

**Advanced potential field data inversion with ℓ_p -norm
regularization**

by

Dominique Fournier

MSc, The University of British Columbia, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL
STUDIES
(Geophysics)

The University of British Columbia
(Vancouver)

September 2019

© Dominique Fournier, 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled

Advanced potential field data inversion with ℓ_p -norm regularization

submitted by Dominique Fournier in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

in Geophysics

Examining Committee:

Douglas W. Oldenburg, Earth, Ocean and Atmospheric Sciences

Supervisor

Eldad Haber, Earth, Ocean and Atmospheric Sciences

Supervisory Committee Member

Phil Austin, Earth, Ocean and Atmospheric Sciences

University Examiner

Ian Mitchell, Computer Science

University Examiner

Additional Supervisory Committee Member:

Craig Hart, Earth, Ocean and Atmospheric Sciences

Supervisory Committee Member

Abstract

The study of gravity and magnetism has a long history in Earth sciences and continues to play an important role in exploration geophysics. Fundamental knowledge about our planet has been gained from the processing of potential field data at all scales: from large tectonic processes, down to the size of mineral deposits. Attempting to model density and magnetization through the inverse process remains challenging however as it is a highly non-unique problem. Conventional unconstrained inversions generally yield either smooth models with poor edge definition or sparse and compact models that are too simplistic to represent subtle features.

In this doctoral dissertation, I provide three technical innovations to enhance the capabilities of potential data inversion. First, I propose to explore a wider range of solutions by independently varying the sparsity assumption imposed on the model values and its gradients. At the core of this research is the use of mixed ℓ_p -norm regularization solved by a scaled Iterative Re-weighted Least Squares algorithm. I provide a path to extract dominant features from the solution space with a Principal Component Analysis and pattern recognition strategy.

Secondly, I provide modification to the magnetic vector inversion in spherical coordinates to process magnetic field data affected by remanent magnetization. I tackle long-lasting issues related to the non-linear trigonometric coordinate transformation with an iterative sensitivity re-weighting strategy. The spherical formulation allows me to impose sparsity constraints on the amplitude and orientation of the magnetization vector independently. The algorithm can recover anomalies with a coherent magnetization direction.

Lastly, I incorporate soft geological information in the inversion through the rotation of the objective function. Combined with sparsity assumptions, the rotated

regularization improves the imaging of orientated geological contacts. Rotation angles are either interpolated from structural measurements or inferred from the data through a learning process. I demonstrate the benefits of structural constraints on gravity and magnetic datasets acquired over the Kevitsa Ni-Cu-PGE deposit, Finland. The recovered density model is compared to a seismic reflection profile for validation and to complement the geological interpretation of the deposit. Accurate modelling of the magnetization vectors yields insights about past tectonic deformations.

Lay Summary

This research focuses on geophysical imaging of Earth's crust based on the density and magnetic properties of rocks. This is done through an inversion process whereby gravity and magnetic field data collected above the surface are converted to a 3D model representing the sub-surface. Due to the nature of the inverse problem, there are infinite solutions that can satisfy the observed data. Conventional inversion algorithms yield smooth models that are difficult to interpret within a geological framework. I tackle these limitations with a flexible algorithm that allows me to generate a suite of possible candidates with variable characteristics. The algorithm is implemented for both scalar density and magnetization vector models. I then provide a learning strategy to extract dominant features from this suite of solutions using Principal Component Analysis. Finally, the efficacy of my approach is demonstrated using gravity and magnetic data acquired over the Kevitsa nickel deposit in Finland.

Preface

The research presented in this thesis was completed by myself, under the guidance of my supervisor Professor Oldenburg. The algorithms and ideas brought forward resulted in multiple articles, either published or in revision.

The mixed norm algorithm presented in Chapter 3 and part of the learning algorithm presented in Chapter 6 resulted in the article “*Inversion using spatially variable mixed ℓ_p -norm*” published in *Geophysical Journal International* (Fournier and Oldenburg, 2019b). I am the primary author of this article with revisions from Dr. Oldenburg. The content of the article had previously been presented at a conference (Fournier et al., 2016). The same algorithm has been used by other researchers and resulted in three research papers to which I am also co-author (Abedi et al., 2018a,b; Miller et al., 2017).

Improvements to the magnetic vector inversion in Chapter 4 will be featured in the accepted paper “*Sparse magnetic vector inversion in spherical coordinates: Application to the Kevitsa Ni-Cu-PGE magnetic anomaly, Finland*” submitted to *Geophysics* (Fournier and Oldenburg, 2019b). The article was written by myself under the guidance of Dr. Oldenburg. I have also co-authored a research paper investigating the geothermal resources at Mount Baker (Schermerhorn et al., 2017). Using the same methodology, I collaborated in the implementation of sparse vector inversion applied to self-potential problems to map hydrothermal circulation at Mount Tongariro, New Zealand (Miller et al., 2018).

The material presented in Chapter 5 and part of Chapter 7 is currently in preparation for publication. The article entitled “*Sparse rotated objective function for stratigraphic constraints: Application to the Kevitsa Ni-Cu gravity anomaly, Finland*” will be submitted within a few weeks.

All programming work done in this thesis builds upon the open-source SimPEG library as well as multiple packages from the Python ecosystem. Accreditation to open-source algorithms have been made wherever necessary.

Table of Contents

Abstract	iii
Lay Summary	v
Preface	vi
Table of Contents	viii
List of Tables	xi
List of Figures	xii
Acknowledgments	xxiv
1 Introduction	1
1.1 Motivation	1
1.1.1 Geophysical processing for geological interpretation . . .	3
1.1.2 Constraining the inversion	7
1.1.3 Exploring the solution space	8
1.1.4 Dealing with remanence	9
1.2 Research objectives	10
1.3 Thesis arrangement	11
2 Forward modeling of potential field data	13
2.1 Discrete systems	13
2.1.1 Choice of discretization	17

2.2	Large scale problems	21
2.2.1	Numerical test	23
3	Inverse Problem	27
3.0.1	Synthetic gravity example	31
3.1	General ℓ_p -norm regularization	34
3.1.1	Synthetic 1D problem	35
3.1.2	Iterative Re-weighted Least Squares algorithm	39
3.1.3	Case 1: ℓ_1 -norm ($p_s = p_x = 1$)	42
3.1.4	Case 2: ℓ_0 -norm ($p_s = p_x = 0$)	46
3.2	Mixed norm regularization	47
3.2.1	Scaled-IRLS steps	48
3.2.2	Threshold ϵ -parameter	55
3.2.3	Summary	60
3.3	Exploring the model space	61
3.3.1	Interpretation	62
4	Sparse Magnetic Vector Inversion	66
4.1	Magnetic Vector Inversion - Cartesian parameters	70
4.2	Magnetic Vector Inversion - Spherical parameters	71
4.2.1	Scaled MVI-S algorithm	81
4.3	Synthesis	82
5	Rotated ℓ_p-norm Regularization	84
5.1	Rotated objective function	85
5.1.1	Diagonal derivatives	92
5.2	Synthetic fold model	94
5.2.1	ℓ_p -norm inversion	94
5.2.2	Directional ℓ_p -norms	96
5.3	Summary	98
6	Automated mixed-norm modeling	100
6.1	Selecting local p -parameters	101
6.1.1	Model space inversions	101

6.1.2	Average PCA model	104
6.1.3	Parameter extraction	106
6.1.4	Summary	108
6.2	Dip and strike estimation	109
6.3	Summary	116
7	Case Study - Kevitsa Ni-Cu-PGE deposit	117
7.0.1	Geological setting	117
7.1	Geophysical data	120
7.1.1	Modeling objectives	124
7.2	Gravity inversion	124
7.3	Magnetic inversion	127
7.4	Summary	129
8	Conclusion	141
8.1	Limitations and future work	143
	Bibliography	145
	Appendix	154

List of Tables

Table 2.1	Summary table for the forward modeling of a magnetized sphere using a Tensor and Octree discretization	21
Table 3.1	Norm values and proportionality ratio obtained for the 1D solution presented in Figure 3.5(c). A proportionality ratio of $\lambda_\infty \approx 1$ indicates that the components of the regularization function are both contributing significantly to the final solution.	39
Table 3.2	IRLS algorithm in pseudo-code made of two stages: Stage 1 Initialization with convex least-squares inversion, Stage 2 IRLS updates with inner β -search steps.	43
Table 3.3	Components of the objective function corresponding to the inversion results presented in Figure 3.12 for $p_s = p_x = 1$	55
Table 3.4	Inversion summary after convergence of the S-IRLS for various cooling rates η as presented in Figure 3.16. Each inversion trial was required to reach the target misfit ϕ_d^* and target $\varepsilon^* = 1e - 6$.	60
Table 7.1	Summary table grouping the various lithological units logged from bore hole. Expected density and magnetic susceptibility contrasts are derived from Figure 7.2.	121
Table 7.2	Intervals along boreholes KV200 and KV297 reporting significant remanent magnetization (Montonen, 2012)	124

List of Figures

Figure 1.1	Cartoon illustrating the (left) gravity and (right) magnetic fields originating from dense and magnetized anomalies. Adapted from (Blakely, 1996).	2
Figure 1.2	Total magnetic map over northern Europe extracted from the global Earth Magnetic Anomaly Grid (2-arc-minute resolution) compiled from satellite, ship and airborne measurements (NCEI, 2017). The case study used in this research, the Kevitsa Ni-Cu-PGE deposit in Finland, sits within a region of strong magnetization. The image overlay in GoogleEarth was generated with the open-sourced geosci.GeoToolkit package.	4
Figure 1.3	(a) Surface geology, (b) Topography (c) ground gravity and (d) airborne magnetic data acquired over the Kevitsa deposit. . . .	5
Figure 1.4	Figure comparing three approaches used to model the Kevitsa deposit. (a) Geological interpretation derived from seismic and borehole data (borrowed from Koivisto et al. (2015)). (b) Vertical section through an unconstrained density inversion using conventional smooth assumptions. (c) Geological surface model guided by geophysical forward modelling using Gocad.	6
Figure 2.1	Parameters describing the spatial relationship between an observation point P and a rectangular prism C	14

Figure 2.2	Simple representation of (a) tensor and (b) Octree meshes for the organization of rectangular prisms within a core domain (red) and padding region over a square domain 4 x 4 m in width. The Tensor mesh can fill the space with 64 cells, compared to only 40 cells with the Octree mesh.	17
Figure 2.3	Three refinement strategies used to discretize a Gaussian curve defined by scattered points (red): (a) box, (b) radial and (c) surface refinement strategy from the SimPEG.discretize library. Cells are coloured by their corresponding Octree level.	19
Figure 2.4	(a) Discretization of a sphere defined by discrete points (red) using the conventional Tensor mesh with a core region and padding cells. Octree meshes refined by (b) box, (c) radial and (d) surface methods from the SimPEG.discretize library.	20
Figure 2.5	(a) Horizontal and (b) vertical sections through the synthetic density model used to test the mesh decoupling strategy. (c) The simulated data contain short and long wavelength information.	24
Figure 2.6	Example of mesh decoupling of the forward problem into nested Octree tiles. (Middle) Local meshes are generated for a pre-defined number of tiles nested inside the global domain. (Bottom) The forward modelling operator for each tile is stored in compressed Zarr file format on solid-state memory. Memory chunks (grey) are accessed in parallel by the Dask library to perform the forward calculations.	25
Figure 2.7	Trade-off curves between the size of the forward problem (blue) and data residual (red) over a range of number of tiles. The total size of the problem is calculated based on the sum of cells in all the local meshes times the number of data. The optimal number of tiles would be at the point of intersection where both the cost of forward calculations and the data residual change significantly.	26

Figure 2.8	Simulated gravity data calculated from (a) the global model and (b) the 12 forward tiled calculations. (c) Data residuals show short wavelength discrepancies between adjacent tiles due to interpolation effects.	26
Figure 3.1	(a) Vertical section through a 25 m cube with density $\rho=0.2$ g/cc placed in a uniform zero density background. (b) Simulated gravity data responses on a 21×21 survey grid placed 5 m above the flat topography. (c) Gravity data with random Gaussian noise added, 10^{-3} mGal standard deviation.	32
Figure 3.2	(a) Vertical section through the inverted density model using the conventional ℓ_2 -norm regularization, (b) predicted and (c) normalized data residual. Outline of the true model (red) is shown for reference.	33
Figure 3.3	Approximated ℓ_p -norm using the Lawson measure (Lawson, 1961) over a range of p -values and for a fixed threshold parameter $\varepsilon = 10^{-1}$	35
Figure 3.4	Linear forward problem made up of: (a) an example kernel function ; (b) model; (c) observed data with assigned standard errors.	36
Figure 3.5	Solution to the 1D inverse problem using (a) an ℓ_2 -norm on the model ($\alpha_x = 0$), (b) the ℓ_2 -norm on model gradients ($\alpha_s = 0$) and (c) combined regularization function ($\alpha_s = 2500$, $\alpha_x = 1$). (d) Convergence curve comparing the misfit (ϕ_d) and the regularization (ϕ_m) as a function of iterations. (e) Comparative plot for the relative contribution of the different components of the objective function measured in terms of maximum absolute gradient ($\ \mathbf{g}_i\ _\infty$)	37

Figure 3.6	(a) Two solutions using an ℓ_1 -norm on the model: (blue) Simplex, and (black) IRLS method. (b) Solution obtained with the approximated ℓ_1 -norm (IRLS) penalty on model gradients alone and (c) with the combined penalty functions ($\alpha_s = 2500$, $\alpha_x = 1$). The calculated proportionality ratio λ_∞ indicates that the combined penalties is dominated by the ϕ_s^1 term. (d) Convergence curve and (e) maximum partial derivatives associated with the components of the objective function as a function of iterations for the inversion in (c). The vertical dotted lines indicate the change in regularization from an ℓ_2 -norm to ℓ_1 -norm measure.	44
Figure 3.7	(a) Solution obtained with the combined penalty functions $\alpha_s \phi_s^1 + \alpha_x \phi_x^1$ after re-adjustment of $\alpha_s = 50$, $\alpha_x = 1$. (b) Convergence curve and (c) maximum partial derivatives associated with the components of the objective function as a function of iteration.	46
Figure 3.8	Solution to the 1D inverse problem using an approximate ℓ_0 -norm (a) on the model, (b) on model gradients and (c) combined penalty functions using the IRLS algorithm ($\alpha_s = 1$, $\alpha_x = 1$). All three solutions honor the data within the target misfit ϕ_d^*	47
Figure 3.9	(a) Recovered model and (b) convergence curves using the conventional IRLS method for $p_s = 0$, $p_x = 2$ and a fixed threshold parameter $\varepsilon = 10^{-3}$ ($\alpha_s = \alpha_x = 1$). (c) Trade-off parameter and maximum gradients for the different components of the objective function. At the start of Stage 2 (iteration 6), the sudden increase in $\ \mathbf{g}_s\ _\infty$ is matched with a decrease in β . Throughout the inversion, $\ \mathbf{g}_x\ _\infty$ remains small in magnitude.	49
Figure 3.10	Derivatives of the Lawson approximation over a range of model values for (a) a fixed threshold parameter $\varepsilon = 10^{-1}$ over a range of p values and for (b) a fixed $p = 0$ over a range of ε values. (c) Applying the γ -scaling to the gradients brings all maximums to be equal irrespective of p and ε	50

Figure 3.11	(a) Recovered model and (b) convergence curves using the Scaled-IRLS approach for $p_s = 0$, $p_x = 2$ and a fixed threshold parameter $\varepsilon = 1e-3$ ($\alpha_s = \alpha_x = 1$). (c) Trade-off parameter and maximum gradients for the different components of the objective function. The scaling procedures preserves the proportionality between $\ \mathbf{g}_s^{p_s}\ _\infty$ and $\ \mathbf{g}_x^{p_x}\ _\infty$ throughout the iteration process. The trade-off β -parameter needed only to be adjusted slightly at the beginning of Stage 2.	52
Figure 3.12	Recovered 1D models for $p_s = p_x = 1$ using different scaling strategies and parameterization. (a) Solution previously shown in Figure 3.7 that uses the standard gradient measure ($\alpha_s=50$). (b) Solution obtained with the finite difference approach ($\alpha_s = \alpha_x = 1$). (c) Model recovered with a different parameterization such that the right half of the domain has cells with half the size. Small discrepancies between the three solutions can be attributed to slight differences in the iterative process.	55
Figure 3.13	Recovered 1D models with variable threshold parameter on the range $10^{-5} < \varepsilon < 10^0$ using a mixed-norm penalty function $\phi_m = \alpha_s \phi_s^0 + \alpha_x \phi_x^2$	57
Figure 3.14	(a)-(d) Recovered 1D models at different iteration steps (k). The value of ε (dash) is shown for reference, highlighting the idea of a progressive thresholding of model values. (e) Scaled partial derivatives ($\gamma^2 \mathbf{g}_s^0$) as a function of the sorted model values and at various iteration stages.	58
Figure 3.15	Partial derivatives of the objective function for iteration (top) $k=15$ and (bottom) $k=55$. The recovered models are shown in red for reference.	59
Figure 3.16	(a-d) Recovered models and (e) convergence curves for the minimization of $\phi_d + \beta \phi_m^P$ with various cooling rates η but with a final $\varepsilon^* = 1e-6$	59
Figure 3.17	(a-i) Vertical section through a suite of density models recovered for varying ℓ_p -norm penalties applied on the model and model gradients for $p_s \in [0, 1, 2]$ and $p_x = p_y = p_z \in [0, 1, 2]$	63

Figure 3.18	Iso-contour values for the 10 th and 90 th percentile of anomalous density calculated from the suite of models shown in Figure 3.17. The outline of the target (red) is shown for reference. Contour lines tightly clustered indicate coherence between inversion trials. Negative anomalies (dash) appear to change significantly, to which I would assign lower confidence.	64
Figure 3.19	(a-i) Residual data map calculated from the suite of density models for varying ℓ_p -norm penalties applied on the model and model gradients for $p_s \in [0, 1, 2]$ and $p_x = p_y = p_z \in [0, 1, 2]$	65
Figure 4.1	(a) Vertical section through a 25 m cube with uniform magnetization \vec{M} [2.0 A/m, I: 45°, D: 90°]. (b) Simulated TMA data response on a 21 × 21 survey grid placed 15 m above the anomaly. (c) Magnetic data with random Gaussian noise added, 1 nT standard deviation.	68
Figure 4.2	Vertical section through the recovered susceptibility model using (a) smooth assumption ($p_s, p_x, p_y, p_z = 2$) and (b) sparse ℓ_p -norms to recover a compact model ($p_s, p_x, p_y, p_z = 0$). Both solutions show an anomaly with a false dip due to the wrong assumption of a vertical magnetization.	69
Figure 4.3	Vertical section through the recovered magnetization vector model using the Cartesian formulation with (a) smooth l_2 -norm assumption and (b) sparsity constraints applied on all three Cartesian components($p_{is}, p_{ix}, p_{iy}, p_{iz} = 0$). Corresponding normalized data residuals are shown in (b) and (d) . The true position and magnetization orientation of the block are shown in red for reference.	72
Figure 4.4	Measure of angle differences $\Delta\theta$ converted to cotermin angle $\Delta\hat{\theta}$	74

Figure 4.5	Vertical section through the (a) starting model and (b) recovered magnetization vector model in Spherical coordinates. The true position and magnetization orientation of the block are shown in red for reference. (c) Normalized data residuals show correlated signal. The inversion stopped after 15 iterations and was enabled to further reduce the objective function.	75
Figure 4.6	Contour map for two objective functions and their gradients (arrows) for a two-parameter inverse problem solved in Cartesian and polar coordinate systems. (a) The non-weighted (gray) problem yields a solution $\mathbf{m}_C = [0.2, 0.4]$ that reflects the size of the forward coefficients. The sensitivity weighted (black) solution is more desirable as it predicts the data with equal model parameters $\mathbf{m}_C^* = [0.33, 0.33]$. (b) Inversion steps performed in the non-linear polar coordinate system (solid red) are highly oscillatory and do not reach the optimal solution $\mathbf{m}_P^* = [0.47, 0.79]$. The same model updates are shown in Cartesian coordinates (red dash) for reference. (green) Inversion steps performed in polar coordinates with iterative sensitivity re-weighting and (blue) with the added Ω -scaling.	77
Figure 4.7	(a) Vertical section through the recovered magnetization vector model and (b) normalized data residual using the spherical formulation with sensitivity based weighting ($p_s = p_x = p_y = p_z = 2$). The true position and magnetization orientation of the block are shown in red for reference.	82
Figure 4.8	(a) Vertical section through the recovered magnetization vector model in Spherical coordinates using sensitivity based weighting with smooth regularization ($p_s = p_x = p_y = p_z = 2$). MVI-C solution is used as a starting model. The true position and orientation of magnetization of the block are shown in red for reference. (c) Recovered vector model in Spherical coordinates using sparse and blocky assumptions on the amplitude and angles of magnetization ($p_{cs} = p_{cx} = p_{cy} = p_{cz} = 0$). (b) and (d) Corresponding normalized data residuals.	83

Figure 5.1	Rotation parameters with respect to the Cartesian axes.	87
Figure 5.2	Two-dimensional representation of finite difference operators using a combination of (a) 3-cell (Li and Oldenburg, 2000), (b) 5-cell (Lelièvre, 2009) and (c) the 7-cell scheme used for the rotation of the objective function. Only the 7-cell strategy allows for the interaction between diagonal neighbors.	90
Figure 5.3	Horizontal cross section through the recovered model with rotation of the objective function using (top) 5-point and (bottom) 7-point gradient operators. Blue and red arrows indicate the rotation frame. (a, d) Recovered models with smooth assumptions ($p_s = p_x = p_y = p_z = 2$) stretched along the rotate x-axis at 45° ($\alpha_{x'} = 100$). (b, e) Sparse solutions ($p_s = p_x = p_y = p_z = 0$) with rotation at 45° . The target orientation of the rotated block is marked by the black dashed line. (c, f) Sparse solutions for a 30° rotation of the objective function.	91
Figure 5.4	Normalized data residuals calculated from the recovered models presented in Figure 5.3(d), (e) and (f) respectively.	93
Figure 5.5	Synthetic density model made up of a folded layer with a N-S oriented axial plane dipping 75° E. The intersection between the fold axial plane and the cross-sections are marked by a red-dashed line: (left) topography draped section and (right) vertical cross-sections along (A-A') and (B-B'). The hinge axis of the fold is plunging 30° S. Two strike and dip measurements are provided at the surface and used as structural constraints.	95
Figure 5.6	Synthetic gravity data generated 1 m above topography over the synthetic density fold model. Gravity station locations (dot) and topographic contours (lines) are shown for reference. Structural data (dip, strike) are provided at two locations on opposite of the syncline. Dip direction of the fold axis (15° S) and limbs are shown with arrows	96

Figure 5.7	Recovered smooth solution with ℓ_2 -norm regularization of the synthetic fold anomaly. The dip of the fold limbs are poorly recovered and there is no indication of the limbs being connected at depth.	97
Figure 5.8	Recovered density model from the mixed norm regularization ($p_s = 0$, $p_{x,y,z} = 2$). Sparse assumptions helped in simplifying the density model over the smooth assumption.	97
Figure 5.9	Interpolated dip and strike information used for the rotation of the objective function. The direction of the arrows define the estimated normal or the folded layer.	98
Figure 5.10	Recovered density model from the mixed-norm regularization. Data locations (dots), topography and the outlines of the true model (black) are shown for reference. Rotated ℓ_p -norm oriented along the folded layer helped in resolving a continuous geological unit at depth.	99
Figure 6.1	(a) Synthetic 2D travel-time tomography problem made up of a rectangular velocity low, and a smooth Gaussian velocity high. Contour lines (25^{th} and 75^{th} percentile) are shown in black for the true position and shape of the velocity anomalies. Transmitters (red) and receivers (green) are positioned on opposite sides of the model domain. (b) Data map for the 143 line integral data calculated for each transmitter-receiver pair.	102
Figure 6.2	Suite of inverted models for various combinations of norms p_s , $p_x = p_y \in [0, 1, 2]$ ($\alpha_s = \alpha_x = \alpha_y = 1$). Contour lines (25^{th} and 75^{th} percentile) are shown in black for the true position and shape of the velocity anomalies.	103
Figure 6.3	(a) to (i) Data residual contour map for the nine inversions presented in Figure 6.2. Contour lines (± 1 standard deviation) are shown in black. (j) Random noise added to the experiment.	104
Figure 6.4	PCA vectors covering 75% of the model variances.	105

Figure 6.5	(a) Averaged PCA model \mathbf{m}_A and (b) corresponding normalized data residual map. Contour lines are shown in (a) for the true position and shape of the velocity anomalies.	106
Figure 6.6	Parametric representation of the nine inverted models derived from the Canny edge detection algorithm.	107
Figure 6.7	Local p -values on the (a) the model norm $\phi_s^{P_s}$ and (b) model gradients $\phi_x^{P_x}, \phi_y^{P_y}$ extracted from the solution space.	107
Figure 6.8	(a) The final SVMN inversion result and (b) normalized data residual map. Contour lines are shown in (a) for the true position and shape of the velocity anomalies.	109
Figure 6.9	Synthetic fold model introduced in Chapter 5 used here to test the dip and strike estimation learning algorithm.	110
Figure 6.10	Example of an automated azimuth estimation using the UBC Crest. (b) The image is first converted to a binary image using the python routine <code>ScikitImage.Skeleton</code> . Image moment calculations are performed on moving windows to determine the position and orientation of dominant features. (c) Vectors are then rotated by 90° to point along the normal of edges. The normal vectors are checked for consistency in preparation for interpolation.	112
Figure 6.11	(a) Image moment calculations over windowed regions of the inverted density model. Arrows are plotted for the principal direction (red) and normal (white) of the main density anomaly. (b) Normal vectors are interpolated in 3D and used to rotate the objective function.	113
Figure 6.12	Suite of models with directional gradients for $\theta \in [-90^\circ, 75^\circ]$	114
Figure 6.13	Sections through the average PCA model and rotation vectors (arrow) derived from the image moment analysis.	115
Figure 6.14	(a) Horizontal and (b, c) vertical sections through the recovered model using rotation angles derived from the average PCA model.	115

Figure 7.1	(a) Geological surface map of the Kevitsa-Satovaara intrusive complex and (b) 2D seismic reflection survey along with the E5 profile. Interpreted seismic reflectors and geological contact (black) from the interpretation of Koivisto et al. (2015) are shown for reference, as well as the outline of the gravity and magnetic data sets analyzed in this chapter.	118
Figure 7.2	Whisker plot of logged (a) density and (b) magnetic susceptibility measured along 279 boreholes. The coloured boxes have a width scaled by the calculated standard deviation and centered on the mean value for all intercepts belonging to the same lithological classification. The black lines on either side define the minimum and maximum values. The different lithologies are colour coded and grouped based on relative age and similarities in physical properties.	122
Figure 7.3	(a) Ground gravity and (b) airborne magnetic data map acquired over the Kevitsa Ni-Cu-PGE deposit, Finland. Color ranges are histogram equalized and sun shaded to highlight details. The location of the Kevitsa mine (red), geological boundaries (black) and faults (dash) identified from surface mapping are shown for reference.	123
Figure 7.4	(a) Horizontal section at ≈ 200 m depth below topography through the recovered density model after convergence of the algorithm for $p_s = p_x = p_y = p_z = 2$. (b) Vertical section of the density model overlaid on the 2D seismic reflection line E5. Interpreted geological contacts (black) are shown for comparison.	126
Figure 7.5	Rotation parameters (normals) derived from (a) surface geology and (b) major reflectors detected on the seismic profile E5.	131
Figure 7.6	Horizontal sections at ≈ 200 m depth below topography through the density models recovered over a range of ℓ_p -norms applied on the model and modeled derivatives for $p_s, p_{x,y,z} \in \{0, 1, 2\}$	132

Figure 7.7	Vertical sections overlaid on the 2D seismic reflection line E5 through the density models recovered over a range of ℓ_p -norms applied on the model and modeled derivatives for $p_s, p_{x,y,z} \in \{0, 1, 2\}$	133
Figure 7.8	Residual data maps for the nine gravity inversions.	134
Figure 7.9	(a) Horizontal and (b) vertical section comparing iso-surfaces for the 5 th and 95 th percentile of anomalous density values recovered from the nine mixed norm inversions for $p_s, p_{x,y,z} \in \{0, 1, 2\}$.	135
Figure 7.10	(a) Horizontal and b) vertical sections through the recovered susceptibility model that ignores the effect of remanence. Lithological contacts (black) identified by Koivisto et al. (2015) are shown for reference. A large dome-shaped zero-susceptibility anomaly is recovered at the center of the Kevitsa intrusion, likely due to remanence. (c) Residual data map shows strong correlation with the negative anomalies.	136
Figure 7.11	Horizontal sections at ≈ 300 m below topography for a suite models using various sparsity assumptions put on the amplitude of magnetization for $p_s, p_{x,y,z} \in [0, : 2]$. Norm measures on the magnetization angle are fixed to $p_{x,y,z} = 0$ in order to promote uniform magnetization	137
Figure 7.12	Vertical sections through the recovered magnetization models. Effective susceptibilities are concentrated within the UDU and on the lower margin of the UPXO unit.	138
Figure 7.13	Residual magnetic data maps for the nine MVI-S inversions.	139
Figure 7.14	(Top) Horizontal and (bottom) vertical sections through iso-contours of magnetization recovered from nine mixed ℓ_p -norm inversions. Magnetization orientation (white) and standard deviation on the angle (red) are shown.	140

Acknowledgments

I would like to first thank my supervisor, Professor Doug Oldenburg, for giving me the opportunity to join the UBC-GIF group and for his tireless excitement for research. Your passion for science is contagious.

I want to acknowledge the outstanding contribution of many colleagues from the open-source community. Thanks to Rowan Cockett, Lindsey Heagy and Seogi Kang for establishing the SimPEG framework and pushing me to learn Python. Special mention goes to Joe Capriotti for his key update to the Octree mesh discretization. The work presented in this thesis would not have been possible without your contribution. Thank you to the rest of the UBC-GIF group for your support, kind words and friendship.

I am grateful to the companies and staff that have supported research at UBC-GIF over all these years: Anglo American, Barrick, BHP, Cameco, CGI, Glencore, Newmont, Rio Tinto, Teck, Vale. Personal thanks to John McGaughey at Mira Geosciences for keeping on staff over all these years, as well as for lending me valuable hardware and software.

Thank you to the many public libraries across Canada where this thesis was written. Cities and towns from east to west (*with great pool nearby): Ottawa, Winnipeg*, Prince George, Fort Saint-James, Smithers*, Vancouver, Richmond*, Steveston, Nanaimo, Courtney, Alert Bay, Port Hardy.

Last but not least, extra special thanks goes to my best friend, my partner in crimes and lover Chantelle Bellrichard. I am done being a student now. Thank you to *sister-wife* Candice Nichol for providing distractions to *first-wife*, allowing me to finish my thesis. So many more BBQ & sunsets at the beach to be enjoyed ...

Chapter 1

Introduction

1.1 Motivation

Density and magnetic permeability are important physical properties used in mineral exploration. Geologists routinely measure these properties on rock samples, along with textural and structural features, to target areas for further investigation. The vast majority of known mineral deposits have been discovered this way (Marjoribanks, 2010). But the conventional approach to mineral exploration is limited to shallow regions of the Earth with exposed bedrock. Large portions of the crust remain untouched, buried under thick sedimentary cover.

In the absence of surface evidence, density and magnetic permeability can still play a role in mineral exploration as a source of geophysical signal. Variations in the local gravity and magnetic fields are used to map geology and identify buried structures such as intrusions, faults, folds and alteration zones (Domzalski, 1966; Grant, 1984). The physical equations for gravity and magnetics are described with the aid of Figure 1.1. For gravity we have the scalar gravitational potential:

$$\phi_g = G \int_V \frac{1}{r} \rho(r) dV \quad (1.1)$$

The gravity field is obtained by taking the gradient of (1.1)

$$\vec{g} = \nabla \phi_g \quad (1.2)$$

which becomes

$$\vec{g} = -G \int_V \frac{\hat{r}}{r^2} \rho(r) dV . \quad (1.3)$$

Equation (1.3) describes the gravity field \vec{g} ($\frac{N}{kg}$) as observed at some arbitrary position P due to an elementary volume dV with density ρ ($\frac{kg}{m^3}$). The field scales as a function of radial distance r between the source and the observer, and by Newton's gravitational constant G ($6.674 \times 10^{11} \frac{Nm^2}{kg^2}$). For the magnetic problem, we

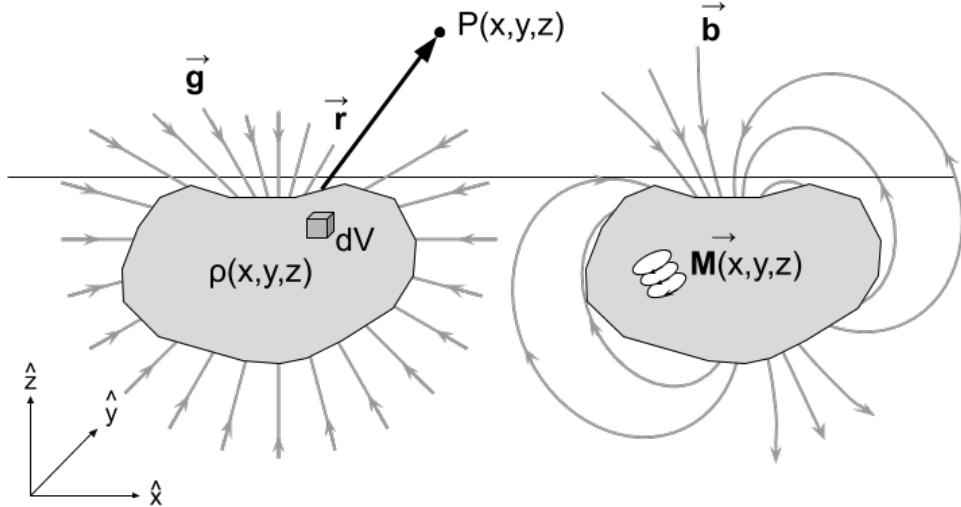


Figure 1.1: Cartoon illustrating the (left) gravity and (right) magnetic fields originating from dense and magnetized anomalies. Adapted from (Blakely, 1996).

follow the notation of Blakely (1996). We consider magnetic minerals as being composed of small dipoles, each with their own magnetic dipole moment. Preferential alignment of these dipoles gives rise to a magnetization per unit volume \vec{M} ($\frac{A}{m}$)

$$\vec{M} = \frac{\text{total magnetic moment}}{\Delta V} . \quad (1.4)$$

The scalar magnetic potential can than be defined as:

$$\phi_b(r) = \frac{\mu_0}{4\pi} \int_V \nabla \frac{1}{r} \cdot \vec{M}(r) dV , \quad (1.5)$$

Taking the gradient of (1.5) yields an expression for the magnetic flux density:

$$\vec{b} = \nabla \phi_b \quad (1.6)$$

and thus

$$\vec{b}(r) = \frac{\mu_0}{4\pi} \int_V \nabla \nabla \frac{1}{r} \cdot \vec{M}(r) dV. \quad (1.7)$$

In this above equation, \vec{b} is the magnetic flux density ($\frac{Wb}{m^2}$ or $\frac{N}{Am}$) and μ_0 is the magnetic permeability of free space ($\frac{H}{m}$ or $\frac{N}{A^2}$). Both equation (1.1) and (1.5) describe a source of geophysical signal that decays radially with distance and that scales linearly with changes in physical property contrasts. As a result, the gravity and magnetic experiments are collectively referred to as *potential field* methods, and there is a field of mathematics devoted to this subject.

A large number of potential field datasets has been compiled over the years and made available to the geoscientific community for analysis: from global satellite and marine measurements, to airborne and ground surveys (Figure 1.2). Processing this wealth of data has encouraged collaboration between geologists and geophysicists in their efforts to map the sub-surface. After nearly five decades of data acquisition and interpretation, gravity and magnetic methods are still an area of active research in geophysics (Nabighian et al., 2005). This Ph.D. project contributes to that body of research.

1.1.1 Geophysical processing for geological interpretation

While the study of potential field data represents a common ground for geologists and geophysicists to investigate the Earth, the two branches of geoscience appear to diverge on the path taken for interpretation. The majority of geological studies have focused on data processing techniques for mapping structures and geological domains in 2D. Fourier filters have been particularly popular, such as the total derivatives (Cordell and Grauch, 2012), analytic signal (Macleod et al., 1993; Roest et al., 1992), tilt angle (Miller and Singh, 1994) or combination of filters (Sanchez et al., 2014; Verduzco et al., 2004). Euler deconvolution has been used to estimate the depth of simple parametric shapes such as planes, dykes, and spherical bodies (Mushayandebvu et al., 2001; Reid et al., 1990; Thompson, 1982).

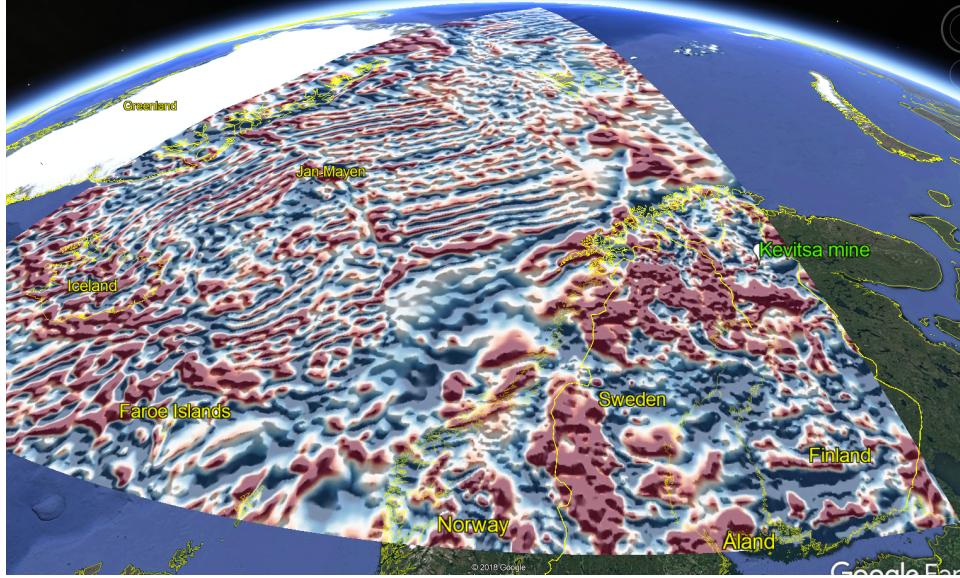


Figure 1.2: Total magnetic map over northern Europe extracted from the global Earth Magnetic Anomaly Grid (2-arc-minute resolution) compiled from satellite, ship and airborne measurements (NCEI, 2017). The case study used in this research, the Kevitsa Ni-Cu-PGE deposit in Finland, sits within a region of strong magnetization. The image overlay in GoogleEarth was generated with the open-sourced `geosci.GeoToolkit` package.

Similarly, the tilt-depth method (Salem et al., 2007) and Source Parameter Imaging (SPI) (Thurston and Smith, 1997) can provide an estimate for the depth and dip of elongated bodies in 2D (Phillips, 2010). A common limitation of these methods is in dealing with data acquired over complex geology. It is challenging to interpret overlapping signals from multiple sources with arbitrary shapes and positions in 3D. It is even more challenging to make sense visually of dipolar magnetic anomalies, especially at low latitudes.

The Kevitsa Ni-Cu-PGE deposit shown in Figure 1.3 is an interesting case study to illustrate the difficulty of interpreting potential field data in a mineral exploration context. The gravity and magnetic maps are both visually challenging to interpret. Attempting to construct a 3D geological model of Kevitsa directly from the data maps would undoubtedly be a difficult task. Yet there is a lot of information that could be extracted to help unravel the geological history of the region.

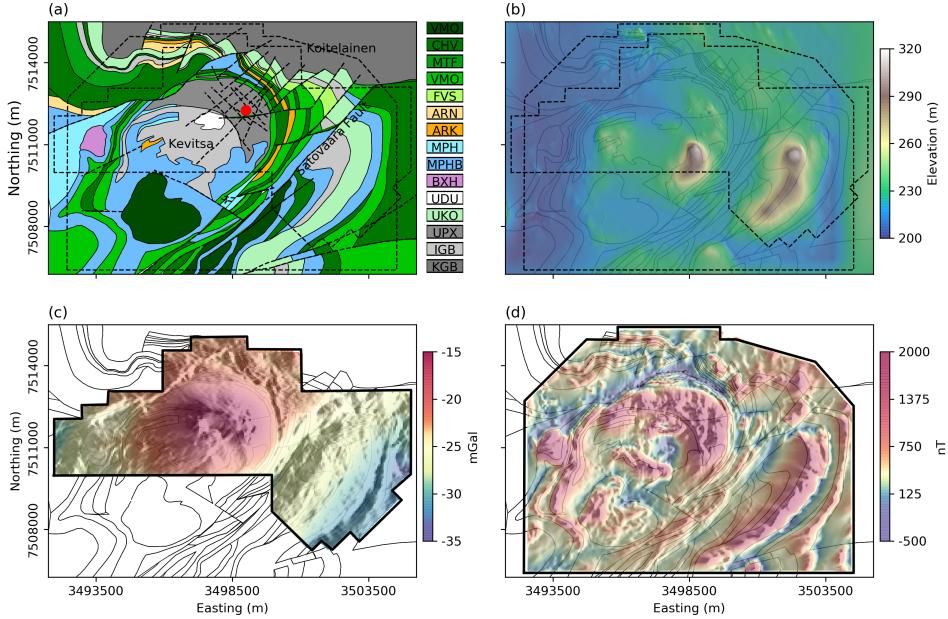


Figure 1.3: (a) Surface geology, (b) Topography (c) ground gravity and (d) airborne magnetic data acquired over the Kevitsa deposit.

To this end, a large body of work from the geophysical community has focused on inversion methodologies to model physical property contrasts in 3D. Among these methods, voxel-based inversions have received considerable attention. One of the main challenges posed by the inverse problem is that many models can fit the observed data; this is referred to as non-uniqueness. Additional information must be provided to isolate geologically sensible solutions. The typical strategy is to assume smooth changes in physical properties. This strategy has become an industry standard that is adopted by several commercial codes such as the UBC-GRAV3D and MAG3D, (Li and Oldenburg, 1996; Li. and Oldenburg, 1998).

The conventional smooth physical property inversion framework is somewhat at odds with the geological mindset however which is geared towards defining discrete geological domains with sharp boundaries. This divide can once again be exemplified with sections through the Kevitsa deposit (Figure 1.4). After years of exploration work in the area, geologists have identified and mapped over a dozen

lithologies along 2D sections. In an ideal case, we would like to use geophysical inversion to complete the modelling work in 3D. Attempting to do so with current methodologies yields the model presented in Figure 1.4(b). The smooth density anomaly is too difficult to interpret in this context as it shows poor correlation with the conceptual model of Kevitsa. It is partially for this reason that the geological community has been slow to adopt 3D inversion methods as part of their interpretation workflow. The apparent disconnect between the two fields of geoscience remains a key issue to be addressed (Li et al., 2019).

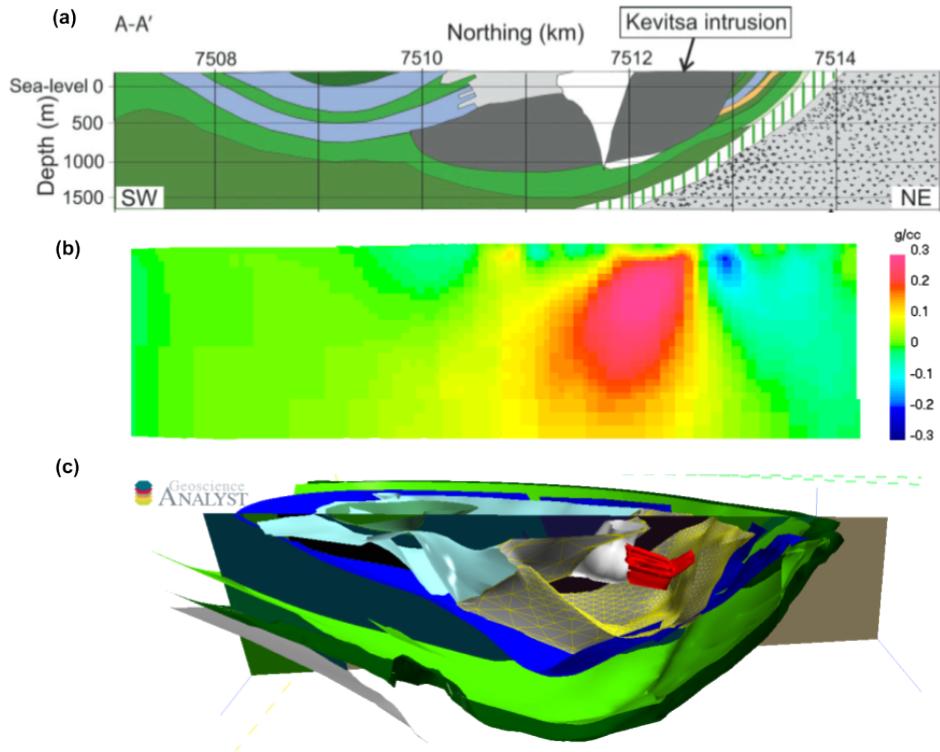


Figure 1.4: Figure comparing three approaches used to model the Kevitsa deposit. (a) Geological interpretation derived from seismic and borehole data (borrowed from Koivisto et al. (2015)). (b) Vertical section through an unconstrained density inversion using conventional smooth assumptions. (c) Geological surface model guided by geophysical forward modelling using Gocad.

1.1.2 Constraining the inversion

A number of inversion strategies have been proposed in the past to help bridge the gap between geologists and geophysicists. For simple scenarios involving a few isolated targets, geophysical anomalies may be approximated by simple parametric bodies such as ellipsoids (McMillan et al., 2016) and dipping dykes (Foss, 2006; Krahenbuhl and Li, 2015). For most complex geological settings, full 3D surface modelling can be used to constrain the inversion (Bosch and McGaughey, 2001; Fullagar et al., 2008). Figure 1.4(c) demonstrates such modelling efforts on the Kevitsa intrusion. Building upon the seismic interpretation of Koivisto et al. (2015), borehole logs and surface mapping, I constructed a 3D surface model using the Gocad-SKUA package. Generating and testing this model against the geophysical data is a laborious process that requires skills, advanced software and that can rapidly become intractable.

On a more data-driven side, hard physical property constraints have been included in the inversion on a cell-by-cell basis (Phillips, 1996; Williams, 2008) or through clustering algorithms (Lin and Zhdanov, 2017; Sun and Li, 2013). Dip and structural information have also been used to reinforce known trends; this has been accomplished by rotating the regularization functions with manual adjustment of model gradients (Davis et al., 2012; Li and Oldenburg, 2000). A common limitation with these approaches is in building and testing different geological scenarios. Time and computational limitations often force practitioners to arbitrarily choose one set of parameters for the regularization function and accept the outcome as a “best” model and have it used in subsequent interpretation.

Few studies in geophysics have attempted to address non-uniqueness within a learning framework. The genetic algorithm of Wijns and Kowalczyk (2007) and supervised learning of Haber and Tenorio (2003) can guide practitioners in the choice of inversion parameters and they could serve as a starting point for advanced 3D modelling work. But the efficacy of these approaches is once again impeded by assuming a smooth physical property distribution. The learning process is limited by the ability to generate solutions that are significantly diverse.

1.1.3 Exploring the solution space

As an alternative to smooth inversions, several studies have employed approximations to ℓ_p -norms in order to recover compact anomalies. Methods such as the Ekblom norm ($p = 1$) (Ekblom, 1973) and the Lawson approximation (Lawson, 1961) ($p = 0$) favor impulsive models with fewer non-zero parameters. Several papers have demonstrated the efficacy of a sparsity assumption in reducing the complexity of physical property models (Ajo-Franklin et al., 2007; Barbosa and Silva, 1994; Blaschek et al., 2008; Chartrand, 2007; Last and Kubik, 1983; Portniaguine and Zhdanov, 2002; Stocco et al., 2009). Likewise, a sparse solution for model gradients can yield blocky (piece-wise constant) models (Daubechies et al., 2010; Farquharson and Oldenburg, 1998; Gorodnitsky and Rao, 1997; Li, 1993; Sun and Li, 2014).

Sparsity assumptions have increased the flexibility of inversion algorithms for the modelling of compact targets with sharp edges. The usual strategy is to define global parameters applied over the entire model space. In most geological settings, however, we can expect to encounter a mix of features, either smooth and sparse in one region but elongated and blocky elsewhere. Determining exactly which assumption to use in a particular geological setting remains largely user-driven. Sun and Li (2014) have made inroads in imposing variable sparsity assumptions, either the ℓ_2 or ℓ_1 -norm measure of the model, to different regions of a 2D seismic tomography problem. They demonstrate that this choice could be automated based on information present in the data. Regions reacting favourably to the sparsity assumption are defined through a learning process.

The next logical step, which I am taking in this research, is to further generalize the use of sparse assumptions for $p \in [0, 2]$ applied to both the model and model gradients in 3D. I provide algorithmic details regarding the mixing of smooth, sparse and blocky assumptions. This general inversion framework can be used to generate diverse solutions and thus help geologists in their interpretation. A suite of models with broad characteristics can also form a basis for learning algorithms.

1.1.4 Dealing with remanence

Contrary to density, magnetization has the added complexity of being a vector quantity defined by strength and orientation. In matter, the total magnetization, which is the magnetic moment per unit volume \vec{M} in (1.7) can be separated in its induced and remanent component such that:

$$\vec{M} = \kappa(\vec{H}_0 + \vec{H}_s) + \vec{M}_r , \quad (1.8)$$

where the magnetic susceptibility κ (SI) is the physical property describing the ability of a rock to get magnetized under an applied field. In nature, this inducing field has two components. The geomagnetic field \vec{H}_0 originates from Earth's core. It is in most cases the dominant component. Secondary fields \vec{H}_s are related to local magnetic anomalies. For highly susceptible material, the secondary fields can oppose the geomagnetic field direction and reduce the total magnetization. It is also referred to as self-demagnetization effects. The remanent magnetization \vec{M}_{rem} is a permanent dipole moment that is preserved in the absence of an inducing field.

It has long been assumed that the induced component of magnetization was dominant. The effect of remanence is often regarded as ‘noise’ and simply ignored by 2D filtering and 3D inversion methods. Recent studies have shown however that remanent minerals, most often magnetite and members of the titanium-hematite series, are commonly associated with mineral deposits such as diamondiferous kimberlites, volcanic massive sulphides and porphyries (Enkin, 2014; Henkel, 1991). The orientation of remanent magnetization can make geologic interpretation more complicated and should not be ignored.

Meanwhile, the same remanent component has been used extensively in paleomagnetic studies and in that field, it has been regarded as geophysical ‘data’. A number of researchers have used the permanent magnetization orientation to map continental block rotation (Kissel and Laj, 1989; Norris and Black, 1961; Vine and Matthews, 1963), for fold and thrust belts reconstruction (Ramon et al., 2012; Vilalain et al., 2015) and in geochronology (Enkin, 2003; Henkel, 1991; Lockhart et al., 2004). While providing valuable information about Earth’s history, these studies have relied primarily on laboratory measurements performed on oriented

cores. The availability and cost to acquire the orientation of magnetization at a point remains a limiting factor for paleomagnetic studies (Pueyo et al., 2016).

Researchers have investigated ways to estimate the magnetization in order to deal with remanence. This includes search algorithms (Dannemiller and Li, 2006; Fedi et al., 1994), magnetic moment analysis (Helbig, 1963; Phillips, 2003), and inversion methods. The most common inversion approach uses simple parametric shapes to approximate elongated and tabular bodies (Clark, 2014; Foss and McKenzie, 2011; Fullagar and Pears, 2013; Pratt et al., 2014). For more complex settings, magnetic data can be transformed into a quantity that is weakly sensitive to the orientation of magnetization, such as magnetic amplitude data, and inverted for a scalar quantity, the effective susceptibility (Shearer, 2005). Lelievre and Oldenburg (2009) introduce the Magnetization Vector Inversion (MVI) for the 3D modelling of magnetization parameters. Because a vector is sought, there are three times the number of parameters compared to conventional inversion. This increases the non-uniqueness, and useful information from an inversion will require more complex regularization functions. Queitsch et al. (2019) have shown the benefits of inverting full magnetic gradient tensor data along with sparsity assumptions imposed on the amplitude of magnetization. Other studies have explored the use of sparsity assumptions for the recovery of compact bodies, but imposing such assumptions on the orientation of magnetization remains difficult (Fournier, 2015; Liu et al., 2015; Zhu et al., 2015).

1.2 Research objectives

The challenges posed by the non-uniqueness of inverse problems and complications posed by remanence form the motivation of my Ph.D. research. The accuracy of potential field inversions depends on our ability to recover the shape of anomalies and, in the case of magnetic inversion, to also accurately predict the magnetization direction. I feel that more knowledge and confidence can be gained from 3D potential field inversion if the algorithm can produce a suite of diverse, yet geologically reasonable solutions with minimal input required from experts. The inversion framework should allow for sharp and smooth geological contacts to fit a broad range of geological scenarios supported by the geophysical data. I want

to extract subtle information present in the data to better define the geometry of geological bodies. More specifically, I want to answer the following questions:

- Can I further generalize the use of sparse norms to generate diverse solutions and determine dominant patterns?
- Can sparse norms be used to constrain the magnetization vector inversion?
- Can we improve the edge definition of geological domains?

In this thesis, I will attempt to bridge the gap between geological and geophysical modelling with three technical innovations. First, I propose to vary sparsity assumptions imposed on the model and model gradients in a semi-automated fashion such that details about sparsity parameters can vary locally without the direct input from the user. At the core of this research is the implementation of a mixed ℓ_p -norms regularization. This algorithm will allow me to generate a suite of possible candidates representing a broad range of characteristics. Secondly, I provide a learning strategy to extract dominant features from the solution space. This aspect is important to help practitioners in their interpretation. Thirdly, I improve the convergence of the non-linear MVI in spherical coordinates. When my improved MVI is combined with sparse norms, I can recover magnetized bodies with well-defined edges and coherent magnetization orientation. The different components are tested on synthetic examples and finally implemented in the Kevitsa case study.

1.3 Thesis arrangement

The chapters in this thesis are meant to build on each other. I progressively add the tools needed for my analysis of the Kevitsa case study in Chapter 7. A number of synthetic examples are provided along the way to showcase specific aspects of inversion. In Chapter 2, I provide theoretical background information related to the forward calculation of potential field data. I provide efficiency improvements over conventional codes by using a nested OcTree mesh-decoupling strategy and out-of-core sensitivity storage.

Chapter 3 introduces the theory needed for potential field inversion. I provide numerical details for the implementation of mixed ℓ_p -norm inversions. This chapter is an improvement of the work presented in my MSc thesis (Fournier et al.,

2016). I use a 1D numerical example to justify changes in the algorithm. I showcase the flexibility of the algorithm on a synthetic 3D gravity problem.

In Chapter 4, I review the work of Lelievre and Oldenburg (2009) on magnetic vector inversion and provide improvements to the non-linear formulation in spherical coordinates by using an iterative sensitivity re-weighting. Sparsity assumptions introduced in Chapter 3 are used to constrain the inversion for the recovery of compact magnetic anomalies with coherent magnetization orientation.

Chapter 5 brings together advances in mixed norm inversion and the directional objective function of Li and Oldenburg (2000). I introduce a 7-cell gradient operator to improve the symmetry of rotated bodies in 3D. I showcase the capability of the rotated norms for imaging a folded layer. Surface dip and strike data are interpolated and used to enforce directionality for the recovery of a continuous layer in 3D.

In Chapter 6, I make inroads in exploring the model space with mixed ℓ_p -norm inversions. Building upon Chapter 3, I propose a learning strategy to extract dominant features from a suite of models with different characteristics. I build an average model with Principal Component Analysis (PCA) and extract local parameters with edge detection and image moment algorithms. I extend the work presented in Chapter 4 to estimate the dip and strike and geological units.

In Chapter 7, I bring together all these components to process gravity and magnetic data acquired over the Kevitsa deposit, Finland. I provide a suite of density and magnetization models to characterize the 2.0 Ga old ultra-mafic intrusion and host stratigraphy. I incorporate geological trends in the modelling using rotation angles derived from a seismic reflection profile and surface structural data. An ensemble of models is analyzed in relation to borehole physical property measurements and conceptual understanding of the deposit. I infer tectonic deformation from the recovered magnetization vector.

I conclude the thesis with a summary chapter assessing the success of the proposed methodology, identifying weaknesses and providing avenues for future research.

Chapter 2

Forward modeling of potential field data

My goal is to characterize the sub-surface in terms of density and magnetization. In Chapter 1, I have introduced the basic physics describing the relation between density and magnetization and their respective anomalous fields in integral form. In this chapter, I provide numerical details about the transformation from the continuous space to a discrete linear system of equations. I also provide efficiency improvements over conventional implementations for the processing of large scale data sets.

2.1 Discrete systems

The usual strategy is to represent the continuous Earth in terms of unit elements each contributing to the total response observed at a given position in space. For the gravity problem, the integral in (1.3) can be evaluated analytically over a discrete prism with uniform density ρ . As derived by Nagy (1966), the integration gives rise to a linear system:

$$\mathbf{g} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \rho , \quad (2.1)$$

where \mathbf{T} relates the cell-centered density value to the observed gravity field \mathbf{g} at some location $P(x_P, y_P, z_P)$:

$$\begin{aligned}
 T_x &= -G \left(\arctan \frac{dy \, dz}{r \, dx} + \log [dz + r] + \log [dy + r] \right) \Big|_{x_L, y_L, z_L}^{x_U, y_U, z_U} \\
 T_y &= -G \left(\arctan \frac{dx \, dz}{r \, dy} + \log [dz + r] + \log [dx + r] \right) \Big|_{x_L, y_L, z_L}^{x_U, y_U, z_U} \\
 T_z &= -G \left(\arctan \frac{dy \, dy}{r \, dz} + \log [dy + r] + \log [dy + r] \right) \Big|_{x_L, y_L, z_L}^{x_U, y_U, z_U} \\
 r &= (dx^2 + dy^2 + dz^2)^{1/2} \\
 dx &= (x_P - x), \quad dy = (y_P - y), \quad dz = (z_P - z)
 \end{aligned} \tag{2.2}$$

and G is Newton's gravitational constant. Parameters needed to define the position and shape of a unit cell are presented in Figure 2.1. Only the lower southwest $L(x_L, y_L, z_L)$ and upper northeast $U(x_U, y_U, z_U)$ corner coordinates are needed to define the relative distance $\mathbf{r}(dx, dy, dz)$ between the observation point and the nodal limits. In this research I use right-handed Cartesian coordinate system such that the \hat{x} , \hat{y} and \hat{z} coordinate axes point along the east, north and vertical (up) direction respectively.

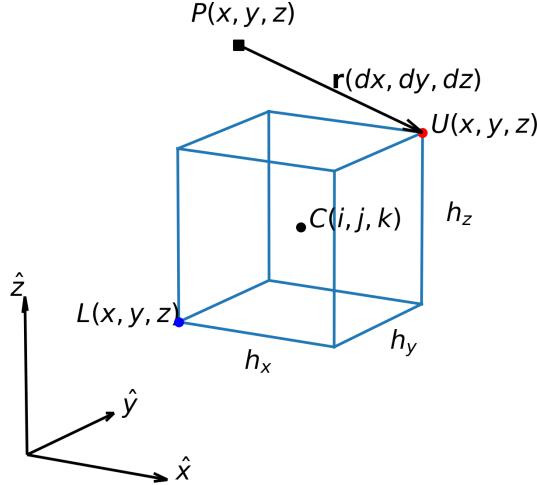


Figure 2.1: Parameters describing the spatial relationship between an observation point P and a rectangular prism C .

For most gravity field experiments only the vertical component of field g_z is measured, such that (2.1) reduces to

$$g_z = T_z \rho \quad (2.3)$$

Equation (2.3) defines the gravity response of a single rectangular prism as observed at a single position in space. I can augment equation (2.3) to describe a gravity experiment conducted over a large volume of earth and at many observation stations

$$\mathbf{g}^{pre} = \mathbf{G}\rho \quad (2.4)$$

such that the linear forward operator $\mathbf{G} \in \mathbb{R}^{N \times M}$ maps the contribution of M number of prisms ($\rho \in \mathbb{R}^M$), each contributing to the response measured over N observation locations ($\mathbf{g}^{pre} \in \mathbb{R}^N$). There are many ways to organize the cells making up this discrete model. In all the work presented in this thesis, I use an Octree-based discretization. More details regarding this choice of parameterization are provided in the following section.

Similarly for the magnetic response, the integral equation in (1.7) can be evaluated analytically for a single prism (Sharma, 1966). This gives rise to a linear system of the form

$$\mathbf{b} = \mathbf{T}\mathbf{m}, \quad (2.5)$$

where \mathbf{T} is a dense 3-by-3 symmetric matrix describing the linear relation between a prism with magnetization $\mathbf{m} = [M_x, M_y, M_z]^T$ to the components of the field $\mathbf{b} = [b_x, b_y, b_z]^T$.

$$\mathbf{T} = \frac{\mu_0}{4\pi} \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{xy} & T_{yy} & T_{yz} \\ T_{xz} & T_{yz} & T_{zz} \end{bmatrix} \quad (2.6)$$

where μ_0 is the magnetic permeability of free-space. It is important to note that the tensor \mathbf{T} is a symmetric matrix with zero trace. Therefore only five of the nine

tensor components need to be calculated.

$$\begin{aligned}
T_{xx} &= -\arctan \frac{dx dy}{dx^2 + r dz + dz^2} \begin{vmatrix} x_U & y_U & z_U \\ x_L & y_L & z_L \end{vmatrix} \\
T_{yy} &= -\arctan \frac{dx dy}{dy^2 + r dz + dz^2} \begin{vmatrix} x_U & y_U & z_U \\ x_L & y_L & z_L \end{vmatrix} \\
T_{xy} &= \log [dz + r] \begin{vmatrix} x_U & y_U & z_U \\ x_L & y_L & z_L \end{vmatrix} \\
T_{xz} &= \log [dy + r] \begin{vmatrix} x_U & y_U & z_U \\ x_L & y_L & z_L \end{vmatrix} \\
T_{yz} &= \log [dx + r] \begin{vmatrix} x_U & y_U & z_U \\ x_L & y_L & z_L \end{vmatrix}
\end{aligned} \tag{2.7}$$

For most geophysical applications, we do not measure the vector field \vec{b} , but rather the Total Magnetic Intensity (TMI) of the field that includes both the geomagnetic and secondary fields. Since we are only interested in the anomalous response from rocks, the approximation is generally made that

$$b^{TMA} = \vec{b} \cdot \hat{\mathbf{H}}_0 - \mu_0 \|\vec{H}_0\|; \tag{2.8}$$

such that the Total Magnetic Anomaly b^{TMA} is assumed to be small and parallel to Earth's field \vec{H}_0 . The simulated magnetic datum in (2.5) simplifies to

$$b^{pre} = [\hat{\mathbf{H}}_0^\top \cdot \mathbf{T}] \mathbf{m} \tag{2.9}$$

Just as for the gravity experiments, this system can be augmented for M number of prisms ($\mathbf{m} \in \mathbb{R}^{3M}$) and N observation locations ($\mathbf{b}^{pre} \in \mathbb{R}^N$)

$$\mathbf{b}^{pre} = \mathbf{F} \mathbf{m} \tag{2.10}$$

such that $\mathbf{F} \in \mathbb{R}^{N \times 3M}$. This linear system has three times the number of parameters compared to the gravity problem as magnetization is a vector property. As part of my contribution to the open-source community, both the gravity and magnetic kernels have been added to the `SimPEG.PF` library.

2.1.1 Choice of discretization

I have so far established the linear equations (2.4) and (2.10) that map the gravity and magnetic response for a collection of rectangular prisms making up a discrete Earth. I have yet to define how these cells are organized in a 3D space. This decision will directly affect the size of the forward (and later inverse) calculations as the size of the linear operators scale linearly with respect M . Approximating the Earth more efficiently will allow me to process large data sets.

The usual strategy is to define a core region of interest with a fine discretization and surround it by coarser cells (padding) to absorb regional signals that may be present in the data. Two options are available to organize rectangular prisms. The simplest implementation uses a regular grid, or tensor mesh shown in Figure 2.2(a). Each unit element shares a face with 6 neighbours. Changes in cell size propagate throughout the domain along the orthogonal direction. The use of tensor meshes has dominated the inversion literature due to the ease of storing and viewing uniformly gridded models.

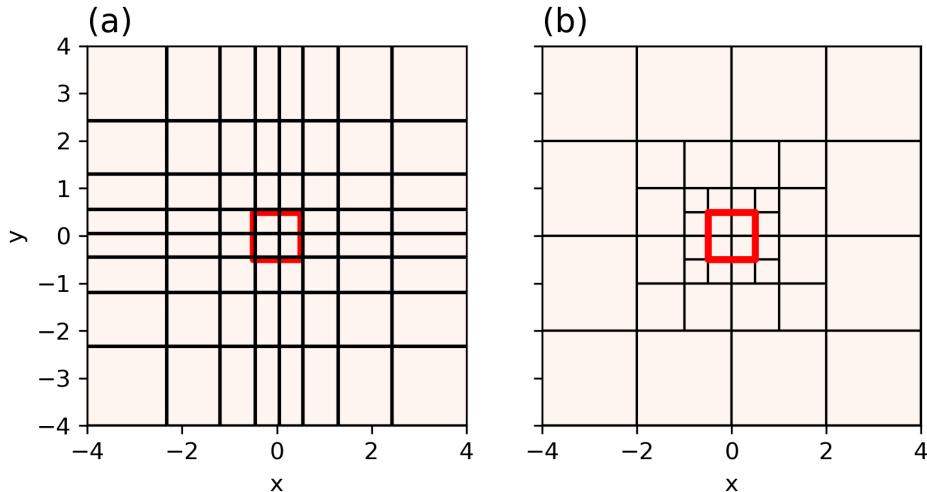


Figure 2.2: Simple representation of (a) tensor and (b) Octree meshes for the organization of rectangular prisms within a core domain (red) and padding region over a square domain 4 x 4 m in width. The Tensor mesh can fill the space with 64 cells, compared to only 40 cells with the Octree mesh.

In an Octree mesh, unit cells are organized in a hierarchical structure as shown in Figure 2.2(b). The resolution of the grid is increased by dividing a parent cell into 8 children (or 4 in 2D). This type of discretization offers the most flexibility for increasing the resolution of the mesh in a specific region without affecting the resolution of boundary cells. The main challenge is in generating a mesh that honours the geometry of the problem in terms of data location, topography and geological contacts. Over the course of my research, I contributed to the open-source community with a suite of refinement functions to facilitate the creation process of Octree meshes. As part of the `SimPEG.discretize` library, I implemented the following three strategies:

- Box refinement includes all cells intersected by a rectangular box containing the input points.
- Radial refinement is performed inside spheres centered on each input points. The radial distance is determined by the user.
- Surface refinement is defined by a continuous Delaunay triangulation of the input points (Barber et al., 1996). The Octree refinement is determined based on the vertical distance between cell centers and the nearest triangle.

Figure 2.3 compares the three refinement methods for the discretization of points (red) placed on a Gaussian surface. The `box` refinement is the simplest but also the least efficient strategy as it yields a uniform grid similar to the Tensor discretization. For the `radial` refinement, I end up with a small number of cells concentrated around the input points. It is an optimal refinement for scattered observation points. The `surface` is well suited to describe continuous features such as topography and geological contacts. It gives, in this case, the most accurate representation of the Gaussian surface.

Forward simulation test

I demonstrate the benefit of an Octree discretization by forward modelling the response of a 1 m sphere shown in Figure 2.4. I want to compare the numerical cost to perform forward simulations using the standard Tensor discretization and an Octree mesh with `surface` refinement. Octree refinement uses scatter points placed

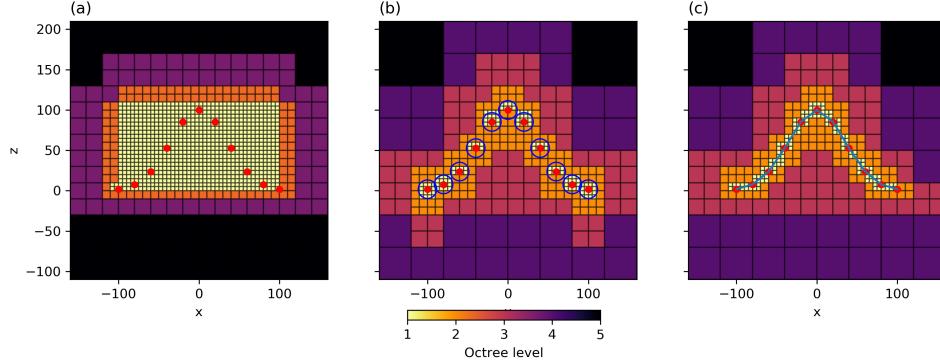


Figure 2.3: Three refinement strategies used to discretize a Gaussian curve defined by scattered points (red): (a) box, (b) radial and (c) surface refinement strategy from the SimPEG.discretize library. Cells are coloured by their corresponding Octree level.

on the outer surface of the sphere. Using equation (2.5), I calculate the magnetic response over an 11-by-11 grid of observations placed 1 m above the anomaly. I repeat the forward simulation over a range of cell sizes. Only cells inside the sphere are considered. The forward calculations are compared to the analytical vertical response (b_z^{ana}) for a vertical magnetization:

$$b_z^{ana} = \frac{\mu_0}{4\pi} \left[-\frac{\mathbf{M}}{r^3} + \frac{3(\mathbf{M} \cdot \mathbf{r})\mathbf{r}}{r^5} \right] \cdot \hat{z} \quad (2.11)$$

where $\mathbf{M} = [0\hat{x}, 0\hat{y}, 1\hat{z}]$ and \mathbf{r} defines the vector between the center of the sphere and the observation location. Table 2.1 summarizes the forward calculations in terms of total number of cells, run time and data residual between the simulation and the analytic solution.

$$\phi_d = \sum_{i=1}^{121} (b_i - b_i^{ana})^2 \quad (2.12)$$

All calculations were performed on a single thread 2.4 GHz Intel processor.

For the Tensor mesh, a reduction in core cell size increases the total number of cells by a factor 8. For the same reduction in cell size using the surface refinement increases the number of cells by a factor 4. This is anticipated as only cells at the surface of the sphere decrease in size compared to a full volume refinement in

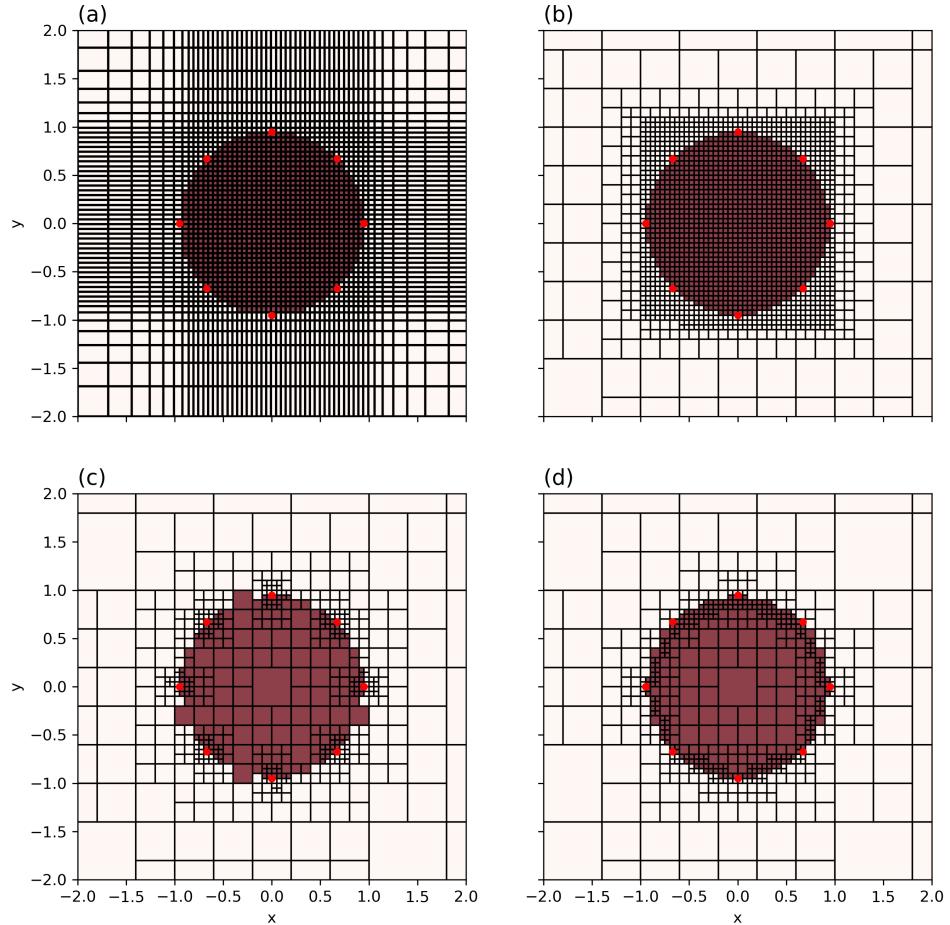


Figure 2.4: (a) Discretization of a sphere defined by discrete points (red) using the conventional Tensor mesh with a core region and padding cells. Octree meshes refined by (b) box, (c) radial and (d) surface methods from the SimPEG.discretize library.

the Tensor mesh. Both discretizations can reproduce the analytical response with roughly the same accuracy. Including padding cells to this problem would further increase the efficiency gap between the two discretization methods as the Octree mesh can rapidly increase the cell size with little influence from the discretization in the core region.

Cell size (m)	Tensor			Octree (Surface)		
	# Cells	Time (s)	ϕ_d	# Cells	Time (s)	ϕ_d
1.00e-01	544	1.8e-01	4.2e+02	496	1.6e-01	7.7e+02
5.00e-02	4196	9.6e-01	4.7e+01	2516	5.1e-01	7.1e+01
2.50e-02	33478	6.5e+00	4.2e+00	10836	2.1e+00	5.2e+00
1.25e-02	268080	6.5e+01	2.0e+00	47276	1.0e+01	2.7e+00

Table 2.1: Summary table for the forward modeling of a magnetized sphere using a Tensor and Octree discretization

2.2 Large scale problems

The amount of memory needed to store the dense linear operators for gravity and magnetics is a limiting factor for the simulation of potential field data with the integral formulation. As prescribed in (2.10), the size of the problem is linearly scaled by the number of model parameters M and the number of the data N . For moderate size problems encountered in exploration geophysics ($M \approx 10^6$, $N \approx 10^4$) the memory requirement to store a dense $M \times N$ matrix can exceed hundreds of gigabytes. While modern supercomputers can handle problems of this size, it is still out of reach for common desktop computers.

A number of strategies have been proposed in the past to reduce the size of the problem. Compression methods, either in the Fourier (Pilkington, 1997) or wavelet domain (Li and Oldenburg, 2003), have proven successful in reducing the size of the forward operators. On the downside, compression methods are prone to introducing modelling artifacts, which can be hard to differentiate from true anomalies. This is especially an issue for the magnetic vector inversion explored in Chapter 3.

Another group of methods takes advantage of the rapid decay of geophysical signals to reduce the size of individual forward simulations. The assumption is made that model parameters in the far-field of measurements contribute little to the total response and can, therefore, be approximated with fewer parameters. In this regard, the concept is loosely related to the Fast Multipole Method (Engheta et al., 1992). This is especially valid for airborne EM experiments where cells beyond roughly 10 times the flight height of airborne EM systems have a negligible impact on the measured response (Reid and Pfaffling, 2006). In its simplest implemen-

tation, Cox et al. (2010) define a footprint approach such that model parameters outside a pre-defined tolerance are simply ignored.

More recently, Yang and Oldenburg (2014) introduced a mesh decoupling approach. Local meshes are designed based on the source location and decay rate of the geophysical signal. Cell-centered conductivity values are homogenized using volumetric-weighted averaging. Similarly, Haber and Schwartzbach (2014) designed a nested Octree mesh strategy. The core region of local meshes is co-located with cells in the global mesh and cover the same lateral extent. Physical property values are homogenized from small cells in the global mesh to larger (or equal) cells in the local meshes. This strategy assures that far-field features can still contribute to the total response. The accuracy of individual forward simulations depends primarily on the interpolation scheme used to transfer physical properties from the global to the local meshes.

In this research, I employ the methodology of Haber and Schwartzbach (2014). I want to divide the full data set into subsets (or tiles), each associated with a local mesh. Physical property values are transferred to local meshes using a volumetric weighted average. In order to further minimize the memory footprint of the simulation, I make use of the parallel Dask package (Dask, 2016). The library was designed by the data science community for out-of-core processing of large datasets (Rocklin, 2015). The `zarr` file format allows the storage of dense arrays on a solid-state drive (SSD) in compressed memory chunks. Out-of-core storage is appealing as it reduces the amount of Random-Access Memory (RAM) needed to store the large forward operators. Read and write operations are done in parallel. The runtime depends largely on the hardware used in terms of processor speed, number of processors and communication speed between the workers and the solid-state memory. The size of individual memory chunks can be adjusted to optimize the processing speed depending on the resources available.

In the case of potential field data, the source of geophysical signals can span a broad range of wavelengths, from tens of meters (local) to hundreds of kilometres (regional). The memory footprint of a problem can be reduced considerably if we manage to approximate far-field features with fewer cells. In the most extreme case, each observation point could have its own local mesh. Choosing the number of tiles becomes a trade-off between accurately capture the heterogeneity of the

model for forward modeling while minimizing traffic between the parallel workers and the SSD memory.

2.2.1 Numerical test

I investigate this trade-off between accuracy and efficiency with a numerical example. I create a synthetic gravity model shown in Figure 2.5. The model contains both short wavelength information from a small block anomaly and long wavelength signal from large domains in the north and southeast quadrants. From equation (2.4), I compute gravity data on 21×21 grid placed 1 m above a flat topography. Forward simulation of the global model required 4.3 Mb of memory and was performed in parallel in 6.7 s with 4 processors (2.4 GHz). It is a relatively small example compared to industry standards, but similar trends are to be expected on larger scale problems.

I will assess the trade-off between efficiency and numerical accuracy by performing a series of five forward simulations using a range of tiling patterns. I break up the dataset into 2, 4, 9, 12 and 16 square tiles. For each tiling experiment, I calculate the total memory footprint (Gb) and the computation time (s) to perform the forward simulations. I also compute the data residual between the global simulation (single mesh) and the tiled simulation using the ℓ_2 -norm measure in equation (2.12).

As seen in Figure 2.7, the memory needed to perform the forward calculations rapidly decreases but eventually levels-off as the number of tiles increases. This is expected as local meshes necessitate a minimum number of cells to fill out the global domain and a minimum number of small padding cells near the edge of the tiles to reduce interpolation artifacts. The reduction in problem size is inversely correlated with an increase in computational time as communication between the workers and the SSD memory becomes a bottleneck.

I also note an increase in data residual as a function of tile size. Figure 2.8 compares the simulated data from the global mesh (single tile) to the combined forward simulation calculated with 12 tiles (last experiment). From the residual map, it is possible to distinguish short wavelength discrepancies between adjacent forward simulations (tiles). These residuals are primarily due to the homogenization of

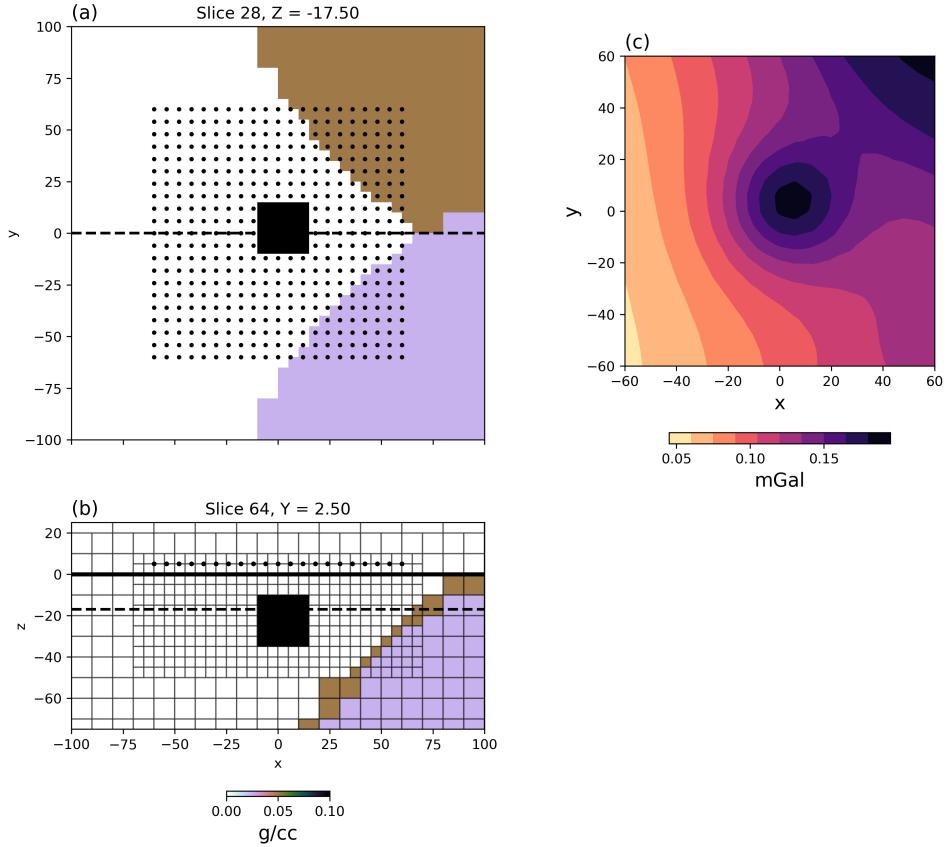


Figure 2.5: (a) Horizontal and (b) vertical sections through the synthetic density model used to test the mesh decoupling strategy. (c) The simulated data contain short and long wavelength information.

anomalies near the edges of tiles, which in turn is a function of the wavelength information contained in the data. As a first pass, I establish experimentally the appropriate padding distance based on the estimated data uncertainties, such that the maximum residual falls below the experimental error. For this experiment, the meshing artifacts are at most 2% of the data amplitude (or 0.004 mGal), which I achieved with a minimum padding distance of 4 cells per Octree level. Determining an optimal padding distance as a function of the geophysical signal would warrant further research.

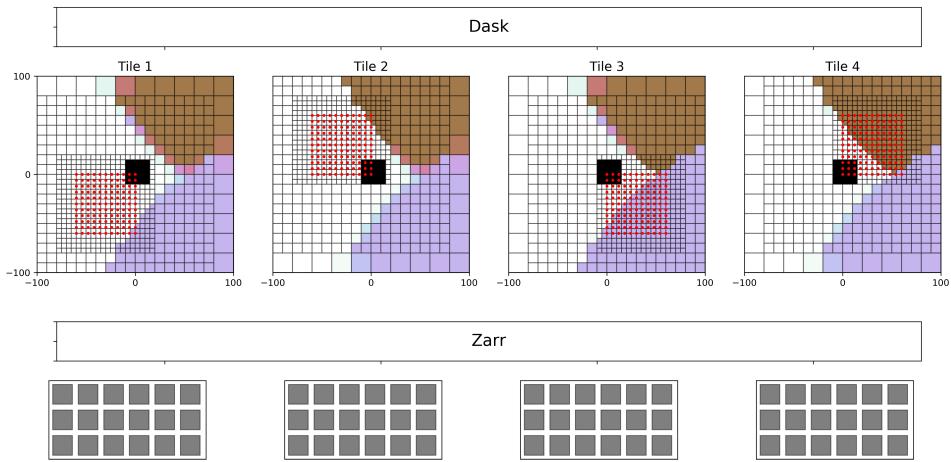


Figure 2.6: Example of mesh decoupling of the forward problem into nested Octree tiles. (Middle) Local meshes are generated for a pre-defined number of tiles nested inside the global domain. (Bottom) The forward modelling operator for each tile is stored in compressed Zarr file format on solid-state memory. Memory chunks (grey) are accessed in parallel by the Dask library to perform the forward calculations.

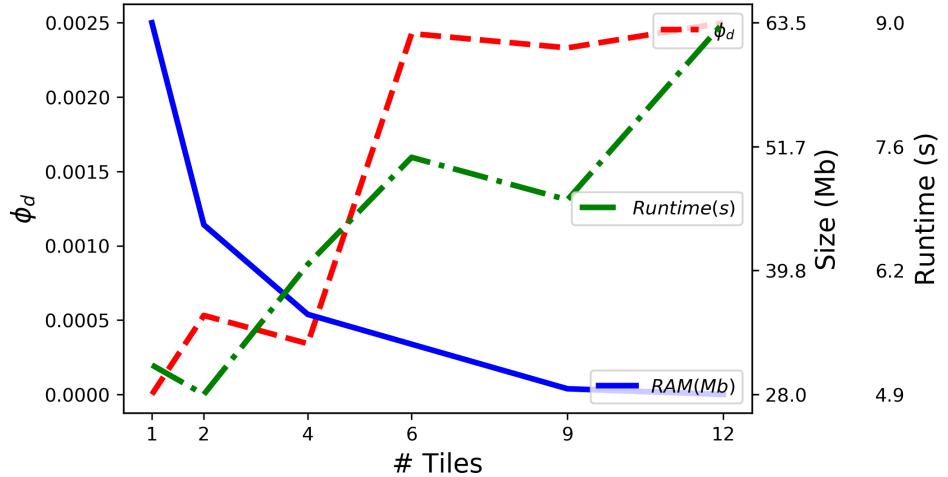


Figure 2.7: Trade-off curves between the size of the forward problem (blue) and data residual (red) over a range of number of tiles. The total size of the problem is calculated based on the sum of cells in all the local meshes times the number of data. The optimal number of tiles would be at the point of intersection where both the cost of forward calculations and the data residual change significantly.

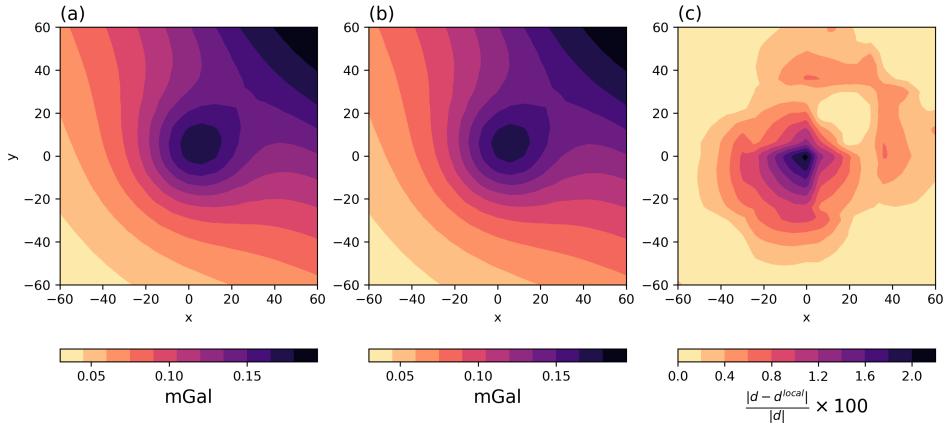


Figure 2.8: Simulated gravity data calculated from (a) the global model and (b) the 12 forward tiled calculations. (c) Data residuals show short wavelength discrepancies between adjacent tiles due to interpolation effects.

Chapter 3

Inverse Problem

In Chapter 2, I defined the linear equations relating density and magnetization to gravity and magnetic data. I now review the theory needed to solve the inverse problem, such that I can recover a 3D representation of the subsurface from the observed data. A key issue related to the inverse problem is that there is an infinite number of possible models that can satisfy the data. Field measurements are generally acquired from the surface resulting in the inverse problem to be ill-posed. The presence of experimental noise further complicates the problem. To circumvent these issues, the inversion is often formulated as an optimization problem of form

$$\begin{aligned} \min_{\mathbf{m}} \phi(\mathbf{m}) &= \phi_d + \beta \phi_m \\ \text{subject to } \phi_d &\leq \phi_d^*. \end{aligned} \tag{3.1}$$

where ϕ_d is the misfit function

$$\phi_d = \sum_{i=1}^N \left(\frac{d_i^{pred} - d_i^{obs}}{\sigma_i} \right)^2, \tag{3.2}$$

that measures the residuals between the observed and predicted data \mathbf{d}^{pre} , normalized by the estimated data uncertainties σ .

The regularization function ϕ_m , or model objective function, serves as a vehicle to introduce *a priori* information in the inversion. Several regularization strategies have been developed over the last decades such that the solution remains geologi-

cally plausible. I focus on the generic ℓ_p -norm regularization of the form

$$\phi_m = \sum_{r=s,x,y,z} \alpha_r \int_V w(m) |f_r(m)|^{p_j} dV . \quad (3.3)$$

The functions f_r can take many forms but most often have been

$$f_s = m - m^{ref}, f_x = \frac{dm}{dx}, f_y = \frac{dm}{dy}, f_z = \frac{dm}{dz} . \quad (3.4)$$

Thus $f_s(m)$ measures the deviation from a reference model m^{ref} and $f_x(m)$, $f_y(m)$ and $f_z(m)$ measure the roughness of the model m along orthogonal directions in 3D. This optimization problem has multiple terms scaled by hyper-parameters. The first parameter is β which controls the balance between misfit and regularization. It is assumed that a value can be found such that the target misfit is reached. The α 's are constants that control the relative influence of the different regularization functions. A larger α -value increases the focus of the optimization on the corresponding penalty function. User-defined weights $w(m)$ are used to incorporate any type of *a priori* information that may be available to guide the solution.

Most often the ℓ_2 -norms measure has been used giving rise to a discrete linear system of form

$$\begin{aligned} \phi_m &= \alpha_s \phi_s + \alpha_x \phi_x + \alpha_y \phi_y + \alpha_z \phi_z \\ &= \sum_{r=s,x,y,z} \alpha_r \|\mathbf{W}_r \mathbf{V}_r \mathbf{G}_r (\mathbf{m} - \mathbf{m}^{ref})\|_2^2 , \end{aligned} \quad (3.5)$$

where ϕ_s measures the deviation of the discrete model \mathbf{m} from a reference model \mathbf{m}^{ref} and ϕ_x , ϕ_y and ϕ_z measure the roughness of the model along Cartesian directions. The reference model is sometimes omitted in the roughness terms. The matrices \mathbf{G}_x , \mathbf{G}_y , and \mathbf{G}_z are discrete gradient operators. For the smallness component, \mathbf{G}_s reduces to the identity matrix. Volumes of integration resulting from the evaluation of (3.3) are applied through diagonal matrices such that

$$\mathbf{V}_s = \text{diag} \left[\mathbf{v}^{1/2} \right] . \quad (3.6)$$

where \mathbf{v} holds the discrete volume elements corresponding to each cell. For the

model derivative terms, the volume of integration is computed over cell interfaces such that

$$\mathbf{V}_x = \text{diag} \left[\left(\mathbf{A}_C^{F_x} \mathbf{v} \right)^{1/2} \right]. \quad (3.7)$$

where the matrix $\mathbf{A}_C^{F_x}$ averages the cell-centered volumes to cell faces. Similar averaging is performed along the orthogonal y and z -direction. Diagonal matrices \mathbf{W}_r hold user-defined weights. More details about these weights are provided in the following section. Lastly, the α parameters control the relative importance given to individual components of the regularization. What is sometimes sought, at least as a first pass, is that if $\phi_s, \phi_x, \phi_y, \phi_z$ have about the same numerical value then they are contributing equally. Dimensional analysis shows that for a uniform discretization h :

$$\frac{[\phi_x]}{[\phi_s]} = [h]^{-2}. \quad (3.8)$$

The common approach is to set α_s accordingly in order to scale the components of the regularization function.

The usual strategy to solve (3.1) is through a gradient descent algorithm, such as a Gauss-Newton approach, where we attempt to find a solution that has zero gradients

$$\mathbf{g} = \nabla_m \phi(\mathbf{m}) = \nabla_m \phi_d + \beta \left[\alpha_s \nabla_m \phi_s + \alpha_x \nabla_m \phi_x + \alpha_y \nabla_m \phi_y + \alpha_z \nabla_m \phi_z \right] = \mathbf{0}. \quad (3.9)$$

where ∇_m stands for the partial derivatives of the function with respect to the discrete parameterization \mathbf{m} . A solution to (3.9) can readily be calculated by gradient descent methods (Hestenes and Stiefel, 1952; Nocedal and Wright, 1999).

A large number of studies have made use of this formulation to incorporate a variety of *a priori* information: physical property data from rock and core samples (Lelièvre et al., 2009), structural knowledge (Lelièvre, 2009; Li and Oldenburg, 2000) and advanced 3D geological modeling ((Bosch and McGaughey, 2001; Fullagar et al., 2008; Phillips, 1996; Williams, 2008). Although successful in identifying imaging anomalies at depth, penalty functions that rely on ℓ_2 -norm measures have a limited range of possible outcomes. The models tend to be smooth and difficult to interpret in relation to known geological domains with discrete boundaries.

Moreover, substantial modelling work is generally required by experts to manually refine these constraints in order to test different geological scenarios.

Sensitivity weighting

The weighting matrices \mathbf{W}_r introduced in (3.5) can take many forms depending on the type of *a priori* information that may be available. For potential fields problems, a sensitivity weighting function is generally used to counteract the rapid decay of the geophysical signal as a function of distance. In the work of Li and Oldenburg (1996), a *distance weighting* approximation is employed and fixed at the onset. In this thesis, I resort to an iterative re-weighting strategy based on the sensitivity of a given inverse problem

$$\mathbf{J} = \frac{\partial F[\mathbf{m}]}{\partial \mathbf{m}}, \quad (3.10)$$

where \mathbf{J} , also referred to as the Jacobian matrix, holds the partial directives of the forward problem $F[\mathbf{m}]$ with respect to \mathbf{m} . Adapted from Haber et al. (1997), I formulate the sensitivity-based weighting function:

$$\begin{aligned} \mathbf{W}_s &= \text{diag} \left[\left[\frac{\mathbf{w}}{\max(\mathbf{w})} \right]^{1/2} \right] \\ w_j &= \left[\sum_{i=1}^N J_{ij}^2 + \delta \right]^{1/2} / v_j, \end{aligned} \quad (3.11)$$

where \mathbf{w} measures the sum square of the columns of the Jacobian, normalized by the cell volumes. The constant δ is a small value (near machine precision) added to avoid singularity. The weights are normalized by the maximum value such that the range of weights are bounded between [0, 1]. The same sensitivity weights can also be applied to the model derivative terms using a cell averaging operation such that

$$\mathbf{W}_x = \text{diag} \left[\left(\mathbf{A}_C^{F_x} \left[\frac{\mathbf{w}}{\max(\mathbf{w})} \right] \right)^{1/2} \right], \quad (3.12)$$

similar to the averaging of volumes used in (3.7). This sensitivity weighting strategy is general and adaptable to any inverse problems where the sensitivity matrix

can be calculated explicitly. While the initial purpose of the sensitivity weighting function of Li and Oldenburg (1996) is to simply counteract the decay of potential fields, I will show numerically in Chapter 4 that the iterative re-scaling process can also be beneficial in improving the convergence of gradient methods applied to non-linear inverse problems.

3.0.1 Synthetic gravity example

As an entry point to the inverse problem, I proceed with a simple synthetic gravity example. I define a volume of interest 600 m wide by 300 m deep, over which I place a uniform survey grid of 21 x 21 stations placed 5 m above a flat topography. The core region directly below the survey grid is discretized at a 5 m resolution as shown in Figure 3.1(a). Within the core region, I build a geophysical target made up of a single dense cube, 25 m in width. I set the density contrast of the prism to 0.2 g/cc in a uniform zero background. From (2.3), I simulate the vertical gravity field of the block and add random Gaussian noise with 10^{-3} mGal standard deviation. Figures 3.1(b) and (c) display the simulated data and ‘noisy’ observations \mathbf{g}_z^{obs} used in the inversion. I revisit this example in Chapter 4 to demonstrate the inversion process on magnetic data.

From the noisy data I will attempt to recover the block anomaly by the inverse process. The objective function to be minimized takes in this case the form:

$$\min_{\mathbf{m}} \phi(m) = \|\mathbf{G} \rho - \mathbf{d}^{obs}\|_2^2 + \beta \sum_{r=s,x,y,z} \alpha_r \|\mathbf{W}_r \mathbf{V}_r \mathbf{G}_r \rho\|_2^2 \quad (3.13)$$

subject to $\phi_d \leq \phi_d^*$

where I set $\rho^{ref} = \mathbf{0}$. Since (3.13) is linear with respect to the density contrast model ρ , I can solve it uniquely for a fixed trade-off parameter β . I repeat this process for variable β values until a find a solution that satisfies $\phi_d \approx N$. Figure 3.2(a) presents a vertical section through the recovered density model. The density anomaly is imaged at roughly the right position, but the edges of the block are poorly defined. As normally obtained with ℓ_2 -norm penalties, the solution is smooth and density values remain near the zero reference model. Hence the need to explore other regularization functions that can better resolve compact objects.

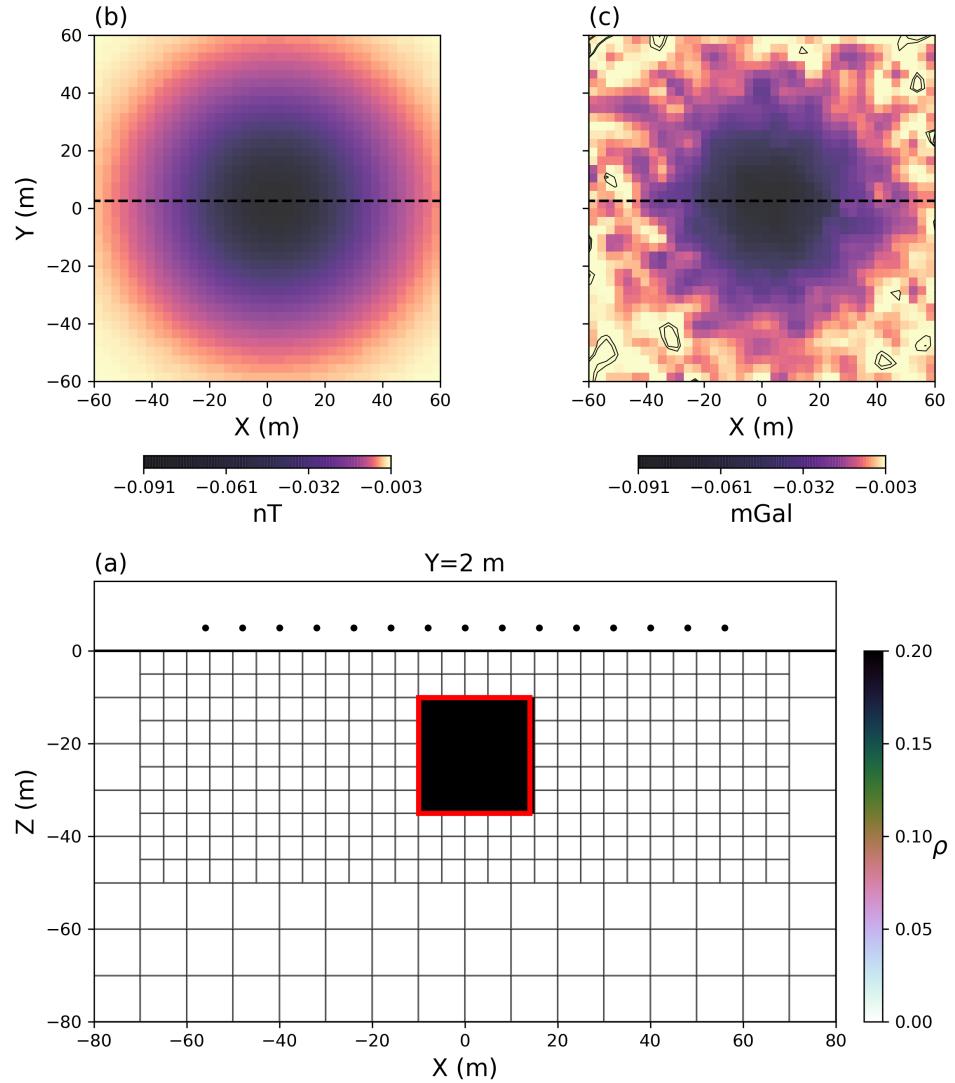


Figure 3.1: (a) Vertical section through a 25 m cube with density $\rho=0.2$ g/cc placed in a uniform zero density background. (b) Simulated gravity data responses on a 21×21 survey grid placed 5 m above the flat topography. (c) Gravity data with random Gaussian noise added, 10^{-3} mGal standard deviation.

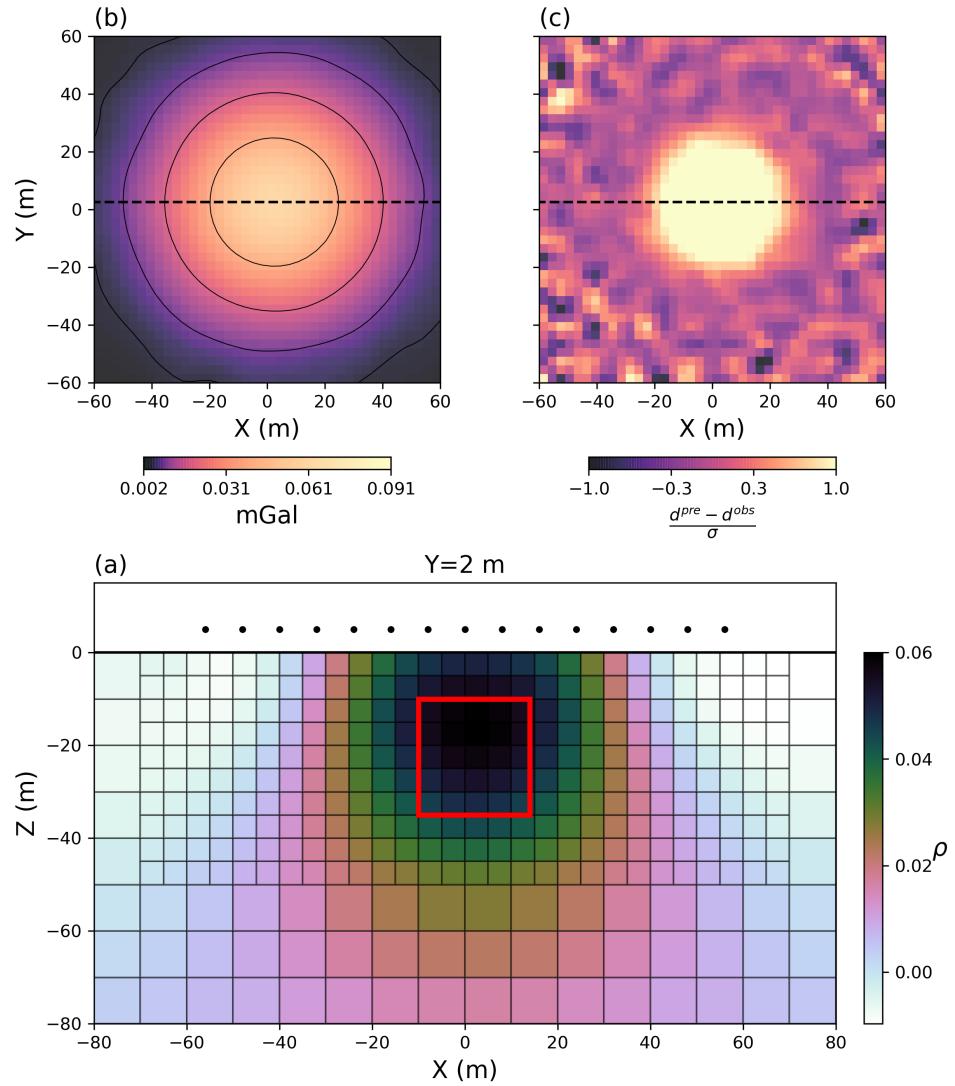


Figure 3.2: (a) Vertical section through the inverted density model using the conventional ℓ_2 -norm regularization, (b) predicted and (c) normalized data residual. Outline of the true model (red) is shown for reference.

3.1 General ℓ_p -norm regularization

Alternatively, researchers have explored the use of non- ℓ_2 measures to promote the recovery of compact anomalies. Approximations to ℓ_1 -norm such as the Huber norm (Huber, 1964)

$$\sum_i |m_i|^p \approx \sum_i \begin{cases} m_i^2, & |m_i| \leq \varepsilon, \\ 2\varepsilon|m_i| - \varepsilon^2, & |m_i| > \varepsilon, \end{cases}$$

and the Ekblom norm (Ekblom, 1973):

$$\sum_i |m_i|^p \approx \sum_i (m_i^2 + \varepsilon^2)^{p/2} \quad (3.14)$$

have received considerable attention in geophysical inversion and signal processing (Daubechies et al., 2010; Farquharson and Oldenburg, 1998; Gorodnitsky and Rao, 1997; Li, 1993; Sun and Li, 2014). Likewise, the Lawson's measure (Lawson, 1961)

$$\sum_i |m_i|^p \approx \sum_i \frac{m_i^2}{(m_i^2 + \varepsilon^2)^{1-p/2}}, \quad (3.15)$$

has been proposed to approximate ℓ_0 -norm and it has proven useful in generating minimum support models. This formulation has received considerable attention in the literature. (Ajo-Franklin et al., 2007; Barbosa and Silva, 1994; Last and Kubik, 1983; Portniaguine, 1999). Figure 3.3 compares the ℓ_p -norms with the Lawson approximation over a range of model values. As $\varepsilon \rightarrow 0$, the approximation approaches the ℓ_p -norm on the complete interval $p \in [0, 2]$. While (3.15) would in theory permit us to explore a wide range of solutions for $0 \leq p \leq 2$, its numerical implementation remains challenging. Most algorithms have been limited to the ℓ_0 , ℓ_1 , and ℓ_2 -norm measure applied evenly to all components of the model objective function.

Recent efforts by Sun and Li (2014) have shown promise in further exploring the model space by varying ℓ_p -norm measures locally. They divided the inversion domain into regions reacting favourably to either the ℓ_1 or ℓ_2 -norm regularization. The automated process could adapt to complex geological scenarios where both

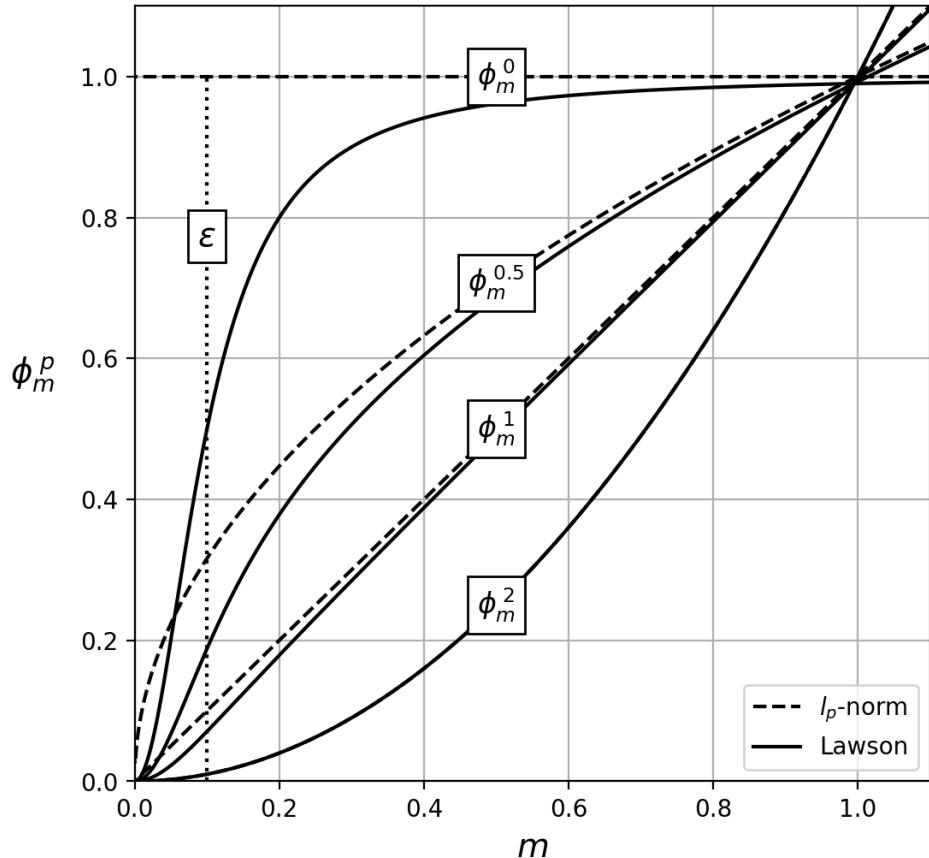


Figure 3.3: Approximated ℓ_p -norm using the Lawson measure (Lawson, 1961) over a range of p -values and for a fixed threshold parameter $\epsilon = 10^{-1}$.

smooth and blocky anomalies are present. Building upon the work I introduced in my M.Sc. Thesis (Fournier, 2015), I want to extend the work of Sun and Li (2014) and further generalized the mixed norm inversion for $p \in [0, 2]$.

3.1.1 Synthetic 1D problem

To develop my methodology it suffices to work with a simple test example. In Figure 3.4(a) I present a synthetic 1D model made up of a boxcar anomaly. The region is divided into 50 uniform cells distributed along the interval $[0 \leq x \leq 1]$. I

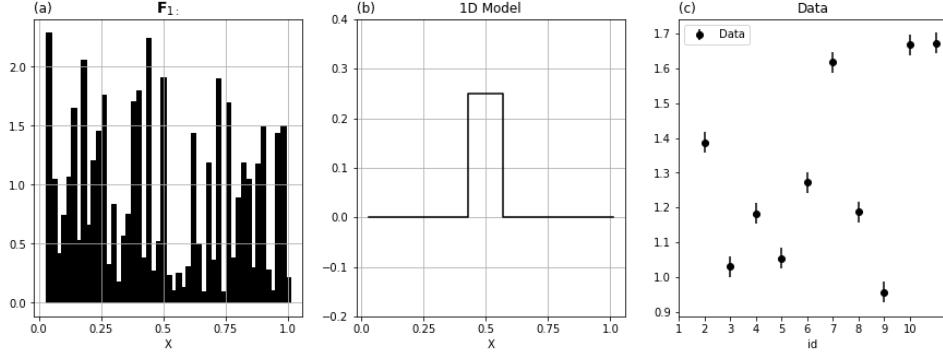


Figure 3.4: Linear forward problem made up of: (a) an example kernel function ; (b) model; (c) observed data with assigned standard errors.

define a synthetic geophysical experiment such that the data (\mathbf{d}^{obs}) are

$$\mathbf{d}^{obs} = \mathbf{F} \mathbf{m}^{true} + \mathbf{e}, \quad (3.16)$$

The kernel coefficients F_{ij} are sampled from a standard normal distribution of positive values multiplied by the discretization intervals. Choosing a stochastic kernel function for a linear inverse problem is unusual. Smooth functions are usually employed (polynomials, decaying exponentials, sinusoids), but my choice will serve to highlight the effects of various regularization functions. I generate 10 data, so $\mathbf{F} \in \mathbb{R}^{N \times M}$ where $M = 50$ and $N = 10$. Random Gaussian error \mathbf{e} ($\sigma=0.025$) is added to simulate noise (Fig. 3.4(c)).

To begin my analysis, I invert my synthetic dataset with two simple regularization functions. For this 1D problem, the objective function takes the form:

$$\min_{\mathbf{m}} \phi(\mathbf{m}) = \|\mathbf{F} \mathbf{m} - \mathbf{d}^{obs}\|_2^2 + \beta \sum_{r=s,x} \alpha_r \|\mathbf{W}_r \mathbf{V}_r \mathbf{G}_r \mathbf{m}\|_2^2 \quad (3.17)$$

subject to $\phi_d \leq \phi_d^*$

To simplify the analysis, I set all weighting terms to unity ($\mathbf{W}_r = \mathbf{I}$) and the reference model to zero. Figure 3.5(a) presents the recovered model after reaching the target misfit ($\phi_d^* = N$) using the smallness term alone ($\alpha_x = 0$). The solution exhibits high variability similar to the stochastic kernel function, but model param-

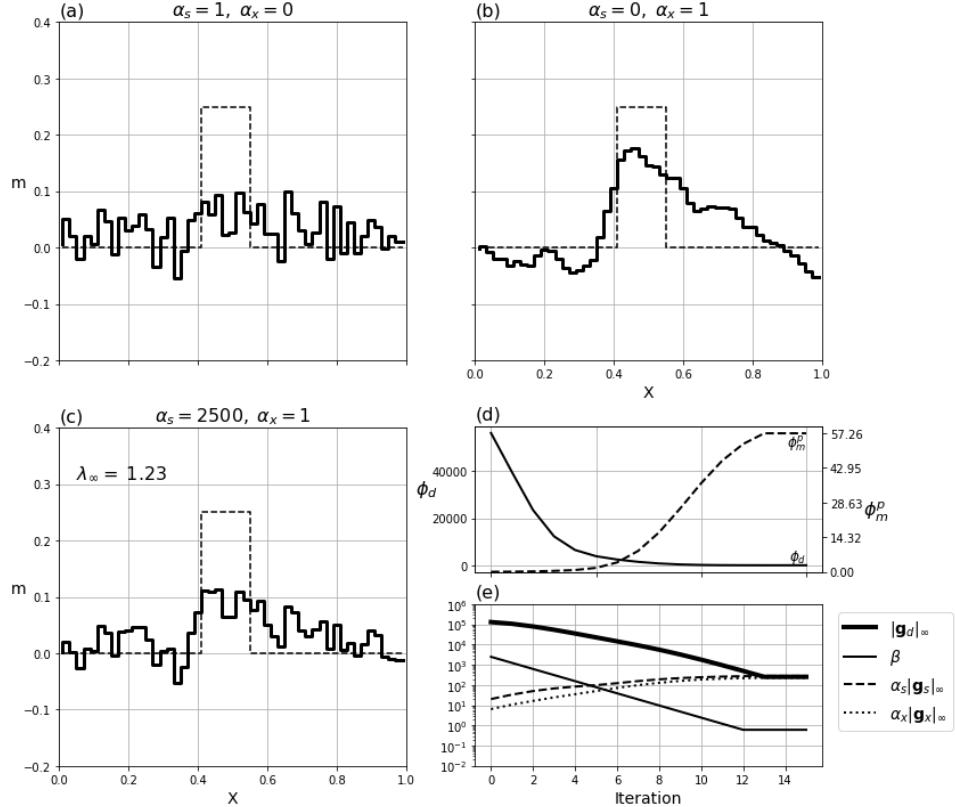


Figure 3.5: Solution to the 1D inverse problem using (a) an ℓ_2 -norm on the model ($\alpha_x = 0$), (b) the ℓ_2 -norm on model gradients ($\alpha_s = 0$) and (c) combined regularization function ($\alpha_s = 2500$, $\alpha_x = 1$). (d) Convergence curve comparing the misfit (ϕ_d) and the regularization (ϕ_m) as a function of iterations. (e) Comparative plot for the relative contribution of the different components of the objective function measured in terms of maximum absolute gradient ($\|\mathbf{g}_i\|_\infty$)

eters remain near the implied zero reference value. Next, I invert the data using the model gradient term ($\alpha_s = 0$); this yields the smoother model presented in 3.5(b). The solution shows less spatial variability and the horizontal position of the boxcar anomaly is better located.

Next, I combine both regularization functions so that the solution remains close to the implied zero reference value and it is smooth. I need to determine the length scale weighting proposed in (3.8). For my problem $h = 0.02$ and hence following

the relationship established in (3.8) I set $\alpha_x = 1$ and $\alpha_s = 2500$. Inverting with these parameter values yields the model in 3.5(c). Convergence curves presented in Figure 3.5(d) show the evolution of ϕ_d and ϕ_m as a function of iteration. As β decreases (shown in Figure 3.5(e)), the misfit, ϕ_d progressively decreases while model complexity, indicated by ϕ_m , progressively increases.

Visually, the solution 3.5(c) exhibits characteristics of remaining near the zero reference value while also attempting to be smooth. Numerical evaluation of the two components of the regularization function, presented in Table 3.1, show that $\alpha_s \phi_s = 3.31$ and $\alpha_x \phi_x = 1.13$. This might suggest that both ϕ_s and ϕ_x are roughly equal in importance.

Rather than working with global norms, in this study, I propose to quantify the relative importance of the terms in the regularization function based on their partial derivatives, or gradients. From (3.9) I expect to find an optimal solution where the sum of the gradients vanishes, either because all components are equal to zero, or because multiple gradients have opposite signs. To quantify the size of the gradients I use the infinity-norm

$$\|\mathbf{g}_r\|_\infty = \|\nabla_m \phi_r\|_\infty , \quad (3.18)$$

corresponding to the maximum absolute value of \mathbf{g}_r . The $\|\mathbf{g}_r\|_\infty$ metric is appealing for a few reasons: (a) it is directly linked to the minimization process because I use gradient descent methods, (b) it does not depend on the dimension M of the parameter space as do other measures that involve a sum of components of the vector, (c) the theoretical maximum can be calculated analytically for any given ℓ_p -norm function. These properties will become useful in the following section when I attempt to balance different norm penalties applied on a cell-by-cell basis.

Figure 3.5(e) compares $\|\mathbf{g}_d\|_\infty$, $\alpha_s \|\mathbf{g}_s\|_\infty$ and $\alpha_x \|\mathbf{g}_x\|_\infty$ over the iterative process. I note that, under the current α -scaling strategy proposed in (3.8), the individual partial derivatives for ϕ_s and ϕ_x also appear to be proportional in magnitude. To quantify this I define a proportionality ratio:

$$\lambda_\infty = \frac{\alpha_s \|\mathbf{g}_s\|_\infty}{\alpha_x \|\mathbf{g}_x\|_\infty} \quad (3.19)$$

α_s	α_x	$\alpha_s \phi_s$	$\alpha_x \phi_x$	λ_∞
2500	1	3.31	1.13	1.23

Table 3.1: Norm values and proportionality ratio obtained for the 1D solution presented in Figure 3.5(c). A proportionality ratio of $\lambda_\infty \approx 1$ indicates that the components of the regularization function are both contributing significantly to the final solution.

I shall use λ_∞ as an indicator to evaluate the relative influence of (any) two terms in the regularization function. For my example $\lambda_\infty = 1.23$, from which I infer that ϕ_s and ϕ_x are contributing nearly equally to the solution (Table 3.1). As I further generalize the regularization function for arbitrary ℓ_p -norm measures, I will attempt to maintain this proportionality ratio ($\lambda_\infty \approx 1$) between competing functions so that my modeling objectives are preserved throughout the inversion process.

3.1.2 Iterative Re-weighted Least Squares algorithm

Solutions obtained with ℓ_2 -norm regularization functions provided some insight about the sought model but better representations can be obtained by employing general ℓ_p -norms:

$$\phi_s^p = \sum_i |m_i|^p \quad (3.20)$$

My main focus is in the regularization function in (3.3) which I approximate with the Lawson norm such that

$$\phi_m = \sum_{r=s,x} \alpha_r \int_V w(m) \frac{f_r(m)^2}{\left(f_r(m)^2 + \varepsilon^2\right)^{1-p_r/2}} dV, \quad (3.21)$$

This measure makes the inverse problem non-linear with respect to the model. The common strategy is to solve the inverse problem through an Iterative Reweighted Least-Squares (IRLS) approach such that (3.21) is expressed as a weighted least-squares problem. The denominator is evaluated for model parameters obtained from the most recent iteration such that

$$\phi_r^{p_r} = \sum_{i=1}^M w_i v_i \frac{f_{r_i}(m)^2}{[(f_{r_i}(m^{(k-1)}))^2 + \varepsilon^2]^{1-p_r/2}} \quad (3.22)$$

where $m_i^{(k-1)}$ are model parameters obtained at a previous iteration. The integral corresponding to the smallest model component can be written as:

$$\phi_s^{p_s} = \sum_{i=1}^M w_{si} v_{si} \frac{m_i^2}{((m_i^{(k-1)})^2 + \epsilon^2)^{1-p_s/2}} \quad (3.23)$$

In (3.23) I have explicitly written the objective function as $\phi_s^{p_s}$ to indicate that I am evaluating a smallest model component with an ℓ_p -norm with $p = p_s$. This approximation of the ℓ_p -norm can be implemented within the same least-squares framework used in (3.5) such that:

$$\phi_s^{p_s} = \|\mathbf{W}_s \mathbf{V}_s \mathbf{R}_s \mathbf{m}\|_2^2. \quad (3.24)$$

where the IRLS weights \mathbf{R}_s are defined as

$$\begin{aligned} \mathbf{R}_s &= \text{diag}[\mathbf{r}_s]^{1/2} \\ r_{s_i} &= \left((m_i^{(k-1)})^2 + \epsilon^2 \right)^{p_s/2-1}. \end{aligned} \quad (3.25)$$

Carrying out the same procedure on the measure of model derivatives yields

$$\phi_x^{p_x} = \sum_{i=1}^{M-1} w_{x_i} v_{x_i} \frac{\left(\frac{m_{i+1}-m_i}{h_i} \right)^2}{\left[\left(\frac{m_{i+1}^{(k-1)}-m_i^{(k-1)}}{h_i} \right)^2 + \epsilon^2 \right]^{1-p_x/2}} \quad (3.26)$$

where h_i defines the cell-center distance between neighboring model parameters. Equation (3.26) can also be expressed in linear form as

$$\phi_x^{p_x} = \|\mathbf{W}_x \mathbf{V}_x \mathbf{R}_x \mathbf{G}_x \mathbf{m}\|_2^2, \quad (3.27)$$

where the gradient operator and the corresponding IRLS weights are calculated by

$$\mathbf{G}_x = \begin{bmatrix} -h_1^{-1} & h_1^{-1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & -h_{M-1}^{-1} & h_{M-1}^{-1} \end{bmatrix}. \quad (3.28)$$

and

$$\mathbf{R}_x = \text{diag} [\mathbf{r}_x]^{1/2}$$

$$r_{x_i} = \left[\left(\frac{m_{i+1}^{(k-1)} - m_i^{(k-1)}}{h_i} \right)^2 + \epsilon^2 \right]^{p_x/2-1}. \quad (3.29)$$

respectively. The final regularization function is thus

$$\phi_m^p = \alpha_s \|\mathbf{W}_s \mathbf{V}_s \mathbf{R}_s \mathbf{m}\|_2^2 + \alpha_x \|\mathbf{W}_x \mathbf{V}_x \mathbf{R}_x \mathbf{G}_x \mathbf{m}\|_2^2. \quad (3.30)$$

The core IRLS procedure described in Table 3.2 involves two main stages:

1. Stage 1 solves the inverse problem using ℓ_2 -norms presented in (3.5). The assumption is made that the globally convex ℓ_2 -norm regularized inversion is a good approximation of the true solution and it is used to form the initial IRLS weights defined in (3.25). The β parameter is controlled by a cooling schedule that starts with a high value and is successively decreased until $\phi_d \approx \phi_d^*$.
2. Stage 2 starts from the solution obtained in Stage 1 and solves the inverse problem iteratively using the regularization in (3.30) and a standard Gauss-Newton procedure. A gradient descent direction $\delta\mathbf{m}$ is found by solving

$$\mathbf{H} \delta\mathbf{m} = \mathbf{g} \quad (3.31)$$

where \mathbf{H} is the approximate Hessian and \mathbf{g} is the gradient of the objective function. I use the Conjugate Gradient method (Hestenes and Stiefel, 1952) to solve this system.

The model update at the k^{th} iteration is

$$\mathbf{m} = \mathbf{m}^{(k-1)} + \alpha \delta\mathbf{m} \quad (3.32)$$

where the step length α is found by a line-search back-stepping method (Nocedal and Wright, 1999). Gradient steps are only performed if the data misfit remains

within the user-defined tolerance η_{ϕ_d} .

$$\frac{|\phi_d - \phi_d^*|}{\phi_d^*} \leq \eta_{\phi_d} \quad (3.33)$$

If outside the tolerance, the algorithm repeats the Gauss-Newton calculation with the previous $\mathbf{m}^{(k-1)}$ and a different β -value, either lower or higher depending on the achieved ϕ_d . This β -search step is an important component in the workflow when the minimization switches between an l_2 to an l_p objective function because ϕ_m^p can vary markedly. This can force a change of β by a few orders of magnitude in some cases. Once an appropriate β has been found such that (3.33) is respected, the model update $\mathbf{m}^{(k)}$ is accepted and used for the next iteration cycle. The IRLS process continues until the change in regularization falls below some pre-defined tolerance η_{ϕ_m}

$$\frac{|\phi_m^{(k-1)} - \phi_m^{(k)}|}{\phi_m^{(k)}} < \eta_{\phi_m} \quad (3.34)$$

I set to $\eta_{\phi_m} = 10^{-5}$ (0.01% change) in all my experiments. Using the above algorithm I now explore specific inversions for a fixed $\varepsilon = 10^{-3}$ and uniform norms, with $p = 1$ and $p = 0$, applied on the model and model gradients.

3.1.3 Case 1: ℓ_1 -norm ($p_s = p_x = 1$)

I first explore the end member of convex functions for $p_s = p_x = 1$ for which optimality can be guaranteed (Daubechies et al., 2010; Osborne, 1985). Using the procedure prescribed in Table (3.2), I invert the 1D problem with three different regularization functions: (a) l_1 -norm measure of the model ($\alpha_x = 0$), (b) l_1 -norm measure of the model gradients ($\alpha_s = 0$) and (c) for the combined penalties using $\alpha_s = 2500$, $\alpha_x = 1$, which I previously used for the l_2 -norm inversion.

As shown in Figure 3.6(a), the first inversion is successful in recovering a sparse solution. From Linear Programming (LP) theory, the expected optimal solution would have as many non-zero parameters as there are linearly independent constraints or 10 values in this case. For comparison, I solve the LP problem by the Simplex routine from the open-source library `Scipy.Optimization.linprog` (Jones et al., 2001). Figure 3.6(a) compares both solutions and shows that my im-

Stage 1: Initialization (ϕ_m^2)	Stage 2: IRLS (ϕ_m^p)
$\min_m \phi_d + \beta \phi_m^2$ <p>s.t. $\phi_d = \phi_d^*$</p> $\beta^{(0)}, \mathbf{m}^{(0)}, \mathbf{R}^{(0)}, \phi_m^{(0)}$	while $\frac{ \phi_m^{(k-1)} - \phi_m^{(k)} }{\phi_m^{(k)}} > \eta_{\phi_m}$ do β-Search $k := k + 1$ $\beta^{(k)}, \mathbf{m}^{(k)}, \mathbf{R}^{(k)}$

β-Search
Solve $\mathbf{H} \delta \mathbf{m} = \mathbf{g}$ $\mathbf{m} = \mathbf{m}^{(k-1)} + \alpha \delta \mathbf{m}$ if $\frac{ \phi_d - \phi_d^* }{\phi_d^*} > \eta_{\phi_d}$ adjust β , re-do else continue

Table 3.2: IRLS algorithm in pseudo-code made of two stages: Stage 1 Initialization with convex least-squares inversion, Stage 2 IRLS updates with inner β -search steps.

lementation of IRLS for l_1 -norm yields a solution in close agreement with the Simplex routine. A better approximation could be obtained (not shown here) by lowering the threshold parameter ε . I will examine this aspect of the algorithm in the following section. Figure 3.6(b) presents the solution for the l_1 -norm applied to the model gradients. The final solution is *blockier* and the general shape of the boxcar model has been improved.

Lastly, the solution obtained with the combined l_1 -norm regularization on the model and model derivative is shown in Figure 3.6(c); it is similar to that in Figure 3.6(a). This shows that the smallest model component has dominated the solution. This is quantified by the evaluated proportionality ratio $\lambda_\infty = 50$; setting $\alpha_s = 2500$ is too large. To understand this result, I can factor out a base cell length

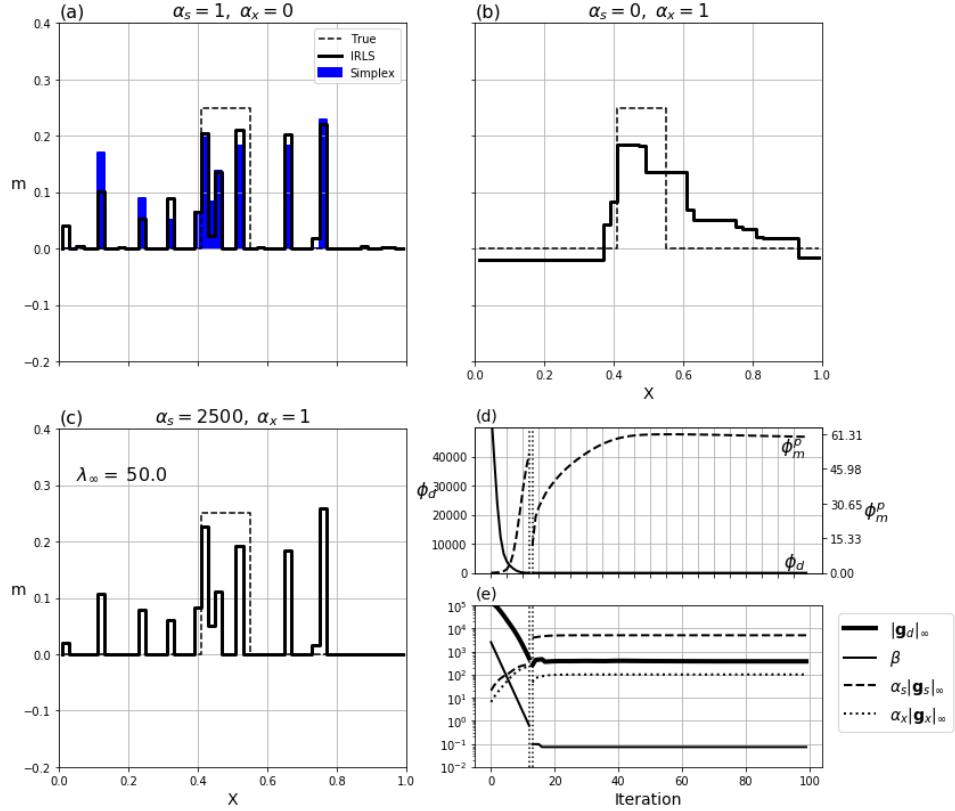


Figure 3.6: (a) Two solutions using an ℓ_1 -norm on the model: (blue) Simplex, and (black) IRLS method. (b) Solution obtained with the approximated ℓ_1 -norm (IRLS) penalty on model gradients alone and (c) with the combined penalty functions ($\alpha_s = 2500$, $\alpha_x = 1$). The calculated proportionality ratio λ_∞ indicates that the combined penalties is dominated by the ϕ_s^1 term. (d) Convergence curve and (e) maximum partial derivatives associated with the components of the objective function as a function of iterations for the inversion in (c). The vertical dotted lines indicate the change in regularization from an ℓ_2 -norm to ℓ_1 -norm measure.

\mathbf{h} from (3.26) such that

$$\begin{aligned}
\phi_x^{p_x} &= \sum_{i=1}^{M-1} w_{x_i} v_{x_i} \cdot \frac{h^{-2} \left(\frac{m_{i+1}-m_i}{\hat{h}_i} \right)^2}{\left[h^{-2} \left(\left(\frac{m_{i+1}^{(k-1)}-m_i^{(k-1)}}{\hat{h}_i} \right)^2 + h^2 \epsilon^2 \right) \right]^{1-p_x/2}} \\
&= \sum_{i=1}^{M-1} w_{x_i} v_{x_i} \cdot \frac{h^{-2} \left(\frac{m_{i+1}-m_i}{\hat{h}_i} \right)^2}{h^{p_x-2} \left[\left(\frac{m_{i+1}^{(k-1)}-m_i^{(k-1)}}{\hat{h}_i} \right)^2 + h^2 \epsilon^2 \right]^{1-p_x/2}} \quad (3.35) \\
&= h^{-p_x} \sum_{i=1}^{M-1} w_{x_i} v_{x_i} \cdot \frac{\left(\frac{m_{i+1}-m_i}{\hat{h}_i} \right)^2}{\left[\left(\frac{m_{i+1}^{(k-1)}-m_i^{(k-1)}}{\hat{h}_i} \right)^2 + h^2 \epsilon^2 \right]^{1-p_x/2}}
\end{aligned}$$

where $h = \min(\mathbf{h})$ represents the core discretization length and $\hat{h}_i = h_i/h$. For a uniform grid, \hat{h}_i simply reduces to unity everywhere. Comparing this expression to (3.22) clearly shows a difference in scales between ϕ_s^p and ϕ_x^p , previously fixed in (3.8), that now depends on the chosen p -value such that

$$\frac{[\phi_x^p]}{[\phi_s^p]} = [h]^{-p_x}. \quad (3.36)$$

It also highlights a dependency between the chosen threshold parameter ϵ and the discretization length h . I will revisit this parameter in later sections.

In accordance to this new relationship, I can re-adjust the importance of ϕ_s^p by setting $\alpha_s = 50$, where $p = 1$ and $h = 0.02$. After applying this change I recover the model presented in Figure 3.7(a). The combined assumption of a piece-wise continuous and sparse model yields a solution that closely resembles the true boxcar model. The recovery of \mathbf{m}^{true} has remarkably improved compared to the ℓ_2 -norm solutions (Fig. 3.5), and this demonstrates the power of customizable objective functions. It is important to notice that the re-adjustment of α_s has brought the partial derivatives of $\phi_s^{p_s}$ and $\phi_x^{p_x}$ to a comparable level, with a final proportionality ratio $\lambda_\infty = 1.01$. Even though I have changed the norms during the inversion, the contribution of both penalty functions has remained at a comparable level during

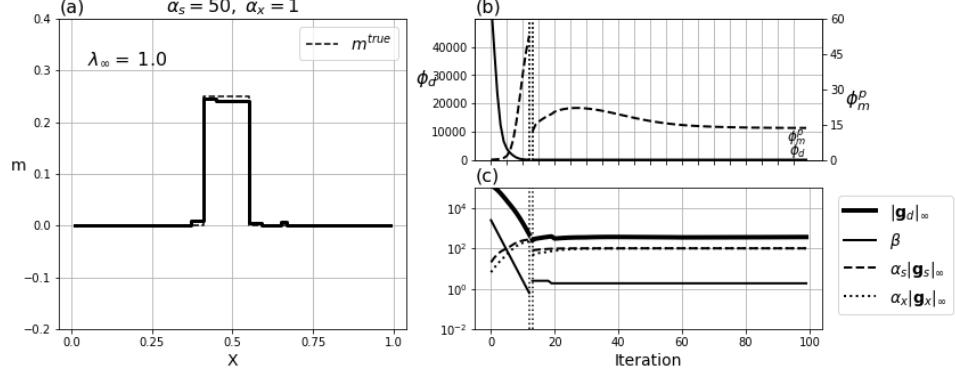


Figure 3.7: (a) Solution obtained with the combined penalty functions $\alpha_s \phi_s^1 + \alpha_x \phi_x^1$ after re-adjustment of $\alpha_s = 50$, $\alpha_x = 1$. (b) Convergence curve and (c) maximum partial derivatives associated with the components of the objective function as a function of iteration.

the transition between Stage 1 and 2 of the algorithm (Fig 3.7(c)).

3.1.4 Case 2: ℓ_0 -norm ($p_s = p_x = 0$)

The main advantage of the IRLS formulation is that it permits, in theory, approximating any norm including the non-linear approximation for $p < 1$. The goal is to potentially recover a model with even fewer non-zero parameters than that obtained by solving the problem with $p = 1$. The IRLS formulation for $p = 0$ has been implemented for various geophysical problems under different names: such as the *compact* inversion (Last and Kubik, 1983), *minimum support* functional (Portnaguine and Zhdanov, 2002), and others (Ajo-Franklin et al., 2007; Barbosa and Silva, 1994; Blaschek et al., 2008; Chartrand, 2007; Stocco et al., 2009).

Following the same IRLS methodology as described in Table 3.2, I invert the synthetic 1D problem with three assumptions: (a) ℓ_0 -norm applied on the model ($\alpha_x = 0$), (b) ℓ_0 on model gradients ($\alpha_s = 0$) and combined penalties ($\alpha_s = 1$, $\alpha_x = 1$). Figure 3.8 presents the solutions for all three cases. I note that in the first case, (a), the approximate ℓ_0 -norm inversion recovers a sparser solution than obtained with the ℓ_1 -norm; there are only eight non-zeros parameters. Similarly for case (b), I recover a model with fewer changes in model values. Finally in case (c) the

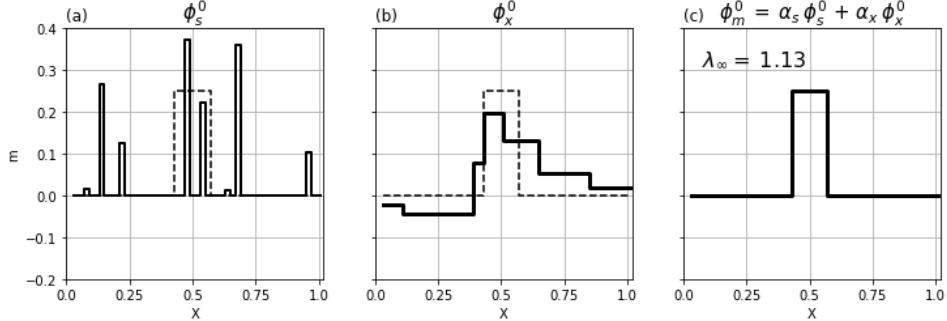


Figure 3.8: Solution to the 1D inverse problem using an approximate ℓ_0 -norm (a) on the model, (b) on model gradients and (c) combined penalty functions using the IRLS algorithm ($\alpha_s = 1$, $\alpha_x = 1$). All three solutions honor the data within the target misfit ϕ_d^* .

solution obtained with the combined ℓ_0 -norm penalties matches almost perfectly the true boxcar anomaly. The final proportionality ratio as calculated from (3.19) indicates once again a good balance between the penalty functions ($\lambda_\infty = 1.13$).

To summarize this section, I have now recovered nine models using different ℓ_p -norm penalties applied on the model and model gradient. All solutions presented in Figure 3.5, 3.6 and 3.8 can reproduce the data within the predefined data tolerance ($\phi_d^{(k)} \approx N$). Without prior knowledge about the true signal, all these solutions would be valid candidates to explain the observed geophysical data.

3.2 Mixed norm regularization

While I was successful in recovering a solution that closely resembles the boxcar model, the same penalty functions might not be appropriate for other models, such as compact targets with smooth edges. I thus explore a broader range of solutions by using the Lawson approximation in (3.23) for any combination of norms on the range $0 \leq p \leq 2$.

The idea of combining different norm measures for the simultaneous recovery of smooth and compact features has partially been explored by Sun and Li (2014) on a 2D seismic tomography problem. They demonstrated the benefits of dividing model space into regions with different ℓ_p -norm penalties. The choice of norms

was limited to be either l_1 or l_2 . Little has been published however on the independent mixing of model and gradient norms on the range $p \in [0, 2]$, although this problem was initially addressed in (Fournier, 2015). I now apply my algorithm to minimize

$$\phi_m^p = \alpha_s \phi_s^0 + \alpha_x \phi_x^2. \quad (3.37)$$

where once again the superscript indicates the ℓ_p -norm measure used in each function. Based upon my previous work, I expect the solution to be sparse, in terms of non-zero model parameters, while smooth with respect to the model gradients. Unfortunately, following the current IRLS strategy, I recover the model presented in Figure 3.9(a). The anomaly is concentrated near the boxcar but appears to be dominated by ϕ_s^0 . There seems to be only marginal influence from ϕ_x^2 . Comparing the partial derivatives of the objective function confirms this. After convergence the calculated proportionality ratio is $\lambda_\infty = 159$. This is a significant change from the end of Stage 1 where $\lambda_\infty \approx 1$. Clearly, iteration 6, at Stage 2 of the IRLS, took the solution away from the proportionality condition (Fig 3.9(c)). I hypothesize that a more desirable solution could be obtained if proportionality was preserved among the components of the objective function throughout the IRLS process. In the following sections, I provide an important modification to the standard IRLS algorithm to achieve this goal.

3.2.1 Scaled-IRLS steps

Since the inverse problem is solved using gradients of the composite objective function $\phi(\mathbf{m})$, the relative magnitude of the individual gradients is a driving force in controlling the iteration step in (3.31). I want to ensure that each penalty term in the objective function is playing a significant role. Taking the partial derivatives of (3.24) with respect to m yields:

$$\mathbf{g}_s^p = \frac{\partial \phi_s^p}{\partial m} = \mathbf{R}_s^\top \mathbf{V}_s^\top \mathbf{W}_s^\top \mathbf{W}_s \mathbf{V}_s \mathbf{R}_s \mathbf{m} \quad (3.38)$$

where I purposely omitted a factor 2 from the differentiations of the ℓ_2 -norm as it gets absorbed by the zero right-end side of (3.9). From Figure 3.10(a), I note that the magnitude of the derivatives increases rapidly for small p values as $m_i \rightarrow 0$.

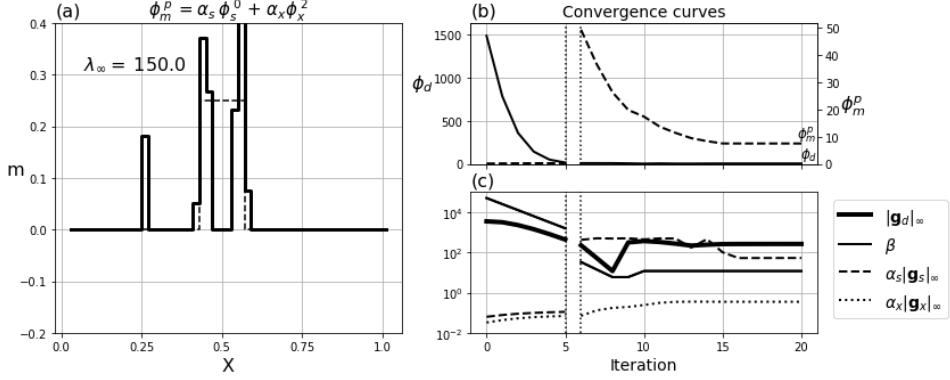


Figure 3.9: (a) Recovered model and (b) convergence curves using the conventional IRLS method for $p_s = 0$, $p_x = 2$ and a fixed threshold parameter $\varepsilon = 10^{-3}$ ($\alpha_s = \alpha_x = 1$). (c) Trade-off parameter and maximum gradients for the different components of the objective function. At the start of Stage 2 (iteration 6), the sudden increase in $\|\mathbf{g}_s\|_\infty$ is matched with a decrease in β . Throughout the inversion, $\|\mathbf{g}_x\|_\infty$ remains small in magnitude.

This trend is accentuated for small ε values as demonstrated in Figure 3.10(b) for $p = 0$. The magnitude of derivatives for $p < 2$ increase rapidly as $m \rightarrow 0$ and $\varepsilon \rightarrow 0$. This results in gradient steps in equation (3.9) that are dominated by sparse norms. This property of the IRLS approximation is important because, when attempting to combine different norm penalties within the same objective function, there will be a systematic bias towards small ℓ_p -norm penalties. To circumvent this bias I propose the re-scale the contribution of each regularization function during the iterative process in order to preserve proportionality. I define the following gradient-based scaling

$$\gamma = \left[\frac{\|\mathbf{g}^2\|_\infty}{\|\mathbf{g}^p\|_\infty} \right]^{1/2}. \quad (3.39)$$

By using this scaling strategy I can equalize the influence of each regularization function. I can evaluate the theoretical maximum of \mathbf{g}_r^p by taking the derivative of (3.21) in terms of $f(m)$, followed by a second derivative after substituting $f(m^{(k-1)})$ for $f(m)$ and setting the expression to zero such that

$$(f(m)^2 + \varepsilon^2)^{p/2-1} + (p-2)f(m)^2(f(m)^2 + \varepsilon^2)^{p/2-2} = 0 \quad (3.40)$$

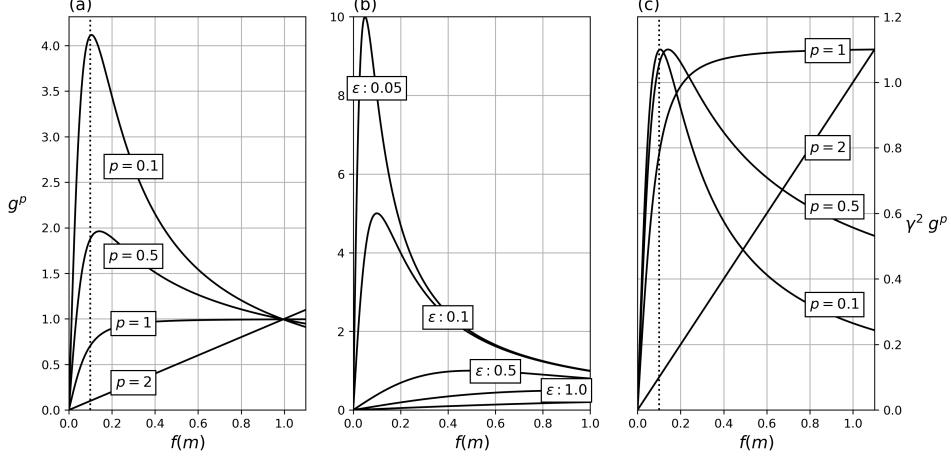


Figure 3.10: Derivatives of the Lawson approximation over a range of model values for (a) a fixed threshold parameter $\epsilon = 10^{-1}$ over a range of p values and for (b) a fixed $p = 0$ over a range of ϵ values. (c) Applying the γ -scaling to the gradients brings all maximums to be equal irrespective of p and ϵ .

The maximum gradient of the Lawson approximation occurs at $f(m)^*$

$$f(m)^* = \begin{cases} \infty \text{ or } f(m)_{\max}, & p \geq 1 \\ \frac{\epsilon}{\sqrt{1-p}}, & p < 1, \end{cases} \quad (3.41)$$

from which I can calculate $\|\mathbf{g}^p\|_\infty$ by substituting $f(m)^*$ into (3.22). I note that for $p < 1$, the maximum gradient does not depend on $f(m)$ but only on the chosen p and ϵ value. Figure 3.10(c) presents the derivatives of different approximated ℓ_p -norms after applying the corresponding γ -scale. The role of γ_s is to reference the partial derivatives of the approximated ℓ_p -norms to the derivatives of its ℓ_2 -norm measure. This re-scaling is done for two reasons. First, at the transition between Stage 1 and 2, it preserves the balance between the misfit and regularization terms and thus no large adjustment in the trade-off parameter β is needed. Secondly, the scaling based on the gradients guarantees that two penalties can co-exist and impact the solution at every step of the IRLS, regardless of the chosen $\{p, \epsilon\}$ -values or the amplitude of $f(m)$.

I therefore define Scaled-IRLS weights such that (3.25) become:

$$\hat{\mathbf{R}}_s = \gamma_s \operatorname{diag} \left[\left((\mathbf{m}^{(k-1)})^2 + \varepsilon^2 \right)^{p_s/2-1} \right]^{1/2} \quad (3.42)$$

where the scaling parameter γ_s is:

$$\gamma_s = \left[\frac{\|\mathbf{g}_s^2\|_\infty}{\|\mathbf{g}_s^{p_s}\|_\infty} \right]^{1/2} \quad (3.43)$$

Two options are possible to compute the γ -scalings: (a) take the maximum absolute gradient directly from the gradient values in (3.38), or (b) calculate $\|\mathbf{g}_j^p\|_\infty$ analytically from (3.41). I have found that Option 2 is more stable since it is based upon a theoretical maximum of the gradient and not on a particular realization of that maximum that arises from the distribution of values in the current model \mathbf{m}^k .

The outcome of the re-scaling strategy is shown in Figure 3.11(a). The solution seems to have my desired properties of being sparse in terms of the number of non-zero model values and the model has smooth edges. The maximum partial derivatives, shown in Figure 3.11(c), confirm that the scaling strategy was successful in balancing the impact of the two components of the regularization. This is quantified by the calculated proportionality ratio $\lambda_\infty = 0.7$. It is an improvement over the previous solution with a ratio of 150 (Figure 3.9), but it appears that the algorithm has reached a steady state solution with slightly more influenced from ϕ_s . In the following section I provide a strategy to better preserve proportionality between each model update through a cooling strategy.

Scaled model derivatives

Applying the same scaling strategy to the model derivative requires additional care as the measure is also dependent on length scales. Following the strategy established in (3.35), I can factor out the length scales from the model derivative term

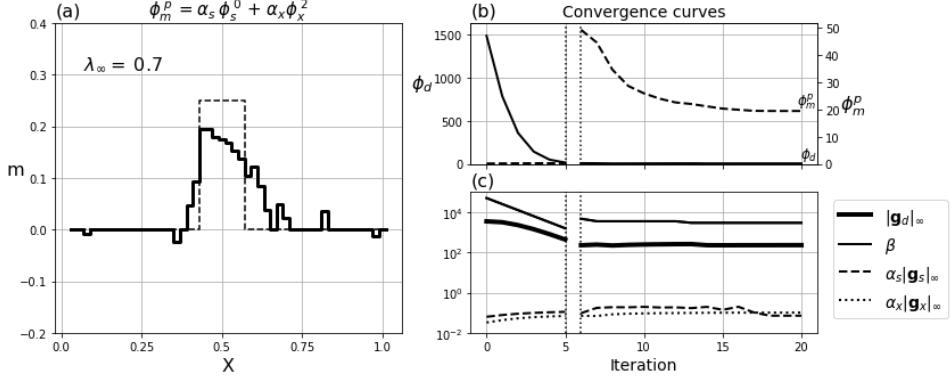


Figure 3.11: (a) Recovered model and (b) convergence curves using the Scaled-IRLS approach for $p_s = 0$, $p_x = 2$ and a fixed threshold parameter $\varepsilon = 1e - 3$ ($\alpha_s = \alpha_x = 1$). (c) Trade-off parameter and maximum gradients for the different components of the objective function. The scaling procedures preserves the proportionality between $\|\mathbf{g}_s^{p_s}\|_\infty$ and $\|\mathbf{g}_x^{p_x}\|_\infty$ throughout the iteration process. The trade-off β -parameter needed only to be adjusted slightly at the beginning of Stage 2.

such that the partial derivative of ϕ_x^p with respect to \mathbf{m} can be written as

$$\begin{aligned} \mathbf{g}_x^p &= \frac{\partial \phi_x^p}{\partial \mathbf{m}} = h^{-p_x} \hat{\mathbf{g}}_x^p \\ &= h^{-p_x} \mathbf{D}_x^\top \hat{\mathbf{R}}_x^\top \mathbf{V}_x^\top \mathbf{W}_x^\top \mathbf{W}_x \mathbf{V}_x \hat{\mathbf{R}}_x \mathbf{D}_x \mathbf{m} \end{aligned} \quad (3.44)$$

where

$$\mathbf{D}_x = \begin{bmatrix} -\hat{h}_1^{-1} & \hat{h}_1^{-1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & -\hat{h}_{M-1}^{-1} & \hat{h}_{M-1}^{-1} \end{bmatrix}. \quad (3.45)$$

measures the model derivatives over normalized length scales \hat{h}_i . The IRLS weights in (3.29) become

$$\begin{aligned} \hat{\mathbf{R}}_x &= \text{diag} [\hat{\mathbf{r}}_x]^{1/2} \\ \hat{r}_{x_i} &= \left[\left(\frac{m_{i+1}^{(k-1)} - m_i^{(k-1)}}{\hat{h}_i} \right)^2 + h^2 \varepsilon^2 \right]^{p_x/2-1}. \end{aligned} \quad (3.46)$$

Since the maximum of \mathbf{g}_x^p scales linearly with the base length scale h^{-p_x} :

$$\|\mathbf{g}_x^p\|_\infty = h^{-p} \|\hat{\mathbf{g}}_x^p\|_\infty , \quad (3.47)$$

I can express the γ_x -scaling as

$$\gamma_x = \left[h^{p-2} \frac{\|\hat{\mathbf{g}}_x^2\|_\infty}{\|\hat{\mathbf{g}}_x^p\|_\infty} \right]^{1/2} . \quad (3.48)$$

By simply factoring out the base length h , I recover a multiplication factor h^{p-2} that is closely related to the α_s scaling strategy specified in (3.36). If I use the γ_x -scaling in (3.48) and the constant α_s scaling prescribed in (3.8) I get that

$$\phi_m = h^{-2} \phi_s + h^{p-2} \phi_x . \quad (3.49)$$

Conversely, if I set the α parameters based on the strategy put forward in (3.36) I get that

$$\phi_m = h^{-p} \phi_s + \phi_x . \quad (3.50)$$

Expression (3.50) and (3.49) are related by a multiplication factor h^{p-2} . Since the partial derivatives of ϕ_m are invariant with respect to a global constant, I can expect to get the same solution with either approach, albeit a re-adjustment of the tradeoff β parameter.

Rather than choosing one approach over the other, and in order to simplify the definition of the hyper-parameters α , I propose to remove the constant factor h from the regularization function altogether. I can do this by directly evaluating the model derivatives with the finite difference operator described in (3.45). This simple change brings both $\phi_s^{p_s}$ and $\phi_x^{p_x}$ to be dimensionally equivalent such that

$$\frac{[\hat{\phi}_x^p]}{[\phi_s^p]} = 1 . \quad (3.51)$$

where $\hat{\phi}_x^p$ denotes the measure of model derivative using the finite difference ap-

proach. The scaled IRLS weights become

$$\hat{\mathbf{R}}_x = \gamma_x \operatorname{diag} \left[\left((\mathbf{D}_x \mathbf{m}^{(k-1)})^2 + \varepsilon^2 \right)^{p_x/2-1} \right]^{1/2}, \quad (3.52)$$

where

$$\gamma_x = \left[\frac{\|\hat{\mathbf{g}}_x^2\|_\infty}{\|\hat{\mathbf{g}}_x^{p_x}\|_\infty} \right]^{1/2}. \quad (3.53)$$

The Scaled-IRLS regularization becomes:

$$\phi_m^P = \alpha_s \|\mathbf{W}_s \mathbf{V}_s \hat{\mathbf{R}}_s \mathbf{m}\|_2^2 + \alpha_x \|\mathbf{W}_x \mathbf{V}_x \hat{\mathbf{R}}_x \mathbf{D}_x \mathbf{m}\|_2^2. \quad (3.54)$$

with default parameters $\alpha_s = \alpha_x = 1$.

In order to demonstrate that this change in length scale does not change the overall objective function, I proceed with two inversions. I repeat the experiment presented in Section 3.1.3 for $p_s = p_x = 1$, also shown in Figure 3.12(a) for comparison. First, I use the regularization function in (3.54) with uniform scaling with the same uniform discretization ($\alpha_s = \alpha_x = 1$). The recovered model shown in Figure 3.12(b) is almost identical to the previous solution. Small discrepancies between the two models can be attributed to slight differences in the iterative process. As presented in Table 3.3, the global objective function $\phi(m)$ remains unchanged with both approaches. Changes in scale between individual components of the objective function ($\phi_d, \phi_s^P, \phi_x^P$) are absorbed by their respective hyper-parameter ($\beta, \alpha_s, \alpha_x$).

The second experiment tests the case of a non-uniform discretization. Using the same noisy data, I refine the mesh on the right-half of the domain by dividing the cell size by a factor 2 ($h_i = 0.01$). I also re-adjust the kernel functions F_{ij} within the refined region such that each random coefficient is sampled twice but over a smaller cell length. The final inversion mesh contains 75 parameters, compared to 50 parameters in the previous experiment. After convergence of the algorithm I recover the model presented in Figure 3.12(c). Once again, the final model and the calculated components of the objective function (Table 3.3) are almost identical to the previous experiments. This demonstrates that the normalization of the model derivatives does not change the global objective function, and that the solution

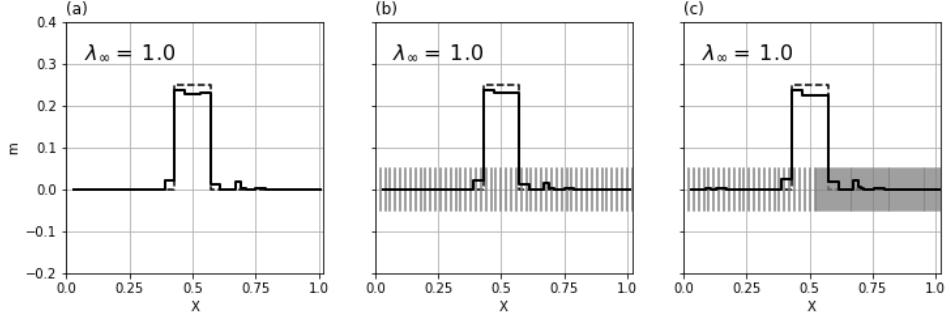


Figure 3.12: Recovered 1D models for $p_s = p_x = 1$ using different scaling strategies and parameterization. (a) Solution previously shown in Figure 3.7 that uses the standard gradient measure ($\alpha_s=50$). (b) Solution obtained with the finite difference approach ($\alpha_s = \alpha_x = 1$). (c) Model recovered with a different parameterization such that the right half of the domain has cells with half the size. Small discrepancies between the three solutions can be attributed to slight differences in the iterative process.

	ϕ_d	β	α_s	α_x	$\beta\alpha_s\phi_s$	$\beta\alpha_x\phi_x$
Figure 3.12(a)	4.87	65.13	50.0	1.0	56.18	8.60
Figure 3.12(b)	5.13	3381.87	1.0	1.0	58.33	8.81
Figure 3.12(c)	4.93	3236.60	1.0	1.0	55.81	8.59

Table 3.3: Components of the objective function corresponding to the inversion results presented in Figure 3.12 for $p_s = p_x = 1$.

remains independent on the choice of discretization. For clarity, I will use ϕ_x^P to denote the measure of model derivatives using the finite difference operator.

3.2.2 Threshold ε -parameter

While I have improved the flexibility of the IRLS algorithm, I have yet to address the threshold ε -parameter which has been held fixed. The choice of threshold parameters remains a subject of disagreement among researchers. In the early work of Last and Kubik (1983), it was suggested that the threshold value should be small or near machine error ($\varepsilon < 10^{-8}$) in order to best approximate the ℓ_p -norm. The same strategy was later adopted by others (Barbosa and Silva, 1994; Stocco et al., 2009). Other researchers, such as in Ajo-Franklin et al. (2007) observed

instabilities with small values, and opted for a wider range ($10^{-4} < \varepsilon < 10^{-7}$).

More recently, Sun and Li (2014) proposed an ε -search phase to highlight regions reacting favourably to the sparsity constraints. A final inversion step was then carried out with a fixed threshold value ($\varepsilon \ll 1e-4$). A similar strategy has also been proposed by Zhdanov and Tolstaya (2004) after selecting an optimal point on a trade-off curve.

Selecting an appropriate threshold value becomes more complicated when combining different penalty functions. I not only need to contend with the range of model values, but also with the relative influence of the components of a non-linear regularization function. To illustrate the challenge, I invert the 1D example again with the mixed norm penalty function ($\phi_m = \alpha_s \phi_s^0 + \alpha_x \phi_x^2$) but this time over a range of threshold values ($10^0 < \varepsilon < 10^{-5}$). The resulting models are shown in Figure 3.13. I identify the following trends:

- For large values ($\varepsilon > 10^{-1}$), no sparsity is achieved and the model resembles the solution previously obtained with $\phi_m = \alpha_s \phi_s^2 + \alpha_x \phi_x^2$.
- With small values ($\varepsilon < 10^{-4}$), $\phi_s^{p_s}$ appears to have little influence on the solution and the model resembles the solution obtained with smooth penalties $\alpha_s = 0$. The proportionality ratio $\lambda_\infty \ll 1$ confirms this bias towards ϕ_x^2 .
- The mid-range values ($\varepsilon^{-1} < \varepsilon < 10^{-3}$) show the most significant variability in the solutions with an achieved proportionality ratio $\lambda_\infty \approx 1$.

From this numerical experiment, there appears to be an optimal ε -parameter in the mid-range ($\varepsilon^{-1} < \varepsilon < 10^{-3}$) that can promote both a sparse and smooth solution. Ideally, I want to automate the selection process.

In this study, I opt for a cooling strategy. Threshold value ε is initialized at a large value then monotonically reduced such that:

$$\varepsilon^{(k)} = \frac{\|\mathbf{m}^{(0)}\|_\infty}{\eta^k}, \quad (3.55)$$

In (3.55), η is a user-defined cooling rate constant and $\|f_j(m)^{(0)}\|_\infty$ denotes the largest function value obtained at the end of Stage 1 of the algorithm. At the start of Stage 2, the Lawson approximation with large ε is effectively an ℓ_2 -norm.

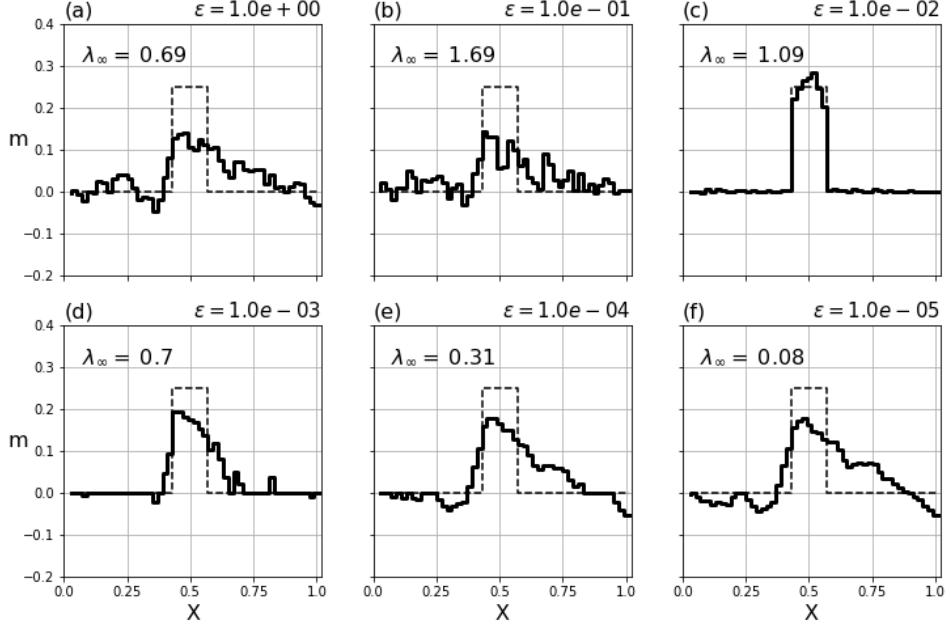


Figure 3.13: Recovered 1D models with variable threshold parameter on the range $10^{-5} < \varepsilon < 10^0$ using a mixed-norm penalty function $\phi_m = \alpha_s \phi_s^0 + \alpha_x \phi_x^2$.

Thus there is only a small change in regularization between Stages 1 and 2 of the algorithm. This is desired since the ℓ_p -norm regularization is highly non-linear and I want to reduce the risk of moving away from my initial proportionality conditions.

I proceed with an inversion with a cooling rate $\eta = 1.25$. Recovered models as a function of iterations are shown in Figure 3.14(a) to (d). As the number of iterations increases and $\varepsilon \rightarrow 0$, the emphasis of the sparse penalties ($\gamma_s^2 \mathbf{g}_s^0$) sweeps through the range of model values, progressively focusing on smaller model parameters. Figure 3.14(e) plots the scaled gradients as a function of absolute model values obtained at iteration $k=10, 15, 24$ and 55 . This plot can be compared to Figure 3.10(c). The gradients associated with sparse penalties (\mathbf{g}_s^0) force small model values ($m \approx \varepsilon$) towards the reference ($m^{ref} = 0$). Large model values ($m \gg \varepsilon$) are free to increase unless penalized by the other competing functions (ϕ_x^{Px} , ϕ_d). Figure 3.15 illustrates the evolution of penalties by plotting the partial derivatives of the objective function for iteration $k=15$ (early stage) and iteration $k=55$ (late stage). As $\varepsilon \rightarrow 0$, small model values are primarily influenced by $\frac{\partial \phi_d}{\partial \mathbf{m}}$ and $\beta \alpha_s \frac{\partial \phi_s}{\partial \mathbf{m}}$,

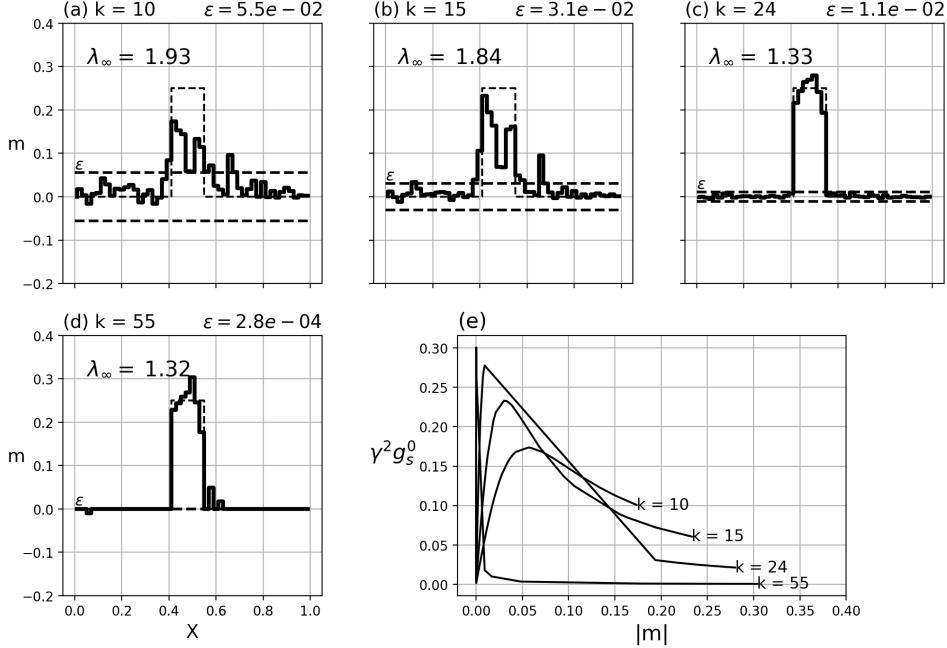


Figure 3.14: (a)-(d) Recovered 1D models at different iteration steps (k). The value of ε (dash) is shown for reference, highlighting the idea of a progressive thresholding of model values. (e) Scaled partial derivatives ($\gamma^2 \mathbf{g}_s^0$) as a function of the sorted model values and at various iteration stages.

while large model values are influenced by $\frac{\partial \phi_d}{\partial \mathbf{m}}$ and $\beta \alpha_x \frac{\partial \phi_x}{\partial \mathbf{m}}$.

From a user standpoint, the cooling strategy is attractive as it eliminates the requirement to predetermine an optimal ε threshold values and instead relies on a cooling rate η . To investigate the impact of the cooling rate on the solution, I solve the inverse problem ($p_s = 0, p_x = 2$) for various cooling rates $\eta = [6, 3, 1.5, 1.125]$. For each independent trial, the iteration process continues until reaching the additional criteria that $\varepsilon^{(k)} = 10^{-6}$ (near machine single precision). Recovered solutions for the 4 cooling rates are presented in Figure 3.16(a-d). I note important differences between the solutions. A summary of the inversions is provided in Table 3.4.

Figure 3.16(e) displays convergence curves for each inversion trial. The final norm ϕ_m^P is evaluated by using the expression (3.30), that is, all γ scaling parame-

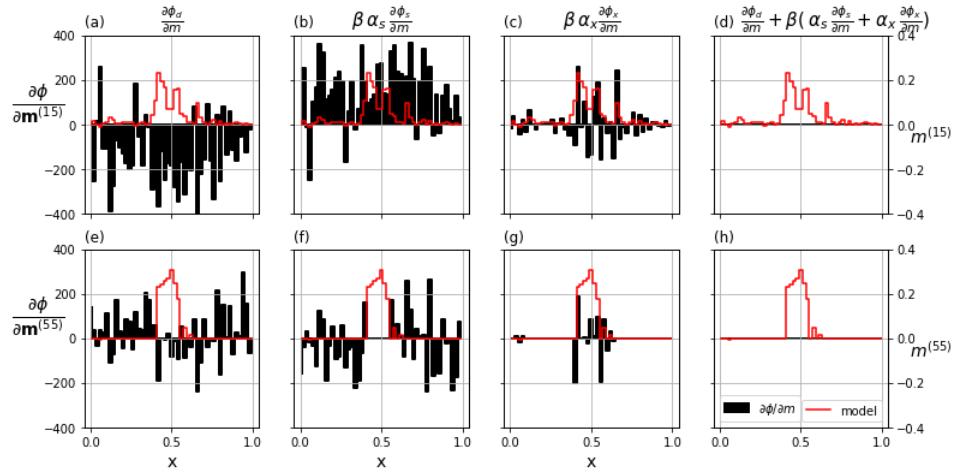


Figure 3.15: Partial derivatives of the objective function for iteration (top) $k=15$ and (bottom) $k=55$. The recovered models are shown in red for reference.

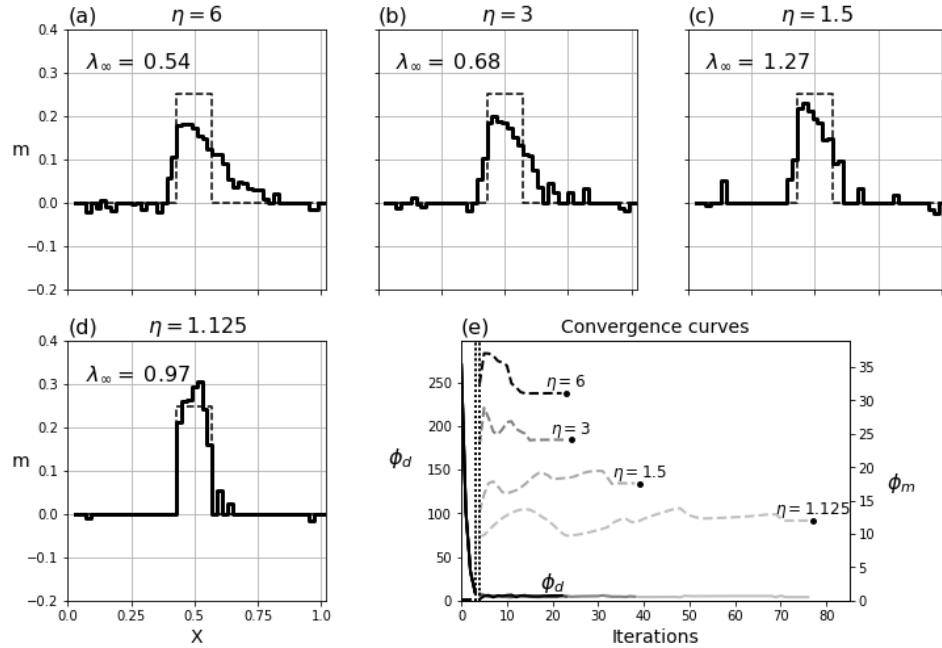


Figure 3.16: (a-d) Recovered models and (e) convergence curves for the minimization of $\phi_d + \beta\phi_m^P$ with various cooling rates η but with a final $\epsilon^* = 1e-6$.

η	# Iterations	ϕ_m^p	$\varepsilon^{(k)}$	$\phi_d^{(k)}$	λ
1.125	77	12.6	1e-06	9.2	0.97
1.5	39	17.8	1e-06	9.1	1.27
3	24	24.1	1e-06	8.5	0.68
6	23	31.1	1e-06	8.9	0.54

Table 3.4: Inversion summary after convergence of the S-IRLS for various cooling rates η as presented in Figure 3.16. Each inversion trial was required to reach the target misfit ϕ_d^* and target $\varepsilon^* = 1e - 6$.

ters have been removed. Thus I am comparing my ability to minimize the global objective function even though the algorithm has used variable scalings to reach that result. This promotes insight into the robustness of the final model as a function of the cooling rate. I note that a lower model norm can be obtained by cooling slowly. Cooling at an even slower rate (not shown here) has reaffirmed this. A rate of $\eta = 1.025$ yielded $\phi_m = 12.5$ in 155 iterations. It appears that $\phi_m^p \rightarrow 12.5$ as $\eta \rightarrow 1$. I found experimentally that for $\eta \approx 1.25$ generally yielded an optimal trade-off between computational time (number of iterations) and convergence to a suitable solution.

3.2.3 Summary

My goal is to solve an inverse problem where the regularization function is composed of multiple terms, each defined as an ℓ_p -norm premultiplied by a scaling parameter. The scaling parameters, α 's, are used to control how much each component contributes to the final solution. The relative influence of these components is quantified by evaluating the proportionality ratio λ_∞ (3.19). If two components contribute equally, then λ_∞ should be close to unity. Unfortunately, when the components of the regularization include model and gradient terms, the scaling is affected by the cell size chosen for discretization. To simplify the implementation I normalize the length scales used in the measure of model derivatives. This makes $\phi_s^{p_s}$ and $\phi_x^{p_x}$ dimensionally equivalent. The default values for obtaining equal contributions are thus $\alpha_s = \alpha_x = 1$ for all combinations of ℓ_p -norms on the gradients.

I solve the inverse problem by replacing the ℓ_p -norms with their Lawson norm

approximations. Thus I search for the model that minimizes

$$\begin{aligned}\phi(m) = & \|\mathbf{W}_d \Delta \mathbf{d}\|_2^2 + \\ & \beta \alpha_s \sum_{i=1}^M w_{si} v_{si} \frac{m_i^2}{(m_i^2 + \varepsilon^2)^{1-p_s/2}} + \\ & \beta \alpha_x \sum_{i=1}^{M-1} w_{xi} v_{xi} \frac{\left(\frac{m_{i+1}-m_i}{\hat{h}_i}\right)^2}{\left[\left(\frac{m_{i+1}-m_i}{\hat{h}_i}\right)^2 + \varepsilon^2\right]^{1-p_x/2}} V_i ,\end{aligned}\quad (3.56)$$

for arbitrarily small ε value. I solve the inverse problem using a two-stage approach. I first find a solution for the ℓ_2 -norm problem and then I change the objective functions to their final desired ℓ_p -norms and solve the optimization problem using IRLS. To keep stability in the iterative process I successively rescale the IRLS weights \mathbf{R} . Thus at each iteration I solve a locally convex problem

$$\begin{aligned}\phi(m^{(k)}) = & \|\mathbf{W}_d \Delta \mathbf{d}\|_2^2 + \\ & \beta (\alpha_s \|\mathbf{W}_s \mathbf{V}_s \hat{\mathbf{R}}_s \mathbf{m}\|_2^2 + \alpha_x \|\mathbf{W}_x \mathbf{V}_x \hat{\mathbf{R}}_x \mathbf{D}_x \mathbf{m}\|_2^2) ,\end{aligned}\quad (3.57)$$

Although the local minimization problems involve scaled gradients, the final desired solution is that which minimizes (3.56) such that all components contribute equally. My ability to achieve this goal depends upon the value of ε and the chosen cooling rate. I find that the best (i.e. minimum norm) solution is obtained when ε is cooled slowly to a final small value. If the cooling is too fast then I obtain a substandard solution in which λ_∞ is not close to unity and my modelling objectives are not satisfied. Slower cooling and the frequent re-scaling of the gradients keeps the proportionality ratio near unity.

3.3 Exploring the model space

The smooth density model presented in Figure 3.2 was a poor approximation of a compact block, but it is one of many possible solutions. Now that I have developed an algorithm that can combine multiple regularization functions with different ℓ_p -norm measure, I can explore the model space by generating a suite of solutions that have variable characteristics. I will demonstrate this on the synthetic gravity

example shown in Figure 3.1. The function to be minimized for the 3D gravity problem becomes

$$\begin{aligned} \min_{\rho} \phi(\rho) = & \| \mathbf{G} \rho - \mathbf{d}^{obs} \|_2^2 + \beta \sum_{r=s,x,y,z} \alpha_r \| \mathbf{W}_r \mathbf{V}_r \hat{\mathbf{R}}_r \mathbf{D}_r \rho \|_2^2 \\ \text{s.t. } & \phi_d \leq \phi_d^* \end{aligned} \quad (3.58)$$

I carry out eight additional inversions. I use a combination of norms on a range of $p_s, p_{[x,y,z]} \in [0, 1, 2]$ values. I set $p_x = p_y = p_z$ in all cases. The solutions, nine models in total, are presented in Figure 3.17. All models have a final misfit $\phi_d^* \approx 441$ and use the same ℓ_2 -norm solution to initiate the IRLS steps. I make the following general observations. There is a progressive transition from a smooth model (upper left) to a blocky solution (lower right) as p_s, p_x, p_y and p_z decrease. The top of the density high is most often recovered at 10 m depth. Away from the anomalous region the density is relatively smooth and close to the background reference model of 0 g/cc. There is also a clear trend in the data misfit map such that the correlated residual decreases as $p_s, p_x, p_y, p_z \rightarrow 0$.

3.3.1 Interpretation

Accessing a range of solutions is important to assess the stability of different features and to avoid over-interpreting one specific realization. The next step requires to compare this ensemble of models and make a geological interpretation. In Chapter 6 I provide a more evolved methodology to extract local parameters, but for now, I will compare the solutions visually. Figure 3.18 presents an overlay for the 10th and 90th percentiles anomalous densities calculated from the suite of models shown in Figure 3.17. I can assess the robustness of features by comparing the iso-contour lines of each model: tight clustering of the contours indicates that several models agree on the position of an edge, while a large spread indicates high variability. At the center of the model, I note that the top of the anomaly (solid) is highly correlated among models, but less so the bottom limit. This is expected as the resolution of the survey decreases with depth. Meanwhile, on the edges of the domain, the shape and sign of density contrasts vary substantially. From this simple analysis, I would assign high confidence on the horizontal and top of

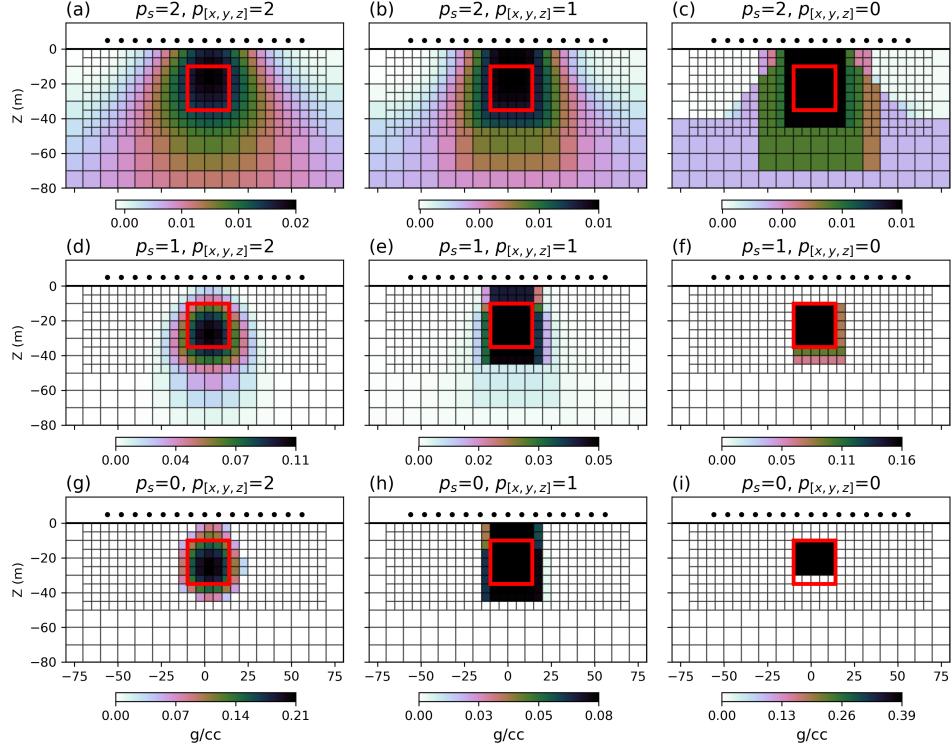


Figure 3.17: (a-i) Vertical section through a suite of density models recovered for varying ℓ_p -norm penalties applied on the model and model gradients for $p_s \in [0, 1, 2]$ and $p_x = p_y = p_z \in [0, 1, 2]$.

the positive density anomaly, and low confidence on features on either side of the inversion domain.

The normalized data residual maps for each inversion are shown in Figure 3.19. The decrease in correlated residual observed on the misfit maps (Figure 3.18) is also an important aspect to consider. While all the inversions have achieved the global target misfit, only after applying the proper constraints (blocky and compact) that the inversion was able to predict the short wavelength information present in the data. This is an important aspect of this research as it further stresses the importance of exploring a range of solutions with broadly different characteristics such that more subtle features can be extracted. It is also a motivation to automate the search for suitable inversion parameters that can better resolve the data.

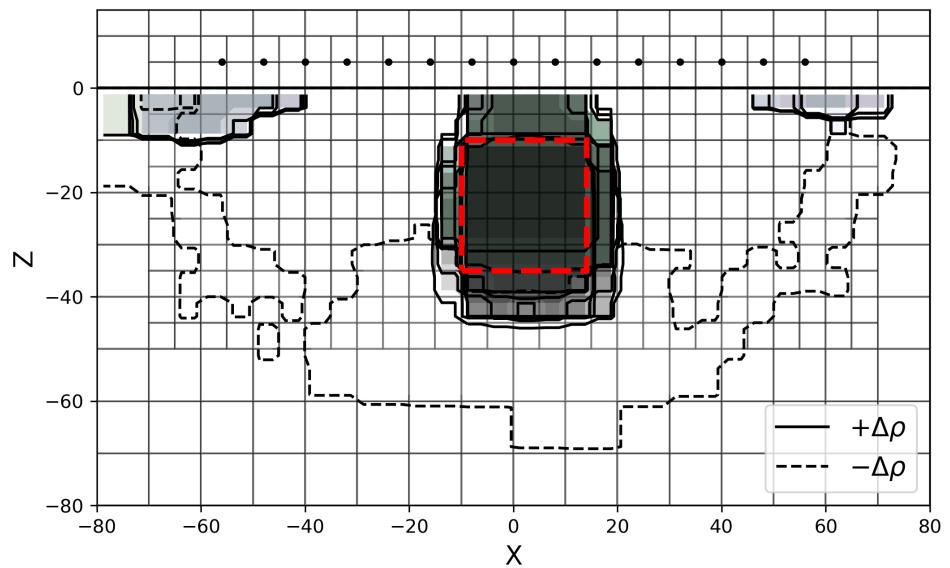


Figure 3.18: Iso-contour values for the 10th and 90th percentile of anomalous density calculated from the suite of models shown in Figure 3.17. The outline of the target (red) is shown for reference. Contour lines tightly clustered indicate coherence between inversion trials. Negative anomalies (dash) appear to change significantly, to which I would assign lower confidence.

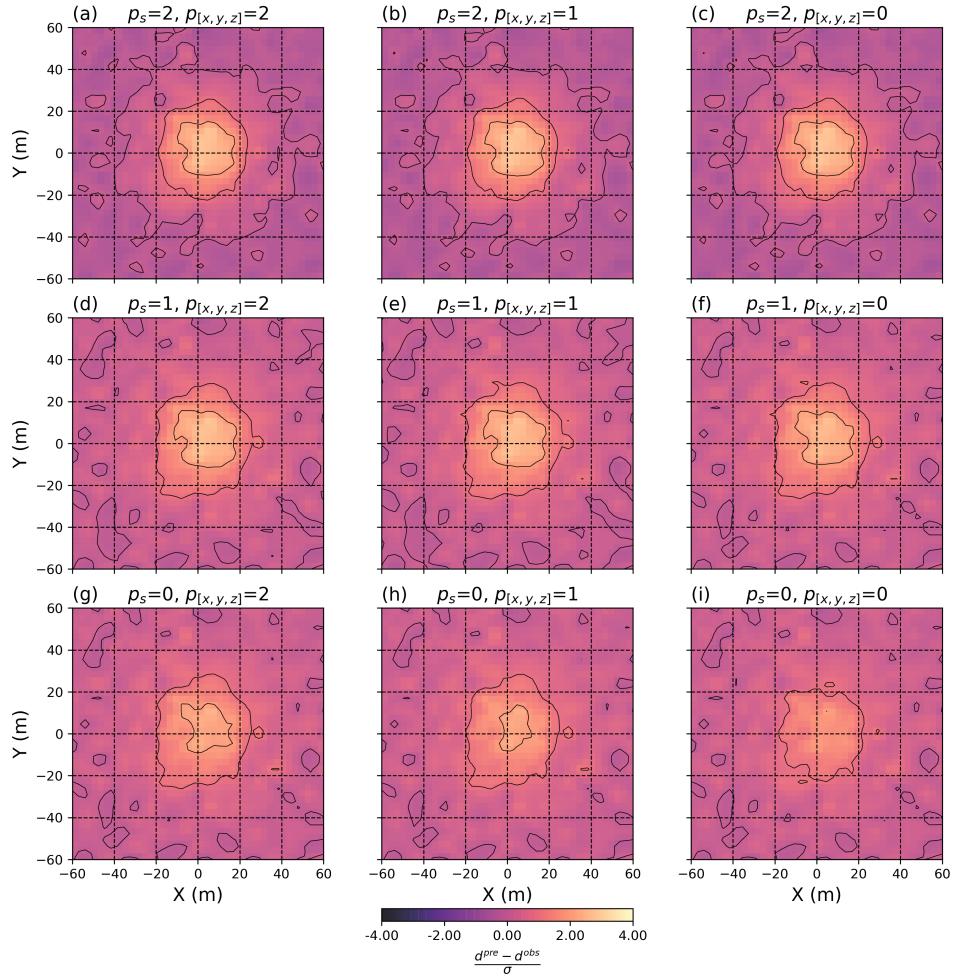


Figure 3.19: (a-i) Residual data map calculated from the suite of density models for varying ℓ_p -norm penalties applied on the model and model gradients for $p_s \in [0, 1, 2]$ and $p_x = p_y = p_z \in [0, 1, 2]$.

Chapter 4

Sparse Magnetic Vector Inversion

I have so far showcased the benefits of employing a mixed-norm inversion to solve the gravity problem. In this chapter, I review and improve inversion methods for the modeling of magnetization parameters. Contrary to the scalar density, magnetization is a vector quantity defined by a magnitude and an orientation as previously defined in (1.8)

$$\vec{M} = \kappa(\vec{H}_0 + \vec{H}_s) + \vec{M}_r . \quad (4.1)$$

Researchers have investigated ways to extract information about magnetization with inversion methods, but this process remains challenging as there are three times the number of unknowns over conventional problems. To simplify the inverse problem, the assumption is often made that Earth's inducing field is much larger than the other component such that secondary fields and the presence of remanence are ignored ($\vec{M}_r = \vec{H}_s = 0$). This assumption has dominated the inversion literature over the past 30 years (Li and Oldenburg, 1996; Pilkington, 1997). Under this assumption, the definition of magnetization (1.8) simplifies to

$$\vec{M} = \kappa\vec{H}_0 ,$$

which gives rise to a linear system relating N data \mathbf{d}^{pre} to the M discrete model cells of magnetic susceptibility κ

$$\begin{aligned}\mathbf{d}^{pre} &= \mathbf{F} \kappa \\ \mathbf{d}^{pre} &\in \mathbb{R}^N, \mathbf{F} \in \mathbb{R}^{N \times M}, \kappa \in \mathbb{R}^M.\end{aligned}\tag{4.2}$$

To illustrate potential issues with this induced assumption I revisit the synthetic block example in terms of magnetic properties. I set the magnetic susceptibility (κ) of the block to 0.035 SI and the vertical inducing flux to \vec{B}_0 [50,000 nT, I: 90°, D: 0°]. I also add a remanent component equal in magnitude and pointing along the x-axis \vec{M}_r [1.4 A/m, I: 0°, D: 90°]. This results in a total magnetization \vec{M} [2.0 A/m, I: 45°, D: 90°]. Using the linear relationship presented in (2.10), I can simulate magnetic data presented in 4.1(b) on which I add random Gaussian noise of 1 nT standard deviation to simulate field conditions. I will attempt to recover the magnetized block from the noisy data shown in Figure 4.1(c).

I begin with the conventional smooth assumptions ($p_s, p_x, p_y, p_z = 2$) and attempt to recover the position and shape of the magnetic block. Since (4.2) is linear with respect to κ , I can use the same inversion methodology established in Chapter 3. The objective function to be minimized becomes

$$\begin{aligned}\min_{\kappa} \phi(\kappa) &= \|\mathbf{F} \kappa - \mathbf{d}^{obs}\|_2^2 + \beta \sum_{r=s,x,y,z} \alpha_r \|\mathbf{W}_r \mathbf{V}_r \mathbf{R}_r \mathbf{D}_r \kappa\|_2^2 \\ \text{s.t. } \phi_d &\leq \phi_d^*\end{aligned}\tag{4.3}$$

As I am dealing with strictly positive magnetic susceptibility κ , I impose bound constraints by the projected gradient method (Vogel, 2002). Model parameters κ_i that become negative are set to zero and ignored for the following Gauss-Newton step in (3.32). After reaching the target misfit criterion in equation (3.33), I recover the susceptibility model shown in Figure 4.2(a). I note that the position of the susceptibility anomaly is shifted to the side of the true block and appears to dip at 45° angle. This is due to the large negative data lobe introduced by the remanent component that I have purposefully ignored. Attempting to improve the solution by solving for a sparse model ($p_s, p_x, p_y, p_z = 0$) yields the solution presented in Figure 4.2(c). The magnetic anomaly is imaged at the right depth and

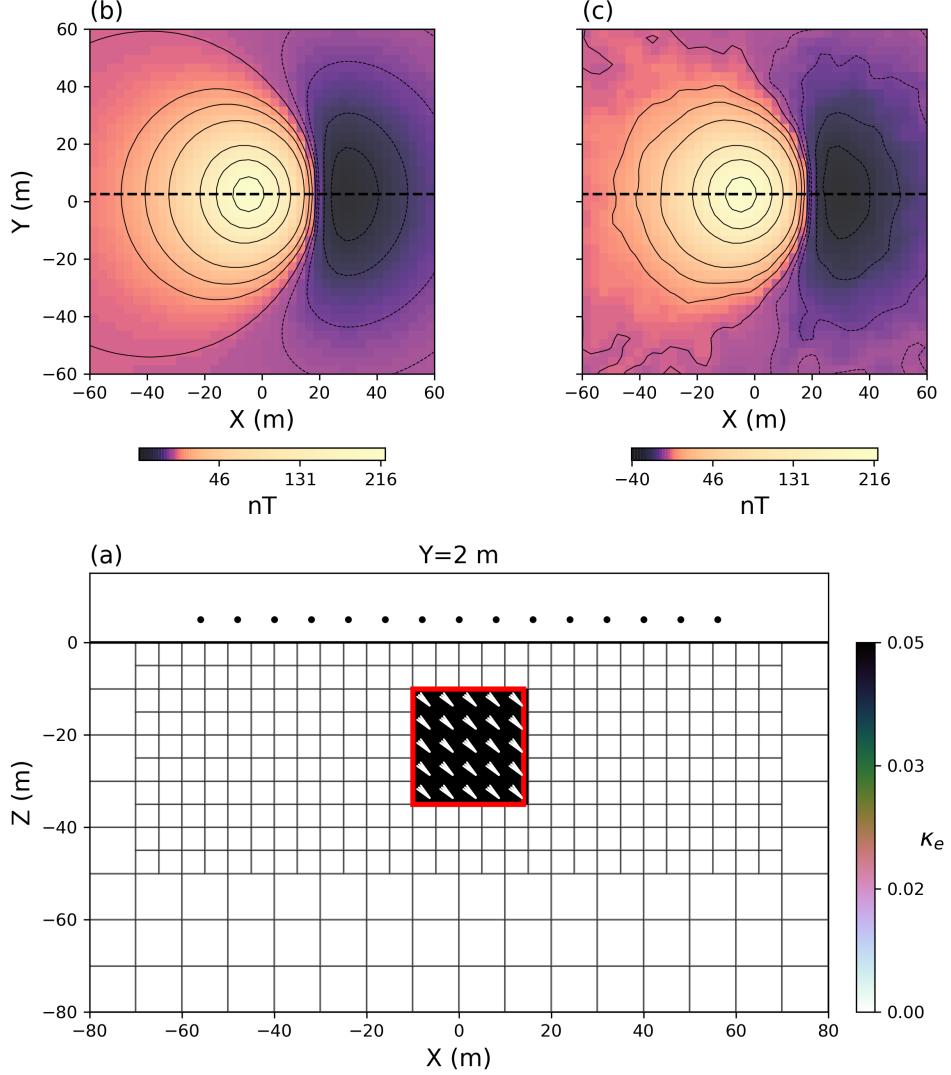


Figure 4.1: (a) Vertical section through a 25 m cube with uniform magnetization \vec{M} [2.0 A/m, I: 45°, D: 90°]. (b) Simulated TMA data response on a 21 × 21 survey grid placed 15 m above the anomaly. (c) Magnetic data with random Gaussian noise added, 1 nT standard deviation.

the vertical extent is better recovered, but position and shape of the anomaly have not improved. It is also important to note the correlated negative data residuals in Figure 4.2(b) and (d). The inversion struggled to reproduce the negative portion of

the magnetic data using strictly positive susceptibility values.

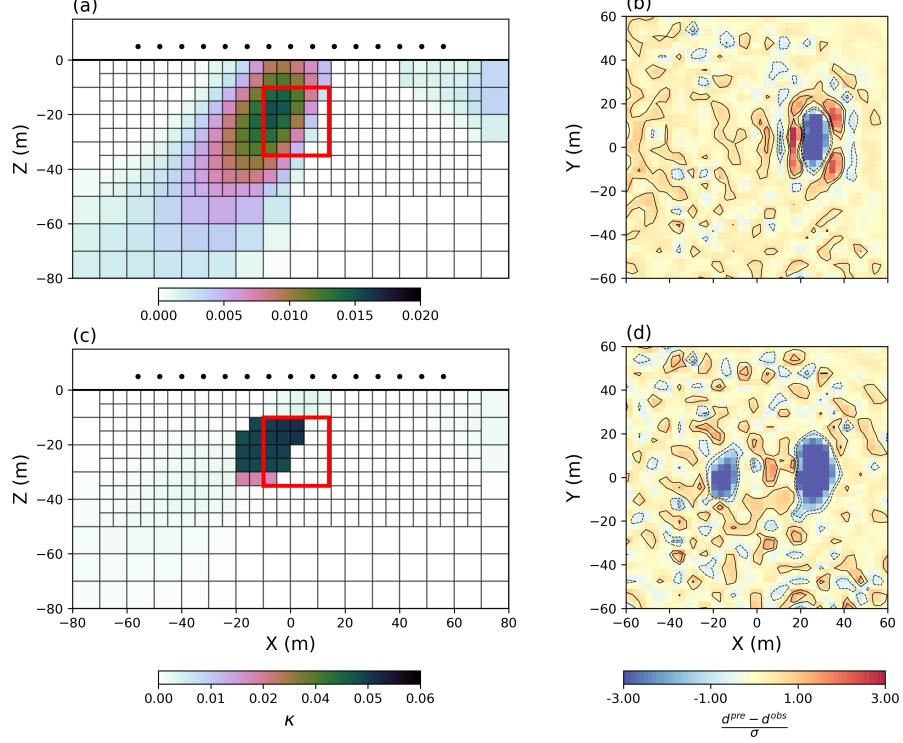


Figure 4.2: Vertical section through the recovered susceptibility model using (a) smooth assumption ($p_s, p_x, p_y, p_z = 2$) and (b) sparse ℓ_p -norms to recover a compact model ($p_s, p_x, p_y, p_z = 0$). Both solutions show an anomaly with a false dip due to the wrong assumption of a vertical magnetization.

The presence of remanence has long been recognized as an obstacle for the geological interpretation of magnetic data. Several types of mineral deposits are associated with remanent magnetization such as diamondiferous kimberlites, volcanic massive sulphides and porphyries (Enkin, 2014; Henkel, 1991). In a mining exploration context, having the wrong image could result in false drilling targets: costly both in time, resources and confidence in geophysical methods. Hence the need for a more robust algorithm that does not require knowledge about the orientation of magnetization. In the following sections, I revisit the Magnetic Vector Inversion introduced in the Ph.D. thesis of Lelièvre (2009). I address numerical

limitations encountered with the spherical formulation. I also leverage advancements made in Chapter 3 and impose sparsity penalties on the magnetization direction and amplitude independently. This allows to recover compact bodies with coherent magnetization direction.

4.1 Magnetic Vector Inversion - Cartesian parameters

Rather than assuming a purely susceptible response, Lelievre and Oldenburg (2009) proposed a strategy to directly invert for the magnetization vector \vec{M} . Re-writing the discrete system (4.2) in terms of vector magnetization:

$$\begin{aligned} \mathbf{d}^{pre} &= \mathbf{F}_C \mathbf{m}_C \\ &= [\mathbf{F}_u \mathbf{F}_v \mathbf{F}_w] \begin{bmatrix} \kappa_u \\ \kappa_v \\ \kappa_w \end{bmatrix} \\ \mathbf{F}_u, \mathbf{F}_v, \mathbf{F}_w &\in \mathbb{R}^{N \times M} \end{aligned} \quad (4.4)$$

where the model $\mathbf{m}_C = [\kappa_u, \kappa_v, \kappa_w]^\top$ describes the strength of magnetization along the Cartesian directions in terms of an *effective susceptibility* parameter,

$$\kappa_e = \frac{\vec{M}}{\|\vec{H}_0\|}, \quad (4.5)$$

The effective susceptibility scales the amplitude of magnetization with respect to the inducing field strength $\|\vec{H}_0\|$. The forward relationship (4.4) is still a linear system of equations but it has three times the number of unknown parameters compared to the susceptibility assumption ($\mathbf{m}_C \in \mathbb{R}^{3M}$). The objective function (3.58) becomes

$$\min_{\mathbf{m}_C} \phi(\mathbf{m}_C) = \|\mathbf{F}_C \mathbf{m}_C - \mathbf{d}^{obs}\|_2^2 + \beta \sum_{c=u,v,w} \sum_{r=s,x,y,z} \alpha_{cr} \|\mathbf{W}_{cr} \mathbf{V}_{cr} \mathbf{R}_{cr} \mathbf{D}_{cr} \mathbf{P}_c \mathbf{m}_C\|_2^2, \quad (4.6)$$

where the projection matrices \mathbf{P}_c select individual Cartesian component of the vector model \mathbf{m}_C . The regularization function is made up of twelve terms. Different norm measures can be applied to each Cartesian component independently.

Keeping the same inversion methodology and smooth assumptions (p_{cs}, p_{cx} ,

$p_{c_y}, p_{c_z} = 2$), I recover the magnetization model presented in Figure 4.3(a). This solution is an improvement over the susceptibility model; the bulk magnetization is recovered at the right position. I note however that the solution is distributed over a large volume and there is a broad distribution in magnetization direction inside and around the block.

In order to reduce the complexity of the solution, I once again resort to ℓ_p -norm measures ($p_{c_s}, p_{c_x}, p_{c_y}, p_{c_z} = 0$). As shown in Figure 4.3(b), the recovery of the block has clearly improved. It is important to point out however that the magnetization vectors are pointing along the Cartesian directions and the final anomaly is slightly wider than the true model. In the Cartesian formulation, both the direction and strength of magnetizations are coupled in the vector components and therefore the method lacks flexibility in recovering sparse vector along arbitrary orientations. This was the main motivation behind recent research investigating advanced regularization methodologies and cooperative approaches (Fournier, 2015; Liu et al., 2015; Zhu et al., 2015). I will attempt to improve on this solution by decoupling the strength and direction of the magnetization vector with the spherical formulation.

4.2 Magnetic Vector Inversion - Spherical parameters

As an alternative to the Cartesian formulation, Lelievre and Oldenburg (2009) also proposed the vector inversion in a spherical coordinate system. The conversion between Cartesian to spherical system follows the relation:

$$\begin{aligned}\kappa_u &= \rho \cos(\theta) \cos(\phi) \\ \kappa_v &= \rho \cos(\theta) \sin(\phi) \\ \kappa_w &= \rho \sin(\theta)\end{aligned}\tag{4.7}$$

where the magnetization vector is defined by parameters of amplitude (ρ) and two angles (θ, ϕ). Performing the forward calculations using spherical parameters \mathbf{m}_S can be expressed as

$$\mathbb{F}[\mathbf{m}_S] = \mathbf{F}_C \begin{bmatrix} \rho \cos(\theta) \cos(\phi) \\ \rho \cos(\theta) \sin(\phi) \\ \rho \sin(\theta) \end{bmatrix}\tag{4.8}$$

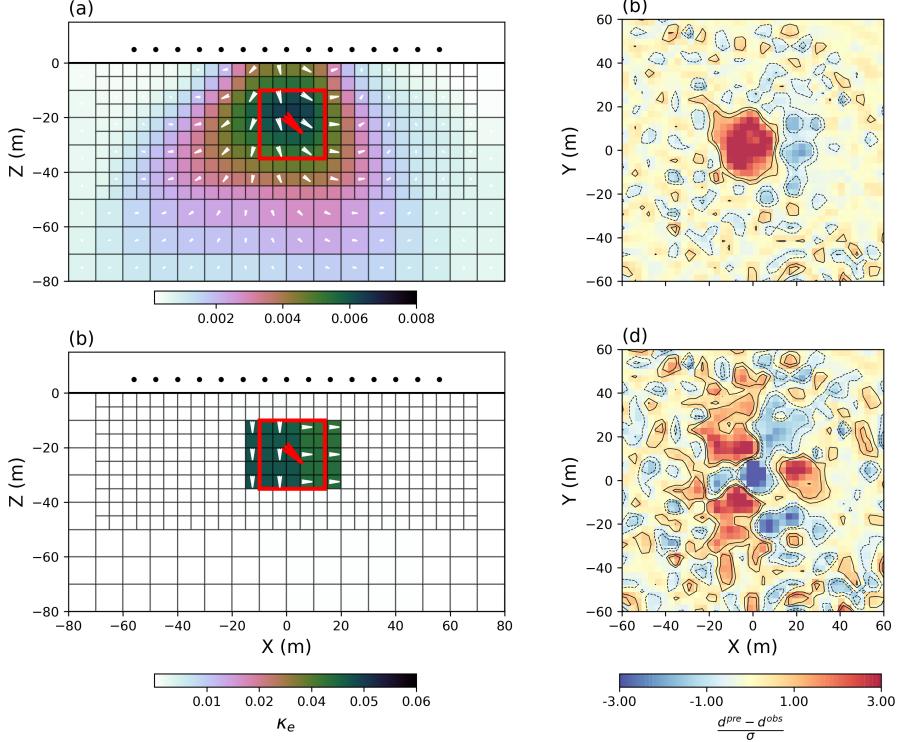


Figure 4.3: Vertical section through the recovered magnetization vector model using the Cartesian formulation with (a) smooth l_2 -norm assumption and (b) sparsity constraints applied on all three Cartesian components ($p_{i_s}, p_{i_x}, p_{i_y}, p_{i_z} = 0$). Corresponding normalized data residuals are shown in (b) and (d). The true position and magnetization orientation of the block are shown in red for reference.

The spherical formulation separates the magnitude and orientation of magnetization vector; this has two advantages. First, physical property constraints can easily be incorporated in the inversion. Magnetization measurements performed on rock samples are often provided in term of susceptibility κ and Koenigsberger ratio Q

$$Q = \frac{\vec{M}_r}{\|\kappa \vec{H}\|} \quad (4.9)$$

measuring the ratio between the induced and remanent component of magnetization. This quantity does not contain directionality but it provides information about

the amplitude of magnetization. Magnetization direction may also be used in the inversion when laboratory measurements are performed on oriented cores. Reference angles can then be introduced as constraints. The second advantage of the spherical coordinate system is that it allows for sparsity assumptions to be imposed on the magnitude and orientation independently. This has the potential of resolving compact bodies with uniform magnetization direction in any orientation not restricted to the Cartesian axes.

Despite its obvious advantages, the MVI-S method has received little attention in the literature because the non-linear transformation between Cartesian and spherical coordinates greatly complicates the inverse problem. I demonstrate challenges encountered with the spherical coordinate system on my synthetic problem. Taking the partial derivatives of (4.4) as a function of $\mathbf{m}_S(\rho, \theta, \phi)$ yields:

$$\mathbf{J} = \frac{\partial \mathbb{F}[\mathbf{m}_S]}{\partial \mathbf{m}_S} = \frac{\partial \mathbb{F}[\mathbf{m}_S]}{\partial \mathbf{m}_C} \frac{\partial \mathbf{m}_C}{\partial \mathbf{m}_S} \quad (4.10)$$

where $\frac{\partial \mathbf{m}_C}{\partial \mathbf{m}_S}$ involves partial derivatives of trigonometric functions prescribed in (4.7). The sensitivity matrix can be linearized before each Gauss-Newton step as:

$$\mathbf{J} = \mathbf{F}_C \mathbf{S} \quad (4.11)$$

where the matrix \mathbf{S} holds the partial derivatives

$$\mathbf{S} = \begin{bmatrix} \cos \theta \cos \phi & -\rho \sin \theta \cos \phi & -\rho \cos \theta \sin \phi \\ \cos \theta \sin \phi & -\rho \sin \theta \sin \phi & \rho \cos \theta \cos \phi \\ \sin \theta & \rho \cos \theta & 0 \end{bmatrix} \quad (4.12)$$

The regularization function becomes

$$\phi_m = \sum_{c=\rho, \theta, \phi} \sum_{r=s, x, y, z} \alpha_{c_r} \|\mathbf{W}_{c_r} \mathbf{V}_{c_r} \mathbf{R}_{c_r} \mathbf{D}_{c_r} \mathbf{P}_c \mathbf{m}_S\|_2^2, \quad (4.13)$$

I also take three additional precautions to deal with the spherical parameterization. First, a zero reference value θ_{ref} and ϕ_{ref} would imply a magnetization direction pointing along the x -axis. In the absence of physical constraints, and since I do not want to assume a specific orientation, I set $\alpha_{\theta_s} = \alpha_{\phi_s} = 0$ in all my experiments.

Thus the regularization only penalizes the change in angle between neighboring cells. Secondly, I perform an adjustment to the measure of angle differences to deal with the discontinuity at $\theta = -\pi, \pi$ in order to prevent over penalizing angles that describe a similar orientation along the x -axis. I convert the difference in angles between adjacent cells to the coterminal angles as shown in Figure 4.4. Thirdly, rather than resorting to a projected gradient approach, I proceed with a double transformation ($\mathbf{m}_S \rightarrow \mathbf{m}_C \rightarrow \mathbf{m}_S$) such that spherical parameter remain bounded as $\rho \in [0, \infty]$, $\theta \in [-\pi, \pi]$ and $\phi \in [-\pi/2, \pi/2]$.

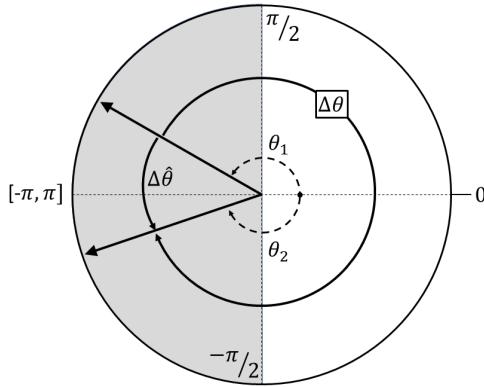


Figure 4.4: Measure of angle differences $\Delta\theta$ converted to coterminal angle $\Delta\hat{\theta}$.

To demonstrate the difficulties encountered with the MVI-S formulation, I invert the synthetic TMI data with a starting magnetization model pointing upward $\mathbf{m}_S^{(0)}(\rho = 10^{-2}, \theta = -45^\circ, \phi = 0^\circ)$ as shown in Figure 4.5(a). This represents a *worst-case* scenario such that the assumed magnetization orientation is at 90° from the true orientation. After convergence of the algorithm, I recover the model shown in Figure 4.5(b). The solution is a poor representation of the true magnetization. Model updates were forced to stop before reaching the target data misfit as the inversion was likely trapped in a local minimum. I note that most of the model updates were performed on the amplitude ρ , with only marginal changes on the angle of magnetization. Similar behaviors have been documented by Lelievre and Oldenburg (2009) and later by Liu et al. (2017). Poor convergence was attributed to an imbalance between the model parameters. Before attempting to implement more advanced constraints, I make inroads in improving the convergence of the

non-linear MVI-S formulation.

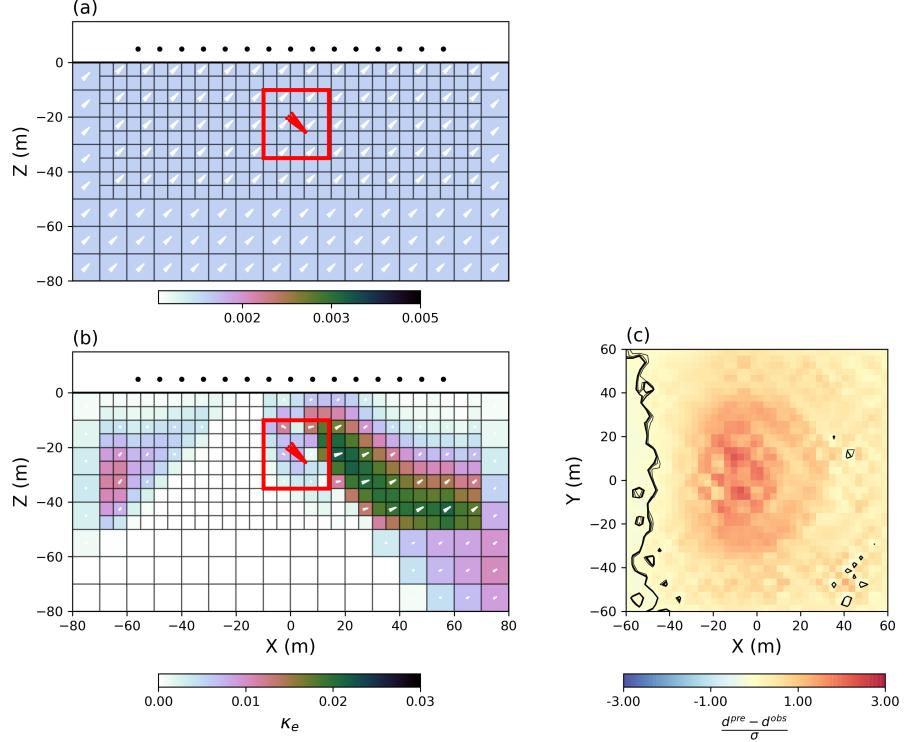


Figure 4.5: Vertical section through the (a) starting model and (b) recovered magnetization vector model in Spherical coordinates. The true position and magnetization orientation of the block are shown in red for reference. (c) Normalized data residuals show correlated signal. The inversion stopped after 15 iterations and was enabled to further reduce the objective function.

To gain some insight about the issues encountered with the MVI-S formulation, I consider a simpler two-parameter linear problem of the form

$$m_x + 2 * m_y = 1 , \quad (4.14)$$

which I can express in matrix form as

$$\begin{aligned} \mathbf{F}_C \mathbf{m}_C &= \mathbf{d}^{obs} , \\ \mathbf{F}_C = [1 \ 2], \ \mathbf{m}_C = \begin{bmatrix} m_x \\ m_y \end{bmatrix}, \ \mathbf{d}^{obs} &= 1 \end{aligned} \quad (4.15)$$

This defines an under-determined linear system of equations. Just as I did for the magnetic inverse problem, I can isolate a solution by minimizing an objective function of the form

$$\phi(m) = \|\mathbf{F}_C \mathbf{m}_C - \mathbf{d}^{obs}\|_2^2 + \beta_C \|\mathbf{m}_C\|_2^2, \quad (4.16)$$

Figure 4.6(a) displays a contour map of the objective function along with its gradients. Following the same methodology as in (3.9), I find a solution such that the gradient of the objective function $\phi(m)$ vanish

$$\frac{\partial \phi}{\partial \mathbf{m}} = \mathbf{g} = (\mathbf{F}_C^\top \mathbf{F}_C + \beta_C \mathbf{I}) \mathbf{m}_C - \mathbf{F}_C^\top \mathbf{d}^{obs} = \mathbf{0} \quad (4.17)$$

where \mathbf{I} is the identity matrix. The factor 2 from the derivative of the ℓ_2 -norm is absorbed by the zero right-hand side. After determining a trade-off parameter β_C such that $\phi_d \leq 1e-3$, I recover the Cartesian model $\mathbf{m}_C = [0.2, 0.4]$. It is the solution that minimizes the distance (evaluated with the ℓ_2 -norm) between the origin and the solution space of \mathbf{F} . I note that the relative magnitudes of model parameters in \mathbf{m}_C reflect the size of the forward coefficients in \mathbf{F} .

As previously discussed for the gravity and magnetic problems, the smallest solution is often not satisfactory as it is strongly influenced by the physics of the experiment. From equation (3.11) in Chapter 3, I can introduce sensitivity based weights to counteract this bias towards a large m_y value

$$\begin{aligned} \mathbf{W}_C &= \text{diag} \left[\left[\frac{\mathbf{w}_C}{\max(\mathbf{w}_C)} \right]^{1/2} \right] \\ w_{Cj} &= \left[\sum_{i=1}^N F_{ij}^2 \right]^{1/2}, \end{aligned} \quad (4.18)$$

where \mathbf{W}_C holds sensitivity weights added to the regularization ($\mathbf{w}_C = [1, 2]^\top$). The new weighted objective function becomes

$$\phi(m) = \|\mathbf{F}_C \mathbf{m}_C - \mathbf{d}^{obs}\|_2^2 + \beta_C \|\mathbf{W}_C \mathbf{m}_C\|_2^2, \quad (4.19)$$

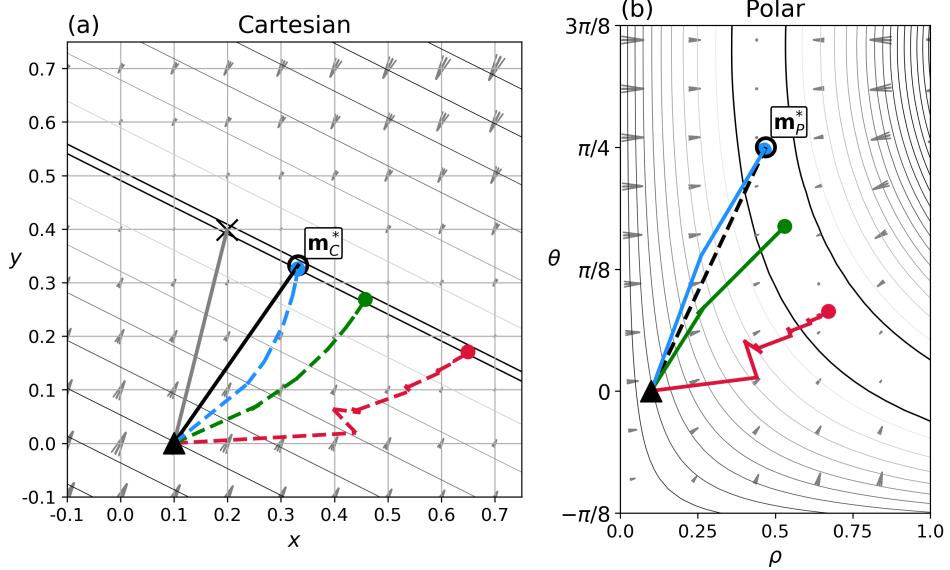


Figure 4.6: Contour map for two objective functions and their gradients (arrows) for a two-parameter inverse problem solved in Cartesian and polar coordinate systems. (a) The non-weighted (gray) problem yields a solution $\mathbf{m}_C = [0.2, 0.4]$ that reflects the size of the forward coefficients. The sensitivity weighted (black) solution is more desirable as it predicts the data with equal model parameters $\mathbf{m}_C^* = [0.33, 0.33]$. (b) Inversion steps performed in the non-linear polar coordinate system (solid red) are highly oscillatory and do not reach the optimal solution $\mathbf{m}_P^* = [0.47, 0.79]$. The same model updates are shown in Cartesian coordinates (red dash) for reference. (green) Inversion steps performed in polar coordinates with iterative sensitivity re-weighting and (blue) with the added Ω -scaling.

and a weighted gradient

$$\mathbf{g}_C = \mathbf{F}_C^\top \mathbf{F}_C \mathbf{m}_C + \beta_C \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_C - \mathbf{F}_C^\top \mathbf{d}^{obs} \quad (4.20)$$

After determining the appropriate β_C^* , I get the solution $\mathbf{m}_C^* = [0.33, 0.33]$ marked with a black circle in Figure 4.6(a). I have reached the only solution with equal contribution from both model parameters that also predicts the data within the tolerance.

Alternatively, I can attempt to solve the same problem in a polar coordinate

system under the transformation

$$\begin{aligned}\mathbf{m}_P &= [\rho, \theta]^\top \\ m_x &= \rho \cos \theta \\ m_y &= \rho \sin \theta,\end{aligned}\tag{4.21}$$

where the polar model \mathbf{m}_P is defined by a radius ρ and an angle θ . This is analogous to the spherical transformation performed for the MVI-S formulation in (4.7). The objective function to be minimized becomes

$$\phi(\mathbf{m}_P) = \|\mathbb{F}[\mathbf{m}_P] - \mathbf{d}^{obs}\|_2^2 + \beta_P \|\mathbf{W}_C \mathbf{m}_P\|_2^2\tag{4.22}$$

The inverse problem is now non-linear with respect to the polar model so I solve it iteratively with the standard Gauss-Newton procedure described in equation (3.31). The partial derivatives of the forward mapping with respect to the polar coordinates are calculated by

$$\mathbf{J} = \frac{\partial \mathbb{F}[\mathbf{m}_P]}{\partial \mathbf{m}_P} = \frac{\partial \mathbb{F}[\mathbf{m}_P]}{\partial \mathbf{m}_C} \frac{\partial \mathbf{m}_C}{\partial \mathbf{m}_P} = \mathbf{F}_C \mathbf{S}\tag{4.23}$$

where the matrix \mathbf{S} holds the partial derivatives of the model with respect to the polar parameters

$$\mathbf{S} = \begin{bmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{bmatrix}.\tag{4.24}$$

The gradient of the objective function becomes

$$\mathbf{g} = \mathbf{S}^\top \mathbf{F}_C^\top \mathbb{F}[\mathbf{m}_P] + \beta_P \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_P - \mathbf{S}^\top \mathbf{F}_C^\top \mathbf{d}^{obs}\tag{4.25}$$

A trade-off parameter β_P is determined through the cooling schedule established in Chapter 3. The inversion is terminated once the data misfit and change in model norm fall below the tolerances η_{ϕ_d} and η_{ϕ_m} defined in equation (3.33) and (3.34) respectively.

Since \mathbf{m}_C^* is a desirable model, I would like to be able to recover a similar

solution in polar parameters ($\mathbf{m}_P^* = [0.47, 0.76]$). Unfortunately, as shown in Figure 4.6(b), the minimization process performed in polar coordinates converges to a different solution (\mathbf{m}_P [$\rho = 0.67, \theta = 0.26$]) and the iterations steps are oscillatory. I display the equivalent iterations (red dash) in the Cartesian space for comparison (\mathbf{m}_P^C [$m_x = 0.65, m_y = 0.17$]).

My main goal is to recover the solution \mathbf{m}_P^* , and I want to reach this solution with only a few model updates. To understand the discrepancy between the two formulations, I compare their respective gradients for a given starting model $\mathbf{m}_C^{(0)}$ and its equivalent polar model $\mathbf{m}_P^{(0)}$. In Cartesian coordinates, the gradient direction is

$$\mathbf{g}_C^{(0)} = \mathbf{F}_C^\top \mathbf{F}_C \mathbf{m}_C^{(0)} + \beta_C^* \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_C^{(0)} - \mathbf{F}_C^\top \mathbf{d}^{obs} \quad (4.26)$$

assuming that I know the optimal trade-off parameter β_C^* . I can convert these gradients to polar coordinate by multiplying (4.26) with the matrix of partial derivatives \mathbf{S} such that

$$\mathbf{g}_C^P = \mathbf{S}^\top \left[\mathbf{F}_C^\top \mathbf{F}_C \mathbf{m}_C^{(0)} + \beta_C^* \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_C^{(0)} - \mathbf{F}_C^\top \mathbf{d}^{obs} \right] \quad (4.27)$$

I want to compare this gradient to the gradient calculated in polar coordinates to find some equivalence between the two systems

$$\mathbf{g}_C^P \simeq \mathbf{S}^\top \mathbf{F}_C^\top \mathbb{F}[\mathbf{m}_P^{(0)}] + \beta_P \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_P^{(0)} - \mathbf{S}^\top \mathbf{F}_C^\top \mathbf{d}^{obs}. \quad (4.28)$$

I can simplify both sides of equation (4.28) by noting that $\mathbf{F}_C \mathbf{m}_C^{(0)} = \mathbb{F}[\mathbf{m}_P^{(0)}]$. Equation 4.28 and 4.29 are therefore the same if

$$\beta_C^* \mathbf{S}^\top \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_C^{(0)} \simeq \beta_P \mathbf{W}_C^\top \mathbf{W}_C \mathbf{m}_P^{(0)}. \quad (4.29)$$

I would like both sides to be roughly equal such that the gradient direction in polar space resemble the gradient direction calculated in the Cartesian space. First, I note from equation (4.29) that the transformation matrix \mathbf{S} is missing from the right-hand side. The current sensitivity weights were designed to compensate for fixed

experimental bias but did not account for a constantly varying sensitivity matrix \mathbf{J} . I address this shortcoming by resorting to an iterative update of sensitivity weights

$$\begin{aligned}\mathbf{W}_P &= \text{diag} \left[\left[\frac{\mathbf{w}_P}{\max(\mathbf{w}_P)} \right]^{1/2} \right] \\ w_{Pj} &= \left[\sum_{i=1}^N J_{ij}^2 \right]^{1/2},\end{aligned}\tag{4.30}$$

where the weights \mathbf{w}_P are updated between each Gauss-Newton step. Inverting once again the non-linear problem with the iterative scaling strategies (green) I recover the model $\mathbf{m}_P[\rho = 0.53, \theta = 0.53]$ (Fig. 4.6(b)). The solution has equal parameters of ρ and θ , and I reached this solution in a few iterations. In most applications, however, obtaining proportionality between the magnitude and angle of the vector is not meaningful. Converted to Cartesian space $\mathbf{m}_P^C[m_x = 0.46, m_y = 0.27]$, I note that the solution is still different from \mathbf{m}_C^* .

To understand this result, I now examine equation (4.29) in terms of the size of the model parameters in \mathbf{m}_P . I have used a regularization function to penalize two parameters with different units: the radius $\rho \in [0, \infty]$ in lengths and angle $\theta \in [-\pi, \pi]$ in radians. The range of values spanned by these parameters differ in scale as depicted by the aspect ratio of Figure 4.6(b). Comparing the largest change in model values (gradient) in relation to the Cartesian space \mathbf{m}_C , it is easy to show that

$$\|\mathbf{g}_\rho\|_\infty \propto \|\mathbf{g}_x\|_\infty + \|\mathbf{g}_y\|_\infty\tag{4.31}$$

such that a change in the radius ρ is proportional in magnitude to a change in components in Cartesian space. The same relation does not hold for the angle parameter as an equivalent change in Cartesian parameters can be achieved with a rotation $\Delta\theta = \pi/2$ independent of length. In order to scale the gradient steps taken along different dimensions, I define a proportionality relation

$$\omega = \frac{\|\mathbf{g}_\rho\|_\infty}{\|\mathbf{g}_\theta\|_\infty} = \|\rho\|_\infty \frac{2}{\pi},\tag{4.32}$$

and scale the regularization as

$$\hat{\mathbf{W}}_P = \text{diag} \left(\begin{bmatrix} 1 & \omega \end{bmatrix}^{(1/2)} \right) \mathbf{W}_P \quad (4.33)$$

My new scaled objective function becomes

$$\phi(\mathbf{m}_P) = \|\mathbb{F}[\mathbf{m}_P] - \mathbf{d}^{obs}\|_2^2 + \beta_P \|\hat{\mathbf{W}}_P \mathbf{m}_P\|_2^2 \quad (4.34)$$

Minimizing this function I get the model (blue) presented in Figure 4.6 ($\mathbf{m}_P^C[m_1 = 0.47, m_2 = 0.74]$). Converted to Cartesian space, this solution $\mathbf{m}_P^C[m_1 = 0.35, m_2 = 0.32]$ closely matches \mathbf{m}_C^* .

4.2.1 Scaled MVI-S algorithm

Now that I have demonstrated the benefit of an iterative sensitivity re-weighting of the regularization, I re-visit my synthetic magnetic problem. I invert the TMI data once more with the starting model oriented at 90° from the true magnetization direction ($\mathbf{m}_P^{(0)}[\rho = 10^{-2}, \theta = -45^\circ, \phi = 0^\circ]$) with smooth assumptions ($p_s, p_x, p_y, p_z = 2$). The recovered magnetization model obtained after convergence of the scaled MVIS-S algorithm is presented in Figure 4.7(a). The inversion took 15 iterations to converge. I note close similarities with the MVI-C solution presented in 4.3(a), with the bulk magnetization centered around the position of the block. This is a clear improvement over the solution previously shown in Figure 4.5(c).

From a practical standpoint, I have found that it is more efficient to initialize the MVI-S algorithm with the Cartesian solution. The linear MVI-C approach allows me to rapidly find a model that fits the observed data and it provides a good approximation to the MVI-S solution. I invert the data once more using the smooth Cartesian solution in 4.3(a) as a starting model. Figure 4.8(a) shows the recovered solution obtained after a single iteration of MVI-S.

Having achieved a stable and reasonable solution with the ℓ_2 -norm, I can now consider applying sparsity constraints to recover a block with a coherent magnetization direction. I vary the regularization measure on the amplitude, derivative of amplitude and derivatives of angles uniformly such that ($p_{c_s}, p_{c_x}, p_{c_y}, p_{c_z} = 0$) where $c \in [\rho, \theta, \phi]$. Figure 4.8(c) presents a section through the magnetic vector model.

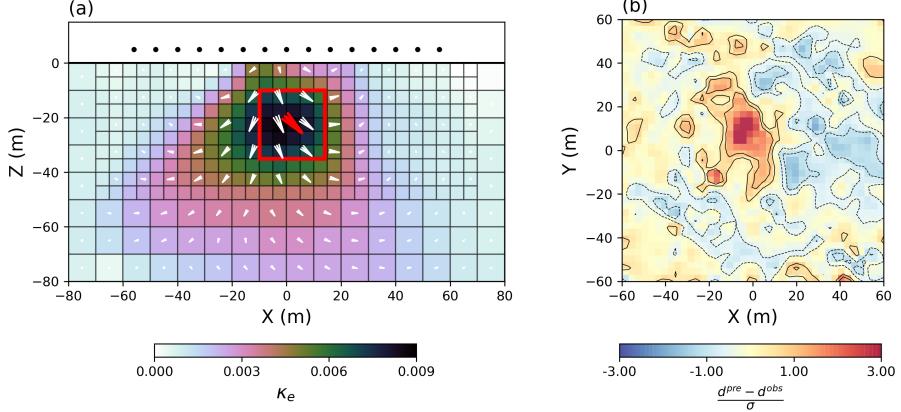


Figure 4.7: (a) Vertical section through the recovered magnetization vector model and (b) normalized data residual using the spherical formulation with sensitivity based weighting ($p_s = p_x = p_y = p_z = 2$). The true position and magnetization orientation of the block are shown in red for reference.

The shape of the anomaly matches that of the magnetic block, with a magnetization direction uniformly orientated at 45° inclination. The uniform effective susceptibility recovered inside the block matches exactly the true model with $\kappa_e = 0.05$ SI. The normalized data residual shows no apparent correlated signal (Fig. 4.8 d). This simple example increases my confidence in my ability to accurately recover the magnetization of geological bodies in 3D.

4.3 Synthesis

In this section, I introduced an iterative sensitivity re-weighting strategy to improve the convergence of non-linear inverse problems. The iterative re-scaling of the regularization function associated with the amplitude and angles of magnetization was crucial in order to get a stable convergence of the MVI-S algorithm. Smoother and more robust solutions allowed me to apply compact norms on the three model parameters independently. This can greatly simplify the solution compared to that from the conventional MVI-C formulation.

Despite this improvement, the MVI problem remains fundamentally under-determined. Incorporating *a priori* information, either through model constraints

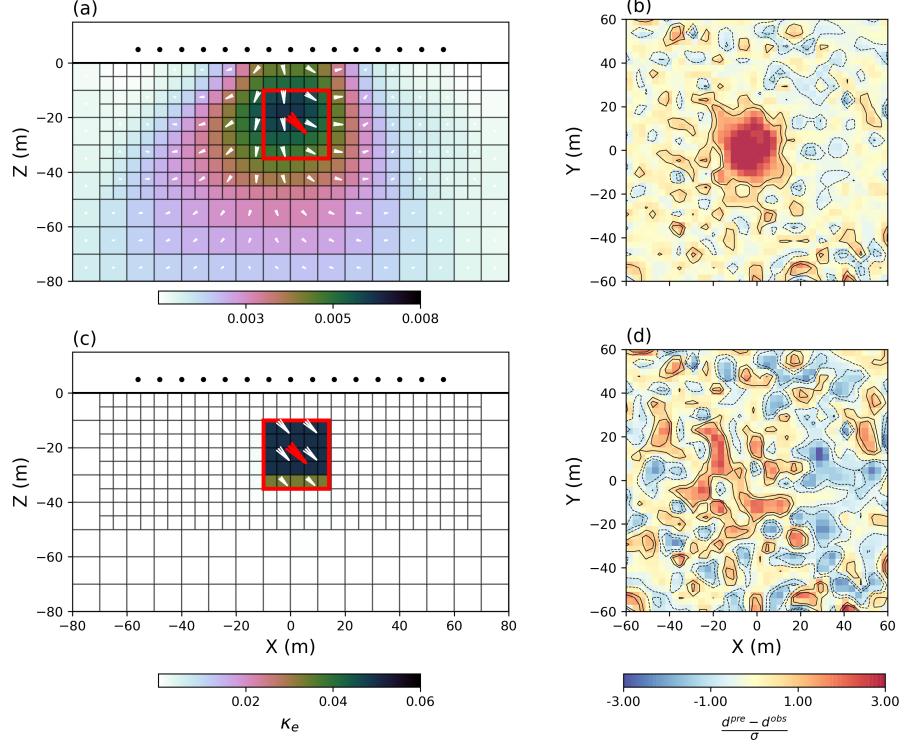


Figure 4.8: (a) Vertical section through the recovered magnetization vector model in Spherical coordinates using sensitivity based weighting with smooth regularization ($p_s = p_x = p_y = p_z = 2$). MVI-C solution is used as a starting model. The true position and orientation of magnetization of the block are shown in red for reference. (c) Recovered vector model in Spherical coordinates using sparse and blocky assumptions on the amplitude and angles of magnetization ($p_{cs} = p_{cx} = p_{cy} = p_{cz} = 0$). (b) and (d) Corresponding normalized data residuals.

or joint physical properties, remains important to accurately represent the geology. In the following section, I introduce structural information to constrain the shape of the recovered magnetic bodies.

Chapter 5

Rotated ℓ_p -norm Regularization

In Chapter 3, I introduced a mixed-norm regularization function that could be used to recover a suite of models that had a broad range of characteristics. This was accomplished by altering the norms on the model and model derivatives independently. I have shown the value in interpreting an ensemble of models to assess the robustness of certain features. When inverting gravity data arising from a buried prism, I managed to recover an anomaly at the right depth and with sharp edges. In this case, the prism was aligned with the Cartesian frame. In most cases, however, the shape of geophysical anomalies can take many forms. Dip and strikes of geological contacts rarely occur at right angles and aligned with the geographic grid. I, therefore, need a more general strategy that is adaptable to any geological scenarios.

Some research has addressed this issue by imposing directionality in the inversion. Barbosa and Silva (2006) introduce a regularization function to force anomalies to be closer to geometric elements in 2D. Their approach can recover compact and elongated bodies, but it requires direct input from the user to define the location of different anomalies. In Li and Oldenburg (2000), directionality is enforced by rotating the objective function along preferential axes. The method was later revised by Lelièvre (2009) and Davis et al. (2012) who advocated the use of forward and backward difference operators to improve symmetry. While successful in recovering elongated bodies, their approach still requires manual adjustment of inversion parameters to control stretching factors along different orientations. As

a natural extension to their work, I am proposing in this chapter to combine the mixed norm regularization function with the rotation strategy for the recovery of sharp and oriented edges.

5.1 Rotated objective function

I begin by reviewing the strategy presented in Li and Oldenburg (2000) for the rotation of the objective function using the conventional ℓ_2 -norm regularization

$$\phi_m = \alpha_s \int_V m^2 dV + \sum_{r=x,y,z} \alpha_r \int_V \left(\frac{dm}{dr} \right)^2 dV . \quad (5.1)$$

Model gradients are traditionally calculated along the Cartesian axes x , y and z . Evaluating the integral over the discrete domain gives rise to linear penalty functions introduced in equation (3.5)

$$\phi_s = \alpha_s \|\mathbf{W}_s \mathbf{V}_s (\mathbf{m} - \mathbf{m}^{ref})\|_2^2 + \sum_{r=x,y,z} \alpha_r \|\mathbf{W}_r \mathbf{V}_r \mathbf{D}_r \mathbf{m}\|_2^2 , \quad (5.2)$$

where \mathbf{W} matrices contain sensitivity or volumetric based weighting and weighting parameters set by the user. As explained in Section 3.1.3, I use finite difference operators \mathbf{D} rather than the conventional gradient operators \mathbf{G} to simplify the implementation. The scaling parameters α_r control the stretching along the three axes. This formulation is restrictive however as the dip and strike of geological boundaries can occur at any orientation. To address this shortcoming, Li and Oldenburg (2000) introduced rotated gradient measures of the form

$$\begin{aligned} \frac{\partial m}{\partial x'} &= \omega_{xx} \frac{\partial m}{\partial x} + \omega_{xy} \frac{\partial m}{\partial y} + \omega_{xz} \frac{\partial m}{\partial z} \\ \frac{\partial m}{\partial y'} &= \omega_{yx} \frac{\partial m}{\partial x} + \omega_{yy} \frac{\partial m}{\partial y} + \omega_{yz} \frac{\partial m}{\partial z} \\ \frac{\partial m}{\partial z'} &= \omega_{zx} \frac{\partial m}{\partial x} + \omega_{zy} \frac{\partial m}{\partial y} + \omega_{zz} \frac{\partial m}{\partial z} \end{aligned} \quad (5.3)$$

Rotation coefficients ω are defined by the triple rotation around the Cartesian axes such that

$$\begin{aligned}\Omega = \Omega_z \Omega_x \Omega_y &= \begin{bmatrix} \omega_{xx} & \omega_{xy} & \omega_{xz} \\ \omega_{yx} & \omega_{yy} & \omega_{yz} \\ \omega_{zx} & \omega_{zy} & \omega_{zz} \end{bmatrix} \\ &= \begin{bmatrix} \cos \phi \cos \psi - \sin \phi \cos \theta \sin \psi & \sin \phi \cos \psi + \cos \phi \cos \theta \sin \psi & \sin \theta \sin \psi \\ -\sin \phi \sin \theta & \cos \phi \sin \theta & -\cos \theta \\ -\cos \phi \sin \psi - \sin \phi \cos \theta \cos \psi & -\sin \phi \sin \psi + \cos \phi \cos \theta \cos \psi & \sin \theta \cos \psi \end{bmatrix} \quad (5.4)\end{aligned}$$

where the matrix Ω_y defines the rotation angle ψ counterclockwise around the y -axis (dip)

$$\Omega_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & \sin(\psi) \\ 0 & -\sin(\psi) & \cos(\psi) \end{bmatrix} \quad (5.5)$$

followed by Ω_x defining the rotation angle θ around the x -axis (plunge)

$$\Omega_y = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \quad (5.6)$$

and a final Ω_z defining the rotation ϕ around the z -axis (strike)

$$\Omega_z = \begin{bmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5.7)$$

Figure 5.1 depicts the rotation angles with respect to the Cartesian system. The regularization functions along the rotated axis x' is defined by substituting (5.3) into (5.1)

$$\phi_{x'} = \alpha_x \int_V \left(\omega_{xx} \frac{dm}{dx} + \omega_{xy} \frac{dm}{dy} + \omega_{xz} \frac{dm}{dz} \right)^2 dV. \quad (5.8)$$

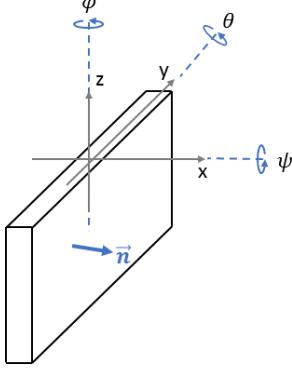


Figure 5.1: Rotation parameters with respect to the Cartesian axes.

which after expanding can be written in discrete form as:

$$\begin{aligned} \phi_x' = \alpha_x \mathbf{m}^\top & \left(\omega_{xx}^2 \mathbf{D}_x^\top \mathbf{W}_x^\top \mathbf{W}_x \mathbf{D}_x + \omega_{xy}^2 \mathbf{D}_y^\top \mathbf{W}_y^\top \mathbf{W}_y \mathbf{D}_y + \omega_{xz}^2 \mathbf{D}_z^\top \mathbf{W}_z^\top \mathbf{W}_z \mathbf{D}_z + \right. \\ & 2 * \omega_{xx} \omega_{xy} \mathbf{D}_x^\top \mathbf{W}_x^\top \mathbf{W}_y \mathbf{D}_y + \\ & 2 * \omega_{xx} \omega_{xz} \mathbf{D}_x^\top \mathbf{W}_x^\top \mathbf{W}_z \mathbf{D}_z + \\ & \left. 2 * \omega_{xy} \omega_{xz} \mathbf{D}_y^\top \mathbf{W}_y^\top \mathbf{W}_z \mathbf{D}_z \right) \mathbf{m} \end{aligned} \quad (5.9)$$

where I absorbed the volumes of integration \mathbf{V}_r in the weighting terms \mathbf{W}_r for clarity. Derivation from the integral to discrete form of (5.9) can be found in Li and Oldenburg (2000). It is important to point out that the difference operators in (5.9) require square matrices such that $\mathbf{D}_x, \mathbf{D}_y, \mathbf{D}_z \in \mathbb{R}^{M \times M}$. This is done by adding a backward difference to the edge cells such that the divergence operator in (3.28) becomes

$$\mathbf{D}_x = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & -1 & 1 \\ \vdots & \dots & 0 & 1 & -1 \end{bmatrix}. \quad (5.10)$$

Carrying out the same procedure for ϕ_y and ϕ_z , I obtain the rotated regularization function $\phi_{m'}$ made up of seven terms:

$$\begin{aligned}\phi_{m'} = & \mathbf{m}^\top \mathbf{W}_s^\top \mathbf{W}_s \mathbf{m} + \\ & \mathbf{m}^\top \left(\mathbf{D}_x^\top \mathbf{W}_x^\top \mathbf{B}_{xx} \mathbf{W}_x \mathbf{D}_x + \mathbf{D}_y^\top \mathbf{W}_y^\top \mathbf{B}_{yy} \mathbf{W}_y \mathbf{D}_y + \mathbf{D}_z^\top \mathbf{W}_z^\top \mathbf{B}_{zz} \mathbf{W}_z \mathbf{D}_z + \right. \\ & \left. 2 * \mathbf{D}_x^\top \mathbf{W}_x^\top \mathbf{B}_{xy} \mathbf{W}_y \mathbf{D}_y + 2 * \mathbf{D}_x^\top \mathbf{W}_x^\top \mathbf{B}_{xz} \mathbf{W}_z \mathbf{D}_z + 2 * \mathbf{D}_y^\top \mathbf{W}_y^\top \mathbf{B}_{yz} \mathbf{W}_z \mathbf{D}_z \right) \mathbf{m}\end{aligned}\quad (5.11)$$

Rotation coefficients ω_\square and α_\square parameters corresponding to each term are collected and added to form the different diagonal matrices \mathbf{B}_\square . Rotation parameters can be defined on a cell by cell basis.

Lelièvre (2009) found experimentally that this formulation resulted in asymmetric solutions due to the cancellation of rotation parameters. Checkerboard patterns were observed for rotation angles at 45° . This was addressed with an averaged combination of forward and backward difference operators. Equation (5.11) becomes

$$\begin{aligned}\phi_{m'} = & \mathbf{m}^\top \mathbf{W}_s^\top \mathbf{W}_s \mathbf{m} + \\ & \mathbf{m}^\top \frac{1}{8} \sum_{i=1}^8 \left(\mathbf{D}_{x,i}^\top \mathbf{W}_x^\top \mathbf{B}_{xx} \mathbf{W}_x \mathbf{D}_{x,i} + \right. \\ & \mathbf{D}_{y,i}^\top \mathbf{W}_y^\top \mathbf{B}_{yy} \mathbf{W}_y \mathbf{D}_{y,i} + \\ & \mathbf{D}_{z,i}^\top \mathbf{W}_z^\top \mathbf{B}_{zz} \mathbf{W}_z \mathbf{D}_{z,i} + \\ & 2 * \mathbf{D}_{x,i}^\top \mathbf{W}_x^\top \mathbf{B}_{xy} \mathbf{W}_y \mathbf{D}_{y,i} + \\ & 2 * \mathbf{D}_{x,i}^\top \mathbf{W}_x^\top \mathbf{B}_{xz} \mathbf{W}_z \mathbf{D}_{z,i} + \\ & \left. 2 * \mathbf{D}_{y,i}^\top \mathbf{W}_y^\top \mathbf{B}_{yz} \mathbf{W}_z \mathbf{D}_{z,i} \right) \mathbf{m}\end{aligned}\quad (5.12)$$

such that the i^{th} components correspond to all the combination of forward and backward difference operators. Figure 5.2(a) and (b) compares the different gradient scheme along the xy-plane. For 3D problems, this formulation results in a regularization function containing 49 terms in total.

I showcase the effect of rotation with three inversions applied to the density

ID	$\mathbf{D}_{x,i}$	$\mathbf{D}_{y,i}$	$\mathbf{D}_{z,i}$
1	Forward	Forward	Forward
2	Backward	Forward	Forward
3	Forward	Backward	Forward
4	Forward	Forward	Backward
5	Forward	Backward	Backward
6	Backward	Forward	Backward
7	Backward	Backward	Forward
8	Backward	Backward	Backward

block model introduced in Chapter 2. I will take advantage of the non-uniqueness of potential fields to accentuate trends in the model, even though the true block anomaly is a single prism orientated parallel to the Cartesian axes. In the first inversion, I use the conventional ℓ_2 -norm regularization for a rotation angle of $\phi = 45^\circ$. Since I am dealing with smooth assumption, I must impose a stretching factor along one of the rotated axes ($\alpha_x = 100$, $\alpha_s = \alpha_y = \alpha_z = 1$). As expected the recovered density anomaly is stretched and symmetric about the rotated axis x' (Figure 5.3(a)). I have managed to replicate the implementation of Lelièvre (2009).

For the second inversion, I attempt to impose sparse assumptions ($p_s = p_x = p_y = p_z = 0$) to recover a compact block anomaly oriented at 45° . As a first pass, I attempt to directly incorporate the IRLS weights into the rotated regularization function such that

$$\begin{aligned}
\phi_m = & \mathbf{m}^\top \mathbf{W}_s^\top \mathbf{R}_s^\top \mathbf{R}_s \mathbf{W}_s \rho + \\
& \mathbf{m}^\top \frac{1}{8} \sum_{i=1}^8 \left(\mathbf{D}_{x,i}^\top \mathbf{R}_{xx}^\top \mathbf{W}_x^\top \mathbf{B}_{xx} \mathbf{W}_x \mathbf{R}_{xx} \mathbf{D}_{x,i} + \right. \\
& \quad \mathbf{D}_{y,i}^\top \mathbf{R}_{yy}^\top \mathbf{W}_y^\top \mathbf{B}_{yy} \mathbf{W}_y \mathbf{R}_{yy} \mathbf{D}_{y,i} + \\
& \quad \mathbf{D}_{z,i}^\top \mathbf{R}_{zz}^\top \mathbf{W}_z^\top \mathbf{B}_{zz} \mathbf{W}_z \mathbf{R}_{zz} \mathbf{D}_{z,i} + \\
& \quad 2 * \mathbf{D}_{x,i}^\top \mathbf{R}_{xy}^\top \mathbf{W}_x^\top \mathbf{B}_{xy} \mathbf{W}_y \mathbf{R}_{xy} \mathbf{D}_{y,i} + \\
& \quad 2 * \mathbf{D}_{x,i}^\top \mathbf{R}_{xz}^\top \mathbf{W}_x^\top \mathbf{B}_{xz} \mathbf{W}_z \mathbf{R}_{xz} \mathbf{D}_{z,i} + \\
& \quad \left. 2 * \mathbf{D}_{y,i}^\top \mathbf{R}_{yz}^\top \mathbf{W}_y^\top \mathbf{B}_{yz} \mathbf{W}_z \mathbf{R}_{yz} \mathbf{D}_{z,i} \right) \mathbf{m}
\end{aligned} \tag{5.13}$$

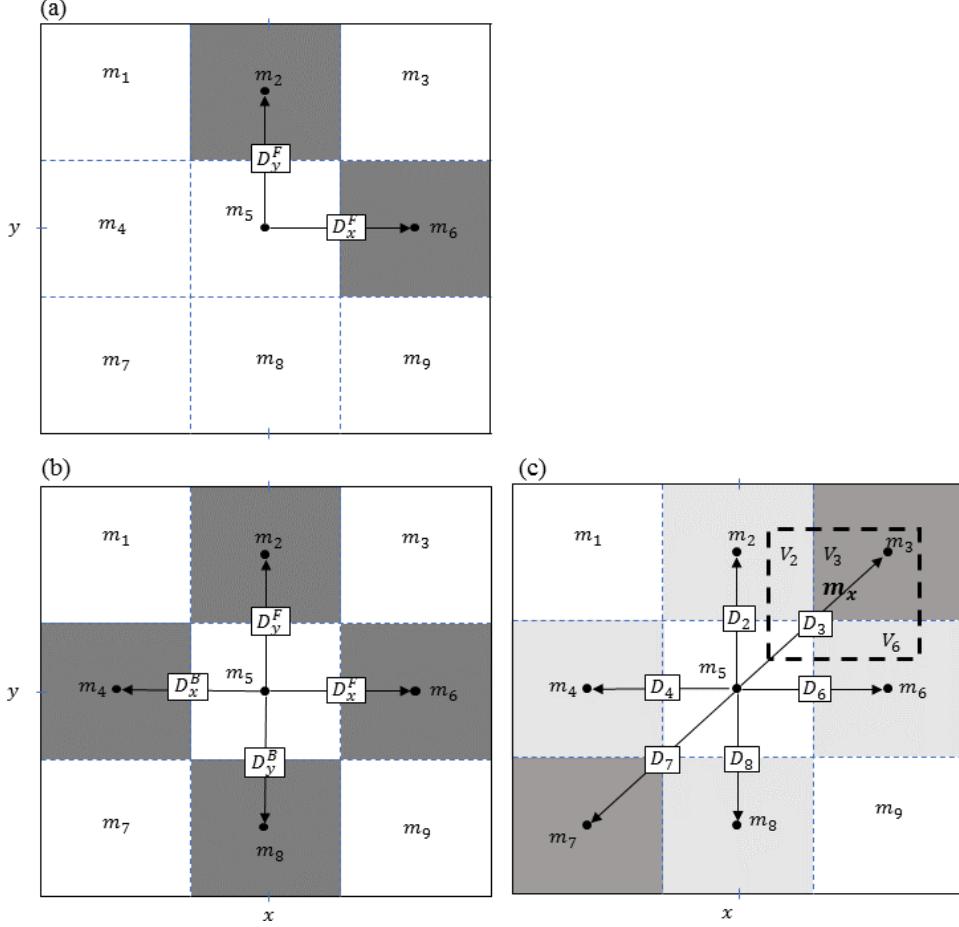


Figure 5.2: Two-dimensional representation of finite difference operators using a combination of (a) 3-cell (Li and Oldenburg, 2000), (b) 5-cell (Lelièvre, 2009) and (c) the 7-cell scheme used for the rotation of the objective function. Only the 7-cell strategy allows for the interaction between diagonal neighbors.

The IRLS weights for the model derivatives are defined as

$$\mathbf{R}_{ij} = \text{diag} \left[\left(|(\mathbf{G}_i \mathbf{m}^{(k-1)}) \odot (\mathbf{G}_j \mathbf{m}^{(k-1)})| + \varepsilon_{ij}^2 \right)^{p_{ij}/2-1} \right]^{1/2}. \quad (5.14)$$

where \odot denotes the Hadamard (element-wise) product between the model derivatives. Figure 5.3(b) shows the recovered density models after convergence of the

IRLS. While the general orientation of the anomaly appears to be aligned at 45° , edges are poorly defined and the solution is not exactly symmetric about the x' axis. Similarly for the third inversion, I rotate the objective function by a smaller angle of 30° . The solution in Figure 5.3(c) shows little difference from the 45° rotation. This is primarily due to the averaging over the 8 combination of backward and forward measures. The regularization function in (5.13) also comes with a significant increase in computational cost as it involves 16 times more operations compared to the conventional Cartesian approach.

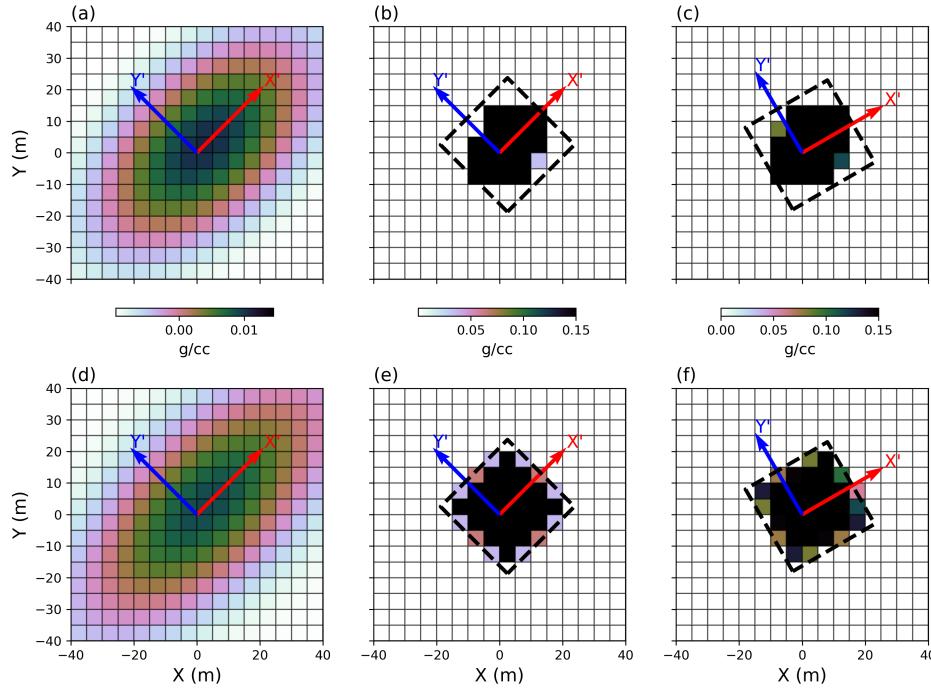


Figure 5.3: Horizontal cross section through the recovered model with rotation of the objective function using (top) 5-point and (bottom) 7-point gradient operators. Blue and red arrows indicate the rotation frame. (a, d) Recovered models with smooth assumptions ($p_s = p_x = p_y = p_z = 2$) stretched along the rotated x -axis at 45° ($\alpha_{x'} = 100$). (b, e) Sparse solutions ($p_s = p_x = p_y = p_z = 0$) with rotation at 45° . The target orientation of the rotated block is marked by the black dashed line. (c, f) Sparse solutions for a 30° rotation of the objective function.

5.1.1 Diagonal derivatives

Building upon the previous results, I will attempt to improve the recovery of oriented bodies by employing a new derivative scheme. Rather than adding and averaging the contribution of gradient operators along the Cartesian axes, I propose to also measure the interaction with diagonal neighbours. For clarity, I first examine the rotated finite difference operator $\mathbf{D}_{x'}$ along the xy-plane (rotation about \hat{z}). For a given cell m_i , there are 8 cells that either share a node or a face. Given a rotation angle θ , I need to determine which of the 8 neighbours participate in the finite difference. I determine the contribution of each cell based on the intersecting volume with a test cell $m_{x'}$ as shown in Figure 5.2(c). To determine the position of $m_{x'}$, I calculate the displacement of its nodes about the center of m_i

$$\mathbf{M}_{x'} = \mathbf{R}(\theta) \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} x_U & x_L \\ y_U & y_L \\ z_U & z_L \end{bmatrix} \quad (5.15)$$

where U and L define the lower southwest and upper northeast node locations of the target cell m_i . The rotated finite difference is calculated by a weighted average

$$d_{x'} = m_i - \frac{\sum_{i=1}^8 v_i m_i}{\sum_{i=1}^8 v_i} \quad (5.16)$$

where the weights v_i are partial volume intersected with $m_{x'}$ as presented in the Appendix. The same calculation can be carried out for rotated test cells $m_{y'}$ and $m_{z'}$ and rotation angle ψ and ϕ . To improve symmetry, this process is repeated for both a forward and backward difference. These calculations are performed for each cell and their neighbors in our 3D domain, giving rise to the regularization function:

$$\begin{aligned} \phi_m = & \| \alpha_s \mathbf{W}_s \mathbf{V}_s \mathbf{R}_s (\rho - \rho^{ref}) \|_2^2 + \\ & \sum_{r=x',y',z'} \| \alpha_r \mathbf{W}_r \mathbf{V}_r \mathbf{R}_r \mathbf{D}_r^F \mathbf{m} \|_2^2 + \| \alpha_r \mathbf{W}_r \mathbf{V}_r \mathbf{R}_r \mathbf{D}_r^B \mathbf{m} \|_2^2, \end{aligned} \quad (5.17)$$

where \mathbf{D}_r^F and \mathbf{D}_r^B denotes the rotated forward and backward difference operators. The upfront cost to build the finite difference operators is higher than for simple

Cartesian operators but the regularization function has only seven terms in total. This translates to about 6 times fewer operations needed during the inversion process compared to the methodology proposed by Lelièvre (2009).

For comparison, I re-invert the synthetic data using this new regularization function. Models are presented in Figure 5.3 for (d) 45° rotation with smooth ℓ_2 -norm ($\alpha_{x'} = 100$), (e) 45° rotation with ℓ_p -norm ($p_s = p_x = p_y = p_z = 0$) and (f) 30° rotation. In the case of the smooth inversion, both formulations give equivalent results with anomalous density stretched along the rotated x' -axis. I note a significant improvement however in the geometry of the rotated sparse models, with good symmetry and better edge definition along the rotation angle.

I have achieved my initial goal of recovering rotated anomalies, even though in this specific case the true model is a simple prism oriented along the Cartesian axes. Figure 5.4 presents the normalized data residual calculated from the last three experiments. It is important to note that all solutions fit the observed data within the tolerance $\phi_d - \phi_d^* \leq \delta d$, yet large correlated residuals are clearly visible. I have forced the solution to exhibit characteristics that are not consistent with the true solution. This reinforces the importance of designing an objective function that can satisfy subtle information that may be present in the data in order to better represent the local geology.

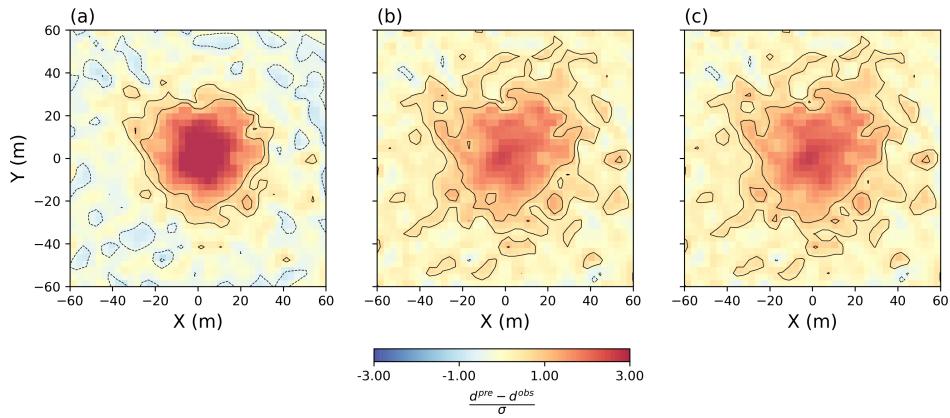


Figure 5.4: Normalized data residuals calculated from the recovered models presented in Figure 5.3(d), (e) and (f) respectively.

5.2 Synthetic fold model

I have so far demonstrated the use of rotated sparse norms for the recovery of oriented blocks by taking advantage of the non-uniqueness of the inverse problem. The true value of this strategy is to promote the recovery of oriented edges to better represent geological units of arbitrary shapes. I demonstrate this by simulating a more realistic scenario. I generate a synthetic fold model made up of a layer with density contrast of 0.1 g/cc placed in uniform background as shown in Figure 5.5. The axial plane of the syncline strikes NS and dips 75°E. The fold axis plunges 15° S. The intersection of the axial plane with the various cross-sections are shown as red dashed lines in Figure 5.5. I divide the ground into an OcTree mesh with core cells that are 20 m meters in width.

I simulate gravity data at 560 stations placed one meter above the topography. The locations are randomly sampled from a grid to emulate a field survey with unevenly spaced data. From the linear relation introduced in (2.3), the density anomaly of the syncline gives rise to the gravity response presented in Figure 5.6. I will attempt to use the rotated mixed-norm regularization to recover the folded layer.

5.2.1 ℓ_p -norm inversion

I begin exploring model space by first generating a density model that is smooth and has minimum structure. Using the regularization function in (5.17) for $p_s = p_x = p_y = p_z = 2$, I recover the density model presented in Figure 5.7. As expected from the smooth regularization function, the thickness of the layer near the surface is not clearly defined; there is a wide transition from a positive to a negative density contrast. More importantly, the dip and extent of the fold limbs at depth are poorly resolved. Without prior knowledge, geological interpretation of this result would be difficult with low confidence about the geometry of the fold.

I follow up with a mixed norm inversion and attempt to reduce the complexity of the solution, with sparsity assumptions applied on the model ($p_s = 0$) and smooth model gradients ($p_{x,y,z} = 2$). Figure 5.8 presents a vertical section at the center of the recovered density model. I note that this solution is an improvement over the smooth model for at least three reasons.

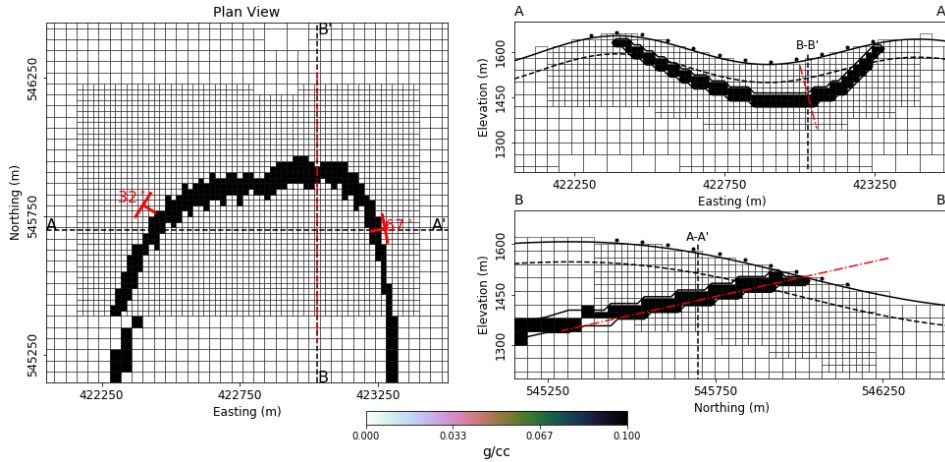


Figure 5.5: Synthetic density model made up of a folded layer with a N-S oriented axial plane dipping 75° E. The intersection between the fold axial plane and the cross-sections are marked by a red-dashed line: (left) topography draped section and (right) vertical cross-sections along (A-A') and (B-B'). The hinge axis of the fold is plunging 30° S. Two strike and dip measurements are provided at the surface and used as structural constraints.

1. The density contrasts values are mostly positive with the maximum density approaching the true density of 0.1 g/cc. The near-surface trace of the fold appears thinner with fewer negative density contrasts seen along the horizontally draped section.
2. The fold is imaged as a continuous body connected at depth along the section A-A'
3. The dip angle along the fold axis (B-B') is improved.

While I have made some progress in imaging a continuous layer connected at depth, there is still space for improvement about the shape and thickness of the fold. This is a motivation for incorporating geological information in the inversion. Rather than building a 3D model to constrain the solution (the usual path), I will attempt to drive the solution with *soft* directional constraints in the form of dip and strike measurements and let the data determine the position and shape of the fold.

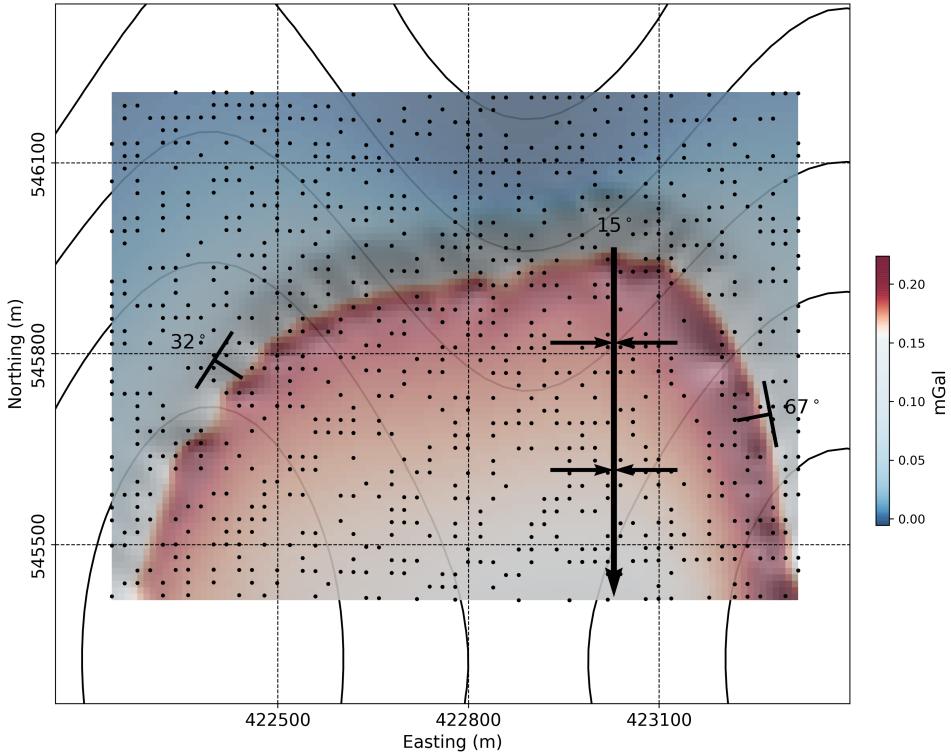


Figure 5.6: Synthetic gravity data generated 1 m above topography over the synthetic density fold model. Gravity station locations (dot) and topographic contours (lines) are shown for reference. Structural data (dip, strike) are provided at two locations on opposite of the syncline. Dip direction of the fold axis (15° S) and limbs are shown with arrows .

5.2.2 Directional ℓ_p -norms

My goal is to introduce stratigraphic constraints in order to recover a continuous layer connected at depth. Apart from the gravity data, structural data are often available in the form of dip and strike information measured on outcrops. I will include three points of structural data: one on each limb and one on the fold axis (Fig. 5.6). I will use this information to build a rotated objective function as prescribed in (5.17).

First I need to extrapolate the structural data to the full 3D domain. I use a minimum curvature approach (Briggs, 1974). The interpolated normal vectors are

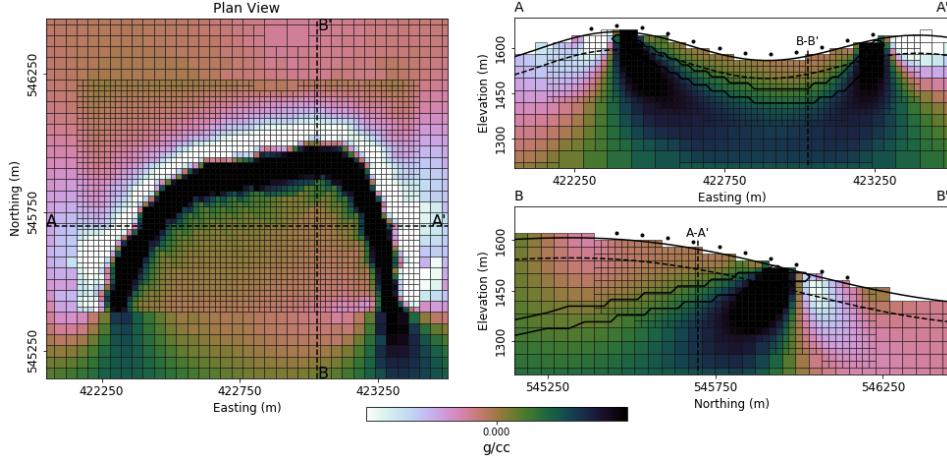


Figure 5.7: Recovered smooth solution with ℓ_2 -norm regularization of the synthetic fold anomaly. The dip of the fold limbs are poorly recovered and there is no indication of the limbs being connected at depth.

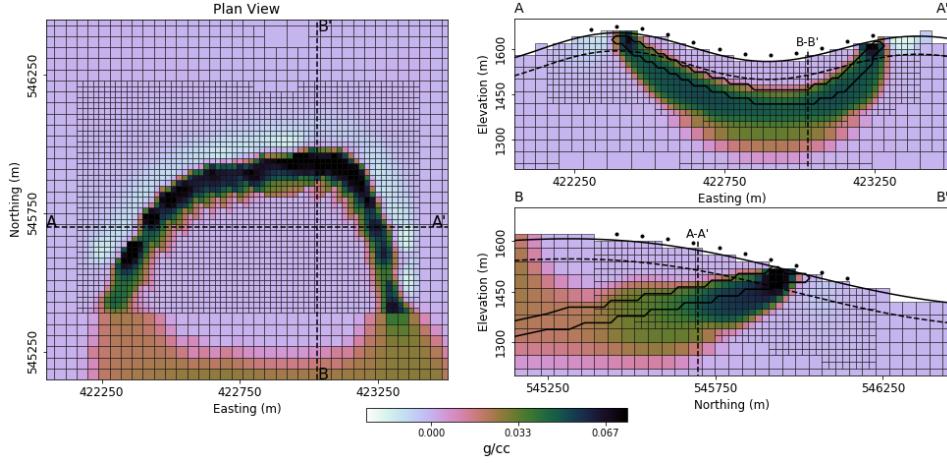


Figure 5.8: Recovered density model from the mixed norm regularization ($p_s = 0$, $p_{x,y,z} = 2$). Sparse assumptions helped in simplifying the density model over the smooth assumption.

shown in Figure 5.9. The orientation is then used to rotate the derivative terms on a cell-by-cell basis such that $\mathbf{D}_{x'}$ points along the normal of the fold, while $\mathbf{D}_{y'}$ and $\mathbf{D}_{z'}$ are parallel to the stratigraphy.

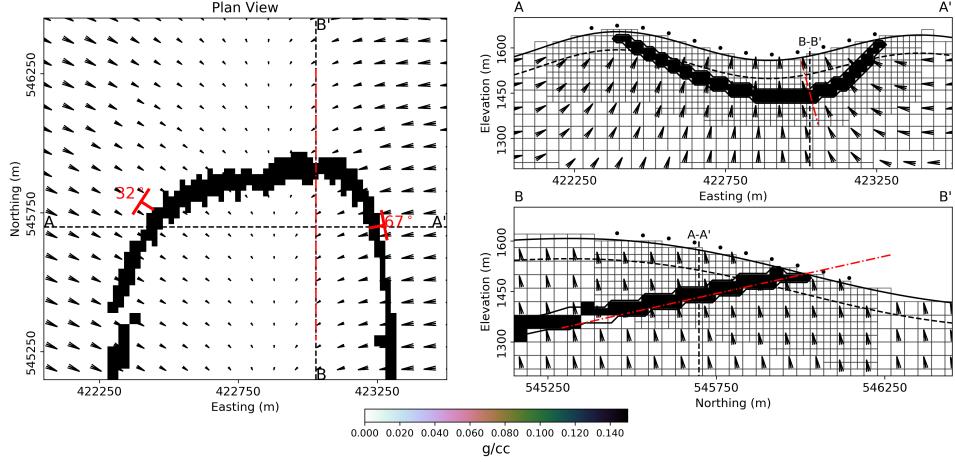


Figure 5.9: Interpolated dip and strike information used for the rotation of the objective function. The direction of the arrows define the estimated normal or the folded layer.

I then invert the gravity data with sparse rotated norms. To re-enforce lateral continuity, I lower the norm applied to model derivatives parallel to the stratigraphy ($p_{x'} = 2$, $p_{y'}, p_{z'} = 0$). Keeping smooth penalties on the normal component ($\mathbf{D}_{x'}$) helps for gradual readjustment of the fold position. Sparsity on the model values are also used ($p_s = 0$). Figure 5.10 presents sections through the final model. The dip and lateral continuity of the layer has improved remarkably, which has been achieved with only three-point constraints at the surface

5.3 Summary

In this chapter, I have successfully combined sparse norms and rotation of the objective function for the modelling of geological bodies with oriented edges. I have shown that it is possible to use point structural measurements, in the form of strike and dip angles, to constrain the model. The recovered fold model closely matched the true solution even though I did not specify the position of the anomaly or its physical property contrast. In this regard, sparse rotated norms can be seen as a *soft* geological constraint that requires only a general understanding of the geometry of the problem. The resulting model, or suite of models, could subsequently be used

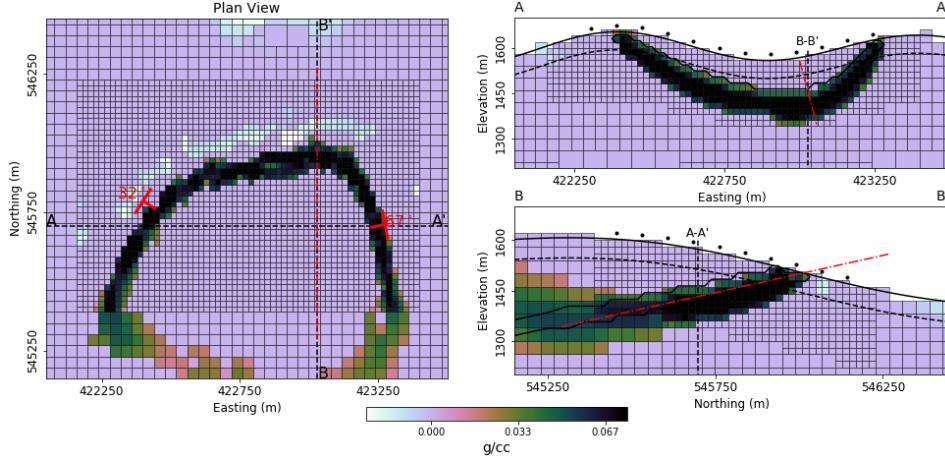


Figure 5.10: Recovered density model from the mixed-norm regularization. Data locations (dots), topography and the outlines of the true model (black) are shown for reference. Rotated ℓ_p -norm oriented along the folded layer helped in resolving a continuous geological unit at depth.

as a reference for more advanced modelling efforts.

The main challenge with the methodology presented so far is in generating rotation parameters in 3D. Simple extrapolation of scarce structural measurement may not accurately capture complex geological settings involving multiple folding and faulting events. For such complicated cases, such as the Kevitsa deposit introduced in Chapter 1, some level of 3D modelling and interpretation may be required by experts. Surface mapping and other types of geophysical data may be available to infer strike and dip information. In the following chapter, I will investigate ways to learn about preferential orientations using geophysical inversion.

Chapter 6

Automated mixed-norm modeling

I have so far developed a methodology that can generate a suite of models that adequately fit the data and, at the same time, sample model space in a consistent fashion. This flexibility promotes two objectives. First, inverting for diverse models provides immediate insight about non-uniqueness and that can prevent practitioners from over-interpreting a single inversion result. This is one of the most important aspects of my work.

The next objective is more challenging. I want to use different combinations of norms on different parts of the model space so that I obtain solutions that are geologically informative. As exemplified with the residual shown in Figure 3.19, certain assumptions may be better suited to fit portions of the data. However, the selection of different norms to be used in different portions of the model can rapidly become overwhelming for large problems in complex geological settings. A semi-automated approach that can capture the heterogeneity of the Earth, with minimal intervention from an expert, is needed. I am proposing the following strategy:

1. Run a suite of inversions over a range of p parameters applied to the model norm and its spatial gradients. I will assume that the suit of models is a representative sampling of diverse solutions from the ℓ_p -space.
2. Form an *average* model (\mathbf{m}_A) that captures most of the variability in the solution space
3. Extract p parameters or rotation angles for windowed portions of model

space based on the correlation between the average \mathbf{m}_A and individual models.

4. Carry out a final SVMN inversion with local parameters.

There are numerous ways to implement the above strategy. Here, I use Principal Component Analysis. I illustrate my approach with a 2D seismic tomography example and on the 3D fold example introduced in Chapter 5.

6.1 Selecting local p -parameters

I will first demonstrate how it is possible to extract local p parameters for the simultaneous recovery of smooth and compact features as introduced in Chapter 3. To demonstrate my methodology I use the 2D travel-time tomography example presented in Sun and Li (2014). The synthetic velocity model shown in Figure 6.1(a) is made up of a smooth velocity high and a blocky velocity low centered at a 1,000 m along the x-axis; the background velocity is 2000 m/s. Contour lines, corresponding to the 25th and 75th percentile values for both anomalies, are also plotted for reference. An array of 11 transmitters and 13 receivers is positioned on either side of the model which is discretized into 32×64 square 25 m cells. First arrival data are calculated for each transmitter-receiver pair by taking the line integral of slowness (reciprocal of velocity) along each ray path; this yields a total of 143 observations. Gaussian noise of 5% is added to the simulated data shown in Figure 6.1(b). I will first attempt to invert this data with mixed-norm penalties that are applied to the entire model.

6.1.1 Model space inversions

Having dealt with scaling issues between ℓ_p -norm measures, I can now exploit the full flexibility of the regularization function in (3.3). My goal is to generate a suite of models under a broad range of assumptions. For this 2D problem, the objective function to be minimized takes the form:

$$\begin{aligned} \min_{\mathbf{m}} \phi(\mathbf{m}) = & \| \mathbf{F} \mathbf{m} - \mathbf{d}^{obs} \|_2^2 + \beta \sum_{r=s,x,y} \alpha_r \| \mathbf{W}_r \mathbf{V}_r \mathbf{R}_r \mathbf{D}_r \mathbf{m} \|_2^2 \\ \text{s.t. } \phi_d \leq \phi_d^* \end{aligned} \tag{6.1}$$

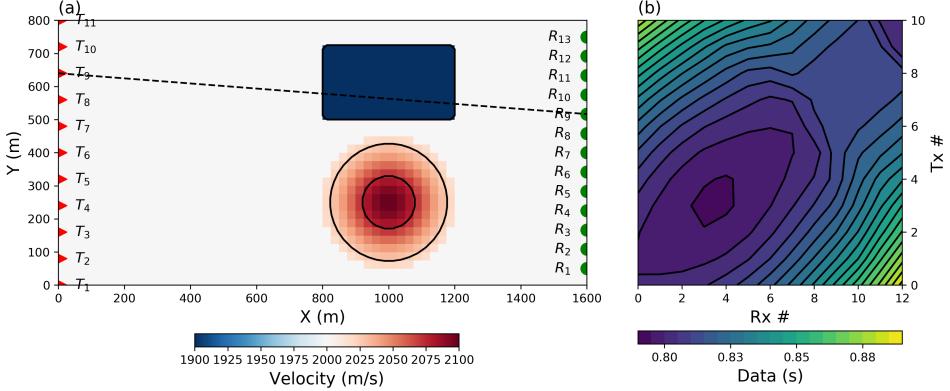


Figure 6.1: (a) Synthetic 2D travel-time tomography problem made up of a rectangular velocity low, and a smooth Gaussian velocity high. Contour lines (25^{th} and 75^{th} percentile) are shown in black for the true position and shape of the velocity anomalies. Transmitters (red) and receivers (green) are positioned on opposite sides of the model domain. (b) Data map for the 143 line integral data calculated for each transmitter-receiver pair.

To demonstrate the flexibility of S-IRLS algorithm, I carry out nine inversions, using a combination of norms on a range of p_s , p_x , $p_y \in [0, 1, 2]$ values ($p_x = p_y$ in all cases). The solutions, nine in total, are presented in Figure 6.2. All models have a final misfit $\phi_d^* \approx 143$ and use the same ℓ_2 -norm solution to initiate the S-IRLS steps. I make the following observations:

- The two velocity anomalies centered at 1000 m on the x -axis are dominant features.
- Anomalies are generally stretched along the ray path due to the experimental bias, even though I have attempted to compensate for it with the sensitivity weighting \mathbf{W}_r . It is less pronounced for $p \leq 1$.
- There is a progressive transition from a smooth model (upper left) to a blocky solution (lower right) as p_s , p_x and p_y decrease.
- The upper body (velocity low) appears to be most often represented as a blocky body with sharp edges.
- The lower anomaly (velocity high) tends to be more smooth.

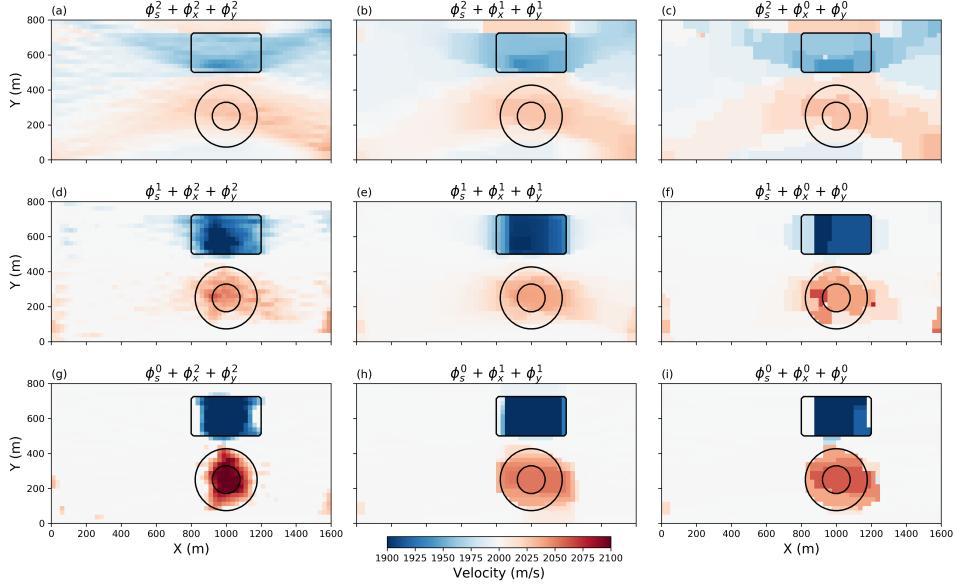


Figure 6.2: Suite of inverted models for various combinations of norms p_s , $p_x = p_y \in [0, 1, 2]$ ($\alpha_s = \alpha_x = \alpha_y = 1$). Contour lines (25^{th} and 75^{th} percentile) are shown in black for the true position and shape of the velocity anomalies.

- Away from the anomalous regions the velocity is relatively smooth and close to the background reference model of 2000 m/s.

Figure 6.3 presents the data residual maps predicted by these nine inversions. The random noise originally added to the data is shown in Figure 6.3(j) for comparison. I note that the strongest correlated residuals appear as *pant-legs*, which corresponds to ray paths travelling through the velocity anomalies. The residual trend associated with the blocky velocity low (low Tx# to high Rx#) is generally more obvious, which indicates that the experiment may be more sensitive to uniform velocity anomalies with sharp edges as it is the source of coherent signal. This pant-leg vanishes as $p, q \rightarrow 0$. Meanwhile, the residual trend associated with the smooth velocity high (low Rx# to high Tx#) diminishes for $p_s = p_x = p_y = 1$. This change in data residual is a strong indicator that a specific combination of norms may be better suited to represent individual velocity anomalies.

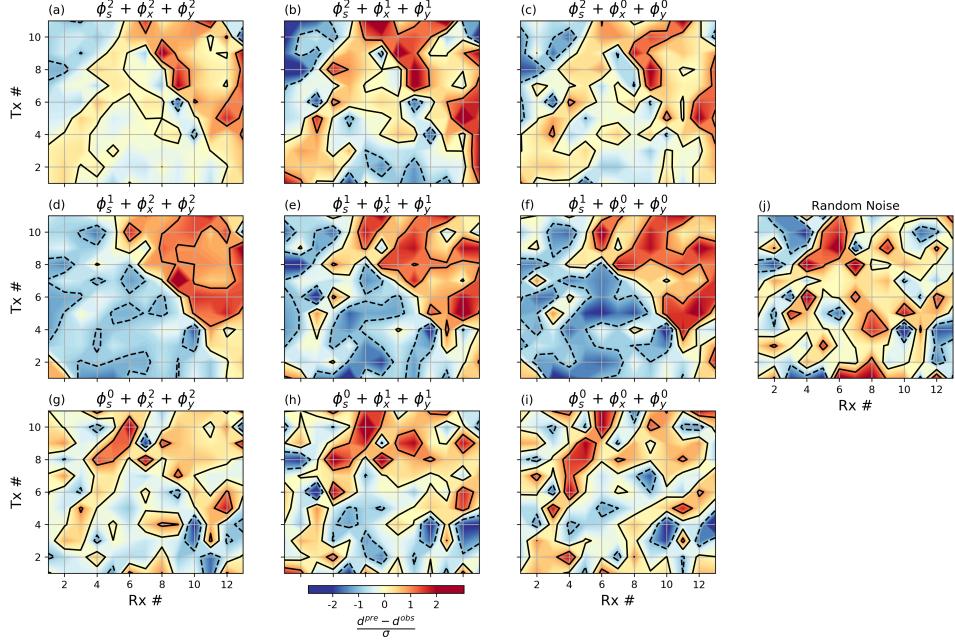


Figure 6.3: (a) to (i) Data residual contour map for the nine inversions presented in Figure 6.2. Contour lines (± 1 standard deviation) are shown in black. (j) Random noise added to the experiment.

6.1.2 Average PCA model

I assume that the suite of models presented in Figure 6.2 contains sufficient variability to be representative of my solution space and I wish to use these to extract the main features. There are numerous approaches to accomplish this and they vary in complexity from simple averaging to advanced machine learning algorithms. I use a Principal Component Analysis (Hotelling, 1933; Pearson, 1901). Considering each model in Figure 6.2 as a *data* vector, the principal components of my solution space can be written as:

$$[\mathbf{m}_1, \dots, \mathbf{m}_{nM}] \approx \mathbf{A} \mathbf{W}. \quad (6.2)$$

such that PCA vectors along the columns of $\mathbf{A} \in \mathbb{R}^{nM \times nV}$ contains a subset of nV eigenvectors spanning the model space. These vectors encode the principal source of variation across the nine models recovered. The corresponding weights $\mathbf{W} \in$

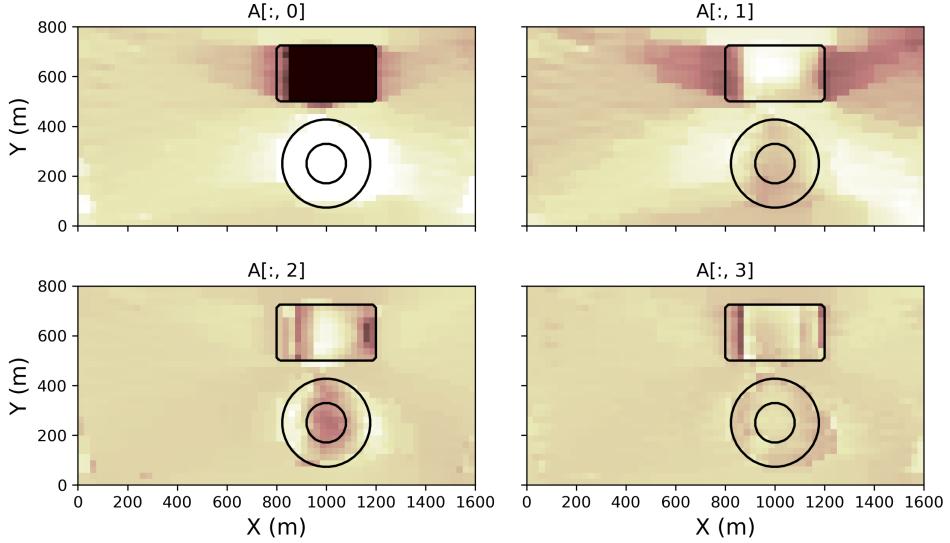


Figure 6.4: PCA vectors covering 75% of the model variances.

$\mathbb{R}^{nV \times nM}$, also known as loadings, are scalars relating to how each principal component contributes to each model. I use the PCA algorithm from the open-source Python library `Scikit-Learn.decomposition.PCA` (Pedregosa et al., 2011). The number of principal components to be used in the analysis is determined by the experimenter. Figure 6.4 presents the four largest principal components, covering in this case over 75% of the variance present in the model space.

Next, I generate a representative model by computing a weighted averaged model based on the positive PCA loadings such that:

$$\mathbf{m}_A = \frac{\sum_{i=1}^{nV} \sum_{j=1}^{nM} W_{ij} \mathbf{m}_i}{\sum_{i=1}^{nV} \sum_{j=1}^{nM} W_{ij}}. \quad (6.3)$$

The average model \mathbf{m}_A is presented in Figure 6.5. I note the close resemblance with the true model. The background is fairly uniform near 2000 m/s; the low velocity anomaly appears to be a block with a velocity near 1900 m/s, and the high velocity anomaly appears is a smooth feature with a maximum velocity near 2100 m/s.

Since \mathbf{m}_A is not the result of inversion, I have no guarantee that this model can satisfy the observed data. The corresponding normalized data residual map shows

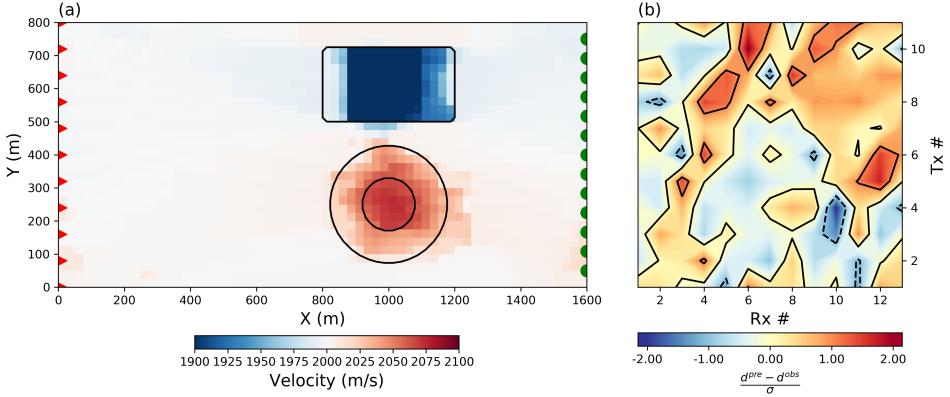


Figure 6.5: (a) Averaged PCA model \mathbf{m}_A and (b) corresponding normalize data residual map. Contour lines are shown in (a) for the true position and shape of the velocity anomalies.

some level of correlation with the anomalies (Figure 6.5(b)). My next objective is to identify inversion parameters that will allow me to fit both the data and features highlighted by the average model.

6.1.3 Parameter extraction

Next, I want to extract optimal inversion parameters on a cell-by-cell basis in order to best describe local features. To do so, I resort to a pattern recognition approach. In order to remove biases towards extreme model values, I transform my model space into a simpler parametric representation. I use a Canny edge detection algorithm from the open-source library `Scikit-Image.feature.Canny` (Pedregosa et al., 2011). Figure 6.6 shows the parametric edges extracted from all nine inversions.

From this simplified representation of each model, I perform a moving window correlation $r_{\mathbf{m}_i \mathbf{m}_P}$ between the average PCA model \mathbf{m}_A and each of the \mathbf{m}_i solutions:

$$r_{\mathbf{m}_i \mathbf{m}_P} = \frac{\sum_{j=1}^n (m_{ij} - \bar{m}_i)(m_{Pj} - \bar{m}_P)}{\sqrt{\sum_{j=1}^n (m_{ij} - \bar{m}_i)^2 \sum_{j=1}^n (m_{Pj} - \bar{m}_P)^2}} \quad (6.4)$$

where \bar{m}_i and \bar{m}_P are the average model and PCA model values inside the window

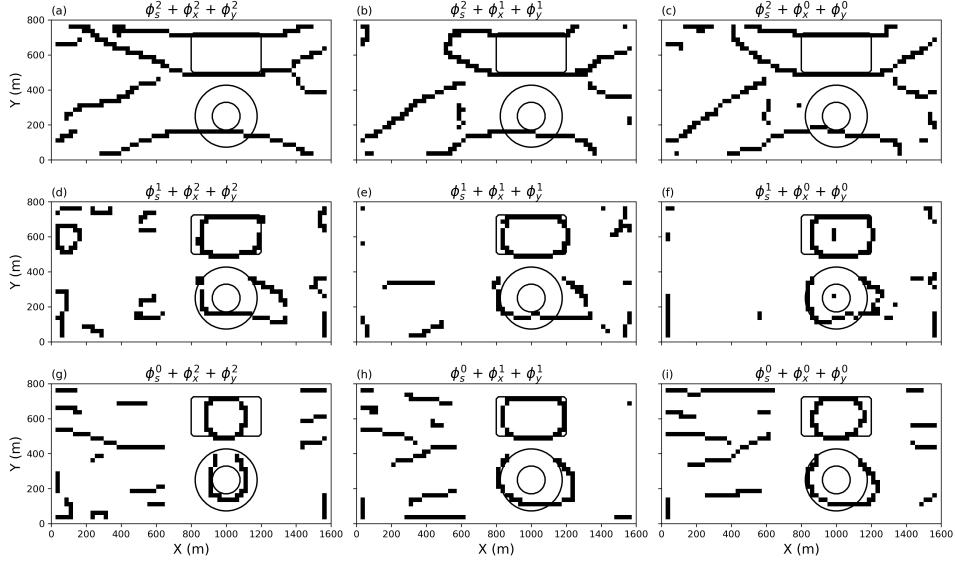


Figure 6.6: Parametric representation of the nine inverted models derived from the Canny edge detection algorithm.

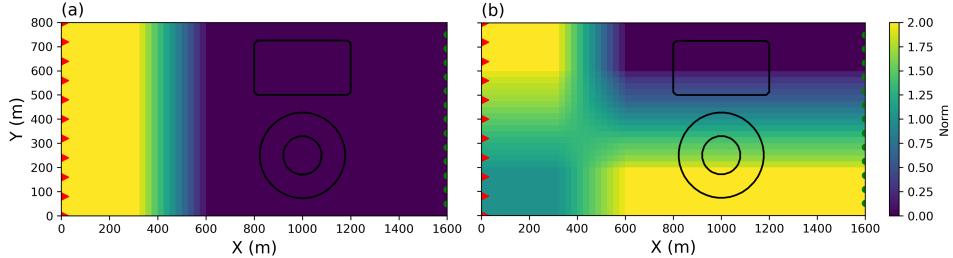


Figure 6.7: Local p -values on the (a) the model norm $\phi_s^{p_s}$ and (b) model gradients $\phi_x^{p_x}, \phi_y^{p_y}$ extracted from the solution space.

denoted by the subscript j . The parameters p_s, p_x, p_y associated with the highest correlation are used in a weighted average as defined in (6.3). The process is repeated over the entire model space. For my example I use 20×20 pixels window. The recovered \mathbf{p}_s and $\mathbf{p}_x, \mathbf{p}_y$ values are presented in Figure 6.7(a) and (b) respectively. I note that the norm on the model gradients is generally larger in the bottom region of the model corresponding to the location of the smooth positive anomaly.

I now define Scaled-IRLS weights on a cell-by-cell basis as:

$$\begin{aligned}\hat{\mathbf{R}}_s &= \text{diag} \left[\gamma_s \odot \left((\mathbf{m}^{(k-1)} - \mathbf{m}^{mref})^2 + \varepsilon_s^2 \right)^{\odot \mathbf{p}_s/2-1} \right]^{1/2}, \\ \hat{\mathbf{R}}_x &= \text{diag} \left[\gamma_x \odot \left((\mathbf{D}_x (\mathbf{m}^{(k-1)} - \mathbf{m}^{mref}))^2 + \varepsilon_x^2 \right)^{\odot \mathbf{p}_x/2-1} \right]^{1/2}, \\ \hat{\mathbf{R}}_y &= \text{diag} \left[\gamma_y \odot \left((\mathbf{D}_y (\mathbf{m}^{(k-1)} - \mathbf{m}^{mref}))^2 + \varepsilon_y^2 \right)^{\odot \mathbf{p}_y/2-1} \right]^{1/2},\end{aligned}\quad (6.5)$$

where $\odot \mathbf{p}_s$, $\odot \mathbf{p}_x$ and $\odot \mathbf{p}_y$ define the element-wise Hadamard power and $\gamma_s \odot$, $\gamma_x \odot$ and $\gamma_y \odot$ are the element-wise scaling multiplications defined in (3.52). In this fashion, the approximated mixed norm regularization can vary on a cell-by-cell basis.

Finally, I proceed with my SVMN inversion. The data are now inverted using the extracted local parameters, applied on a cell by cell basis, and the result is shown in Figure 6.8(a). There is good correspondence with the true model: both the low-velocity blocky anomaly and the high-velocity smooth anomaly are imaged at the right location, with the right shape and near their respective seismic velocity. The normalized data residual in Figure 6.8(b) do not clearly show pant-legs associated with the targets. This is a good indication that most of the important structures have been recovered.

6.1.4 Summary

This example shows that, as long as the data are sufficiently sensitive to the geometry of causative bodies, it is possible to extract preferential inversion parameters. In the original study of Sun and Li (2014), inversion parameters (either ℓ_2 or ℓ_1 -norm on the model gradients) were extracted based on the data residual. In this study, I proposed a different path using modelling trends identified in the solution space. I would argue that the modelling route is a more robust approach and easily applicable to other geophysical methods (EM, seismic) as complex signals from multiple sources can be unravelled by the inverse process.

It is also important to note that I have achieved my final result without direct input from the user, other than setting tuning parameters used in the pattern recogni-

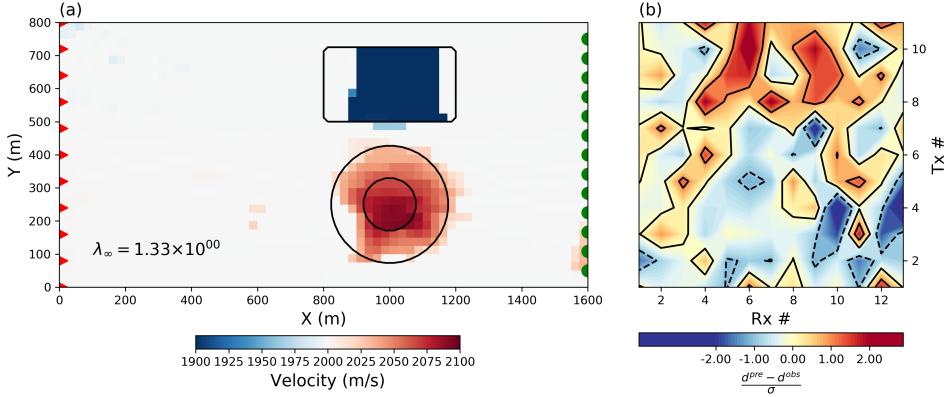


Figure 6.8: (a) The final SVMN inversion result and (b) normalized data residual map. Contour lines are shown in (a) for the true position and shape of the velocity anomalies.

tion phase. The modelling process is therefore completely replicable. The learning process remains sensitive however to choices made by the user. The size and shape of the averaging window is an important aspect that can significantly impact the outcome. Choosing a window that is too large can lump smaller features together, while a window that is too small might introduce unwanted variability. The work presented here does provide however an interesting basis for more advanced work in machine learning.

6.2 Dip and strike estimation

In Chapter 5, I have demonstrated the benefits of using rotated sparse gradient to accentuate geological trends and better image continuous bodies at depth. I have used the fold model shown in Figure 6.9 to showcase the use of surface structural data. In most greenfield exploration settings however such data may either not be available or too scarce for accurate interpolation. Practitioners would benefit from being able to estimate the strike and dip of geological features in a semi-automated way. In this section, I propose to repeat the rotated sparse norm experiment through a learning process. I will make use of a similar strategy elaborated in the previous section such that preferential orientations are extracted from the solution space with a pattern recognition approach.

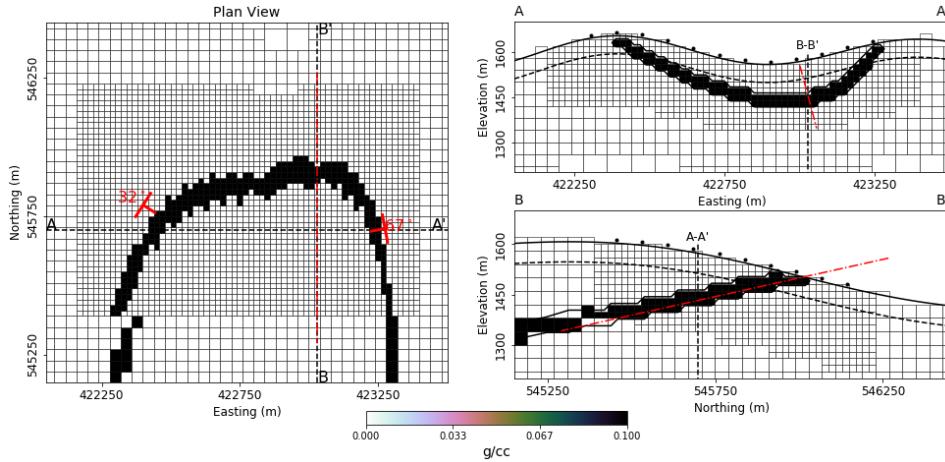


Figure 6.9: Synthetic fold model introduced in Chapter 5 used here to test the dip and strike estimation learning algorithm.

I propose to estimate the angle of rotation in two stages: first along a horizontal plane to determine the strike, then on vertical sections to estimate the dips. I create a workflow divided into five steps:

1. Run an unconstrained inversion with sparse and smooth assumptions ($p_s = 0$, $p_{[x,y,z]} = 2$)
2. Estimate horizontal trends and rotate the objective function on the xy -plane
3. Run a suite of inversions over a range of dip angles ($p_s = 1$, $p_x = 2$, $p_{[y,z]} = 0$).
4. Form an *average* model (\mathbf{m}_A) and extract dip angles locally.
5. Carry out a final SVMN inversion with local p parameters and rotation angles.

I apply this procedure to recover the folded density layer used in Chapter 5 (Figure 5.5). I will attempt to do this without any external input from the user other than general tuning parameters in Step 2 and 4.

Step 1: Unconstrained inversion

Potential field data are most sensitive to lateral changes in density and magnetic properties, so it is generally easier to determine horizontal trends related to geology. I am using a mixed norm inversion ($p_s = 0$, $p_{[x,y,z]} = 2$) to get a first estimate of the geometry of the problem such that physical property contrasts are increased without enforcing assumptions about directionality. While similar analysis could be performed directly on a gridded data map at a much smaller computational cost, the inversion route has proven experimentally to be more robust. It can deal with topographic effects and it can isolate long wavelength trends from near surface anomalies. Since I already have inverted this synthetic example with sparsity assumptions in Chapter 5 (Figure 5.8), I can advance directly to Step 2.

Step 2: Azimuth estimation

From the recovered model, I want to extract directional information. I resort to an image moment algorithm to extract dominant patterns on the windowed portion of the model. I first demonstrate the procedure on the UBC Crest shown in Figure 6.10.

2.1: Density values are converted to a binary image using the python routine `ScikitImage.Skeleton` to extract dominant feature (Lee et al., 1994). The 2D image is reduced to one pixel wide representations by recursively identifying and removing border pixels. The process is carried out until no pixels can be removed without breaking the connectivity of a given feature.

2.2: Position and orientation of dominant features are extracted with an image moment approach (Hu, 1962). For a set window of the image with pixel intensity defined by $\mathbf{I}(x,y)$, the image moment is given by

$$m_{ij} = \sum_x \sum_y x^i y^j I(x,y) . \quad (6.6)$$

The angle defining the principal axis is calculated by

$$\theta = \frac{1}{2} \arctan \left(\frac{2m_{11}}{m_{20} - m_{02}} \right) \quad (6.7)$$

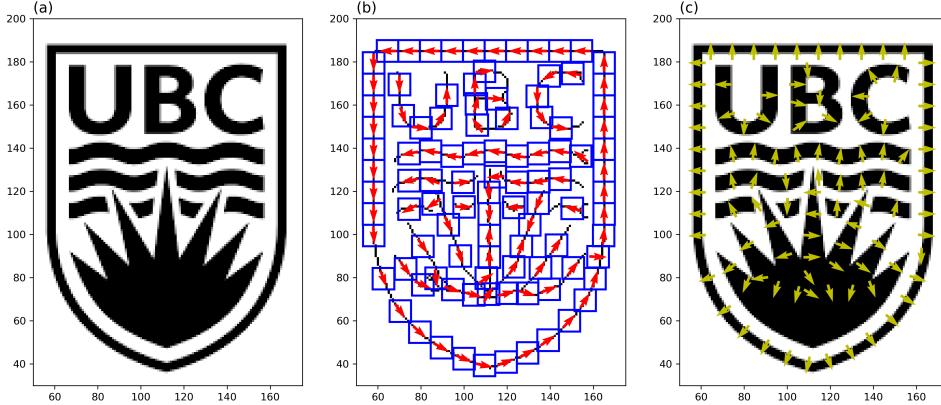


Figure 6.10: Example of an automated azimuth estimation using the UBC Crest.
 (b) The image is first converted to a binary image using the python routine `ScikitImage.Skeleton`. Image moment calculations are performed on moving windows to determine the position and orientation of dominant features.
 (c) Vectors are then rotated by 90° to point along the normal of edges. The normal vectors are checked for consistency in preparation for interpolation.

and the center of mass of the image is given by

$$x_c = \frac{m_{10}}{\sum_x \sum_y I(x,y)} \quad (6.8)$$

$$y_c = \frac{m_{01}}{\sum_x \sum_y I(x,y)},$$

Calculations are only performed for windows that contain at least 5% of non-zero pixels. This process is repeated across the entire section by incrementally moving the window location. Efficiency can be gained by displacing the window based on the calculated orientation. This results in a collection of vectors pointing along with the main features, as shown in Figure 6.10(b).

2.3: Orientation vectors are rotated by 90° to point normal to the edge of features (Figure 6.10(c)). Neighbouring normals are compared to each other for consistency, such that vectors pointing against the general trend are rotated by 180°. This final step is important for a smooth interpolation of angles in 3D space. In complex geological settings, this step may require quality control checks from the user.

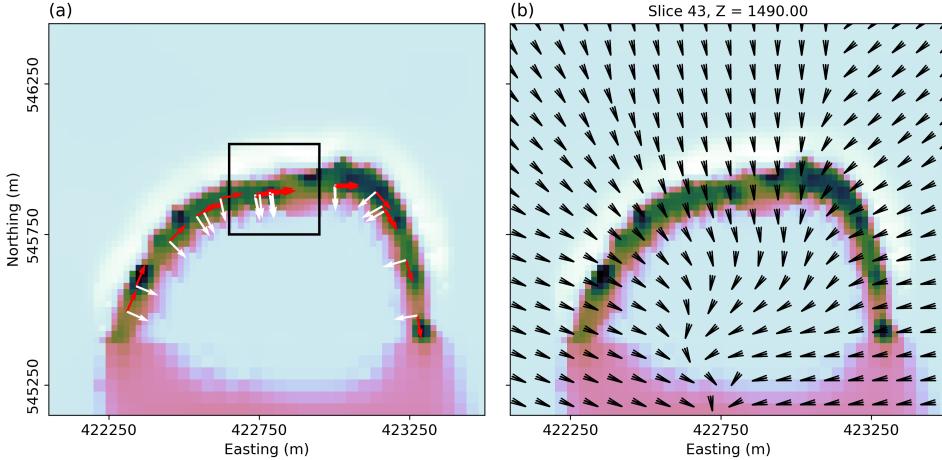


Figure 6.11: (a) Image moment calculations over windowed regions of the inverted density model. Arrows are plotted for the principal direction (red) and normal (white) of the main density anomaly. (b) Normal vectors are interpolated in 3D and used to rotate the objective function.

Carrying out the process described above on a horizontal section of the density model, I recover the orientation vectors presented in Figure 6.11. In this case, I choose a square window 200 m in width that sweeps through a draped section of the density model at 40 m depth (two vertical cells). This choice depends on the discretization and wavelength information present in the data. Going down a few cells below topography generally offers a good compromise between damping the near-surface variations while preserving major trends. Just as for the structural data in Chapter 5, I interpolate the normal vectors in 3D using minimum curvature approach (Briggs, 1974).

Step 3: Model space inversion

Now that I have determined the preferential orientation (normal) of geological bodies, I will attempt to extract preferential dip directions. I proceed with a series of 12 inversions with rotated sparse norms over a range of rotation angles for $\theta \in [-90^\circ, 75^\circ]$ at 15° . Figure 6.12 compares the recovered models along the EW section B-B'.

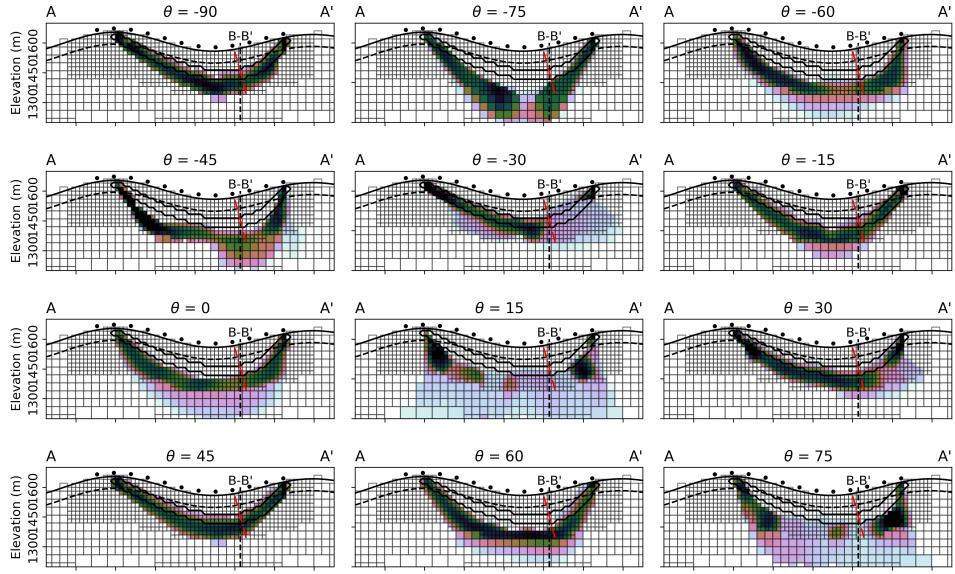


Figure 6.12: Suite of models with directional gradients for $\theta \in [-90^\circ, 75^\circ]$.

Step 4: Dip estimation

From the twelve recovered models, I proceed with PCA and form an average model as prescribed in Section 6.1.2. The calculated average model is shown in Figure 6.13. From this average model, I extract dip information along vertical sections using the image moment strategy presented in Step 2. The resulting angles are extrapolated in 3D and used to rotate the objective function.

Step 5: SVMN inversion

In the final step, I re-invert the dataset with rotated sparse norms. The final solution is shown in Figure 6.14. I note the good match between the solution and the true position of the dense layer. I have obtained this solution without external input, other than the window size and threshold value used in the image moment estimation.

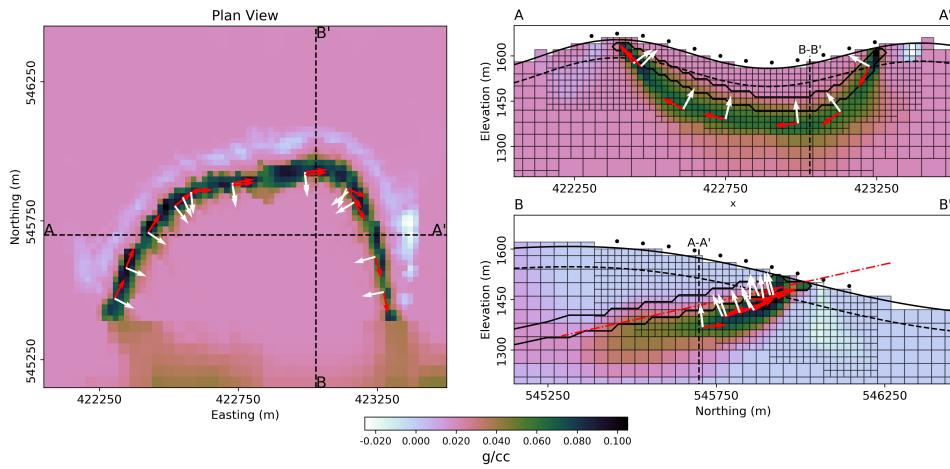


Figure 6.13: Sections through the average PCA model and rotation vectors (arrow) derived from the image moment analysis.

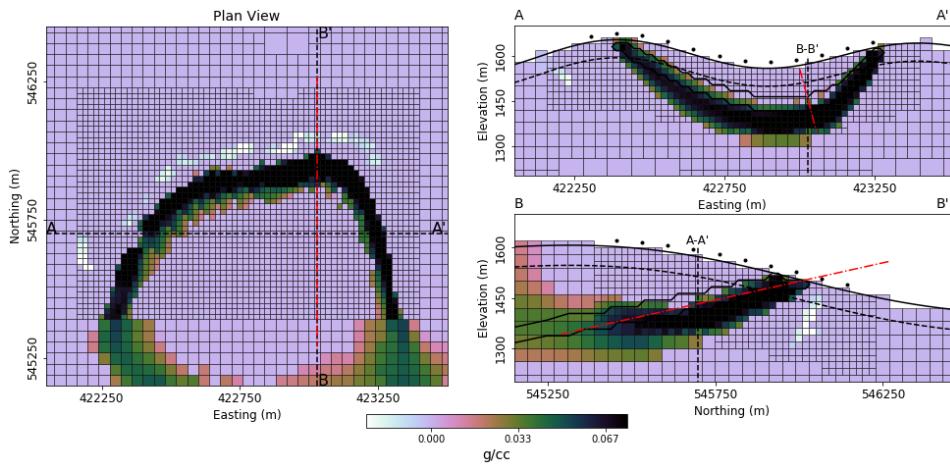


Figure 6.14: (a) Horizontal and (b, c) vertical sections through the recovered model using rotation angles derived from the average PCA model.

6.3 Summary

Leveraging the mixed norm regularization function of Chapter 3 and pattern recognition algorithms, I have explored ways to extract inversion parameters from the solution space. The assumption is made that the shape and position of robust features can be highlighted, as long as the data is sufficiently sensitive to their geometry. I have shown that gravity surveys could be used to determine the strike and dip of geological features with detectable edges. I have demonstrated this by successfully imaging a continuous and folded density layer. This process was only possible because the edges of the anomaly were detectable from the surface. Similar procedures would not have been successful in a strictly layered environment to which potential fields are not sensitive.

The rotation parameters are *soft* geological constraints that reinforce modelling trends without having to specify the location of physical property contrasts. This is an appealing property of the algorithm for greenfield settings as the implementation requires little input from practitioners and easily be automated.

The methodology developed in this chapter is general and can easily be extended to other geophysical methods, such as electromagnetic and seismic data. The rotation parameters could potentially be derived from multiple physical properties. Preferential orientations could form a basis for joint or cooperative physical property inversion.

Chapter 7

Case Study - Kevitsa Ni-Cu-PGE deposit

In the previous chapters, I have provided technical developments to explore the model space, interpret a suite of models and extract dominant inversion parameters. The different components were tested on synthetic examples that showcased specific aspects. In this final chapter, I demonstrate the practical implementation of these advances on gravity and magnetic data sets acquired over the Kevitsa Ni-Cu-PGE deposit. The complex geology of Kevitsa makes it an ideal candidate to showcase my research.

7.0.1 Geological setting

The Kevitsa deposit was discovered in the mid-1980's through extensive exploration programs sponsored by the Geological Survey of Finland. The proven 160 million tons of nickel is economically significant in the region as it is expected to become the largest mine in the country. Figure 7.1 presents a simplified geological map of the Kevitsa-Satovaara intrusive complex adapted from Koivisto et al. (2015). The ore deposit is hosted in a funnel-shaped ultra-mafic olivine pyroxenite (UPXO) unit dipping towards the southwest, dated to approximately 2054 Ma (Gregory et al., 2011; Mutanen, 1997). Directly adjacent to the southwest margin is a large gabbro (IGB) unit formed during a late phase of the intrusion. Similar gabbro, found west of the Satovaara Fault, is likely related temporally. The body

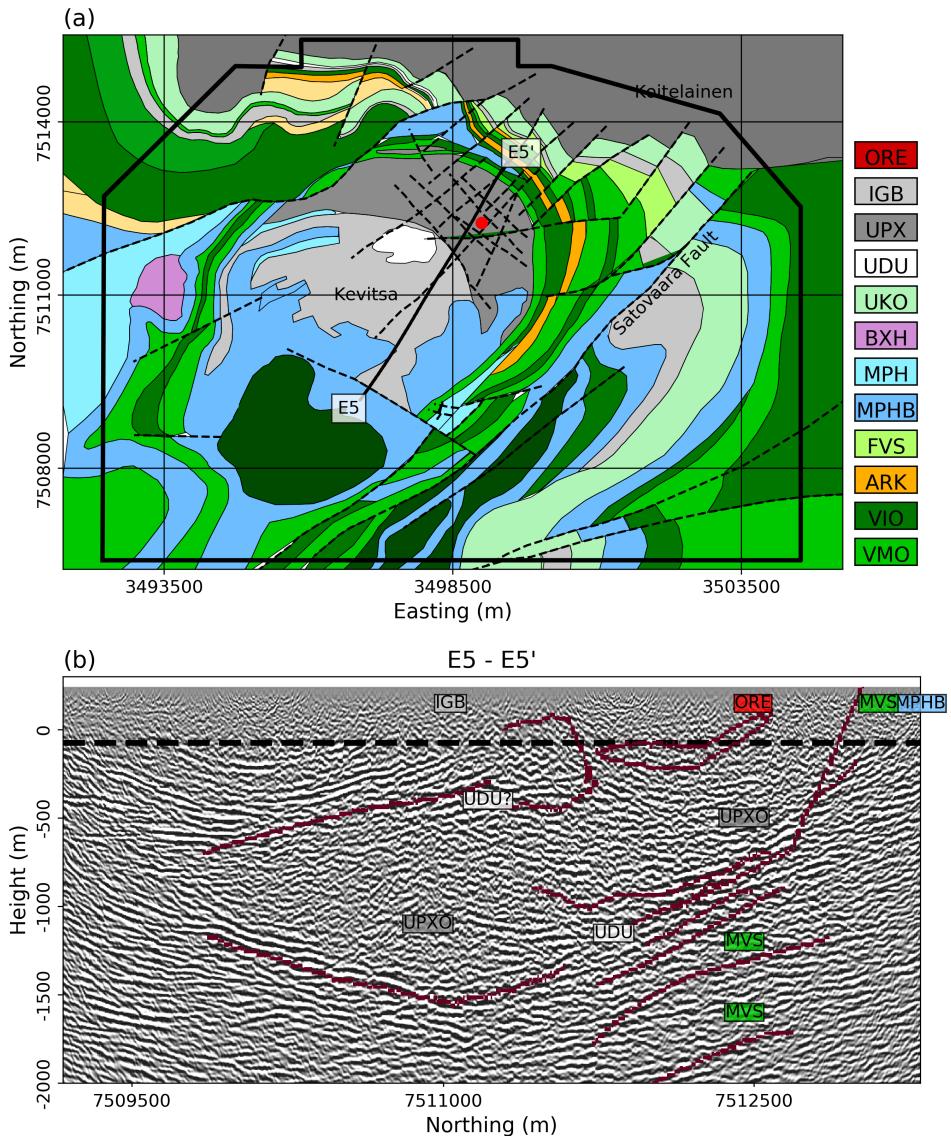


Figure 7.1: (a) Geological surface map of the Kevitsa-Satovaara intrusive complex and (b) 2D seismic reflection survey along with the E5 profile. Interpreted seismic reflectors and geological contact (black) from the interpretation of Koivisto et al. (2015) are shown for reference, as well as the outline of the gravity and magnetic data sets analyzed in this chapter.

intruded a mostly layered sequence of mafic volcanic (MVS) and carbonaceous phyllites (MPH) units. Discontinuous layers of arkose (ARK), arenite (ARN) and felsic volcanic (FVS) units are interbedded within the volcanic sequence. The current folded configuration of the geology is due to large tectonic events that deformed the host stratigraphy around the more competent UPXO block.

It is believed that the disseminated sulphide mineralization within the UPXO was precipitated from the dissolution of Ni-Cu rich proterozoic metasediments, referred to as black schist (MPHB) (Mutunen, 1997). Mineralogical and textural changes observed on core samples within the intrusion reveal the highly heterogeneous nature of the deposit. A highly altered dunite raft (UDU), referred to as the *central dunite*, is located between the IGB and UPXO. It has likely been partially assimilated by the UPXO intrusion.

Seismic surveys acquired over the deposit have been studied extensively and yielded the most complete picture of the deposit to date. Koivisto et al. (2015) interpreted the profiles along side drill hole data to produce a partial 3D surface model of the deposit. The bottom limit of the intrusion is marked by a series of strong reflectors, interpreted as *footwall dunite* units and host stratigraphy (Figure 7.1). Additional UDU units have been intercepted by two deep holes, but it is unclear if this is related to the *central dunite* or to the komatiite (UKO) layers found in the host stratigraphy. The southern continuation of the Kevitsa intrusion can clearly be seen along the E5 seismic profile, as a thick 1km non-reflective zone, but no bore hole is available to confirm the lateral extent.

Petrophysical data

Along with geophysical data, physical property measurements collected along 279 bore holes were made available for analysis; they include density, magnetic susceptibility, conductivity, seismic velocity measurements. Figure 7.2 presents whisker plots summarizing the measured density and susceptibility grouped by lithological units. The interpretation of core logs is challenging due to the complex geology and broad terminology used to describe the rocks. In order to focus my analysis, I generalize the rock classification into 10 main geological groups as summarized in Table 7.1. These groups were defined in terms of relative age and expected physical

property contrasts. I note a few obvious trends:

- Intermediate and felsic volcanic rocks (VIO) show low density and moderate to low magnetic susceptibility
- The komatiite (UKO) and dunite (UDU) show significantly higher susceptibility but low density
- The intrusive rocks (UPX) and (IGB) are both dense and susceptible

Overall I should expect high physical property contrasts between the different lithologies. While I will not directly incorporate this information in the inversion, it will be used to define research questions and for the final interpretation.

7.1 Geophysical data

Kevitsa is an interesting case study considering the large collection of data sets that were acquired and made available to researchers: bore hole petrophysical measurements (Montonen, 2012), seismic refraction (Koivisto et al., 2015), direct-current resistivity and magnetotelluric (Le et al., 2016), airborne time-domain EM surveys (VTEM 2009, SkyTEM 2010). In this section, I focus on the ground gravity dataset and airborne magnetic data collected by the VTEM survey shown in Figure 7.3(a) and (b) respectively.

From simple visual inspection, I note some obvious connections between the potential field data and the surface geology:

- High gravity and magnetic data correlated with the UPXO and hydrothermal BXH
- Moderate response from the VMO
- Weak fields over most of the MPH units.
- Magnetic high and gravity low over the komatiite (UKO),
- Large negative magnetic anomaly within the UDU and near the southern edge of UPXO

	Code	Description	Density	Susceptibility
■	OVB	Overburden	Low	Medium
■	IGB	IGB: Gabbro IPG: Pegmatite IGBO: Olivine gabbro IDI: Diorite IDB: Diabase IMO: Intrusive (mafic) IGBM: Magnetite gabbro	Medium Medium	Medium High
■	UPX	UPXO: Olivine peroxinitite UOO: Ultramafic (Undiff.) MPE: Metaperidotite UWB: Websterite	Medium	Medium
■	UDU	UDU: Dunite UPE: Peridotite USP: Serpentinite	Low	High
■	UKO	Komatiite	Low	High
■	BXH	BXO (undiff.) BXHC: Hydrothermal (crackle)	Low High	Medium
■	SOO	MAB: Albítite MAM: Amphibolite MQZ: Quartzite MSCSD: Schist	Low	Low
■	MPH	MPH: Phyllite MPHB: Black Phyllite MSCBK: Black Schist MHF: Hornfels	High	Low
■	VMO	VMO: Volcanic Mafic VBA: Basalt VTUM: Volcanic tuff	High	Low
■	VIO	VIO: Volcanic Intermediate VTUI: Volcanic tuff VAN: Andesite VOO: Volcanic (undiff.)	Low	Medium

Table 7.1: Summary table grouping the various lithological units logged from bore hole. Expected density and magnetic susceptibility contrasts are derived from Figure 7.2.

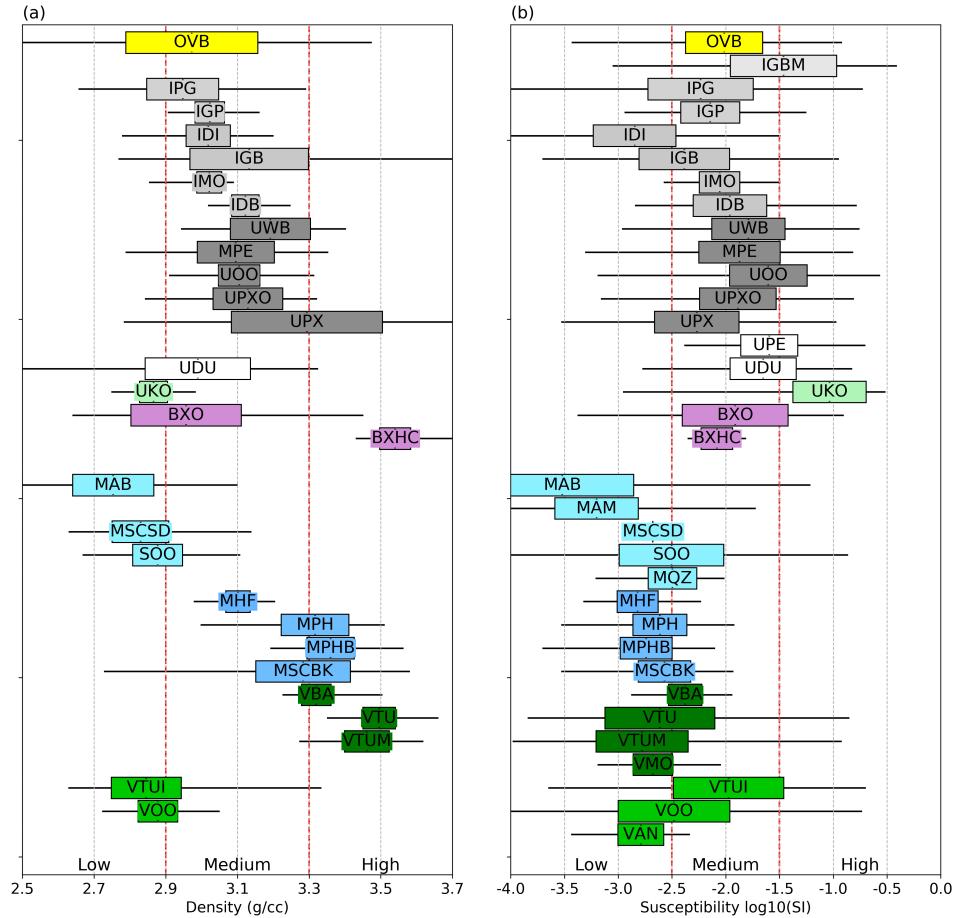


Figure 7.2: Whisker plot of logged (a) density and (b) magnetic susceptibility measured along 279 boreholes. The coloured boxes have a width scaled by the calculated standard deviation and centered on the mean value for all intercepts belonging to the same lithological classification. The black lines on either side define the minimum and maximum values. The different lithologies are colour coded and grouped based on relative age and similarities in physical properties.

The strong negative magnetic fields observed over the dunite unit are of particular interest for this study, as it is likely related to remanent magnetization. Analysis from core samples by Montonen (2012) reported large Koenigsberger ratios and reversed magnetization direction in the UDU unit as summarized in Table 7.2. It is important to note that large Koenigsberger ratios were also measured in the

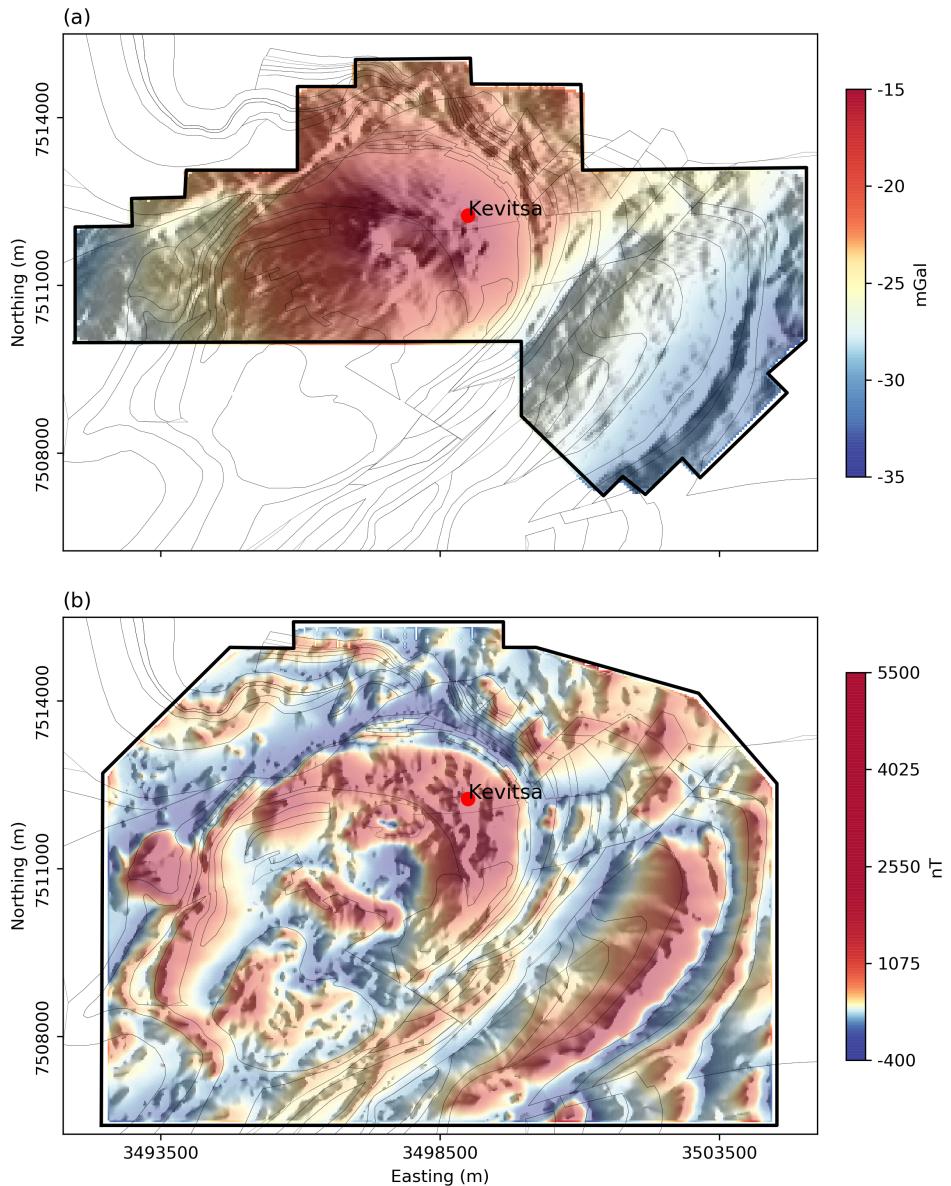


Figure 7.3: (a) Ground gravity and (b) airborne magnetic data map acquired over the Kevitsa Ni-Cu-PGE deposit, Finland. Color ranges are histogram equalized and sun shaded to highlight details. The location of the Kevitsa mine (red), geological boundaries (black) and faults (dash) identified from surface mapping are shown for reference.

Hole ID	Interval	κ (SI)	Inc. (\circ)	Q
KV297	0-52.9	0.034	$-42.4 \pm 18.9^\circ$	[2, 10]
KV200	29.9	0.038	-50.9	5.4

Table 7.2: Intervals along boreholes KV200 and KV297 reporting significant remanent magnetization (Montonen, 2012)

lower UPXO unit, although susceptibility values remained small. In the absence of oriented core, no magnetic declinations were provided. From forward modelling of magnetized sheets, Montonen (2012) estimated that a magnetized unit with effective susceptibility 0.82 SI and orientated [$I = -42.5^\circ, D = 240^\circ$] could be responsible for the observed negative magnetic anomaly.

7.1.1 Modeling objectives

Due to the complex and disseminated nature of the mineralization, substantial modelling efforts continue to be invested to better understand the intrusion. Most of the drill holes made public are concentrated within the UPXO, which leaves a few questions unanswered

1. Can potential fields be used to define the 3D geometry of the intrusion?
2. Can we distinguish variations in physical properties within the UPXO related to mineralization?
3. Can we characterize the central dunite and possible extension at depth?
4. Can the magnetic response be used to infer tectonic deformation?

7.2 Gravity inversion

I first invert the ground gravity data set acquired over the Kevitsa deposit shown in Figure 7.3(a). I will attempt to model the 3D geometry of the intrusion and host stratigraphy based on density contrasts. The survey consists of 28,860 gravity stations collected at 20 m intervals along east-west and north-south lines spaced at 100 m. Data were terrain corrected with a reference density of 2.67 g/cc. I design

an OcTree mesh for the inversion with 25 m cells in the core region extending 2 km at depth. After several inversion trials, data uncertainties are set to 0.1 mGal.

As a first pass, I invert the data with the linear ℓ_2 -norm regularization. Figure 7.4(a) presents a section through the recovered density model at ≈ 200 m depth. I also extract a vertical section along the seismic profile E5 in Figure 7.4(b). I note that the bulk of the density anomaly is located inside the upper portion of the UPXO. As expected from smooth assumptions, the edges of the anomaly are poorly defined.

Clearly the assumption of a smooth density distribution is not sufficient to establish a clear correlation between the known lithology and density. Taking advantage of the methodology developed in this thesis, I will attempt to improve the definition of geological domains. First I need to define rotation angles in 3D to better represent the folded geology of Kevitsa. In this case, I do not have to learn the orientation as I have access to geological interpretation at the surface and a seismic profile at depth. The folded geology of Kevitsa appears to be more or less symmetric radially around the center of the intrusion. I will, therefore, assume that vector components extracted from the vertical section can be interpolated in 3D. Using the method of image moment introduced in Chapter 6, I extract orientation of major contacts over the study area. Figure 7.5 presents the orientation vectors used for the rotation of the objective function.

I then proceed with nine inversions over a range of mix-norm penalty functions for $p_s, p_{y',z'} \in \{0, 1, 2\}$, $p_{x'} = 2$. Only the normal component of the geological contacts is fixed to the ℓ_2 -norm. Horizontal and vertical sections through the nine recovered density models are shown in Figure 7.6 and Figure 7.7 respectively. Residual data maps are shown in Figure 7.8. Correlated residuals indicate that some of the short wavelength information is under fitted.

In order to simplify the interpretation, I overlay iso-contour values of low and high density using the 5th and 95th percentile of density values for each of the nine independent inversions. From Figure 7.9 I make a preliminary interpretation.

- The recovered density within the UPXO unit appears to be highly variable. The known ore deposit appears to be sitting within a region of moderate density.

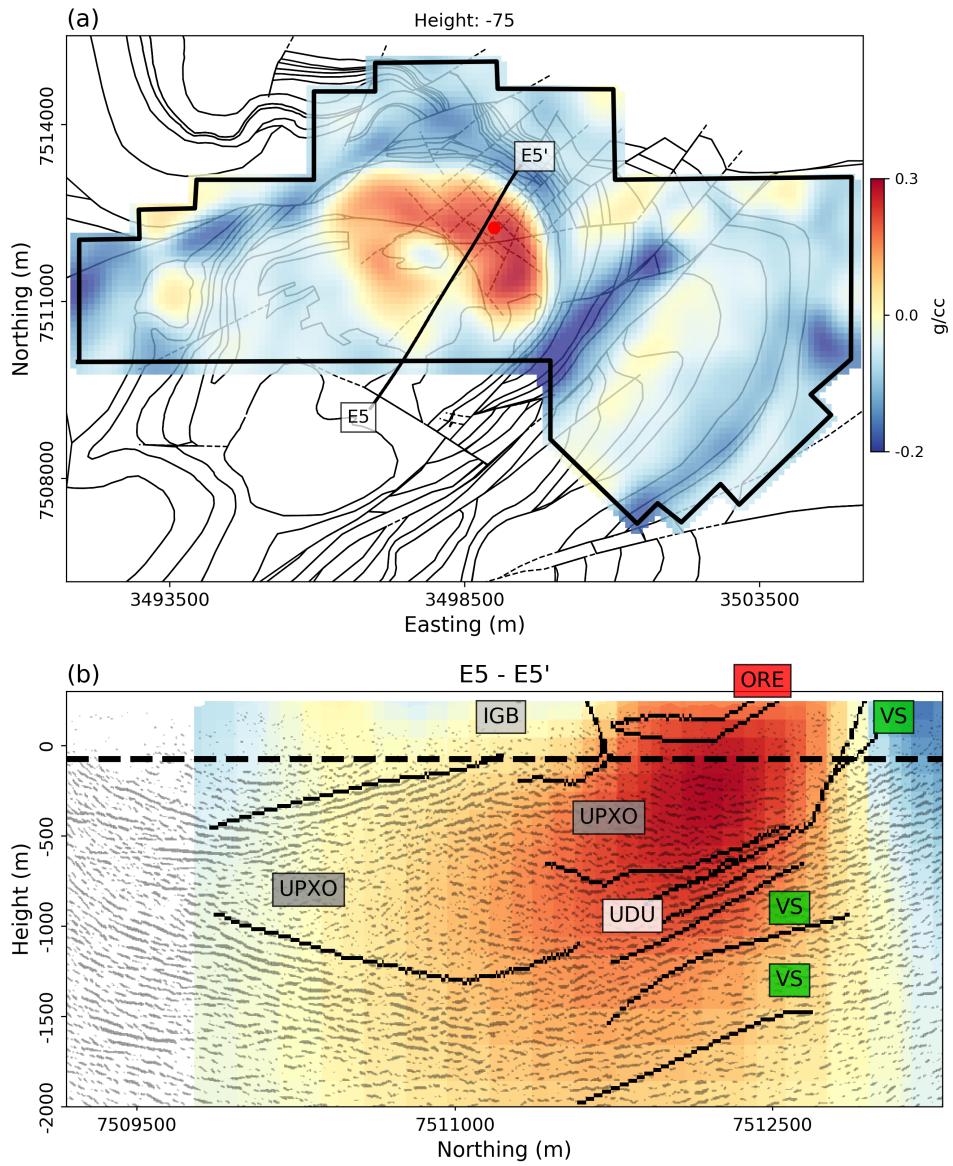


Figure 7.4: (a) Horizontal section at ≈ 200 m depth below topography through the recovered density model after convergence of the algorithm for $p_s = p_x = p_y = p_z = 2$. (b) Vertical section of the density model overlaid on the 2D seismic reflection line E5. Interpreted geological contacts (black) are shown for comparison.

- Density lows are strongly correlated with mapped schist (SOO) and phyllite (MPH) units on the outer perimeter of the intrusion. The same does not apply to the *black schist* unit (MPHB) found immediately south of Kevitsa. While MPH and MPHB appear to have been used interchangeably to build the geological map, the recovered density appears to be a distinguishing property.

The most important aspect of this result is the general trend observed for the main intrusion that extends south-east below the gabbro. This result is interesting as I managed to recover models that are in good agreement with the seismic interpretation of Koivisto et al. (2015), even though no information about the depth of the intrusion was included in the inversion. *Soft* geological constraints, in terms of general trend, were sufficient to push the lateral extent of the density anomaly in accordance with the gravity data.

7.3 Magnetic inversion

I follow my analysis with the processing of magnetic data collected during the 2009 VTEM survey (Figure 7.3(b)). The inducing field parameters at the time of acquisition were $B_0 = [A : 52,800 \text{ nT}, I : 77.5^\circ, D : 12.2^\circ]$. Similar to the gravity inversion, I design an OcTree mesh with 50 m core cells extending 2 km at depth. I determined experimentally a 10 nT uncertainty floor value to the data.

I first invert the TMI data with the conventional smooth susceptibility assumption ($p_s = p_x = p_y = p_z = 2$) and ignore the effect of remanence. From sections through the recovered susceptibility model, presented in Figure 7.10(a) and (b), I note discrepancies with the known geology:

- In plan view, the arc shaped anomaly, SW of the deposit, is recovered outside the mapped UKO unit.
- Along the E5 seismic section, no susceptibility anomaly is recovered over the *central dunite*. This directly contradicts the core sample measurements made by Montonen (2012).
- The shape and extent of the large anomaly correlates poorly with the UPXO unit interpreted by Koivisto et al. (2015) from seismic reflectors.

- The residual data map shows strong correlation with the location of strong negative magnetic fields.

I conclude that the magnetic response observed at Kevitsa cannot solely be attributed to an induced magnetization.

To address issues posed by remanence, I proceed with the MVI-S algorithm. I perform a series of nine inversions with varying sparsity measures to assess the variability in the magnetization model. Starting from a common ℓ_2 -norm MVI-C model, I sequentially vary the combination of norms applied to the amplitude and its derivatives for ($p_{\rho_s}, p_{\rho_{xyz}} \in [0, 2]$) and I apply the same rotation parameters as previously used for the gravity inversions. In this regard, the sparsity and rotation parameters tie the density and magnetization inversions together. While not jointly inverted, I expect some correlation between the two physical properties. Horizontal and vertical sections through the recovered nine magnetization models are shown in Figure 7.11 and 7.12 respectively. Residual data maps are shown in Figure 7.13 for all nine inversions.

To simplify the analysis, I superimpose the 90th percentile iso-value of amplitude for each of the nine models (Figure 7.14). I calculate an average magnetization direction (white) and standard deviation on the angle (red). I observe the following:

- The known ore deposit appears to sit within a volume of low magnetization.
- Parts of the central dunite unit appear to be reversely magnetized [$\kappa_e = 0.09 \text{ SI}$, $I = -52^\circ \pm 15^\circ$, $D = 246^\circ \pm 5^\circ$] This is in excellent agreement with the laboratory results published by Montonen (2012)
- The vertical magnetic anomaly between the IGB and UPXO unit, likely related to the central dunite unit, appears to be plunging towards SE, potentially extending below the UPXO unit as hypothesized by Koivisto et al. (2015).
- Strong magnetization recovered along the outer-shell of the ultra-mafic intrusion appears to be pointing radially outward. Largest magnetization appears to originate below the UPXO unit.

- Similar magnetization direction pointing normal to the arc-shaped peridotite unit.

The last two remarks are interesting for a few reasons. First, strong magnetization below the base of the ultra-mafic supports the presence of magnetic UDU layers. Secondly, the orientation of magnetization pointing normal to the base of the unit may be indicative of past tectonic deformation. Under the assumption that the remanent magnetization component had been fairly uniform within the layered UDU, UKO and UPXO unit at the time of formation, then the current radiating magnetization pattern would be caused by subsequent folding of the units. If this is the case, then this is one of the first times that airborne magnetic data would have been used as a marker for tectonic deformation of elongated and folder geological units.

While my modelling of the central dunite unit agrees with published laboratory measurements, the cause for this reverse magnetization direction remains unclear. No other lithological units appear to share the same orientation. Re-magnetization after the emplacement of the ultra-mafic intrusion is unlikely as a similar reversed polarity pattern would also be expected elsewhere at Kevitsa. I speculate that the dunite block could be related to the lower UDU unit, which would have been rotated to its current sub-vertical location.

7.4 Summary

The inversion of gravity and magnetic data over the Kevitsa deposit yielded valuable information. First from the gravity inversion, I have imaged the UPXO at depth and extending below the gabbro unit. Second, variations in density and magnetization inside the UPXO unit is of interest as it appears to correlate well with the known mineralization. This could potentially be used for future exploration work.

Third, the average magnetization model confirms that the central dunite unit is associated with strong reversed magnetization oriented roughly [$\kappa_e = 0.09$ SI, $I = -52^\circ \pm 15^\circ$, $D = 246^\circ \pm 5^\circ$]. There is also a strong indication that it could be connected to the lower dunite units, although the magnetization direction would require the central dunite to have undergone a rotation of almost 180° .

Potentially the most significant outcome of this case study is the recovered

magnetization pointing normal to the base of the UPXO unit. A similar radiating pattern is observed in the komatiite unit east of the deposit. If confirmed by laboratory measurements, this result would be one of the first cases of paleomagnetism based on the inversion of airborne magnetic data over folded geology.

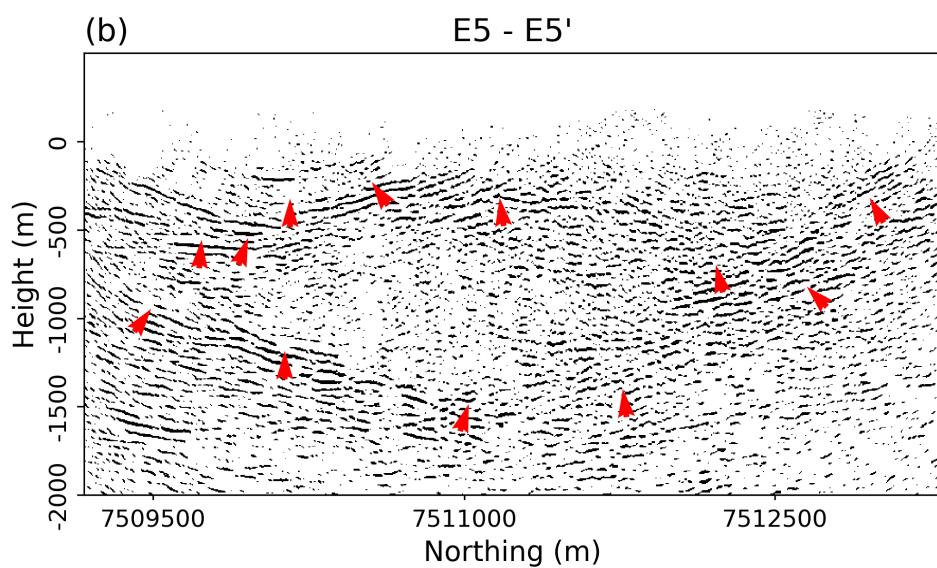
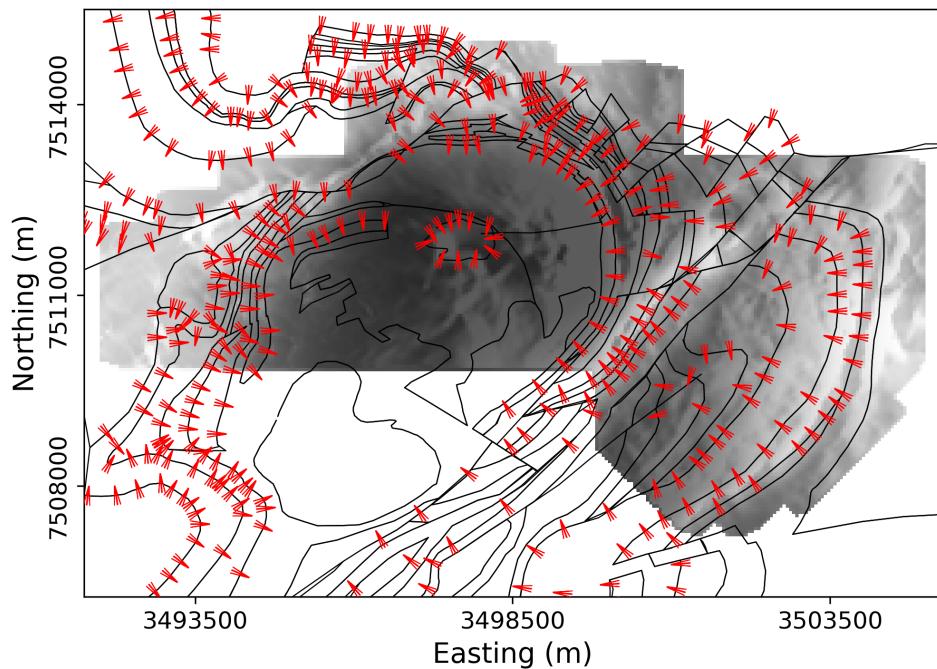


Figure 7.5: Rotation parameters (normals) derived from (a) surface geology and (b) major reflectors detected on the seismic profile E5.

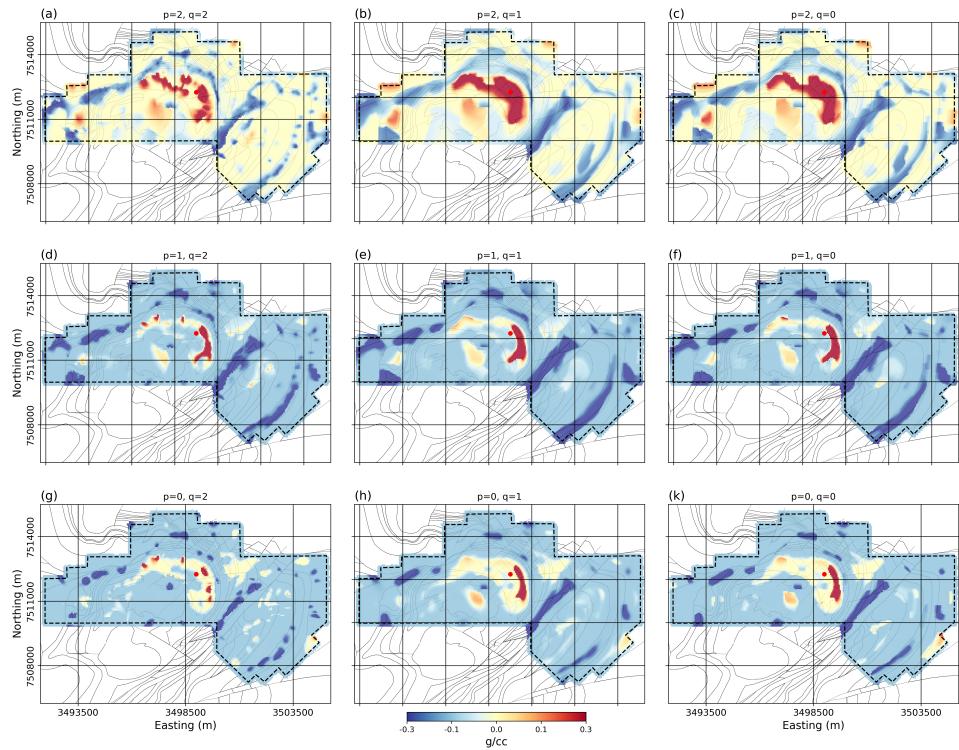


Figure 7.6: Horizontal sections at ≈ 200 m depth below topography through the density models recovered over a range of ℓ_p -norms applied on the model and modeled derivatives for $p_s, p_{x,y,z} \in \{0, 1, 2\}$

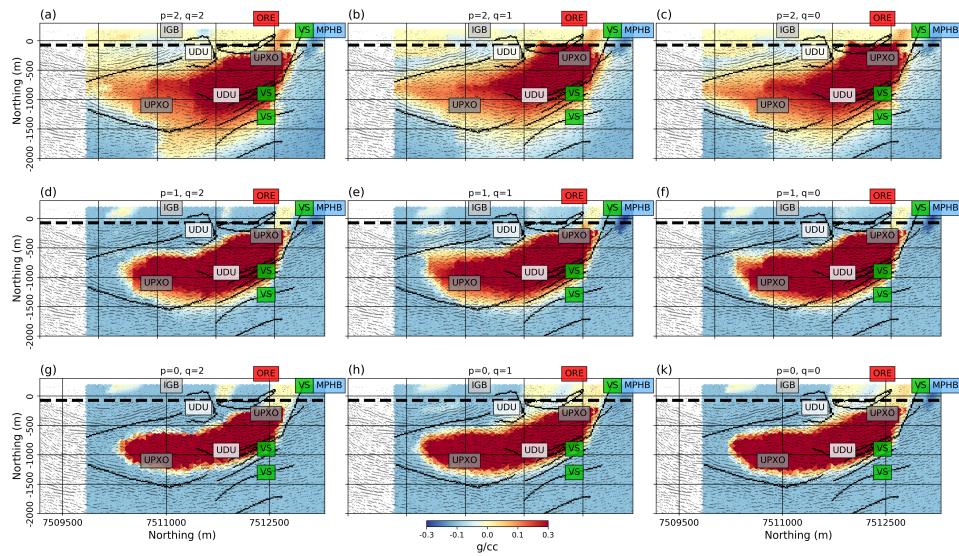


Figure 7.7: Vertical sections overlaid on the 2D seismic reflection line E5 through the density models recovered over a range of ℓ_p -norms applied on the model and modeled derivatives for $p_s, p_{x,y,z} \in \{0, 1, 2\}$

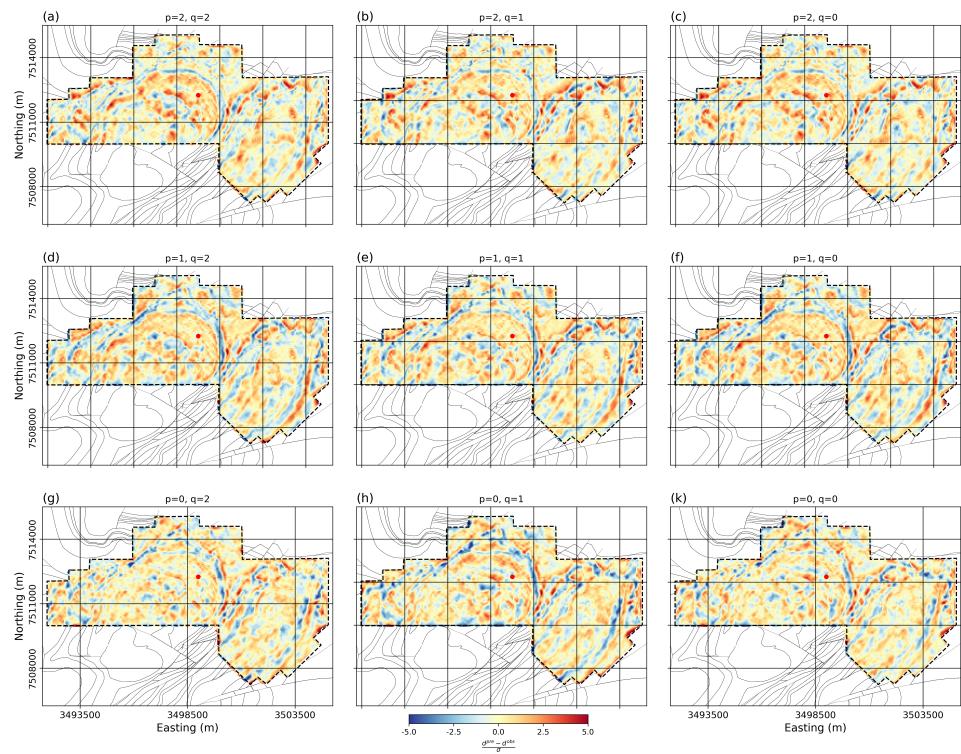


Figure 7.8: Residual data maps for the nine gravity inversions.

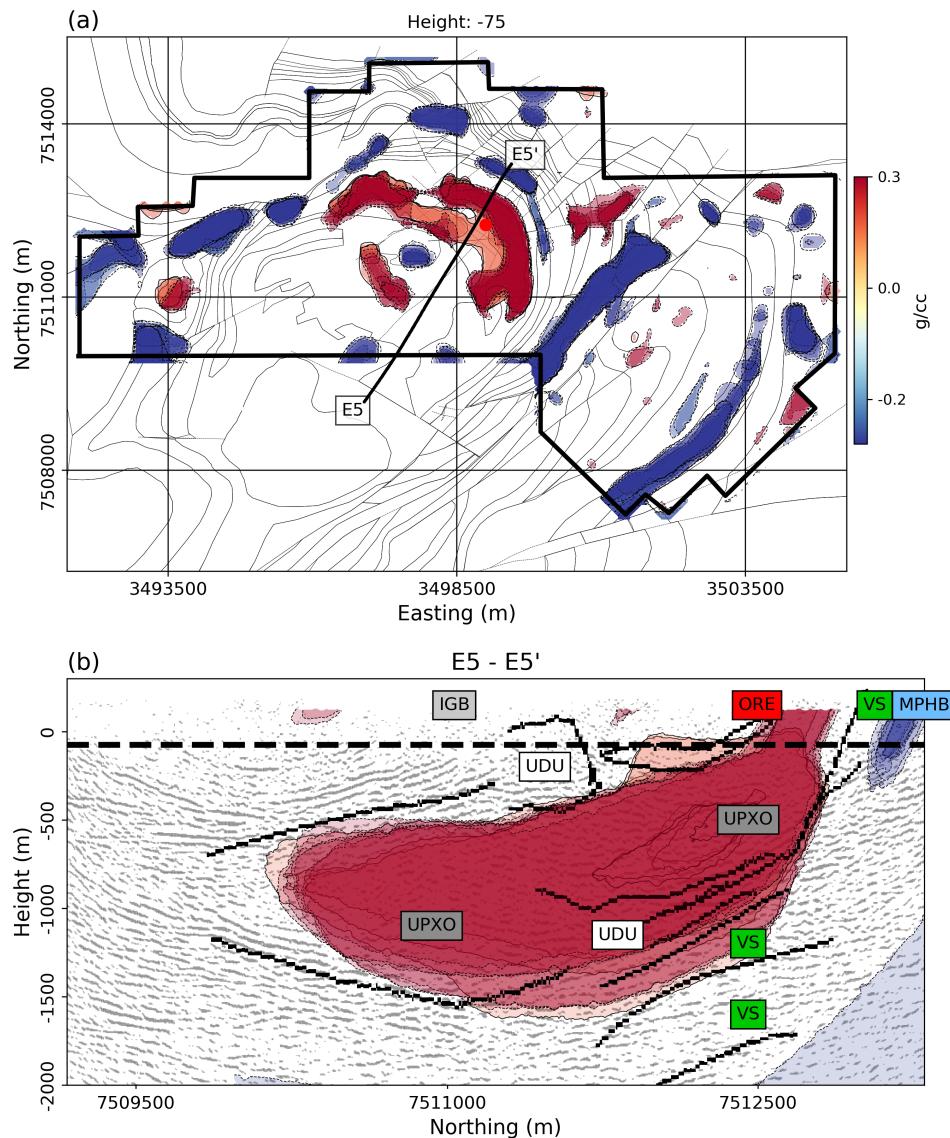


Figure 7.9: (a) Horizontal and (b) vertical section comparing iso-surfaces for the 5th and 95th percentile of anomalous density values recovered from the nine mixed norm inversions for $p_s, p_{x,y,z} \in \{0, 1, 2\}$.

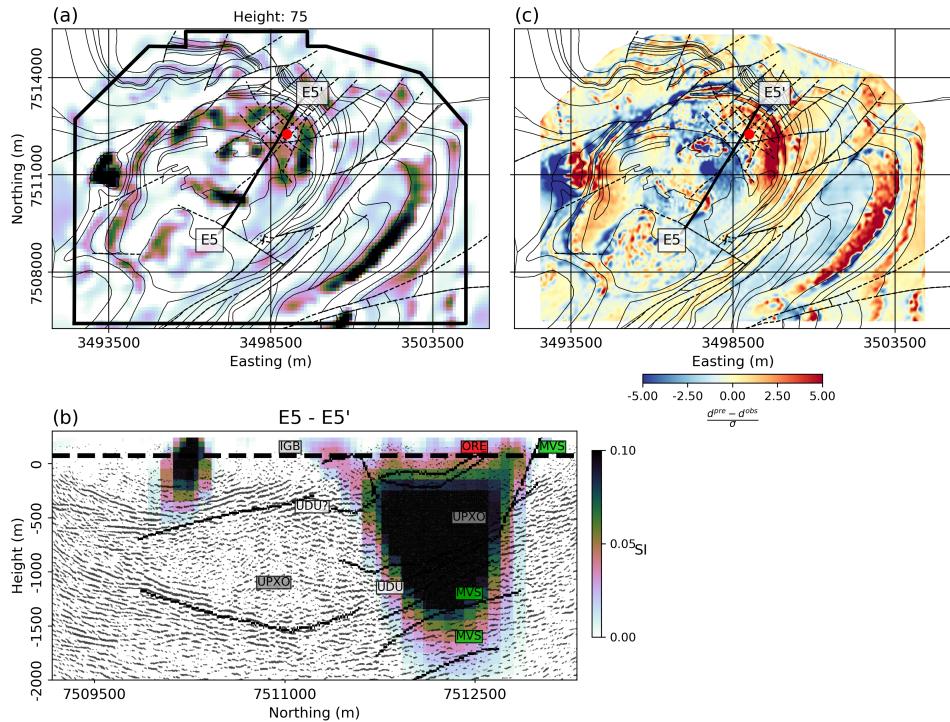


Figure 7.10: (a) Horizontal and b) vertical sections through the recovered susceptibility model that ignores the effect of remanence. Lithological contacts (black) identified by Koivisto et al. (2015) are shown for reference. A large dome-shaped zero-susceptibility anomaly is recovered at the center of the Keivitsa intrusion, likely due to remanence. (c) Residual data map shows strong correlation with the negative anomalies.

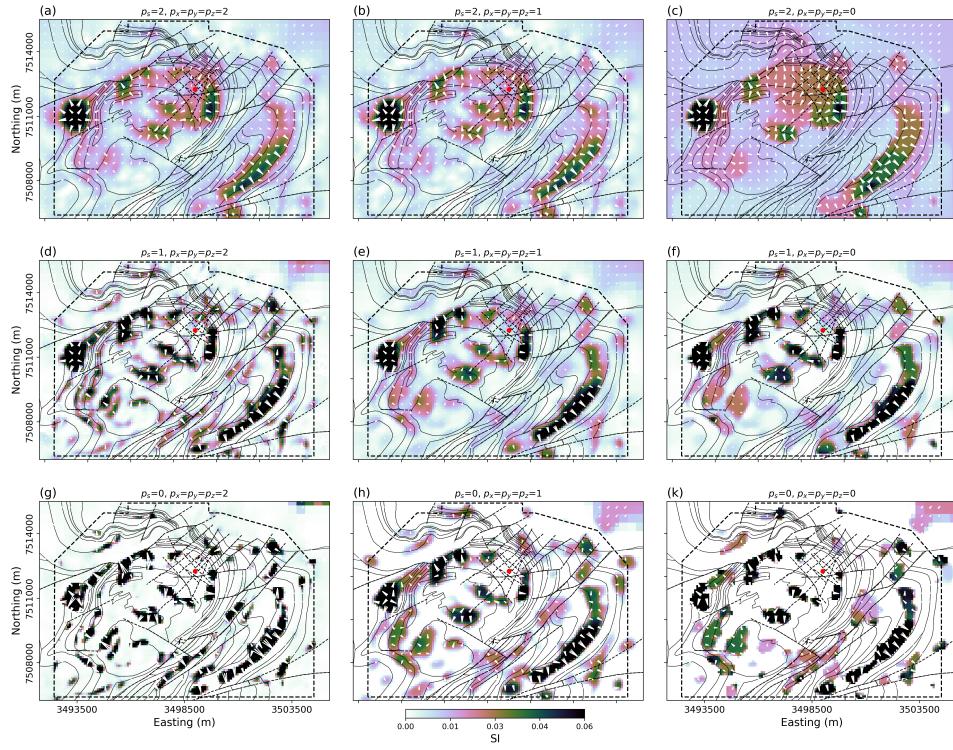


Figure 7.11: Horizontal sections at ≈ 300 m below topography for a suite of models using various sparsity assumptions put on the amplitude of magnetization for $p_s, p_{x,y,z} \in [0; : 2]$. Norm measures on the magnetization angle are fixed to $p_{x,y,z} = 0$ in order to promote uniform magnetization

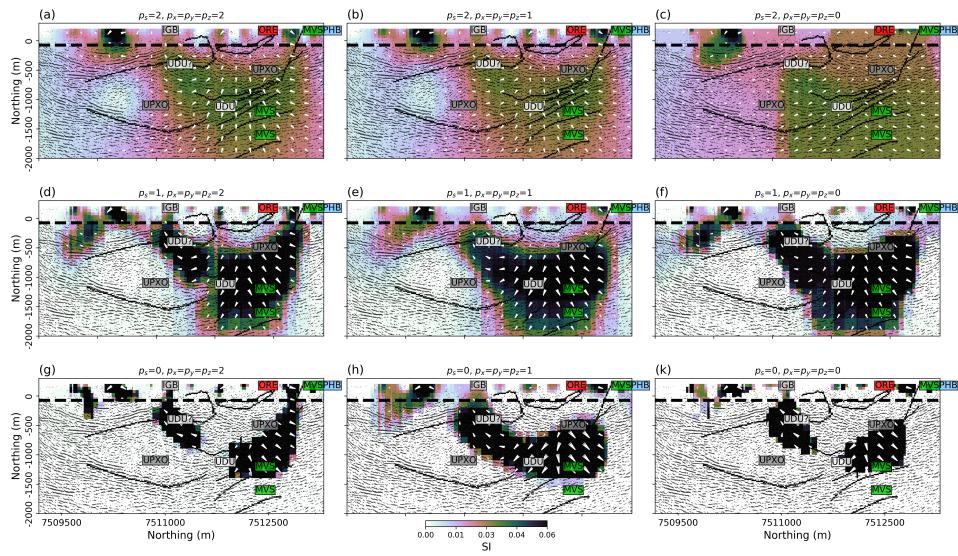


Figure 7.12: Vertical sections through the recovered magnetization models. Effective susceptibilities are concentrated within the UDU and on the lower margin of the UPXO unit.

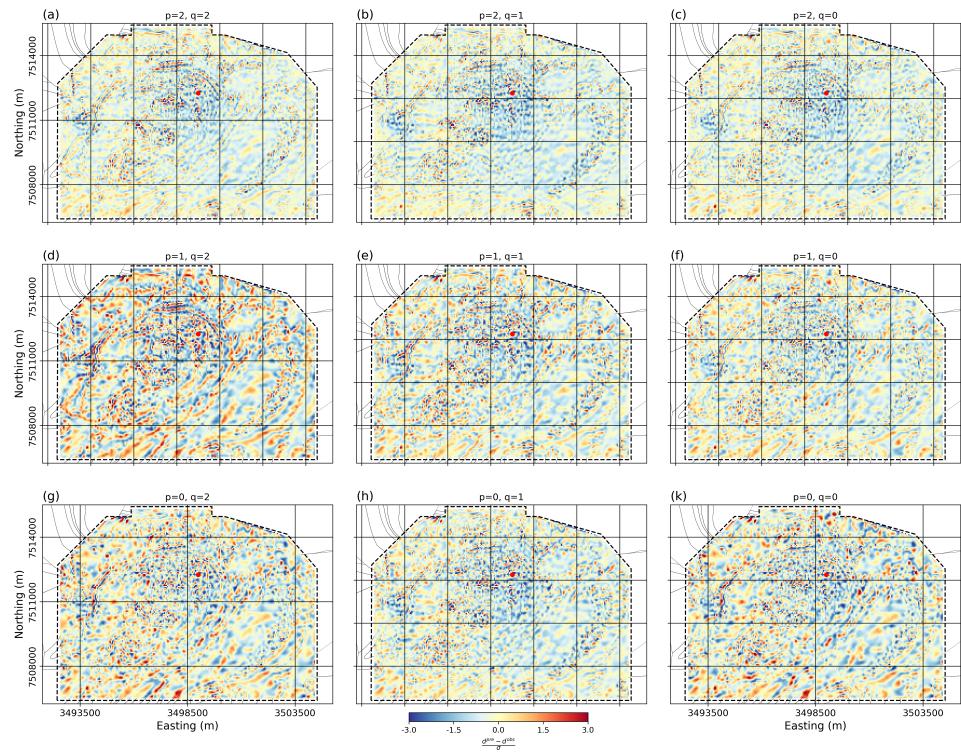


Figure 7.13: Residual magnetic data maps for the nine MVI-S inversions.

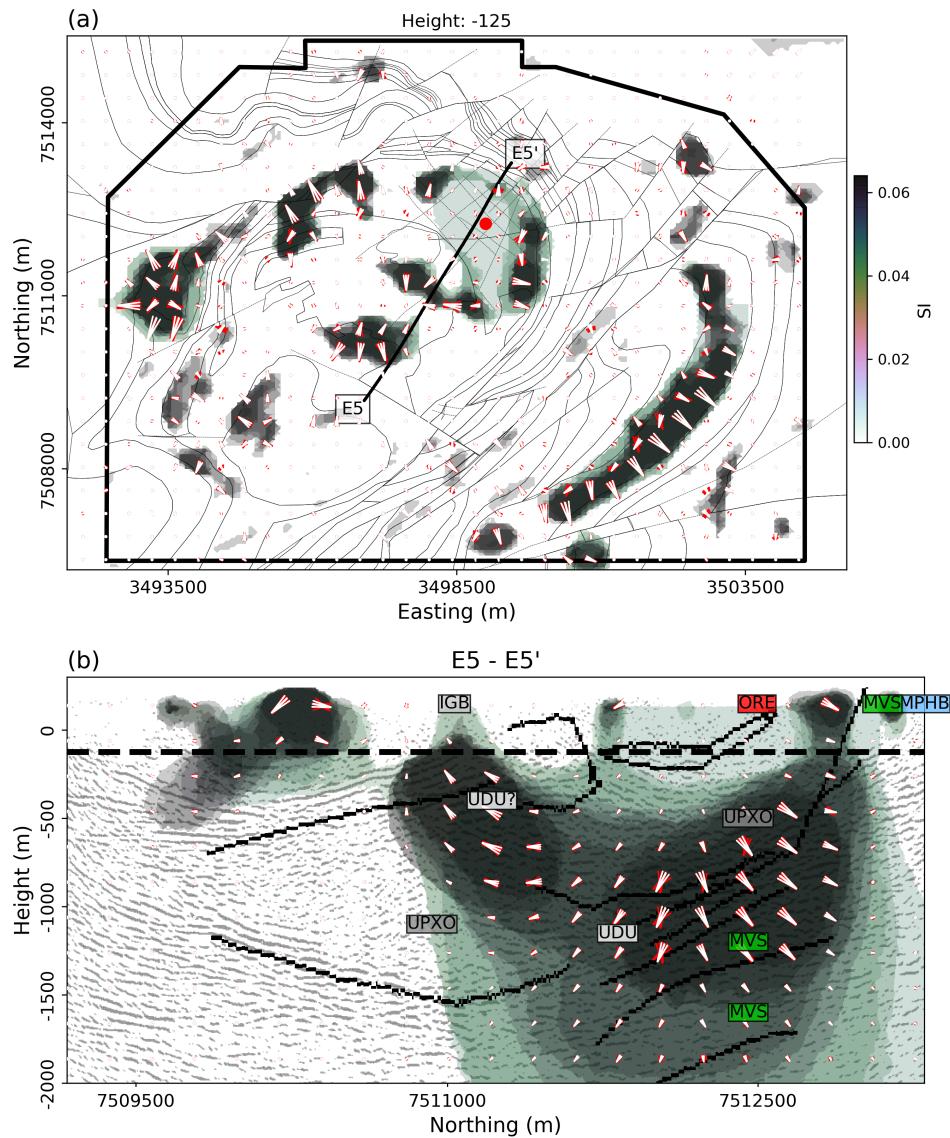


Figure 7.14: (Top) Horizontal and (bottom) vertical sections through iso-contours of magnetization recovered from nine mixed ℓ_p -norm inversions. Magnetization orientation (white) and standard deviation on the angle (red) are shown.

Chapter 8

Conclusion

The overarching goal of this research thesis was to facilitate the interpretation of potential field data acquired over complex geology and to extract as much information out of the data through a semi-automated learning process. These goal was motivated by technical limitations encountered with conventional inversion methodologies. Geophysical inverse problems are inherently non-unique and the character of the solution depends on assumptions set by the user. Smooth physical property inversions yield models that poorly represent sharp geological contact, while sparsity assumptions generally yield simplistic blocky anomalies. The range of possible solutions available for interpretation has generally been fairly limited. More experienced users may be able to improve the interpretation with geological constraints, but building and testing different scenarios remains a laborious process that is difficult to track. In this regard, the methodology presented in this thesis sit somewhere between the blind unconstrained inversion and the expert-driven geological inversion. I have developed a methodology and software that can generate a suite of models that honor *soft* geological assumptions.

In Chapter 2, I reviewed the numerical implementation of gravity and magnetic forward modeling in integral form. I tackle numerical limitations associated with the storage and manipulation of large dense matrices. The memory footprint of potential fields problem is reduced at two levels. First, I borrow the mesh decoupling strategy previously used in electromagnetic modeling. The global forward problem is broken down into tiles each associated with nested Octree mesh. Secondly,

I leveraged open-sourced technologies for out-of-core storage of dense matrices. Used in concert, the two advancements allowed to run large forward problems without the need for compression.

In Chapter 3 I reviewed some of the work I had previously investigated during my Master's thesis regarding the implementation of a mixed ℓ_p -norm regularization. I identified a new scaling strategy based on the maximum partial derivatives of individual regularization functions such that multiple penalties could impact the solution. The robust implementation of mixed norm assumptions is at the core of this research as it allows me to generate a suite of solutions and explore the model space in a consistent manner. This chapter resulted in the research paper Fournier and Oldenburg (2019a).

Chapter 4 builds upon the magnetic vector inversion in spherical coordinates first introduced by Lelièvre (2009). The algorithm has received little attention due to the difficulty in solving the non-linear inverse problem. With knowledge gained in Chapter 3, I developed an iterative rescaling strategy based on the maximum partial derivatives of the sensitivity function. The decoupling of the magnetization strength and orientation allowed me to apply sparsity assumptions for the recovery of well-defined anomalies with coherent magnetization direction. It is an improvement over methods previously published as it can deal with complex geological settings comprising multiple anomalies with arbitrary shape.

Chapter 5 is a generalization of the methodology introduced in Chapter 3 for the recovery of oriented edges. The measure of model gradients along the Cartesian axes was a choice of convenience but it poses a major limitation for the modeling of folded and dipping geological contacts. I introduced a 7-point gradient operator measuring the model gradients with diagonal cells. I have also shown how scarce structural measurements can be interpolated and used to guide the inversion.

Chapter 6 takes the model space inversion further by attempting to learn from the suite of solutions. I investigated ways to extract optimal inversion parameters based on recurrence of dominant features. I utilized basic machine learning tools such as PCA, edge detection and image moment algorithms to identify patterns and build constraints. I elaborated a strategy to automate dip and strike estimation of isolated geological units.

All the technology brought forward in this thesis were put to test in Chapter 7

on the Kevitsa Ni-Cu-PGE deposit. I inverted ground gravity and airborne magnetic data using rotated sparse norms. The resulting density and magnetization model highlighted important features of the deposit. First from the density model, I confirmed the seismic interpretation of an extended ultra-mafic unit under cover. The magnetization model also corroborated previous studies in modeling the remanent component of the central dunite. The inversion also indicated a possible connection with other units found at the base of the intrusion. Potentially the most significant outcome of this case study is the orientation of magnetization recovered in several folded units. While other studies have attempted to use magnetization inversion to gain insight about the geological history of a deposit, it is likely the first time that this was done at such a large scale and over complex geology.

8.1 Limitations and future work

The work presented in this thesis represents a significant gain in flexibility for inverting gravity and magnetic data, but this comes at an added computational cost needed to a single inversion and also for generating a suite of models. This is exacerbated by the cooling strategy of the threshold parameter introduced in Chapter 3 such that sparsity assumptions are slowly phased in. It is important to note that each model is a valid candidate with respect to the geophysical data. These models could potentially be used earlier in the learning process to avoid repetitive iteration steps.

Attempting to use magnetic vector inversion for large scale paleomagnetic studies is an appealing approach that warrants further investigation. Sparsity assumptions have helped in recovering coherent magnetization direction inside isolated anomalies. This process is less stable when performed over multiple magnetic anomalies with overlapping signal. The use of magnetic gradient data may help in reducing the non-uniqueness and in better defining the boundaries of anomalies. Another aspect of magnetic vector inversion that I have not addressed in this thesis is to isolate the remanent component from the total magnetization. This could potentially be accomplished by jointly inverting electromagnetic and magnetic data for the modeling of susceptibility, remanence and conductivity.

The learning algorithms presented in this thesis have served their purpose in

identifying trends in the model space and extracting local parameters. A critical component of this process is the pattern recognition phase. This step still relies heavily on the user to determine tuning parameters that are currently found by trial and error. Some level of quality control is also needed when defining rotation parameters such that strike and dip angles are consistent with known dip directions. More advanced machine learning algorithm could potentially improve these procedures. The learning process could be performed on multiple physical properties and serve as a link for joint and cooperative inversions.

Research dedicated to potential field data is likely to continue and grow as the quality and quantity of surveys continues to increase. While I have made inroads in reducing the computation cost of potential fields inversions, large (continental) scale inversion remains difficult. One of the main tuning parameters that I have mostly ignored in this project is the reference model. At a large scale, choosing a single reference value is undoubtedly a gross generalization that can adversely affect the solution. Future work should investigate the use of model decomposition techniques such that the inversion is performed on both the reference and anomalous properties. Large scale model decomposition inversion could be done at various resolutions with the use of spherical Octree discretization.

The methodology presented in this thesis will be extended to other geophysical data. Work is currently underway to test the model space inversion on electrical resistivity problems and airborne electromagnetics.

Bibliography

- Abedi, M., D. Fournier, S. G. Devriese, and D. W. Oldenburg, 2018a, Integrated inversion of airborne geophysics over a structural geological unit: A case study for delineation of a porphyry copper zone in Iran: *Journal of Applied Geophysics*, **152**, 188 – 202. → pages vi
- , 2018b, Potential field signatures along the Zagros collision zone in Iran: *Tectonophysics*, **722**, 25 – 42. → pages vi
- Ajo-Franklin, J. B., B. J. Minsley, and T. M. Daley, 2007, Applying compact constraints to differential traveltome tomography: *Geophysics*, **72**, 67–75. → pages 8, 34, 46, 55
- Barber, C. B., D. P. Dobkin, and H. Huhdanpaa, 1996, The Quickhull algorithm for convex hulls: *ACM Transactions on Mathematical Software*, **22**, 469–483. → pages 18
- Barbosa, V. C., and J. B. Silva, 2006, Interactive 2D magnetic inversion: A tool for aiding forward modeling and testing geologic hypotheses: *Geophysics*, **71**, L43–L50. → pages 84
- Barbosa, V. C. F., and J. B. C. Silva, 1994, Generalized compact gravity inversion: *Geophysics*, **59**, 57–68. → pages 8, 34, 46, 55
- Blakely, R., 1996, Potential theory in gravity and magnetic applications: Cambridge University Press. → pages xii, 2
- Blaschek, R., A. Hordt, and A. Kemna, 2008, A new sensitivity-controlled focusing regularization scheme for the inversion of induced polarization data based on the minimum gradient support: *Geophysics*, **73**, 45–54. → pages 8, 46
- Bosch, M., and J. McGaughey, 2001, Joint inversion of gravity and magnetic data under lithologic constraints: *The Leading Edge*, **20**, 877–881. → pages 7, 29
- Briggs, I. C., 1974, Machine contouring using minimum curvature: *Geophysics*, **39**, 39–48. → pages 96, 113
- Chartrand, R., 2007, Exact reconstruction of sparse signals via nonconvex minimization: *IEEE Signal Processing Letters*, **76**, 167–188. → pages 8, 46
- Clark, D. A., 2014, Methods for determining remanent and total magnetisation of magnetic sources - a review: *Exploration Geophysics*, **45**, 271–304. → pages

- Cordell, L., and V. J. S. Grauch, 2012, *in* Mapping basement magnetization zones from aeromagnetic data in the San Juan Basin, New Mexico: 181–197. → pages 3
- Cox, L. H., G. A. Wilson, and M. S. Zhdanov, 2010, 3D inversion of entire airborne electromagnetic surveys using a moving footprint: ASEG Extended Abstracts, 1–10. → pages 22
- Dannemiller, N., and Y. Li, 2006, A new method for estimation of magnetization direction: *Geophysics*, **71**, L69–L73. → pages 10
- Dask, D. T., 2016, Dask: Library for dynamic task scheduling: <https://dask.org>. → pages 22
- Daubechies, I., R. DeVore, M. Fornasier, and C. S. Gunturk, 2010, Iteratively re-weighted least squares minimization for sparse recovery: *Communication on Pure and Applied Mathematics*, **63**, 1–35. → pages 8, 34, 42
- Davis, K., D. W. Oldenburg, and M. Hillier, 2012, Incorporating geological structure into the inversion of magnetic data: Presented at the ASEG 22th Conference and Exhibition. → pages 7, 84
- Domzalski, W., 1966, Importance of aeromagnetics in evaluation of structural control of mineralization: *Geophysical Prospecting*, **14**, 273–291. → pages 1
- Ekblom, H., 1973, Calculation of linear best l_p -approximation: *BIT*, **13**, 292–300. → pages 8, 34
- Engheta, N., W. D. Murphy, V. Rokhlin, and M. S. Vassiliou, 1992, The Fast Multipole Method (FFM) for the electromagnetic scattering problem: *IEEE Transactions on Antenna and Propagation*, **40**, 634641. → pages 21
- Enkin, R., 2003, Paleomagnetic analysis of selected Ekati kimberlite. Unpublished report for BHP Billiton Diamonds. Geological Survey of Canada. → pages 9
- Enkin, R. J., 2014, The rock physical property database of British Columbia, and the distinct petrophysical signature of the Chilcotin basalts.: *Canadian Journal of Earth Sciences*, **51**, 327–338. (NRC Research Press). → pages 9, 69
- Farquharson, C. G., and D. W. Oldenburg, 1998, Non-linear inversion using general measures of data misfit and model structure: *Geophysical Journal International*, **134**, 213–227. → pages 8, 34
- Fedi, M., G. Florio, and A. Rapolla, 1994, A method to estimate the total magnetization direction from a distortion analysis of magnetic anomalies1: *Geophysical Prospecting*, **42**, 261–274. → pages 10
- Foss, C., 2006, *in* Parametric inversion as an advanced technique for magnetic source depth estimation: 943–947. → pages 7
- Foss, C., and B. McKenzie, 2011, Inversion of anomalies due to remanent magnetisation: an example from the Black Hill Norite of South Australia:

- Australian Journal of Earth Sciences, **58**, 391–405. → pages 10
- Fournier, D., 2015, A cooperative magnetic inversion method with l_p -norm regularization: Msc. thesis, The University of British Columbia. → pages 10, 35, 48, 71
- Fournier, D., K. Davis, and D. W. Oldenburg, 2016, Robust and flexible mixed-norm inversion: SEG Technical Program Expanded Abstracts, 1542–1547. → pages vi, 11
- Fournier, D., and D. Oldenburg, 2019a, Sparse magnetic vector inversion in spherical coordinates: Application to the Kevitsa Ni-Cu-PGE magnetic anomaly, Finland: Geophysics. (In revision). → pages 142
- Fournier, D., and D. W. Oldenburg, 2019b, Inversion using spatially variable mixed l_p -norms: Geophysical Journal International, **218**, 268–282. → pages vi
- Fullagar, P. K., and G. A. Pears, 2013, 3D magnetic modelling and inversion incorporating self-demagnetisation and interactions: ASEG Extended Abstracts, **2013**, 1–4. → pages 10
- Fullagar, P. K., G. A. Pears, and B. McMonnies, 2008, Constrained inversion of geologic surfaces - pushing the boundaries: The Leading Edge, **27**, 98–105. → pages 7, 29
- Gorodnitsky, I. F., and B. D. Rao, 1997, Sparse singal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm: IEEE Transactions on Signal Processing, **45**, 600–616. → pages 8, 34
- Grant, F., 1984, Aeromagnetic, geology and ore environments, I. Magnetite in igneous, sedimentary and metamorphic rocks: An overview: Geoexploration, **23**, 303–333. → pages 1
- Gregory, J., N. Journet, G. White, and M. Lappalainen, 2011, Kevitsa copper nickel project Finland: 43-101 report, First Quantum Mineral. → pages 117
- Haber, C., and C. Schwartzbach, 2014, Parallel inversion of large-scale airborne time-domain electromagnetic data with multiple OcTree meshes: Inverse Problems, **30**, 1–28. → pages 22
- Haber, E., D. W. Oldenburg, T. Farncombe, and A. Cellar, 1997, Direct estimation of dynamic parameters in SPECT tomography: IEEE Transaction on Nuclear Science, **44**, 2425–2429. → pages 30
- Haber, E., and L. Tenorio, 2003, Learning regularization functionals: a supervised training approach: Inverse Problems, **19**. → pages 7
- Helbig, K., 1963, Some integrals of magnetic anomalies and their relation to the parameters of the disturbing body: Zeitschrift fr Geophysik, **29**, 83–96. → pages 10
- Henkel, H., 1991, Petrophysical properties (density and magnetization) of rocks from the northern part of the baltic shield: Tectonophysics, **192**, 1–19. → pages 9, 69

- Hestenes, M. R., and E. Stiefel, 1952, Method of Conjugate Gradient for solving linear systems: *Journal of Research of the National Bureau of Standards*, **49**, 409–436. → pages 29, 41
- Hotelling, H., 1933, Analysis of a complex of statistical variables into principal components: *Journal of Educational Psychology*, **24**, 417–441. → pages 104
- Hu, M.-K., 1962, Visual pattern recognition by moment invariants: *IRE Transactions on Information Theory*, **8**, 179–187. → pages 111
- Huber, P. J., 1964, Robust estimation of a location parameter: *The Annals of Mathematical Statistics*, **35**, 73–101. → pages 34
- Jones, E., T. Oliphant, and P. Peterson, 2001, SciPy: Open source scientific tools for Python: <http://www.scipy.org/>. → pages 42
- Kissel, C., and C. Laj, 1989, Paleomagnetic rotation and continental deformation: Kluwer Academic Publishers, volume **254** of Series C: Mathematical and Physical Sciences. → pages 9
- Koivisto, E., A. Malehmir, N. Hellqvist, T. Voipio, and C. Wijns, 2015, Building a 3D model of lithological contacts and near-mine structures in the Kevitsa mining and exploration site, Northern Finland: constraints from 2D and 3D reflection seismic data: *Geophysical Prospecting*, **63**, 754–773. → pages xii, xxii, xxiii, 6, 7, 117, 118, 119, 120, 127, 128, 136
- Krahenbuhl, R., and Y. Li, 2015, Robust parametric inversion using adaptive quenched simulated annealing: SEG Technical Program Expanded Abstract, 1526–1530. → pages 7
- Last, B. H., and K. Kubik, 1983, Compact gravity inversion: *Geophysics*, **48**, no. 8, 713–721. → pages 8, 34, 46, 55
- Lawson, C. L., 1961, Contribution to the theory of linear least maximum approximation: PhD Thesis, University of California, Los Angeles. → pages xiv, 8, 34, 35
- Le, C. V. A., B. D. Harris, and A. M. Pethick, 2016, Magnetotelluric inversion, carbonaceous phyllites and an ore zone: Kevitsa, finland: 25th Geophysical Conference and Exhibition, 525–529. → pages 120
- Lee, T.-C., R. Kashyap, and C.-N. Chu, 1994, Building skeleton models via 3-d medial surface/axis thinning algorithms: *Computer Vision, Graphics, and Image Processing*, **56**, 462–478. → pages 111
- Lelièvre, P. G., 2009, Integrating geological and geophysical data through advanced constrained inversion: PhD Thesis, The University of British Columbia. → pages xix, 29, 69, 84, 88, 89, 90, 93, 142
- Lelievre, P. G., and D. W. Oldenburg, 2009, A 3D total magnetization inversion applicable when significant, complicated remanence is present: *Geophysics*, **74**, L21–L30. → pages 10, 12, 70, 71, 74
- Lelièvre, P. G., D. W. Oldenburg, and N. Williams, 2009, Integrating geologic and

- geophysical data through advanced constrained inversions: Extended Abstracts, 20th Geophysical conference and exhibition, ASEG, 1–6. → pages 29
- Li, Y., 1993, A globally convergent method for l_p problems: SIAM J. Sci. Stat. Comput., **3**, 609–629. → pages 8, 34
- Li, Y., A. Melo, C. Martinez, and J. Sun, 2019, Geology differentiation: A new frontier in quantitative geophysical interpretation in mineral exploration: The Leading Edge, **38**, 60–66. → pages 6
- Li, Y., and D. W. Oldenburg, 1996, 3-D inversion of magnetic data: Geophysics, **61**, 394–408. → pages 5, 30, 31, 66
- , 1998, 3-D inversion of gravity data: Geophysics, **63**, 109–119. → pages 5
- , 2000, Incorporating geological dip information into geophysical inversions: Geophysics, **65**, 148–157. → pages xix, 7, 12, 29, 84, 85, 87, 90
- , 2003, Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method: Geophysical Journal International, **152**, 251–265. → pages 21
- Lin, W., and M. Zhdanov, 2017, Joint inversion of seismic and gravity gradiometry data using gramian constraints: SEG Technical Program Expanded Abstracts, 1734–1738. → pages 7
- Liu, S., X. Hu, Y. Xi, T. Liu, and S. Xu, 2015, 2D sequential inversion of total magnitude and total magnetic anomaly data affected by remanent magnetization: Geophysics, **80**, K1–K12. → pages 10, 71
- Liu, S., X. Hu, H. Zhang, M. Geng, and B. Zuo, 2017, 3D magnetization vector inversion of magnetic data: improving and comparing methods: Pure and Applied Geophysics, **174**, 4421–4444. → pages 74
- Lockhart, G., H. Grutter, and J. Carlson, 2004, Temporal, geomagnetic and related attributes of kimberlite magmatism at Ekati, Northwest Territories, Canada: Lithos, **77**, 665–682. → pages 9
- Macleod, I. N., K. Jones, and T. F. Dai, 1993, 3-D Analytic Signal in the Interpretation of Total Magnetic Field Data at Low Magnetic Latitudes: Exploration Geophysics, **24**. → pages 3
- Marjoribanks, R., 2010, Geological methods in mineral exploration and mining, second ed.: Springer. → pages 1
- McMillan, M., C. Schwarzbach, E. Haber, and D. Oldenburg, 2016, Multiple body parametric inversion of frequency- and time-domain airborne electromagnetics: SEG Technical Program Expanded Abstracts 2016, 846–851. → pages 7
- Miller, C., S. Kang, D. Fournier, and G. Hill, 2018, Distribution of and condensate in a hydrothermal system: Insights from selfpotential inversion at Mount Tongariro, New Zealand: Geophysical Research Letters, **45**, 8190–8198. → pages vi
- Miller, C. A., G. Williams-Jones, D. Fournier, and J. Witter, 2017, 3D gravity

- inversion and thermodynamic modelling reveal properties of shallow silicic magma reservoir beneath Laguna del Maule, Chile: Earth and Planetary Science Letters, **459**, 14 – 27. → pages vi
- Miller, H. G., and V. Singh, 1994, Potential field tilta new concept for location of potential field sources: Journal of Applied Geophysics, **32**, 213–217. → pages 3
- Montonen, M., 2012, Induced and remanent magnetization in two boreholes of the Kevitsa intrusion: Master's thesis, University of Helsinki. → pages xi, 120, 122, 124, 127, 128
- Mushayandebvu, M. F., P. van Driel, A. B. Reid, and J. D. Fairhead, 2001, Magnetic source parameters of two-dimensional structures using extended Euler deconvolution: Geophysics, **66**, 814–823. → pages 3
- Mutanen, T., 1997, Geology and ore petrology of the Akanvaara and Koitelainen mafic layered intrusions and the Keivitsa-Satovaara layered complex, northern Finland: Geological Survey of Finland, **395**. → pages 117, 119
- Nabighian, M. N., M. E. Ander, V. J. S. Grauch, R. O. Hansen, T. R. LaFehr, Y. Li, W. C. Pearson, J. W. Peirce, J. D. Phillips, and M. E. Ruder, 2005, Historical development of the gravity method in exploration: Geophysics, **70**, 63ND–89ND. → pages 3
- Nagy, D., 1966, The gravitational attraction of a right rectangular prism: Geophysics, **31**, 361–371. → pages 13
- NCEI, 2017, Emag2v3: Earth magnetic anomaly grid (2-arc-minute resolution): Online. (geomag.models@noaa.gov). → pages xii, 4
- Nocedal, J., and S. J. Wright, 1999, Numerical optimatization: Springer Science. → pages 29, 41
- Norris, D. K., and R. F. Black, 1961, Application of paleomagnetism to thrust mechanics: Nature, **192**, 933–935. → pages 9
- Osborne, M. R., 1985, Finit algorithms in optimization and data analysis: John Wiley & Sons Lts. → pages 42
- Pearson, K., 1901, On lines and plane of closest fit to systems of points in space: Philosophical Magazine, Series 6, **2**, 559–572. → pages 104
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, 2011, Scikit-learn: Machine Learning in Python: Journal of Machine Learning Research, **12**, 2825–2830. → pages 105, 106
- Phillips, J. D., 2003, Can we estimate total magnetization directions from aeromagnetic data using Helbig's formulas: Abstracts Week B, IUGG 2003, B261. → pages 10
- , 2010, Estimating structural dip from gravity and magnetic profile data:

- SEG Technical Program Expanded Abstracts 2010, 1202–1206. → pages 4
- Phillips, N. D., 1996, Geophysical inversion in an integrated exploration program: examples from the San Nicolas deposit: Master's thesis, The University of British Columbia. → pages 7, 29
- Pilkington, M., 1997, 3-D magnetic imaging using conjugate gradients: *Geophysics*, **62**, 1132–1142. → pages 21, 66
- Portniaguine, O., and M. S. Zhdanov, 2002, 3-D magnetic inversion with data compression and image focusing: *Geophysics*, **67**, 1532–1541. → pages 8, 46
- Portniaguine, O. N., 1999, Image focussing and data compression in the solution of geophysical inverse problems: PhD Thesis, Department of Geology and Geophysics, The University of Utah. → pages 34
- Pratt, D. A., B. K. McKenzie, and T. S. White, 2014, Remote remanence estimation (RRE): *Exploration Geophysics*, **45**, 314–323. → pages 10
- Pueyo, E. L., F. Cifelli, A. J. Sussman, and B. Oliva-Urcia, 2016, Introduction: Palaeomagnetism in fold and Thrust belts: New perspective: Geological Society, London, **425**, 1–6. (Special Publications). → pages 10
- Queitsch, M., M. Schiffler, R. Stolz, C. Rolf, M. Meyer, and N. Kukowski, 2019, Investigation of three-dimensional magnetization of a dolerite intrusion using airborne full tensor magnetic gradiometry (FTMG) data: *Geophysical Journal International*, **217**, 1643–1655. → pages 10
- Ramon, M. J., E. L. Pueyo, J. L. Briz, A. Pocovi, and J. C. Ciria, 2012, Flexural unfolding of horizon using paleomagnetic vectors: *Journal of Structural Geology*, **35**, 28–39. → pages 9
- Reid, A. B., J. M. Allsop, H. Granser, A. J. Millett, and I. W. Somerton, 1990, Magnetic interpretation in three dimensions using Euler deconvolution: *Geophysics*, **55**, 80–91. → pages 3
- Reid, J. E., and A. Pfaffling, 2006, Airborne electromagnetic footprint in 1d earths: *Geophysics*, **71**, 63–72. → pages 21
- Rocklin, M., 2015, Dask: Parallel computation with blocked algorithms and task scheduling: Proceedings of the 14th Python in science conference, 126–132. → pages 22
- Roest, W. R., J. Verhoef, and M. Pilkington, 1992, Magnetic interpretation using the 3-D analytic signal: *Geophysics*, **57**, 116–125. → pages 3
- Salem, A., S. Williams, J. D. Fairhead, D. Ravat, and R. Smith, 2007, Tilt-depth method : A simple depth estimation method using first-order magnetic derivatives: *The Leading Edge*, **26**, 1502. → pages 4
- Sanchez, M. G., M. M. Allan, C. J. R. Hart, and J. K. Mortensen, 2014, Extracting ore-deposit-controlling structures from aeromagnetic, gravimetric, topographic, and regional geologic data in western Yukon and eastern Alaska: *Interpretation*, **2**, 75–102. → pages 3

- Schermerhorn, W. D., B. Ritzinger, M. Anderson, J. Peacock, J. Witter, J. Glen, A. Steely, C. Forson, P. Stelling, and D. Fournier, 2017, Geophysical investigation of the Mount Baker geothermal play: Presented at the GSA Annual Meeting in Seattle. → pages vi
- Sharma, P. V., 1966, Rapid computation of magnetic anomalies and demagnetization effects caused by bodies of arbitrary shape: *Pure Appl. Geophys.*, **64**, 89–109. → pages 15
- Shearer, S., 2005, Three-dimensional inversion of magnetic data in the presence of remanent magnetization: Master's thesis, Colorado School of Mines, Golden, Colorado. → pages 10
- Stocco, S., A. Godio, and L. Sambuelli, 2009, Modelling and compact inversion of magnetic data: A Matlab code: *Computer and Geosciences*, **35**, 2111–2118. → pages 8, 46, 55
- Sun, J., and Y. Li, 2013, A general framework for joint inversion with petrophysical information as constraints: SEG Houston Annual Meeting, 3093–3097. → pages 7
- , 2014, Adaptive l_p inversion for simultaneous recovery of both blocky and smooth feature in geophysical model: *Geophysical Journal International*, **197**, 882–899. → pages 8, 34, 35, 47, 56, 101, 108
- Thompson, D. T., 1982, EULDPH: A new technique for making computer-assisted depth estimates from magnetic data: *Geophysics*, **47**, 31–37. → pages 3
- Thurston, J. B., and R. S. Smith, 1997, Automatic conversion of magnetic data to depth, dip, and susceptibility contrast using the SPI (TM) method: *Geophysics*, **62**, 807–813. → pages 4
- Verduzco, B., D. J. Fairhead, C. M. Green, and C. MacKenzie, 2004, New insights into magnetic derivatives for structural mapping: *The Leading Edge*, **23**, 116. → pages 3
- Villalain, J. J., A. M. Casas-Sainz, and R. Soto, 2015, Reconstruction of inverted sedimentary basins from syntectonic remagnetizations. a methodological proposal: *Palaeomagnetism in Fold and Thrust Belts: New Perspectives*. → pages 9
- Vine, F. J., and D. H. Matthews, 1963, Magnetic anomalies over the oceanic ridges: *Nature*, **199**, 947–949. → pages 9
- Vogel, C. R., 2002, Computational Methods for Inverse Problems (Frontiers in Applied Mathematics): Society for Industrial Mathematics. → pages 67
- Wijns, C., and P. Kowalczyk, 2007, Interactive geophysical inversion using qualitative geological constraints: *Exploration Geophysics*, **38**, 208–212. → pages 7
- Williams, N., 2008, Geologically constrained inversion: PhD Thesis, University

- of British Columbia, Vancouver, British Columbia. → pages 7, 29
- Yang, D., and D. W. Oldenburg, 2014, 3D inversion of airborne electromagnetic data parallelized and accelerated by local mesh and adaptive soundings: *Geophysical Journal International*, **196**, 1492–1507. → pages 22
- Zhdanov, M., and E. Tolstaya, 2004, Minimum support nonlinear parametrization in the solution of a 3D magnetotelluric inverse problem: *Inverse Problems*, **20**, 937–952. → pages 56
- Zhu, Y., M. S. Zhdanov, and M. Cuma, 2015, Inversion of TMI data for the magnetization vector using Gramian constraints: Presented at the SEG Annual Meeting. → pages 10, 71

Appendix

This section describes the partial volume calculation defined by the intersection of two upright prisms. Let us define the extent of a prism \mathbf{m}_i by the lower southwest and upper northeast nodal coordinates:

$$\mathbf{m}_i = [\mathbf{n}_L \ \mathbf{n}_U] = \begin{bmatrix} x_L & x_U \\ y_L & y_U \\ z_L & z_U \end{bmatrix} \quad (\text{A.1})$$

To calculate the volume of intersection of two prisms \mathbf{m}_i and \mathbf{m}_j I first calculate the length L_x intercepted along the x -axis as

$$\begin{aligned} L_x &= \max \left[x_{max} - x_{min}, 0 \right] \\ x_{min} &= \min \left[n_{x_{U_i}}, n_{x_{U_j}} \right] \\ x_{max} &= \max \left[n_{x_{L_i}}, n_{x_{L_j}} \right] \end{aligned} \quad (\text{A.2})$$

The same calculation is done along the y and z axes (L_y, L_z). The total volume is then calculated as

$$V = L_x * L_y * L_z \quad (\text{A.3})$$

Intersecting length calculations in (A.2) can easily augmented for M pairs of cells.