specifically, the difficulties encountered with the iterative algorithms. Chapter 8 — Crack Detection — reports on a study of feature extraction and classification algorithms for distinguishing cracks in the measurement response from conductivity variations.

1.4 UK Research Centre in Non-Destructive Evaluation (RCNDE)

This research project is funded by the RCNDE Centre for Doctoral Training (CDT), grant number EP/L022125/1, and Électricité de France (EDF) Energy. The RCNDE is concerned with the generation of research relating to the quantitative assessment of the condition of components, plant and engineering structures of all kinds during manufacture and in-service. Moreover, it is concerned with the design of novel sensors and algorithms for the capture and processing of data. The sensor technology can be from any number of modalities, with the two main being electromagnetic and ultrasonic.

There are two arguments for the development and continued research of Non-Destructive Evaluation (NDE): public safety and reduced downtime if monitoring must be performed online. Therefore, it is key that NDE be as efficient as possible, with reliable, quantitative results and robust sensor technology. The RCNDE is connected to industries such as aerospace, power generation, defence and oil and gas; through improving safety, new engineering

designs are facilitated as the risk of catastrophic failure is lowered. Therefore, the benefits are multi-faceted. The complete mission statement of this grant can be found in [1].

Over the duration of the project the RCNDE provides a number of compulsory and optional training courses. These courses are typically offered at one of the member universities, where corresponding expertise are held. The key benefit of the CDT is that students in the cohort can share ideas, support each other and have access to high quality research and teaching. These courses are usually divided into a theory constituent and practical laboratory sessions, providing an opportunity to apply and test the understanding of the theory within a controlled environment. For each course an assessment must be completed, either through coursework, presentation or exam. The courses completed to date are:

- Intro to NDE
- NDE of Composites
- Practical NDE
- Acoustics and Ultrasonics
- Electromagnetic NDE
- Ultrasonic Transduction
- Signal Processing
- Intro to Matlab
- Finite Element Modelling
- Radiography

Chapter 2

Background and Industrial Motivation

2.1 Overview

This chapter begins with a discussion of historic and projected energy and electricity usage, in the context of achieving net zero greenhouse gas emissions by the year 2050. The implications of achieving net zero are a vast shift in the way electricity is used and sourced, and consequently, there is a large amount research and development of technologies such as carbon capture, hydrogen and heat pumps.

To facilitate understanding of the next generation of nuclear reactors, basic concepts of reactor designs are explained, where key differences in reactor designs are: the fuel cycle, the point of operation in the neutron spectrum (neutron energy), the coolant used and the temperature of the core. Introduced are the promising generation IV reactor designs capable of heat, hydrogen and electricity generation; these capabilities make nuclear extremely attractive for achieving net zero. The UK will likely adopt the high temperature gas-cooled reactor, which is capable of heat, electricity and hydrogen production. These

are considered an evolution of the UK's current reactor technology; therefore, there already exists the expertise and regulatory standards for such a reactor.

The relevance of the work presented is then given in the safety case, where it is essential that core components are monitored for failure and operation; the loss of density of a moderator affects both the structural integrity of the core and reactivity. The economic value of Non-Destructive Testing (NDT) is then provided in terms of service and equipment industries, where growth in both of these areas is expected with a projected combined market value of $11.6 billion dollars in 2022.

The AGR is briefly described, specifically the role of graphite and its degradation mechanisms. Further detail is provided on the manufacturing process of graphite and how this process affects the material properties and degree of isotropy — this is important in modelling the graphite,

Current inspection technologies are described, which are: visual inspection with a camera, displacement measurements to measure channel distortion and eddy current inspection for crack detection and depth profiling. The harsh environment and single sided access limits the tools available. Additionally, inspection mostly concerns the fuel channels; there is scope and motivation for developing tools for the inspection of the control rod channels. In this thesis, only the fuel channels are considered.

A literature review is presented with regards to eddy current inspection. In the graphite application, there are no analytic solutions to the inverse problem; therefore, iterative techniques are used. This requires a forward model which accurately describes the experimental problem. To tune a forward model, a grid search can be used to determine some optimal model parameters. However, this process is slow, and thus, optimisation techniques are more favourable. Although there are no analytic inverse solutions in this application, Machine Learning (ML) methods can be used to approximate the

mapping.

2.2 Industrial Context

2.2.1 Energy and Electricity Demand

The UK is at the beginning of a vast shift in the way that energy is both used and sourced. Recently, an independent public body, the Committee on Climate Change (CCC), released its target of reaching net-zero greenhouse gases by 2050 [2]. It is expected that in order to reach net-zero, extensive electrification of current energy sources is required; the electricity generated must be derived from low-carbon sources. A cautious estimate of double the current electricity demand is given by the year 2050 [2]. Further reductions in emissions will come from the use of hydrogen, heat pumps and carbon capture technology [2], [3]. The current plan set by the CCC is to source at least 59% the electrical energy in 2050 from renewables [4]. The remainder is to be sourced from nuclear, natural gas and bioenergy, with emissions balanced using carbon capture technology. The full 10 point plan laid out by the government can be found in [5].

Fig. 2.1 shows the total breakdown of energy usage into four sectors: industry, transport, domestic and other — other includes agriculture, commercial and services. The most energy intensive sector is transport, followed closely by domestic, industry and then other. There is a clear low utilisation of electricity within the transport sector; in 2020 almost 95% of the energy used was sourced from petroleum, mainly owed to the reliance of petroleum in aviation and road transport [6]. This is in sharp contrast to the other sectors where electricity and natural gas are much more substantial constituents. In the UK, natural gas is currently the largest source of energy used to generate

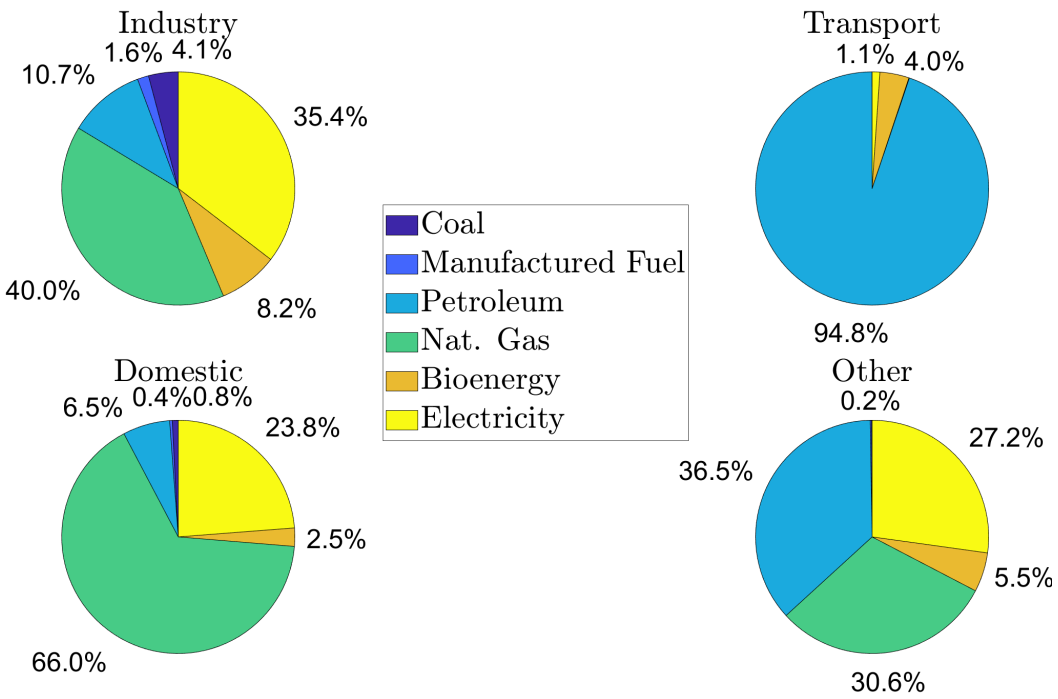

Figure 2.1: The breakdown of energy consumption across the four sectors in 2020 [6].

electricity. Recently, wind has overtaken nuclear to become the second largest source, the complete breakdown is shown in fig. 2.2.

The path taken towards achieving net zero is a highly divisive subject. This path can be chosen according to the current lowest carbon sources of energy or the potential to be the lowest. The best approach to achieving net zero is complex and requires consideration of a technologies maturity, deciding where best to direct investment and the social and environmental impact of action or inaction. It is prudent to remember the words of Watson-Watt:

> *Give them the third best to go on with; the second best comes too late, the best never come.*

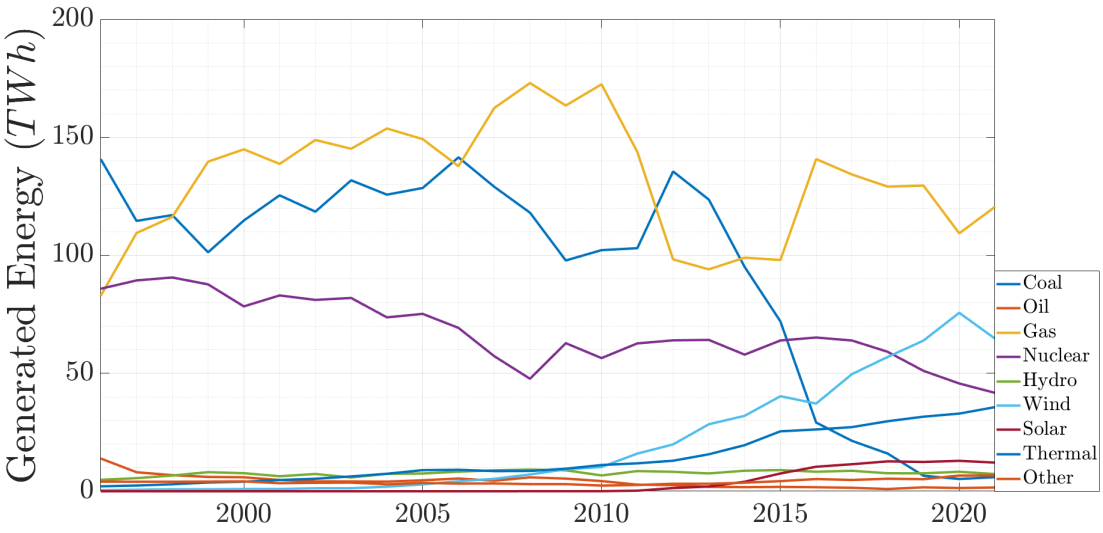

Figure 2.2: The sources of fuel used for the generation of electricity over time [7].

2.2.2 Development of Nuclear Energy in the UK

The civil nuclear industry was born out of military applications of weaponising atomic energy during the second world war. In the UK, the first nuclear reactors were built in Harwell, Oxfordshire at the Atomic Energy Research Establishment, starting operation in 1947. The success of the initial research programs led to the construction of the air-cooled Windscale Piles site in Cumbria, followed by the construction of Calder Hall and later the Windscale AGR. The expansion of the nuclear program was, in part, a response to the Suez crisis in 1956; as such, the use of nuclear energy is both a strategic and low emissions asset.

The world's first commercial nuclear power station was built at Calder Hall, UK. The reactor built at Calder Hall used fuel elements clad in a magnesium alloy; hence, these were called the Magnox reactors. Following the success of the Magnox reactors (in producing commercial levels of power), attention shifted to the next generation of reactors. There was some debate surrounding whether the UK should adopt a design based on its own AGR

Figure 2.3: The timeline for decommissioning the current AGR fleet [9].

technology or American water-cooled reactors. The decision was to use AGR technology which could attain much larger thermal efficiencies compared to the Magnox reactors and Pressure Water Reactor (PWR) designs. The thermal efficiency of various reactor designs is given in table 2.1 — the thermal efficiency is defined as the percentage of input thermal energy that is converted to useful work in the electrical generator output. AGRs have since

Table 2.1: Summary of different reactor types, considering thermal efficiency, coolant and fuel [8].

Reactor	Fuel	Moderator	Coolant	Thermal efficiency
Magnox	Natural uranium	Graphite	Gas	$\approx 31\%$
AGR	Enriched uranium	Graphite	Gas	$\approx 42\%$
PWR	Enriched uranium	Water	Water	$\approx 32\%$

been built at 7 different sites, currently with a combined total of 14 reactor cores, where the majority of these sites will be closed in the next few years (see fig. 2.3). Now that the AGRs are due to be decommissioned, the question of future reactor designs has reappeared.

2.2.3 The Future of Nuclear Energy

Nuclear power stations have a lifetime measured in decades, and consequently, nuclear reactor designs are split into generations. Generation I reactors refer to the experimental reactor designs, such as Calder Hall, and generation II are the reactors subsequently built, designed for increased economy and reliability; for example, the AGR and PWR. Generation III reactors are essentially the same designs as the second generation but with improvements to safety and efficiency; the European PWR is a generation III reactor. The generation IV reactors are currently being developed, with an international consortium researching the different reactor designs. The focus of the generation IV reactor designs is on safety, sustainability, cost-effectiveness and resistance to weapons proliferation.

Nuclear Role in Combating Climate Change

The use of nuclear energy for the reduction of greenhouse gas emissions is an ongoing source of debate in academic literature. It has been argued that nuclear power does not associate with lower carbon emissions, and additionally, that nuclear and renewables show a tendency to be mutually incompatible [10]. Within [10], it is posed that nuclear and renewables tend to "crowd each other out", where the larger scale requirements of nuclear energy conflict with the infrastructure required for smaller scale renewables sites; for example, the two may require different energy grid designs, with renewables being more distributed. However, this is not necessarily true of all nuclear technologies; for example, the modular reactor may have a higher compatibility with renewable technologies and this is given as a limitation of the study by the authors: in the study all nuclear and renewable technologies were grouped into two classes, and therefore, the relative merit of different reactor designs was overlooked. These findings have been directly challenged in the literature

and evidence has been provided that both are environmentally beneficial in reducing national Carbon Dioxide (CO_2) levels [11]. Further, it is argued that, in some countries, the geography and weather may not be conducive to renewable generation [11]. Additionally, there are some promising next generation reactor designs that have broader capabilities in electricity, hydrogen and heat generation: comparing these with renewables is more complex.

Nuclear Fission Fundamentals

"Fission" is derived from the Latin word *findere*, meaning 'to split'. Nuclear fission is the splitting of an atom into smaller atoms by inducing instability, typically with neutron capture (an alternative is photofissioning). This process occurs as the nuclei of large atoms decay into a more stable form, in which there is a more favourable balance of the nuclear and electric force exerted by neutrons and protons respectively. The fissioning of a single uranium-235 isotope releases approximately 162 MeV in kinetic energy of its daughter atoms, with residual energy in emitted gamma rays, neutrons etc. [12]. Uranium-235 has an atomic mass of 235.043930131 u [13], meaning that in one gram of uranium-235 there is a potential 1.59201×10^{13} Joules of kinetic (thermal) energy. This value is subject to the probability of capture and fissioning. In a realistic design, nuclear can achieve up to 28 GJ/g and is considerably larger than natural gas alternatives which generate approximately 0.05 MJ/g [14].

The fissioning process is shown in fig. 2.4. The elements of concern in nuclear engineering are the actinides, a group of rare earth metals which are all radioactive [15]. Over time, radioactive isotopes decay and convert from one type of element to another (known as transmutation). There is considerable research into the separation and transmutation of waste products from nuclear fuel [16]–[18]. There is potential to extract further energy from these elements and also to reduce the amount of radioactive waste that is sent to a geological

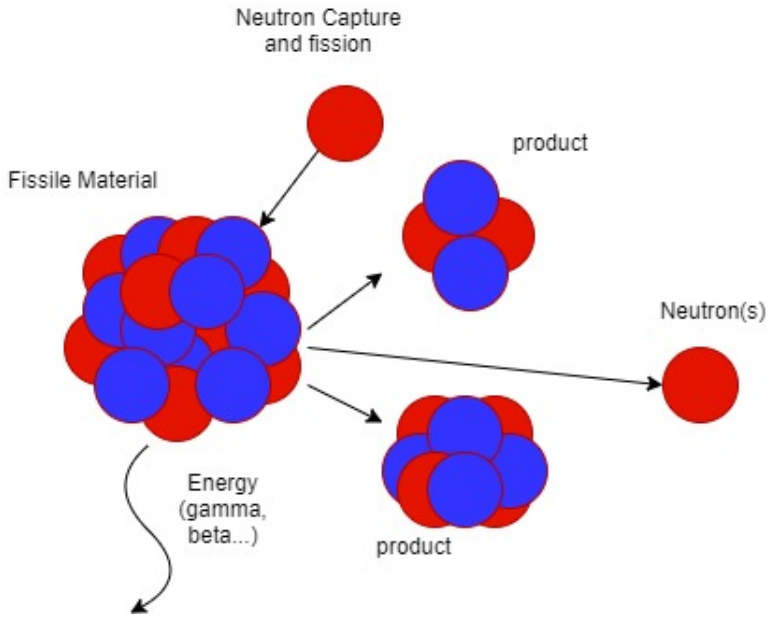

Figure 2.4: The fissioning of an element through neutron capture.

repository.

Consider fig. 2.5, in which the number of neutrons emitted per neutron absorbed is plotted against the incident neutron energy. There are two characteristic regions of operation, referred to as the fast neutron regime and thermal regime. The difference between the two regimes is in the probability of neutron absorption and subsequent fission. In the fast regime, the probability of neutron absorption is lower but if absorbed it has a higher probability of fissioning than in the thermal regime [12]. The regime of operation is one of the key distinctions in nuclear reactor designs, where those that operate in the fast regime are referred to as fast reactors and those which use the thermal regime thermal reactors.

Due to the differences in the probability of neutron absorption and fissioning, the thermal regime generally requires less fuel but has a poorer fuel economy. Additionally, the fast reactor is better at reducing the amount of radioactive waste products because the larger neutron energy increases the probability of the minor actinides fissioning [16]. The recycling of the waste products is an active area of research for both regimes [17], [18]. The recy-

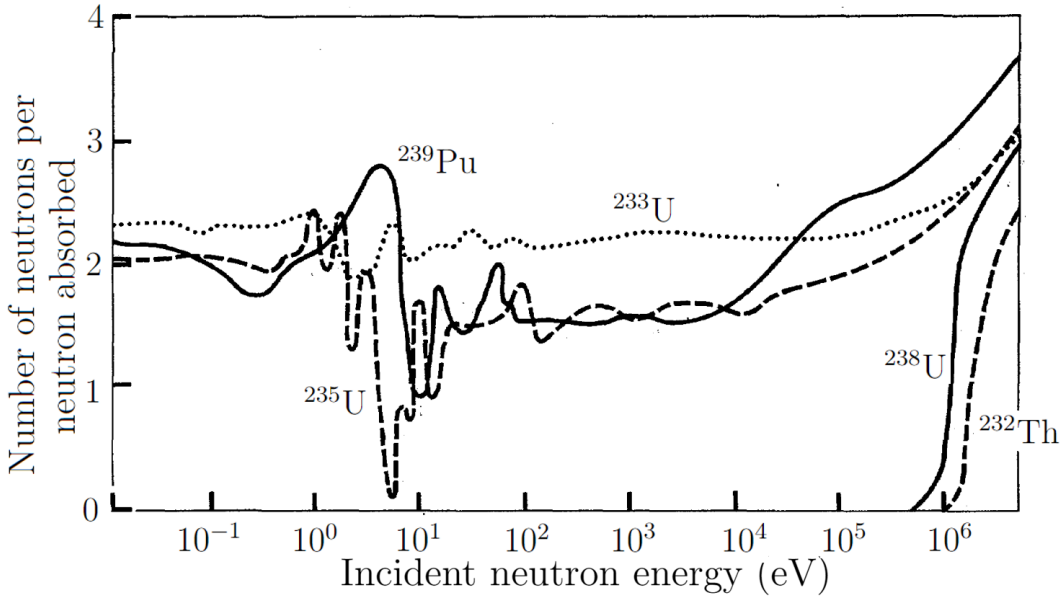

Figure 2.5: The average number of neutrons released per neutron capture, as a function of incident neutron energy [20].

cling of the spent fuel is referred to as the closed fuel cycle and its complexities and cost in reprocessing the fuel make it less favourable than simply using it once and safely disposing, i.e. the open cycle [19]. Further, there are increased short term proliferation threats from reprocessing in the closed fuel cycle because of the extraction of plutonium and other usable fissile material; this increased risk is from the transport of materials and the extraction of weapons grade material, such as plutonium-239. However, in the long term, the burn up of the recycled material means that these would not be present in spent fuel, and therefore, there would be a lower risk associated with the disposal of spent fuel [19].

Further differences between reactors include the type of fuel used, such as uranium-235, uranium-233 or plutonium-239; these isotopes can be found naturally or by transmutation of a fertile material, such as thorium-232 or uranium-238. Some reactor fuels use a mixture of recycled plutonium and uranium, called mixed oxide. Transmutation of thorium-232 has a number of benefits: thorium is more abundant than uranium, it has a higher fuel burn up (more efficient), it has less long-lived radioactive waste and is considered more

difficult to make nuclear weapons [21] [22]. However, the National Nuclear Laboratory released an independent review of fuel technologies, concluding that a thorium fuel cycle does not yield much benefit over a fast reactor using uranium-238 to breed fissionable plutonium [23]. Further, they discuss the over-statement of proliferation resistance of thorium, since it needs a fissionable uranium isotope to induce the reaction. Although they do not dismiss thorium based reactors in the future, it is widely considered an immature technology with further research and development required. Additional distinguishing features include the use of a solid or liquid fuel, the coolant used and the heat engine cycle — some potential reactors have a single primary circuit directly driving a steam turbine. Reactors which use a gas or water coolant require pressurisation of the core, and thus, have a larger footprint in components concerning safety.

Not all of the nuclear fission products are waste: the decay products bismuth-213 and actinium-225 are examples of a widely researched alpha emitting isotopes for the treatment of cancers, such as prostate and leukemia, in a process known as targeted alpha therapy [24], [25]. Similarly, radium-223 (a product of the uranium-235 decay chain) is currently being researched for the therapy of bone cancer [24]. The production of these isotopes remains a challenge. The bismuth and actinium radionuclides are largely sourced from stockpiles from previous nuclear research programs. The sustainable production of these isotopes is being pursued, without reliance on the nuclear fuel cycle.

Proposed Designs

In designing any power generation system that relies on the transduction of thermal energy, the maximum efficiency is dictated by the Carnot theorem: $\zeta_{max} = \frac{T_H - T_C}{T_H}$, where T_H is the temperature of the hot reservoir and T_C the

cold. This means that the maximum efficiency is dictated by the difference of inlet and outlet temperature in the primary heat exchanger, not withstanding the losses in the thermal circuits. Therefore, to increase the thermal efficiency, reactors should be run at higher temperatures, but this has consequences in designing a system that can withstand such temperatures. As an example of limited efficiency, the AGR reactor has a gas inlet and outlet temperature of approximately 339°C and 669°C, respectively [26] and therefore has a maximum thermal efficiency of $\zeta_{max} = 49\%$, whilst the actual efficiency is $\zeta \approx 40\%$. Besides thermal efficiency, the total efficiency of a reactor is driven by the fuel economy: how much energy is extracted from the fuel.

There are six reactor types being considered:

- Super-critical water-cooled reactor: this design is an advancement of the light-water reactor, with several improvements to safety and efficiency. The water coolant is operated at super critical condition, meaning it is indistinguishable as a gas or liquid. This requires operating the reactor at high pressure of 22.1 MPa (218.11 atmospheres) and temperature in the range of 510-625 °C, resulting in a thermal efficiency of approximately 44% [27]. The need for steam dryers, steam separators, recirculating pumps, and steam generators is eliminated, reducing the size of the core. The operation of the coolant at super-critical condition also avoids the formation of bubbles within the circuit, which can otherwise affect the moderator abilities of the core. The concept of the super-critical water cooled reactor can be used in either a fast or thermal reactor design with an open or closed fuel cycle [28] and has the possibility of being used with a thorium fuel [27], [29]. However, this reactor would need significant research and development before demonstration [28].

- Very high-temperature gas reactor: this is conceptually very similar

to the UK's AGRs. It is a helium cooled, graphite moderated reactor, operating in the thermal neutron spectrum with an open cycle [28]. Because of the open cycle this design has a larger number of waste products when compared with the other generation IV designs; however, it does have the adaptability of burning plutonium and minor actinides [28]. Operating at upwards of 1000 °C, it has a high thermal efficiency of approximately 50% [27]. It is attractive as a means of both power and thermal generation for process industries such as steel and aluminium, chemical and oil. It can also produce hydrogen as a bi-product and is, therefore, attractive for the broader aims for achieving net zero emissions. It can operate with a uranium, plutonium and thorium fuel types, allowing potential design modifications as research into thorium progresses [28].

- Molten salt reactor: the molten salt reactor has been researched for many decades, with the first experimental reactor operating in the 1960's at the Oak Ridge National Laboratory. Instead of using a solid fuel it effectively uses a liquid fuel by dissolving the nuclear fuel into the coolant — a molten salt. The reactor can be operated as either a fast neuron, thermal or breeder reactor. In the latter two reactors, the moderator is a matrix of graphite cylinders in which the coolant flows through; however, modern research in this generation IV design is focused on the fast reactor [28]. Due to the high output temperature, this reactor can be used for process manufacturing and for the production of hydrogen [27]. The thermal efficiency is approximately 45%. It has inherent safety properties such as operating at atmospheric pressure and passive draining of the fuel from the core if required [30]. This reactor would be operated under a closed fuel cycle [28]. A unique characteristic of the molten salt reactor fuel-coolant is the ability to process the fuel online without handling, making it attractive for waste

minimisation and lower proliferation risks (compared with other closed cycle designs). However, there are significant barriers in research and development and the lower level of expertise of such a reactor design within the UK.

- Gas-cooled fast reactor: similar to the very high-temperature gas reactor, this design uses helium but is operated as a fast reactor with a closed fuel cycle [28]. This type of reactor is a direct-cycle, meaning there is no separation between the turbines driving the electrical generator and the reactor core, with the advantage being fewer losses and higher efficiency [27]. With a high outlet temperature of 850 °C it is suitable for heat, electricity and hydrogen generation and has a thermal efficiency of approximately 45%. However, this technology requires extensive research and development before demonstration.

- Lead-cooled fast reactor: This design operates in the fast region of the neutron spectrum, with a closed fuel cycle. This reactor uses molten lead as a primary coolant, which is highly abundant, and therefore, inexpensive to source [28]. Further, there are a number of safety features which make this attractive: atmospheric pressurisation of the core, high boiling point of lead (boiling would produce voids in the coolant) and passive natural circulation safety systems [31]. Moreover, the inherent safety of the design lowers the footprint of the plant and systems within it, reducing the cost. However, lead is corrosive above 500 °C and has complications in terms of fuel and coolant inspection due to being optically and electromagnetically opaque. Due to the lower outlet temperature of approximately 550 °C, this reactor is suitable only for electricity generation and has a moderate thermal efficiency of 41% [28].

- Sodium-cooled fast reactor: a fast reactor which uses liquid sodium as a coolant, operating at atmospheric pressure and a temperature of

approximately 550°C [27], [32]. This reactor uses a closed fuel cycle, and therefore, has fewer waste products [28]. This reactor is conceptually similar to the lead-cooled fast reactor in that it would use a liquid metal as a coolant; however, sodium is a highly reactive metal which reacts with air and water. Therefore, additional safety features are required, such as the use of three separate coolant circuits to minimise problems with interactions with water [32]. The primary attraction to a sodium-cooled fasted reactor is that it is a mature technology, capable of having an extremely high fuel utilization and low waste products — it is considered the nearest term deployable system for waste management [28] . The lower temperatures make this technology attractive for the production of electricity but not for process manufacturing; the lower temperature means the thermal efficiency of this design is approximately 38% [28].

Considering the following factors: required research and development, UK expertise, safety, waste products, proliferation resistance, generation of heat and electricity and the ability of the UK to manufacture the components — the nuclear innovation and research office has recommended that the UK government pursue a high temperature gas-cooled thermal reactor design [28]. This would benefit from the decades worth of experience the UK has in AGR and Magnox reactors, alongside the infrastructure already in place to regulate such a reactor.

Within the future it is perhaps best to develop and demonstrate the feasibility of other reactor designs, such as the molten-salt reactor where the radiotoxicity of waste products can be greatly reduced. There could be an intermediate research and development phase where the molten-salt reactor recycles waste from other rectors, perhaps treated as a waste and reprocessing facility; disadvantages of such a setup would be in the handling and repro-

cessing of the fuel. However, it would be beneficial to the UK to expand its expertise and skills in different reactor technologies; something which is currently a hindrance to adopting other reactor technologies. In the future, designs such as the super-critical water reactor or molten salt reactor may be significantly improved, in which case the UK would have to rely on international partners to supply knowledge and expertise in designing, constructing and operating the nuclear technology.

2.2.4 The Safety Case

The use of nuclear energy is a highly divisive subject, particularly in the social and political context. The main discussion point is that of safety because any nuclear emissions are highly hazardous to both human health and the environment. To minimise risk, there is a large amount of regulation and research into the operation, maintenance, and present and future condition of a nuclear reactor power station. The hazards associated with nuclear energy form a large part of the public debate on operating a nuclear facility. Therefore, safety is paramount to enable any company to continue operating as a business. Nuclear safety concerns installations, transport of radioactive material, security and safeguards [33]. The installations criterion includes the construction, operation and decommissioning of nuclear power stations [33].

To ensure nuclear safety in the UK, the nuclear industry is regulated by the Office for Nuclear Regulation (ONR). The ONR is a member of the Western European Nuclear Regulators Association (WENRA). WENRA was initiated in 1999 as a means of providing a baseline in nuclear safety using standards from the international atomic energy agency; membership allows applicant countries to exchange knowledge and experience, enriching nuclear safety. The ONR is a government body that evaluates whether any business related to nuclear energy is safe, where a license is issued based off the ONR's evaluation.

However, it is the responsibility of the applicant to show that risks are As Low As Reasonably Practicable (ALARP) and to provide this evidence to the ONR. The demonstration that the risks are ALARP is achieved through good practice in engineering, operation and safety management [33]. This requires an understanding of the condition of the facility, presently and in the future, and also its behaviour in a variety of situations. The ONR state that [34]:

> *the onus is on the dutyholder to implement measures to the point where it can demonstrate that the costs of any further measures would be grossly disproportionate to the reduction in risks achieved by their adoption*

Regarding the graphite core, the ONR specify that the safety case should demonstrate that the graphite core is free of defects that could impair safety or that the safety of the core is tolerant to the possible defects, where a defect is defined as a deviation from the design specification. The main concern relating to graphite is if defects prejudice the delivery of a safety function, such as impeding the control rod, fuel element or coolant flow. Accordingly [35]:

> *There should be appropriate monitoring systems to confirm the graphite structures are within their safe operating envelope (operating rules) and will remain so for the duration of the life of the facility.*

In addition to measurement systems, analytic models of the core are required to enable predictions of the graphite core material properties, displacements, stresses, loads and overall condition. These analytic models can be used to simulate the behaviour of the core in a fault condition given the current condition of the core; therefore, the accuracy of the measurement

has a direct impact on evaluating the risk of operating or decommissioning a station. The complete set of the ONR's safety assessment principles can be found in [33].

Safety cases need to adopt what the ONR call a 'multi-legged' approach, using independent and diverse arguments to provide defence in depth. This PhD is to aid the safety case through the development of an inspection system which provides data that will enhance the assessment of the risks involved in operating or decommissioning EDF's AGRs.

2.3 The Non-Destructive Testing Market

The use of NDT is primarily to ensure safety but it also provides a mechanism for manufacturers to ensure the quality of their goods and to reduce waste. The NDT services and equipment market is expected to grow, with the expansion of infrastructure, increase in safety standards and the development of alternative materials such as carbon fibre composites.

Services

The global market revenue for third party NDT services in 2020 was $8.9 billion USD and is projected to grow to over $10 billion by 2025 [36]. The growth will, in part, come from a need to adhere to domestic and foreign safety standards (exports). However, growth is mainly driven by global investments in infrastructure such as rail and power generation; for example, the Chinese belt and road initiative [36]. The main industries supporting NDT are oil and gas, manufacturing and transportation; the COVID-19 pandemic significantly impacted the NDT sector, but this is only projected to be superficially (see fig. 2.6). More concerning to NDT will be the long term decline of oil and gas and the need for inspecting pipelines. This vulnerability to the dependence on oil

and gas can be seen in the recent volatility in the oil and gas market, which directly affected the NDT services sector [36]. This will be partly negated through the development of technologies such as wind turbines, which use carbon fibre composite blades, and perhaps an increase of the use of hydrogen, which will require pipelines.

Growth of NDT services is partly fuelled by companies outsourcing NDT services to remove the expense associated with expert training for operation of equipment. There are additional factors such as the internet of things, industry 4.0 and artificial intelligence which will no doubt influence the way in which NDT is carried out; for example, permanent monitoring systems can be interconnected and evaluations made using artificial intelligence, particularly useful for structural monitoring. This is an example of a push-pull situation between safety regulations and NDT, where improved NDT methods will enable stricter safety regulations, in turn supporting the NDT market. The market share according to the different NDT modalities is shown in fig. 2.7, where it can be seen that ultrasonic test equipment has the largest share.

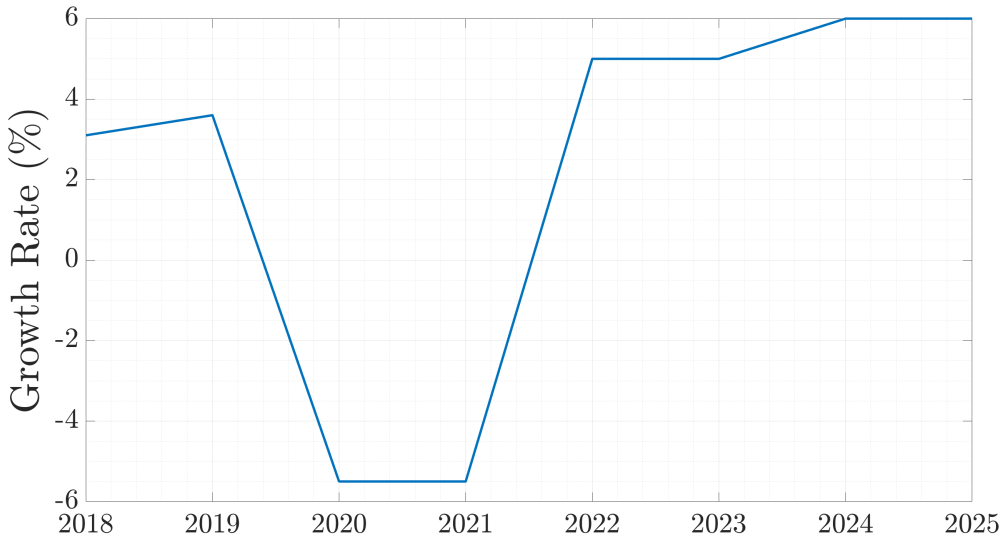

Figure 2.6: The growth of NDT services, which is forecast to bounce back from the pandemic [36]. The forecast period is from 2021–2025, where prior dates relate to the actual revenue.

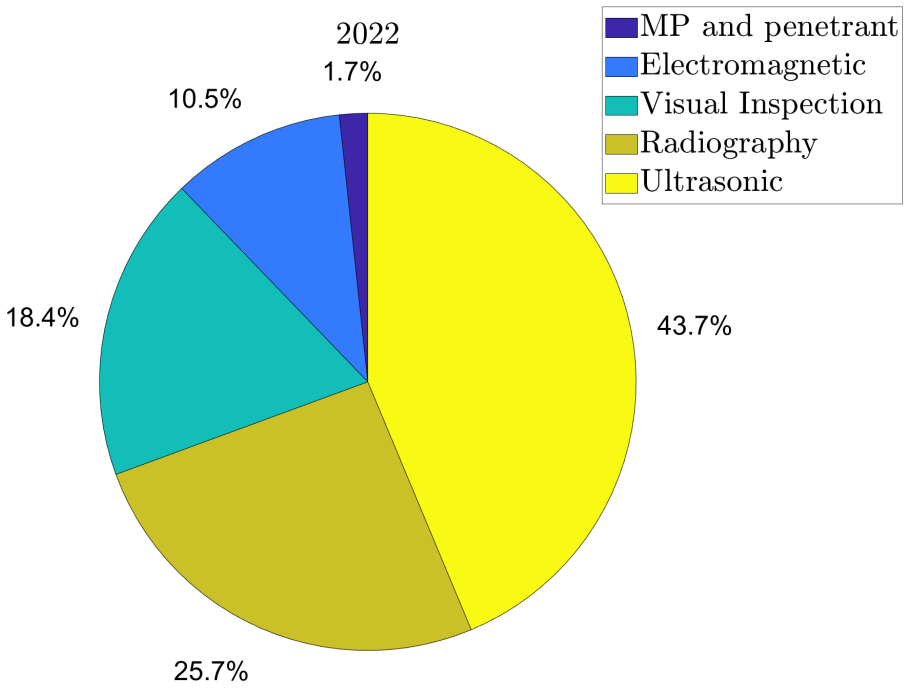

Figure 2.7: The forecast market share of different NDT service modalities in 2022 [36].

Equipment

Unfortunately, at the time of writing there has not been an update of the forecast for the equipment market since 2017; therefore, the forecast growth and revenue may not be reliable and does not reflect the circumstances of the pandemic. For this reason, only brief comments are made about the equipment market.

The total market value in 2017 was $1.955 billion USD and was expected to grow to $2.718 billion USD in 2022 [37]. Similar to the services market, growth is expected because of the global increase in safety standards. Additionally, the shift to carbon fibre composite material is given as a driver for developing new NDT techniques and equipment. Using the data from 2017, the forecast market share according to NDT modality is plotted in fig. 2.8. It can be seen that radiography has the largest share, likely due to the cost of the equipment relative to other methods..

Overall, growth is expected in NDT services and equipment, with the development of new materials to inspect, the increase in safety standards and the expansion of infrastructure. Therefore, research and development will continue to be instrumental to the NDT market.

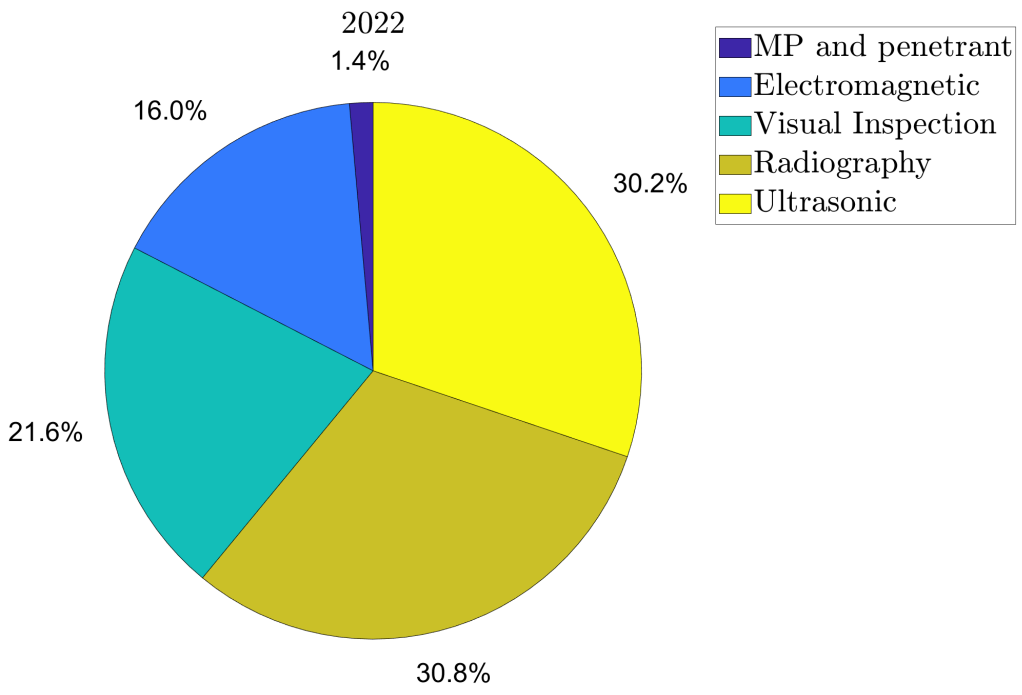

Figure 2.8: The forecast market share of different NDT equipment modalities in 2022 [36].

2.4 Advanced Gas-Cooled Reactors

In AGRs, CO_2 gas is used to transfer heat from the core to the boilers. CO_2 is used because it has a low probability of absorbing neutrons, does not become radioactive and is relatively abundant [38]. Graphite was initially chosen as a moderator since it has a high scattering cross-section and a low absorption cross-section [39]. Additionally, the graphite moderator bricks provide a path for the coolant to flow and housing for the fuel elements and control rods to slide into. The interconnections between lattice of graphite bricks is shown in fig. 2.9 and a channel specific labelled diagram in fig. 2.10. The inspection of the fuel channels is of main concern in this work.

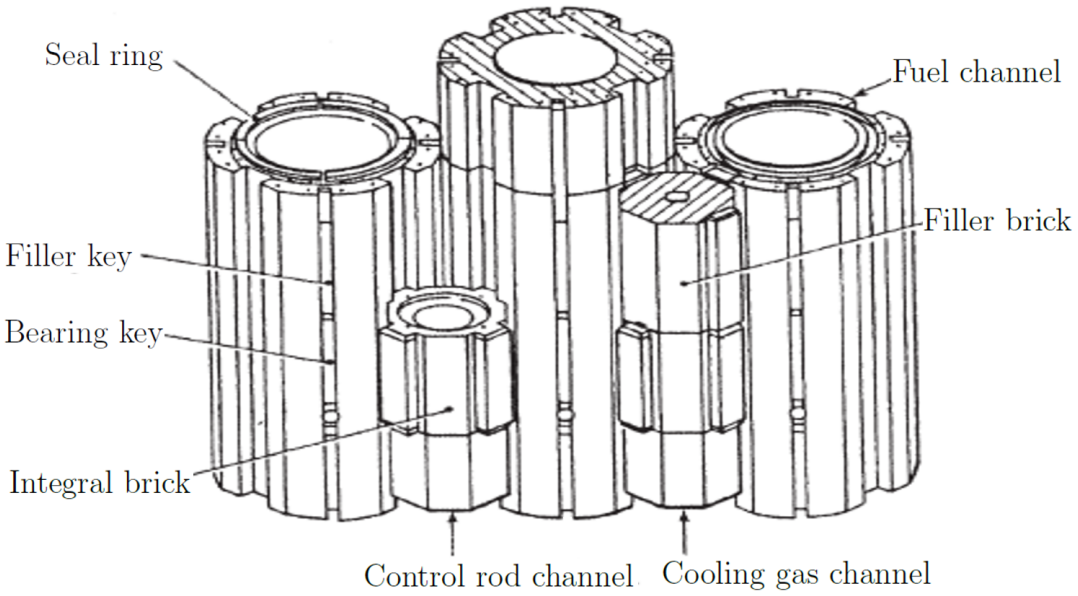

Figure 2.9: Typical lattice of graphite channels [38]. There are three channels types: fuel, control rod and cooling channels. The latter two are generally grouped together and termed interstitial channels.

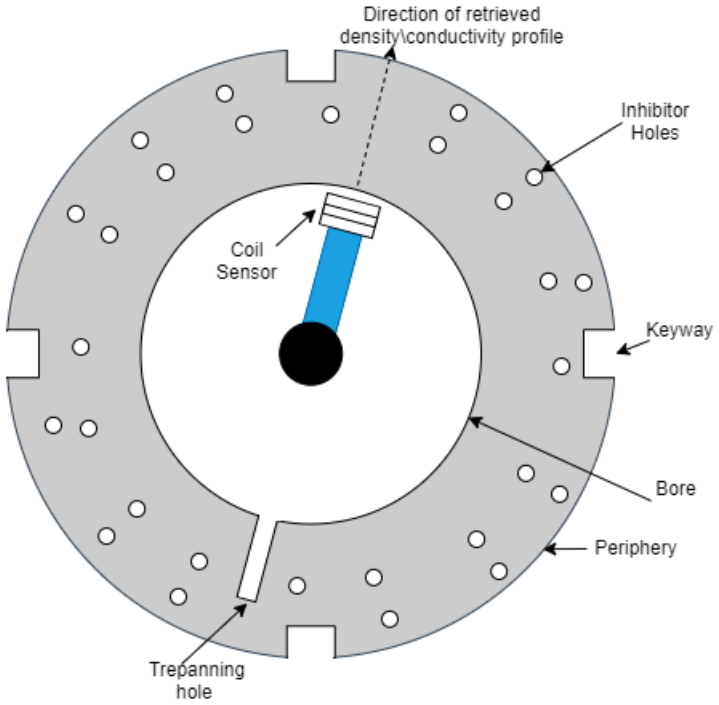

Figure 2.10: Key characteristics of graphite channel.

2.5 Nuclear Graphite

2.5.1 Manufacturing

The Magnox reactors used a type of graphite known as Pile Grade-A (PGA); this was changed in the AGR design to Gilsocarbon. Although Eatherly et al. describe the manufacturing process for PGA [40], there are only a few differences to the way that Gilsocarbon graphite was produced, namely: the type of coke (Gilsonite was used for manufacturing Glisocarbon) and that PGA was produced using extrusion and Gilsocarbon by vibration-moulding [41]. The difference between the type of coke results in a needle shaped particle in PGA and spherical particles in Gilsocarbon [41].

The petroleum coke used to make PGA is a by-product of cracking (breaking down hydro-carbon molecules) crude oil. After the initial cracking of crude oil, the heavier residuals are sent to be further refined until the desired coke is obtained. There are several degrees of freedom to how coke is manufactured at different oil refineries; therefore, there may be some variability in petroleum coke just from the manufacturing stage.

The next stage in producing graphite is calcination, where the raw coke is heat treated to remove volatile hydro-carbons and effect a shrinkage in the coke. This stage is crucial because $\approx 25\%$ of the weight of raw coke is lost in calcination. If this step is missed then it can result in extremely weak or cracked material later in the manufacturing process. Due to variability of the crystal alignment of the raw coke, because of the different refinery practices and in the material used, the calcined coke can vary in shape. An elongated particle shape results in a highly anisotropic graphite body.

The calcined coke is then bound together with coal-tar pitch, the residue left from distillation of coal tar. After mixing, the material is then sent

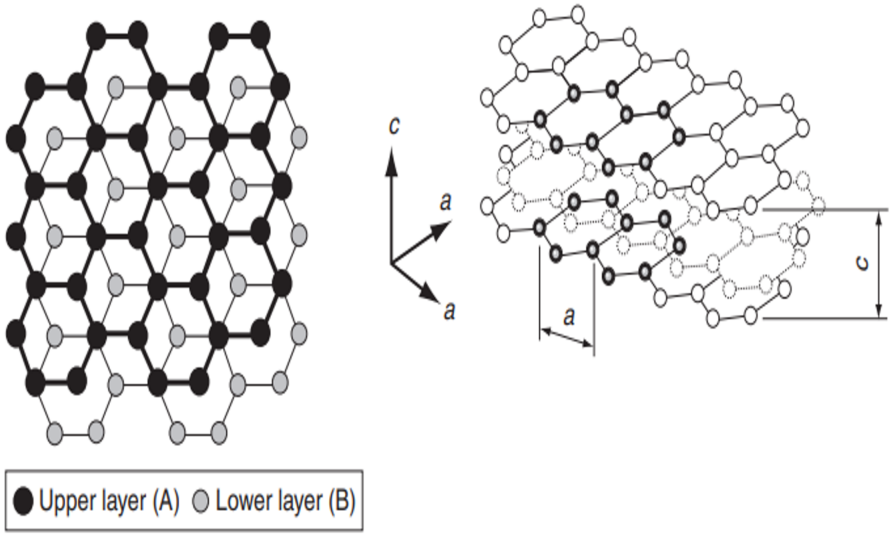

Figure 2.11: Crystal structure of graphite [42].

to be formed (in the case of PGA by extrusion), after which the formed 'green article' is left to cool down and is then baked. In extruded graphite, the filler coke particles become aligned with their long dimensions parallel to the direction of extrusion, resulting in the strong anisotropy of graphite; properties such as the thermal and electrical conductivities are greatest in the direction of extrusion. After baking, the article is then impregnated with a substance which deposits carbon in the voids upon reheating, reducing the porosity of the material. Finally the impregnated article is heat treated in order to grow the crystals inherent in the raw coke. The resultant graphite crystal has the structure shown in fig. 2.11.

The properties of the nuclear graphite depend on the raw materials used and the mode of fabrication; nuclear graphite was produced to have high thermal conductivity and resistance to radiation damage. When enriched uranium became available the irradiation levels increased, necessitating new strength and density requirements. These new requirements were met with the use of gilsonite coke as the raw material. The difference in formation and type of particle means that PGA has a much higher degree of anisotropy than Gilsocarbon. The degree of conductive isotropy of Gilsocarbon graphite

blocks is important because it is a potential source of modelling error. In this work, the graphite is modelled as isotropic.

2.5.2 Irradiation

There are two types of radiation of concern with respect to graphite: ionising gamma radiation and neutron radiation. The ionising radiation dissociates the CO_2 coolant into carbon monoxide and oxygen that chemically reacts with the graphite in a process known as Radiolytic Oxidation (RO), resulting in loss of graphite density: this is referred to as weight loss [41], [42]. Generally, weight loss is not expected to be uniform because the dosage is attenuated by the graphite [42].

The result of RO is a change in various material properties [41]. For example, Marsden et al. note that the compressive and tensile strength, Young's modulus and thermal conductivity is significantly modified by RO [42]. RO is reduced by adding a methane inhibitor; however, the amount of inhibitor is limited as it can lead to carbon deposition around the fuel assembly [43]. The methane travels through holes drilled throughout the brick to inhibit the gasification of the carbon across the graphite volume (see fig. 2.10).

Emitted neutrons in the reactor have an average energy of 2 MeV, far greater than the 60 eV required to permanently displace a carbon atom from the Crystal lattice [42]. Collisions between the neutrons and carbon atoms produce clusters of displaced carbon atoms between the basal planes [42]. This is the consequence of Fast Neutron Radiation (FNR).

2.5.3 Stresses and Keyway Root Cracking

Over the course of the reactor lifetime, the difference in radiation dosages and temperature throughout the graphite creates stresses within the bricks [41]. The stresses within the graphite are dependent on the dimensional change and material properties; for example, the coefficient of thermal expansion and Young's modulus [44].

Early in the reactor life, the bore experiences tensile stress and the periphery compressive; this reverses later in life and is called stress reversal [41][45]. Further, the stresses generated are dependent on temperature: the difference in the coefficient of thermal expansion across the brick means that the thermal stresses become significant at reactor shutdown. The normalised stresses are plotted as a function of lifespan in fig. 2.12. Prior to stress reversal, there is a high possibility of cracks initiating from the bore and post stress reversal a high possibility of cracks initiating from the keyways at the periphery [45].

The theoretical explanation for the dimensional change and the stress reversal is a key part of the safety case. Under irradiation, the graphite crystal expands in the c direction and shrinks in the a direction (see fig. 2.11) [44]. The model used by EDF explains the reversal in stress by means of accommodation porosity, in which the expansion in the c direction is unconstrained to a point; once the accomodation porosity is depleted, the expansion in the c direction dominates and leads to overall expansion [44].

Keyway root cracking occurs later in the reactor life when the periphery of the bricks experience tensile stress and compression at the bore [46]. The keyway contains sharp angles which are stress concentration sites and when placed under stress cracks are likely to propagate from these sites [47]. It is likely that once a crack is initiated it will propagate through the width of the material, terminating at either the bore or at one of the methane holes

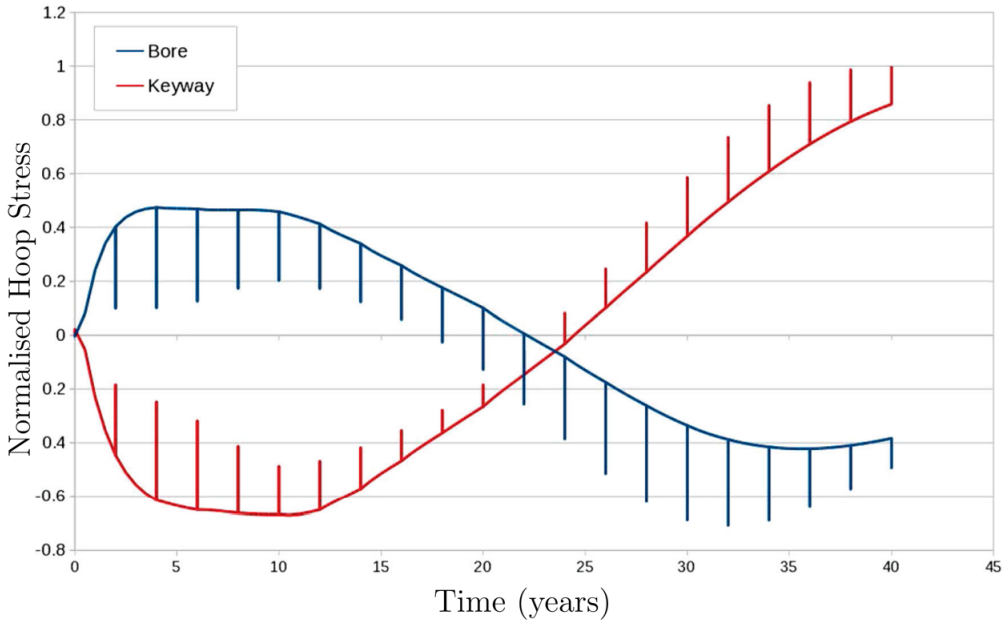

Figure 2.12: The normalised hoop (circumferential) stress at the graphite bore and periphery [41]. The spikes are thermal stress induced at reactor shutdown.

[46]. Detection of these cracks is crucial for the safety case, as the number of keyway root cracks is limited under the operational license.

2.6 Literature Review

2.6.1 Current Inspection Systems

In service inspection of the reactor is carried out using cameras, a treppaning tool, an instrument called the Channel Bore Inspection Unit (CBMU) and eddy current probes.

The variation in radiation dose results in uneven shrinkages within the graphite channels, causing channel distortion. The channel distortion is traced by measuring the diameter, ovality, channel bow and channel tilt using the CBMU. With this information the CBMU is capable of detecting the presence of significant axial cracks originating at the bore or periphery, as well as

circumferential cracks [48]. The system is deployed vertically into the fuel channels and it remains centred within the channel using two sets of four self adjusting wheels. There are linear displacement sensors placed at every 90 degrees around the device which measure the displacement from the centre.

The New in-Core Inspection Equipment (NICIE) was developed to provide a faster method of inspecting the graphite channels using cameras; it was then upgraded in the mark 2 NICIE by fitting the cameras onto a CBMU.

Inspection capabilities within the interstitial channels is much more limited. However, at several stations there is the core-restraint viewing winch, which has a control rod viewing camera and an eddy current probe; this is still a prototype arrangement. The development of the interstitial eddy current inspection would be a good area of research to develop.

The use of cameras and linear displacement sensors is limited to surface measurements. To make volumetric measurements of the core, a trepanning tool is used. The trepanning tool cuts a cylindrical sample in the radial direction of the channels. A trepanned sample does not extend the full width of the brick, only between $40 - 60$ mm; a new version recently put into service can retrieve longer samples. These samples are used to determine many material properties, such as the density of the brick, Young's modulus and the coefficient of thermal expansion. As the sample does not extend the full width of the brick, the behaviour of the graphite at the periphery is less understood than the bore.

Eddy current techniques were initially developed for the detection of cracks within the graphite core [49]. Research in eddy current inspection methods resumed in 2006 with the development of the Proof of Principle Eddy Current Tool (PoPECT). Following the initial iteration, the Prototype Eddy Current Inspection Tool (PECIT) was developed, where this tool had three eddy current probes instead of one and had the ability to perform helical scans [49].

A key issue in inspecting graphite is in discriminating variations in a signal due to cracking and electrical conductivity; however, the sensitivity of the measurement to conductivity variations demonstrated that the electrical conductivity distribution could be reconstructed, and therefore, porosity [50][51]. The post-processing capabilities were successfully extended by Fletcher and Tesfalem in reconstructing the conductivity profile of the graphite channels and in providing methods of mitigating noise originating from conductivity variations in crack detection [52], [53]. Additionally, the eddy current sensor itself was developed to have a greater sensitivity at a further depth, using an elliptical probe geometry [54]. This Prototype Elliptical Probe (PEP) was manufactured and integrated into the inspection tool, where this iteration was called the Eddy Current Inspection Tool (ECIT).

2.6.2 Eddy Current Inspection

Inversion techniques are based on either direct or iterative methods. Within eddy current inspection, there are few direct methods for the reconstruction of the conductivity profile. In [55], a direct solution was obtained by linearising eddy current problem via the Born approximation and explicitly solving a decoupled Laplace-Fourier transform in the time domain to recover conductivity. This technique considered a specific coil geometry, and it is not clear how this translates to complex elliptical geometries. Tamburrino et al. developed a direct inversion technique for imaging anomalous regions [56]; however, this technique does not recover the depth profile of electrical conductivity. Depth profiling using optimisation techniques has been applied for some time and in different modalities [57][58]; therefore, there is a large amount of experience and knowledge currently in the literature. It has been previously demonstrated that non-linear iterative algorithms can be used to reconstruct the conductivity depth profile of a nuclear graphite brick [51]. However, to date,

the algorithms used in inverting nuclear graphite data have been limited in scope: inversion using Gauss-Newton (GN) type directions continued up until the work of Tesfalem [52], where trust region methods were employed instead.

The objective function used for graphite inversion has always considered the absolute error. Beyond the graphite application, the absolute error has been dominantly used in electromagnetic tomography, where some examples are: electrical impedance tomography for medical imaging [59], depth profiling in inductance tomography [60][61] and permittivity reconstruction in electrical capacitance tomography [62]. However, the inductance spectrum spans orders of magnitude, meaning the residual error is typically much larger at higher frequencies than at low frequencies. Therefore, the higher frequency information dominates the step direction. Instead, the relative error can be used, where this places an equal weighting to each spectral component: the step direction is then equally informed by all the spectral data [63]. However, the Signal-Noise Ratio (SNR) may be poorer at low frequencies, and therefore, the relative error objective function may be more susceptible to noise.

Iterative reconstruction algorithms require a forward model. For problems which consider simple geometries there are analytic solutions [64][65] but there is no analytic formulation of the forward model for the complex graphite geometry; therefore, a Finite Element Model (FEM) is used. However, with the development of ML techniques, the forward and inverse model can be found empirically from data. More generally, outside of the nuclear graphite application, ML techniques in eddy current inspection have been investigated for several decades for inverse modelling [66]–[69]. Additionally, they have been used in forward modelling, where an example is using a ML forward model surrogate in place of a FEM [70] [71]. Other applications considered the theory developed in neural networks to solve finite element equations by re-formulating the equations in a form in which they can be represented by a neural network [72][73][74]. However, in safety critical systems (such as

inspecting nuclear graphite) the use of ML methods must be supported by a deductive method embedded in the physics.

A forward model must accurately describe the experimental problem and be properly calibrated. In electromagnetic inspection of the graphite core, the FEM has been previously calibrated using a grid search of the coil diameter [52] and by a transfer function approach [53]. In the former technique, the time to tune the FEM increases rapidly with the number of coil parameters, and the latter approach assumes that the calibration factors are invariant to conductivity. Instead, optimisation techniques can be used to tune the coil parameters. Optimisation techniques have been previously used to calibrate an eddy current model; in [75] the forward model was calibrated by minimising the difference between the measured and simulated responses by solving a least squares problem via the update of model lift off. There are numerous constraints that must be adhered to in calibration, such as lift off being strictly positive and non-zero; therefore, a constrained optimisation algorithm must be used.

A crack can be viewed as a sharp change in density, and therefore, the depth profiling algorithm may be capable of imaging such a defect. Research has been published in comparing penalty terms in an iterative algorithm for imaging sharp contrasts in conductivity [62]. However, the resolution of the FEM may not be fine enough to resolve cracks. The FEM used in this thesis is discretised only in the radial direction and not azimuthal. Alternatively, direct methods can be used; for example, using the monotonicity principle or ML algorithms. In the monotonicity principle an anomalous region can be rapidly imaged using an impedance matrix characterised by a set of coils [56][76]. Inferential techniques offer an alternative methodology for crack detection. ML classification algorithms have been widely used for several decades in eddy current NDE, where examples include support vector machines [77] and neural networks [78][79]. Crack detection within a nuclear graphite brick has histori-

cally used averaging techniques to remove the effect of conductivity variations in the measurement; averaging can be seen as a feature extraction technique in which it is assumed that the underlying conductivity profile is derived from the same probability distribution in the azimuthal direction. Other feature extraction techniques have been used on eddy current data: Principal component analysis (PCA) [80], the Hilbert transform [81], Independent Component Analysis (ICA) [82] [77] and k-means clustering on Lissajous curves [83].

2.6.3 Other Non-Destructive Testing Methods

The eddy current response can be measured using a frequency sweep or transient response, where analysis can be performed in the time or frequency domain. In pulsed eddy current, attributes such as the time to peak height, the peak height, the rising point time and the zero-crossing point can be used to determine defects, lift off and material thickness [84], [85]. A key disadvantage of pulsed eddy current techniques is that the energy in the excitation current is spread across the spectrum, and therefore, SNR is generally lower than in frequency sweeping.

If ultrasonic methods are used, they must be non-contact and capable of operation in radioactive environments. Electromagnetic acoustic transducers can be used to induce ultrasonic waves in a radioactive environment but these require sufficient electrical conductivity to produce a strong Lorentz force. Generally, ultrasonic attenuation is much higher in graphite compared with steel; it has been found that there is a loss of 0.93 dB/mm, whereas for stainless steel 304 it is 0.2 dB/mm [86]. Furthermore, ultrasonic techniques may be unsuitable because of the graphite geometry; the methane holes may considerably scatter the ultrasonic signal.

2.7 Chapter Summary

- The electricity market demand is set to dramatically increase because of the net zero carbon emissions target. Additionally, the electricity is to be sourced from low carbon sources, where nuclear is expected to be a key part of the net zero strategy.

- The commercial power stations built in the UK have all used graphite as the moderator and a gas-coolant to transfer heat; this will be unchanged for the next generation, which will use graphite and a helium gas coolant, operated at a higher temperature for improved efficiency.

- Irradiation of the graphite moderator results in weight loss and dimensional changes within it. The irradiation of graphite changes its material properties, such as the co-efficient of thermal expansion, Young's modulus and density. Understanding and measuring weight loss is important for modelling and the safety case.

- Direct solutions to the inverse problem are not available for the graphite inversion problem; although there are techniques such as the monotonicity principle (for anomaly detection) and the decoupled Fourier-Transform technique (for depth profiling), these would need either a re-design of the sensor system or generalising to the graphite inverse problem.

- Iterative reconstruction algorithms typically consider the absolute relative error. For the eddy current problem, the relative error may be more suitable. Further, search directions have historically used second order methods, where there is no broader comparison of search directions for graphite inversion.

- Calibration of the forward model is essential in iterative algorithms.

Typically, experimental data is mapped to simulated data using a transfer function approach or some other scaling factor. Alternatively, the forward model can be tuned by adjusting some model parameter, such as coil diameter, until there is sufficient agreement between the experimental and simulated data. The search for the ideal model parameters can be done using a grid search but this is slow and computationally expensive. Instead, optimisation algorithms can be used to adjust the model parameters, which are faster to converge and are more feasible with a larger number of model parameters.

- To detect the presence of cracks, previous attempts used averaging to minimise the effect of conductivity variations in the measurement. This assumed that the measurement is a linear function of conductivity and crack height; moreover, it assumes that conductivity variations are derived from the same probability distribution. This is perhaps sufficient, but there may be other feature extraction and classification algorithms that are better suited for crack identification.

Chapter 3

Inductance Spectroscopy and the Inverse Problem

3.1 Overview of Chapter

The chapter opens with a review of basic electromagnetic theory, which is necessary for describing the classes of electromagnetic tomography and the material affects on electromagnetic signals; for example, the skin effect. Inductance Spectroscopy (IS) is an ideal imaging modality for the nuclear graphite application because it requires only passive components in the sensor and can be operated at some distance from the point of inspection. The principle in IS is to measure the multi-frequency eddy current response of the material under test, where different frequencies correspond to different eddy current distributions.

In the following sections, the need for a numeric forward solver is explained. In certain applications, there are analytic formulations relating electrical conductivity to mutual inductance [65][64]; however, the complex geometry of the graphite makes an analytic solution impractical. Instead, Finite Element Analysis (FEA) is used to solve the problem numerically.

To solve the inverse problem, optimisation techniques are used and are introduced in this chapter. There are other methods, such as the HyperBand [87] algorithm or grid search, which are global optimisers but these typically require many more iterations to find a solution, and therefore, are infeasible with a FE forward model. There are a wide range of iterative solvers, where those that are introduced are based on Newton or Gradient Descent (GD) directions.

The Multi-Variable Polynomial Regression (MVPR) algorithm is described, which is a simple inferential technique mapping the input and target data using a polynomial function approximation. There are only two hyper-parameters in the model and it is, therefore, a simple regression technique that can be easily optimised. It is useful for bench marking more powerful regression techniques such as neural networks.

3.2 Fields in Matter

An electric field strength $\vec{E}$ is, by definition, the electric force experienced by a unit charge. When an electric field is applied on free charge a current arises, with a current density $\vec{J_f}$. In the time domain the current density is defined as eq. 3.1 and in the frequency domain eq. 3.2, where σ is the electrical conductivity [88].

$$\vec{J_f}(t) = \int_{-\infty}^{\infty} \vec{E}(\tau)\sigma(t-\tau)d\tau \tag{3.1}$$

$$\vec{J_f}(\omega) = \vec{E}(\omega)\sigma(\omega) \tag{3.2}$$

When an electric field is applied on a material there is a polarisation of bound charge. The polarisation vector, $\vec{P}$, is defined as the dipole moment per unit volume [89] and it encapsulates the polarisation of charge within an atom.

The electric flux density, $\vec{D}$, is then given by eq. 3.3 [89]. The constant ϵ_0 is the permittivity of free space.

$$\vec{D}(\omega) = \epsilon_0 \vec{E}(\omega) + \vec{P}(\omega) \tag{3.3}$$

A magnetic analogue of the electric polarisation vector is the magnetisation vector $\vec{M}$. The magnetic flux density is then given by eq. 3.4, where $\vec{H}$ is the magnetic field strength and μ_0 is the permeability of free space.

$$\vec{B}(\omega) = \mu_0(\vec{H}(\omega) + \vec{M}(\omega)) \tag{3.4}$$

These two constitutive relations can be condensed into eq. 3.5 and 3.6, where the polarisation and magnetisation vector have been subsumed. The constants μ, ϵ and σ are characteristic of the material under test and determine the resulting fields $\vec{B}, \vec{D}$ and $\vec{J}_f$ within it. In graphite it is expected that $\mu = \mu_0$ and $\epsilon = \epsilon_0$; the effects of the magnetic and electric polarisation are minimal.

$$\vec{B}(\omega) = \mu \vec{H}(\omega) \tag{3.5}$$

$$\vec{D}(\omega) = \epsilon \vec{E}(\omega) \tag{3.6}$$

To understand the effect of the material on incident electric and magnetic fields, it is helpful to use the monochromatic plane wave (more complicated wave descriptions can be made from this using Fourier analysis). The monochromatic wave has an electric and magnetic field of the form eq. 3.7 and 3.8 [90].

$$\vec{E} = \left\| \vec{E} \right\| \left(\frac{e^{j(\vec{k}\cdot\vec{r}-\omega t)} + e^{-j(\vec{k}\cdot\vec{r}-\omega t)}}{2} \right) \tag{3.7}$$

$$\vec{H} = \left\| \vec{H} \right\| \left(\frac{e^{j(\vec{k}\cdot\vec{r}-\omega t)} + e^{-j(\vec{k}\cdot\vec{r}-\omega t)}}{2} \right) \tag{3.8}$$

The wave vector $\vec{k}$ describes how the wave spatially changes. In dispersive matter, the amplitude of the wavevector can be written as eq. 3.9 [88].

$$\left\|\vec{k}(\omega)\right\| = \omega\sqrt{\mu(\epsilon_0 + j\frac{\sigma(\omega)}{\omega})} \tag{3.9}$$

This can be expanded further by using the Drude model of electrical conductivity: $\sigma(\omega) = \frac{\sigma_{DC}}{1-j\omega\tau}$, where τ is the mean-free time of collisions. At low frequencies $\lim_{\omega\to 0}\sigma(\omega) = \sigma_{DC}$ and in a conductor such as graphite $\epsilon_0 \ll \frac{\sigma_{DC}}{\omega}$; therefore, $\left\|\vec{k}\right\|$ is approximately imaginary and the electric and magnetic fields are damped. The damping term is given by $e^{-\delta}$, where δ is called the skin depth. However, this is considering a monochromatic plane wave; the fields emanating from the coil and the corresponding penetration depend on the coil geometry [91][92].This skin depth equation is valid for any frequency of a plane wave transmitted into a conductive halfspace but may not be representative of the decay of more complex fields. Nevertheless, the monochromatic wave decay is useful for understanding how the material properties affect the decay of the fields within a conductor. The skin depth of a plane wave transmitted into a halfspace is given by eq. 3.10.

$$\delta = \sqrt{\frac{1}{\mu\sigma_{DC}\pi f}} \tag{3.10}$$

Defining a new term, the plasma frequency ω_p, at high enough frequencies the wavevector can be approximated by

$$\left\|k(\omega)\right\| = \omega\sqrt{\mu\epsilon_0(1 - \frac{\omega_p^2}{\omega^2})}$$

When the frequency of the electromagnetic wave is above the plasma frequency of the material, the wave vector is essentially real and the medium becomes transparent to the wave [88]. This is what defines classes of tomography. High penetrating frequencies (e.g. x-ray and gamma rays) are used in

hard field tomography which generate high resolution localised images, displaying sharp contrasts in the attenuation of the wave/photon intensity. Hard field tomography typically requires the source and detector either side of the material under test (but not exclusively, Compton scattering is sometimes used in x-ray imaging). In soft field tomography, a low frequency range is used, typically below 10MHz. In this regime the displacement term in Ampere's equation can be ignored, the effect being that the oscillating electromagnetic fields no longer conform to the wave equation, but to the diffuse equation. The key advantage of using soft field tomography is that both the source and detector can sit on the same side of the material and it only requires passive components. The large amount of high frequency ionising radiation in the core would destroy active components unless they are radiation hard. For the inspection of the graphite channels in this project, soft field tomography is employed with IS as the imaging modality.

3.3 Depth Profiling

Depth profiling with IS has been used in a range of applications, such as imaging electrical conductivity and permeability in metal [60], [93], thickness evaluation of turbine thermal barrier coating [94] and in nuclear graphite inspection [51]. In all of these examples, non-linear optimisation techniques are used to solve the inverse problem iteratively. Iterative techniques are typically required since there are no analytic inverse solutions available in this modality. Norton et al. [61] used an analytic forward model and the Levenberg-Marquart (LM) optimisation algorithm for image reconstruction of a "pin in sleeve" problem, where a cylindrical metal core was surrounded by a sleeve of differing conductivity.

The problem to be solved is the spatial reconstruction of electrical con-

ductivity. Ideally, this would be a 3D reconstruction of the continuous profile. This could be done by describing the conductivity distribution using a set of basis functions [51], [52]. The parameters of the function can then by tuned rather than the conductivity itself. For example, consider $\sigma(r, \theta, z) = ar^3 + br^2 + cr + d$, the conductivity distribution in cylindrical co-ordinates with no azimuthal or height variation and only radial; this axisymmetric conductivity distribution can be described with only four parameters. An example of this is given in fig. 3.1. However, there is difficulty in choosing what the functional should be, and in safety critical systems the justification of the functional. Instead, the conductivity profile can be modelled as piecewise constant and the conductivity of each region solved for directly. To reduce the complexity further, the conductivity can be assumed to vary in one direction only and this is the approach used; the conductivity is assumed to vary with depth only as a piecewise constant function. The conductivity of each layer is tuned until there is sufficient agreement between the model response and measured response and such an example is given in fig 3.2.

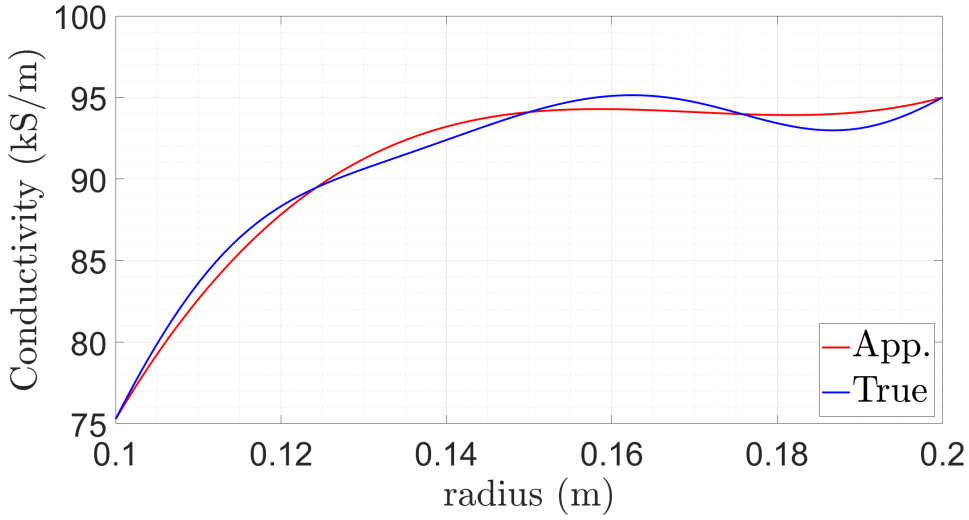

Figure 3.1: An example of what a continuous depth profile reconstruction might look like.

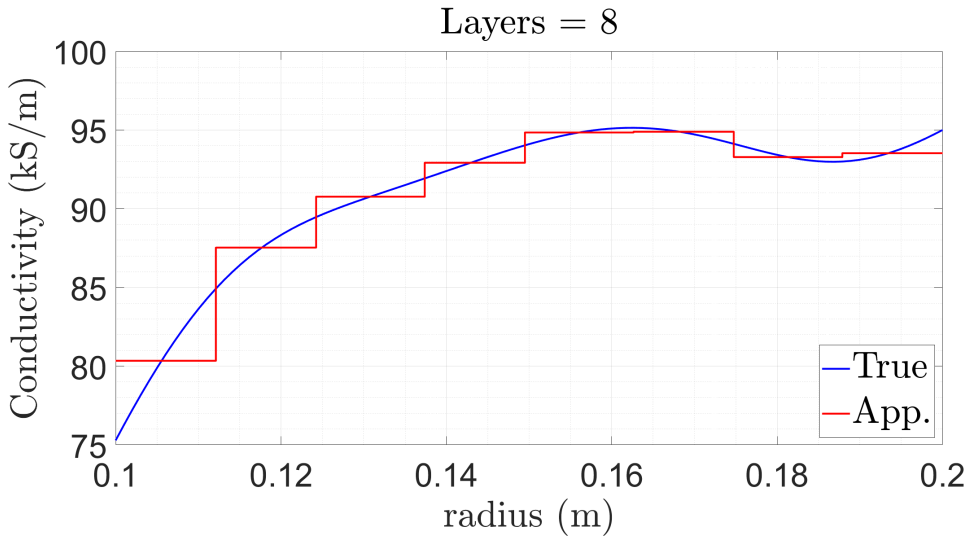

Figure 3.2: An example of what a discrete depth profile reconstruction might look like.

3.4 Eddy Current Formulation

Electromagnetic problems can be formulated in a number of different ways. The appropriate formulation depends on the problem and solving method [95] [96]. For example, the V formulation is used when no magnetic field is present [97], where only two of Maxwell's equations are required and the differential equation can be expressed in terms of the electric potential: this is not appropriate for the eddy current problem. In the solver used in the thesis, the A formulation is used to encapsulate the eddy current problem [97], where Maxwell's equations are expressed using the magnetic vector potential.

The magnetic field can be written in terms of the magnetic vector potential, $\vec{A}$, with

$$\nabla \times \vec{A} = \vec{B}$$

Substituting this into Ampere's law and ignoring the displacement current yields

$$\nabla \times \frac{1}{\mu} \nabla \times \vec{A} = \vec{J_f}$$

The free current density can be separated into the current density due to the applied magnetic field, $\vec{J_{eddy}}$, and the source current, $\vec{J_{source}}$. Further, using Faraday's law the electric field can be written as $-\frac{\partial \vec{A}}{\partial t}$. Combining these gives

$$\nabla \times \frac{1}{\mu} \nabla \times \vec{A} = \vec{J_{source}} - \frac{\partial \vec{A}}{\partial t} \sigma$$

For a single frequency excitation, this simplifies to eq. 3.11.

$$\nabla \times \frac{1}{\mu} \nabla \times \vec{A} + \sigma j \omega \vec{A} = \vec{J_{source}} \tag{3.11}$$

In [96], analytic solutions to this equation are given for a few simple geometries, such as a straight filament parallel to a conductive slab. The well

known paper written by Dodd and Deeds provides an analytic solution of the magnetic vector potential for an axi-symmetric air cored coil above a two conductor half space [64], where this was extended upon by Theodoulidis for a continuous depth profile of a conductive halfspace [65]. Kriezis et al. state that problems that have inhomogeneous material parameters, non-linear media or complex geometries cannot be solved analytically except in simple two-dimensional problems. Thus, to solve eq. 3.11 a numeric method is needed. In this work the FEA package COMSOL® is used.

3.5 Numeric Solution

The COMSOL® package uses FEA to numerically solve a wide range of physics problems. In FEA, the problem domain is discretised into sub-domains referred to as elements. Each element conforms to the given differential equation of the problem and is solved according to the boundary and initial values.

In two dimensions, a single triangular element can be used to approximate the magnetic vector potential within the element, $\vec{A}^e$, where such a triangular element is show in fig. 3.3. Using a set of basis vectors, $\vec{N}_j^e$, the magnetic vector potential over the element can be written as

$$\vec{A}^e = \sum_{j=1}^{3} a_j \vec{N}_j^e$$

The vector $\vec{N}_j^e$ is a vector field and is called the vector shape function for edge j over the element e. The scalar term a_j corresponds to the line integral of the field along edge j [98].

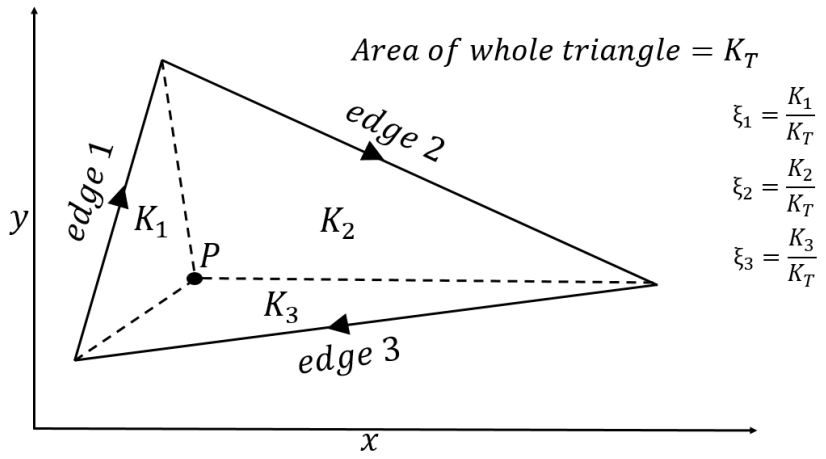

Figure 3.3: A 2-D finite element.

The co-ordinates of a point within the triangle can be described as weightings of the vertices

$$x = \xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3$$

$$y = \xi_1 y_1 + \xi_2 y_2 + \xi_3 y_3$$

with the condition

$$\xi_1 + \xi_2 + \xi_3 = 1$$

If a point P sits in the triangle and lines are drawn from P to each vertex, the triangle would be split into a further three sub-triangles. The ratio of the area of a sub-triangle, K, to the area of the entire triangular element is given by ξ_k. The vector shape function, $\vec{N_j}$, using the area co-ordinates is given by eq. 3.12, where L_j is the length of the edge [99][100]. These are plotted in fig. 3.4.

$$\vec{N_j} = (\xi_i \nabla \xi_k - \xi_k \nabla \xi_i) L_j, \quad j \neq i \neq k \tag{3.12}$$

The vector in eq. 3.12 has a number of useful properties. It has no tangential component along the other edges; therefore, if P is on edge 1, $a_2\vec{N_2^e}$ and $a_3\vec{N_3^e}$ have no tangential components, and therefore, the $\vec{A}$ field along edge j can be written as $a_j\vec{N_j^e}$ [98]. Furthermore, edge element shape functions are well suited for modelling electromagnetic interface conditions because only tan-

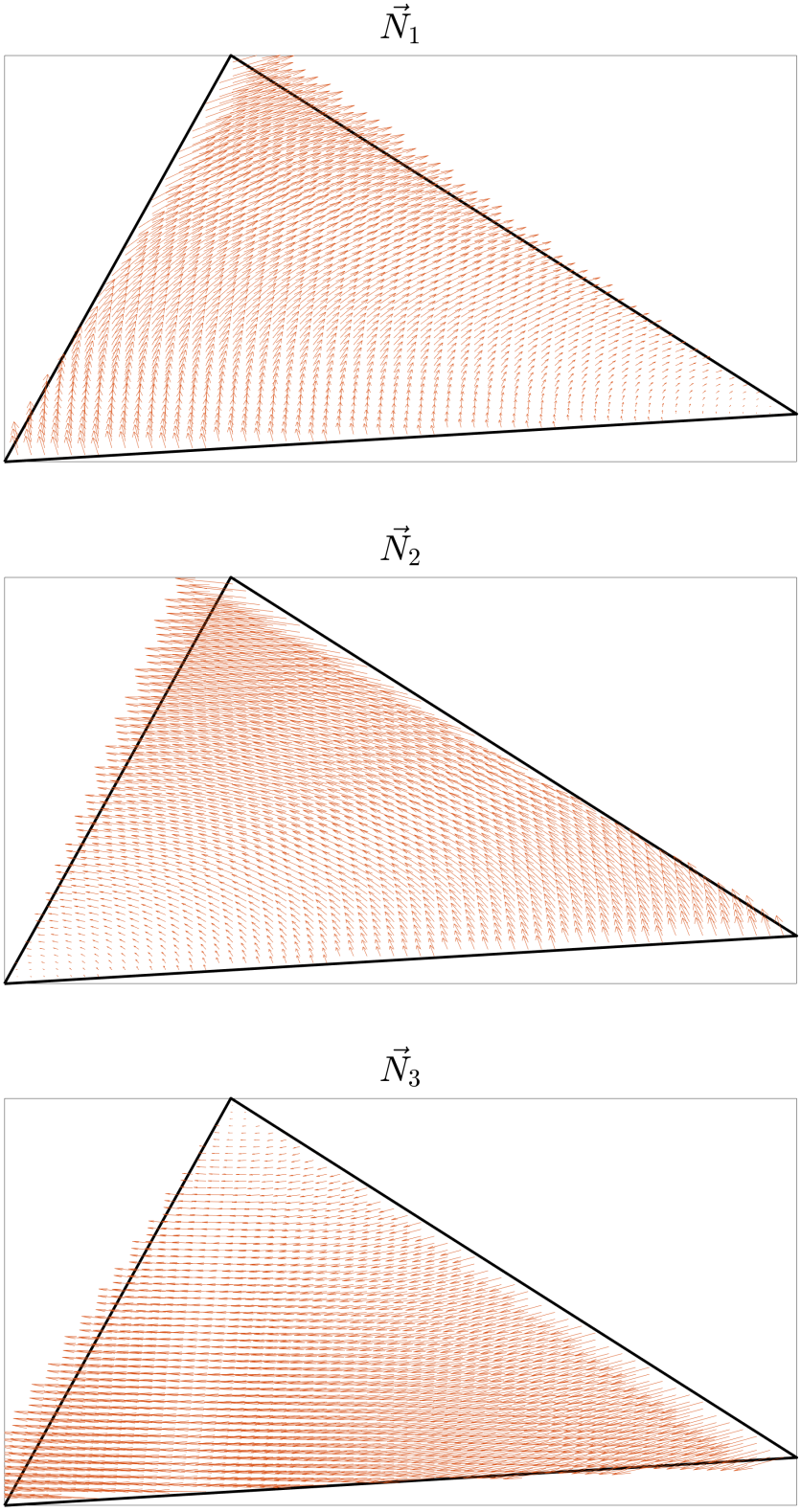

Figure 3.4: The shape functions using eq. 3.12. These were computed numerically using the finite difference method to compute the gradient.

gential continuity is enforced across the element boundaries but continuity of normal components is not [98].

There are a selection of elements available in COMSOL®, such as the tetrahedron, pyramid, prism and hexahedron. The vector shape functions for all of these elements are defined in a similar way as the 2-D triangular element already discussed, the only difference is volume co-ordinates are used instead [99].

To solve eq. 3.11 there are a variety of methods. In COMSOL® a weighted residual method is used, where

$$\nabla \times \frac{1}{\mu} \nabla \times \vec{A}_{app}(x,y,z) + \sigma j \omega \vec{A}_{app}(x,y,z) - \vec{J}_{source} = \vec{e}(x,y,z)$$

The approximate solution is found by solving $\int \vec{e}(x,y,z) \cdot \vec{v}(x,y,z) dV = 0$, where $\vec{v}(x,y,z)$ is a vector field which weights the error [101]. This can be understood from a geometric point of view as minimising the projection of the error onto the vector field $\vec{v}$. In FEA, both the weight function and A field are expressed in terms of basis functions. In COMSOL®, the Galerkin method is used [102] where the basis functions for both A and the weight function are the vector shape functions, N_j [101].

COMSOL® has several linear and non-linear numerical solvers of differing types, either direct or iterative. The numeric solvers differ in the amount of memory usage, the speed of convergence, stability and the computation time. If a solver is not selected COMSOL® automatically sets what it thinks is the best solver for the problem defined.

3.6 Mutual Inductance

The voltage, ψ, produced over a loop is given by 3.13, the integral form of Faraday's law.

$$\psi(t) = -\frac{d}{dt}\int_S \vec{B}\cdot\vec{n}ds \tag{3.13}$$

Electric circuits that are not directly connected, are coupled by a mutual magnetic and electric field between them. The induced voltage in one circuit due to a current (I) in another can be expressed as a lumped component model and this is given in 3.14. The mutual inductance, $\mathfrak{M}$, is a constant that describes how much of the magnetic field is linked between one circuit and another, independent of the driving current.

$$\psi(t) = \mathfrak{M}\frac{dI}{dt} \tag{3.14}$$

The measurement made is always the open circuit induced voltage across a receive coil. In air, the total flux linked depends on the geometry of the problem, orientation of the coils, the frequency and current of the driving signal and the conductivity, permeability and permittivity of the air. In the presence of a graphite target the same is true, but with the addition that now the coupled magnetic field is altered by the eddy currents induced within the material itself. The change in the magnetic field alters the voltage measured on the receiver.

When the sensor coils are in the vicinity of the graphite target, the mutual inductance variable is a complex quantity because of eddy current losses, where the real part pertains to the phase change and imaginary the eddy current losses. Furthermore, $\mathfrak{M}$ becomes a frequency dependent quantity. This is encapsulated by eq. 3.15, which considers a single frequency excitation.

$$\psi(\omega) = j\omega\mathfrak{M}(\omega)I(\omega) \tag{3.15}$$

The difference between the mutual inductance between two coils in the presence of graphite and in air, is useful in practice for mitigating background noise; to a first order approximation, the differential measurement is due entirely to the presence of the target.

$$m = (\mathfrak{M}^{\mathrm{c}}_{\mathrm{Rx}}) - (\mathfrak{M}^{\mathrm{air}}_{\mathrm{Rx}}) \tag{3.16}$$

In the same manner, a gradiometer can be used to produce a response which is ideally due to the presence of graphite: when the gradiometer is balanced the measurement in air is (ideally) equal to zero. However, gradiometer coils are generally not perfectly balanced, and therefore, the gradiometer response can be differenced with air to eliminate residual mis-balance. The gradiometer measurement is given by eq. 3.17. The measurement is broadly described as the Differential Mutual Inductance (DMI) for both eq. 3.16 and 3.17; which measurement is being referred to will be explicitly stated.

$$m = (\mathfrak{M}^{\mathrm{c}}_{\mathrm{Rx}_1} - \mathfrak{M}^{\mathrm{c}}_{\mathrm{Rx}_2}) - (\mathfrak{M}^{\mathrm{air}}_{\mathrm{Rx}_1} - \mathfrak{M}^{\mathrm{air}}_{\mathrm{Rx}_2}) \tag{3.17}$$

In practice, the air measurements are taken using a nylon ring of the same diameter as the graphite channel. A nylon ring is used because the gradiometer probes are deployed from an inspection tool made of metal, and therefore, the response is a function of distance from the probe. Ideally, the effect of the metal casing would be minimal or would be mitigated through coil balancing; unfortunately, the coils were not balanced in the inspection tool with the cables attached, and therefore, there is a reliance on the subtraction of the free space response to reduce the effect of casing etc. to at least a first order.

3.6.1 Gradiometer Coil

The PECIT and the recently deployed ECIT, deploy a gradiometer sensor. The PECIT gradiometer was symmetric and flat, meaning that the two receive coils were geometrically symmetric about the transmitter, having the same number of turns and having no curvature. A flat and symmetric gradiometer coil is simple to balance but has the consequence that it does not optimally fit against the curvature of the graphite. The optimal fit would have the transmitter and receiver as close to the graphite as possible, to maximise the transmitted and received magnetic fields. Furthermore, a key disadvantage of a flat coil is that a larger diameter would necessitate a larger lift-off, creating a dilemma in increasing coil sensitivity.

The limitations of the PECIT sensor were studied by Tesfalem et al. and it was shown that an asymmetric and curved gradiometer configuration can be tuned for increased coil sensitivity to eddy currents at larger depths; the initial study showed that an optimised asymmetric probe had an average of 43.7% sensitivity improvement compared with the PECIT probe, and five times higher sensitivity at a maximum depth [103]. This work led to a deeper study into asymmetric gradiometers for the nuclear application and the increased sensitivity to deeper cracks [54]. In both of these studies, the gradiometer size was constrained, such that the maximum diameter corresponded to negligible effects of the steel casing of the inspection tool — the larger the diameter of the coil, the more susceptible to field interactions from the steel casing. Further, the consequence of a larger coil is a trade off between resolution and penetration. The differences between the induced eddy currents of the two different coils is shown in fig. 3.5.

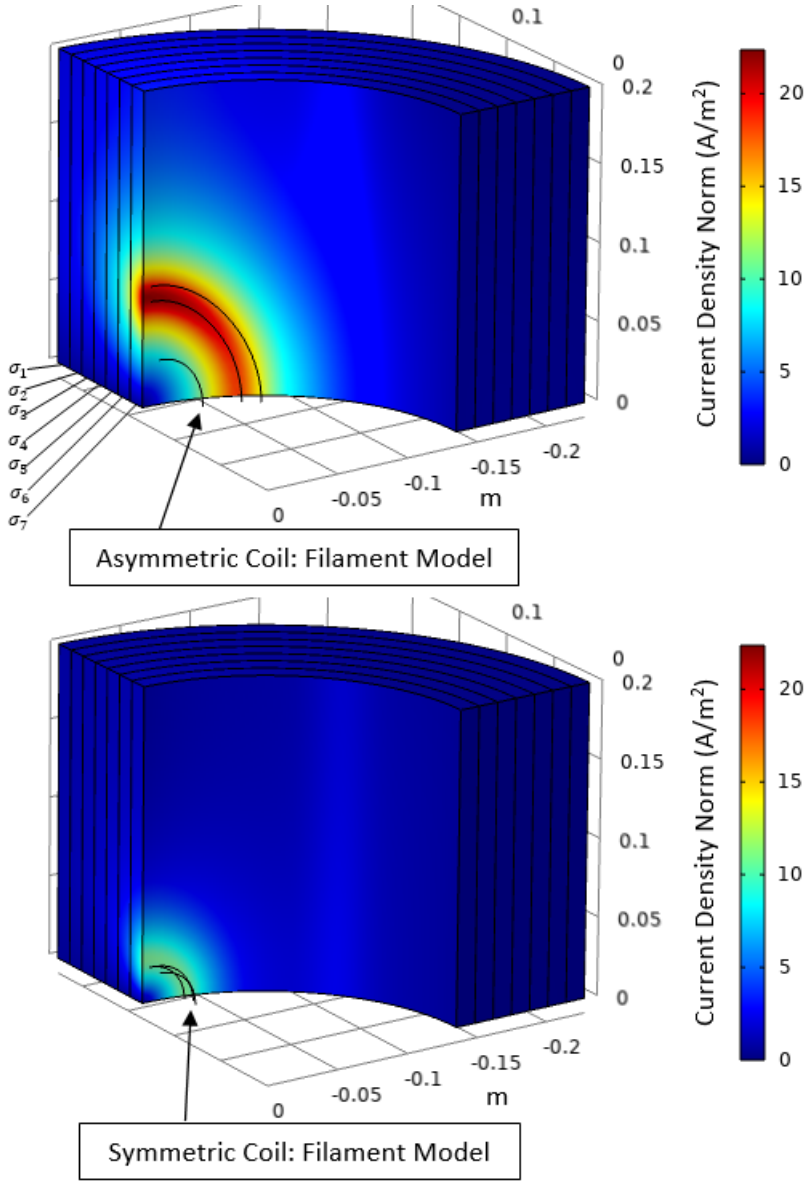

Figure 3.5: The induced current density for a single filament in the asymmetric and symmetric probes. The colour maps are on the same scale. It can be seen in the larger probe that there is a much stronger field produced and at a greater depth but over a greater volume. Ideally, the sensitivity would be of the same magnitude in each layer and the results in [103] shows the asymmetric sensor is closest to this.

3.6.2 Sensitivity Equations

Theory

To compute the coil sensitivities the perturbation method can be used. In the perturbation method, the sensitivities are computed numerically using a technique such as the forward difference method [60]. In COMSOL® the measurement spectrum is computed in a parametric sweep; therefore, to compute the sensitivities this would require at least $n+1$ executions of the forward problem, where n is the number of layers. Thus, it is computationally expensive.

The sensitivity of the receiver coil voltage to some electromagnetic variables can be found using Tellegen's reciprocity theorem [104]. Tellegen's theorem is a statement of quasi-power for two topologically identical systems; thus, it is an extremely general reciprocity theorem. The power flow into a closed volume can be found as eq. 3.18 by integrating the Poynting vector over a closed surface [105].

$$\oint_S (\vec{E} \times \vec{H}) \cdot \vec{n} dS = -\oint_V (\vec{E} \cdot \frac{\partial \vec{D}}{\partial t} + \vec{H} \cdot \frac{\partial \vec{B}}{\partial t}) dV - \oint_V (\vec{J} \cdot \vec{E}) dV \tag{3.18}$$

Tellegen's theorem is given as eq. 3.19 [106]. The subscripts A and B denote the field quantities for two different, but topologically identical systems, illustrated in fig. 3.6.

$$\begin{aligned} &\oint_S (\vec{E}_B \times \vec{H}_A) \cdot \vec{n} dS - \oint_S (\vec{E}_A \times \vec{H}_B) \cdot \vec{n} dS \\ = \quad &-\oint_V (\vec{E}_B \cdot \frac{\partial \vec{D}_A}{\partial t} + \vec{H}_A \cdot \frac{\partial \vec{B}_B}{\partial t}) dV - \oint_V (\vec{J}_A \cdot \vec{E}_B) dV \\ + \quad &\oint_V (\vec{E}_A \cdot \frac{\partial \vec{D}_B}{\partial t} + \vec{H}_B \cdot \frac{\partial \vec{B}_A}{\partial t}) dV + \oint_V (\vec{J}_B \cdot \vec{E}_A) dV \end{aligned} \tag{3.19}$$

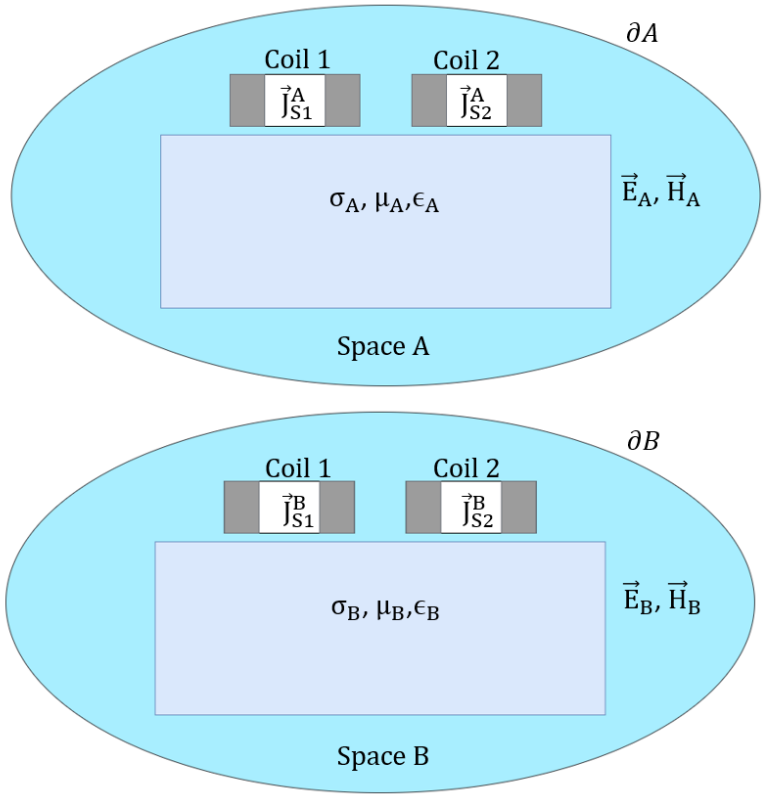

Figure 3.6: Topology of a typical measurement using IS

To obtain the sensitivities, Dyke and Freeman considered three systems, given in table 3.1 [104]. After various substitutions of these systems into eq. 3.19 and after some manipulation, eq. 3.20 is derived.

Table 3.1: Three Systems used to determine the analytic sensitivity equations

System 1	System 2	System 3
μ	$\mu + \delta\mu$	μ
ϵ	$\epsilon + \delta\epsilon$	ϵ
σ	$\sigma + \delta\sigma$	σ
$\vec{J}_{s_A}$	$\vec{J}_{s_A} + \delta\vec{J}_{s_A}$	$\vec{J}_{s_B}$
$\vec{E}_A$	$\vec{E}_A + \delta\vec{E}_A$	$\vec{E}_B$
$\vec{B}_A$	$\vec{B}_A + \delta\vec{B}_A$	$\vec{B}_B$

$$\begin{aligned}
&\oint_V \vec{J}_{s_B} \cdot \delta\vec{E}_A dV \\
\approx &\oint_V \vec{E}_A \cdot \vec{E}_B \delta\sigma + \vec{E}_A \cdot \vec{E}_B j\omega\delta\epsilon + \vec{E}_B \cdot \delta\vec{J}_{S_A} - j\omega\vec{H}_A \cdot \vec{H}_B \delta\mu dV \\
&+ \oint_S (\vec{E}_B \times \delta\vec{H}_A + \vec{H}_B \times \delta\vec{E}_A) \cdot \vec{n} dS
\end{aligned} \tag{3.20}$$

Assuming the fields themselves are generated by the source currents and that the fields generating these source currents are internal to the domain boundary (there is no power flow across the boundary), collecting terms and rearranging gives eq. 3.21. In COMSOL®, the outer boundary is defined with the normal components of the magnetic field set such that $\vec{n} \times \vec{A} = 0$; therefore, the condition of no power flow across the boundary is enforced.

$$\begin{aligned}
&\oint_V \delta\vec{E}_A \cdot \vec{J}_{S_B} dV \\
\approx &\oint_V \vec{E}_A \cdot \vec{E}_B \delta\sigma + \vec{E}_A \cdot \vec{E}_B i\omega\delta\epsilon + \vec{E}_B \cdot \delta\vec{J}_{S_A} - i\omega\vec{H}_A \cdot \vec{H}_B \delta\mu dV
\end{aligned} \tag{3.21}$$

It follows that if system B corresponds to a receiver being excited (with other coils open circuit) and system A the transmitter excited (with other coils open circuit) both with current I, then the change in receiver voltage due some perturbation in an electromagnetic variable is given by eq. 3.22 [104].

$$\begin{aligned}
&\delta\psi \\
\approx &\frac{1}{I} \oint_V \vec{E}_{Tx} \cdot \vec{E}_{Rx} \delta\sigma + \vec{E}_{Tx} \cdot \vec{E}_{Rx} j\omega\delta\epsilon + \vec{E}_{Rx} \cdot \delta\vec{J}_{Tx} + j\omega\vec{H}_{Tx} \cdot \vec{H}_{Rx} \delta\mu dV
\end{aligned} \tag{3.22}$$

In the limit of small perturbations, this is the definition of the total differential of a function. Through inspection, the sensitivities to the material properties

are given in eq. 3.23-3.26.

$$\frac{\partial \psi}{\partial \sigma} = \frac{1}{I} \oint_V \vec{E}_{Tx} \cdot \vec{E}_{Rx} dV \tag{3.23}$$

$$\frac{\partial \psi}{\partial \epsilon} = \frac{1}{I} \oint_V j\omega \vec{E}_{Tx} \cdot \vec{E}_{Rx} dV \tag{3.24}$$

$$\nabla_{\vec{J}_{Tx}} \psi = \frac{1}{I} \oint_V \vec{E}_{Rx} dV \tag{3.25}$$

$$\frac{\partial \psi}{\partial \mu} = \frac{1}{I} \oint_V j\omega \vec{H}_{Tx} \cdot \vec{H}_{Rx} dV \tag{3.26}$$

Fletcher [52] looked at three types of dot product and found experimentally that the product,

$$\vec{E}_{Tx} \cdot \vec{E}_{Rx} = \sum_{i=1}^{k} \vec{E}^i_{Tx} \vec{E}^i_{Rx}$$

where $\vec{E}^i$ is the i^{th} component of the vector $\vec{E}$, to be the most accurate. For a gradiometer coil configuration, the sensitivity to conductivity is

$$\begin{aligned} \frac{\partial(\psi_{Rx1} - \psi_{Rx2})}{\partial \sigma} &= \frac{1}{I} \oint_V \vec{E}_{Tx} \cdot \vec{E}_{Rx1} dV - \frac{1}{I} \oint_V \vec{E}_{Tx} \cdot \vec{E}_{Rx2} dV \\ &= \frac{1}{I} \oint_V \vec{E}_{Tx} \cdot (\vec{E}_{Rx1} - \vec{E}_{Rx2}) dV \end{aligned}$$

This was first shown by Fletcher and applies to both a symmetric and asymmetric coil configuration. Tesfalem modelled the coils as filaments with a single turn and scaled the voltages accordingly [53]. From Faraday's law, the voltage across the ends of a coil is equal to the total flux penetrating a turn (a loop) multiplied by the number of loops. If the pickup has T_{Rx1} turns and the backing off has T_{Rx2}, the scaled sensitivities are

$$\frac{\partial(\psi_{Rx1} - \psi_{Rx2})}{\partial \sigma} = \frac{T_{Rx1}}{I} \oint_V \vec{E}_{Tx} \cdot \vec{E}_{Rx1} dV - \frac{T_{Rx2}}{I} \oint_V \vec{E}_{Tx} \cdot \vec{E}_{Rx2} dV$$

This is provided that the flux through each turn is equal. Tesfalem compared the modelled and measured DMI and showed some divergence between the two at higher frequencies.

Yin et al. derived the sensitivity equations that include velocity effects of the sensor [57], where the sensitivity to conductivity is found by including the current due to the relative movement of the conductor. In this thesis, the measurements are statically taken, and therefore, there are no velocity effects.

Practical Computation

The sensitivity equation is broad its generality: it is derived without any assumptions on the geometry of the problem, presence of other materials (as long as they are included in the model) and can be used with other electromagnetic probes such as hall effect sensors [57]. In the discretised model, the sensitivity to any-one layer can be approximated by dividing each layer into a three dimensional grid with L number of voxels. The sensitivity to a layer conductivity can then be numerically approximated by eq. 3.27.

$$\frac{\partial \psi}{\partial \sigma_n} \approx \frac{1}{I} \sum_{i=1}^{L} (\vec{E}_{Tx}(i) \cdot \vec{E}_{Rx}(i)) V(i) \tag{3.27}$$

where $\vec{E}(i)$ is the electric field strength evaluated at at the centre of voxel i, with volume $V(i)$.

For complex geometries, it can be difficult to analytically define exact points in which to evaluate sensitivity and voxels which perfectly tessellate over the geometry. Instead, a numeric approach is used. Since all station bricks are machined from cylindrical geometries, the evaluation grid is defined in cylindrical co-ordinates. To determine which grid points are within the material under inspection, the conductivity map is first extracted from the model. The next stage is to find the points where the conductivity is equal to air and to remove them. This process is demonstrated in fig. 3.7; it can be seen that with more points, there will be a smaller voxel size and greater accuracy. This method of selecting points to evaluate is simpler and more

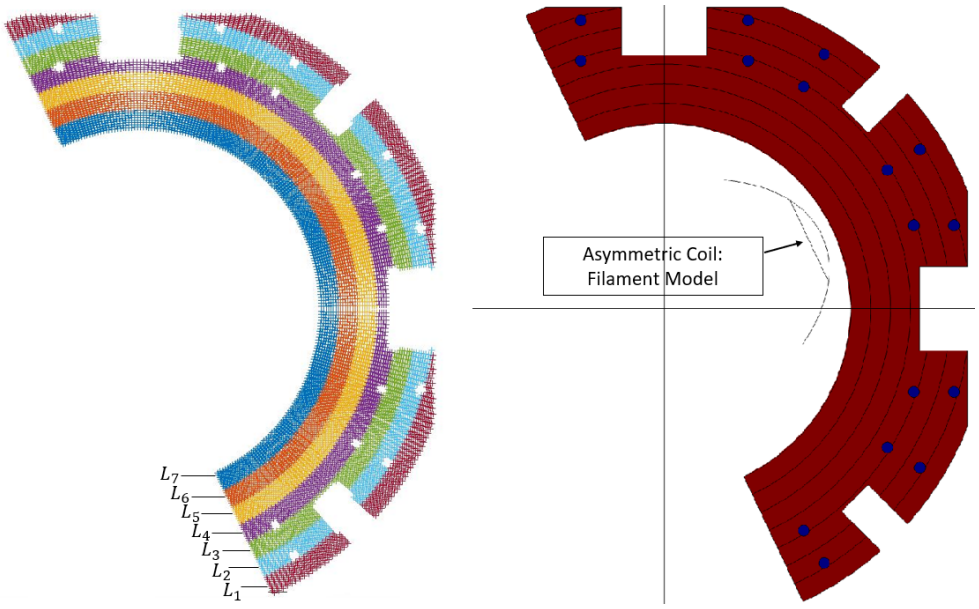

Figure 3.7: A slice of the evaluation grid of a Heysham-type brick. It can be seen that the methane holes can easily be resolved with a small enough voxel size.

robust to different geometries than analytic methods.

3.7 Inverse Problem

3.7.1 Jacobian

The total differential of some measurement due to a number of different conductive regions is given in eq. 3.28, where $d\sigma_n$ is an infinitesimal change in a layer conductivity.

$$dm = \frac{\partial m}{\partial \sigma_1} d\sigma_1 + \frac{\partial m}{\partial \sigma_2} d\sigma_2 \cdots + \frac{\partial m}{\partial \sigma_n} d\sigma_n \tag{3.28}$$

When $\vec{\sigma}$ is perturbed by some amount $\delta\vec{\sigma}$, the approximate change in measurement is given by eq. 3.29.

$$\Delta m \approx \frac{\partial m}{\partial \sigma_1} \Delta\sigma_1 + \frac{\partial m}{\partial \sigma_2} \Delta\sigma_2 \cdots + \frac{\partial m}{\partial \sigma_n} \Delta\sigma_n \tag{3.29}$$

A simple inversion technique is to use eq. 3.29 to map changes in the measurement to changes in the material conductivity (relative to some starting point). For p different interrogating frequencies, the approximate change in the measurements can be written as the following system of equations

$$\begin{bmatrix} \frac{\partial m_1}{\partial \sigma_1} & \frac{\partial m_1}{\partial \sigma_2} & \frac{\partial m_1}{\partial \sigma_3} & \cdots & \frac{\partial m_1}{\partial \sigma_n} \\ \frac{\partial m_2}{\partial \sigma_1} & \frac{\partial m_2}{\partial \sigma_2} & \frac{\partial m_2}{\partial \sigma_3} & \cdots & \frac{\partial m_2}{\partial \sigma_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial m_{2p}}{\partial \sigma_1} & \frac{\partial m_{2p}}{\partial \sigma_2} & \frac{\partial m_{2p}}{\partial \sigma_3} & \cdots & \frac{\partial m_{2p}}{\partial \sigma_n} \end{bmatrix} \begin{bmatrix} \Delta\sigma_1 \\ \Delta\sigma_2 \\ \vdots \\ \Delta\sigma_n \end{bmatrix} \approx \begin{bmatrix} \Delta m_1 \\ \Delta m_2 \\ \vdots \\ \Delta m_{2p} \end{bmatrix}$$

The number of rows is twice the number of frequencies because the real and imaginary parts of the data are concatenated. The matrix is called the Jacobian ($\boldsymbol{J}$). The system of equations is written in simpler notation as eq. 3.30, where the residual vector here is defined as $\vec{r}(\vec{\sigma}) = \vec{m}_{model} - \vec{m}_{actual}$. This approximation is exact if the relationship between the conductivity profile and measurement is linear.

$$\boldsymbol{J}\vec{p} \approx \vec{r}(\vec{\sigma}) \tag{3.30}$$

3.7.2 Over-determined system

An over-determined system of equations refers to when there are a greater number of rows than columns in the Jacobian: the number of equations is greater than the number of unknowns, and therefore, the column vectors do not span the full space they are described in [107]. This is simplest to understand with two vectors describing a plane in 3D space, where any vector within the plane of the two vectors can be expressed but not outside the plane. Typically, in the Jacobian the number of measurements (rows) is greater than the number of columns (unknowns). For an over-determined problem, any small amount of noise (or approximation error, such as in eq. 3.30) may make the problem unsolvable. However, a useful solution can still be found by instead

finding the combination of column vectors which has the smallest distance between them and $\vec{r}(\vec{\sigma})$. Geometrically, this is the orthogonal projection of $\vec{r}(\vec{\sigma})$ onto the column space of $\boldsymbol{J}$. This solution is called the Moore-Penrose pseudo-inverse, given by eq. 3.31.

$$\vec{p} = (\boldsymbol{J}^T\boldsymbol{J})^{-1}\boldsymbol{J}^T\vec{r}(\vec{\sigma}) \tag{3.31}$$

3.8 Non-Linear Optimisation

The inverse problem is highly non-linear. This can be recognised by simply considering that a highly conductive surface layer will affect the incident magnetic field, and thus, the eddy current distribution in lower layers. This means for some matrix $\boldsymbol{A}$ a linear mapping does not exist between $\vec{\sigma}$ and $\vec{m}$, such that $\boldsymbol{A}\vec{\sigma} = \vec{m}$. To determine the underlying conductivity distribution, the minimum of some objective function $f(\vec{\sigma})$ is determined with respect to the conductivity profile; this is denoted by $\text{argmin}_{\vec{\sigma}}(f(\vec{\sigma}))$.

3.8.1 Generalised Search Directions

It is helpful to re-write the conductivity variable as some step $\vec{p}$ from a point $\vec{\sigma}_k$, such that $\vec{\sigma} = \vec{\sigma}_k + \vec{p}$. This equation can be used to define the recursive relation $\vec{\sigma}_{k+1} = \vec{\sigma}_k + \vec{p}$. The step $\vec{p}$ is chosen such that a minimum is approached. The choice of algorithms is extremely rich in optimisation literature and is

Algorithm 1:

```
define σ⃗_0 ;
while Some Condition do
    compute p⃗ ;
    σ⃗_{k+1} = σ⃗_k + p⃗ ;
end
```

subject to the scale of the problem, the required accuracy, how much *a priori*

knowledge there is of the solution and the speed of the inversion.

The choices of step can all be derived from the second order Taylor series approximation of $f(\vec{\sigma})$, which is given by eq. 3.32, where $\nabla_{\vec{\sigma}} f(\vec{\sigma}_k)$ is the gradient at some point $\vec{\sigma}_k$ and $\boldsymbol{H}_f(\vec{\sigma}_k)$ is the Hessian matrix at that same point. The Hessian matrix contains the second order derivative information of the objective function. The function $q(\vec{p})$ is the quadratic approximation of the objective function.

$$f(\vec{\sigma_k} + \vec{p}) \approx q(\vec{p}) = f(\vec{\sigma}_k) + \nabla_{\vec{\sigma}} f(\vec{\sigma}_k)^T \vec{p} + \frac{1}{2} \vec{p}^T \boldsymbol{H}_f(\vec{\sigma}_k) \vec{p} \tag{3.32}$$

The step can be decomposed as $\vec{p} = \alpha \vec{d}$, where α is the step length and $\vec{d}$ the search direction. If $f(\vec{\sigma}) = q(\vec{p})$, the optimal step length in the direction $\vec{d}$ can be found analytically, otherwise a numeric search for the step length would have to be performed.

The simplest search direction is to move in the opposite direction to the gradient, given by eq. 3.33. This is the GD direction. This direction is always perpendicular to the contours of the objective function and is, therefore, slow to converge when the current iterate is near a minimum.

$$\vec{d} = -\frac{\nabla_{\vec{\sigma}} f(\vec{\sigma}_k)}{\left\| \nabla_{\vec{\sigma}} f(\vec{\sigma}_k) \right\|} \tag{3.33}$$

A method which has a faster rate of convergence is the Newton step, which uses the local curvature to yield a more direct step towards a minimum. This step can be derived by finding the gradient of eq. 3.32 with respect to conductivity, where equating this to zero and re-arranging yields the Newton step given in eq. 3.34.

$$\vec{p} = -\boldsymbol{H}_f(\vec{\sigma}_k)^{-1} \nabla_{\vec{\sigma}} f(\vec{\sigma}_k) \tag{3.34}$$

In trust region methods, a constraint region is applied around the objective function, such that $\|\vec{p}\| < \|\Delta\|$. The step towards the minimum of $q(\vec{p})$ in this region can then be derived using Lagrange multipliers as eq. 3.35 [108], where γ is the damping parameter. This step can be seen as a combination of the Newton and GD methods, where as $\gamma \to 0$ the trust region direction converges towards the Newton method, and as $\gamma \to \infty$ the direction converges to GD with $\alpha \to \left\|\nabla_{\vec{\sigma}} f(\vec{\sigma}_k)\right\| / \gamma$.

$$\vec{p} = -(\boldsymbol{H}_f(\vec{\sigma}_k) + \gamma \boldsymbol{I})^{-1} \nabla_{\vec{\sigma}} f(\vec{\sigma}_k) \tag{3.35}$$

The size of the constraint region can be altered depending on the quadratic model agreement with the actual objective function. This agreement can be quantified by defining a function called the gain ratio, given in eq. 3.36. If the quadratic model of the objective function accurately represents the problem then $\rho \approx 1$ and the solution of $\text{argmin}_{\vec{p}}(q(\vec{p}))$ in this constraint region is trusted to be the step towards a minimum of $f(\vec{\sigma})$. The size of the constraint region can be altered depending on user defined ranges of ρ, i.e. levels of trust in the quadratic model.

$$\rho = \frac{f(\vec{\sigma}_k + \vec{p}) - f(\vec{\sigma}_k)}{q(\vec{p}) - q(0)} \tag{3.36}$$

3.8.2 Scaling

An alternative problem can be solved by defining $\boldsymbol{D}\vec{p} = \vec{\omega}$, or equivalently $\vec{p} = \boldsymbol{S}\vec{\omega}$, where commonly the choice of the matrix $\boldsymbol{D}$ is to scale the components of $\vec{p}$ in order to create equal sensitivities across the component parts of $\vec{\omega}$. In [109], the transformed problem is considered by having

$$q^s(\vec{\sigma}_k + \vec{\omega}) \approx f(\vec{\sigma}_k + \boldsymbol{S}\omega) = f^s(\vec{\omega})$$

The current point is then given by setting $\vec{\omega} = 0$, such that $f^s(0) = f(\vec{\sigma}_k)$. The gradient and Hessian of the scaled function at the current iterate is then given by eq. 3.37 and 3.38, respectively [109]. The direction towards the minimum in the scaled problem can then be transformed into the original co-ordinates as $\vec{p} = \boldsymbol{S}\vec{\omega}$. Note that for GD methods that $\vec{p} = \alpha \boldsymbol{S}\boldsymbol{S}^T \nabla_{\vec{\sigma}} f(\vec{\sigma}_k)$, meaning that the GD direction is not invariant under co-ordinate transformation, therefore, convergence can be improved using an appropriate scaling matrix.

$$\nabla_{\vec{\omega}} f^s(0) = \boldsymbol{S}^T \nabla_{\vec{\sigma}} f(\vec{\sigma}_k) \tag{3.37}$$

$$\boldsymbol{H}_{f^s}(0) \approx \boldsymbol{S}^T \boldsymbol{H}_f(\vec{\sigma}_k) \boldsymbol{S} \tag{3.38}$$

For the trust region method, using the scaled version of the problem, the minimum in the constraint region can be found by substituting the scaled gradient and Hessian into eq. 3.35, and simplifying to give eq. 3.39 [108]. Substitution of the modified Hessian and gradient into the Newton equation will not yield any benefit as the Newton method is invariant under re-scaling [110].

$$\vec{p} = -(\boldsymbol{H}_f(\vec{\sigma}_k) + \gamma \boldsymbol{D}^T \boldsymbol{D})^{-1} \nabla_{\vec{\sigma}} f(\vec{\sigma}_k) \tag{3.39}$$

A common choice for the matrix $\boldsymbol{D}$ is to use the diagonal values of the Hessian, such that $\boldsymbol{D}^T \boldsymbol{D} = diag\{\boldsymbol{H}_f(\vec{\sigma}_k)\}$ [111], [112], this choice of scaling matrix is known as the Jacobi pre-conditioner. However, non-diagonal matrices can be used; for example, in consideration of the eigenvalue analysis of the Hessian [109]. Scaling the variables in this way was discussed in Marquardt's seminal paper [113] and has been used in inverse problems for many decades [114]. In [51] and [115] the diagonals of the Hessian were used to solve an unregularised objective function, similar to the method used by Cheney et al. [58]. The matrix $\boldsymbol{D}$ is referred to as the scaling matrix in this work.

3.8.3 Objective Function

The function to be minimised must consider the error between the simulated and measured response, given by $\vec{r}(\vec{\sigma})$. This term alone is not a sufficient formulation since an inversion algorithm would move towards the largest negative residual. A norm must be applied. There are many different norms, such as the L1 (Manhattan) where the absolute values of the components of $\vec{r}(\vec{\sigma})$ are summed. The L1 norm is a piecewise linear function, and therefore, minimising objective functions using this norm requires non-smooth optimisation methods. In this work the L2 norm is used to formulate the objective function. Fig. 3.8 and 3.9 show the difference between the two norms. The objective function used is given in eq. 3.40, where for mathematical convenience the L2 norm is squared and multiplied by a factor of 1/2.

$$f(\vec{\sigma}) = \frac{1}{2}\left\|\vec{r}(\vec{\sigma})\right\|_2^2 \tag{3.40}$$

An alternative objective function is the weighted least squares, given in eq. 3.41. The gradient and Hessian are given by eq. 3.42 and 3.43, where the matrix $\boldsymbol{W}$ has only diagonal components with $W_{ii} = P_i^2$ [116]. The weighting could be chosen to be the measurement data, such that the objective is to minimise the sum of the squared relative error.

$$f(\vec{\sigma}) = \frac{1}{2}\sum (P_i r_i(\vec{\sigma}))^2 \tag{3.41}$$

$$\nabla_{\vec{\sigma}} f(\vec{\sigma}_k) = \boldsymbol{J}_k^T \boldsymbol{W} \vec{r}(\vec{\sigma}_k) \tag{3.42}$$

$$\boldsymbol{H}_f(\vec{\sigma}_k) \approx \boldsymbol{J}_k^T \boldsymbol{W} \boldsymbol{J}_k \tag{3.43}$$

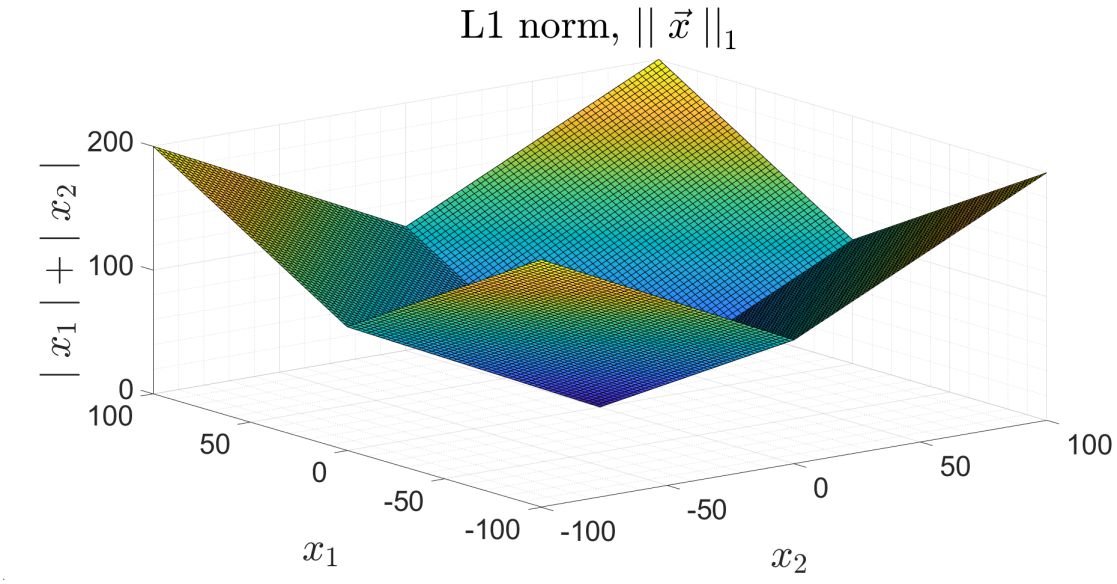

Figure 3.8: An example of the L1 norm for a two dimensional vector, $\vec{x}$.

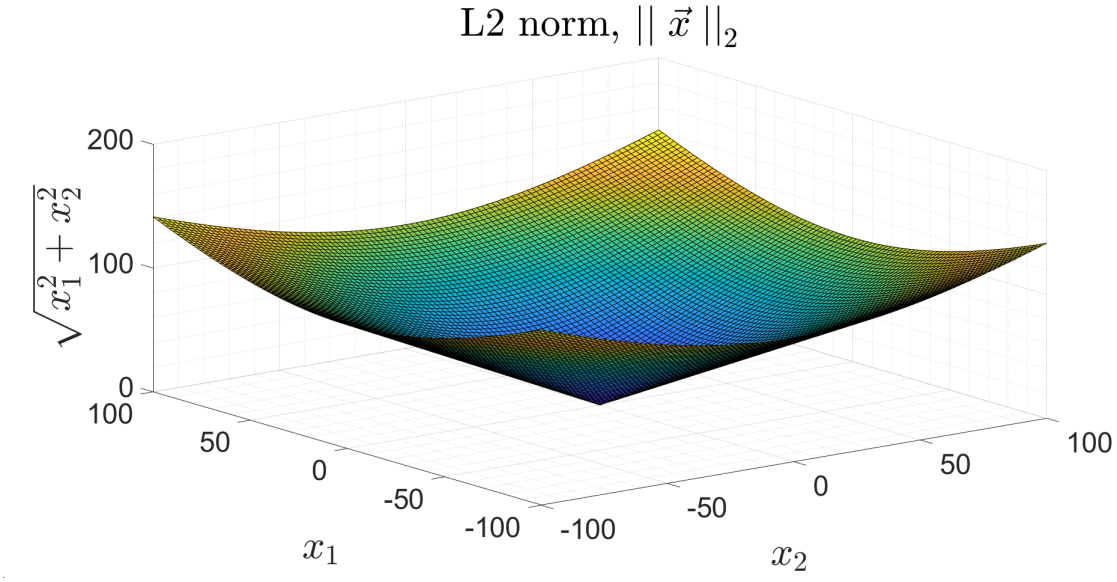

Figure 3.9: An example of the L2 norm for a two dimensional vector, $\vec{x}$.

3.9 Regularisation

The Singular Value Decomposition (SVD) of the Moore-Penrose pseudo-inverse is given by eq. 3.44

$$(\boldsymbol{J}^T\boldsymbol{J})^{-1}\boldsymbol{J}^T = \sum_{i=1}^{rank(\boldsymbol{J})} v_i \sigma_i^{-1} u_i^T \tag{3.44}$$

where the vector $\vec{u}_i$ is the i^{th} eigenvector of the matrix $\boldsymbol{J}\boldsymbol{J}^T$, and $\vec{v}_i$ is the i^{th} eigenvector of the matrix $\boldsymbol{J}^T\boldsymbol{J}$ and σ_i are the singular values of the matrix $\boldsymbol{J}$ [117]. It can be seen in the SVD of the Jacobian that small singular values close to zero can cause erroneous results; this type of problem is referred to as ill-conditioned because of numerical instability. It is expected that the Jacobian matrix will be ill-conditioned and small errors will result in large errors in $\vec{p}$. The purpose of regularisation is to improve on numerical stability and to prevent over-fitting to noise. Hansen describes the purpose of regularisation as follows [118]

> *it is necessary to incorporate further information about the desired solution in order to stabilize the problem and to single out a useful and stable solution. This is the purpose of regularization.*

The physical reason for the ill-conditioning can be understood as a direct consequence of the skin effect, where the sensitivity is typically greatest near the coils and decays rapidly with increasing distance. The conditioning of the Jacobian matrix can be improved through careful selection of the frequency range and in the discretisation of the problem. However, these considerations alone are not sufficient for correcting the ill-conditioning of the problem.

There are many types of regularisation, where different types of regularisation have different filtering actions [118], [119]. The most common types of regularisation use penalty functions and have been historically used in graphite inversion [51], [52], [120]. The regularisation of the solution can be explicitly defined by incorporating a penalty $\mathfrak{R}$ within the objective function, such as in eq. 3.45.

$$f(\vec{\sigma}) = \frac{1}{2}\left\|\vec{r}(\vec{\sigma})\right\|_2^2 + \mathfrak{R}(\vec{\sigma}) \tag{3.45}$$

In Tikhonov regularisation an L2 penalty norm is added to the objective function. There are other regularisation methods based on different norms of

the penalty term; for example, the total variation method uses the L1 norm of the depth profile gradient to produce a non-smooth penalty. The advantage of total variation is in reconstructing images with discontinuities present such as voids. Soleimani et al. were the first to apply total variation to electrical capacitance tomography, showing the superior performance to Tikhonov in reconstructing images with sharp edges [62].

3.9.1 Truncated Singular Value Decomposition

One regularisation technique is to not sum over the rank of $\boldsymbol{J}$ in eq. 3.44, but to a smaller index; this is called Truncated Singular Value Decomposition (TSVD). In TSVD, a matrix with improved conditioning is extracted without too much loss of information from the original matrix. TSVD is an alternative to penalty methods that explicitly lowers the condition number of a matrix through the elimination of small singular values. The point of truncation must be determined according to the acceptable minimum size of the singular values.

3.9.2 Generalised Tikhonov Regularisation

The Tikhonov regularised least squares objective function used in this work is given by eq. 3.46. The parameter λ controls the amount of biasing towards the penalty. The quantity $\vec{\sigma}_0$ is an *a priori* estimate of the solution, which could be derived from a typical conductivity profile or a ML algorithm. In this work, $\vec{\sigma}_0$ is used as the initial guess of the conductivity depth profile in the iterative algorithms.

$$f_\lambda(\vec{\sigma}) = \frac{1}{2}\left\|\vec{r}(\vec{\sigma})\right\|_2^2 + \frac{\lambda}{2}\left\|\boldsymbol{\Gamma}(\vec{\sigma} - \vec{\sigma}_0)\right\|_2^2 \tag{3.46}$$

The gradient of eq. 3.46 is given by eq. 3.47 and the Hessian approximated as eq. 3.48, where these are derived in the appendix. Note that eq. 3.48 gives the approximate Hessian and when this is used in the Newton step it is referred to as the GN method.

$$\nabla_{\vec{\sigma}} f_\lambda(\vec{\sigma}_k) = \boldsymbol{J}_k^T \vec{r}(\vec{\sigma}_k) + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma}(\vec{\sigma}_k - \vec{\sigma}_0) \tag{3.47}$$

$$\boldsymbol{H}_{f_\lambda}(\vec{\sigma}_k) \approx \boldsymbol{J}_k^T \boldsymbol{J}_k + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma} \tag{3.48}$$

There are numerous choices for the regularisation matrix $\boldsymbol{\Gamma}$, which reflects how the penalty is to be chosen. The simplest to use is the identity matrix and another is the finite difference operator. When the finite difference operator is employed the approximate gradient information is used and the penalty term is scaled by the size of the gradient or change in gradient relative to some a priori solution.

3.10 Regularised Least Squares Directions

3.10.1 Modified Gauss-Newton

The GN direction can be modified such that the direction is multiplied with a constant α, this method is called the Modified Gauss-Newton (MGN) or Hartley's method [111]. The regularised MGN direction is then given by eq. 3.49.

$$\vec{p} = -\alpha(\boldsymbol{J}_k^T \boldsymbol{J}_k + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma})^{-1}(\boldsymbol{J}_k^T \vec{r}(\vec{\sigma}_k) + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma}(\vec{\sigma}_k - \vec{\sigma}_0)) \tag{3.49}$$

It can be seen that with regularisation, the GN direction is the Moore-Penrose pseudo-inverse of the linear least squares problem in eq. 3.50, where $\vec{p}_0 = \vec{\sigma}_k - \vec{\sigma}_0$ (more details can be found in [121] and the appendix).

$$\vec{p} = \text{argmin}_{\vec{p}}\left(\frac{1}{2}\left\|\boldsymbol{J}_k \vec{p} + \vec{r}(\vec{\sigma}_k)\right\|_2^2 + \frac{\lambda}{2}\left\|\boldsymbol{\Gamma}(\vec{p} - \vec{p}_0)\right\|_2^2\right) \tag{3.50}$$

3.10.2 Levenberg-Marquardt

The regularised LM step is given by eq. 3.51, this is the direction used in Tesfalem's inversion algorithm [120].

$$\vec{p} = -(\boldsymbol{J}_k^T \boldsymbol{J}_k + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma} + \gamma \boldsymbol{D}^T \boldsymbol{D})^{-1} (\boldsymbol{J}_k^T \vec{r}(\vec{\sigma}_k) + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma} (\vec{\sigma}_k - \vec{\sigma}_0)) \tag{3.51}$$

Without regularisation of the objective function ($\lambda = 0$), the LM direction is the solution of the linear problem eq. 3.52

$$\vec{p} = \text{argmin}_{\vec{p}} (\frac{1}{2} \left\| \boldsymbol{J}_k \vec{p} + \vec{r}(\vec{\sigma}_k) \right\|_2^2 + \frac{\gamma}{2} \| \boldsymbol{D} \vec{p} \|_2^2) \tag{3.52}$$

At first glance of eq. 3.50 and 3.52, the scaling matrix and regularisation matrix appear to be performing the same operation, since both regularise the step direction. However, relating these back to the objective function, it can be seen that although the LM regularises the step, it does not regularise the objective function. In particular, the trust region algorithm is a Tikohnov regularisation of the step, but with $\vec{p}_0 = 0$. In this view of the trust region approach, the matrix $\boldsymbol{D}$ can be chosen in consideration of the conditioning of the problem, and the regularisation matrix $\boldsymbol{\Gamma}$ can independently be chosen with regards to preventing over-fitting to the noise. In practice the regularisation of the step and objective function both contribute towards conditioning and preventing over-fitting. The scaling matrix is not guaranteed to produce approximately equal sensitivities, the shortfall of which is made up through the regularisation of the objective function. Further, if there is too much noise near the solution then the gain ratio will not be near unity, and the damping parameter will be increased, preventing too much deviation from the current iterate and preventing over-fitting.

3.10.3 Quasi-Newton

Quasi-Newton (QN) methods use the Newton direction with an approximation of the Hessian. It is especially useful when the Hessian is not known or when the approximation in eq. 3.48 no longer holds, i.e. the so called large-residual problem [122]. Two examples of QN approximations are the Broyden-Fletcher-Goldfarb-Shanno (BFGS) and the NL2SOL, the former makes an approximation of the inverse of the Hessian, whilst the latter is a hybrid technique which predicts the terms not included in eq. 3.48. Specifically, in the NL2SOL algorithm $\boldsymbol{H}_{f_\lambda}(\vec{\sigma}_k) \approx \boldsymbol{J}_k^T \boldsymbol{J}_k + \lambda \boldsymbol{\Gamma}^T \boldsymbol{\Gamma} + \boldsymbol{S}_k$, which is a more accurate approximation to the Hessian. Conversely, the BFGS algorithm makes a prediction of the approximate inverse, $\boldsymbol{B}_k \approx \boldsymbol{H}_{f_\lambda}(\vec{\sigma}_k)^{-1}$. The BFGS algorithm uses a Newton step, whilst the NLS2OL algorithm uses a trust region method [123].

3.11 Multi-Variable Polynomial Regression (MVPR)

Forward modelling using finite element software is slow and limits image reconstruction to be offline. Instead, data driven approaches can be employed using many samples of modelling data. Both an inverse model and forward model can be derived for direct and iterative inversion, respectively.

In MVPR, the mapping between conductivity and inductance is approximated by fitting polynomial curves. The data used for determining the model parameters is referred to as the training data. A further two sets are reserved, called the validation and test. The validation set is used for finding the hyper-parameters of the ML algorithm, and the test is an unused dataset used to evaluate the ML algorithm performance.

Given a set of input and target data, the multi-variable polynomial equation can be written in matrix form as eq. 3.53. The matrix $\boldsymbol{P}$ contains the terms of each sample polynomial expression, where the number of rows is equal to the number of samples and the number of columns is the number of polynomial terms. The matrix $\boldsymbol{C}$ contains the desired coefficients of the polynomial terms, and $\boldsymbol{T}$ is a matrix containing the target data. Each row of a given column in $\boldsymbol{P}$ contains the different sample evaluations of the same term. To help with conditioning, each column can be standardised by subtracting the mean and dividing through by the standard deviation, or alternatively, the maximum absolute value.

$$\boldsymbol{PC} = \boldsymbol{T} \tag{3.53}$$

To find the coefficient matrix the TSVD of the Moore-Penrose pseudo-inverse can be used. MVPR using TSVD is a simple ML technique consisting of a single model hyper-parameter (the polynomial order) and a single inversion hyper-parameter (the truncation point). Finding the optimal results for this algorithm is simple, which makes it extremely attractive for use as a benchmark for more complicated ML regression models, such as neural networks.

3.12 Chapter Summary

- Basic electromagnetic concepts were introduced, in particular the material response to a plane wave. This provided an opportunity to introduce the skin effect and explain the division of electromagnetic inspection into soft and hard field modalities.

- The differential equation governing the eddy current problem was introduced. There are several different formulations, depending on the problem to be solved. The A formulation is used by the numeric solver

COMSOL®.

- FEA for electromagnetic problems was described; the problem domain is discretised into elements, in which the vector field is defined in terms of area or volume co-ordinates and the A-field within interpolated. The system of equations is then assembled and solved for.

- A gradiometer coil is used to collect reactor measurement data. The gradiometer coil may not be perfectly balanced, and therefore, residual mis-balancing can be reduced by subtracting the response with the graphite target absent. The sensitivity of the measurement to conductivity can be found by the forward difference method or the $\vec{E} \cdot \vec{E}$.

- A description of the Jacobian was given in §3.7, in which the Moore-Penrose pseudo-inverse was presented. This pseudo-inverse is essential when there are more equations than unknowns. For linear problems, the Jacobian would be constant, and therefore, the minimiser can be found in a single step.

- For non-linear problems, the solution is found iteratively by locally approximating the objective function and bounding the step size. This is the subject of §3.8, where the methods described are derived from the second order Taylor series expansion of the objective function. The nuclear graphite inverse problem is ill-conditioned, meaning there is large differences in sensitivity between the underlying parameters.

- Regularisation is important for stopping over-fitting and for improving conditioning. There are different types of regularisation, which correspond to different filtering actions on the step direction. Tikhonov regularisation is to be used.

- MVPR is a simple inferential technique, in which there are few hyper-parameters, and therefore, it is simple to optimise. This makes it ideal

to use as a benchmark for more complex algorithms.

Chapter 4

Comparison of Iterative Algorithms

4.1 Overview of Chapter

The iterative algorithms historically applied in graphite inversion used either GN or LM directions. The work presented in this chapter investigates a broad class of first order and second order methods for the inversion of nuclear graphite data, whilst considering different levels of noise and prior estimates.

There are two different prior estimates employed in the study: one from a ML algorithm and the other using a homogeneous conductivity of 45 kS/m. The ML algorithm used was a Convolutional Neural Network (CNN), trained using synthetic data; the procedure used to generate the synthetic data is outlined and is repeatedly used throughout the thesis.

The results show that the NL2SOL direction is in the top class for all cases; in particular, the improved performance of QN methods indicates a possible performance limit due to the residual error in the Hessian. In general, the LM algorithm is sufficient and competitive with other techniques for graphite

inversion.

4.2 The Study

In [124], a study was conducted on a wide range of algorithms. These algorithms were tested on five different problems and assigned into classes of performance; no such study has been completed for the nuclear graphite application (and more generally in eddy current NDE). Dekdouk et al. presented a small scale study of three GN type algorithms for reconstructing the depth profile of a conductive nuclear graphite block. In other work, the LM algorithm has been used for inversion and, again, the GN direction [52][53][120][125].

This chapter presents a wider study into search direction algorithms for the inversion of nuclear graphite data and considers both trust region and line search directions: the GN, LM, QN, GD, non-linear Augmented Conjugate Gradient (ACG) [126] and Powell's Dogleg (PD) [112]. Further, different regularisation methods are investigated. The study was not extended to different objective functions.

4.3 Coil and Problem Geometry

A cylindrical conductive block was used as the graphite target and this is shown in fig. 4.1. The cylindrical geometry was to allow a 2D axi-symmetric model to be used in the FEM, decreasing computation times. The conductive block model was discretised into ten layers, with every layer assigned its own independent conductivity variable. The sensor was composed of a single receive coil embedded in a transmit coil. The measurement made was the DMI (corresponding to eq. 3.16), where this measurement was made at 10

different frequencies from 600 Hz to 10 kHz. The bounds of the measurement frequency are typically determined from a calibration stage, wherein there is ideally good agreement between the simulated and experimental data: the frequency range determined is dependent on modelling accuracy, where 600 Hz to 10 kHz is appropriate for graphite inversion; for example, in previous work 100 Hz to 10 kHz has been successfully used [53].

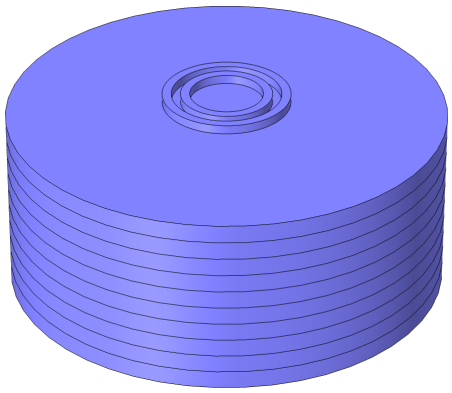

Figure 4.1: Inspection of a cylindrical block of material, that can be simulated using 2D FEM with an axi-symmetric model.

4.4 Direct Solution

4.4.1 Generation of Synthetic Data

Generating Conductivity Profiles

Each conductivity variable was assumed to belong to a uniform distribution between 0 kS/m and 95 kS/m; this conductivity range was used because conductivity cannot be negative and the conductivity of virgin Gilsocarbon is approximately 95 kS/m (loss of density in graphite means that electrical conductivity cannot be larger than its virgin value). The profile was then smoothed by fitting an n^{th} order polynomial. The algorithms were tested on 10 profiles, where each profile has a different order from $n = 1$ to $n = 10$.

As the number of layers is also 10, the highest order polynomial amounts to no smoothing. This method of modelling the conductivity profile has few assumptions about the behaviour of the graphite. An example of a generated profile is shown in Figure 4.2.

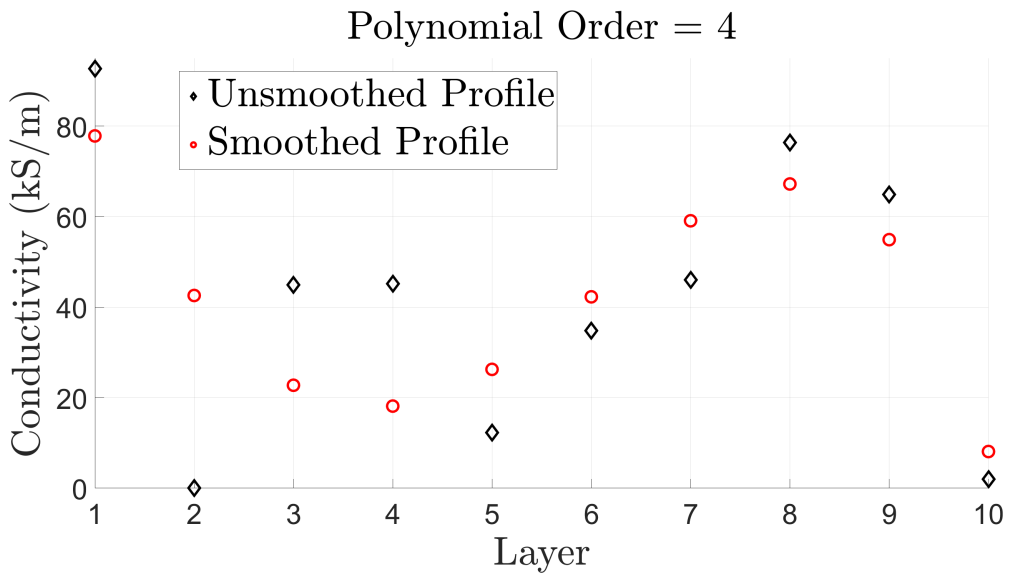

Figure 4.2: Smoothed conductivity profile, generated by fitting a polynomial to unsmooth profiles.

Measurement Noise

In order to avoid an inverse crime, it is critical to include realistic levels of noise in the synthetic data [127]. After a conductivity profile was generated, the DMI response was computed using the FEM forward solver. For the samples used as training data, noise was added such that the SNR was in the range [20 60] dB of the weakest amplitude component in the DMI response, where the noise distribution was also modelled as a uniform random variable. For the test profiles, two SNR levels were investigated: 20 dB and 60 dB.

4.4.2 Machine Learning Algorithm

Each algorithm was tested using two different *a priori* estimates: one using a ML prediction and the other using a homogeneous conductivity of 95 kS/m. The algorithm used for determining the ML prior was a CNN.

CNNs are commonly used on images, where a filter is convolved with the image. Electromagnetic data can be arranged in a similar format to image data for use in a CNN [69][128]. In this manner, the real and imaginary parts of the DMI signals were arranged into an array with a depth of two. Following this, each feature of the data was standardised by subtracting the mean and dividing by the standard deviation, where the mean and standard deviation correspond to the training dataset.

After pre-processing the data, the network hyperparameters were selected. The structure of the network was fixed: the input data was convolved with filters, down-sampled and then inputted into a Fully Connected Network (FCN). The rest of the network hyperparameters were free to be selected, such as the number of layers and the amount of down-sampling. These hyperparameters were chosen by randomly selecting combinations and manually testing. The Keras library was used, which has a large amount of functionality for building and training networks (including the hyperparameter random search). In total, 10,000 sets of DMI data were used to train the network.

4.5 Performance Metric

To compare the performance of the algorithms, the Mean Percentage Error Norm (MPEN) was used, denoted by ϵ. The MPEN as a function of iteration, $\epsilon(i)$, is given by eq. 4.1, where n is the number of different conductivity profiles and $\vec{\sigma}_p^k(i)$ is the predicted conductivity profile of the i^{th} iteration and k^{th} test

profile.

$$\epsilon(i) = \frac{1}{n}\sum_{k=1}^{n}\frac{\left\|\vec{\sigma}_t^k - \vec{\sigma}_p^k(i)\right\|}{\left\|\vec{\sigma}_t^k\right\|} \times 100 \tag{4.1}$$

Each algorithm performance was split into three classes, this approach to evaluating algorithm performance was used in [124], in which classification was done qualitatively. Here, the algorithms were grouped into clusters with k-means clustering, providing a more robust classification procedure.

4.6 Algorithm Selection

The types of search direction used were MGN, GD, LM, PD, BFGS, NL2SOL and ACG, most of which were introduced in §3.10. All of the algorithms, except those that use the GN or BFGS direction, used the diagonal of the Hessian as a scaling matrix. Tikhonov regularisation was used, where the regularisation matrix was either the finite difference operator or the identity matrix. All methods updated the regularisation parameter every iteration. TSVD of the MGN direction was also included, denoted by MGN†. The algorithms studied are summarised in table 4.1.

Table 4.1: Algorithms compared for graphite inversion.

Algorithm	Search direction	Γ
1	MGN	I
2	MGN	FD
3	MGN†	-
4	LM	I
5	LM	FD
6	PD	I
7	PD	FD
8	NL2SOL	I
9	NL2SOL	FD
10	BFGS	I
11	BFGS	FD
12	GD	I
13	GD	FD
14	ACG	I
15	ACG	FD

The ideal regularisation and damping parameters were found using the knowledge of the optimal step direction; this is not possible in practice but allowed for the comparison of the optimal results of each algorithm. These parameters were selected by using a grid search corresponding to the step closest to the optimal direction. The total number of iterations executed was 10, where more iterations may have resulted in greater accuracy but at the expense of increased computation times: 10 iterations was chosen as a compromise such that the simulation times were not excessive.

4.7 Results and Discussion

The results are discussed with reference to the four cases shown in table 4.2. The top 5 algorithms for cases 1 and 2 are plotted in fig. 4.3 and, similarly, cases 3 and 4 in fig. 4.4. These figures are included to provide some visual appreciation of an algorithm as a function of iteration.

Table 4.2: Cases in which the inversion algorithms were applied.

Case	SNR (dB)	Selection of $\vec{\sigma}_0$
1	60	Direct Inverse Solution
2	20	Direct Inverse Solution
3	60	Homogeneous Conductivity
4	20	Homogeneous Conductivity

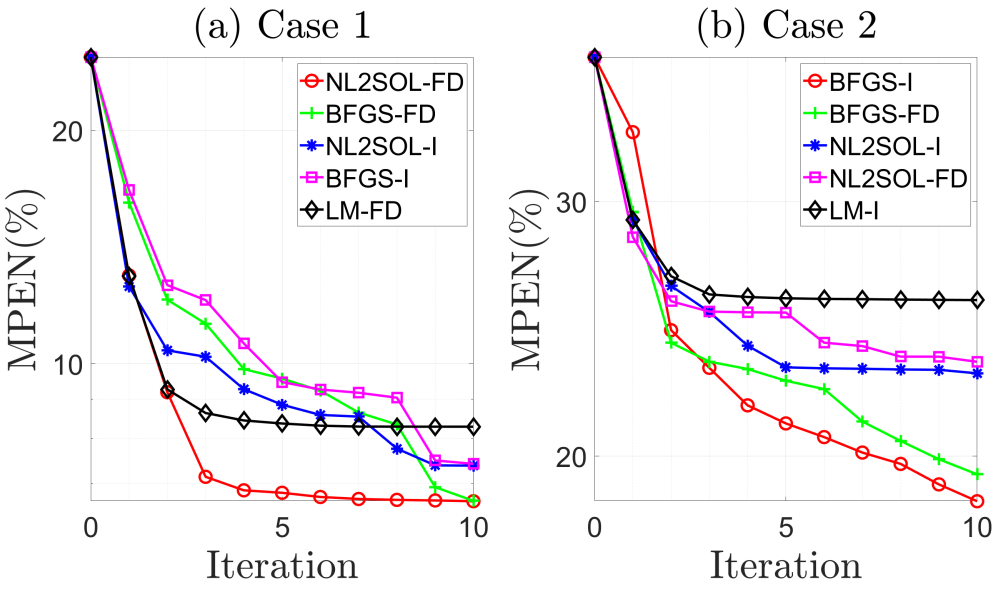

Figure 4.3: Comparing iterative algorithms: Top 5 algorithms for (a) Case 1 (b) Case 2

The first notable details are that the BFGS and NL2SOL algorithms perform extremely well for cases 1 and 2. This provides evidence that the Hessian approximation suffers from large-residual errors at the minimum. This can be seen in fig. 4.3 (a), where the NL2SOL algorithm has a notably lower error than the LM, with an equivalent regularisation matrix. The NL2SOL direction is the same as the LM direction, but approximates the residual terms in the Hessian; therefore, there is a notable difference in this case due to the large-residual error.

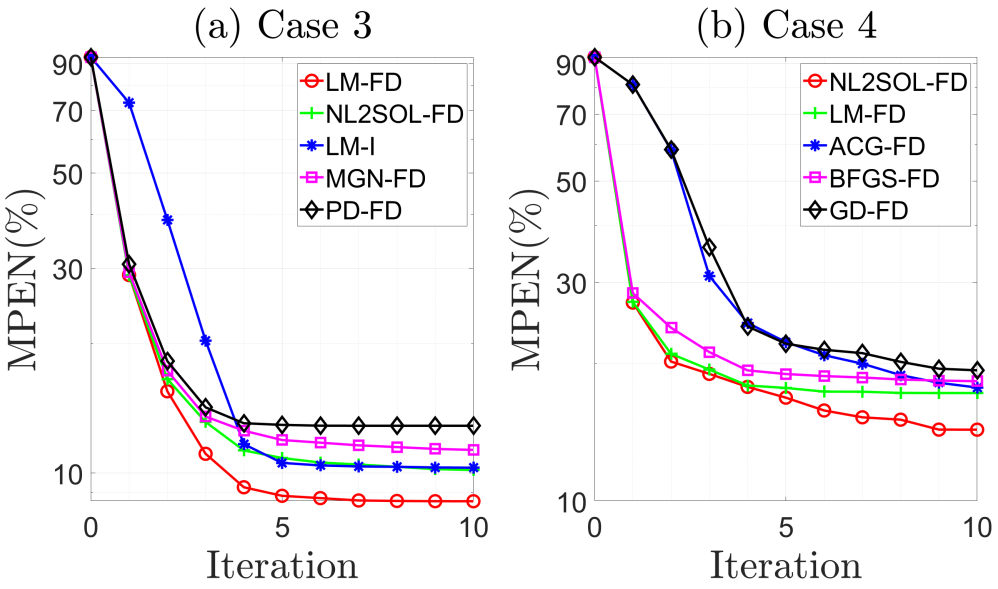

Figure 4.4: Comparing iterative algorithms: Top 5 algorithms for (a) Case 3 (b) Case 4

For cases 3 and 4 the QN performances were not as dominant, perhaps because the initial estimate was further away from the minimum (possibly affecting the Hessian approximation) or because the prior estimate (the penalty) was limiting convergence. With a poor prior estimate, the BFGS and NL2SOL algorithms may not get close enough towards the minimum where the residual terms of the Hessian become more important.

For cases 1 and 3 it can be seen that the ML prior estimates facilitate a lower final error; however, in cases 2 and 4 it can be seen the final error is larger when the prior is an ML prediction. This shows that the noise is limiting the benefit of the ML prior estimate.

Table 4.3: Case 1, where $\epsilon(0) = 24.8606\%$.

SD	$\mathbf{\Gamma}$	Class	$\epsilon(N_f)$ (%)
NL2SOL	FD	I	6.6414
BFGS	FD	I	6.6494
NL2SOL	I	I	7.3824
BFGS	I	I	7.4219
LM	FD	I	8.288
LM	I	I	8.8147
MGN	FD	II	12.2979
MGN†	-	II	12.7794
MGN	I	II	13.0636
PD	I	II	15.1056
PD	FD	II	15.7665
ACG	FD	III	19.5224
GD	FD	III	19.8481
ACG	I	III	20.1281
GD	I	III	20.2562

Case 1 has six class I algorithms, with no notable pattern in regularisation types across all classes. For this case there were clear groupings of types of search direction, with clustering of Newton, QN and first order methods; therefore, for this case the type of search direction was most important.

Case 2 has three class I algorithms. In this case it can be seen that the choice in regularisation matrix becomes more important than in case 1, with more mixing of the search direction and a tighter grouping of the choices in regularisation.

Table 4.4: Case 2, where $\epsilon(0) = 37.7533\%$.

SD	$\boldsymbol{\Gamma}$	Class	$\epsilon(N_f)$ (%)
BFGS	I	I	18.6098
BFGS	FD	I	19.4251
NL2SOL	I	I	22.8146
NL2SOL	FD	II	23.2381
LM	I	II	25.6369
LM	FD	II	26.1631
ACG	FD	II	26.486
PD	FD	II	26.7359
MGN	FD	III	28.7902
ACG	I	III	29.0787
GD	FD	III	29.1565
PD	I	III	29.2477
MGN†	-	III	31.0074
GD	I	III	31.1869
MGN	I	III	32.885

Table 4.5: Case 3, where $\epsilon(0) = 93.1785\%$.

SD	$\boldsymbol{\Gamma}$	Class	$\epsilon(N_f)$ (%)
LM	FD	I	8.5943
NL2SOL	FD	I	10.159
LM	I	I	10.2874
MGN	FD	I	11.3183
PD	FD	I	12.8849
MGN†	-	I	13.0246
MGN	I	I	13.0857
NL2SOL	I	I	13.7814
PD	I	II	15.1985
ACG	FD	II	16.5015
BFGS	FD	II	18.3074
GD	FD	II	18.3488
ACG	I	III	24.6189
GD	I	III	25.8199
BFGS	I	III	30.1816

Case 3 has eight class I algorithms. Notably, the QN methods were poorer than in case 1 and 2: with a penalty or initial estimate further from the minimum the benefit of including the residuals in the Hessian decreases and so it should be expected that QN methods are not as good.

Case 4 has seven class I algorithms. There was a greater mix of directions in this class, with a greater weighting on the importance of regularisation

compared to the other cases. The choice of regularisation matrix was tighter than in case 3. GD and ACG perform particularly well for this case.

Table 4.6: Case 4, where $\epsilon(0) = 93.1785\%$.

SD	$\mathbf{\Gamma}$	Class	$\epsilon(N_f)$ (%)
NL2SOL	FD	I	14.3276
LM	FD	I	17.2233
ACG	FD	I	17.7059
BFGS	FD	I	18.2785
GD	FD	I	19.3066
MGN	FD	I	19.4388
PD	FD	I	20.4059
LM	I	II	22.2847
PD	I	II	23.6394
NL2SOL	I	II	24.4398
ACG	I	II	24.7409
BFGS	I	II	25.2689
GD	I	II	27.9869
MGN	I	III	38.6137
MGN†	-	III	39.4183

It can be seen when comparing cases 1-2 and 3-4 that when the *a priori* estimate was the homogeneous conductivity, there were a greater number of algorithms in class 1. This was perhaps because the penalty chosen was either in the zeroth or first order deviation in some constant profile, and therefore, there was less flexibility in the reconstructed profile; thus, the algorithms converge to some limit in case 3-4. The choice of search direction is less relevant if there is a poor *a priori* estimate than if a good prior estimate is known. Generally, it can be seen that when there was a lower SNR that the regularisation matrix becomes as important as the choice in search direction.

The results show that there was not one particular search direction or choice in regularisation which dictates the performance. The choice was also dependent on the prior and number of iterations to be run. However, a strong candidate would be the hybrid technique NL2SOL, an algorithm which performs well across each case.

The criterion for a top performing algorithm was chosen to be the lowest

error at the 10th iteration, therefore, other algorithms could be more or less superior depending on a number of design choices. For other applications, the computation of the Hessian may be costly and therefore, GD or ACG may be more appropriate.

4.8 Chapter Summary

- To date, there has been no comprehensive study into the different search algorithms available for graphite inversion and this extends to eddy current NDE in general. Typically, second order methods have been used for graphite inversion.

- The study investigated two different prior estimates: one of homogeneous conductivity and the other a ML prediction. The ML algorithm was trained using synthetic data, where the methodology for generating the synthetic data is used throughout the thesis. First, a conductivity depth profile is randomly generated and smoothed using a polynomial. The conductivity profile is then used to obtain the corresponding inductance data from a forward model. Noise is then added to the inductance data.

- The performance of algorithms was classified by a k-means clustering algorithms. There is no single algorithm which is optimal for all cases — the search direction and regularisation type depend the prior estimate and noise level. However, the augmented LM is consistently in the top. Generally, the LM algorithm with the finite difference regularisation matrix is a sufficient and competitive algorithm.

Chapter 5

Inversion by Inference

5.1 Overview of Chapter

In the previous chapter, a CNN was used to obtain a direct solution to the inverse problem for use as a prior estimate in an iterative algorithm. A CNN algorithm may be more complicated than necessary and requires extensive tuning of hyper-parameters. This chapter considers the suitability of a CNN, by comparing it with a simple MVPR algorithm. In the comparison, a number of statistical measures are used to assess performance: the correlation coefficient (between error and SNR) and error range, standard deviation and mean.

Currently, there is little information about the behaviour of the graphite at the periphery; therefore, three different studies are contained, one on smooth conductivity profiles (up to order four), another on unsmooth profiles and the last on reactor type profiles (derived from previously trepanned graphite and are extrapolated for the purpose of the study). In all of these cases, the two algorithms are mostly comparable in performance. Although a CNN has the capacity to model complicated functions, the MVPR algorithm is much simpler to design and faster to train.

5.2 Direct Inversion

The application of neural networks to electromagnetic data has been reported for several decades, with some of the earliest work done by Udpa et al. in defect characterisation [78]. In the following years, research was published showing the applicability of neural networks to recovering electromagnetic parameter distributions, such as conductivity depth profiles [66]–[69]. The use of neural networks in electromagnetic NDE, ranges from medical imaging [66] to monitoring CO_2 emissions [69].

Using data driven techniques, it is possible to obtain a solution to the inverse problem directly. This solution could be used as a prior estimate in an optimisation algorithm or it could be trusted as a sufficiently accurate approximation to the true solution.

5.3 The Study

The methodology for generating the synthetic data was the same as in §4.4.1, with the same SNR bounds. The conductivity profiles were limited to an order between zero and four. The same coil was used and the same measurement. The key difference was the use of a cuboid block instead of the previously used cylindrical block. In total, 5000 synthetic samples were generated for training, 500 reserved for testing and 500 for validation.

Three different cases are investigated: smooth conductivity profiles (up to order four), unsmooth and reactor type profiles. The reactor profiles are not expected to be random; however, the unsmooth case provides insight into the ability of the ML algorithms in mapping conductivity to inductance.

To compare the algorithms, a number of statistical measures are used to

assess performance: the correlation coefficient (between error and SNR) and error range, standard deviation and mean. If each of these these statistics is represented by s_i, then the mean ratio of the statistics is

$$r = \frac{1}{4}\sum_{i=1}^{4}\left\|\frac{s_i^{CNN}}{s_i^{MVPR}}\right\| \tag{5.1}$$

This ratio can be used to encapsulate the performance of each algorithm with a single number. If $r < 1$ then the CNN can be said to have a better performance, if $r > 1$ then the MVPR is better and if $r = 1$ they are equal.

5.4 Hyper-parameter Determination

5.4.1 Convolutional Neural Network

The CNN was trained as in §4.4.2. First, the data was arranged into a 3-dimensional matrix, similar to an image; the real part of the data formed one channel and the imaginary another. The data was standardised in the same way. The same CNN structure and tuning process was used, with the hyper-parameters found by manually tuning.

5.4.2 Multi-variable Polynomial Regression

The disadvantage in MVPR is that the number of polynomial terms grows rapidly with order. Table 5.1 shows the memory requirements to store the matrix P with 5000 training samples and 20 dependent variables. For the machine used and code written, an order of 5 was the greatest achievable, where orders greater than this would crash the program when computing the matrix inverse.

In MVPR, the data is transformed by a polynomial expansion, where the

Order	Memory (GB)
1	8.4×10^{-4}
2	9.2×10^{-3}
3	7.1×10^{-2}
4	4.3×10^{-1}
5	2.1×10^{0}
6	9.2×10^{0}
7	3.5×10^{1}
8	1.2×10^{2}

Table 5.1: Memory requirements of the MVPR algorithm for type double.

coefficients are to be determined. To determine the polynomial coefficients, TSVD was used to regularise the Moore-Penrose pseudo-inverse of the transformed training data, with the ideal truncation point found using a golden-section search. The golden-section search is a derivative free optimisation method and is given in algorithm 2. The polynomial order and truncation point are found according to the minimum error corresponding to the validation data. This is a simple regularisation technique in which the regularisation

Algorithm 2: Univariate search algorithm for minima. It is the golden-section search algorithm when $\frac{1}{a} = \frac{1+\sqrt{5}}{2}$

```
define x_l and x_u;      // x_u is the upper bound and x_l the lower
define a, s.t. 0.5 < a < 1;
while Some Condition do
    d = |x_u − x_l|;                          // distance between bounds
    x_tu = x_l + d × a;                       // x_tu is the trial upper bound
    x_tl = x_u − d × a;                       // x_tl is the trial lower bound
    if f(x_tu) > f(x_lu) then
        x_u = x_tu;
    else
        x_l = x_tl;
    end
end
```

parameter (the truncation point) is bounded. The metric used for tuning the hyper-parameters was the MPEN. Fig. 5.1 shows the golden-section search in finding the ideal truncation index. The red markers are the updated upper limits in the golden-section search and the black are the updated lower limits.

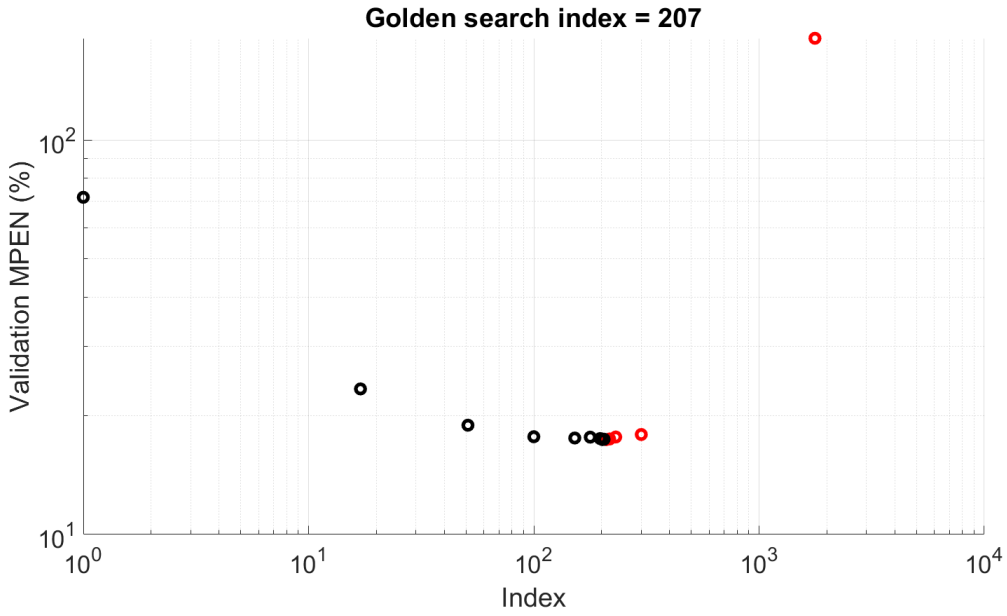

Figure 5.1: Golden section search for truncation point in TSVD of the Moore-Penrose pseudo-inverse (used for finding the coefficients in MVPR)

The above procedure is repeated for the different polynomial orders: the validation error in fig. 5.2 corresponds to the minimum error found in the truncation point search for a given order. For this work an order of 3 was found to be optimal.

MVPR using TSVD is a simple ML technique consisting of two hyper-parameters: the polynomial order and regularisation parameter (the truncation point). Finding the optimal results for this algorithm is simple, making it extremely attractive for use as a benchmark for more complicated regression models, such as neural networks.

Figure 5.2: Determining the polynomial order in MVPR: the validation error as a function of polynomial order.

5.5 Results and Discussion

5.5.1 Smooth Depth Profiles

The results of testing on 500 smoothed samples are shown in fig. 5.3, with four example profiles shown in fig. 5.4 to provide some visual context for the error data. The correlation coefficient measures the dependence of the error on the SNR. It can be seen that the CNN has a larger correlation coefficient, a smaller error bound, a larger MPEN and a smaller standard deviation. In this case $r = 0.9123$, indicating that the CNN is marginally better.

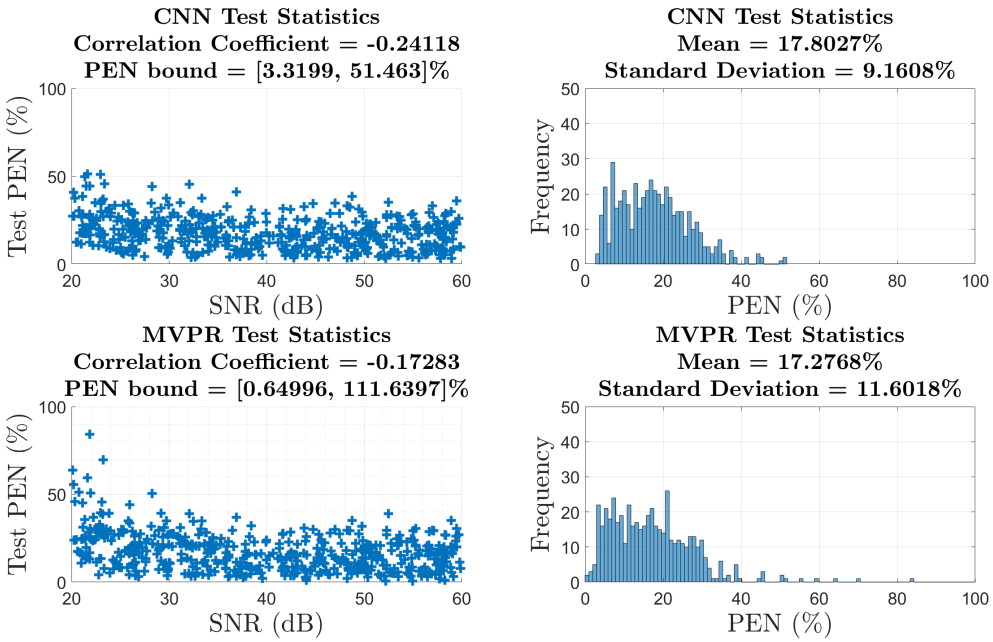

Figure 5.3: Test statistics of two regression models for direct inversion of inductance data corresponding to smooth conductivity profiles: CNN Vs MVPR.

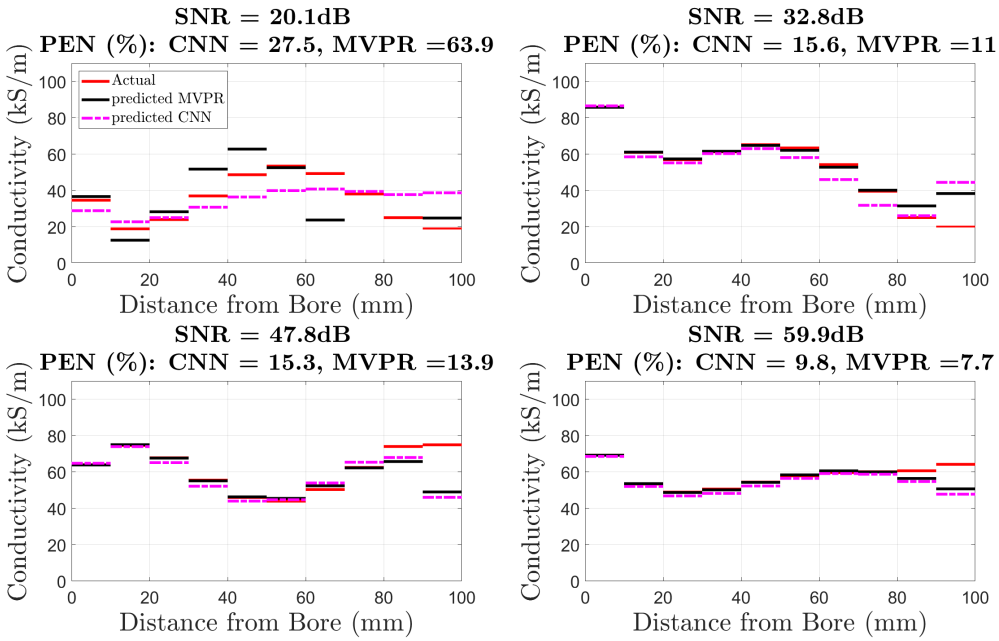

Figure 5.4: Examples of recovered depth profiles for direct inversion of inductance data corresponding to smooth conductivity profiles.

5.5.2 Unsmooth Depth Profiles

The performance of the MVPR and CNN models could be seen as unrealistic because the profiles were smoothed beforehand, i.e. a predetermined relationship was built between the inductance data and conductivity profile. This is an issue of data generation, and the only way to mitigate this is to lower the smoothing and to generate more training data so that the models learn an improved understanding of the fundamental physical relationship between the conductivity and DMI data. This is a more global issue in data-driven approaches, where bias can exist in the input training data. Therefore, it is beneficial to see how both models perform on profiles that are completely random.

Figures 5.5 and 5.6 show the results on 500 unsmoothed test profiles. In this case, the MVPR algorithm has a larger mean, standard deviation, error bound but a smaller correlation coefficient. The mean ratio of the statistics is $r = 1.0784$, indicating that the MVPR is narrowly better.

Even though both models have been trained with smoothed profiles they can still both provide a useful approximation of profiles of a higher order. Although the behaviour of the graphite at the periphery is unknown, the profiles are also not expected to be random, and therefore, the assumption that the profiles have some smoothness is likely valid. Moreover, the assumption that these profiles can be approximated by a fourth-order polynomial is a reasonable one.

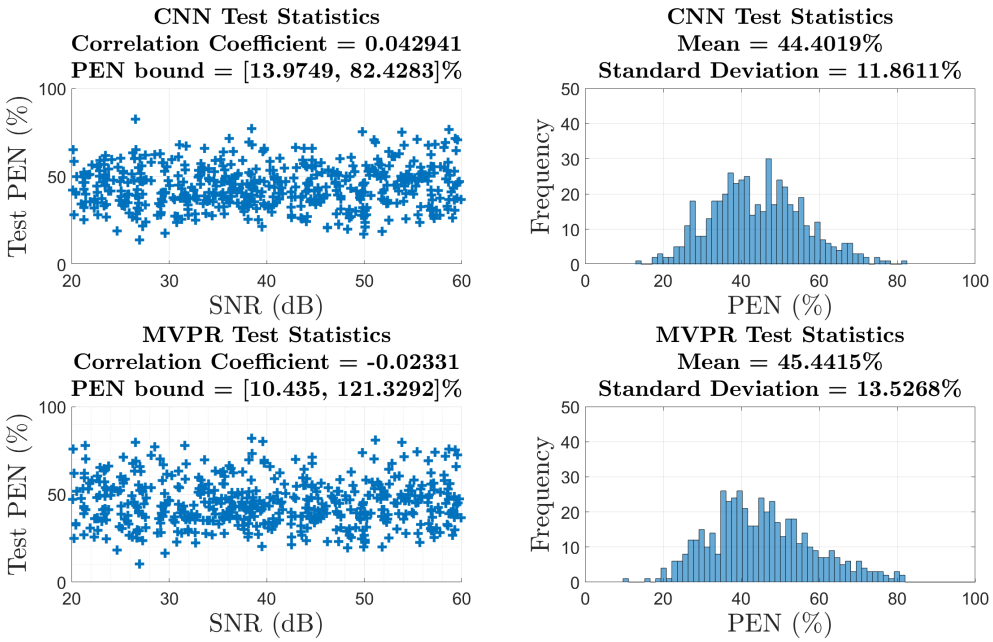

Figure 5.5: Test statistics of two regression models for direct inversion of inductance data corresponding to unsmooth conductivity profiles: CNN Vs MVPR.

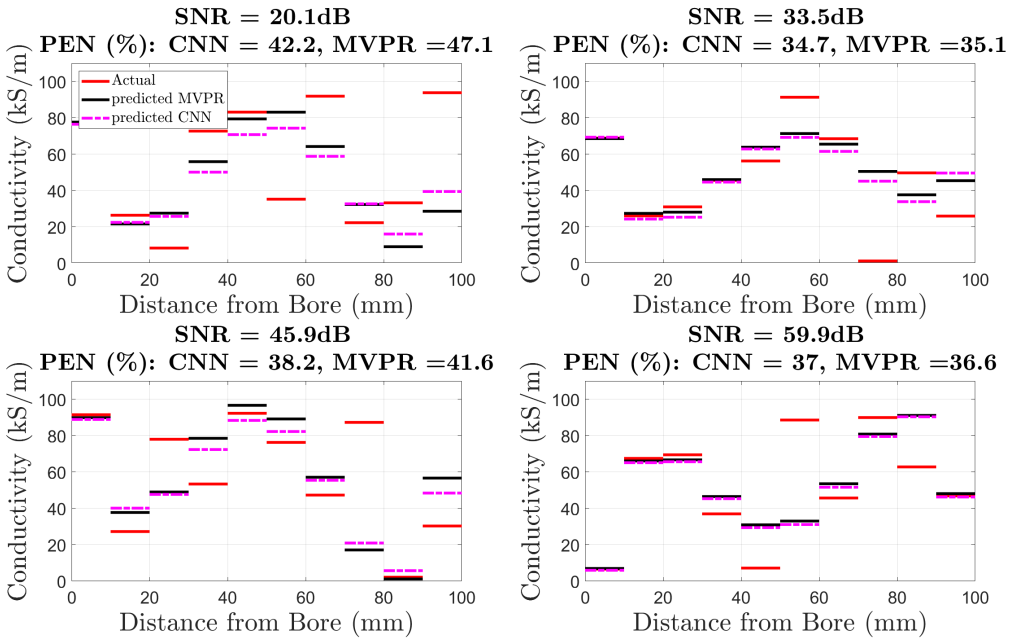

Figure 5.6: Examples of recovered depth profiles for direct inversion of inductance data corresponding to unsmooth conductivity profiles.

5.5.3 Reactor-type Depth Profiles

The previously discussed test profiles are not indicative of actual reactor conductivity depth profiles. To model typical reactor-type profiles, trepanned density measurements were used. As mentioned, the trepanning machine does not penetrate through the entire through-wall thickness of the graphite channels. The length of a trepanned sample is approximately 2/3 of the thickness of the channel wall. To estimate the behaviour beyond this point, a curve can be fitted to the trepanned data and the remainder of the depth profile extrapolated. Fig. 5.7 shows the estimated depth profiles from 6 trepanned samples. In total, 193 extrapolated profiles were generated by fitting a second-order polynomial to the trepanned data.

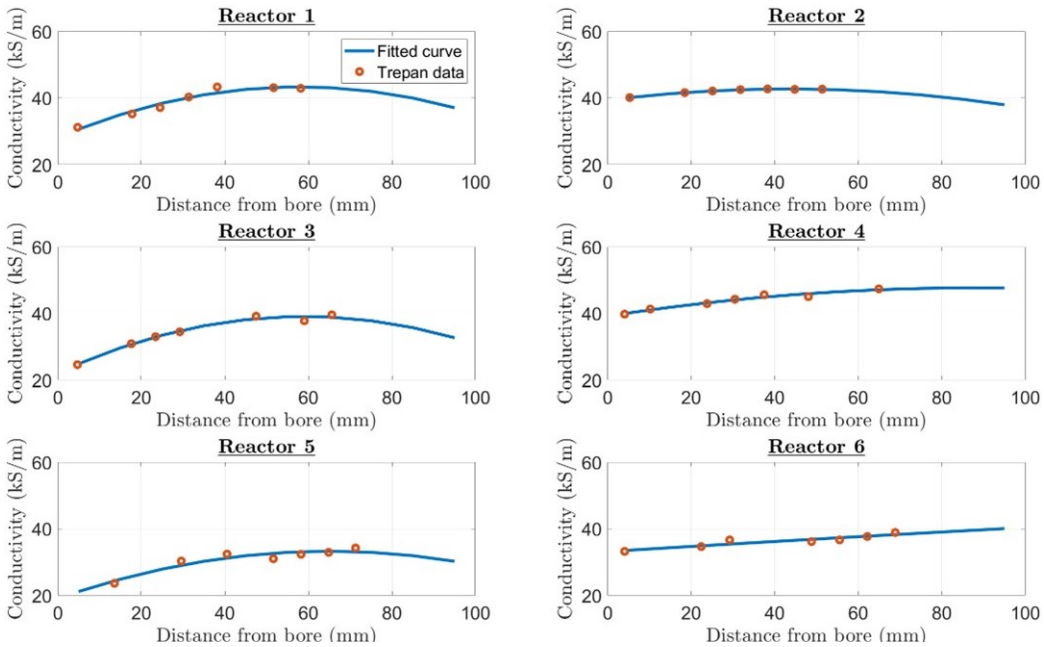

Figure 5.7: Conductivity profiles based on reactor data, extrapolated by fitting a quadratic curve.

Fig. 5.8 shows the test statistics of the two regression models on reactor-type data, with example inversions shown in fig. 5.9. It can be seen from these statistics that the CNN has a larger MPEN, a smaller standard deviation, a smaller error bound and a marginally larger correlation coefficient. In this case $r = 0.6708$, indicating a notable performance difference in the CNN and

MVPR.

Generally, it can be seen that the CNN reconstruction has better noise immunity compared with the MVPR. The mean ratio across the smooth, unsmooth and reactor-type data is $\bar{r} = 0.8871$, indicating that overall the CNN is better; however, the MVPR is competitive with the CNN algorithm and it is concluded that the MVPR algorithm is sufficient for graphite inversion. For larger sets of training data, the neural network may be more appropriate because of the memory requirements in MVPR.

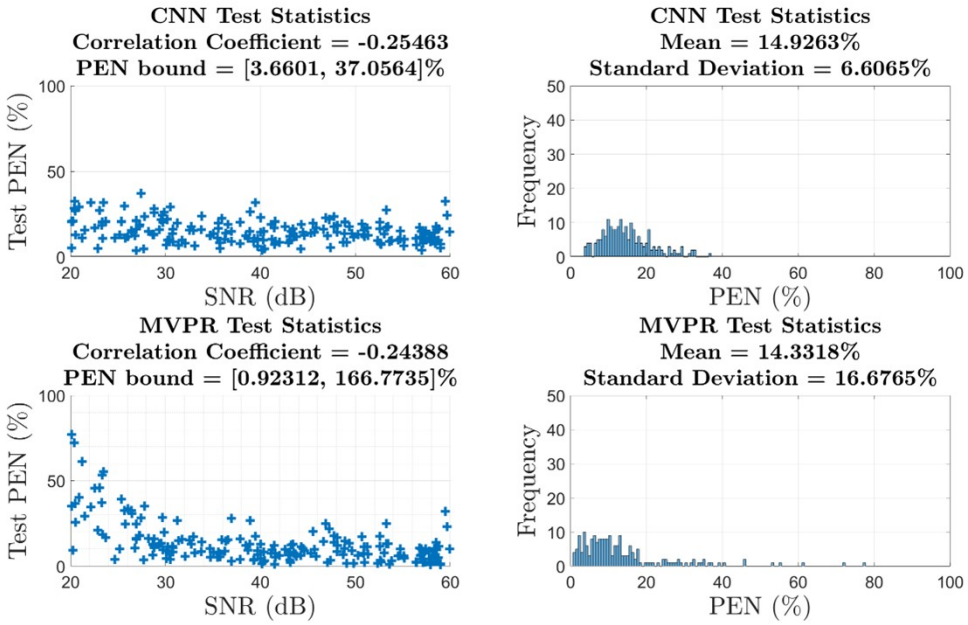

Figure 5.8: Test statistics of two regression models for direct inversion of inductance data corresponding to reactor-type conductivity profiles: CNN Vs MVPR.

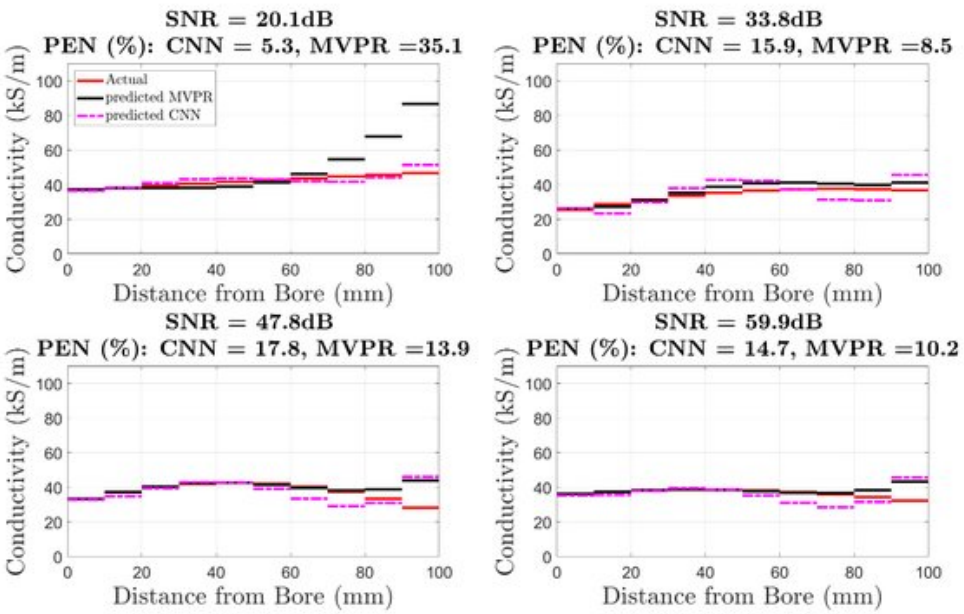

Figure 5.9: Examples of recovered depth profiles for direct inversion of inductance data corresponding to reactor-type conductivity profiles.

5.6 Chapter Summary

- Data-driven techniques can be used to obtain a direct inverse solution. A CNN has been previously used in inverting graphite data; however, it may not be the simplest and most efficient solution because they have many hyper-parameters to tune.

- A comparison of a CNN against a simple MVPR algorithm has shown that they are competitive against a number of test statistics: correlation coefficient (error and SNR), mean error, error bounds and error standard deviation. Therefore, an MVPR algorithm is used throughout the rest of this work. However, the memory requirements for a MVPR algorithm are significant, which may limit the polynomial order with larger datasets. It is noted that the CNN seems to have better noise immunity when reconstructing the depth profile.

Chapter 6

Sensor Characterisation

6.1 Overview of Chapter

The previous studies in inferential and iterative techniques considered synthetic data. To apply the algorithms on measured data, it is necessary to calibrate the forward model. This is the subject of this chapter.

This chapter begins by observing that there are low frequency resonant effects in the deployed measurement system. In the DMI measurement, the air response is subtracted and the resonant effects nulled to a first order degree. This may not be sufficient for inversion, and therefore, damping circuitry is used to carefully extend the operating frequency range.

Optimisation techniques are used to calibrate the forward model, where there are constraints on coil sequence, lift off, turns etc. The constrained optimisation algorithm, used for both calibration and inversion, is presented. This algorithm is based on the LM trust region algorithm, which inherently finds a step towards a minimum of an objective function within a spherical or elliptical constraint region. The feasibility of the algorithm is demonstrated on the calibration of the reactor and laboratory probes. It is shown how

the calibration procedure can be used to determine the coil model, frequency range and mesh. For this application, a filament model is as accurate as a 3D coil model.

It is concluded from the calibration of the reactor probes, that damping does reduce the resonant effects and improves the forward model approximation. Therefore, damping goes beyond visibly reducing resonant effects.

6.2 Resonance and Damping

The work done in studying and characterising the resonance effects was led by A. Fletcher, where credit and thanks are given. The contribution of the author was in collecting the data, tuning the lumped equivalent circuit and on-site testing of damping components.

The eddy current formulation in the magnetic fields interface in COMSOL® does not include the displacement current. Further, the FEM does not model the capacitance between each turn of a coil. A more accurate description would include the capacitance of the coils and the cables connecting them to the measurement instrument. As the frequency is increased, the capacitance between coil windings and cables becomes increasingly important. This can be better understood with a lumped circuit model of the problem. Fig. 6.1 shows the equivalent circuit for a gradiometer coil, where the self-inductance, $\mathfrak{L}$, describes how an individual circuit's magnetic field is coupled to itself.

The resonant characteristics of the coil are not necessarily problematic if they are included in the modelling. To model the resonance in COMSOL®, both the electric and magnetic fields interface would be required. There are then challenges in coil modelling, since each turn would need to be individually defined — this then raises complications with meshing because the model computation time would significantly increase. Another option is to use the

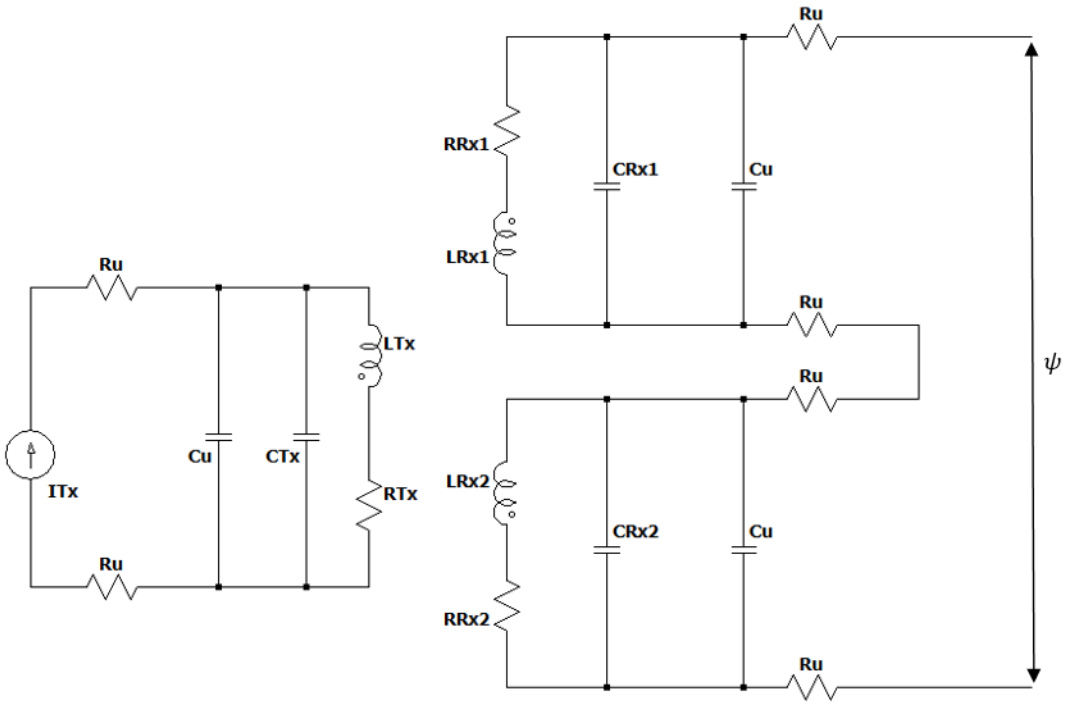

Figure 6.1: The lumped equivalent circuit model of a gradiometer coil and connecting cables.

electric circuit interface in COMSOL® and to apply this in conjunction with the magnetic fields interface. In either case, the modelling of resonance is not guaranteed to provide suitable inversion results because the resonant problem is significantly ill-conditioned. Instead, it is simpler to truncate the spectrum to a suitable range to avoid resonant effects altogether. Inversion of inductance data with resonant effects is perhaps an avenue for future research.

There are three PEP sensors which are discussed, all manufactured with the same specification. In reactor conditions, the sensor is connected to the instrumentation device with a long umbilical cable. The umbilical cable is approximately 60 m long and adds a large amount of capacitance into the circuit. The deployed reactor probe is compared with a replica system at the James Fisher Nuclear Laboratory (JFNL) site, which is intended to closely model the reactor setup. The measured resistance of the umbilical at the JFNL site was $R_u = 19.8\ \Omega$ and the capacitance 2.05 nF. The free space

(air) measurement of the three PEP sensors is shown in fig. 6.2. The JFNL and reactor probe were connected to the console with different umbilical cables and the laboratory probe had no umbilical cable. The reactor and JFNL probes are deployed from a tool with metal casing, and therefore, two different probe extensions are included with a diameter of 263 mm and 270 mm. The differences between the various responses could be due to manufacturing variations and/or balancing of the sensors and/or the umbilical and hoist. It can be seen that there is little sensitivity to the metal casing of the inspection tool because the two different extensions of the probe are almost identical in the response and this is expected since it was designed to have little sensitivity to the casing [54]. Comparing the reactor and JFNL probes, it is observed that there is a possible defect in the reactor system (hoist or probe) and that the differences between the two systems is not related to the probe distance from the metal casing. Overall, resonant effects in the reactor probe are visible at a significantly lower frequency than the JFNL and laboratory probes.

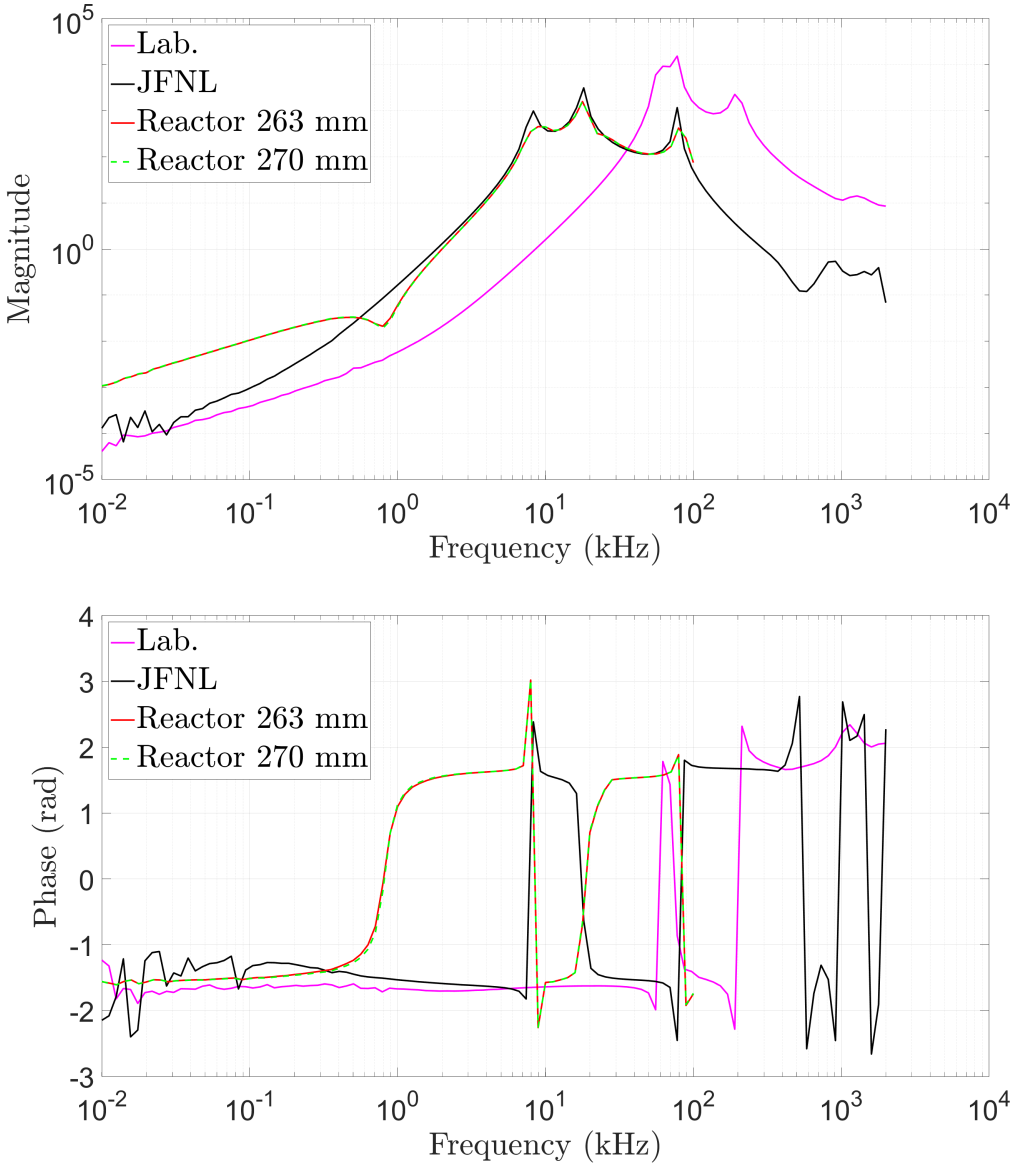

Figure 6.2: The measurement of three different PEP sensors in air. The laboratory sensor is not in the ECIT housing and does not have an umbilical connected. The JFNL and reactor probes are both in the ECIT housing and are connected to the measurement device via an umbilical. Included are two extensions of the reactor probe, at a distance of 131.5 mm diameter and 135 mm from the centreline of the tool.

Determining Lumped Component Model

To extend the bandwidth of the deployed reactor probe, damping circuitry was added. To choose the damping components, the equivalent circuit model of the undamped gradiometer must first be determined. The model parameters (such as coil resistance) were found according to the minimum error between the experimental and analytic measurement. Tuning the model parameters is difficult for resonant circuits because the difference between the largest and smallest measurement spans several orders of magnitude. Therefore, optimisation techniques are not employed and instead a grid search was used. The gradiometer measurement can be segmented for coil tuning; for example, the low frequency data will have little sensitivity to coil capacitance. This can be used to simplify the fitting problem. First, a curve was fitted to the low frequency data up to 100 Hz by tuning the mutual inductance between the transmit and receive coils. The frequency range was then increased beyond the dip at approximately 600 Hz and the self-inductance and resistance of each coil tuned, in addition with the mutual inductance between the two receive coils. Following this, the capacitances of all the coils were tuned, using the full measurement bandwidth.

The tuned model and experimental response, using the reactor probe, are plotted in fig. 6.3. This free space experimental measurement was made using a nylon section within the hoist (shown in fig. 6.4) and not in air. The resonant peaks are not well modelled. There is a large amount of error in the second peak, likely due to the limited data available; however, this is not as important because it is the low frequency resonant effects that are to be smoothed.

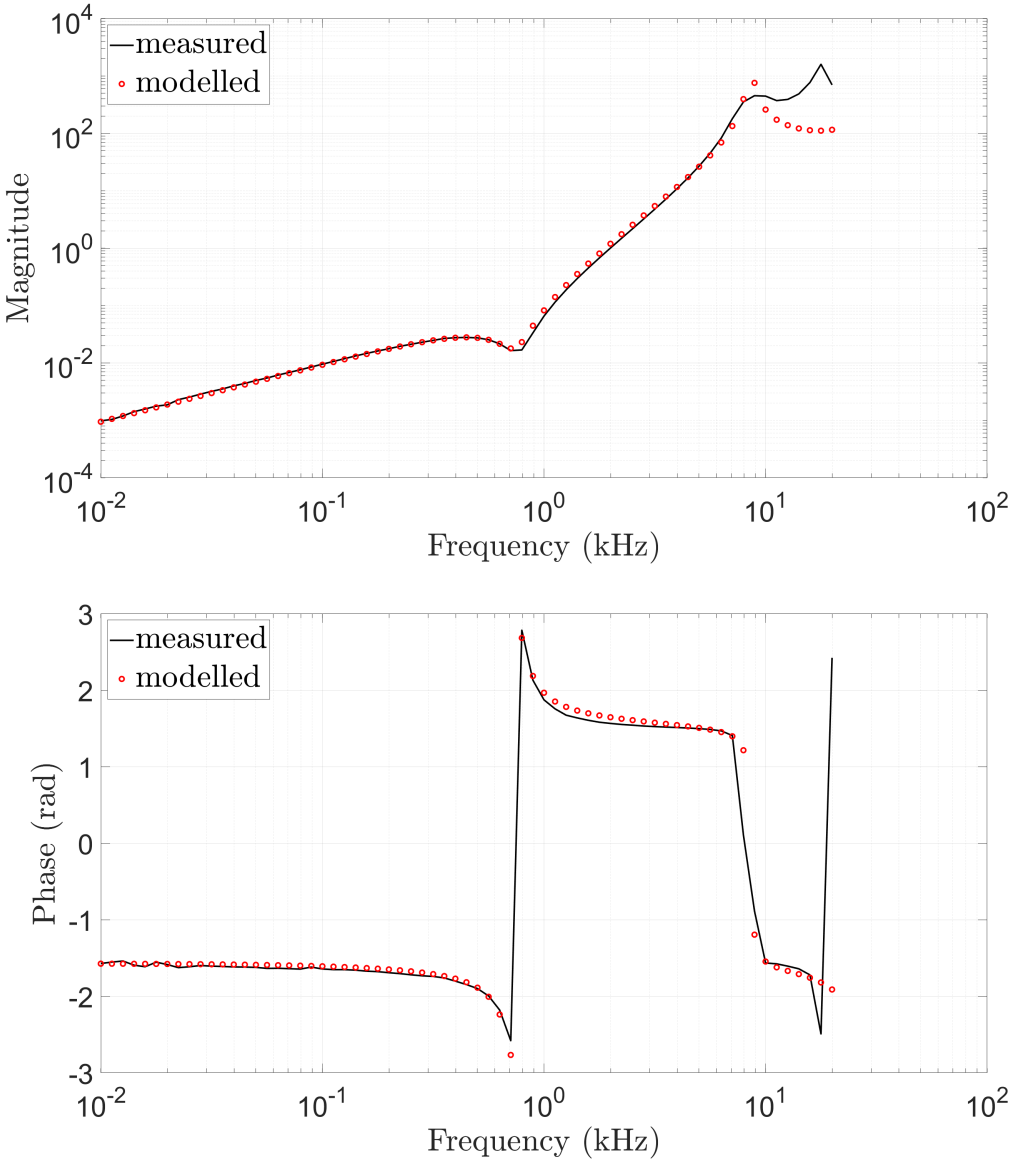

Figure 6.3: The measured response of the nylon brick using the deployed reactor probe and the tuned equivalent circuit of this measurement.

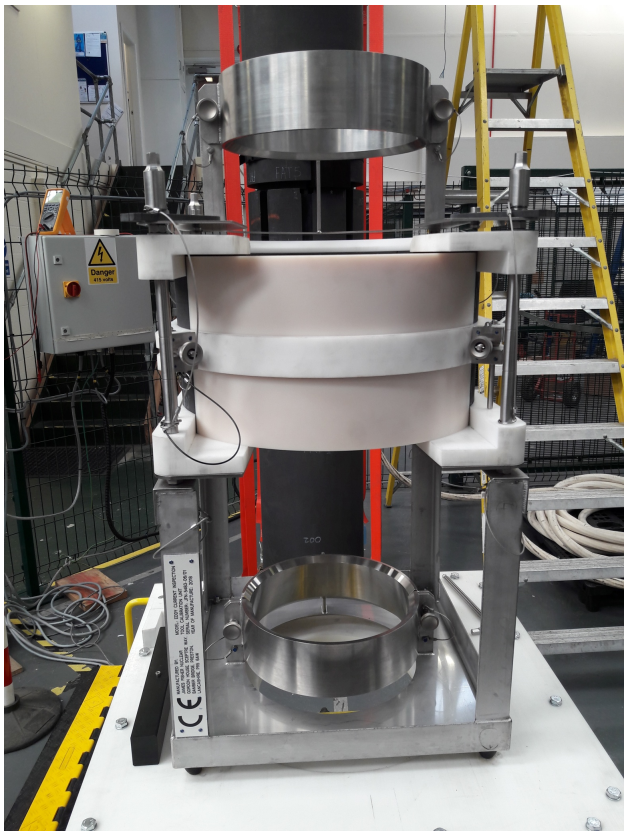

Figure 6.4: The calibration brick that is placed within the hoist.

Selecting The Damping Components

To smooth out the 'dip', damping circuitry is added and this is shown in fig. 6.5. The damping components were determined by first finding the lumped equivalent circuit equation *with* the additional damping circuitry. The previously determined equivalent circuit values are then used to define the coils in the damped circuit equation. The damping components are chosen such that the 'dip' is smoothed out and the phase approximately equal to $\pi/2$ radians. These modelled damping component values are used as initial values for the experimental work, where a selection of components are taken on site. The result of this process is shown in fig. 6.6. On site, the damping parameter values chosen were: $C_{d1} = 0$ pF, $C_{d2} = 0$ nF, $R_{d1} = 11.5$ kΩ and $R_{d2} = 10.5$ kΩ. As a heuristic observation, the magnitude response was most sensitive to the damping resistors and the phase response to the capacitors. There were difficulties with tuning the phase because there was very little

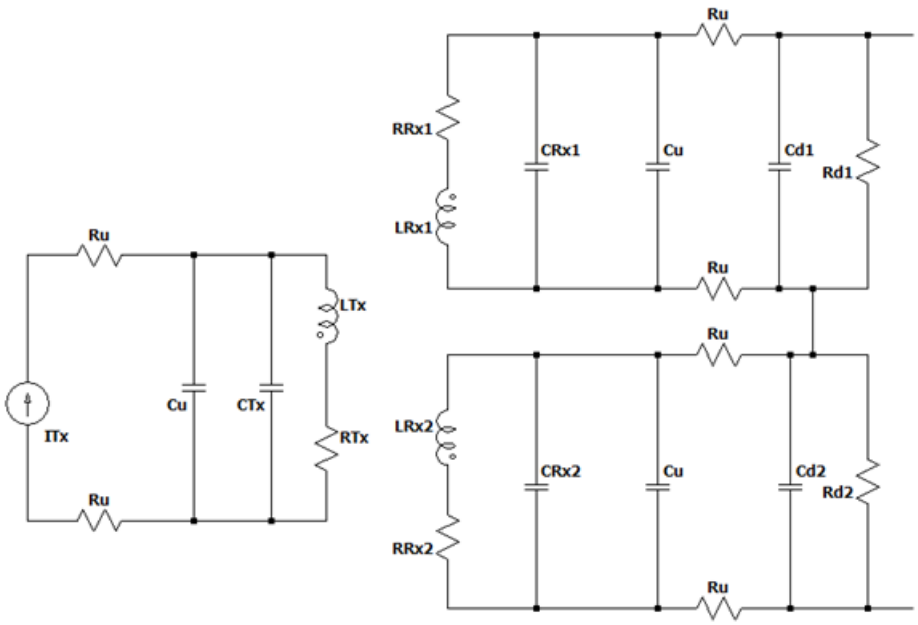

Figure 6.5: Additional damping components are added to smooth out resonant effects.

sensitivity to any of the capacitors. Additionally, the phase response of the damped nylon measurement is significantly different to the damped response of the same gradiometer in air. The equivalent circuit was found using the free space measurements from the calibration brick in the hoist, and therefore, the damped and undamped circuits were tuned to the same experimental system used for on-site testing of the damping components: it is not clear why there was little sensitivity to the capacitors on site, whilst in the model they were sufficient — perhaps the lumped equivalent circuits have a very small range in which the components accurately model the experimental system.

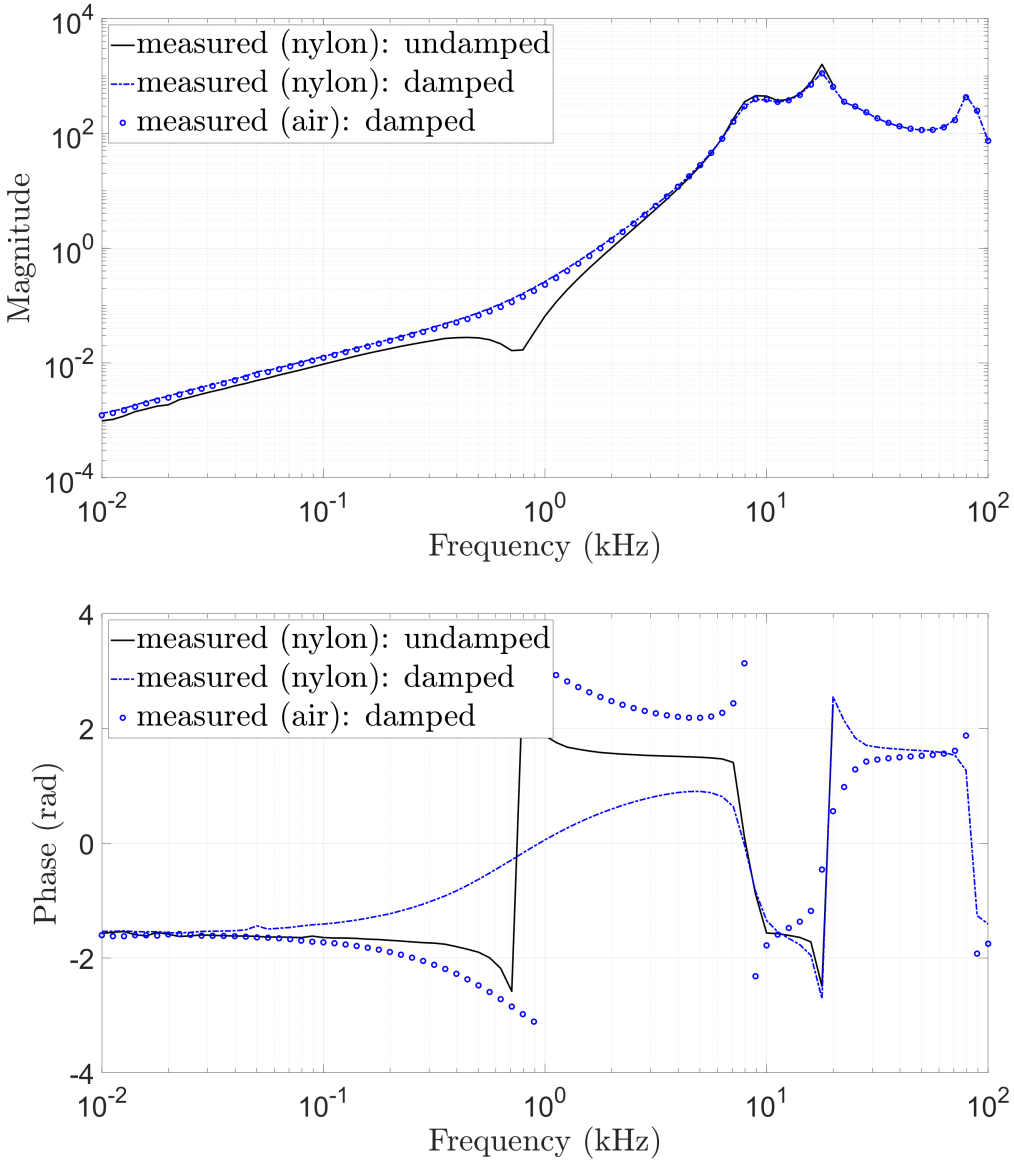

Figure 6.6: The measured free space responses with and without the damping circuitry.

6.3 Coil Characterisation and Noise

Prior to employing a FEM in an inversion algorithm, the forward model parameters need to be calibrated, such that the FEM accurately encapsulates the physical system. Further, ML methods are becoming increasingly popular for image reconstruction, wherein large amounts of training data are required. This training data can be acquired experimentally, numerically or analytically. If a numeric or analytic model is to be used to produce synthetic training data, then good agreement between model and simulated data is essential.

Calibration procedures vary in eddy current inspection, a common approach is to obtain a series of reference measurements for materials of known characteristics. Properties of the measurements can then be inferred from the reference samples; for example, in determining the presence of a notch [129]. In methods which require forward models, the calibration stage must minimise modelling inaccuracies. In electromagnetic inspection of the graphite core, calibration procedures for tuning forward models have included various techniques: varying coil diameter and finding the diameter corresponding to the best model agreement [52] or mapping the FEM and physical system using a transfer function approach [53] or a mixture of both. In [75], the forward model was calibrated by first accounting for losses and stray capacitance between coil windings and then minimising the difference between the measured and simulated responses by solving a least squares problem via the update of model lift off.

6.3.1 Constrained Optimisation Algorithm

An algorithm is needed to calibrate a set of model parameters $\vec{\eta}$ subject to constraints. This section introduces an algorithm for solving a constrained optimisation problem using a LM trust region method. Ideally, a matrix $\boldsymbol{D}$

exists which satisfies some arbitrary boundary conditions, such as in fig. 6.7. However, for most cases where there are boundaries this would require non-linear, piecewise transformations. Further, it is not practical to determine the matrix $\boldsymbol{D}$ for arbitrary constraint boundaries.

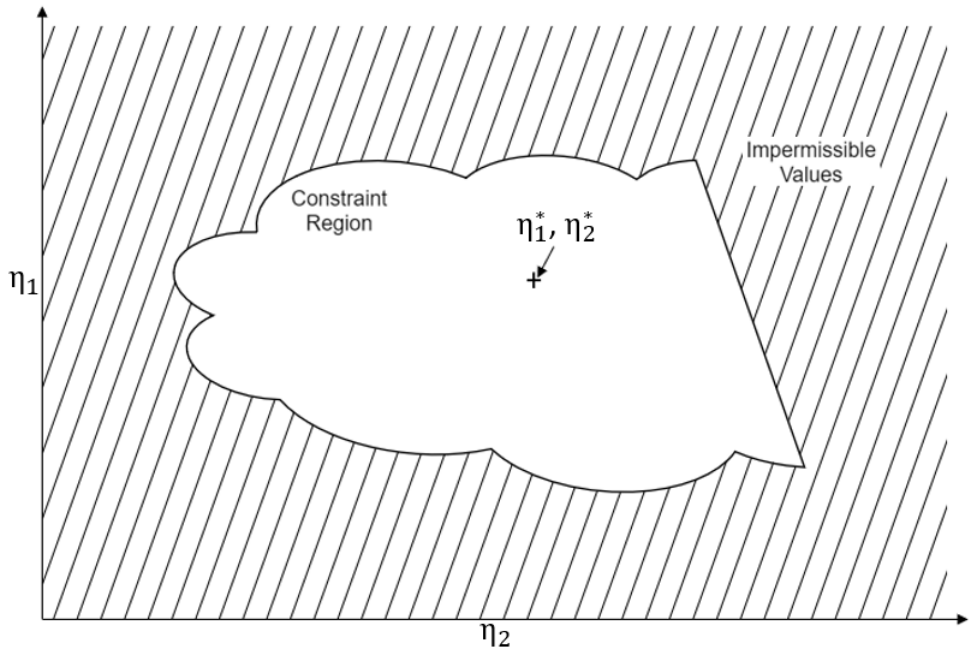

Figure 6.7: An arbitrary constraint region in an optimisation problem. The algorithm must find the minimiser of the objective function within the constraint region (in this example η_1^* and η_2^*).

Consider a set of coil parameters which are to be tuned to the experimental data (for example, lift off or turns), the update of any-one parameter, $\eta_k^j + p^j$, must be within the constraint boundary. In the algorithm developed, if a predicted value $\eta_k^j + p^j$ has a constraint violation and the previous value η_k^j was within the permissible boundary, then a bisection search was used to place $\eta_k^j + p^j$ within or on the boundary. If η_k^j is already on some boundary and $\eta_k^j + p^j$ is in an impermissible set of values, then ρ cannot be computed. To proceed, the j^{th} parameter is omitted from the step for the remainder of the iteration by eliminating the variable from the gradient and Hessian; the damping parameter is then altered until the conditions on ρ are satisfied, as normal. During the next iteration, any previously omitted parameters are re-introduced into the inversion and the full process above repeated. This

freezing step ensures that other parameters can take on new values towards a minimum and, thus, the inversion can progress. In practice it has been observed that values that have been frozen later move from the boundary into the permissible space, as the algorithm navigates around some boundaries in the parameter space towards a minimum. The set of variables that are within the feasible region are referred to as $\boldsymbol{F}_1$, those that are to be used in a bisection as $\boldsymbol{F}_2$ and $\boldsymbol{F}_3$ for those that are to be frozen — these three sets are disjoint and are formally written in eq. 6.1.

$$\begin{aligned}
(l_j < \eta_k^j < u_j) \wedge (l_j \leq \eta_k^j + p^j \leq u_j), j \in \boldsymbol{F}_1 \\
(l_j < \eta_k^j < u_j) \wedge (\eta_k^j + p^j < l_j \vee \eta_k^j + p^j > u_j), j \in \boldsymbol{F}_2 \\
(\eta_k^j \leq l_j \vee \eta_k^j \geq u_j) \wedge (\eta_k^j + p^j \leq l_j \vee \eta_k^j + p^j \geq u_j), j \in \boldsymbol{F}_3
\end{aligned} \tag{6.1}$$

The algorithm used for the bisection search is shown in algorithm 3. Within the bisection search, the γ corresponds to a step such that the resultant point is close to the boundary because it is numerically difficult to find a step landing exactly on the boundary. Due to this, $\gamma = 10^{\text{low}}$ is returned, corresponding to a violation slightly beyond the boundary and into the impermissible region, where this ensures that the $j \in \boldsymbol{F}_2$ will correctly be considered in the variables to be frozen in the next iteration. Alternative search algorithms could be used; however, the relative speedup between algorithms is negligible and a bisection is used for its simplicity.

The constrained algorithm used is shown in algorithm 4. The first stage is to perform a bisection search to find the damping parameter corresponding to all the variables in $\boldsymbol{F}_2$ moving on or within the boundary; this stage is completed first because the indices in $\boldsymbol{F}_3$ may change with γ. The variables in $\boldsymbol{F}_3$ are then frozen by truncating the gradient and Hessian. Two bisection searches in a single iteration are not permitted. The next stage is to compute ρ and determine whether the quadratic model constraint region requires alter-

Algorithm 3: Bisection Search for the damping parameter exponent.

low = $\log_{10}(\gamma)$;
upper = $\log_{10}(\gamma)$;
determine $\boldsymbol{F}_2$ with $\vec{p}(\gamma = 10^{\text{upper}})$;
while $\boldsymbol{F}_2$ *is not empty* **do**
    upper = upper $* 2$;
    determine $\boldsymbol{F}_2$ with $\vec{p}(\gamma = 10^{\text{upper}})$;
end
$i = 1$;
for $i <$ *bisection iterations* **do**
    mid = (low + upper)/2 ;
    compute $\vec{p}(\gamma = 10^{\text{mid}})$;
    determine $\boldsymbol{F}_2$ with $\vec{p}(\gamma = 10^{\text{mid}})$;
    if $\boldsymbol{F}_2$ *is not empty* **then**
        low = mid;
    else
        upper = mid;
    end
    $\gamma = 10^{\text{low}}$;
    $i+ = 1$;
end
return γ;

ing; an important stage is to check if a bisection search has been performed in the check for the over-constrained ρ — this is required because the damping parameter cannot be lowered as it will potentially place the variable further into the impermissible region. However, the damping parameter is free to be increased.

For the damping parameter selection if $0.8 < \rho < 0.95$ or $1.05 < \rho < 1.2$ then the step is accepted, otherwise the region is contracted by increasing the damping parameter. The value of $\rho = 1$ was avoided as it is possible the constraint region is contracted to the point that the numerator and denominator tend to zero (producing numerical instability) and also provides the ability to lower γ.

Algorithm 4: Computing the step direction

ρ = Nan;
count = 1;
already bisected this iteration = False;
while ρ *does not satisfy constraints or count* < *count limit* **do**
    compute $\vec{p}(\gamma)$;
    find $\boldsymbol{F}_2$ corresponding to $\vec{p}(\gamma)$;
    if $\boldsymbol{F}_2$ *is not empty and not already bisected this iteration* **then**
        γ = compute bisection();
        already bisected this iteration = True;
        compute $\vec{p}(\gamma)$;
    find $\boldsymbol{F_3}$ corresponding to $\vec{p}(\gamma)$;
    if $\boldsymbol{F_3}$ *is not empty* **then**
        Truncate gradient and Hessian;
        Compute reduced step, $\vec{p}_r$, corresponding to indices in $\boldsymbol{F}_1 \cup \boldsymbol{F}_2$;
        Assemble full step $\vec{p}$ using $\vec{p}_r$, where $p^j = \begin{cases} 0, & j \in \boldsymbol{F}_3 \end{cases}$
    compute ρ using step, gradient and Hessian corresponding to $\boldsymbol{F}_1 \cup \boldsymbol{F}_2$;
    if *over-constrained* **then**
        if *already bisected this iteration* **then**
            $\vec{\eta}_{k+1} = \vec{\eta}_k + \vec{p}$;
            break;
        $\gamma = \gamma/2$;
    else if *under-constrained* **then**
        $\gamma = \gamma \times 3$;
    else if *Not under-constrained or over-constrained* **then**
        $\vec{\eta}_{k+1} = \vec{\eta}_k + \vec{p}$;
        break;
    count += 1;
end

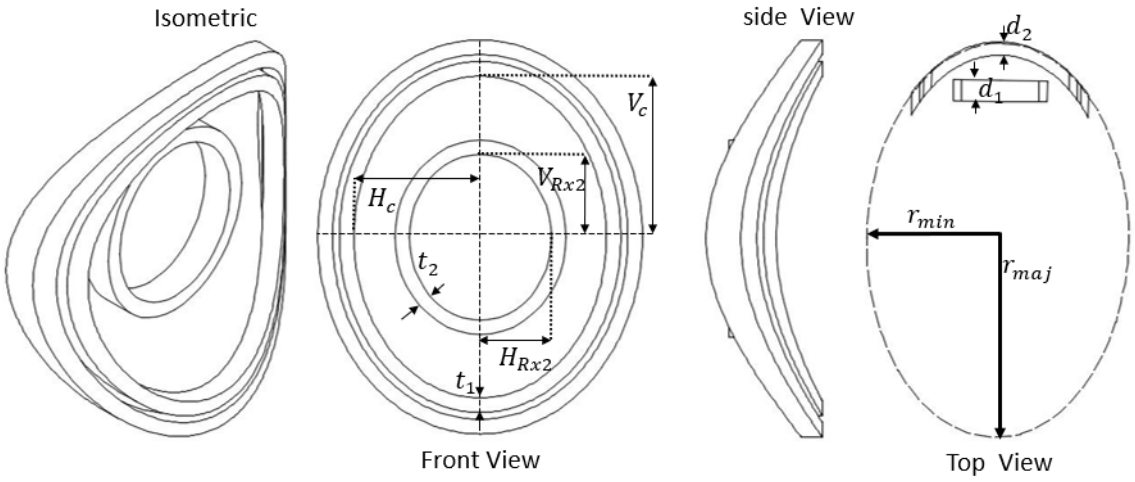

Figure 6.8: The gradiometer coil used to collect and invert reactor data, characterised by the inset parameters.

6.3.2 Constraints

The elliptical coils are defined using the parameters in fig. 6.8. The sequence of the coils is as follows: Rx_2-Tx-Rx_1, from smallest to largest. The large receive (Rx_1) and transmit coil (Tx) lie on the same plane, where the curvature of the surface is described by the major and minor radii of a 2D ellipse. In contrast, the small receive coil (Rx_2) lies in a flat plane. All coils are elliptical in shape, as seen in the front view. In order to simplify the problem, the depth (d) and thickness (t) of Rx_2, Rx_1 and Tx are fixed; the wire cross section is set so that the turns fit exactly according to the fixed depth and thickness. The fixed values are $d_1 = 14$ mm, $d_2 = 9$ mm, $t_1 = 6.35$ mm and $t_2 = 6.5$ mm.

The diameters of Rx_1 and Tx are not updated individually but instead defined by the inner diameters of the curved plane, V_c and H_c. The advantage of which is a reduced set of variables. For example, the horizontal inner diameter of Rx_1 is $H_{Rx_1} = H_c + t_1 + 3.3 \times 10^{-3}$ and $H_{Tx} = H_c$, where 3.3 mm is added because this is the measured separation between the Rx_1 and Tx coil. Similarly, a single lift off parameter is defined, L_c, for the Tx-Rx_1 pair. The smaller backing off coil has its own lift off parameter, L_{Rx2}. The number of turns, T, on each coil is treated individually.

The constraints used are given in table 6.1. The constant s is the minimum separation between the inner diameters of the Rx_2 and Tx and is defined as the Rx_2 coil thickness plus some offset. The offset used was 3 mm to ensure the coils can be suitably meshed. Additionally, the upper limit of Rx_2 is tied to the lower of Tx; therefore, if one is in violation then so is the other, enforcing the sequence of the coils. The lift off for each coil has no upper bound, but has a lower constraint of 3 mm, again to ensure the coils can be suitably meshed. The rest of the variables are constrained using the measured values with a 25% tolerance.

Table 6.1: Constraints used in coil modelling

Parameter	l^j	u^j
H_{Rx2}	0.001	$\frac{H_c+H_{Rx2}-s}{2}$
H_c	$\frac{H_c+H_{Rx2}+s}{2}$	∞
V_{Rx2}	0.001	$\frac{V_c+V_{Rx2}-s}{2}$
V_c	$\frac{V_c+V_{Rx2}+s}{2}$	∞
T_{Rx2}	562.5	937.5
T_{Tx}	75	125
T_{Rx1}	112.5	187.5
L_{Rx2}	0.003	∞
L_c	0.003	∞
r_{min}	0.1	0.12
r_{maj}	0.1	0.12

To compute the Jacobian, the perturbation method was used. In this approach, the sensitivities are found using the forward difference method, given by eq. 6.2, where f is the objective function. For large values of γ (decreasing constraint region radius), the gain ratio can become numerically erratic; this is because the scale of the constraint region radius approaches that of the perturbation. Therefore, the damping parameter was limited to $\gamma < 10$.

$$\frac{\partial f(\vec{\eta})}{\partial \eta_k^j} \approx \frac{f(\eta_k^j + \Delta\eta_k^j) - f(\eta_k^j)}{\Delta\eta_k^j} \tag{6.2}$$

6.4 Results and Discussion

6.4.1 Laboratory System

In the calibration procedure, it is prudent to use a variety of calibration blocks with different conductivities, such that the calibration error would be ideally conductivity invariant. Further, the use of multiple calibration bricks helps

prevent over-fitting. In the laboratory, three blocks were available that have Low Conductivity (LC), Medium Conductivity (MC) and High Conductivity (HC), where high and low are respective to the nuclear graphite application. The conductivity of the calibration bricks was physically measured using a four point probe [130]. The conductivities determined were 11.0 kS/m, 39.6 kS/m and 85.3 kS/m. A homogeneous cuboid block of graphite (see fig. 6.9) was used and provided a basic geometry which could be easily and accurately be defined and was faster to compute, compared with station geometry. A template was used to fix the position of the coil for measurement consistency. The objective function which was used is given by eq. 6.3. The term $\vec{r}^k(\vec{\eta})$ is the residual error between the measured and simulated response of the k^{th} calibration block and n is the number of calibration blocks. No regularisation was used.

$$f = \frac{1}{2}\sum_{k=1}^{n}\left\|\vec{r}^k(\vec{\eta})\right\|_2^2 \tag{6.3}$$

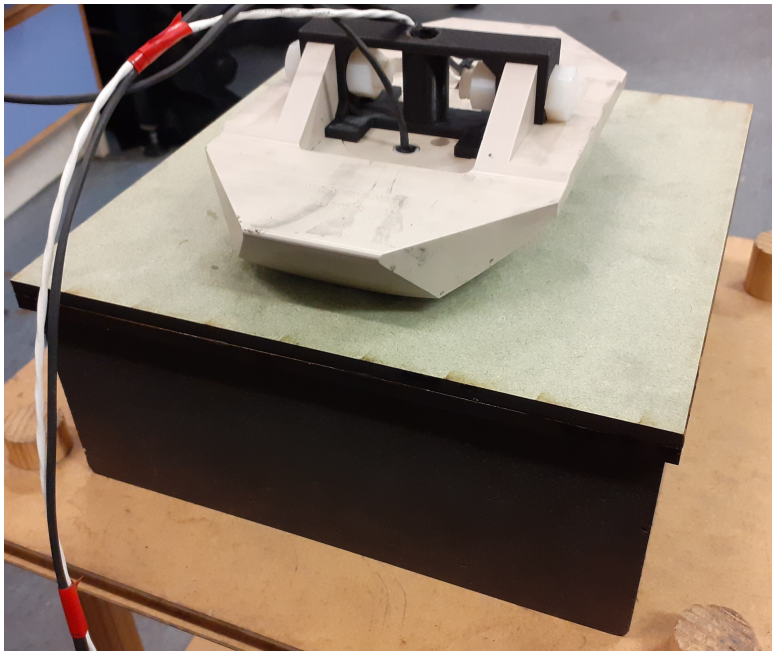

Figure 6.9: The setup used for calibration of the laboratory system; a cuboid block of graphite with a template to fix the sensor position.

To model the experimental system, a filament or 3D coil model can be used. In the filament model, each receive and transmit coil is approximated with a wire filament, which has zero cross sectional area. Within the filament model, the total voltage across a coil can be found by integrating the flux penetrating the filament and multiplying by the number of turns, assuming the flux through each coil is the same. However, the field is spatially varying and the amount of flux penetrating each turn of the physical coil is not necessarily uniform. This could be improved through the use of more filaments, capturing the spatial variation of penetration. The use of a filament model is simplistic but results in faster computation times because the number of finite elements in modelling the coil is reduced. In contrast, the 3D model captures the variation in flux between turns. Therefore, this model is generally more accurate but has longer computation times.

The agreement of the FEM with the physical measurements was quantified using the SNR (θ), given by eq. 6.4, where m_s^i is the simulated measurement, m_p^i the actual at the i^{th} frequency measurement. The frequency range was initially 10 Hz - 10kHz. After tuning to this it was truncated according to the range of frequencies with the largest SNR.

$$\theta^i = 20 \times \log_{10} \left(\frac{\left| m_s^i \right|}{\left| m_p^i - m_s^i \right|} \right) \tag{6.4}$$

The experimental and simulated filament model responses *prior* to optimisation are shown in fig. 6.10 and the SNR given in table 6.2. Similarly, the measurements *post* optimisation are shown in fig. 6.11 and the SNR given in table 6.3.

From fig. 6.11 it can be seen that the experimental and simulated responses begin to clearly deviate at 215 Hz in the real part of the measurement of the LC block; the low frequency error is expected because the measurement

Table 6.2: Laboratory system: SNR of the filament model using the full spectrum, prior to optimisation.

frequency (Hz)	HC (dB)	MC (dB)	LC (dB)
10	0.62	-4.58	-5.39
21.54	5.54	4.22	7.27
46.42	5.27	3.52	2.91
100	5.29	3.67	3.96
215.44	5.24	3.7	3.95
464.16	5.17	3.73	3.95
1000	5.05	3.73	4.02
2154.43	5.09	3.83	4.13
4641.59	5.25	4.28	4.52
10000	6.43	5.79	6.25

signal decreases in accordance with Faraday's law and the limiting noise floor of the measurement device. Therefore, the spectra are truncated to 464 Hz - 10 kHz and optimisation procedure repeated.

The results for the truncated spectra are shown in table 6.4. From table 6.3 and 6.4 it can can see that there is not much improvement in the SNR's, thus, the lower frequency measurements were not necessarily limiting the convergence of the model parameters. Further, there may be coil modelling or measurement inaccuracies which are limiting convergence.

The procedure was repeated with the 3D coil model, where the truncated results are shown in table 6.5. The mean SNR across the truncated spectrum and across the three calibration bricks is 29.43 dB for the 3D model and 28.27 dB for the filament. From this, it is clear that there is little difference between the two coil models, and therefore, the non-uniformity of the magnetic field between turns and coil losses were not the deciding factors between the two models. This could indicate that there is a common modelling error in both cases or that there is systematic error in the measurement data.

Table 6.3: Laboratory system: SNR of the filament model using the full spectrum, post optimisation

frequency (Hz)	HC (dB)	MC (dB)	LC (dB)
10	11.45	-0.72	-3.77
21.54	32.29	18.26	22.84
46.42	39.19	21.03	16.83
100	43.29	22.46	22.73
215.44	41.32	22.82	23.73
464.16	36.43	22.95	24.25
1000	32.88	22.84	24.82
2154.43	31.23	22.99	25.71
4641.59	31.36	24.75	28.75
10000	31.97	46.16	29.24

Table 6.4: Laboratory system: SNR of the filament model using the truncated spectrum, post optimisation

frequency (Hz)	HC (dB)	MC (dB)	LC (dB)
464.16	32.54	26.71	22.44
1000	30.28	26.08	22.84
2154.43	29.82	25.49	23.46
4641.59	31.59	27.64	26.01
10000	30.99	35.03	33.09

A four point probe was used to measure the electrical conductivity of the blocks; therefore, the accuracy is potentially limited by the four point probe measurement. This would also go some way in explaining the asymptotic nature of the accuracy in both coil models. The assumption made was that the electrical conductivity distribution throughout the bricks is homogeneous and this is not necessarily a very good assumption, since there can be some spatial variations. To gain accuracy, it may be better to obtain bricks which have greater homogeneity or an electrical conductivity measurement more indicative of the brick average; the latter can be obtained by using a different

four point geometry (greater distance between probes) and a larger number of measurements at different locations. Furthermore, accuracy may be limited by the finite element mesh. The optimisation procedure was repeated with a finer mesh within the calibration block, using 111810 elements instead of the 40811 previously used; the optimisation on the truncated spectrum with the filament model achieved a mean SNR of 29.09 dB, where previously this was 28.27 dB. This indicates that the mesh size was not significantly limiting the accuracy, in this case.

Finally, the 3D model took $\approx$ 89 s to compute and the filament $\approx$ 54 s — this is for the truncated spectrum where the computations were performed on a Lenovo ThinkStation P520, with 128GB RAM. This is subject to change with coil parameters and model geometry studied, but provides a useful gauge for the relative speed up of a filament model.

Table 6.5: Laboratory system: SNR of the 3D Model using the truncated spectrum, post optimisation

frequency (Hz)	HC (dB)	MC (dB)	LC (dB)
464.16	36.96	24.13	25.65
1000	33.68	24.05	26.35
2154.43	31.28	24.11	27.31
4641.59	31.54	25.45	31.04
10000	30.61	41.93	27.4

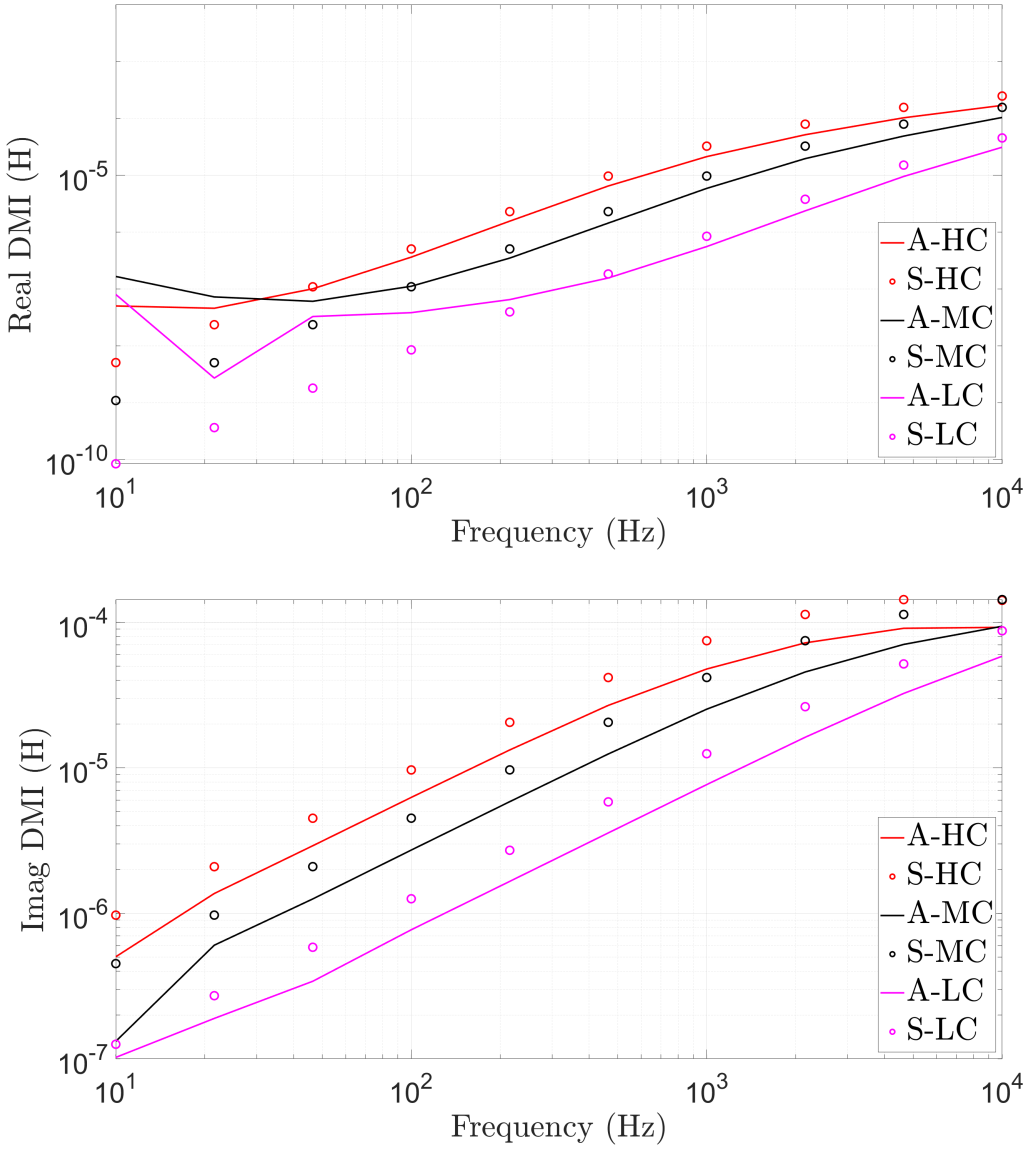

Figure 6.10: Laboratory system: the simulated filament model and actual measurement of the three calibration bricks, prior to optimisation. The prefixes A and S stand for actual and simulated, respectively.

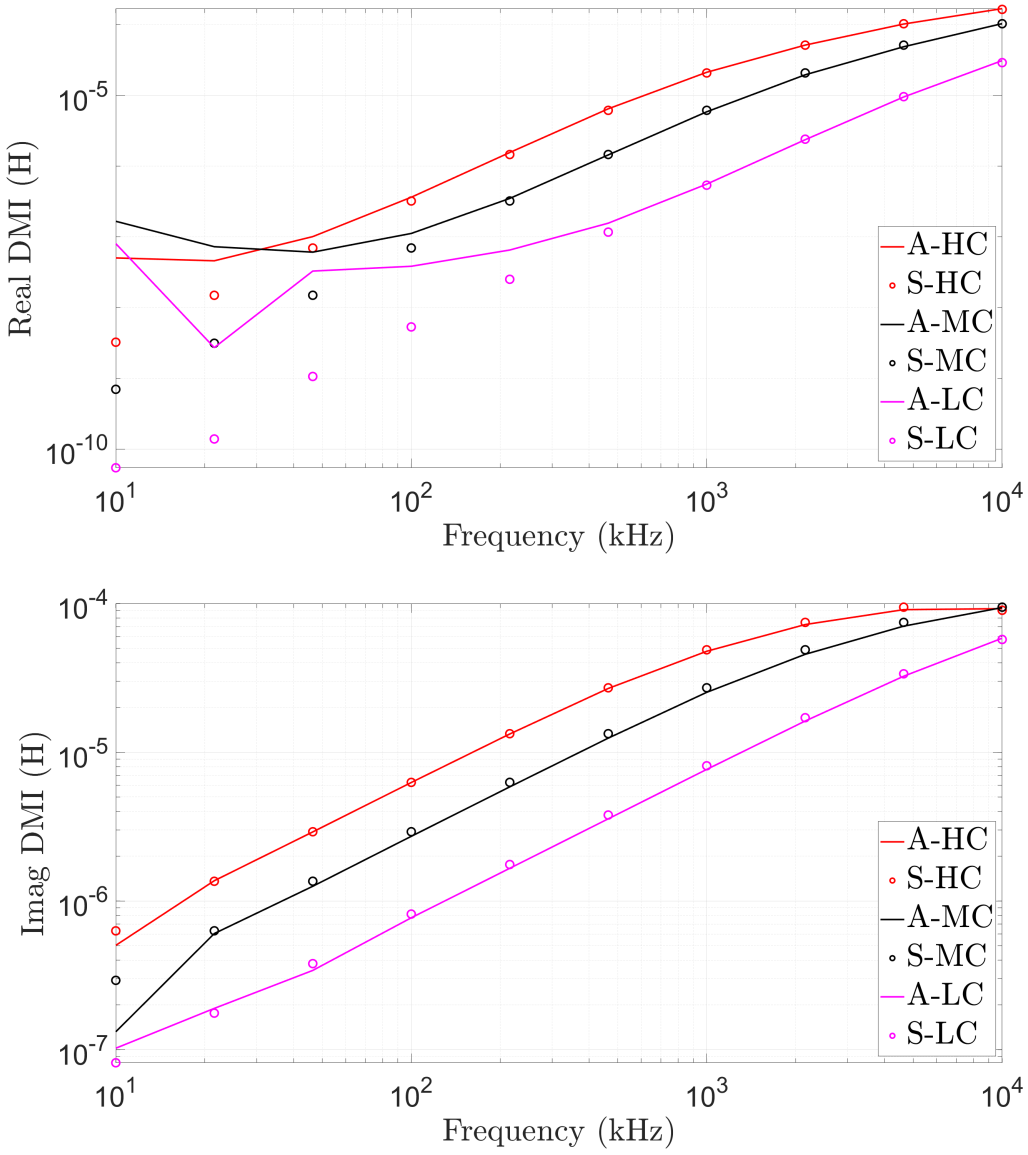

Figure 6.11: Laboratory system: the simulated filament model and actual measurement of the three calibration bricks, post optimisation. The prefixes A and S stand for actual and simulated, respectively.

6.4.2 Reactor Systems

The reactor measurement system is shown in fig. 6.12. The reactor data was collected at two sites; the two sites have different channel geometries, with different diameters, methane hole spacing and keyway size. The data at one site was collected with the damped gradiometer probe and the other site the undamped, where the gradiometer coil, hoist and umbilical were the same: the only differences between the two sites were the graphite channel geometry and the presence of damping circuitry. This provided an ability to study the effectivity of damping, and specifically, if the FEM accuracy is improved upon. There was only one set of calibration data for each of the deployed damped and undamped gradiometer coils. In the laboratory, it was simple to obtain small graphite samples of differing conductivities. Practical issues complicate the choice of calibration brick when on site. These issues are: the availability or transport of graphite, the extended distance of the probe from the tool and the fixed vertical position of the gradiometer because it is deployed from a hoist.

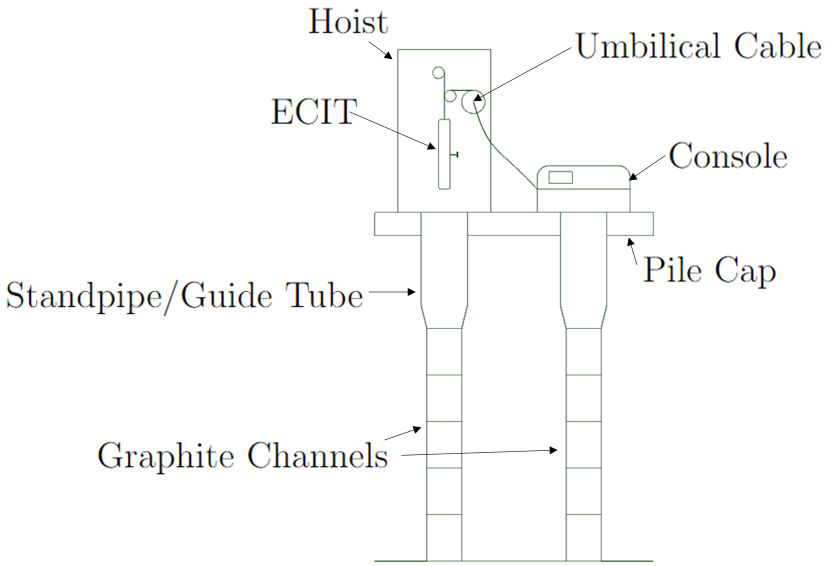

Figure 6.12: Schematic of reactor measurement system, including the hoist, pile cap, the ECIT and console where the data is collected.

The objective function used to calibrate the laboratory data was the least squared absolute error. This was suitable since the COMSOL® physics was adequate for modelling the experimental setup. However, although damping was added to mitigate resonance, there are still possibly resonant effects in the damped system. The least squared absolute error objective function would be most sensitive to the frequency point with the largest absolute error, i.e. the high frequency data which would be affected by resonance to a greater degree than the low frequency. Therefore, instead of the absolute error, the relative error was used given by eq. 6.5, where k is the number of real and imaginary components.

$$f = \frac{1}{2}\sum_{i=1}^{k}(\frac{m_s^i - m_p^i}{m_p^i})^2 \tag{6.5}$$

Damped System Calibration

The damped gradiometer was deployed on graphite channels with an inner bore diameter of 270 mm. The conductivity of the calibration brick was determined, from four point probe measurements, to be 90.5 kS/m.

In fig. 6.13, the amplitude response is plotted from 10 Hz to 100 kHz, using the measured coil parameters. It is seen that the COMSOL® simulation and experimental response cross at 4.46 kHz, diverging thereafter. Initially, the model was calibrated from 400 Hz — 4.46 kHz. The spectrum was then further truncated to the range 631 Hz - 1.26 kHz in order to improve modelling accuracy. The SNR spectrum is plotted in fig. 6.14. The frequency range cannot be lowered any further because of limited available data. The experimental and calibrated FEM DMI response is plotted in fig. 6.15, where the peak in the SNR is due to the crossing of the experimental and simulated DMI data. The mean SNR across the spectrum is 33.99 dB, which is higher than the laboratory system.

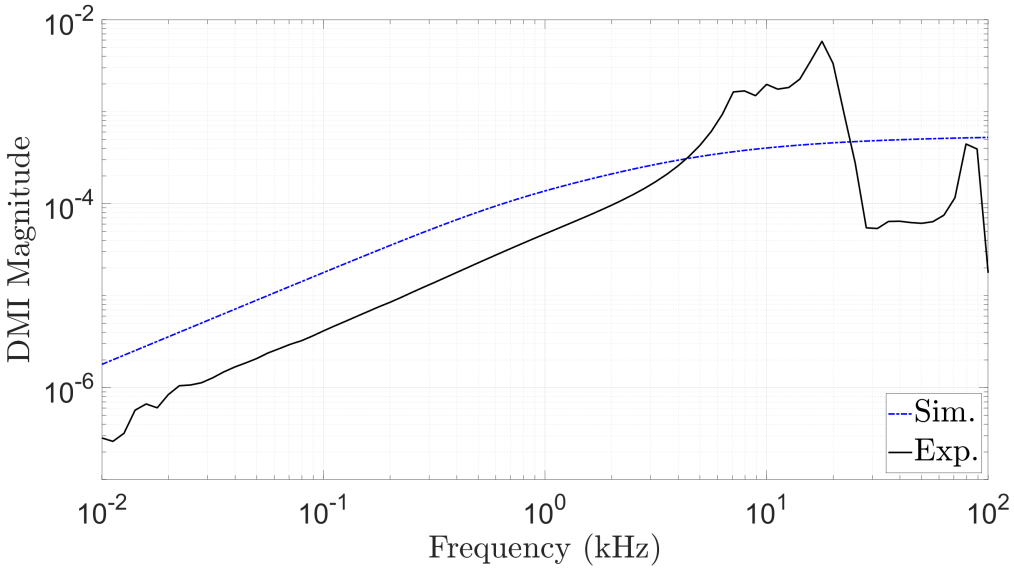

Figure 6.13: Damped reactor probe: experimental and simulated measurement amplitude

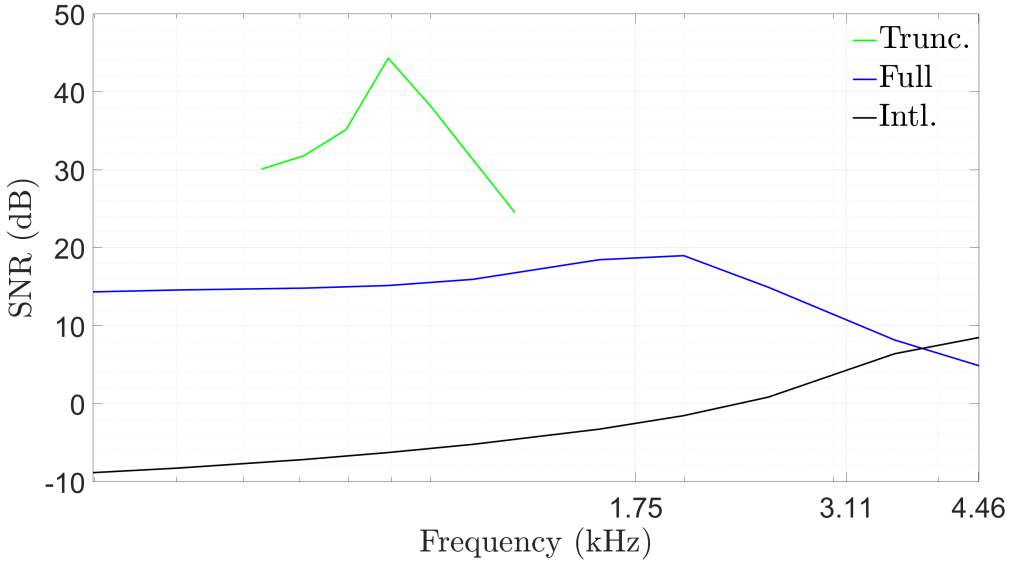

Figure 6.14: Damped reactor probe: the SNR spectra after tuning to the full frequency range 400 Hz to 4.46 kHz and truncated frequency range 631 Hz to 1.26 kHz. Included is the initial SNR spectrum when FEM is not tuned.

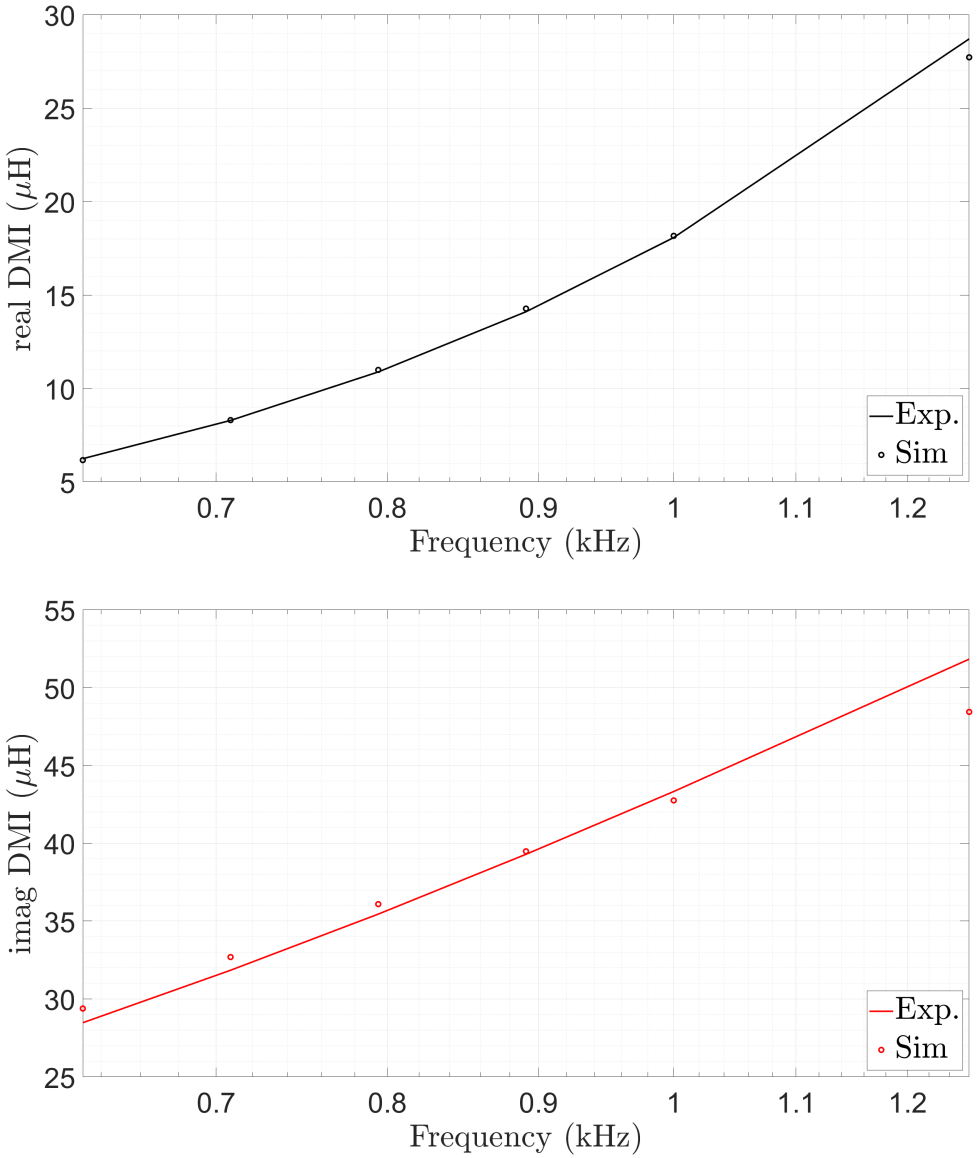

Figure 6.15: The experimental and the tuned FEM DMI response: damped reactor probe.

Undamped System Calibration

The undamped gradiometer was deployed on graphite channels with an inner bore diameter of 263 mm. The conductivity of the calibration brick was determined, from four point probe measurements, to be 80.5 kS/m.

The frequency ranges used are the same as previously. The SNR spectrum is given in fig. 6.16: comparing this to the damped gradiometer coil, it can be seen that damping does help with removing resonance effects and is not just superficial. The mean SNR across the spectrum is 20.59 dB, considerably lower than the damped and laboratory system.

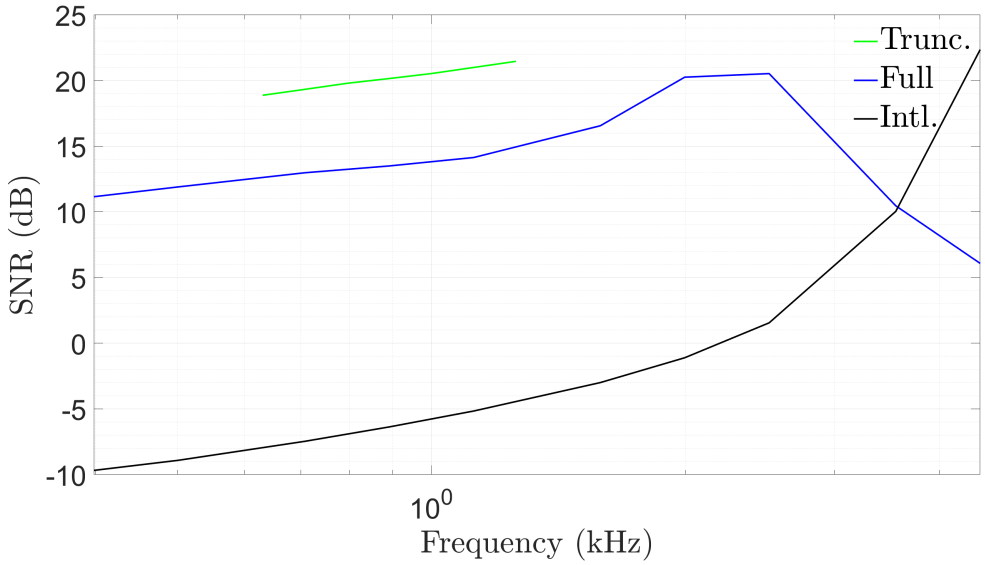

Figure 6.16: Undamped reactor probe: the SNR spectra after tuning to the full frequency range 400 Hz to 4.46 kHz and truncated frequency range 631 Hz to 1.26 kHz. Included is the initial SNR spectrum when FEM is not tuned.

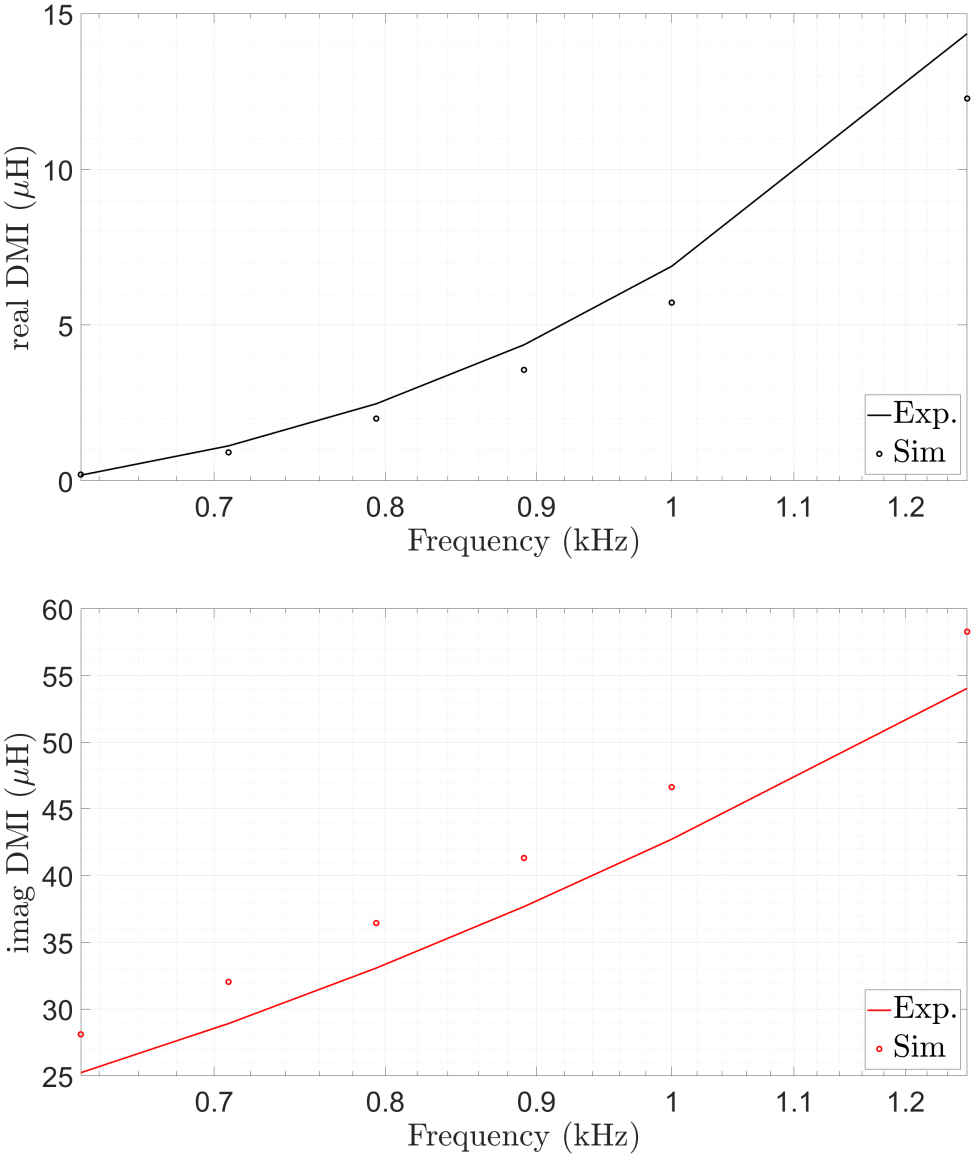

Figure 6.17: The experimental and the tuned FEM DMI response: undamped reactor probe.

6.4.3 Discussion

The average SNR of the reactor and laboratory system indicates that the damped probe is possibly modelled better than the laboratory probe, but fig. 6.18 shows that it is more nuanced. There is a peak in the SNR spectrum of the damped system because there is a crossing but the SNR drops considerably thereafter, possibly due to remaining resonance issues. However, the damped system does yield a clear increase in modelling accuracy compared with the

undamped.

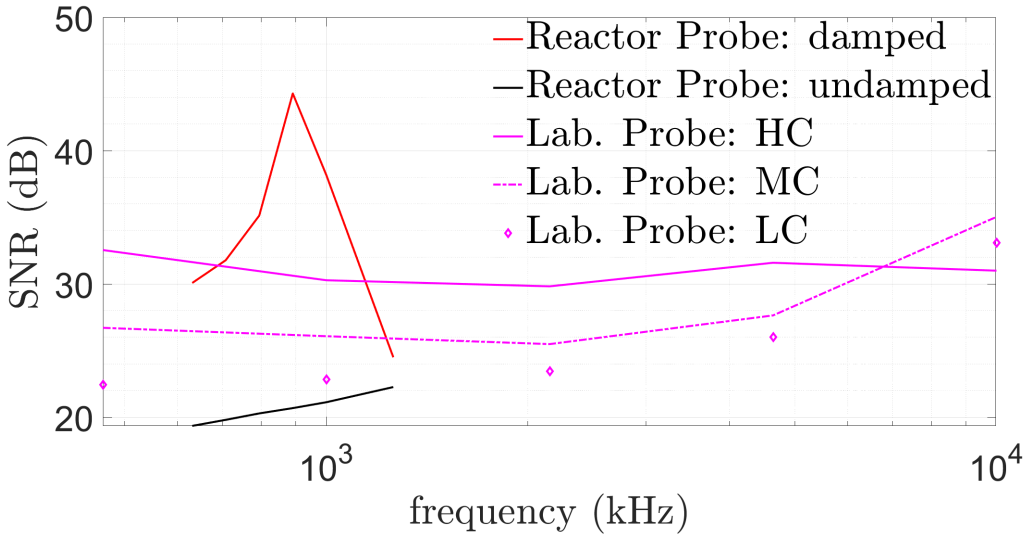

Figure 6.18: Comparing the SNR of the calibrated models of the laboratory and reactor probe.

6.5 Chapter Summary

- There are visible low frequency resonant effects in the deployed gradiometer probe air measurement. The DMI measurement would remove these effects to a first order degree, but this may not be sufficient. Damping was used to lower the severity of resonant effects.

- It may be possible to model resonanance within COMSOL®, using the electric fields or circuits and magnetic fields interface. However, inverting with resonant effects is unproved, and therefore, the strategy used was to truncated the measurement spectrum.

- A constrained optimisation algorithm has been presented and demonstrated. This calibration algorithm was used to determine that a filament model was as accurate as a 3D coil model; therefore, a filament model is used in the proceeding chapters.

- It has been shown that damping does improve the forward model accuracy, and therefore, that damping goes beyond visibly remove resonant effects.

Chapter 7

Depth Profiling the Graphite Core

7.1 Overview of Chapter

This chapter concerns the design and application of inversion algorithms for reconstructing the electrical conductivity radial depth profile of a nuclear reactor fuel channel. The reactor data was collected using the undamped gradiometer probe. There is no discussion of the damped probe within this chapter because the trepanned data was not available for the damped system at the time of writing.

The reactor sensor was calibrated using the relative error to avoid overfitting to resonant effects. However, it is unclear whether the absolute or relative error is preferential for reconstructing the depth profile; therefore, the relative and absolute errors are compared. The search direction used was the LM algorithm, using the diagonal of the Hessian as the scaling matrix and the finite difference operator as the regularisation matrix. The initial damping and regularisation parameters were determined from testing on a reserved set of synthetic data. The discrepancy principle was used for updating

the regularisation parameter, where the noise floor was determined from the calibration stage. The damping parameter was updated according to standard trust region theory. The prior estimate was determined from a MVPR algorithm. The use of the MVPR algorithm was decided upon due to the small amount of data collected.

The reconstruction of the depth profiles is first demonstrated on noise free data, where it is shown that there is a trend in increasing error in the depth profile from numerical error and ill-conditioning alone. This is the base error of the algorithm. The application of the algorithm on data with noise indicates that the relative error in the objective function is more susceptible to noise; this is because the low frequency data has a higher weighting and the low frequency data generally has a lower SNR. The inversion of reactor data is compared with trepanned data, in which there is close agreement at the bore but otherwise diverging with increasing depth.

It is seen that the algorithm is overly restrictive of the step direction. To study the restrictive movements of the iterative algorithms, the inversion of graphite inductance data is repeated with a more informed prior and varying the discrepancy multiplier. The regularisation matrix used for the repeated inversion was the identity matrix. The results show that the restricted steps are not from over-regularisation but that the steps are typically rejected because the residual norm is lower than the discrepancy. Moreover, the residual of the prior may never be large enough for the step to be accepted. To combat this, there are a number of options: a step can be accepted even if it is below the noise floor, another prior could be used in which the residual norm is larger than the discrepancy or the modelling error can be lowered.

7.2 Model Geometry Reduction

The measurement sensitivity to a graphite channel reduces with angular and vertical distance from the gradiometer. It is unnecessary to model a full graphite fuel channel, where these channels have a full height of approximately 10 metres. Further, modelling the full cylinder is not necessary. A model is sought in which the full graphite geometry can be represented with reduced dimensions with minimal loss of accuracy.

To investigate the reduced model, a brick of height 1 m was used to represent the full graphite stack, with the coil positioned in the centre. The complex measurement of the "full" model is denoted by m_{full}. The error is then given by eq. 7.1.

$$\text{Percentage Error} = \frac{100}{2}\left(\left\|\frac{\text{Re}(m)-\text{Re}(m_{full})}{\text{Re}(m_{full})}\right\|+\left\|\frac{\text{Im}(m)-\text{Im}(m_{full})}{\text{Im}(m_{full})}\right\|\right) \tag{7.1}$$

This percentage error is a function of frequency and graphite conductivity. Therefore, the percentage error is determined according to two frequencies, 400 Hz and 10 kHz, and two conductivities, 10 kS/m and 100 kS/m. The Mean Percentage Error (MPE) is then found. A filament model was used to model the coil. The graphite geometry was of the reactor type.

The MPE of the reduced model is plotted in fig. 7.1 as a function of the distance between the outer filament and the edge of the graphite. Similarly, the MPE of the reduced model is plotted in fig. 7.2 as a function of the sector angle. It can be seen that a graphite cylinder can be conservatively approximated with a 180° graphite sector. However, the MPE as a function of height has no obvious corner point in which the error suddenly rises. In previous studies, EDF Energy determined that the distance between the top of the coil and the brick edge must be no smaller than $1.5d$, where d the

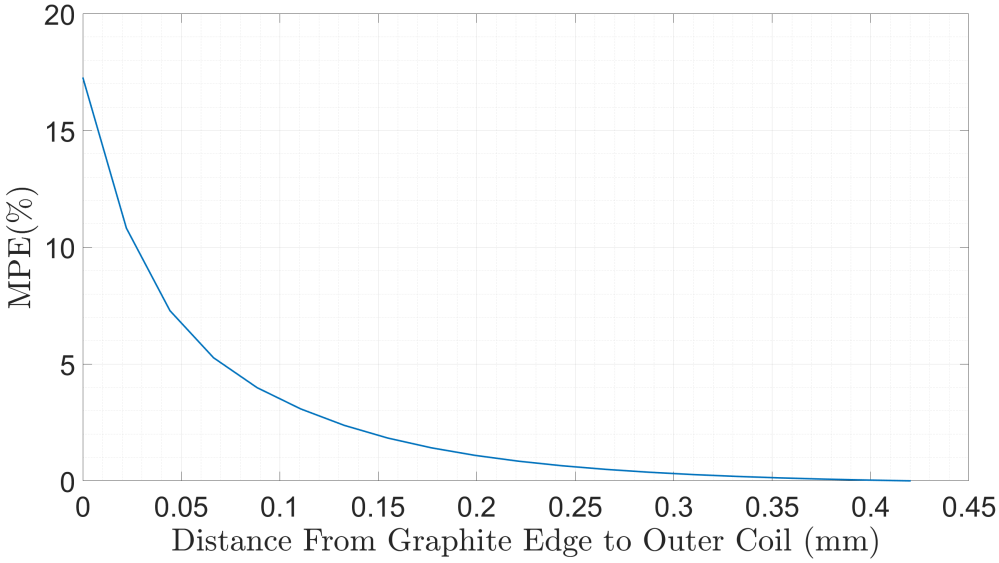

Figure 7.1: Determining the reduced model dimensions: the error with respect to graphite cylinder height.

diameter; therefore, the edge to coil distance should be no less than 230 mm — this seems reasonable according to fig. 7.1. The reduced model is shown in fig. 7.3 and 7.4.

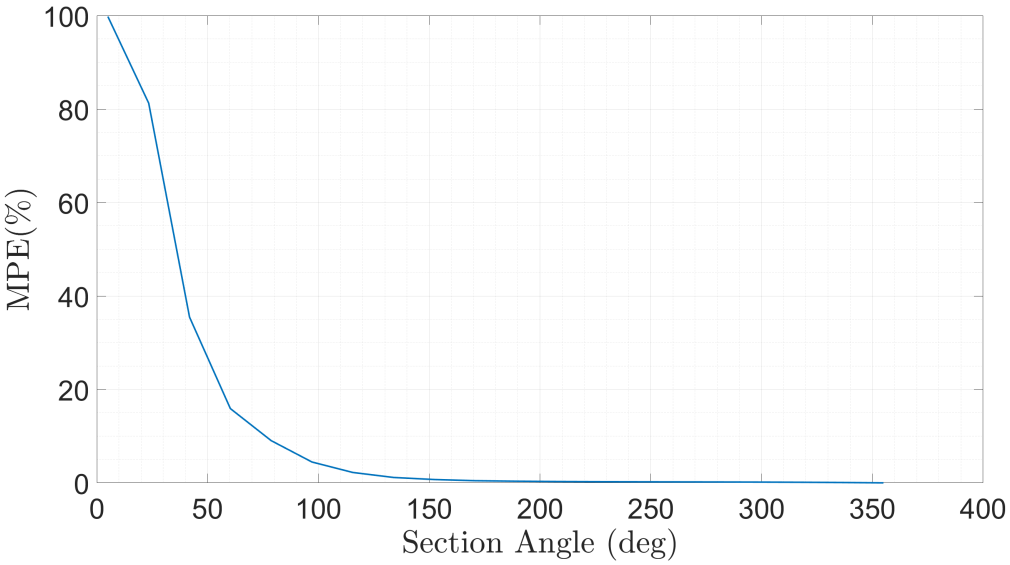

Figure 7.2: Determining the reduced model dimensions: the error with respect to the section angle.

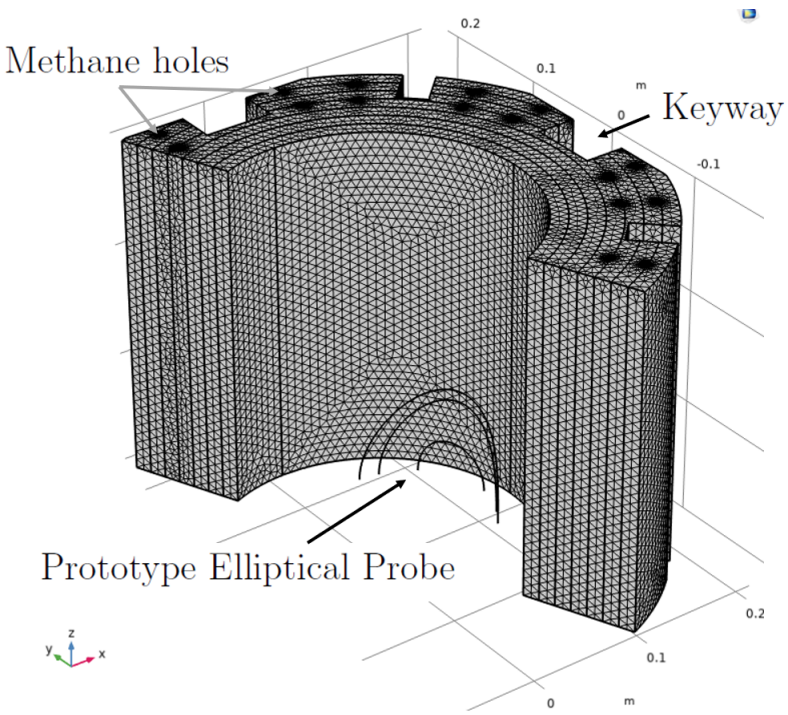

Figure 7.3: The reduced model: isometric view.

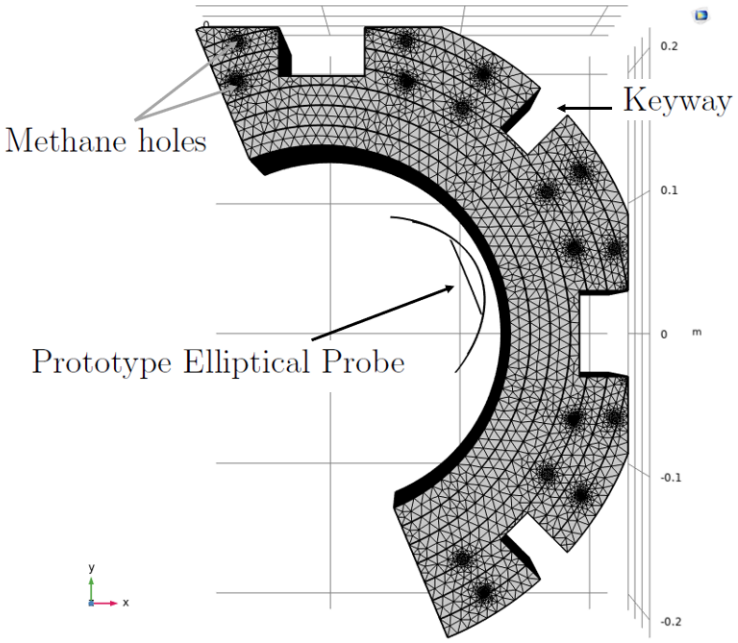

Figure 7.4: The reduced model: top view.

7.3 Direct Algorithm for Prior

7.3.1 Generating Synthetic Data

The studies in the previous chapters did not use noise profiles derived from experimental data. In these studies, the noise amplitude was arbitrarily set such that the SNR of the lowest amplitude signal was in the range of [20 60] dB. To improve this, realistic noise in the experimental data is considered; specifically, the noise profile was derived from the discrepancy between calibrated FE models and the corresponding experimental measurement.

The fewest assumptions were made about the probability distribution of the noise: the uniform random distribution was used. The synthetic noise is formulated in eq. 7.2 and 7.3, where n_i is the complex noise at the i^{th} frequency point, U is the uniform distribution and $\bar{r}_i^{cal}$ defined by eq. 7.4. The residual, r_i^m, is the residual at the i^{th} frequency for the m^{th} calibration

block.

$$\|n_i\| \sim U\{\min(\vec{r}_i^{cal}), \max(\vec{r}_i^{cal})\} \tag{7.2}$$

$$arg(n_i) \sim U\{0, 2\pi\} \tag{7.3}$$

$$\vec{r}_i^{cal} = \begin{bmatrix} \|r_i^1\| & \|r_i^2\| & \dots & \|r_i^m\| \end{bmatrix} \in \mathbb{R}^{1\times m} \tag{7.4}$$

The noise profile generated for the laboratory system is shown in fig. 7.5 for two frequency points. The true discrepancy is included to illustrate eq. 7.2 and to show the noise bounds. It can be seen that the true discrepancy across the three calibration bricks is grouped closely together, typically within a single quadrant. However, a conservative estimate was used and the noise assumed to be anywhere within the lower and upper bounds of the discrepancy amplitudes, which is the reasoning being eq. 7.3. In the case where there is only a single calibration brick, the uniform distribution reduces to the generation of a constant amplitude value. However, the phase is still randomly sampled from 0 to 2π. The discrepancy according to the undamped gradiometer is shown in fig. 7.6.

The conductivity depth profiles were determined as in §4.4.1. In total, 1000 samples were generated for training and 20 for testing. The low number of samples for testing is because of the considerable amount of time it takes to reconstruct a profile.

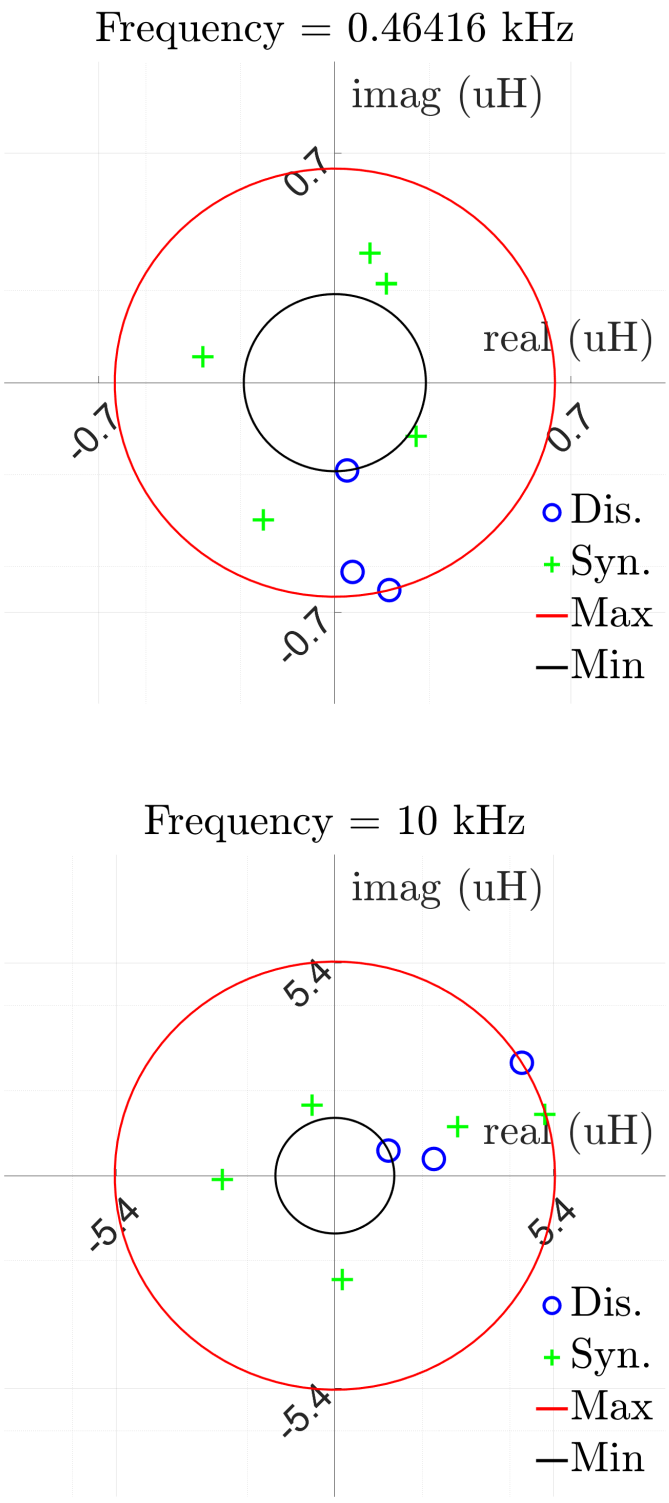

Figure 7.5: Laboratory System: noise characterisation using the calibration discrepancy from multiple bricks. The true discrepancies are denoted by dis., the synthetic noise by syn. and the bounds of the noise by max and min.

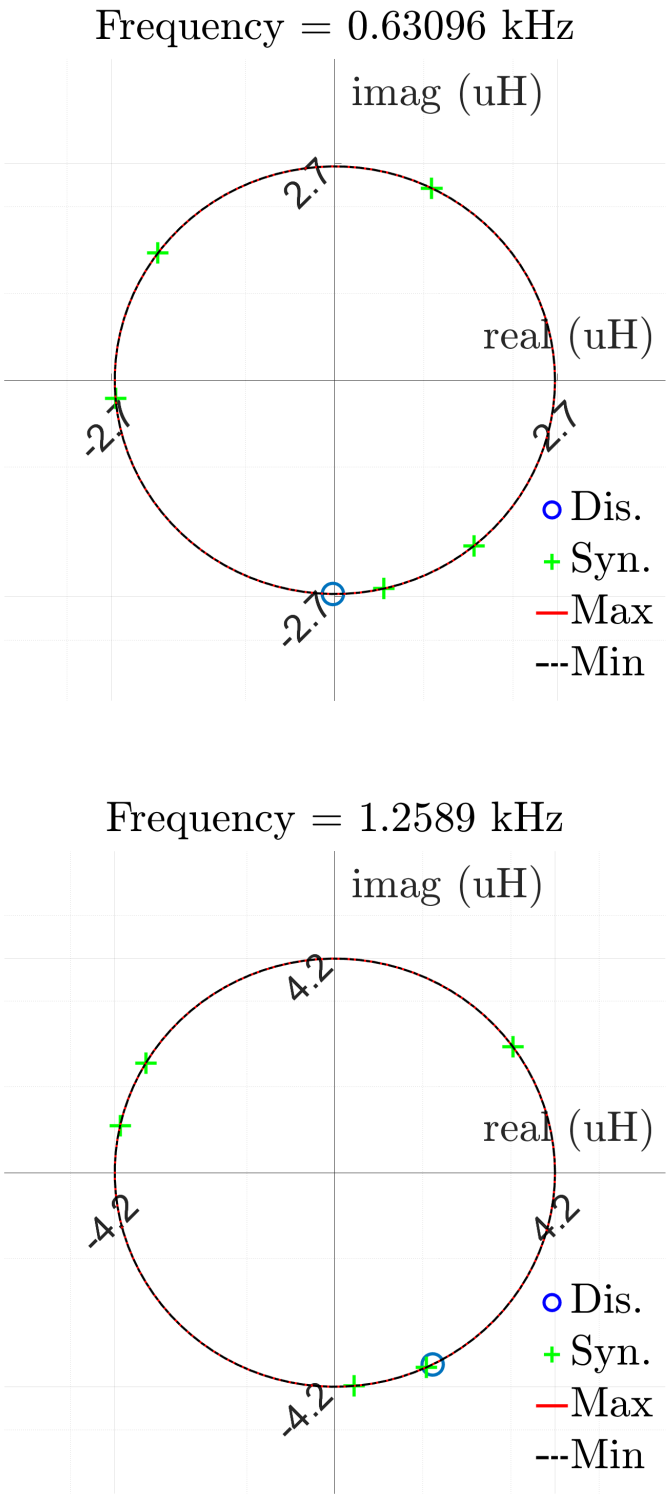

Figure 7.6: Damped reactor probe: noise characterisation using the calibration discrepancy.

7.3.2 Machine Learning Algorithm

The MVPR algorithm was used because of the low amount of training data. TSVD was used as the regularisation method, using a golden-section search to determine the truncation point. The polynomial order was determined using the same methodology as in §5.4.2.

7.4 Iterative Algorithm

7.4.1 Objective function, Constraints and Search Direction

In the calibration of the FEM, a weighted error objective function was used. The weights were chosen such that the objective function was minimising the least squared *relative* error; this gives an equal importance to the real and imaginary parts of the measurement for each spectral component. It is not immediately clear whether the absolute or relative error produces better results for depth profiling. Therefore, in this chapter, both are used and compared, where synthetic and trepanned data is used to perform error analysis.

The modelling error is expected to be too large for the missing residual terms in the Hessian to have an affect. Therefore, the search direction used was the LM algorithm. The regularisation matrix was the finite difference operator and the scaling matrix the diagonal of the Hessian.

In imaging conductivity the following constraint must be satisfied, $0 < \sigma < 100$ kS/m; therefore, the constrained optimisation approach described in §6.3.1 was used.

7.4.2 Regularisation Parameter Selection

In practice, the damping and regularisation parameters each help to prevent over-fitting to noise and in constraining the step size. Therefore, the starting value of the regularisation and damping parameters should be selected carefully; a reserved set of data (consisting of five samples) and a grid search was used to determine these initial parameters. In this search, only the initial step was computed and not the full algorithm. The initial values found to be optimal are given in table 7.1.

Table 7.1: Initial damping and regularisation parameters.

Error	λ_0	γ_0
Absolute	$5.62 \cdot 10^{-21}$	$7.19 \cdot 10^{-19}$
Relative	$1 \cdot 10^{-11}$	$1.12 \cdot 10^{-08}$

There are many methods for updating the regularisation parameter, two of which are the L-curve and discrepancy principle [131]. In the L-curve method, $\|\vec{r}\|$ is plotted against the solution norm, $\|\vec{\sigma}_k\|_2^2$, and the regularisation parameter corresponding to the corner point of this plot is used [131][132]. The corner point can be determined from the maximum curvature and can be found using optimisation techniques such as the golden-section search [133]. Tesfalem et al. found the regularisation parameter by observation of the L-curve and fixed it throughout the inversion algorithm [120][52]. In this thesis, the discrepancy principle was used to control the regularisation parameter.

There is no knowledge of the true discrepancy between the FEM and any given measurement. The discrepancy across the real and imaginary parts of the spectrum was predicted using the discrepancy in the calibrated data. The real and imaginary parts of the calibration discrepancy were concatenated into the vector $\vec{d}$. The discrepancy principle is given in eq. 7.5 [131].

$$\left\|\vec{r}(\vec{\sigma}_k)\right\| > \left\|c \cdot \vec{d}\right\| \tag{7.5}$$

The discrepancy multiplier, c, is user defined and is typically $c = 1$. If the residual curve is below the discrepancy then the regularisation parameter is increased and vice versa for above. This procedure is given by algorithm 5. It can be seen that the conductivity depth profile is only updated if the residual norm is above the discrepancy norm, otherwise the step is rejected. Fig. 7.7 and 7.8 each illustrate the principle of the algorithm: if the residual norm is below the discrepancy, then no step is taken and the regularisation parameter is increased. By increasing the regularisation parameter the objective function is weighted more towards the penalty and it is assumed that the residual norm increases.

Algorithm 5: Regularisation parameter update

if $\left\|\vec{r}(\vec{\sigma}_k + \vec{p})\right\| < \left\|c \cdot \vec{d}\right\|$ **then**
    $\lambda_{k+1} = \lambda_k * m_k$;
    $m_{k+1} = m_k * 2$;
else if $\left\|\vec{r}(\vec{\sigma}_k + \vec{p})\right\| > \left\|c \cdot \vec{d}\right\|$ **then**
    $\lambda_p = \lambda_k/3$;
    $m_{k+1} = 2$;
    $\vec{\sigma}_{k+1} = \vec{\sigma}_k + \vec{p}$;
    if $\lambda_p > \lambda_0$ **then**
        $\lambda_{k+1} = \lambda_p$;
    else
        $\lambda_{k+1} = \lambda_0$;

For the damping parameter update, if $0.8 < \rho < 0.99$ or $1.01 < \rho < 1.2$ then the step was accepted, otherwise the trust region was contracted. To update the damping parameter the Nielson method was used [134].

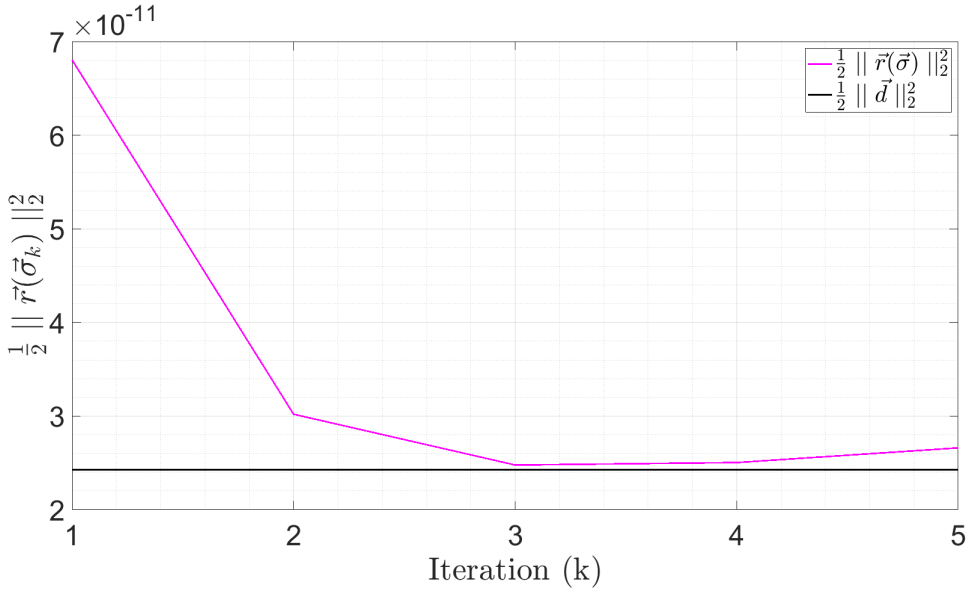

Figure 7.7: Updating the regularisation parameter: initial residual norm above discrepancy norm.

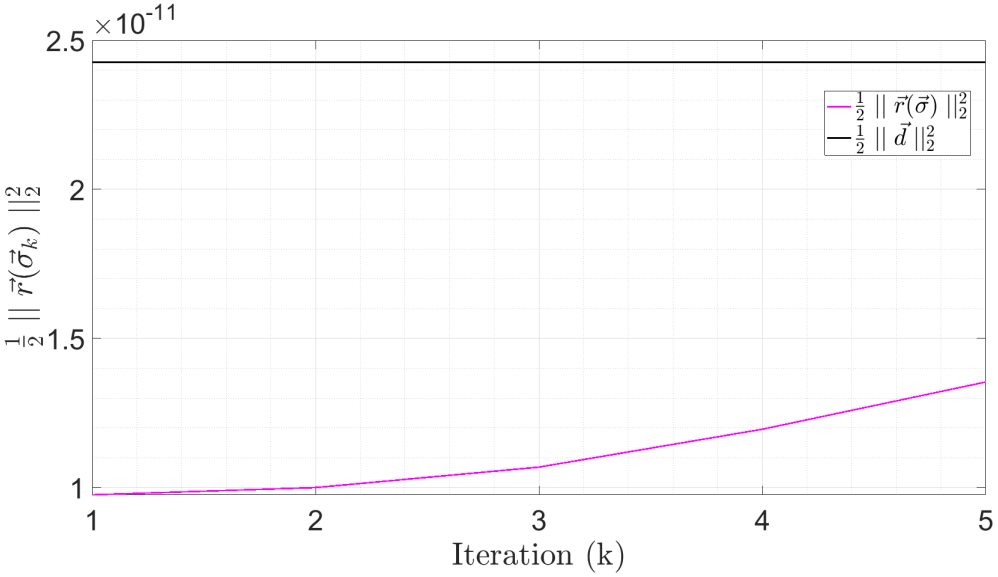

Figure 7.8: Updating the regularisation parameter: initial residual norm below discrepancy norm

7.4.3 Exit Conditions

There was a timeout condition on the search for the damping parameter — if ρ was not satisfied after 5 updates of the damping parameter then the current iteration was aborted, not the full inversion. This allows updates to the regularisation parameter, as it may be noise causing poor model agreement. There was also a timeout condition if $\rho > 10$; this was required not because of being under constrained but because ρ can diverge if the constraint region was too small (error in the Hessian, gradient and measurement will begin to dominate). Finally, the optimisation algorithm stops if $f(\vec{\sigma}_k + \vec{p}) > f(\vec{\sigma}_k)$.

7.4.4 Laboratory Demonstration

This algorithm has been successfully demonstrated on laboratory data, where a reconstructed profile is shown in fig. 7.9. The noise was quantified from the respective calibration data and the regularisation parameter adjusted as described by algorithm 5. This plot corresponds to the squared absolute error objective function. The initial damping and regularisation damping parameters were different, but obtained via the same procedure.

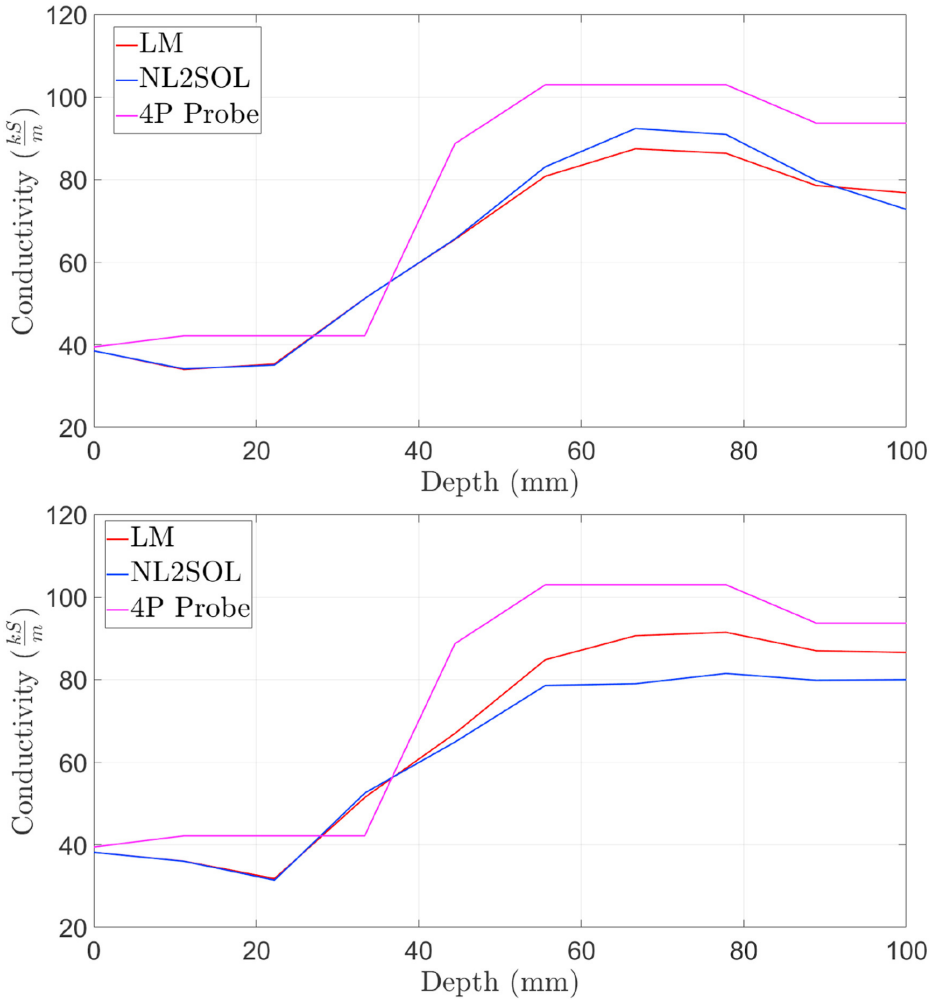

Figure 7.9: Demonstration of algorithm on laboratory data: Top). ML prior. Bottom). Homogeneous conductivity prior

7.5 Inversion Results

7.5.1 Synthetic Results and Discussion

Without Noise

Although a trivial inverse problem, using synthetic data without noise offers the ability to observe the ill-conditioning of the problem and to check the algorithms work as expected. The error in the reconstructed profiles is shown in the boxplot in fig. 7.10, where the interquartile range, median, maxima and minima of the error are shown. The trend in error is clear: error increases

with depth, where this is expected in accordance with the ill-conditioning. This shows that the algorithms are highly accurate near the bore and more erroneous at the periphery. This error can be used to define the base performance of the algorithm, where the expected error can be no smaller than this with realistic measurement data.

Overall, the optimisation algorithms only marginally improve on the ML prior. This could be because the ML prior is sufficiently accurate in this simple case (prior estimate being close to the solution) or possibly over-regularisation. It is suspected to be the former. It is difficult to make comments regarding the two different objective functions as there are no obvious differences.

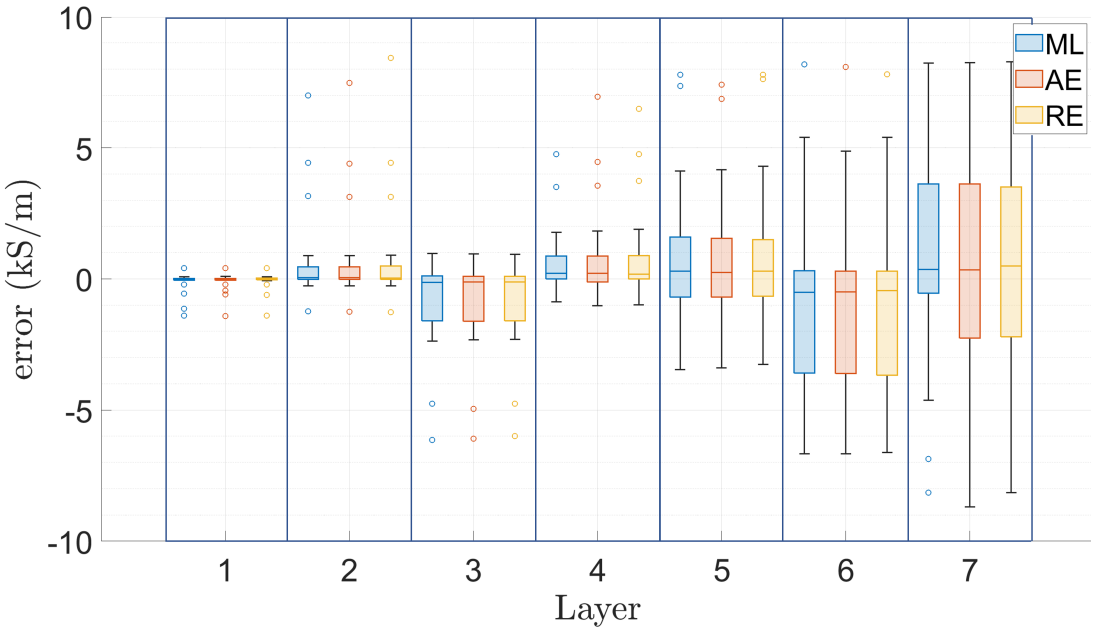

Figure 7.10: The reconstruction error using the ML prior and the two iterative algorithms: synthetic data without noise.

With Noise

The error is presented in fig. 7.11, again using box plots for each layer. Examples of reconstructed profiles are given in fig. 7.12 and 7.13; for readability, the depth profiles in these plots are interpolated from the trepanned sample and the discrete conductivity profile.

It can be see that there is no consistent reduction in the error in the iterative reconstructions relative to the ML prior. However, there is a trend in increased error in the relative error objective function compared with the absolute error. This may be explained by the increased weighting of the low frequency data in the relative function, where the low frequency data is more susceptible to noise. Fig. 7.13 gives an example where the step is not too restricted: for these particular profiles, the algorithm makes a good step towards the true solution.

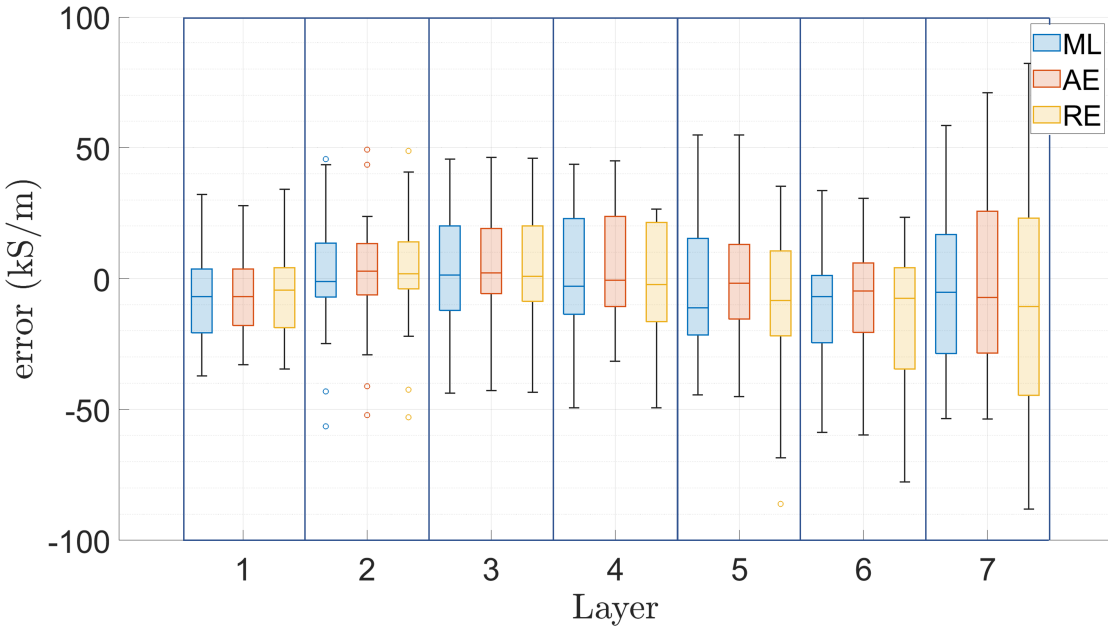

Figure 7.11: The reconstruction error using the ML prior and the two iterative algorithms: synthetic data with noise.

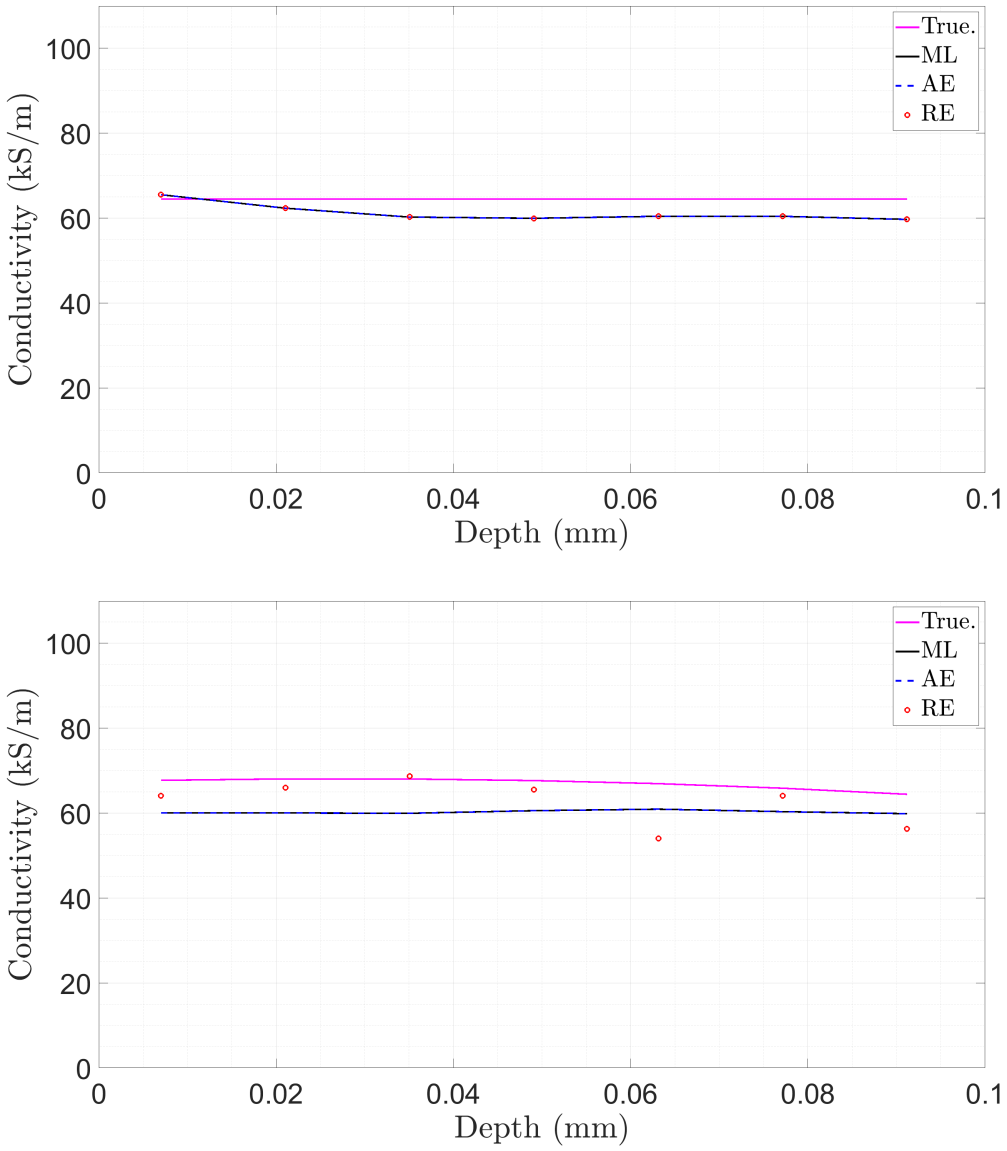

Figure 7.12: Examples of reconstructed profiles using synthetic measurement data with noise

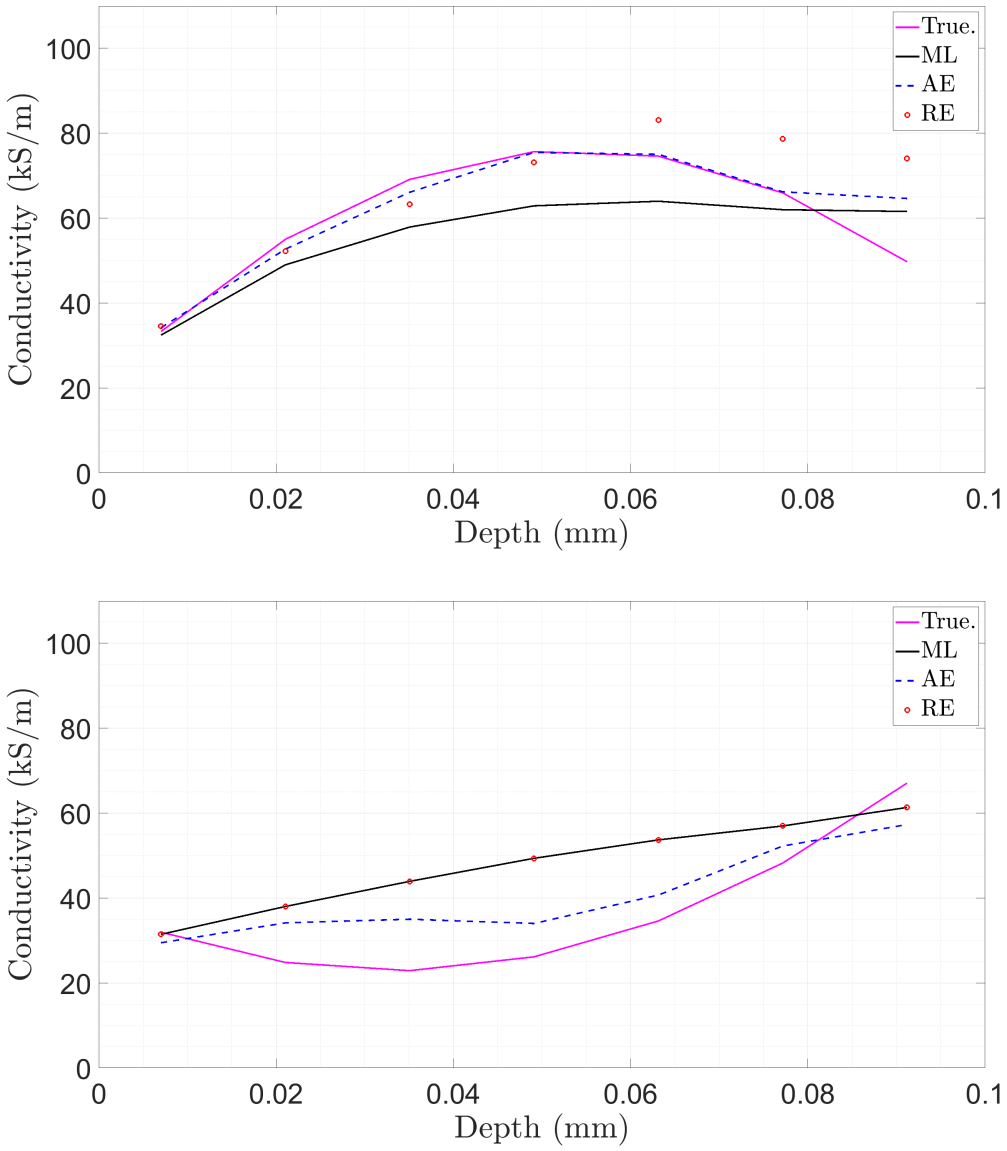

Figure 7.13: Examples of reconstructed profiles using synthetic measurement data with noise

7.5.2 Reactor Results and Discussion

The inverted reactor data is shown in fig. 7.14-7.15, where these plots correspond to different trepanning locations. It can be seen that the bore value is accurately predicted, but the profile diverges with depth; this trend has been previously shown [53][52]. Qualitatively, there is some agreement regarding the trend of the conductivity depth profile: the conductivity initially increases and plateaus.

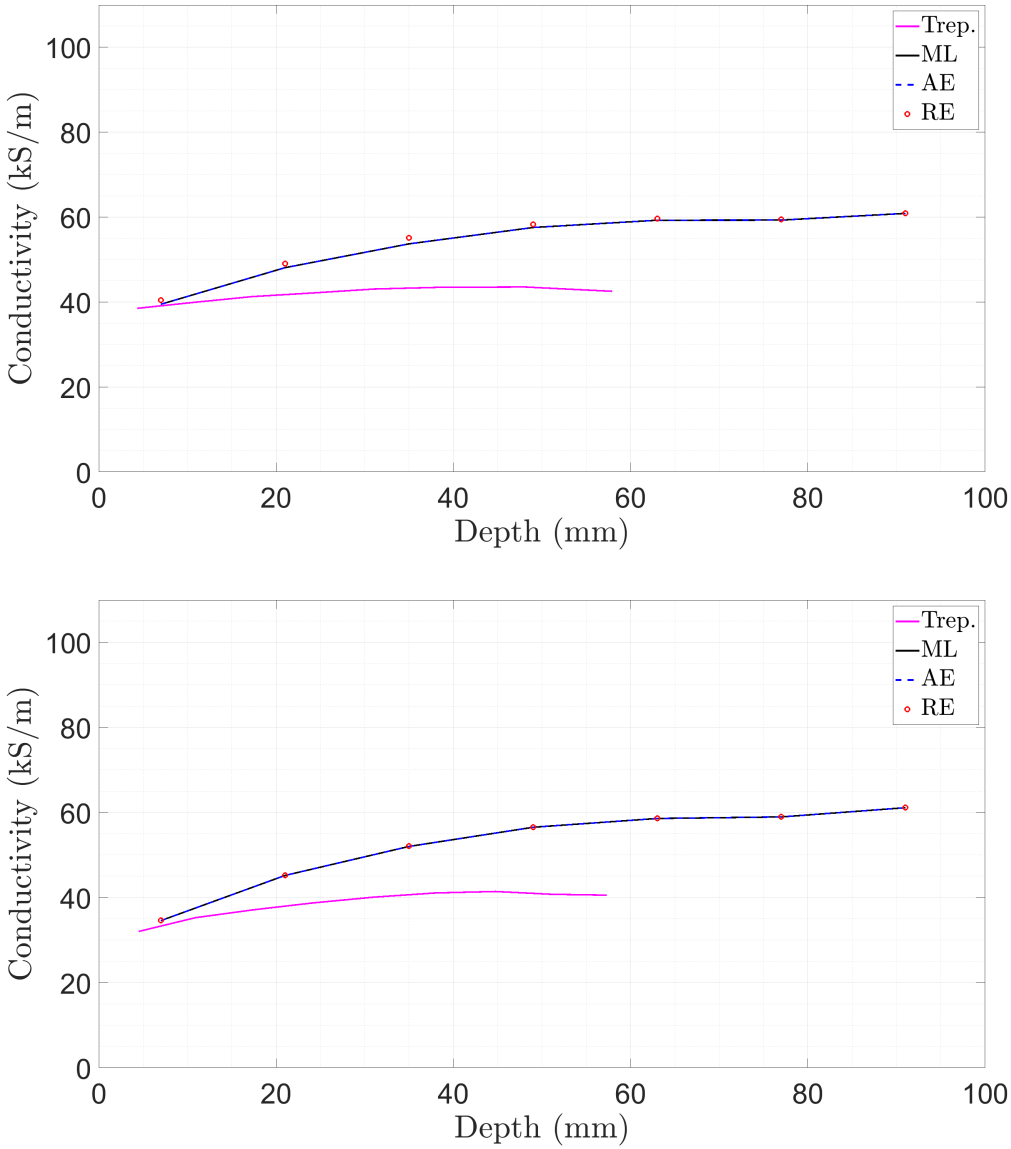

Figure 7.14: The trepanned depth profiles and reconstructed depth profiles from reactor data using a ML prior estimate.

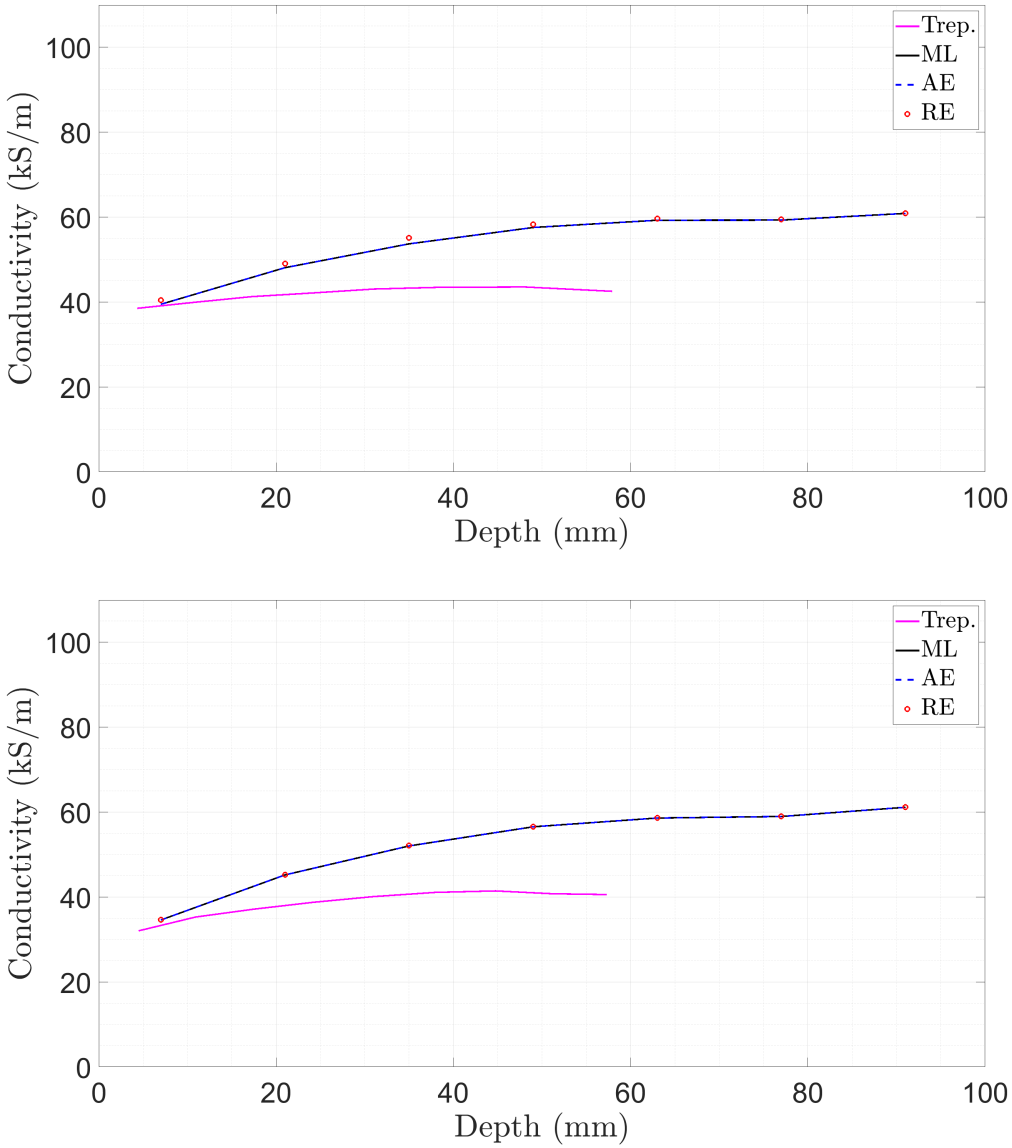

Figure 7.15: The trepanned depth profiles and reconstructed depth profiles from reactor data using a ML prior estimate.

Again, the conductivity depth profiles reconstructed from the iterative algorithms only marginally move from the ML prediction. This could be due to over-regularisation or the residual of the prior estimate being lower than the discrepancy: the regularisation parameter could be continually increased and it may not push the residual norm above the discrepancy. This is encapsulated by eq. 7.6, where $f_\lambda(\vec{\sigma})$ is the regularised objective function and $\vec{\sigma}^*$ is the

minimum of the objective function.

$$\lim_{\lambda\to\infty} \operatorname{argmin}(f_\lambda(\vec{\sigma})) = \begin{cases} \vec{\sigma}^* \to \vec{\sigma_0} \text{ and } \vec{r}(\vec{\sigma}^*) \to \vec{r}(\vec{\sigma_0}), & \text{if } \Re(\vec{\sigma}) = \frac{\lambda}{2}\|\vec{\sigma} - \vec{\sigma}_0\|_2^2 \\ \vec{\sigma}^* \to \vec{0} \text{ and } \vec{r}(\vec{\sigma}^*) \to \vec{r}(\vec{0}), & \text{if } \Re(\vec{\sigma}) = \frac{\lambda}{2}\|\vec{\sigma}\|_2^2 \end{cases} \tag{7.6}$$

According to algorithm 5, a step is rejected if the resultant residual norm is below a threshold: it is possible that a step is never accepted. This is a possible explanation for the lack of spread in the reconstructed profiles. This could be mitigated by lowering the discrepancy multiplier, using a prior with a residual above the noise floor or by allowing the conductivity profile to be updated when the residual is below the noise floor. The latter option could cause errors to propagate through the algorithm, resulting in early termination or divergence; therefore, this option is not favourable.

The possibility of modelling error means that the ML algorithm may not be the most appropriate prior (this was discussed in §4.7). It may be better to not predict the depth profile but instead the coefficient σ_c, σ_d and σ_s in eq. 7.7. This has the advantage of having fewer parameters to predict and resembles the expected behaviour of the graphite.

$$\sigma(r) = \begin{cases} \sigma_c - \sigma_d e^{-\sigma_s \cdot r}, & \text{if } r \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{7.7}$$

To study the restrictive movements of the algorithms, the inversions are repeated with a more informed prior and with different discrepancy multipliers. The relative error is not used because of the susceptibility to noise. For the benefit of time, instead of using an ML algorithm to predict the coefficients in eq. 7.7 from inductance data, eq. 7.7 was fitted to the trepanned depth profiles. These fitted curves were used as priors in the same iterative algorithm but with the identity as the regularisation matrix.

Figs. 7.16—7.19 illustrate the sensitivity of the algorithm to the discrepancy multiplier. From this, it can be seen that the initial damping and regularisation parameters are not too large because the iterative algorithm moves away from the prior estimate when the discrepancy multiplier is lowered. Further, this demonstrates that over-regularisation is not preventing movement, but that a step is rejected because the residual is below the noise floor. A further study should be conducted to obtain a suitable value of c, in the same manner as the initial damping and regularisation parameters; from these profiles, it seems a value between 0.5 and 0.6 would be ideal. It should be noted, that although the residual norm decreases with the lower discrepancy multiplier (as shown in fig. 7.20) the reconstructed profile diverges from the true solution. This demonstrates the issue of lowering the discrepancy multiplier and the increased susceptibility to error.

Instead of determining some optimal values for c, the error in the reconstructed profiles can be reduced by reducing modelling error. This could be achieved by using the damped probe or with a further calibration stage using a transfer function method.

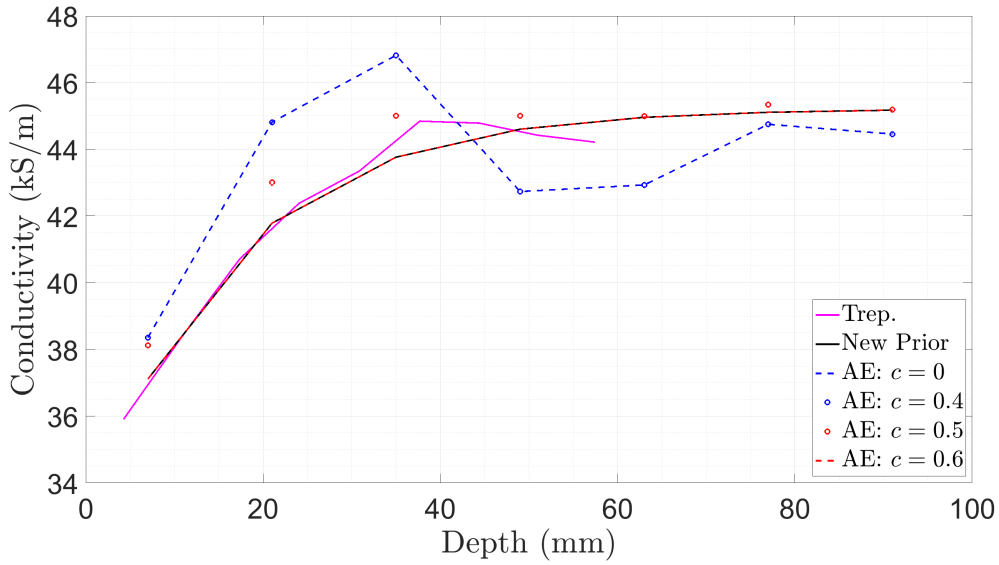

Figure 7.16: The trepanned depth profile and reconstructed depth profiles with a more informed prior and various discrepancy multipliers.

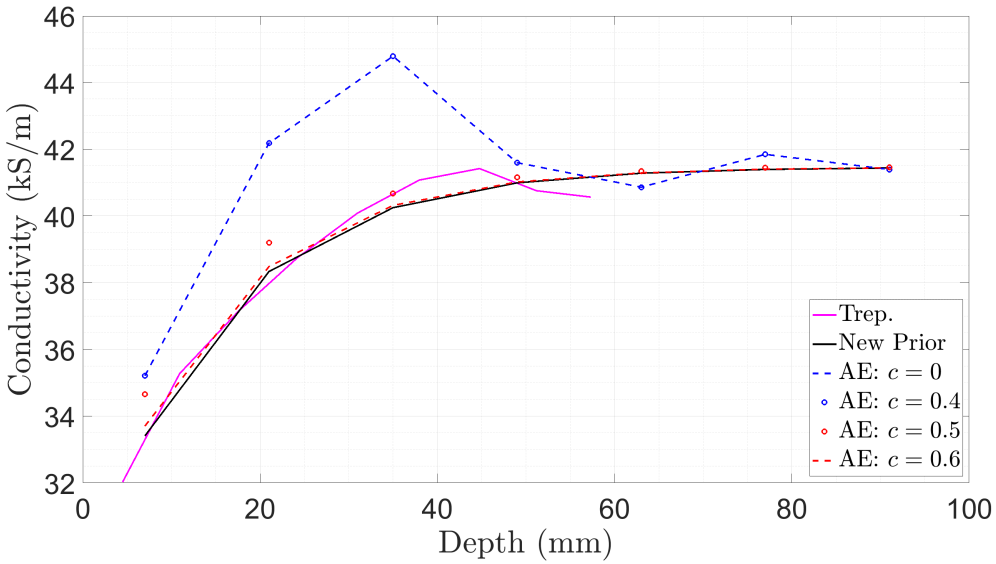

Figure 7.17: The trepanned depth profile and reconstructed depth profiles with a more informed prior and various discrepancy multipliers.

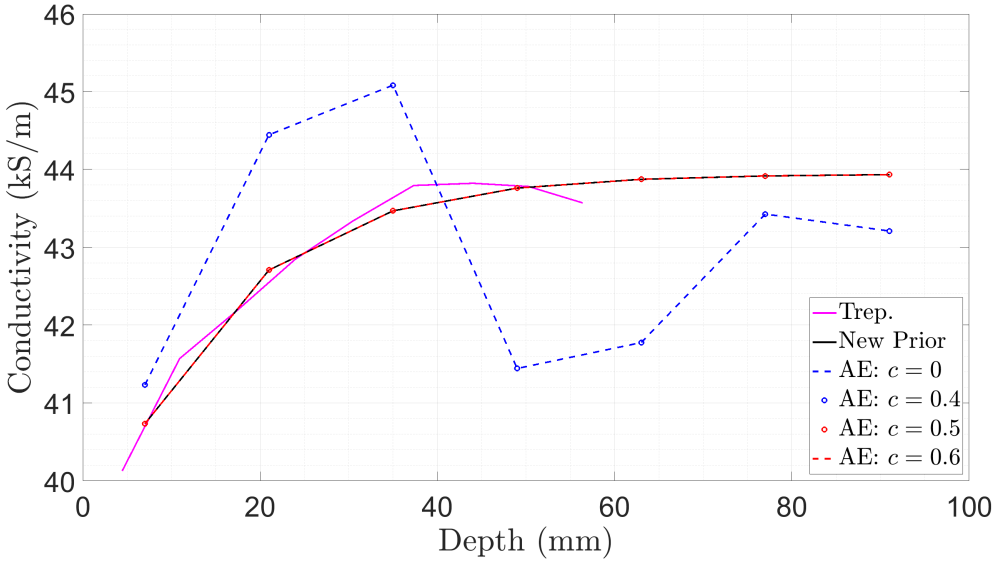

Figure 7.18: The trepanned depth profile and reconstructed depth profiles with a more informed prior and various discrepancy multipliers.

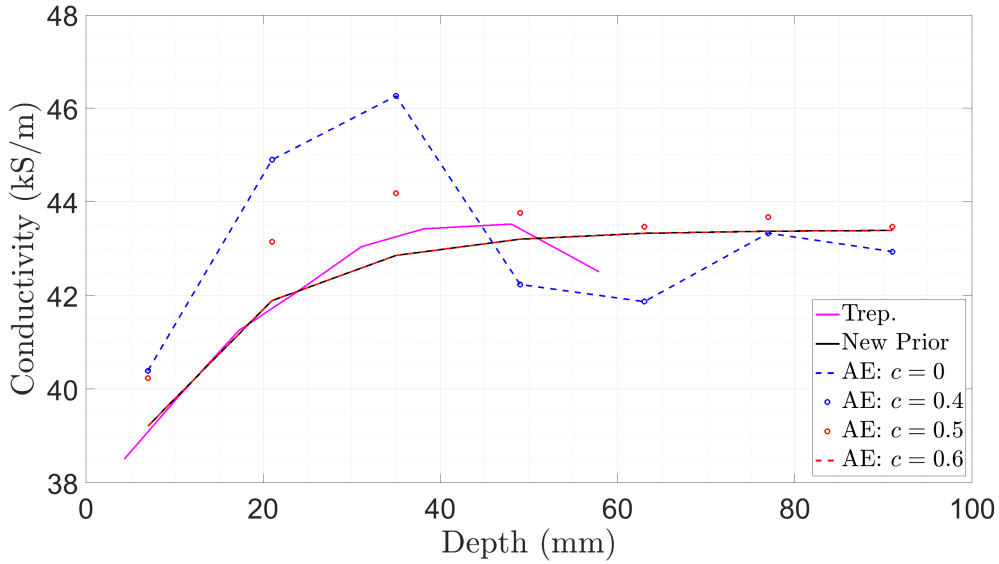

Figure 7.19: The trepanned depth profile and reconstructed depth profiles with a more informed prior and various discrepancy multipliers.

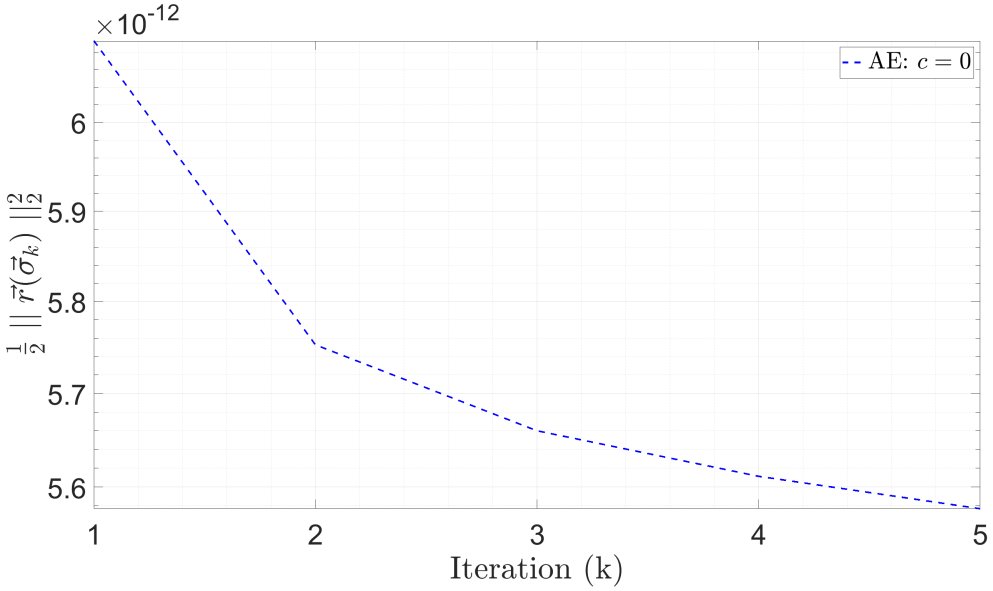

Figure 7.20: The residual norm of the objective function with no regularisation parameter update. This corresponds to the conductivity depth profile in fig. 7.16.

7.6 Chapter Summary

- The inversion algorithm consisted of a prior estimate from a ML algorithm and a step derived from the LM direction. Two errors were compared: the absolute and relative error.

- The inversion algorithms were tested on synthetic data, with and without noise, and reactor data. The study on data without noise defined the base performance of the algorithms; the ill-conditioning of the problem was illustrated using the error of the reconstructed profiles. The study on data with noise demonstrated the increased susceptibility of the relative error to noise compared with the absolute error. In both of these cases, the iterative algorithms only marginally deviate from the prior estimate.

- The inversion of reactor data was similar to the study on synthetic data with noise. Specifically, the step was too restrictive. This was not because of over-regularisation but because a step was rejected if the residual norm was below the discrepancy norm. This could be rectified by decreasing the discrepancy multiplier or by allowing updates to the conductivity profile even if below the noise floor; however, this would introduce error into the reconstructed profile. Instead, it would be favourable to lower modelling error.

Chapter 8

Crack Detection

8.1 Overview of Chapter

Crack detection is complicated by the fact that a crack can be viewed as a loss in electrical conductivity, and therefore, there are significant overlaps in features in the inductance measurement. Averaging has previously been used as an attempt to increase the sensitivity to a crack.

The objective of this chapter is to investigate data driven techniques for feature extraction and classification, without using averaging techniques. It is shown that for measurements with a SNR in the range of approximately 40-55 dB the balanced classification accuracy is no larger than 71.45%.

A notch was used to characterise a crack; this is generally not representative of crack propagation in a nuclear graphite channel, but it does simplify the methodology and results. A limitation given by Fletcher et al., was that stresses in a nuclear reactor brick force the crack faces together, significantly lowering sensitivity; this was not overcome in this study. A more accurate investigation would include the crack width and boundary conditions, perhaps modelling the crack boundary by varying porosity.

8.2 Averaging Techniques

The investigation of crack detection by Fletcher et al. was accomplished by varying the depth of a machined notch, emulating the keyway root cracks [135]. In the experimental study, the measurement made was $m - m_{av}$, where m_{av} was the average of the responses over the spatial extent of the brick. The phase was most sensitive to the presence of a notch change, where notches were detectable at a distance greater than 10 mm from the keyway. Fletcher et al. gave two limitations to this study of crack detection:

1. The width of the crack was assumed to have negligible affect on the phase compared to the radial width. In the core, surrounding bricks may force the crack open or close it. A crack which is closed may have sufficient electrical contact that the eddy current has little disturbance, producing a very small phase response.

2. The study used virgin graphite blocks that have not been irradiated. Therefore, the graphite blocks were homogeneous and the electrical conductivity profile could be modelled using a single probability distribution.

Tesflalem investigated more realistic cracks than the machined slot model. This was achieved by using a wedge fitted in a keyway to provide a tensile force to push the corners apart. This process would initiate a crack originating at the keyway corners extending preferentially in the radial direction towards the bore [53]. When inspecting, the wedge was left in place leaving the crack open. Tesfalem compared three different averaging methods for detecting cracks:

1. Use a spatial average of the measurement using the same block the defect lies in (m_{av}^{def}) and subtract this average from the measurement at

a given position (m^{def})

$$m^{def} - m_{av}^{def}$$

2. Use an additional defect free reference brick and determine the spatial average for both the reference (m_{av}^{ref}) and material under test (m_{av}^{def}). Next, difference each measurement and the spatial average, $\Delta m^{ref} = m^{ref} - m_{av}^{ref}$ and $\Delta m^{def} = m^{def} - m_{av}^{def}$. Follow this with subtracting the change in the reference brick data from the change in the defect brick data

$$\Delta m^{def} - \Delta m^{ref}$$

3. Use an additional defect free reference brick, calculate a spatial average just on the reference brick, subtract this average from the measurement of the defected brick at a given position

$$m^{def} - m_{av}^{ref}$$

Method two was found to provide the best circumferential positioning of subsurface cracks, possibly due to the reduction in both measurement variation owed to noise and non-uniform conductivity distributions [53]. Tesfalem found that a machined slot with an extent of 15mm from the keyway could be located, whereas the realistic crack was only detectable at a distance of 32mm from the keyway.

The use of a spatial average in these methods is important for isolating changes in the data due to the crack. However, the effectiveness of this depends on the similarity between the spatial average of conductivity and the underlying conductivity profile at the defected location (i.e. azimuthal conductivity variations are more or less constant). This can be overcome by using a local average instead, but this would require a study into the azimuthal conductivity variations.

8.3 Machine Learning Techniques

In determining the existence of a crack, the unknown electrical conductivity distribution can be seen as noise in the measurement. Ideally, the inverse function would output a binary 1 or 0 to indicate the presence of a crack, without the use of averaging

$$f(m) = \begin{cases} 1 & \text{Crack present} \\ 0 & \text{No Crack} \end{cases}$$

This function is not known but can be approximated by ML techniques.

8.3.1 Feature Extraction

Principal Component Analysis

PCA is one of the simplest techniques for extracting features from measurement data. PCA compresses the data by finding the projection of samples along some directions, called the principle components; these directions are determined such that the variables in the new feature space are uncorrelated and have maximal variance [136]. The principal components can be found by evaluating the data's mean-centred co-variance matrix, given in eq. 8.1 [137]. Alternatively, the correlation matrix can be used, by dividing through by a variable's standard deviation; the benefit of which is to scale the variables such that they are of a similar magnitude and the choice of principal direction is not then dominated by any one variable [138]. The standardised observation is then placed in the row space of $\boldsymbol{X} \in R^{n \times m}$, where n is the number of samples and m is the number of variables.

$$\boldsymbol{C} = \frac{1}{n} \boldsymbol{X}^T \boldsymbol{X}. \tag{8.1}$$

The principal components are the eigenvectors of the matrix $\boldsymbol{C}$ in eq. 8.1. These eigenvectors can then be normalised such that $\langle \vec{m}, \vec{v} \rangle = \|\vec{m}\| \, cos(\theta)$, where $\vec{v}$ is an eigenvector, $\vec{m}$ is a sample of measurement data and θ is the angle between these two vectors.

Kernel Principal Component Analysis

Kernel Principal Component Analysis (KPCA) is conceptually the same as PCA, but with the measured data transformed prior to compression. Moreover, the projection of the transformed data with its principle components can be computed without computing the transform. The transformation is usually a higher dimensional view of the measured data. This is useful when the data cannot be linearly separated. PCA is then applied on the transformed data, $\phi(\vec{m})$, such that the co-variance matrix can be written as eq. 8.2.

$$\boldsymbol{C}' = \frac{1}{n} \sum_{i=1}^{n} \phi(\vec{m}) \phi(\vec{m})^T \tag{8.2}$$

The eigenvectors can be found by solving $\boldsymbol{C}'\vec{v} = \lambda\vec{v}$ [139]. The projection of the data on the principal axes is then given by eq. 8.3, where the vector $\vec{v}_j$ can be described in terms of $\phi(\vec{m})$ as $\vec{v}_j = \sum_{l=1}^{n} \alpha_l \phi(\vec{m}_l)$ [139].

$$\langle \phi(\vec{m}_i), \vec{v}_j \rangle = \langle \phi(\vec{m}_i), \frac{1}{\lambda} \boldsymbol{C}' \vec{v}_j \rangle \tag{8.3}$$

The trick in KPCA is to use the kernel defined by eq. 8.4. There are additional steps in centring the data, details of which can be found in [137], [139] and [140].

$$k_{ij}(\vec{m}_i, \vec{m}_j) = \langle \phi(\vec{m}_i), \phi(\vec{m}_j) \rangle \tag{8.4}$$

The polynomial, gaussian and sigmoid kernels are common transformations; the polynomial, Gaussian and sigmoid kernels are given in eq. 8.5, 8.6 and

8.7 respectively.

$$k(\vec{m}_i, \vec{m}_j) = (\gamma\langle\vec{m}_i, \vec{m}_j\rangle + c_0)^{\beta_0} \tag{8.5}$$

$$k(\vec{m}_i, \vec{m}_j) = e^{-\gamma\left\|m_i - \vec{m}_j\right\|^2} \tag{8.6}$$

$$k(\vec{m}_i, \vec{m}_j) = tanh(\gamma\langle\vec{m}_i, \vec{m}_j\rangle + c_0) \tag{8.7}$$

Independent Component Analysis

ICA is conceptually very similar to PCA but uses different principle components. In ICA, the measured data is modelled as some mixing of independent variables. The components are determined by maximising the statistical independence of the variables [141]. Once the independent components are determined the projection of the data can be found.

Auto-Encoder

An Auto-Encoder (AE) is a type of neural network. In this type of network there are two distinct components: an encoder and a decoder, where the input is also used as the output target. The advantage of such a network is when the encoder output dimension is less than the number of input variables, this is the undercomplete AE [142]. The aim of the network is to extract the most relevant features of the input data such that the input can be accurately reconstructed.

8.3.2 Classification

K-Nearest Neighbours

In k-nearest neighbours (KNN) the test data is compared with the k-nearest training samples. A classification is made based on the classification of the neighbours. The classification of the neighbours is averaged and a decision made based on some threshold.

Logistic Regression

In Logistic Regression (LR), a continuous function is desired in which a probability that a feature of interest is present can be associated with a measurement. If the measurement is represented by the random variable vector $\vec{m}$, then the probability desired is $P(c = 1|\vec{m})$, where c is a binary variable indicating the presence of a particular feature. In this work $c = 1$ is given to bricks that are defective. The probability function is not known and it is the aim of LR to derive it, details can be found in [143].

Support Vector Machine

A Support Vector Machine (SVM) splits data by explicitly defining a hyperplane in the feature space, i.e. the decision boundary. The hyperplane normal is described by the vector $\vec{\omega}$ and constant b, where these unknowns are to be determined. This hyperplane is found such that the maximum margin is found between the decision boundary and the closest data point from each class, where the margin is the perpendicular distance from a data point to the hyperplane given by eq. 8.8. For the hyperplane defined using SVM, the distance between the closest data point in either class and the hyperplane

is to be maximised.

$$\langle \vec{\omega}, \vec{m_i} \rangle + b = \delta \tag{8.8}$$

If the data is linearly separable then the hard-margin case can be used, where there is a constraint condition that no samples in either class can be on the wrong side of the decision boundary. For cases which are not linearly separable the soft margin is used, where a penalty term is included in the objective function for data-points on the wrong side of the margin. See [144] and [145] for further details.

Similar to KPCA, kernel methods can be used for using a SVM in alternative representations without directly computing the transformation.

8.4 Geometry and Measurement

The problem geometry considered is shown in fig. 8.1. All modelled notches had a fixed width of 5 mm, whilst the notch height was varied. The graphite brick had a width and length of 240 mm and height of 100 mm. A cuboid geometry considerably simplified the methodology for changing the conductivity depth profile in the experimental study.

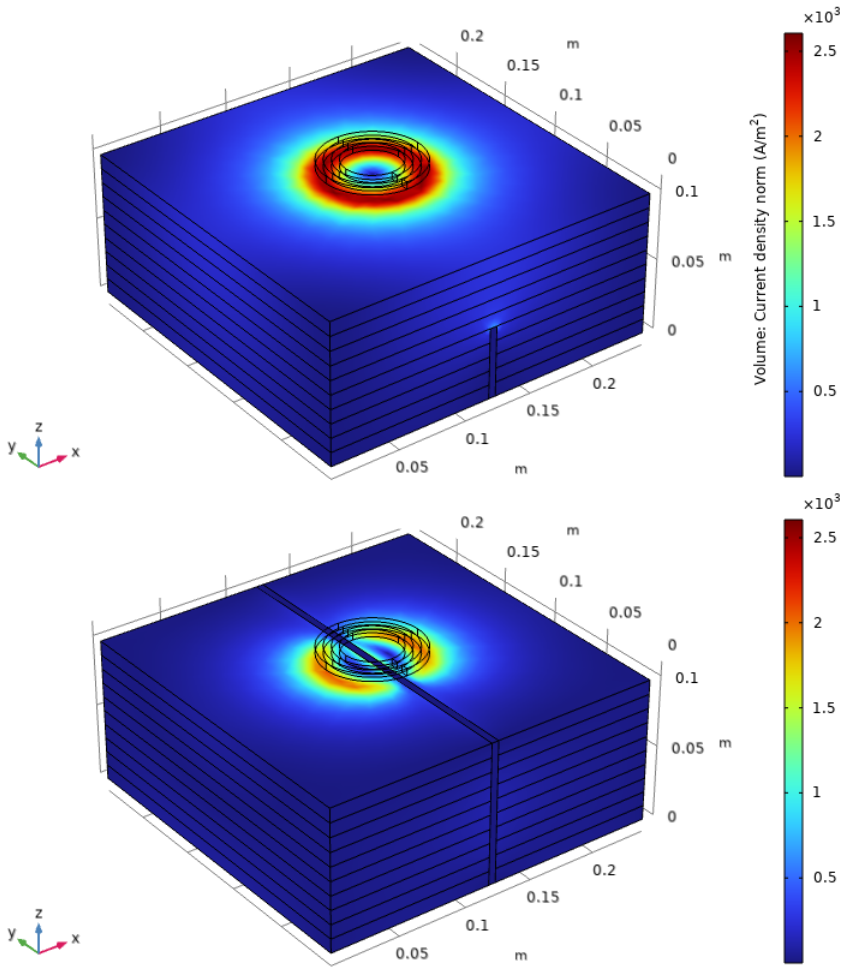

Figure 8.1: The problem investigated: conductive layers stacked together to obtain a conductivity depth profile, $\vec{\sigma}$. The notch height was defined with respect to the base of the brick.

The sensor consisted of a receive coil embedded within a transmit coil, where the measurement was made according to eq. 3.16. The FEM was calibrated as in §6.4.1.

8.5 Generating the Synthetic Data

The conductivity depth profile was generated as described in §4.4.1. The height of a notch was selected from the tuple $(5, 10, 15, \ldots, 95, 100)$ mm, using 148 samples of each. The total number of datasets generated was 5920 samples, with 2960 having a notch and 2960 not having a notch.

The measurement noise was determined from the calibrated data and the noise profile was modelled by the methodology in §7.3.1.

8.6 Generating the Experimental data

8.6.1 Conductivity Depth Profile

The graphite grades which were available are given in table 8.1, where there are two of each density in any-one-stacking sequence. There are a large number of permutations for the possible conductivity profiles; there are 3.6288×10^6 permutations. The number of unique permutations is 1.134×10^5.

Table 8.1: Grades of graphite available - each graphite layer was 10 mm thick.

Graphite Grade	Electrical Conductivity (kS/m)
Z645	9.82E+03
No.26	1.66E+04
IVC032	2.41E+04
R7660	3.79E+04
NC	9.82E+04

In practice, a smooth trend would be expected in the conductivity profile, such as linear or quadratic. To Find profiles of a smooth trend, the number of maxima, n_{max}, and number of minima, n_{min}, in a given permutation was counted and the order computed by $1 + n_{min} + n_{max}$ (there were not enough plates for order 0). Profiles which had an approximate order of 1, 2, and 3 were then selected. This initial step provided a fast way to filter out profiles. However, some filtered profiles are better than others; for example, some profiles may have a greater degree of linearity than others. Therefore, out of these selected profiles, those which have the smallest error with a fitted polynomial curve of the same order were selected. Fig. 8.2 shows an, approximately, linear experimental stacking sequence and its fitted curve. Two profiles of each order were used; therefore, there were 6 different profiles to choose from

for a given notch height.

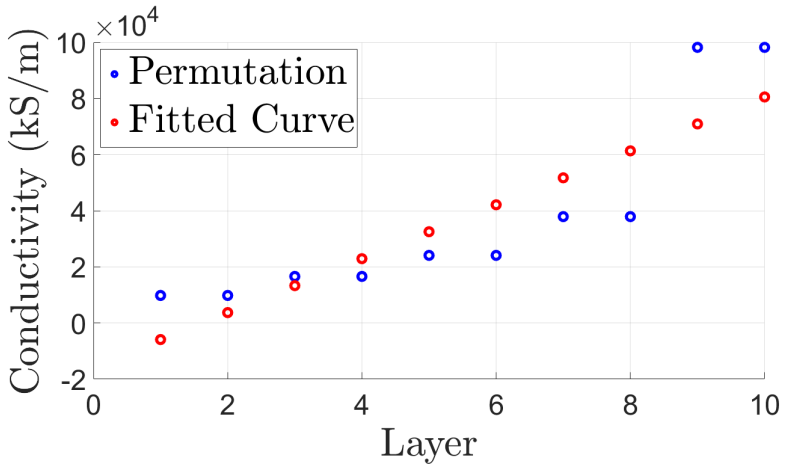

Figure 8.2: An experimental profile with an order 1 and its equivalent fitted profile. Profiles with the smallest error between them and the fitted curves were selected.

8.6.2 Notch Profile

For each conductivity profile, a notch height was selected from the tuple $(0, 10, 20, \ldots, 90, 100)$; in total this amounted to 66 experimental samples. An example of a stacking sequence with a notch of 90mm is shown in fig 8.3.

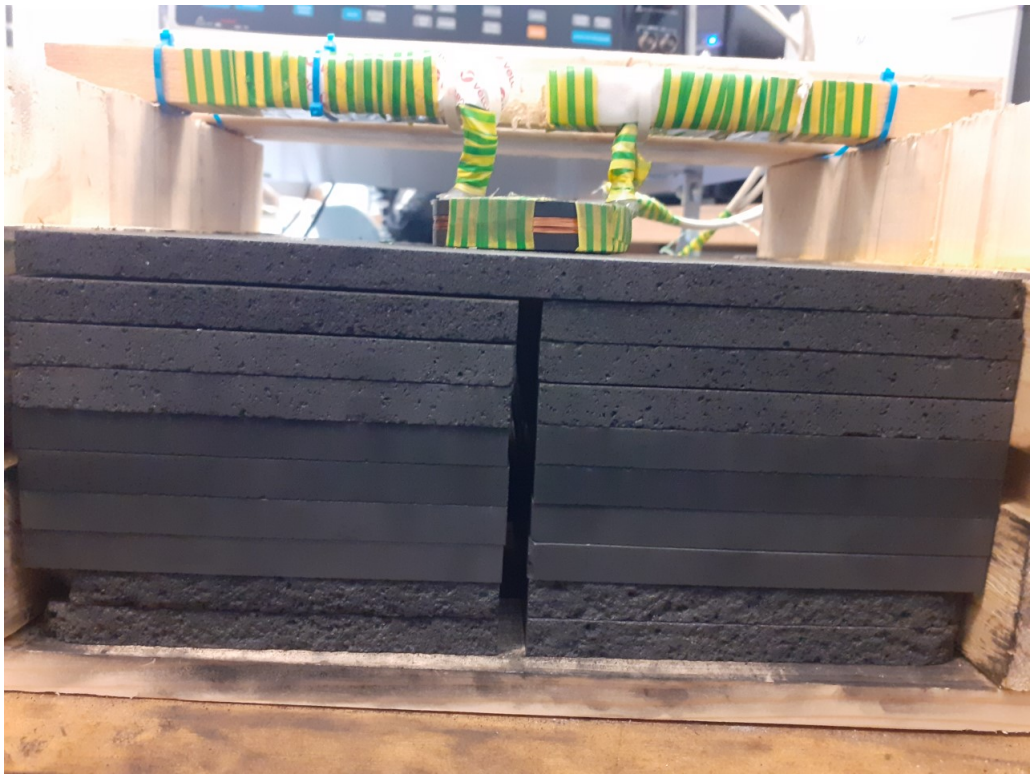

Figure 8.3: An example of a stacking sequence.

8.6.3 The Measurement Function

To see how the measurement behaves as a function of frequency and notch height, the average relative change in the measurement was found; this is defined by eq. 8.9, where n is the number of different depth profiles. This is shown in fig. 8.4 using the experimentally measured data, where $n = 6$.

$$\delta(f,h) = \frac{1}{n}\sum_{k=1}^{n} \frac{\mid m(f,h) - m(f,h=0) \mid}{\mid m(f,h=0) \mid} \times 100 \tag{8.9}$$

This shows the difficulty of the problem, as the sensitivity to the notch height drops orders of magnitude with decreasing height.

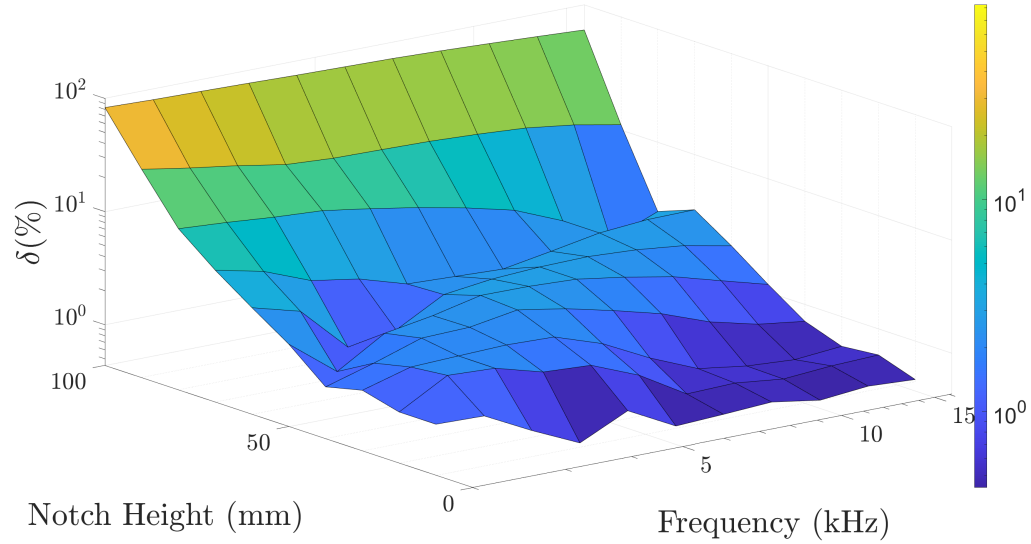

Figure 8.4: The change in measurement with respect to a defect free brick as a function of notch height.

8.7 Algorithm Selection

The feature extraction and classification algorithms used are given in tables 8.2 and 8.3, where each classification method is tested on each feature extraction method; therefore, 49 different algorithms were tested. Both linear and non-linear methods are included for feature extraction. The polynomial, gaussian and sigmoid kernels are used in both the KPCA and KSVM type algorithms, denoted by P-KPCA, G-KPCA, S-KPCA and P-KSVM, G-KSVM, S-KSVM, respectively. The linear (untransformed) PCA and SVM algorithms are denoted by LPCA and LSVM, respectively.

Table 8.2: Feature extraction algorithms used.

Feature Extraction
None
ICA
LPCA
P-KPCA
S-KPCA
G-KPCA
AE

Table 8.3: Classification algorithms used.

Classifier
KNN
LR
LSVM
P-KSVM
G-KSVM
S-KSVM
FCN

8.7.1 Parameter Selection

The Hyperband algorithm [87] was used to find the hyper-parameters in the neural networks, and the halving grid search for the other feature extraction and classifier algorithms. These algorithms are similar in that they iteratively restrict the search space, rather than sample the entire search space such as in a grid search. Both of these algorithms are available through the open source python packages, Keras and scikit-learn [146].

There were several methods used to speed up the parameter selection process. The first was in recognising that the AE hyper-parameters can be determined independently of the classifier, meaning the AE network architecture only has to be found once for each encoder output dimension. This same trick cannot be used for the FCN classifier because the architecture will depend on the input from the feature extraction algorithm; therefore, a Hyperband search would need to be run in every single trial in the halving grid search search. Thus, the second method for speed up was to initially use a low number of epochs (number of runs through entire dataset) etc. within the Hyperband algorithm; when called upon in the halving grid search the initial Hyperband algorithm is fast and a good set of hyper-parameters quickly determined outside of the neural network. Once this initial search was completed, the hyper-parameters outside of the neural networks was fixed and those within the neural networks found using a finer search by increasing the number of epochs etc. in the HyperBand algorithm.

8.7.2 Performance Metric

When selecting the performance metric, it is useful to know the distribution of samples according to the presence of a notch. The distribution of samples is shown in fig. 8.5 for the experimental and synthetic training data. The exact

number of samples in each class within the validation and training set is not known, since the data is shuffled before splitting for validation. Further, the number of each class in the experimental set is imbalanced. This imbalance requires the use of a metric other than overall accuracy, such as the F-score or balanced accuracy.

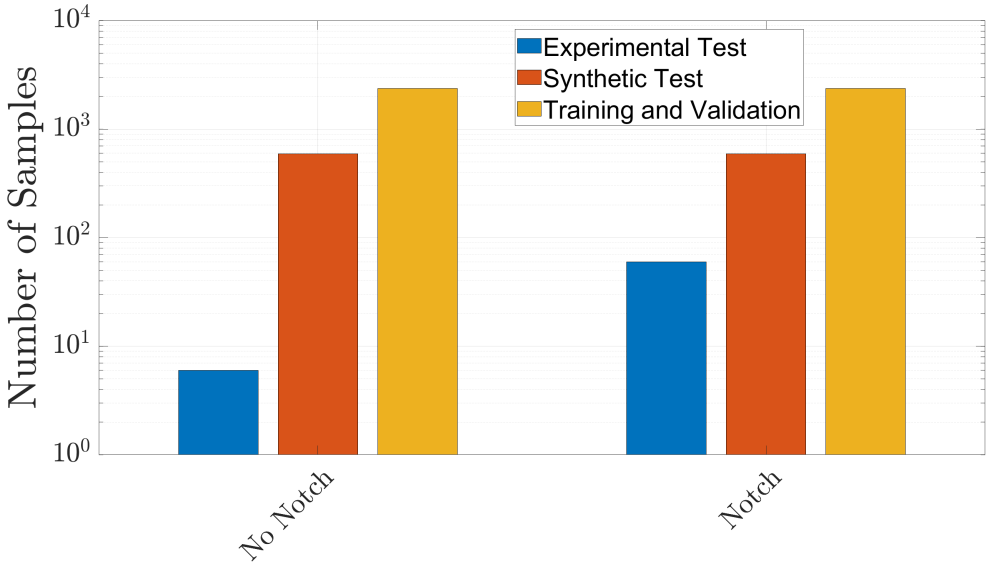

Figure 8.5: The distribution of defected and undefected samples in the dataset.

The balanced accuracy was used to train and test the different ML algorithms, i.e. the average accuracy of each class. This is given by eq. 8.10, where T_0 and T_1 correspond to the number of correctly labelled samples from each class and F_0 and F_1 the number of incorrectly labelled samples from each class.

$$A_b = \frac{1}{2}\left(\frac{T_0}{T_0 + F_1} + \frac{T_1}{T_1 + F_0}\right) \tag{8.10}$$

8.8 Application of Algorithms, Results and Discussion

8.8.1 Synthetic Data with Noise

The results on the synthetic test and experimental data are presented in two parts. First, the accuracy is shown according to the binary classification; i.e. if there is a notch present or not. Following this, the breakdown of the accuracies within the defected bricks is shown, such that the performance of the ML algorithms with varying height can be reported.

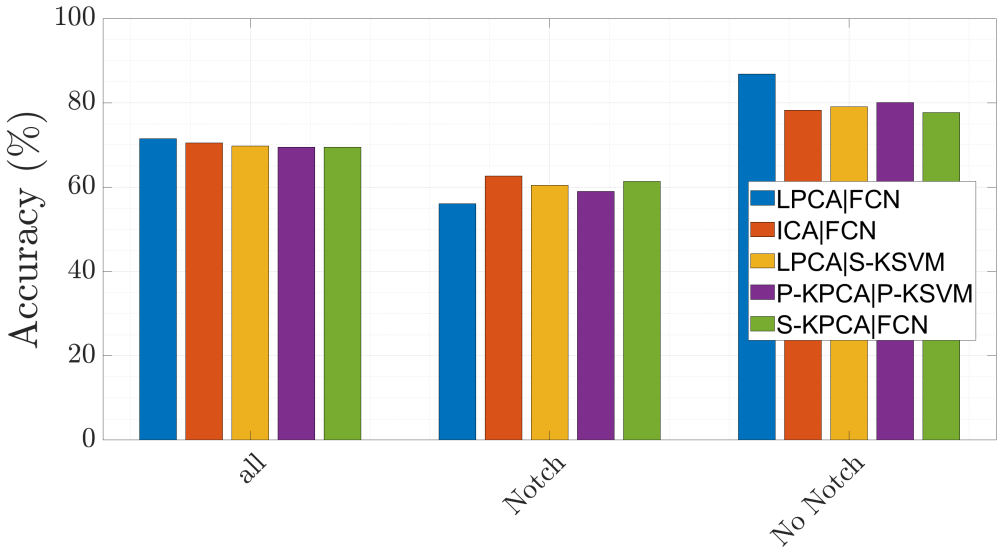

Figure 8.6: Results on synthetic test data. The accuracy of 'all' is the balanced accuracy of the binary classes.

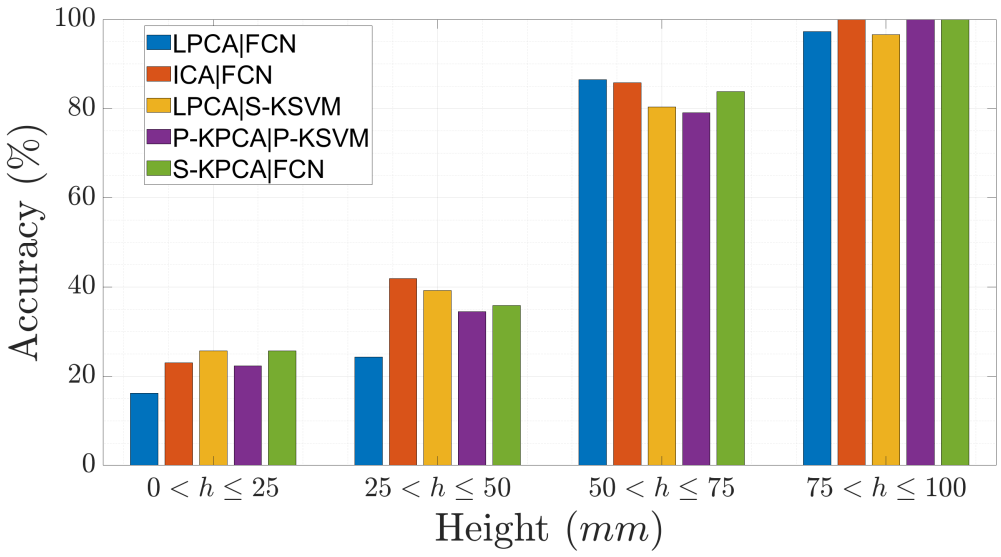

Figure 8.7: Breakdown of synthetic test results in terms of the accuracy of the samples belonging to each range given.

The results on the top five algorithms, corresponding to the synthetic test data, are shown in fig. 8.6 and fig. 8.7. The best balanced accuracy of all the synthetic test data was 71.45%. It can be seen that the PCA feature extraction algorithm correlates with a top performing algorithm and similarly the SVM and FCN algorithms for classification. However, there is no obvious trend in the different kernels. The breakdown of the notch accuracies, in terms of ranges of notch height, reflects the decay in the eddy currents with increasing depth, and therefore, decay in sensitivity.

8.8.2 Experimental Data

The results on the top five algorithms, corresponding to the experimental test data, are shown in fig. 8.8 and fig. 8.9. The first notable difference is that the best balanced accuracy on the experimental data is larger than the synthetic: 80.83%. This possibly indicates that the simulated noise was too large and could, perhaps, be lower. Further, the breakdown shows a generally more even distribution of accuracies and an accuracy of $\approx 58\%$ of those with a notch in the range of 0 to 25mm. However, there were only 12 experimental samples

within this region, and therefore, this is not very informative in general.

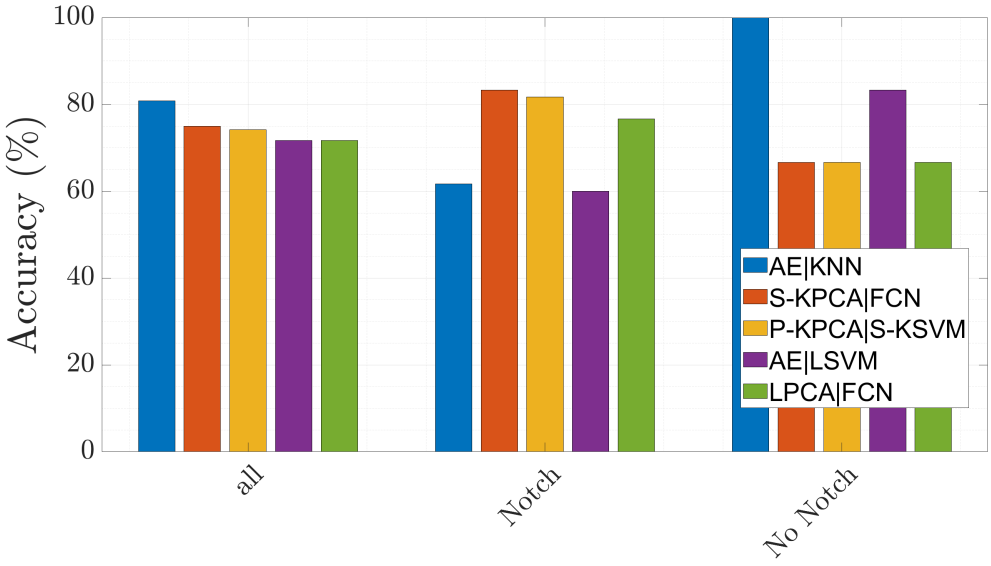

Figure 8.8: Results on experimental data. The accuracy of 'all' is the balanced accuracy of the binary classes.

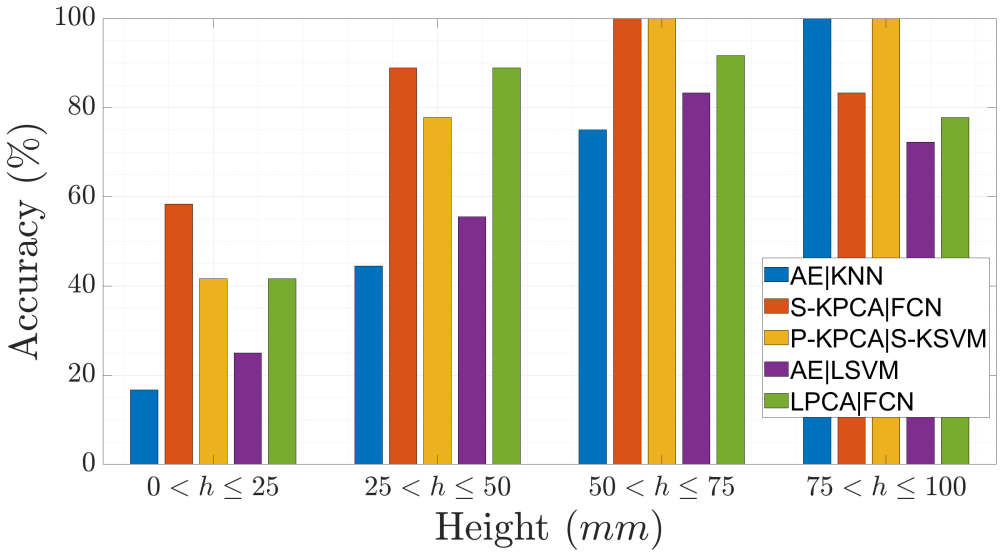

Figure 8.9: Breakdown of experimental results in terms of the accuracy of the samples belonging to each range given.

8.8.3 Synthetic Data without Noise

Finally, in order to see the limiting factors, the results of synthetic test data with no noise in the measurement response are presented; these algorithms have been re-trained and tested on the same set of data but with no noise.

This affords the ability to see the limiting factor in using ML techniques for notch detection and observe whether the limiting factor is the SNR of the measurement response or the underlying conductivity variation. For this, a significantly higher accuracy was achieved, approximately 99.5% using the FCN classifier and ICA feature extraction algorithm. Moreover, the ICA algorithm is the only feature extraction algorithm in the top five. Again, the FCN and SVM classification algorithms are in the top five for this case. However, the top performing algorithms are different from the case using samples with realistic SNRs; this is possibly because some feature extraction algorithms are better at noise suppression than others. These results also show that ML algorithms can discriminate between notches and conductivity variations with a high accuracy. However, when combined with realistic noise in the measurement this discrimination is much poorer, probably because delicate features are overwhelmed.

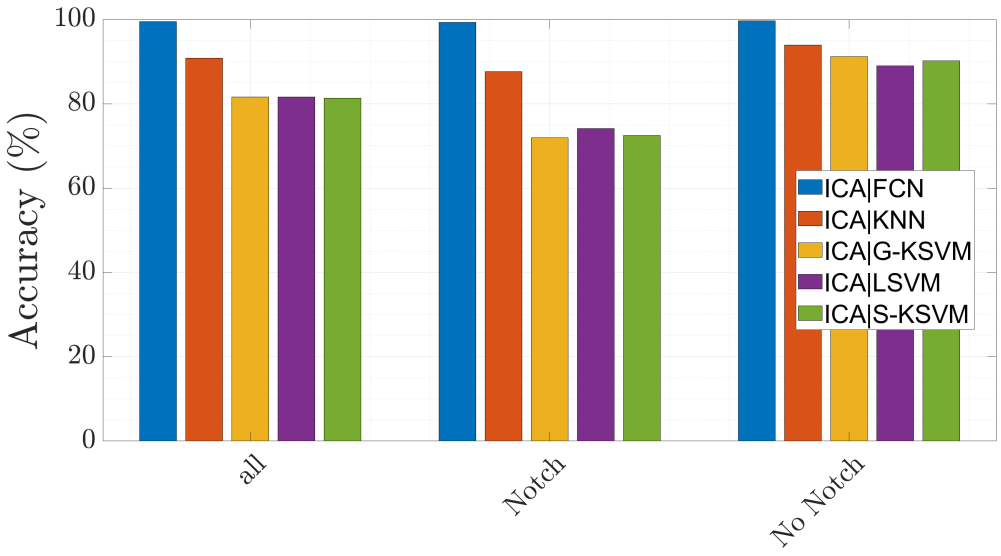

Figure 8.10: Results on synthetic test data with no noise. The accuracy of 'all' is the balanced accuracy.

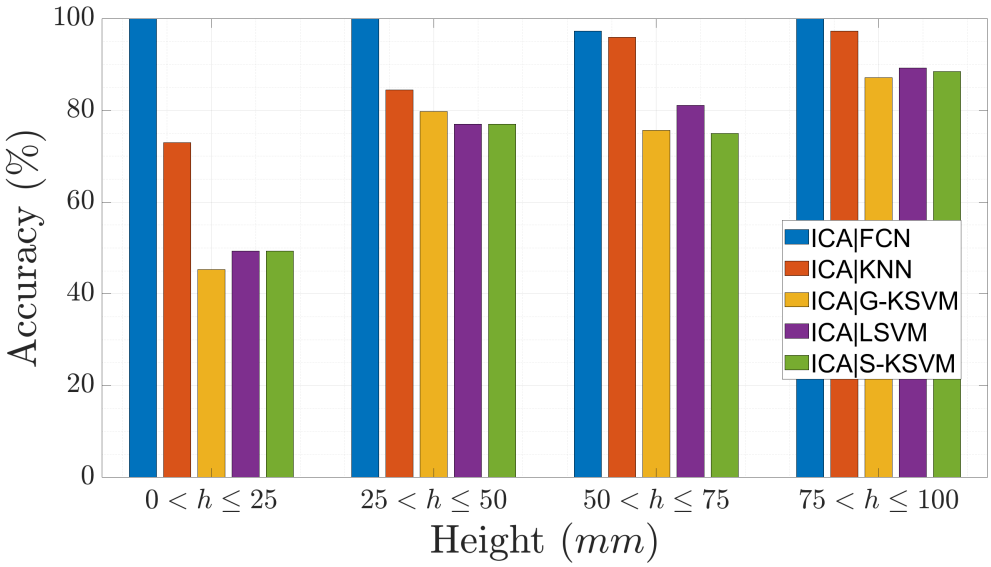

Figure 8.11: Breakdown of synthetic test results with no noise, in terms of the accuracy of the samples belonging to each range given.

8.9 Chapter Summary

- Past work in crack detection used spatial averaging to separate out the response due to conductivity variations and the presence of a crack. The study in this chapter assumed no prior knowledge in the measurement: no averaging was used. The aim of the study was to investigate whether ML algorithms could extract distinguishing features from the data, which could be attributed to the presence of a notch.

- 49 different ML algorithms were investigated for binary classification of a notch within a conductive block of material, in which there were variations in electrical conductivity. It was shown that for a SNR in the range of approximately 40-55 dB that the balanced classification accuracy was no larger than 71.45%; this accuracy was not limited by the underlying conductivity variations since a significantly higher accuracy can be achieved when there is no noise present.

- Alternative measurements that use averaging may result in improved ac-

curacies for a similar SNR level. Further, the accuracy may be sensitive to the frequency range used and coil geometry.

Chapter 9

Conclusions and Future Work

9.1 Contribution

This thesis has investigated and developed existing eddy current inspection methodologies for depth profiling and crack detection of nuclear reactor bricks. The areas studied and developed were: the choice of step direction (including first and second order methods), direct inversion using ML, calibration using optimisation, an iterative LM algorithm using the discrepancy principle and ML feature extraction and classification methods for crack detection.

Overall, the project aims have been met: the time to complete an inversion has been lowered, without loss of accuracy, with the use of ML algorithms; the inversion of reactor data showed that the iterative algorithms only marginally improved on the ML estimate. Further, a robust modelling methodology has been demonstrated, in which a constrained optimisation algorithm tunes the forward model. Additionally, ML methods for crack detection were studied for feature extraction and classification; it was shown that ML methods could separate the effect of conductivity variations and cracks on the measurement response, but requires a low amount of noise in the system and good modelling accuracy. The comparison of algorithms consolidates previous work in

the graphite area and supports the past use of the LM algorithm. The study included variations in the choice of search direction, regularisation matrix, prior estimate and noise level in the inductance data. Generally, QN methods and the LM algorithm performed well. However, with increased levels of noise the type of regularisation becomes as important as the search direction. Further, if there is considerable noise in the data, there may be no benefit to the QN methods because the residual error in the Hessian does not become significant. The study used knowledge of the optimal step to determine the hyper-parameters, and therefore, in a realistic algorithm the classifications may be different; however, it is a useful comparison of the baseline performance of the different algorithms.

ML methods for graphite inversion, and specifically, their use in tandem with deductive iterative techniques has been reported. Direct solutions obtained using inference were investigated; two ML algorithms were compared, one using a MVPR model and the other a CNN. The MVPR algorithm is suitable when there is a low amount of training data, in which it remains competitive and computationally inexpensive compared with a more advanced CNN model. With respect to the results on smoothed test data, the CNN had a MPEN of 17.8% and the MVPR 17.3%. For reactor-type profiles the CNN had a MPEN of 14.9% and the MVPR 14.3%. Generally, the CNN seemed to have a greater noise immunity compared to the MVPR algorithm.

Beyond the nuclear graphite application and in electromagnetic tomography in general, there is a lack of literature surrounding the calibration of a forward model. Typically, the forward model is mapped to the measured experimental data using either a scalar value or a transfer function. These methods assume that the scaling factor or transfer function is conductivity invariant and independent of the measurement and this may not be justifiable, particularly if there are resonant effects present. Instead, the model can be tuned by varying the properties of the coil until there is sufficient agreement

between the model and experimental data. In previous work, the diameter of the coil has been tuned by using a grid search. This is feasible for a low number of parameters but the computation time exponentially increases with respect to the parameter number. An optimisation method has been presented for the calibration of a FEM for use in an eddy current inspection application. The algorithm requires no further theory beyond trust region algorithms; further, it can be used in both constrained and unconstrained optimisation. It has been demonstrated how the choice of FEM and measurement can be determined through this calibration procedure; for example, the coil model, mesh size and operating frequency range. It has been shown that a filament coil model is as accurate as a 3D model of the eddy current probe. In the deployed rector probe, there were visible resonance effects in the low frequency measurements. This is problematic if resonance is not included in the modelling; therefore, damping circuitry was used to reduce resonance effects. The calibration of the reactor probe with and without damping, showed that the damping circuitry did improve modelling accuracy. Using a filament model, the mean SNR of the laboratory system was 28.27 dB, the undamped reactor system 20.59 dB and the damped 33.99 dB. The resonant issues necessitated the use of the relative error in the objective function because the errors due to resonance would dominate the choice of coil parameters if the absolute error was used.

In the inversion of graphite inductance data, the iterative algorithm was highly restrictive of the step direction. This was not because the regularisation and damping parameters were too large but because the residual norm was below the discrepancy norm. The repeated inversions with a lower discrepancy multiplier revealed this to be case. However, lowering the discrepancy multiplier increases susceptibility to modelling errors. Therefore, to improve the accuracy the modelling error should be lowered. This modelling error may be sufficiently reduced with the damped gradiometer probe; it is to be

seen whether this is the case or not. Generally, the bore values were highly accurate, but the error increases with depth. This is partially due to the ill-conditioning of the problem, where this same trend is shown on synthetic data without noise. It was demonstrated that the reconstruction using the relative error is more erroneous than the reconstruction using the absolute error; this is due to the increased weighting of the low frequency data, in which there is a lower SNR.

Crack detection has previously relied on a running average of the experimental measurement, in an attempt to smooth out the effect of conductivity variations on the measurement data. This is inadequate if the azimuthal conductivity profile is not derived from the same probability distribution. A possibility is to not use averaging but instead use a feature extraction technique to predict the presence of a crack. The difficulties with this is that a crack can be understood as a sharp loss of conductivity, and therefore, there are a significant overlap in features in the measurement data from conductivity variations and cracks. A wide range of ML algorithms were studied for binary classification of a notch within a conductive block of material, in which there are variations in electrical conductivity. It was shown that for synthetic data with realistic levels of noise that the balanced classification accuracy was no larger than 71.45%. This accuracy was not limited by the underlying conductivity variations since a significantly higher accuracy can be achieved on synthetic data without noise. It is difficult to compare these ML feature extraction and classification algorithms to techniques that use averaging because the conductivity variations need to be well understood to study averaging. However, alternative measurements that use averaging in the ML algorithms may result in improved accuracies for a similar SNR level.

9.2 Future Work

Although there were many contributions to previous work in the inspection of graphite, there are incomplete areas of research. This section discusses some of these incomplete areas, where it is hoped that this would be useful to not just the graphite application but the wider community in eddy current inspection.

The ML prior estimate used in this thesis predicted the conductivity depth profile directly. Further, the algorithm was trained using polynomial conductivity profiles of up to order 5 (for the inversion of reactor data). This may be too general, since the conductivity is typically expected to be monotonically increasing. It would perhaps be better to define the functional explicitly and to use an ML algorithm to predict the coefficients of the functional. However, this may be unfavourable for the safety case. Alternatively, the inspections could be limited to trepanned locations only, where the trepanned data could then be used to determine the prior and not the measured inductance data. This would make the use of an ML prior redundant, greatly simplifying the problem.

In characterising the measurement noise, it may be best to conduct a wider study across a greater number of calibration bricks to see if the noise profile can be modelled more precisely than using a uniform random distribution. This may improve the accuracy of direct solutions using inferential techniques.

Although ML methods may not be necessary for predicting the prior, there is significant scope for replacing the FEM forward model with a ML forward model. Moreover, there may be a possibility of combining the two for a hybrid methodology. The ML algorithms designed in this thesis make no consideration of the physics of the problem. There may be great performance benefits to training a ML algorithm subject to the constraints of the physics.

In this thesis, the conductivity map in the forward model was approximated with a discrete depth profile. Similar to the ML prior comments, it may be simpler to use a functional to define a continuous depth profile. A possible benefit to this is in solving a problem with improved conditioning because the functional could regularise the solution; for example, a linear function could be used in which there are no high frequency variations in the depth profile. It would be sufficient to obtain the sensitivities using a forward difference approximation; however, analytic solutions may be obtained using the $E \cdot E$ formulation and the chain rule, an example derivation is given in the appendix. It is suspected that a further benefit of using a continuous depth profile is a simplification of the meshing in the numeric solver, since the material under test would not need to be discretised (the depth profile can be analytically specified in COMSOL®). This would lead to a faster solution time.

Another possibility is to use a global optimisation technique, rather than the local iterative methods. Iterative methods find a local minimum and this may not be the optimum solution. With the development of ML, there has been an expansion of research into fast global optimisation algorithms, such as the HyperBand technique. It is envisioned that a fast ML forward model could be embedded in such an algorithm.

Although this is the third iteration of the project, there are many possible areas of development which remain relevant to the wider eddy current modality. It is difficult to see where significant reductions in time can be gained from the design of the iterative algorithm; however, there may be gains in speedup with the use of a continuous depth profile in the FEM. The use and development of iterative techniques remains important as a deductive technique, where they can be used to validate batches of inverted reactor data derived from data-driven techniques.

Bibliography

[1] P. Cawley. "UK Research Centre in Non-Destructive Evaluation (RCNDE) 2014-2020." (2014), [Online]. Available: https://gtr.ukri.org/projects?ref=EP%5C%2FL022125%5C%2F1#/tabOverview (visited on Jul. 14, 2022).

[2] Committee on Climate Change, "Net Zero — The UK's contribution to stopping global warming," 2019. [Online]. Available: https://www.theccc.org.uk/publication/net-zero-the-uks-contribution-to-stopping-global-warming/ (visited on Jul. 14, 2022).

[3] Committee on Climate Change, "Hydrogen in a low-carbon economy," 2018. [Online]. Available: https://www.theccc.org.uk/publication/hydrogen-in-a-low-carbon-economy/ (visited on Jul. 14, 2022).

[4] Committee on Climate Change, "Net Zero Technical report," 2019. [Online]. Available: https://www.theccc.org.uk/publication/net-zero-technical-report/ (visited on Jul. 14, 2022).

[5] HM Government, "The Ten Point Plan for a Green Industrial Revolution," 2020. [Online]. Available: https://www.gov.uk/government/publications/the-ten-point-plan-for-a-green-industrial-revolution (visited on Jul. 14, 2022).

[6] Department for Business, Energy and Industrial Strategy, *Digest of UK Energy Statistics (DUKES): energy*, data retrieved from: Aggregate energy balances (DUKES 1.1-1.3), Available: https://www.gov.

uk/government/statistics/electricity-chapter-5-digest-of-united-kingdom-energy-statistics-dukes, 2021.

[7] Department for Business, Energy and Industrial Strategy, *Digest of UK Energy Statistics (DUKES): electricity*, data retrieved from: Electricity fuel use, generation and supply (DUKES 5.6), Available: https://www.gov.uk/government/statistics/electricity-chapter-5-digest-of-united-kingdom-energy-statistics-dukes, 2022.

[8] International Atomic Energy Agency. "Other Designs of Nuclear Power Stations." (2022), [Online]. Available: https://nucleus.iaea.org/sites/graphiteknowledgebase/wiki/Guide_to_Graphite/Other%20Designs%20of%20Nuclear%20Power%20Stations.aspx (visited on Jul. 14, 2022).

[9] EDF. "Extending the operating lives of Advanced Gas-cooled Reactors." (2022), [Online]. Available: https://www.edfenergy.com/energy/nuclear-lifetime-management (visited on Jun. 30, 2022).

[10] B. K. Sovacool, P. Schmid, A. Stirling, G. Walter, and G. MacKerron, "Differences in carbon emissions reduction between countries pursuing renewable electricity versus nuclear power," *Nature Energy*, vol. 5, pp. 928–935, 2020. DOI: 10.1038/s41560-020-00696-3.

[11] F. Wagner, "CO2 emissions of nuclear power and renewable energies: a statistical analysis of European and global data," *The European Physical Journal Plus*, vol. 136, no. 562, 2021. DOI: 10.1140/epjp/s13360-021-01508-7.

[12] J. Turner, "Neutrons, Fission, and Criticality," in *Atoms, Radiation, and Radiation Protection*, Wiley, 2007, pp. 230–233. DOI: 10.1002/9783527616978.

[13] Nuclear Data Center at KAERI. "Table of Nuclides." (2021), [Online]. Available: https://atom.kaeri.re.kr/nuchart/ (visited on Oct. 3, 2021).

[14] World Nuclear Association. "Heat Values of Various Fuels." (2016), [Online]. Available: https://world-nuclear.org/information-library/facts-and-figures/heat-values-of-various-fuels.aspx (visited on Jul. 14, 2022).

[15] libretexts. "General Properties and Reactions of The Actinides." (2021), [Online]. Available: https://chem.libretexts.org/@go/page/31905 (visited on Oct. 3, 2021).

[16] T. Kooyman, "Current state of partitioning and transmutation studies for advanced nuclear fuel cycles," *Annals of Nuclear Energy*, vol. 157, 2021. DOI: 10.1016/j.anucene.2021.108239.

[17] Nuclear Energy Agency, "Introduction," in *Minor Actinide Burning in Thermal Reactors: A Report by the Working Party on Scientific Issues of Reactor Systems*, Organisation for Economic Co-operation and Development, 2014, pp. 9–18. DOI: 10.1787/9789264208537-en.

[18] International Atomic Energy Agency, "Conclusion," in *Status of Minor Actinide Fuel Development*, 2010, p. 63, ISBN: 978-92-0-106909-2.

[19] B. Taebi, "Moral Dilemmas of Uranium and Thorium Fuel Cycles," in *Radioactivity in the Environment*, vol. 19, Elsevier, 2013, pp. 259–280. DOI: 10.1016/B978-0-08-045015-5.00015-0.

[20] National Energy Agency, "The Occurrence and Properties of Plutonium and its Compounds," in *Plutonium Fuel: An Assessment*, 1989, p. 29, ISBN: 978-92-0-106909-2.

[21] M. B. Schaffer, "Abundant thorium as an alternative nuclear fuel: Important waste disposal and weapon proliferation advantages," *Energy Policy*, vol. 60, pp. 4–12, 2013. DOI: 10.1016/j.enpol.2013.04.062.

[22] U. E. Humphrey and M. U. Khandaker, "Viability of thorium-based nuclear fuel cycle for the next generation nuclear reactor: Issues and prospects," *Renewable and Sustainable Energy Reviews*, vol. 97, pp. 259–275, 2018. DOI: 10.1016/j.rser.2018.08.019.

[23] National Nuclear Laboratory, "The Thorium Fuel Cycle," National Nuclear Laboratory, 2010. [Online]. Available: https://www.nnl.co.uk/wp-content/uploads/2019/01/nnl__1314092891_thorium_cycle_position_paper.pdf (visited on Jul. 14, 2022).

[24] R. Eychenne, M. Chérel, F. Haddad, F. Guérard, and J. Gestin, "Overview of the Most Promising Radionuclides for Targeted Alpha Therapy: The "Hopeful Eight"," *Pharmaceutics*, vol. 13, no. 6, 2021. DOI: 10.3390/pharmaceutics13060906.

[25] B. Nelson, J. Andersson, and F. Wuest, "Targeted Alpha Therapy: Progress in Radionuclide Production, Radiochemistry, and Applications," *Pharmaceutics*, vol. 13, no. 1, 2020. DOI: 10.3390/pharmaceutics13010049.

[26] E. Nonbel, "Description of the Advanced Gas Cooled Type of Reactor (AGR)," Riso National Laboratory, 1996.

[27] J. K. Shultis and R. E. Faw, "Nuclear Power," in *Fundamentals of Nuclear Science and Engineering*, Taylor and Francis, 2016, pp. 395–400. DOI: 10.1201/9781315183183.

[28] Nuclear Innovation and Research Office, "Advanced modular reactors (AMRs): technical assessment," Department for Business, Energy and Industrial Strategy, 2021. [Online]. Available: https://www.gov.uk/government/publications/advanced-modular-reactors-amrs-technical-assessment (visited on Jul. 14, 2022).

[29] T. Schulenberg, L. Leung, D. Brady, *et al.*, "Supercritical Water-Cooled Reactor (SCWR) Development through GIF Collaboration," in *Conf.*

on Opportunities and Challenges for Water Cooled Reactors in the 21st Century, International Atomic Energy Agency, 2010.

[30] World Nuclear Association. "Molten Salt Reactors." (2021), [Online]. Available: https://world-nuclear.org/information-library/current-and-future-generation/molten-salt-reactors.aspx (visited on Jul. 14, 2022).

[31] A. Alemberti, K. Tuček, M. Takahashi, *et al.*, "Lead-cooled Fast Reactor (LFR) System Safety Assessment," Generation IV International Forum, 2020. [Online]. Available: https://www.gen-4.org/gif/upload/docs/application/pdf/2020-06/gif_lfr_ssa_june_2020_2020-06-09_17-26-41_202.pdf (visited on Jul. 14, 2022).

[32] M. Joyce, "Advanced Reactors and Future Concepts," in *Nuclear Engineering: A Conceptual Introduction to Nuclear Power*, Butterworth-Heinemann, 2018, pp. 263–295. DOI: 10.1016/B978-0-08-100962-8.00011-1.

[33] "Safety Assessment Principles for Nuclear Facilities," Office for Nuclear Regulation, 2014, Revised 2021. [Online]. Available: https://www.onr.org.uk/saps/saps2014.pdf (visited on Jul. 14, 2022).

[34] "Safety Assessment Principles for Nuclear Facilities," Office for Nuclear Regulation, 2014, Revised 2021, p. 9. [Online]. Available: https://www.onr.org.uk/saps/saps2014.pdf (visited on Jul. 14, 2022).

[35] "Safety Assessment Principles for Nuclear Facilities," Office for Nuclear Regulation, 2014, Revised 2021, p. 87. [Online]. Available: https://www.onr.org.uk/saps/saps2014.pdf (visited on Jul. 14, 2022).

[36] Frost and Sullivan, "Third-party NDT Technologies to Drive the Global NDT Inspection Services Market Growth, 2020," Frost and Sullivan, 2020.

[37] Frost and Sullivan, "Global Non-destructive Test Equipment Market, Forecast to 2022," Frost and Sullivan, 2017.

[38] A. G. Steer, "Reactor Designs and operation: AGR core design, operation and safety functions," in *Management of Ageing in Graphite Reactor Cores*, Royal Society of Chemistry, 2007, pp. 11–18. DOI: 10.1039/9781847557742-00001.

[39] B. J. Marsden, A. N. Jones, G. Hall, and M.Treifi, "Graphite as a core material for Generation IV nuclear reactors," in *Structural Materials for Generation IV Nuclear Reactors*, Woodhead Publishing, 2017, pp. 495–532. DOI: 10.1016/B978-0-08-100906-2.00014-8.

[40] W. Eatherly and E. Piper, "Manufacture," in *Nuclear Graphite*, Academic Press, 1962, pp. 21–51. DOI: 10.1016/B978-1-4832-2854-9.50006-7.

[41] B. J. Marsden, M. Haverty, W. Bodel, *et al.*, "Dimensional change, irradiation creep and thermal/mechanical property changes in nuclear graphite," *International Materials Reviews*, vol. 61, no. 3, pp. 155–182, 2016. DOI: 10.1080/09506608.2015.1136460.

[42] B. J. Marsden, G. N. Hall, and A. N. Jones, "Graphite in Gas-Cooled Reactors," in *Comprehensive Nuclear Materials*, vol. 4, Elsevier Publishing, 2012, pp. 325–390. DOI: 10.1016/B978-0-08-056033-5.00092-6.

[43] J. Best, W. Stephen, and A. Wickham, "Radiolytic graphite oxidation," *Progress in Nuclear Energy*, vol. 16, no. 2, pp. 127–178, 1985. DOI: 10.1016/0149-1970(85)90002-2.

[44] "Assessment of EDF's graphite materials properties model in NP/SC 7766 V10 SS1 (Operation of HNB R3 to a burn up of 16.425TWd) and implications for the predicted brick stresses, dimensions and overall damage tolerance assessment," Office for Nuclear Regulation, 2020,

pp. 1–68. [Online]. Available: https://www.onr.org.uk/documents/2020/ofd-ar-19-093.pdf (visited on Jul. 14, 2022).

[45] M. Fahad, K. McNally, G. N. Hall, N. Warren, B. J. Marsden, and P. Mummery, "Variation of internal stresses in nuclear graphite brick in AGRs," in *Conf. on Structural Mechanics in Reactor Technology*, 2015. [Online]. Available: https://repository.lib.ncsu.edu/bitstream/handle/1840.20/33806/SMiRT-23_Paper_289.pdf?sequence=1&isAllowed=y (visited on Jul. 14, 2022).

[46] M. Treifi, B. J. Marsden, G. N. Hall, and P. M. Mummery, "Keyway Root Crack Arrest in an AGR Graphite Brick," in *Conf. on Structural Mechanics in Reactor Technology*, 2015. [Online]. Available: https://repository.lib.ncsu.edu/bitstream/handle/1840.20/33815/SMiRT-23_Paper_190.pdf?sequence=1&isAllowed=y (visited on Jul. 14, 2022).

[47] S. Kyaw, D. Tanner, A. Becker, W. Sun, and D. Tsang, "Modelling Crack Growth within Graphite Bricks due to Irradiation and Radiolytic Oxidation," in *Proc. of the 20th European Conference on Fracture*, vol. 3, 2014, Pages 39–44. DOI: 10.1016/j.mspro.2014.06.009.

[48] A. Cole-Baker and J. Reed, "Measurement of AGR Graphite Fuel Brick-Shrinkage and Channel Distortion," in *Management of Ageing in Graphite Reactor Cores*, Royal Society of Chemistry, 2007, pp. 169–248. DOI: 10.1039/9781847557742-00125.

[49] M. Brown, J. Williams, T. Bloodworth, and M. Anderson, "Eddy current technique development - inspection results from AGR graphite reactor cores," in *The 4th EDF Energy Nuclear Graphite Symposium: Engineering Challenges Associated with the Life of Graphite Reactor Cores*, EMAS publishing, 2015, ISBN: 9780957673052.

[50] T. Bloodworth, M. Anderson, M. Brown, and J. Williams, "Eddy current inspection of AGR fuel channels: keyway cracks and density mapping," in *Proc. NDT*, BINDT, 2012. [Online]. Available: https://www.bindt.org/downloads/NDT2012_5A3.pdf (visited on Jul. 14, 2022).

[51] B. Dekdouk, R. Chapman, M. Brown, and A. J. Peyton, "Evaluating the conductivity distribution in isotropic polycrystalline graphite using spectroscopic eddy current technique for monitoring weight loss in advanced gas cooled reactors," *NDT and E International*, vol. 51, pp. 150–159, 2012. DOI: 10.1016/j.ndteint.2012.06.011.

[52] A. Fletcher, "Non-Destructive Testing of the Graphite Core within an Advanced Gas-Cooled Reactor," Ph.D. dissertation, University of Manchester, 2014.

[53] H. Tesfalem, "Eddy Current Based Non-Destructive Testing of the Advanced Gas-Cooled Reactor Core," Ph.D. dissertation, University of Manchester, 2018.

[54] H. Tesfalem, A. D. Fletcher, M. Brown, B. Chapman, and A. J. Peyton, "Study of asymmetric gradiometer sensor configurations for eddy current based non-destructive testing in an industrial environment," *NDT and E International*, vol. 100, pp. 1–10, 2018. DOI: 10.1016/j.ndteint.2018.07.007.

[55] S. M. Nair and J. H. Rose, "Reconstruction of three-dimensional conductivity variations from eddy current (electromagnetic induction) data," *Inverse Problems*, vol. 6, no. 6, pp. 1007–1030, 1990. DOI: 10.1088/0266-5611/6/6/010.

[56] A. Tamburrino, F. Calvano, S. Ventre, and G. Rubinacci, "Non-iterative imaging method for experimental data inversion in eddy current tomography," *NDT and E International*, vol. 47, pp. 26–34, 2012. DOI: 10.1016/j.ndteint.2011.11.013.

[57] W. Yin and A. J. Peyton, "Sensitivity Formulation Including Velocity Effects for Electromagnetic Induction Systems," *IEEE Transactions on Magnetics*, vol. 46, no. 5, pp. 1172–1176, 2010. DOI: 10.1109/TMAG.2009.2038275.

[58] M. Cheney, D. Isaacson, J. Newell, S. Simiske, and J. Goble, "NOSER: An Algorithm for Solving the Inverse Conductivity Problem," *International Journal of Imaging Systems and Technology*, vol. 2, no. 2, pp. 66–75, 1990. DOI: 10.1002/ima.1850020203.

[59] M. Cheney, D. Isaacson, and J. Newell, "Electrical Impedance Tomography," *Society for Industrial and Applied Mathematics Review*, vol. 41, no. 1, pp. 85–101, 1999. [Online]. Available: https://www.jstor.org/stable/2653168.

[60] W. Yin, S. J. Dickinson, and A. J. Peyton, "Imaging the Continuous Conductivity Profile Within Layered Metal Structures Using Inductance Spectroscopy," *IEEE Sensors Journal*, vol. 5, no. 2, pp. 161–166, 2005. DOI: 10.1109/JSEN.2004.842637.

[61] S. Norton, A. Kahn, and M. Mester, "Reconstructing Electrical Conductivity Profiles from Variable-Frequency Eddy Current Measurements," *Research in Nondestructive Evaluation*, vol. 1, pp. 167–179, 1989. DOI: 10.1007/BF01606513.

[62] M. Soleimani and W. R. B. Lionheart, "Nonlinear image reconstruction for electrical capacitance tomography using experimental data," *Measurement Science and Technology*, vol. 16, no. 10, pp. 1987–1996, 2005. DOI: 10.1088/0957-0233/16/10/014.

[63] J. R. Bowler and S. J. Norton, "Eddy Current Inversion for Layered Conductors," *Research in Nondestructive Evaluation*, vol. 4, no. 4, pp. 205–219, 1992. DOI: 10.1007/BF01616488.

[64] C. V. Dodd and W. E. Deeds, "Analytical Solutions to Eddy-Current Probe-Coil Problems," *Journal of Applied Physics*, vol. 39, no. 6, pp. 2829–2838, 1968. DOI: 10.1063/1.1656680.

[65] T. P. Theodoulidis, "Impedance of a coil above a planar conductor with an arbitrary continuous conductivity depth profile," *International Journal of Applied Electromagnetics and Mechanics*, vol. 59, no. 4, pp. 1179–1185, 2019. DOI: 10.3233/JAE-171122.

[66] A. Adler and R. Guardo, "A Neural Network Image Reconstruction Technique for Electrical Impedance Tomography," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 594–600, 1994. DOI: 10.1109/42.363109.

[67] I. T. Rekanos, T. P. Theodoulidis, S. M. Panas, and T. D. Tsiboukis, "Impedance inversion in eddy current testing of layered planar structures via neural networks," *NDT and E International*, vol. 30, no. 2, pp. 69–74, 1997. DOI: 10.1016/S0963-8695(96)00047-3.

[68] C. Glorieux, J. Moulder, J. Basart, and J. Thoen, "The determination of electrical conductivity profiles using neural network inversion of multi-frequency eddy-current data," *Journal of Physics D: Applied Physics*, vol. 32, no. 5, pp. 616–622, 1999. DOI: 10.1088/0022-3727/32/5/018.

[69] V. Puzyrev, "Deep learning electromagnetic inversion with convolutional neural networks," *Geophysical Journal International*, vol. 218, no. 2, pp. 817–832, 2019. DOI: 10.1093/gji/ggz204.

[70] L. Liang, M. Liu, C. Martin, and W. Sun, "A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis," *Journal of the Royal Society Interface*, vol. 15, no. 138, 2018. DOI: 10.1098/rsif.2017.0844.

[71] P. V. Badarinath, M. Chierichetti, and F. D. Kakhki, "A Machine Learning Approach as a Surrogate for a Finite Element Analysis: Status of Research and Application to One Dimensional Systems," *Sensors*, vol. 21, no. 5, 2021. DOI: 10.3390/s21051654.

[72] P. Ramuhalli, L. Udpa, and S. Udpa, "Finite-element neural networks for solving differential equations," *IEEE Transactions on Neural Networks*, vol. 16, no. 6, pp. 1381–1392, 2005. DOI: 10.1109/TNN.2005.857945.

[73] A. Reynaldi, S. Lukas, and H. Margaretha, "Backpropagation and Levenberg-Marquardt Algorithm for Training Finite Element Neural Network," in *in Proc. 6th European Modelling Symposium*, 2012. DOI: 10.1109/EMS.2012.56.

[74] J. Takeuchi and Y. Kosugi, "Neural network representation of finite element method," *Neural Networks*, vol. 7, no. 2, pp. 389–395, 1994. DOI: 10.1016/0893-6080(94)90031-0.

[75] M. Fan, B. Cao, P. Yang, W. Lia, and G. Tian, "Elimination of liftoff effect using a model-based method for eddy current characterization of a plate," *NDT and E International*, vol. 74, pp. 66–71, 2015. DOI: 10.1016/j.ndteint.2015.05.007.

[76] A. Tamburrino and G. Rubinacci, "Fast methods for quantitative eddy-current tomography of conductive materials," *IEEE Transactions on Magnetics*, vol. 42, no. 8, pp. 2017–2028, 2006. DOI: 10.1109/TMAG.2006.877542.

[77] Y. He, M. Pan, F. Luo, D. Chen, and X. Hu, "Support vector machine and optimised feature extraction in integrated eddy current instrument," *Measurement*, vol. 46, no. 1, pp. 764–774, 2013. DOI: 10.1016/j.measurement.2012.09.014.

[78] L. Udpa and S. Udpa, "Neural networks for the classification of non-destructive evaluation signals," *IEE Proceedings F - Radar and Signal Processing*, vol. 138, no. 1, pp. 41–45, 1991. DOI: 10.1049/ip-f-2.1991.0007.

[79] L. S. Rosado, F. M. Janeiro, P. M. Ramos, and M. Piedade, "Defect Characterization With Eddy Current Testing Using Nonlinear-Regression Feature Extraction and Artificial Neural Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5, pp. 1207–1214, 2013. DOI: 10.1109/TIM.2012.2236729.

[80] A. Sophian, G. Tian, D. Taylor, and J. Rudlin, "A feature extraction technique based on principal component analysis for pulsed Eddy current NDT," *NDT and E International*, vol. 36, no. 1, pp. 37–41, 2003. DOI: 10.1016/S0963-8695(02)00069-5.

[81] T. Chen, G. Y. Tian, A. Sophian, and P. W. Queb, "Feature extraction and selection for defect classification of pulsed eddy current NDT," *NDT and E International*, vol. 41, no. 6, pp. 467–476, 2008. DOI: 10.1016/j.ndteint.2008.02.002.

[82] W. Q. Yang and L. Peng, "Image reconstruction algorithms for electrical capacitance tomography," *Measurement Science and Technology*, vol. 14, no. 1, pp. 1–13, 2002. DOI: 10.1088/0957-0233/14/1/201.

[83] L.Yin, B. Ye, Z. Zhang, *et al.*, "A novel feature extraction method of eddy current testing for defect detection based on machine learning," *NDT and E International*, vol. 107, pp. 1–7, 2019. DOI: 10.1016/j.ndteint.2019.04.005.

[84] G. Y. Tian and A. Sophian, "Defect classification using a new feature for pulsed eddy current sensors," *NDT and E International*, vol. 38, no. 1, pp. 77–82, 2005. DOI: 10.1016/j.ndteint.2004.06.001.

[85] A. Sophian, G. Tian, and M. Fan, "Pulsed Eddy Current Non-destructive Testing and Evaluation: A review," *Chinese Journal of Mechanical Engineering*, vol. 30, no. 3, pp. 500–514, 2017. DOI: 10.1007/s10033-017-0122-4.

[86] R. Himawan, F. Lie, P. D. Basoeki, and M. Haryanto, "Applicabilty Study of Ultrasonic Flaw Detector For Nuclear Grade Graphite Examination," *Journal of Physics: Conference Series*, vol. 1198, no. 2, 2019. DOI: 10.1088/1742-6596/1198/2/022018.

[87] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, "Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization," *Journal of Machine Learning Research*, vol. 18, no. 185, pp. 1–52, 2018. DOI: 10.48550/arXiv.1603.06560.

[88] A. Zangwill, "Waves in Dispersive Matter," in *Modern Electrodynamics*, Cambridge University Press, 2013, pp. 624–659. DOI: 10.1017/CBO9781139034777.

[89] R. Feynman, "Dielectrics," in *The Feynman Lectures on Physics, Volume II: mainly electromagnetism and matter*, vol. 2, Basic Books, 2010, ISBN: 9780465024940. [Online]. Available: https://www.feynmanlectures.caltech.edu/II_toc.html.

[90] A. Zangwill, "Waves in Simple Matter," in *Modern Electrodynamics*, Cambridge University Press, 2013, pp. 584–617. DOI: 10.1017/CBO9781139034777.

[91] Z. Mottl, "The quantitative relations between true and standard depth of penetration for air-cored probe coils in eddy current testing," *NDT International*, vol. 23, no. 1, pp. 11–18, 1990. DOI: 10.1016/0308-9126(90)91444-X.

[92] P. Stucky and W. Lord, "Skin Depth Considerations in Eddy Current NDT," *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 28, pp. 299–306, 1992. DOI: 10.1007/978-1-4615-3344-3_37.

[93] W. Yin, S. Dickinson, and A. Peyton, "Evaluating the Permeability Distribution of a Layered Conductor by Inductance Spectroscopy," *IEEE Transactions on Magnetics*, vol. 42, no. 11, pp. 3645–3651, 2006. DOI: 10.1109/TMAG.2006.880992.

[94] Y. Li, B. Yan, W. Li, and D. Li, "Thickness Assessment of Thermal Barrier Coatings of Aeroengine Blades Via Dual-Frequency Eddy Current Evaluation," *IEEE Magnetics Letters*, vol. 7, pp. 1–5, 2016. DOI: 10.1109/LMAG.2016.2590465.

[95] O. Bíró, "Edge element formulations of eddy current problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 169, no. 3-4, pp. 391–405, 1999. DOI: 10.1016/S0045-7825(98)00165-0.

[96] E. Kriezis, T. Tsiboukis, S. Panas, and J. Tegopoulos, "Eddy currents: theory and applications," *Proceedings of the IEEE*, vol. 80, no. 10, pp. 1559–1589, 1992. DOI: 10.1109/5.168666.

[97] COMSOL, "Formulations, Methods, Technologies," in *ACDC Module Users Guide*, Version 6.0 — available with package installation, COMSOL, 2022, pp. 50–54.

[98] Y. Li, "Edge based finite element simulation of eddy current phenomenon and its application to defect characterization," Ph.D. dissertation, Iowa state University, 2002. DOI: 10.31274/rtd-180813-13188.

[99] J. L. Volakis, "Shape functions for scalar and vector finite elements," in *Finite Element Method Electromagnetics: Antennas, Microwave Circuits, and Scattering Applications*, Wiley-Interscience, 1998, pp. 48–62, ISBN: 978-0-780-33425-0.

[100] M. Chari and S. Salon, "The Finite Element Method," in *Numerical Methods in electromagnetism*, Wiley-Interscience, 2000, pp. 353–357, ISBN: 9780080512891.

[101] N. Ottosen and H. Petersson, "choice of weight function- weighted residual methods," in *Introduction to the finite element method*, Prentice Hall, 1992, pp. 142–156, ISBN: 9780134738772.

[102] COMSOL. "The Finite Element Method (FEM)." (2017), [Online]. Available: https://www.comsol.com/multiphysics/finite-element-method?parent=physics-pdes-numerical-042-62 (visited on Jul. 14, 2022).

[103] H. Tesfalem, A. J. Peyton, A. D. Fletcher, M. Brown, and B. Chapman, "Eddy Current Sensor and Inversion Technique for Estimation of the Electrical Conductivity Profile of the Graphite Bricks in an Advanced Gas-Cooled Reactor Core," *Studies in Applied Electromagnetics and Mechanics: Electromagnetic Non-Destructive Evaluation*, vol. 42, pp. 253–264, 2017. DOI: 10.3233/978-1-61499-767-2-253.

[104] D. Dyck, D. Lowther, and E. Freeman, "A method of computing the sensitivity of electromagnetic quantities to changes in materials and sources," *IEEE Transactions on magnetics*, vol. 30, no. 5, pp. 3415–3418, 1994. DOI: 10.1109/20.312672.

[105] R. Feynman, "Field Energy and Field Momentum," in *The Feynman Lectures on Physics, Volume II: mainly electromagnetism and matter*, vol. 2, Basic Books, 2010, ISBN: 9780465024940. [Online]. Available: https://www.feynmanlectures.caltech.edu/II_toc.html.

[106] P. Penfield, R. Spence, and S. Duinker, "Extensions to other physical systems," in *Tellegen's theorem and electrical networks*, The MIT Press, 1970, pp. 109–111, ISBN: 9780262160322.

[107] G. Strang, "Least Squares Approximations," in *Introduction to Linear Algebra*, Wellesley-Cambridge Press, 2016, ISBN: 978-09802327-7-6.

[108] A. Conn, M. Gould, and P. Toint, "The Trust-Region Subproblem," in *Trust-Region Methods*, Society for Industrial and Applied Mathematics, 2000, pp. 169–248. DOI: 10.1137/1.9780898719857.

[109] A. Conn, M. Gould, and P. Toint, "Global Convergence of the Basic Algorithm," in *Trust-Region Methods*, Society for Industrial and Applied Mathematics, 2000, pp. 162–166. DOI: 10.1137/1.9780898719857.

[110] R. Fletcher, "Newton-like Methods," in *Practical Methods of Optimization*, 2nd ed., Wiley, 2000, pp. 44–74. DOI: 10.1002/9781118723203.

[111] G. A. F. Seber and C. J. Wild, "Computational Methods for Nonlinear Least Squares," in *Nonlinear Regression*, Wiley, 1989, pp. 619–660. DOI: 10.1002/0471725315.

[112] J. Nocedal and S. J. Wright, "Trust-Region Methods," in *Numerical Optimization*, Springer, 2006, pp. 66–98. DOI: 10.1007/978-0-387-40065-5.

[113] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963. [Online]. Available: https://www.jstor.org/stable/2098941.

[114] R. Fletcher, "A modified Marquardt subroutine for non-linear least squares," *United Kingdom Atomic Energy Authority*, 1971. [Online]. Available: https://epubs.stfc.ac.uk/manifestation/6683/AERE_R_6799.pdf.

[115] T. Dai, M. Soleimani, and A. Adler, "Four-Dimensional Regularization for Electrical Impedance Tomography Imaging," in *in Proc. IFMBE*, vol. 17, Springer, 2007, pp 408–411. DOI: 10.1007/978-3-540-73841-1_106.

[116] A. A. Nielsen, "Least Squares Adjustment: Linear and Nonlinear Weighted Regression Analysis," Technical University of Denmark, 2013. [Online]. Available: https://www2.imm.dtu.dk/pubdb/edoc/imm2804.pdf.

[117] G. Strang, "The Singular Value Decomposition (SVD)," in *Introduction to Linear Algebra*, Wellesley-Cambridge Press, 2016, ISBN: 978-09802327-7-6.

[118] P. C. Hansen, "Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems," *Numerical Algorithms*, vol. 6, pp. 1–35, 1994. DOI: 10.1007/BF02149761.

[119] G. Golub and C. Loan, "Modified Least Squares Problems and Methods," in *Matrix Computations*, 4th ed., John Hopkins University Press, 2013, pp. 303–346, ISBN: 978-1421407944.

[120] H. Tesfalem, A. J. Peyton, A. D. Fletcher, M. Brown, and B. Chapman, "Conductivity Profiling of Graphite Moderator Bricks From Multifrequency Eddy Current Measurements," *IEEE Sensors Journal*, vol. 20, no. 9, pp. 4840–4849, 2020. DOI: 10.1109/JSEN.2020.2965201.

[121] J. Nocedal and S. J. Wright, "Least-Squares Problems," in *Numerical Optimization*, Springer, 2006, pp. 245–269. DOI: 10.1007/978-0-387-40065-5.

[122] J. Nocedal and S. J. Wright, "Line Search Methods," in *Numerical Optimization*, Springer, 2006, pp. 30–63. DOI: 10.1007/978-0-387-40065-5.

[123] J. E. Dennis, D. M. Gay, and R. E. Welsch, "An Adaptive Nonlinear Least Square Algorithm," National Bureau of Economic Research, Tech. Rep. 196, 1977. DOI: 10.3386/w0196.

[124] Y. Bard, "Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems," *SIAM Journal on Numerical Analysis*, vol. 7, no. 1, pp. 157–186, 1970. DOI: 10.1137/0707011.

[125] H. Tesfalem, A. J. Peyton, A. D. Fletcher, M. Brown, and B. Chapman, "Resistivity Profiling of the Graphite Bricks in an Advanced Gas-Cooled Reactor Core using a Non-linear Inverse Eddy Current Method," in *in Proc. 12th European Conf. on Non-Destructive Testing*, 2018. [Online]. Available: https://www.ndt.net/search/docs.php3?id=22848.

[126] J. Nocedal and S. J. Wright, "Conjugate Gradient Methods," in *Numerical Optimization*, Springer, 2006, pp. 121–131. DOI: 10.1007/978-0-387-40065-5.

[127] A. Wirgin, *The inverse crime*, 2004. DOI: 10.48550/ARXIV.MATH-PH/0401050.

[128] A. Fakhrulddin, X. Fei, and H. Li, "Convolutional neural networks (CNN) based human fall detection on Body Sensor Networks (BSN) sensor data," in *4th International Conf. on Systems and Informatics*, 2017. DOI: 10.1109/ICSAI.2017.8248516.

[129] G. Tian and A. Sophian, "Defect classification using a new feature for pulsed eddy current sensors," *NDT and E International*, vol. 38, no. 1, pp. 77–82, 2005. DOI: 10.1016/j.ndteint.2004.06.001.

[130] I. Miccoli, F. Edler, H. Pfnür, and C. Tegenkamp, "The 100th anniversary of the four-point probe technique: the role of probe geometries in isotropic and anisotropic systems," *Journal of Physics: Condensed Matter*, vol. 27, no. 22, 2015. DOI: 10.1088/0953-8984/27/22/223201.

[131] P. C. Hansen, "Parameter-Choice Methods," in *Rank-Deficient and Discrete Ill-Posed Problems*, Society for Industrial and Applied Mathematics, 1998, pp. 175–208. DOI: 10.1137/1.9780898719697.

[132] P. C. Hansen, "The L-Curve and its Use in the Numerical Treatment of Inverse Problems," in *Computational Inverse Problems in Electrocar-*

diology, ed. P. Johnston, Advances in Computational Bioengineering, WIT Press, 2000, pp. 119–142.

[133] A. Cultrera and L. Callegaro, "A simple algorithm to find the L-curve corner in the regularisation of ill-posed inverse problems," *IOPSciNotes*, vol. 1, no. 2, pp. 1–6, 2020. DOI: 10.1088/2633-1357/abad0d.

[134] H. B. Nielsen, "Damping Parameter in Marquart's Method," Department of Mathematical Modelling - Technical University of Denmark, 1999, Report IMM-REP-1999-05. [Online]. Available: http://www2.imm.dtu.dk/documents/ftp/tr99/tr05_99.pdf.

[135] A. Fletcher, A. Peyton, D. Armitage, M. Brown, and R. Chapman, "Subsurface cracks in AGR graphite bricks," *NDT and E International*, vol. 66, pp. 72–81, 2014. DOI: 10.1016/j.ndteint.2014.05.001.

[136] I. Goodfellow, Y. Bengio, and A. Courville, "Machine Learning Basics," in *Deep Learning*, MIT Press, 2016, pp. 95–151, ISBN: 9780262035613.

[137] B. Schölkopf, A. Smola, and K. Müller, "Kernel principal component analysis," in *Artificial Neural Networks — ICANN'97*, Springer Berlin Heidelberg, 1997, pp. 583–588. DOI: 10.1007/BFb0020217.

[138] B. Flury, "Linear Principal Component Analysis," in *A First Course in Multivariate Statistics*, Springer, 1997, pp. 563–637. DOI: 10.1007/978-1-4757-2765-4.

[139] B. Schölkopf and A. J. Smola, "Kernel Feature Extraction," in *Learning with Kernels : Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, 2001, pp. 427–452, ISBN: 9780262536578. DOI: 10.7551/mitpress/4175.003.0019.

[140] M. Welling, "Kernel Principal Components Analysis," in *Advances in neural information processing systems 18*, 2005, pp. 70–72.

[141] A. Hyvärinen, "Independent component analysis: recent advances," *Philosophical transactions - Royal Society. Mathematical, Physical and engineering sciences*, vol. 371, no. 1984, 2013. DOI: 10.1098/rsta.2011.0534.

[142] I. Goodfellow, Y. Bengio, and A. Courville, "Autoencoders," in *Deep Learning*, MIT Press, 2016, pp. 502–524, ISBN: 9780262035613.

[143] L. Wasserman, "Linear and Logistic Regression," in *All of Statistics: A Concise Course in Statistical Inference*, Springer Texts in Statistics, 2010, pp. 209–229, ISBN: 978-0-387-21736-9. DOI: 10.1007/978-0-387-21736-9_13.

[144] M. Welling, "Lecture notes in Support Vector Machines," Department of Computer Science University of Toronto, 2004.

[145] J. Shawe-Taylor and N. Cristianin, "Support Vector Machines," in *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, 2000, pp. 93–124. DOI: 10.1017/CBO9780511801389.008.

[146] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

Appendix A

Sensitivities of Regularised Objective Function

A.1 Gradient

$$f(\vec{x}) = \frac{1}{2}\vec{r}^T\vec{r} + \frac{1}{2}(\mathbf{\Gamma}(\vec{x} - \vec{x}_0))^T\mathbf{\Gamma}(\vec{x} - \vec{x}_0)$$

Make the substitution $\Delta = \vec{x} - \vec{x}_0$

$$\Longrightarrow$$

$$f(\vec{x}) = \frac{1}{2}(r_1^2 + r_2^2 \cdots + r_n^2) + \frac{1}{2}\left(\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_m \end{bmatrix}\right)^T \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_m \end{bmatrix}$$

$$\Longrightarrow$$

$$f(\vec{x}) = \frac{1}{2}(r_1^2 + r_2^2 \cdots + r_n^2) + \frac{1}{2}\begin{bmatrix} \gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 + \cdots + \gamma_{1m}\Delta_m \\ \gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2m}\Delta_m \\ \vdots \\ \gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mm}\Delta_m \end{bmatrix}^T \begin{bmatrix} \gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 + \cdots + \gamma_{1n}\Delta_m \\ \gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2n}\Delta_m \\ \vdots \\ \gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mm}\Delta_m \end{bmatrix}$$

$\implies$

$$\begin{aligned} f(\vec{x}) = &\frac{1}{2}(r_1^2(\vec{x}) + r_2^2(\vec{x}) \cdots + r_n^2(\vec{x})) + \\ &\frac{1}{2}(\gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 + \cdots + \gamma_{1n}\Delta_n)^2 + \\ &\frac{1}{2}(\gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2n}\Delta_n)^2 + \\ &\frac{1}{2}(\gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mn}\Delta_n)^2 \end{aligned}$$

Taking the gradient

$$\begin{aligned} \nabla f(\vec{x}) = &\begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_m} \end{bmatrix} (\frac{1}{2}(r_1^2 + r_2^2 \cdots + r_n^2) + \\ &\frac{1}{2}(\gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 + \cdots + \gamma_{1m}\Delta_m)^2 + \\ &\frac{1}{2}(\gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2m}\Delta_m)^2 \cdots + \\ &\frac{1}{2}(\gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mm}\Delta_m)^2) \end{aligned}$$

Using the chain rule $\implies$

$$\nabla f(\vec{x}) = \boldsymbol{J}^T\vec{r}+$$

$$\begin{bmatrix} \gamma_{11}(\gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 \cdots + \gamma_{1m}\Delta_m) \\ \gamma_{12}(\gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 \cdots + \gamma_{1m}\Delta_m) \\ \vdots \\ \gamma_{1m}(\gamma_{11}\Delta_1 + \gamma_{12}\Delta_2 \cdots + \gamma_{1m}\Delta_m) \end{bmatrix}$$

$$+$$

$$\begin{bmatrix} \gamma_{21}(\gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2m}\Delta_m) \\ \gamma_{22}(\gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2m}\Delta_m) \\ \vdots \\ \gamma_{2m}(\gamma_{21}\Delta_1 + \gamma_{22}\Delta_2 \cdots + \gamma_{2m}\Delta_m) \end{bmatrix}$$

$$\cdots +$$

$$\begin{bmatrix} \gamma_{m1}(\gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mm}\Delta_m) \\ \gamma_{m2}(\gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mm}\Delta_m) \\ \vdots \\ \gamma_{mm}(\gamma_{m1}\Delta_1 + \gamma_{m2}\Delta_2 + \cdots + \gamma_{mm}\Delta_m) \end{bmatrix}$$

$$\Longrightarrow$$

$$\nabla f(\vec{x}) = \boldsymbol{J}^T\vec{r}(\vec{x}) + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}(\vec{x} - \vec{x}_0)$$

A.2 Hessian

The Hessian is populated with second order derivatives:

$$H_f(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 x_1} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_m} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2 x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m x_1} & \frac{\partial^2 f}{\partial x_m x_2} & \cdots & \frac{\partial^2 f}{\partial x_m x_m} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_m} \end{bmatrix} \nabla f(\vec{x})^T$$

$\Longrightarrow$

$$H_f(\vec{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_m} \end{bmatrix} (\vec{r}^T \boldsymbol{J} + (\vec{x} - \vec{x}_0)^T \boldsymbol{\Gamma}^T \boldsymbol{\Gamma})$$

Using the Gauss-Newton approximation [121], this simplifies to

$$H_f(\vec{x}) \approx \boldsymbol{J}^T \boldsymbol{J} + \boldsymbol{\Gamma}^T \boldsymbol{\Gamma}$$

A.3 Linear Mapping

In the case where the residual can be written with a linear mapping, $\vec{r} = \boldsymbol{A}\vec{x} + \vec{b}$, the regularised Gauss-Newton direction is an exact solution of the step towards the minimiser of $\text{argmin}_{\vec{x}}(\frac{1}{2}\left\|\boldsymbol{A}\vec{x} + \vec{b}\right\|_2^2 + \frac{1}{2}\left\|\boldsymbol{\Gamma}(\vec{x} - \vec{x}_0)\right\|_2^2)$, with $\boldsymbol{J} = \boldsymbol{A}$

$$\vec{p} = -(\boldsymbol{A}^T \boldsymbol{A} + \boldsymbol{\Gamma}^T \boldsymbol{\Gamma})^{-1}(\boldsymbol{A}^T \vec{r}(\vec{x}_k) + \boldsymbol{\Gamma}^T \boldsymbol{\Gamma}(\vec{x}_k - \vec{x}_0))$$

Instead of computing a step from a current iterate, the direct solution can be obtained by setting the initial point as $\vec{x}_k = \vec{0}$. The optimum solution is then

$$\begin{aligned}\vec{x^*} = \vec{p} &= -\left(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\right)^{-1}\left(\boldsymbol{A}^T\vec{r}(\vec{0}) + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}(\vec{0} - \vec{x}_0)\right) \\ &= -\left(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\right)^{-1}\left(\boldsymbol{A}^T\vec{b} - \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\vec{x}_0\right)\end{aligned}$$

Appendix B

Continuous Depth Profile of a Conductive Cuboid Block

Consider the discretisation of a cuboid block with a thickness L, shown in fig. B.1. The number of layers, n, controls the thickness of any-one layer, such that $d = L/n$. If the number of layers increases, the thickness of each layer

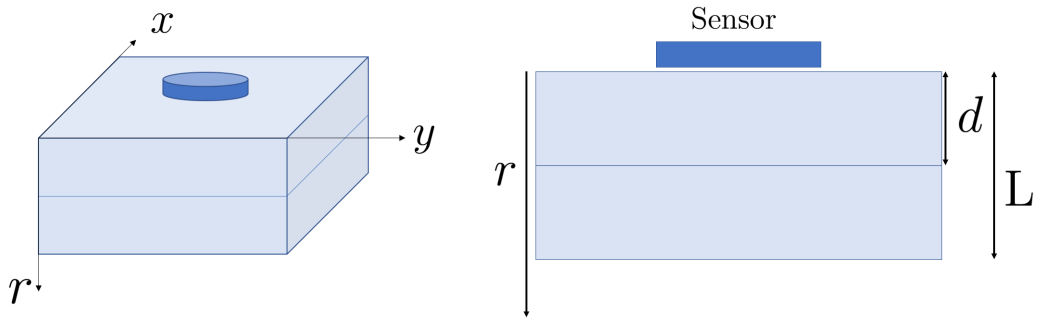

Figure B.1: Left: Discretisation and co-ordinate system. Right: Cuboid block with two layers: $d = L/2$.

decreases, as shown in fig. B.2. Consider the limit as the number of layers goes to infinity: $d \to 0$ and the volume of a layer reduces to a surface, S. In the simple case of a cuboid block, the surface is a flat plane described by the function $S(x, y, r) = x + y + r$, simply denoted as $S(r)$. This surface has a

conductivity $\sigma(r)$. The sensitivity of the measurement voltage to a surface is

$$\frac{\partial \psi}{\partial \sigma(r)} = \lim_{d \to 0} \int \vec{E}_{Tx} \cdot \vec{E}_{Rx} dV = \int \vec{E}_{Tx} \cdot \vec{E}_{Rx} dS(r) \tag{B.1}$$

In the discrete case, the voltage is a function of each conductivity vari-

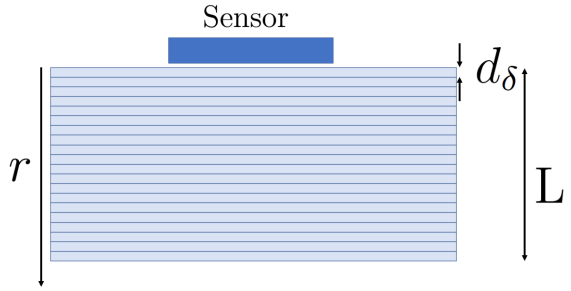

Figure B.2: Discretisation with 20 layers.

able, $\psi(\sigma_1, \sigma_2 \ldots, \sigma_n)$, and if each of these is a parametric function, such that $\psi(\sigma_1(t), \sigma_2(t) \ldots, \sigma_n(t))$, then the voltage sensitivity to t is

$$\frac{\partial \psi}{\partial t} = \sum_{i=1}^{n} \frac{\partial \psi}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial t} \tag{B.2}$$

Analogously, the continuum version is

$$\frac{\partial \psi}{\partial t} = \int_{r_0}^{r_1} \frac{\partial \psi}{\partial \sigma(r)} \frac{\partial \sigma(r)}{\partial t} dr \tag{B.3}$$

If the continuous depth profile in the material under test is written in terms of a sum of basis functions as $\sigma(r) = \sum_{i=1}^{m} b_i(\vec{k}_i, r), \quad s.t. \quad 0 \leq r \leq L$, where $\vec{k}_i$ are the coefficients of the i^{th} basis, then the sensitivity of the voltage to the k_i^j coefficient is denoted by

$$\begin{aligned} \frac{\partial \psi}{\partial k_i^j} &= \int_0^L \frac{\partial \psi}{\partial \sigma(r)} \frac{\partial b_i(\vec{k}_i, r)}{\partial k_i^j} dr \\ &= \int_0^L \int \vec{E}_{Tx} \cdot \vec{E}_{Rx} dS(r) \frac{\partial b_i(\vec{k}_i, r)}{\partial k_i^j} dr \end{aligned} \tag{B.4}$$

For example, if $b(r) = \{\sigma_c, -\sigma_d e^{-\sigma_s \cdot r}\}$, then

$$\begin{aligned} \frac{\partial \psi}{\partial \sigma_c} &= \int_0^r \frac{\partial \psi}{\partial \sigma(r)} \frac{\partial \sigma(r)}{\partial \sigma_c} dr \\ &= \int_0^L \int \vec{E}_{Tx} \cdot \vec{E}_{Rx} dS(r) dr \end{aligned} \tag{B.5}$$

Similarly

$$\begin{aligned} \frac{\partial \psi}{\partial \sigma_d} =& \frac{\partial \psi}{\partial \sigma(r)} \frac{\partial \sigma(r)}{\partial \sigma_d} \\ =& -\int_0^L \int \vec{E}_{Tx} \cdot \vec{E}_{Rx} dS(r) \cdot e^{-\sigma_s \cdot r} dr \end{aligned} \tag{B.6}$$

and

$$\begin{aligned} \frac{\partial \psi}{\partial \sigma_s} =& \frac{\partial \psi}{\sigma(r)} \frac{\partial \sigma(r)}{\partial \sigma_s} \\ =& \int_0^L \int \vec{E}_{Tx} \cdot \vec{E}_{Rx} dS(r) \cdot r\sigma_d e^{-\sigma_s \cdot r} dr \end{aligned} \tag{B.7}$$

The Jacobian is then

$$\begin{bmatrix} \frac{\partial \psi_1}{\partial \sigma_c} & \frac{\partial \psi_1}{\partial \sigma_d} & \frac{\partial \psi_1}{\partial \sigma_s} \\ \frac{\partial \psi_2}{\partial \sigma_c} & \frac{\partial \psi_2}{\partial \sigma_d} & \frac{\partial \psi_2}{\partial \sigma_s} \\ \vdots & \vdots & \vdots \\ \frac{\partial \psi_m}{\partial \sigma_c} & \frac{\partial \psi_m}{\partial \sigma_d} & \frac{\partial \psi_m}{\partial \sigma_s} \end{bmatrix}$$

where ψ_m is the m^{th} voltage measurement. This theory can be generalised further, with variations in all directions such that $\sigma(x, y, r)$. This would be an exciting area to develop further.