

INTEGRATING GEOLOGIC AND GEOPHYSICAL DATA THROUGH ADVANCED CONSTRAINED INVERSIONS

by

Peter George Lelièvre

M.Sc., The University of British Columbia, 2003
B.Sc.(hon.), Acadia University, 1999

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Geophysics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2009

© Peter George Lelièvre 2009

Abstract

To be reliable, Earth models used for mineral exploration should be consistent with all available geologic and geophysical information. There are two areas of research that are important to help achieve the goal of more reliable Earth models: development of geophysical inversion methods that 1) increase the kinds of geologic information that can be incorporated, and 2) can combine several complimentary types of geophysical data collected over the same Earth region.

Discussions with geologists connected to exploration studies have identified several types of geologic information which could help improve results if incorporated into geophysical inversions. A specific goal of my research was to develop methods for incorporating all of that information into inversion algorithms within a deterministic framework (i.e. minimization of an objective function). I developed several methods to reach this goal.

Inversion of magnetic data is complicated by the presence of remanent magnetization. Typical magnetic inversion routines assume no remanence exists and erroneous results can be obtained if this assumption is made incorrectly. To solve this problem I developed two total magnetization vector inversion routines that allow for the incorporation of geologic information that is commonly available regarding remanence.

Another important form of geologic information is orientation information. I developed methods for incorporating orientation information into geophysical inversions. This information can be incorporated by performing a rotation of an orthogonal system of smoothness operators and through the addition of linear constraints to the inverse problem.

Lastly, I developed an iterative procedure for cooperatively inverting multiple types of geophysical data (from surveys responsive to different physical properties). The procedure creates models that are structurally similar and geologically realistic, in that they involve sharp interfaces (physical property jumps) between rock units (regions of nearly constant physical properties).

My research provides functional methods for applying geophysics to exploration problems. None of the methods I developed involve significant complications to the inverse problem and they can be applied to current exploration problems without requiring additional computing resources.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	vii
List of Figures	viii
Acknowledgements	xvi
Dedication	xviii
Co-Authorship Statement	xix
1 Introduction	1
1.1 Research motivation	1
1.2 Existing methods for incorporating geologic information	2
1.3 Specific objectives	4
1.4 Thesis arrangement	5
1.5 Mathematical nomenclature	6
1.6 References	7
2 Integrating geologic and geophysical data through advanced constrained inversions	9
2.1 Introduction	9
2.2 Types of geologic information available	10
2.3 An illustrative synthetic scenario	11
2.4 Geologically unconstrained inversion	12
2.5 Incorporating physical property value information	13
2.6 Incorporating physical property trend information	14
2.7 Incorporating orientation and aspect ratio information	15
2.8 Recovering rock type models	16
2.8.1 Recovering sharp boundaries between rock units	16
2.8.2 Cooperative inversion	17
2.9 Conclusion	18
2.10 References	24

3 A 3D total magnetization inversion applicable when significant, complicated remanence is present	25
3.1 Introduction	25
3.1.1 The problem of remanence in magnetic interpretation	25
3.1.2 Previous approaches to dealing with remanence	29
3.2 Inversion for magnetization	30
3.2.1 Discretization of the TMV formulations	31
3.2.2 An underdetermined optimization approach	33
3.2.3 Inversion functionality and important practical considerations	37
3.2.4 A simple synthetic test	38
3.3 Reducing the nonuniqueness in TMV inversions	39
3.3.1 Total magnetization known exactly	39
3.3.2 Measurements of χ and Q	40
3.3.3 Known or assumed remanence direction	41
3.3.4 Comparison of constraints in the Cartesian and spherical formulations .	42
3.4 A complicated synthetic scenario	42
3.5 A real data example with multiple bodies and complicated remanence	45
3.6 Application to regions containing high susceptibility	50
3.7 Conclusion	51
3.8 References	55
4 A comprehensive study of including structural orientation information in geophysical inversions	57
4.1 Introduction	57
4.2 Incorporating orientation information as soft constraints	59
4.2.1 Our deterministic inversion framework	59
4.2.2 Specifying preferred elongation directions and aspect ratios	60
4.2.3 A 2D dipping model objective function	61
4.2.4 Extension to 3D: specification of strike, dip and tilt	63
4.2.5 The choice of difference operators	65
4.2.6 Specifying orientations at point locations	74
4.2.7 A practical issue: depth and distance weighting	74
4.3 Incorporating orientation information as hard constraints	76
4.3.1 Bounding the spatial gradient direction	76
4.3.2 Specifying directions of increase or decrease	77
4.3.3 Practical application of linear constraints	78
4.3.4 Linear inequality constrained inverse problems	78
4.3.5 Applying hard orientation constraints to a 2D example	81
4.4 Incorporating different forms of orientation information	82
4.4.1 Incorporating surface orientation information	84
4.4.2 Incorporating interpreted volumetric orientation information	85
4.4.3 Incorporating orientation information from a rock model	85
4.5 Application to the San Nicolás deposit	89
4.6 Application to the Hislop deposit	99

4.7 Conclusion	104
4.8 References	105
5 An iterative cooperative inversion strategy for obtaining geologically realistic models	107
5.1 Introduction	107
5.2 An iterative approach for recovering geologically realistic models	108
5.2.1 Our deterministic inversion framework	108
5.2.2 An overview of the iterative procedure	110
5.2.3 Measuring model structure	113
5.2.4 The iterative weighting procedure	114
5.2.5 Stopping criteria	115
5.2.6 Applying topological rules	116
5.3 A more complicated example	119
5.4 Extension to allow cooperative use of multiple data types	121
5.4.1 Joint and cooperative inversion	121
5.4.2 An overview of our cooperative strategy	121
5.4.3 Application of our cooperative strategy to a 2D synthetic example . .	123
5.4.4 Comparison to cross-gradient joint inversion	127
5.5 Application to the San Nicolás deposit	129
5.6 Conclusion	137
5.7 References	139
6 Conclusion	141
6.1 Summary and primary significance of results	141
6.2 A note on the San Nicolás examples	143
6.3 Advantages of increased regularization options	143
6.3.1 Feature appraisal	143
6.3.2 Learned regularization	144
6.4 Critical assessment of methods and future research	145
6.4.1 Chapter 2	145
6.4.2 Chapter 3	145
6.4.3 Chapter 4	146
6.4.4 Chapter 5	147
6.4.5 Deterministic and stochastic methods	147
6.5 The last word	148
6.6 References	149

Appendices

A Inversion primer	151
A.1 Introduction	151
A.2 The forward problem	152

Table of Contents

A.3	The inverse problem	152
A.4	Fitting the data	153
A.5	Specifying desired model characteristics	153
A.6	Optimization	154
B	Derivation of the Hessian for the spherical magnetization inversion formulation	155

List of Tables

2.1	The anomalous physical properties of each rock type for the synthetic example.	11
4.1	Difference options for the four sets of finite difference operators for the 2D dipping model objective function.	72
4.2	Strike and dip values (in degrees) assigned to each region indicated in Figure 4.26.	99

List of Figures

2.1	Image (a) shows the gravity data used in the inversions. Image (b) shows the inversion mesh in black, the locations of the gravity data are indicated with blue dots (above the surface), and the cross-well seismic tomography ray paths are indicated with red lines. Image (c) shows the seismic tomography travel-time data used in the inversions.	19
2.2	True and inverted models for the gravity and cross-well seismic tomography synthetic example. Density models are on left, slowness models on right. To remove clutter, we do not show axis labels on plots (a) through (h) but we provide the inversion mesh below (g), which is a duplication of that shown in Figure 2.1(b). The colour scale for all models is $[-0.7, 0.7]$, using the colour bar below (h). Superimposed black lines in (a) through (h) indicate the rock unit boundaries in the true model. From top to bottom the models are: (a,b) the true models; (c,d) geologically unconstrained results; (e,f) results with bounds applied along the surface; (g,h) results with additional linear inequality constraints applied to enforce expected depth trends. Models from subsequent inversions are shown in Figure 2.3.	20
2.3	Inverted models for the gravity and cross-well seismic tomography synthetic example. Density models are on left, slowness models on right. To remove clutter, we do not show axis labels on plots (a) through (h) but we provide the inversion mesh below (g), which is a duplication of that shown in Figure 2.1(b). The colour scale for all models is $[-0.7, 0.7]$, using the colour bar below (h). Superimposed black lines in (a) through (h) indicate the rock unit boundaries in the true model. The location of a drill-hole is indicated with a white line in (a) and (b). From top to bottom the models are: (a,b) results with additional drill-hole bounds, (c,d) results with additional orientation information incorporated; (e,f) results of our structural cooperative inversion strategy after incorporating only surface bounds and expected depth trends; (g,h) independent results using approximate ℓ^1 -norms and the same geologic information as in (e,f). Models from previous inversions are shown in Figure 2.2.	21
2.4	Histograms showing the distribution of physical property values in the corresponding models in Figure 2.2.	22
2.5	Histograms showing the distribution of physical property values in the corresponding models in Figure 2.3.	23

3.1	A vertical cross-section at northing = 0m through the true 3D synthetic model. The mesh cell spacing is identical in the easting and northing directions. A small magnetic body with $\chi = 0.1$ (plotted in grey here) resides in the centre of the mesh within a nonmagnetic background. The induced component of the magnetization is labelled J_{ind} , the remanent component J_{rem} and the total magnetization J_{tot}	26
3.2	Map views of the (a) induced and (b) remanent components of the TMI response, and (c) total combined TMI response 10m above the mesh in Figure 3.1.	27
3.3	Vertical cross-sections at northing = 0m through 3D models recovered from inversion of the data in Figure 3.2(c): (a) susceptibility inversion with positivity imposed and assuming no remanence exists; (b) TMVC inversion with $\gamma = 1.0$ and no bounding applied; (c) TMVS inversion with $\theta = 90^\circ$ in the initial halfspace model; (d) TMVS inversion with $\theta = -45^\circ$ in the initial halfspace model. No reference models were included in these inversions. The colour scale used in (a) is SI susceptibility (unitless), in all others it is magnetization amplitude in units of effective susceptibility (i.e. normalized by the Earth's field strength). Total magnetization vectors are indicated by black and white lines with white at the head. Vectors with zero amplitude are not displayed. The white squares indicate the position of the block in the true model.	28
3.4	Views of the synthetic two body model: (a) shows a 3D perspective view, (b) shows a vertical cross-section at easting = 30m and (c) shows a horizontal cross-section at depth = 30m. Mesh cell spacing is identical in the easting and northing directions. The induced component of the magnetization is indicated in (b) and the remanent component in (c).	43
3.5	Map views of the (a) induced and (b) remanent components of the TMI response, and (c) total combined TMI response 10m above the mesh in Figure 3.4.	44
3.6	3D perspective views of models recovered from inversion of the data in Figure 3.5(c). Plot (a) shows a $\chi_{eff} = 0.10$ isosurface of the model from a magnetization amplitude inversion (i.e. the method of Shearer and Li (2004)). Plot (b) shows a $\chi_{eff} = 0.032$ isosurface of the model from a TMVC inversion with $\gamma = 1.0$, no reference models included and a depth weighting power of 2.0. Plot (c) shows a $\chi_{eff} = 0.11$ isosurface of the model from a TMVS inversion with $\gamma = 1.0 \times 10^{-3}$, no reference models included and bounds applied.	46
3.7	Plots (a) and (b) show horizontal cross-sections at depth = 30m through the models in Figures 3.6(b) and 3.6(c) respectively. Total magnetization vectors are indicated by black and white lines with white at the head. Vectors with zero amplitude are not displayed. The black rectangles indicate the positions of the blocks in the true model. Note that some padding cells have been removed in these plots.	47
3.8	A map of the real TMI survey data. The data ranges from 726nT to 1062nT but we have altered the colour scale here to better represent the larger scale data features.	48

3.9	3D perspective views from the southeast of the model recovered from a TMVC inversion of real survey data with $\gamma = 1.0$. Plot (a) shows $\chi_{eff} = -0.009$ and $\chi_{eff} = +0.005$ isosurfaces of the p component. Plot (b) shows a $\chi_{eff} = 0.006$ isosurface of the st component model ($= \sqrt{s^2 + t^2}$).	49
3.10	Vertical cross-sections at northing = 0m through 3D models for the high susceptibility example. The inversion mesh is identical in the northing and easting directions. This cross-section runs along the long axis of the true spheroid body, shown schematically in (a). The spheroid has a susceptibility of 10 (SI units). The TMVC inversion result with $\gamma = 1.0$ and a depth weighting power of 2.0 is shown in (b). The TMVS inversion result with $\gamma = 100$ and the Earth's field direction used in the initial model is shown in (c). The Earth's field is indicated in (a), labelled H_0 , and the resulting magnetization for the spheroid is labelled J . In (b) and (c), magnetization models are shown with the amplitude in units of effective susceptibility indicated on the colour bars. Total magnetization vector directions are indicated by black and white lines with white at the head. Vectors with zero amplitude are not displayed. The position of the spheroid is indicated in (b) and (c) with a black outline.	52
3.11	Total-field data maps calculated 4.0m above the surface for the spheroid model in Figure 3.10(a): (a) shows the observed data (forward modelled data with noise added); (b) shows the predicted data for the model in Figure 3.10(b); (c) shows the predicted data for the model in Figure 3.10(c). The locations of the data are indicated by black dots.	53
4.1	The mesh-orthogonal and rotated coordinate systems for a 2D problem.	61
4.2	The mesh-orthogonal and rotated coordinate systems for a 3D problem (this figure was reproduced from Li and Oldenburg (2000)).	65
4.3	A 1D mesh with four cells.	65
4.4	A 3-by-3 2D mesh. The shaded cells indicate those involved in the finite differences defined at the centre of cell 5 when forward differences are employed.	66
4.5	The true 2D density model is in (a). The recovered 2D density model with no preferred elongation direction specified ($w_{x'} = w_{z'} = 1.0$) is in (b).	68
4.6	The recovered 2D density models using a dipping-objective function with forward differences used in the model objective function. The specified dips are (a) $+45^\circ$ and (b) -45° . The weights used were $w_{x'} = 1.0$ and $w_{z'} = 0.001$ across the entire mesh.	69
4.7	A chess-board pattern is in (a). The recovered 2D density model using node-centred finite-difference operators in the model objective function with $\theta = 45^\circ$ and $w_{x'} = w_{z'}$ is in (b).	71
4.8	A 3-by-3 2D mesh. The central cell 5 is split into 4 quadrants. The shaded regions across cell faces indicate the regions across which x - and z -direction forward finite differences are calculated for cell 5 (those regions overlap in quadrant 5d).	72

4.9	The recovered 2D density models using a dipping-objective function with four sets of differences used in the model objective function. The specified dips are (a) $+45^\circ$ and (b) -45° . The weights used were $w_{x'} = 1.0$ and $w_{z'} = 0.001$ across the entire mesh.	73
4.10	A 3-by-3 2D mesh. The shaded cells indicate those involved in the finite differences defined at the centre of cell 5 when four sets of finite difference operators are employed.	74
4.11	The recovered 2D density model from a similar inversion to that for Figure 4.9(a) but with additional linear inequality constraints in equation (4.58) added into the inversion.	82
4.12	The true 2D susceptibility model is in (a). The recovered 2D susceptibility model with no preferred elongation direction specified ($w_x = w_z = 1.0$) is in (b). The location of the layers in the true model are indicated with a thick black line in (b).	83
4.13	The recovered 2D susceptibility models with the following information incorporated into the inversions: (a) surface bounds; (b) surface bounds and surface orientations.	84
4.14	Interpreted dips (angles in degrees) for the 2D synclinal model are in (a). The $w_{x'}/w_{z'}$ ratios used across the volume for the result in Figure 4.15(b) are in (b).	86
4.15	The recovered 2D susceptibility models with the following information incorporated into the inversions: (a) surface and drill-hole bounds; (b) surface and drill-hole bounds, and interpreted orientations at depth.	87
4.16	Calculated and interpolated dips (angles in degrees) for the 2D synclinal rock model are in (a). Cells where dip can not be calculated or interpolated (i.e. below the bottom-most susceptible layer) have been set to zero. The $w_{x'}/w_{z'}$ ratios used across the volume for the result in Figure 4.17(a) are in (b).	88
4.17	The recovered 2D susceptibility models with surface and drill-hole bounds, and interface location information incorporated into the inversions. The inversion for (a) had bound constraints only. Additional linear constraints were incorporated for (b).	89
4.18	A top view of the mesh used to invert the San Nicolás gravity data, a map of which is overlayed. The mesh has 53-by-33-by-35 cells (easting-northing-depth). The locations of the cross-sections shown in the figures that follow are indicated with green lines.	91
4.19	A map of the San Nicolás gravity data (units of mGal) is in (a); a regional component has been removed from the original data used in Phillips (2001). The predicted data for the model in Figure 4.22 is in (b). The predicted data for the model in Figure 4.23 is in (c). All three data maps are plotted on the same scale for comparison. The locations of the 422 data are indicated by black dots.	92
4.20	The anomalous density model (units of g/cc) created from the geologic model and physical property information for the San Nicolás deposit: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.	94

4.21 The w_x' smoothness weights used in the inversion of the San Nicolás data with orientation information incorporated: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section. Weights of 1.0 are shown in dark grey and weights of 0.01 are shown in light grey.	95
4.22 The recovered density model (units of g/cc) for the default, unconstrained inversion of the San Nicolás data: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.	96
4.23 The recovered density model (units of g/cc) for the inversion of the San Nicolás data with orientation information incorporated: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.	97
4.24 A map of the gravity data (units of mGal) forward modelled for the model in Figure 4.20. The locations of the 422 data are indicated by black dots.	98
4.25 This figure shows the same information as in Figure 4.23 but the colour scale has been altered to better emphasize the keel structure at depth.	98
4.26 A top view of the mesh used to invert the Hislop magnetic data, a map of which is overlaid. The mesh has 73-by-55-by-40 cells (easting-northing-depth). Three regions are numbered and their boundaries indicated with red lines. The outline of the volumes shown in Figure 4.28 is indicated with a green rectangle.	101
4.27 A map view of the Hislop magnetic data is shown in (a). The predicted data for the model in Figure 4.28(a) are in (b). The predicted data map for the model in Figure 4.28(b) is visibly indistinguishable from that in (b) so we do not show it. Both data maps are plotted on the same scale for comparison and the units are nT . The locations of the 1725 data are indicated by black dots.	102
4.28 Vertical cross-sections at easting = 552850m through the susceptibility models (SI units) recovered from inversion of the Hislop data. The default, unconstrained inversion result is shown in (a). The result with orientation information incorporated is shown in (b). A geologic map, adapted from an interpretation by Power et al. (2004), has been overlaid on top of the mesh: the legend on top right in (a) indicates the colours of the magnetic mafic volcanic rocks (VMF) and ultramafic volcanic rocks (VUO). Interpreted depth traces for two faults are indicated with white and black lines: the Porcupine-Destor Fault is white and the local fault (no official name) black.	103
5.1 A flow diagram illustrating our iterative procedure. The number of mesh cells in the discretized Earth is denoted M	111

- 5.2 Plots for the first synthetic example: (a) shows the true slowness model (units of s/km) with the discrete mesh shown in black; (b) shows the ray-paths through the volume; (c) shows the inversion result with $w_s = 0$ and ℓ^2 -norms on the smoothness measures; (d) shows the inversion result with an approximate total variation measure (Ekblom with $p = 1.1$ applied to $\vec{\nabla}m$); (e) and (f) show the magnitude of the model gradient for the models in (c) and (d) respectively. To remove clutter, we do not show axis labels in (a) though (f) but we provide a labeled version of the inversion mesh in (g) with the ray-paths overlayed in red. 112
- 5.3 Results of our iterative inversion procedure for the first synthetic example with no smoothing applied. The recovered slowness models (units of s/km) are in the left column and weights applied in the inversions are in the right column. The numbers on the left indicate iteration number. The model at iteration 0 is the initial model recovered from an inversion with uniform weighting across the volume. The colour bar used is the same as that in Figure 5.2 (red high, blue low) and the colour scale for the models is on $[0, 1]$. The low weighting values are 0.01, indicated in white. To remove clutter, we do not show axis labels but we refer the reader to the labeled version of the inversion mesh in Figure 5.2(g). 117
- 5.4 Progression curves for the iterative inversion procedure results in Figures 5.3 (black line with point markers) and 5.9 (grey line with cross markers). The progression measure is the ℓ^1 -norm of the difference between the model at each iteration and the true model. 118
- 5.5 An illustrative schematic of our procedure. The volume of interest contains two rock units, R1 and R2. We are confident that the shaded regions in (a) contain the indicated rocks. Between those we are unsure and the essence of our procedure is to carefully expand those regions towards each other to obtain an improved estimate for the true interface location. In our inversions, the shaded regions would contain high smoothness weights and the white regions low weights. Two possible weighting scenarios with different topological characteristics are shown in (b) and (c). 118
- 5.6 Results of our iterative inversion procedure for the first synthetic example with smoothing applied. The recovered slowness models (units of s/km) are in the left column and weights applied in the inversions are in the right column. The numbers on the left indicate iteration number. The model at iteration 0 is the initial model recovered from an inversion with uniform weighting across the volume. The colour bar used is the same as that in Figure 5.2 (red high, blue low) and the colour scale for the models is on $[0, 1]$. The low weighting values are 0.01, indicated in white. To remove clutter, we do not show axis labels but we refer the reader to the labeled version of the inversion mesh in Figure 5.2(g). 120

5.7	Plots for the second synthetic example: (a) shows the true density model (units of g/cc); (b) shows the inversion result with ℓ^2 -norms; (c) shows the inversion result with an approximate total variation measure (Ekblom with $p = 1.1$ applied to $\vec{\nabla}m$); (d) shows the result of applying our iterative procedure on the region to the left of the dotted white line in (a). Solid white lines in (b), (c) and (d) indicate the location of features in the true model.	122
5.8	Plots for the third synthetic example: (a) shows the true density model (units of g/cc) and mesh; (b) shows the true slowness model (units of s/km) and mesh; (c) and (d) show the gravity and seismic tomography inversion results with ℓ^2 -norms; (e) and (f) show the inversion results with an approximate total variation measure (Ekblom with $p = 1.1$ applied to $\vec{\nabla}m$); (g) and (h) show the results for our iterative cooperative procedure; (i) and (j) show the results for our iterative cooperative procedure when the scenario is altered slightly to remove the body at depth in the density model only; (k) and (l) show the cross-gradient joint inversion results (again using the Ekblom measure with $p = 1.1$ applied to $\vec{\nabla}m$). The colour bar used is the same as that in Figure 5.2 (red high, blue low). The colour scales are on $[0, 1]$ for the density models and $[-1, 1]$ for the slowness models. The seismic sources and receivers are indicated in (b) by red dots on the left and right of the mesh. To remove clutter, we only show axis labels in (a) and (b): units are metres.	125
5.9	Results of the iterative cooperative procedure for the third synthetic example. The recovered density models (units of g/cc) are in the left column, the slowness models (units of s/km) are in the centre column, and the weights applied in the inversions are in the right column. The numbers on the left indicate iteration number. The models at iteration 0 are the initial models recovered from inversions with uniform weighting across the volume. The colour bar used is the same as that in Figure 5.2 (red high, blue low). The colour scales are on $[0, 1]$ for the density models and $[-1, 1]$ for the slowness models. The low weighting values are 0.01, indicated in white. To remove clutter, we do not show axis labels but we refer the reader to the labeled version of the inversion mesh in Figure 5.8(a) and 5.8(b).	126
5.10	Progression curves for the iterative cooperative procedure results in Figure 5.9. The progression measure is the ℓ^1 -norm of the difference between the model at each iteration and the true model.	127
5.11	A top view of the mesh used to invert the San Nicolás gravity data, a map of which is overlaid. The mesh has 53-by-33-by-35 cells (easting-northing-depth). The locations of the cross-sections shown in the figures that follow are indicated with green lines.	130
5.12	A map of the San Nicolás gravity data (units of mGal) is in (a); a regional component has been removed from the original data used in Phillips (2001). The predicted data for the model in Figure 5.14 is in (b). The predicted data for the model in Figure 5.15 is in (c). All three data maps are plotted on the same scale for comparison. The locations of the 422 data are indicated by black dots.	131

5.13 The anomalous density model (units of g/cc) created from the geologic model and physical property information for the San Nicolás deposit: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.	133
5.14 The recovered density model (units of g/cc) for the default, unconstrained inversion (ℓ^2 -norms) of the San Nicolás data: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.	134
5.15 The recovered density model (units of g/cc) after applying our iterative procedure to the San Nicolás data: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.	135
5.16 The final smoothness weights after applying our iterative procedure to the San Nicolás data: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section. Weights of 1.0 are shown in dark grey and weights of 0.01 are shown in light grey. . . .	136
5.17 A map of the gravity data (units of mGal) forward modelled for the model in Figure 5.13. The locations of the 422 data are indicated by black dots. . . .	137

Acknowledgements

Quite simply put, this thesis would not exist without Roman Shekhtman. Many thanks and twice as many apologies for all you have helped me with over the years. You deserve the first mention, I don't think Doug would disagree.

I feel lucky to have had Doug Oldenburg as a supervisor. He kept me in the loop, ensuring I was working along ticketyboo, like gangbusters, on those canonical show-stopper problems at the forefront of the science. Always there to help me develop a staw-man for the way ahead, to ensure my efforts were germane to the problem at hand, to help me benchmark my copacetic achievements, and ensure I never threw the baby out with the bath water. Doug, without you this thesis would have been hooped, and despite this tongue-in-cheek tribute (remember what they say about the sincerest form of flattery), I'm sincerely grateful for your tireless support, interest, advice, and enthusiasm for all things geophysics.

My humble gratitude to the following people for discussions and contributions to the research: Colin Farquharson, Eldad Haber, Ken Hickey, Richard Lane, Yaoguo Li, Dianne Mitchinson, Nigel Phillips, Nick Williams, and my committee members Greg Dipple and Michael Friedlander. Thanks also to those I shared the office and hallway with over the years, but special mention to Laurens Beran, Rob Eso and Megan Sheffer for acting as research backboards when required.

Specific mention should go to Ken Hickey, Richard Lane and Nick Williams for their help developing the list of geologic information found in Chapters 1 and 2; to Yaoguo Li for providing the magnetization amplitude inversion code used in Chapter 3; to Ken Hickey for his help inverting and interpreting the real survey data in Chapter 3; to Dianne Mitchinson for her help inverting and interpreting the Hislop data in Chapter 4; and to Nigel Phillips for his help inverting the San Nicolás data in Chapters 4 and 5.

Ben Polzer at Vale INCO, Alan Herring and Bob Johnson at EDCON-PRJ, Michael Michaud at St. Andrew Goldfields Ltd., and Teck Corporation provided real geophysical survey data with which to test my methods. Richard Waltz and Rich Cohen at Ziena Optimization Inc. were

Acknowledgements

kind enough to provide a limited version of the KNITRO software package for my use.

This thesis was formatted using the \LaTeX class *ubctheses* developed by Michael McNeil Forbes. The manuscripts in this thesis were significantly improved by reviews from John Paine (Chapter 2), Olivier Boulanger and Mark Pilkington (Chapter 3), and Michel Chouteau and Antonio Guillen (Chapter 4).

Financial support was generously provided by Doug Oldenburg and the Department of Earth and Ocean Sciences. Emotional support was generously provided by family, friends, pets, and pets of friends (you all know who you are).

Dedication

To my wife Melissa. I have infinite respect for your integrity and dedication. I don't think a thesis dedication does you enough justice, but here you are, and here we go ...

Co-Authorship Statement

Chapters 2 through 5 of this thesis are versions of research articles submitted or in preparation to be submitted for publication in peer-reviewed journals. The following provides specific contribution statements for each co-author. As my supervisor, Doug Oldenburg oversaw the research and provided ideas and advice for all chapters. I was principally responsible for documenting the research (manuscript preparation), although all co-authors had editorial roles for the chapters to which they are attached.

Chapter 2: Integrating geologic and geophysical data through advanced constrained inversions

Authors: Peter Lelièvre, Doug Oldenburg and Nick Williams

In this chapter I combine my methods into a single demonstrative example. Nick Williams provided advice when designing the synthetic example, which is a simplified version of one he created based on a real exploration scenario. Other than minor edits from co-authors, I wrote the bulk of the manuscript.

Chapter 3: A 3D total magnetization inversion applicable when significant, complicated remanence is present

Authors: Peter Lelièvre and Doug Oldenburg

Doug Oldenburg had the original idea to develop a vector magnetization inversion algorithm and he provided an initial formulation for the problem. I was principally responsible for performing the subsequent research, including a complete formulation of the problem, coding of algorithms, developing solutions to practical application issues, and testing. I developed the inversion code for the Cartesian formulation with programming help from Roman Shekhtman (during my initial

foray into the Fortran programming language). I developed code for the spherical framework independently. I also independently designed, coded and tested a variant of the Gradient-Projection-Reduced-Newton method (also mentioned in Chapter 2) that is used to deal with bound constraints. Other than minor edits from co-authors, I wrote the bulk of the manuscript.

Chapter 4: A comprehensive study of including structural orientation information in geophysical inversions

Authors: Peter Lelièvre and Doug Oldenburg

As indicated in this chapter, the method involving a rotation of orthogonal smoothness operators was initially developed in a paper by Yaoguo Li and Doug Oldenburg (see the chapter for reference). I took that preliminary development to its conclusion by developing, implementing and testing new discretization strategies while completely re-writing pre-existing computer code used for 3D inversions. Code for the logarithmic barrier algorithm existed previously for 3D inverse problems but was applied to bound constraints; I extended the code to apply the logarithmic barrier method to the linear constraints I developed in this chapter. I wrote all 2D computer code independently. Other than minor edits from co-authors, I wrote the bulk of the manuscript.

Chapter 5: An iterative cooperative inversion strategy for obtaining geologically realistic models

Authors: Peter Lelièvre, Doug Oldenburg and Nigel Phillips

The iterative inversion procedure I developed in this chapter is based in part on an idea by Nigel Phillips but I was entirely responsible for its design, development and testing. I wrote all the required 2D and 3D computer code independently. Other than minor edits from co-authors, I wrote the bulk of the manuscript.

Chapter 1

Introduction

1.1 Research motivation

To be reliable, Earth models used for mineral exploration should be consistent with all available geologic and geophysical information. Due to data uncertainty and other aspects inherent to the underdetermined geophysical inverse problem, there are an infinite number of models that can fit the geophysical data to the desired degree (i.e. the problem is nonunique). Additional information is essential to obtain a unique and useful solution. Incorporating prior geologic knowledge, and combining several complimentary types of geophysical data collected over the same Earth region, can reduce ambiguity and enhance inversion results, leading to more reliable Earth models. There are two areas of research that are important to help achieve the goal of more reliable Earth models: development of geophysical inversion methods that 1) increase the kinds of geologic information that can be incorporated, and 2) can combine several complimentary types of geophysical data collected over the same Earth region.

The research of this thesis is aimed at reducing model uncertainty. Considering the nonuniqueness of the inverse problem and the resulting uncertainty in the results, more than one model is required to allow reliable conjecture regarding the subsurface. Once several models of the Earth have been recovered through inversion, one should assess the results (i.e. feature appraisal) before using them to guide further exploration (e.g. spotting holes for drilling programs). Feature appraisal strategies are discussed in more detail in Chapter 6, but before model uncertainty can be quantified via those strategies, an important first step is to develop methods to incorporate constraining geological information into geophysical inversions. The research in this thesis deals primarily with that first step.

Geophysical inversion technology has generally provided some basic functionality for incorporating geologic information into geophysical inversions. This functionality has shown great promise for practical use. Williams (2006) uses a real-world-based synthetic example and comprehen-

sively demonstrates how some types of geologic information can be incorporated to provide far superior results than default, geologically unconstrained inversions. Phillips (2001) and Farquharson et al. (2008) applied similar methods to their inversion investigations of the San Nicolás copper-zinc deposit and the Voisey's Bay nickel copper-cobalt deposit respectively and had similar conclusions. However, the models obtained through typical inversion methods can be inconsistent with prior geologic knowledge that can not currently be included.

1.2 Existing methods for incorporating geologic information

Many researchers have provided functionality for incorporating different types of geologic information into their particular inversion frameworks. Below I discuss several relevant references to provide a general overview, and I place my methods within the context of those. In Appendix A I provide a brief primer on geophysical inversion that presents enough background information to provide entry to the remaining material in this thesis. Readers unfamiliar with inversion are encouraged to read that material before continuing below.

Inversion algorithms fit into two categories: deterministic and stochastic inversions. In a deterministic inversion a function is designed that, once minimized, the resulting model (i.e. that which minimizes the function) will adequately recreate the geophysical survey data and contain desired compositional and structural characteristics. In a stochastic inversion the overall goal is the same, i.e. the model must fit the data and contain desired characteristics. However, instead of strictly minimizing an objective function and obtaining a single result, a probability density function is designed and the model space (i.e. all possible models) is investigated through some statistical sampling technique. This creates many model estimates rather than a single result and provides statistical information regarding the model space.

There are benefits and limitations associated with both deterministic and stochastic approaches. The objective function in a deterministic algorithm is typically designed to be well-behaved (e.g. differentiable, convex) such that optimization strategies can quickly find a minimum without much issue. In contrast, the probability density function and sampling rules found in a stochastic algorithm are not required to be differentiable. Hence, a stochastic approach has the potential to be more flexible in the form of regularization applied and, therefore, in the types of geologic information that can be incorporated. However, stochastic approaches rely on sampling methods that lead to much heavier computational costs than deterministic approaches.

Li and Oldenburg (1996, 1998) designed a well-behaved quadratic objective function containing “smallness” and “smoothness” regularization. Under default conditions, their inversions can be considered minimum-structure inversions as they seek models that are smooth in all spatial directions and do not contain unnecessarily large values. When geologic information is available, it can be incorporated through the addition of a reference model, representing the best guess at the subsurface. In Li and Oldenburg (2003), additional bound constraints are added to the inverse problem to allow specification of lower and upper bounds on particular model elements such that field measurements of physical properties can be incorporated.

Typical smooth-model regularization, such as that used by Li and Oldenburg (1996, 1998), generates smoothly varying models. Sharper discontinuities (i.e. blocky structures) have been created through several strategies that involve less well-behaved (e.g. nonlinear and nonquadratic) functions, creating challenges in the subsequent minimization. In an extension of the method of Last and Kubik (1983), Portniaguine and Zhdanov (1999) developed a method based on a minimum gradient support functional that helps lead to the construction of models with sharp features. A total variation measure for smoothness regularization, an approach developed by Rudin et al. (1992) and Vogel and Oman (1998), can also help to create sharper features. Farquharson and Oldenburg (1998) implement further general measures of model smoothness which can do the same.

Li and Oldenburg (2000) extended the approach of Li and Oldenburg (1996, 1998) to allow specification of elongations in arbitrary orientations within the model. Chasserau and Chouteau (2003) introduce regularization through a parameter covariance matrix, the elements of which are estimated statistically using variograms. The covariance matrix can be estimated from physical property data (e.g. measurements taken at surface or down drill-holes) or using some initial model(s) representing the best guess at the subsurface distribution. The variogram calculations involve three specified ellipsoid axes in any spatial directions and as such, their method allows construction of structures with different shapes and orientations. Estimation of the parameter covariance matrix requires significant memory and computation time. Last and Kubik (1983) developed a compact (minimum volume) gravity inversion and Guillen and Menichetti (1984) extended that to allow specification of a dip line along which the moment of inertia is minimized. Barbosa et al. (1994) extended the method further to allow specification of mass concentration information along several arbitrary axes, allowing specification of elongation direction and scale.

Barbosa and Silva (2006) apply the method of Barbosa et al. (1994) within an interactive environment in which the interpreter can adjust the arbitrary inertial axes as desired in order

to aid geologic hypothesis testing. As such, their approach is similar to interactive forward modelling, where the interpreter investigates the model space, but differs in that the algorithm automatically fits the data. Another novel approach is that of Wijns and Kowalczyk (2007) who, similarly to the approach of Barbosa and Silva (2006), allow for input from the interpreter to help ensure a geologically reasonable solution. Several inversions are performed with random values for several control parameters (e.g. regularization parameters). The resulting suite of recovered models are then visually inspected by the interpreter and ranked by how geologically reasonable they are (based on the interpreter's prior knowledge). A genetic algorithm then takes that ranking into account and modifies the control parameter set to generate a new suite of models. This procedure progressively converges towards a reasonable set of solutions. The strategies of Barbosa and Silva (2006) and Wijns and Kowalczyk (2007) rely on a significant increase in the amount of inversions performed.

Bosch et al. (2001) and Guillen et al. (2008) work in a stochastic inversion framework that directly recovers rock type (i.e. a lithologic inversion) from a list of those assumed present. Prior information is placed into the problem through probability density functions and topology rules (relationships between rock units) applied during the sampling. The model space is investigated (sampled) through a random walk process, an approach proposed by Mosegaard and Tarantola (2002). Their approach is attractive as it directly recovers lithologic models, removing the interpretation step required to go from physical property models to lithology.

There is significant computational cost associated with most of the approaches mentioned above. I choose to follow work by Li and Oldenburg (1996, 1998, 2000, 2003) and work in a relatively computationally efficient deterministic framework in which a well-behaved function is minimized subject to optional constraints. Although this is only one of many possible inversion frameworks, many of those mentioned above contain mathematical commonality and advances made in one framework can carry over to others.

1.3 Specific objectives

Discussions with geologists connected to exploration studies have identified several types of geologic information which could help improve results if incorporated into geophysical inversions. Those types of information are as follows and are discussed in more detail in Chapter 2.

- physical property measurements on rock samples;

- lithology observations (combined with petrophysical information);
- structural orientations;
- structural contacts between rock units;
- physical properties change sharply (e.g. across offset faults) or smoothly (e.g. across zones of alteration) between rock units;
- relative positions of rock units (e.g. a particular rock type is known to lie above another but the location of the contact is unknown);
- physical properties increase or decrease in particular directions (e.g. density often increases with depth);
- expected target shapes and aspect ratios (e.g. an intrusive body should be disk-like or pipe-like).

A specific goal of my research was to develop methods for incorporating all of the above information into geophysical inversion algorithms within a deterministic framework (i.e. minimization of an objective function) in order to enhance inversion functionality for use in mineral exploration. Another specific goal was to develop a method, in that same framework, that could combine several complimentary types of geophysical data to obtain improved subsurface models.

1.4 Thesis arrangement

This is a manuscript-based thesis and the main body (Chapters 2 through 5) consists of four research articles. A version of Chapter 2 has been submitted for publication in a peer-reviewed journal and two other versions have been published in conference extended abstracts. Versions of Chapters 3 and 4 have been accepted for publication (both are in press at the time of submission of this thesis). A version of Chapter 5 will also be submitted for publication in a peer-reviewed journal. See the footnotes at the start of each chapter for publication information.

Chapter 2 provides an overall view of the work by combining methods into a single demonstrative example. It is not meant to provide critical analysis of the methods but, rather, it functions as an introductory reading before moving to the development of theory and methods provided in the subsequent chapters. Chapter 3 describes research on a magnetic inverse problem. Although that research spends significant focus on the inverse problem itself, the paper holds true to the overall theme of the thesis mentioned above; specifically, the methods allow incorporation of geologic information regarding remanence. Chapter 4 is involved with the incorporation of

orientation information into geophysical inversions. Chapter 5 deals with recovery of geologically realistic models and cooperatively inverting multiple types of geophysical data. Chapter 6 concludes the thesis by summarizing the results, discussing their significance, providing a critical assessment of the methods developed, and mentioning potential future research related to the methods.

Due to the nature of the manuscript-based thesis, there will be some duplication of information between chapters. To reduce manuscript length, some important mathematical details were left out of one manuscript comprising this thesis (Chapter 3). That material is not crucial to an understanding of the methods developed and is provided in Appendix B.

1.5 Mathematical nomenclature

The text of this thesis is typeset in a sans-serif font; the mathematics are typeset in a serif font to help the reader pick them out. In the mathematics of this thesis I use the following conventions. Scalars are represented by plain Roman or Greek letters (e.g. U and Φ). Field vectors (of two or three components) are indicated with an arrow (e.g. \vec{J}). Other vectors (i.e. lists of values) are represented by bold lowercase letters (e.g. \mathbf{m}); unless stated otherwise, a vector is a column vector. Matrices are represented by bold uppercase letters (e.g. \mathbf{A}).

1.6 References

- Barbosa, V. C. F., and J. B. C. Silva, 1994, Generalized compact gravity inversion: *Geophysics*, **59**, 57–68.
- Barbosa, V. C. F., and J. B. C. Silva, 2006, Interactive 2D magnetic inversion – a tool for aiding forward modeling and testing geologic hypotheses: *Geophysics*, **71**, L43–L50.
- Bosch, M., A. Guillen, and P. Ledru, 2001, Lithologic tomography – an application to geophysical data from the Cadomian belt of northern Brittany, France: *Tectonophysics*, **331**, 197–227.
- Chasserau, P., and M. Chouteau, 2003, 3D gravity inversion using a model of parameter covariance: *Journal of Applied Geophysics*, **52**, 59–74.
- Farquharson, C. G., M. R. Ash, and H. G. Miller, 2008, Geologically constrained gravity inversion for the Voisey's Bay ovoid deposit: *The Leading Edge*, **27**, 64–69.
- Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure: *Geophysical Journal International*, **134**, 213–227.
- Guillen, A., Ph. Calcagno, G. Courrioux, A. Joly, and P. Ledru, 2008, Geological modelling from field data and geological knowledge – Part II. Modelling validation using gravity and magnetic data inversion: *Physics of the Earth and Planetary Interiors*, **171**, 158–169.
- Guillen, A., and V. Menichetti, 1984, Gravity and magnetic inversion with minimization of a specific functional: *Geophysics*, **49**, 1354–1360.
- Last, B. J., and K. Kubik, 1983, Compact gravity inversion: *Geophysics*, **48**, 713–721.
- Li, Y., and D. W. Oldenburg, 1996, 3-D inversion of magnetic data: *Geophysics*, **61**, 394–408.
- Li, Y., and D. W. Oldenburg, 1998, 3D inversion of gravity data: *Geophysics*, **63**, 109–119.
- Li, Y., and D. W. Oldenburg, 2000, Incorporating geologic dip information into geophysical inversions: *Geophysics*, **65**, 148–157.
- Li, Y., and D. W. Oldenburg, 2003, Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method: *Geophysical Journal International*, **152**, 251–265.
- Mosegaard, K., and A. Tarantola, 2002, Probabilistic approach to inverse problems: *International Geophysics*, **81**, 237–265.

- Phillips, N. D., 2001, Geophysical inversion in an integrated exploration program – examples from the San Nicolás deposit, Master's Thesis, University of British Columbia, Canada.
- Portniaguine, O., and M. S. Zhdanov, 1999, Focusing geophysical inversion images: *Geophysics*, **64**, 874–887.
- Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms: *Physica D*, **60**, 259–268.
- Vogel, C. R., and M. E. Oman, 1998, Fast total variation based reconstruction of noisy, blurred images: *IEEE Transactions on Image Processing*, **7**, 813–824.
- Wijns, C., and P. Kowalczyk, 2007, Interactive geophysical inversion using qualitative geological constraints: *Exploration Geophysics*, **38**, 208–212.
- Williams, N. C., 2006, Applying UBC-GIF potential field inversions in greenfields or brownfields exploration: *Australian Earth Sciences Convention, ASEG/GSA, Abstracts*.

Chapter 2

Integrating geologic and geophysical data through advanced constrained inversions¹

2.1 Introduction

Geophysical inversion seeks to recover models of the Earth's physical properties (e.g. density, conductivity) that can adequately reproduce anomalies in geophysical survey data (e.g. gravity, DC resistivity) while being consistent with geologic information. The physical properties are related to mineral structure and, hence, the models are an important source of information for understanding subsurface geology as it applies to mineral exploration.

Due to data uncertainty and other aspects inherent to the underdetermined geophysical inverse problem, there are an infinite number of models that can fit the geophysical data to the desired degree (i.e. the problem is nonunique). Additional information is essential for a unique solution. Incorporating prior geologic knowledge, and combining several complimentary types of geophysical data collected over the same Earth region, can reduce ambiguity and enhance inversion results, leading to more reliable Earth models. Phillips (2001), Williams (2006) and Farquharson et. al. (2008) provide examples of incorporating physical property information to dramatically improve inversion results. Below we outline some additional types of geologic information that are commonly available and can be incorporated into our inversion algorithms.

¹Two versions of this chapter have been published: Lelièvre, P. G., D. W. Oldenburg and N. C. Williams, 2008, Constraining geophysical inversions with geologic information: 78th SEG Annual Meeting Extended Abstracts, 1223–1227; Lelièvre, P. G., D. W. Oldenburg and N. C. Williams, 2009, Integrating geologic and geophysical data through advanced constrained inversions: ASEG Annual Meeting Extended Abstracts.

2.2 Types of geologic information available

The geologic information available can come from many sources: surface mapping, drill-holes, hand samples, in situ measurements, preliminary mining or any other manner in which geologic data are collected. We categorize geologic information as either located (spatially tied) or non-located. Below we list some types of geologic information that we can incorporate using our methods.

Located information:

- physical property measurements on rock samples;
- lithology observations (combined with petrophysical information);
- structural orientations;
- structural contacts between rock units.

Non-located information:

- physical properties change sharply (e.g. across offset faults) or smoothly (e.g. across zones of alteration) between rock units;
- relative positions of rock units (e.g. a particular rock type is known to lie above another but the location of the contact is unknown);
- physical properties increase or decrease in particular directions (e.g. density often increases with depth);
- expected target shapes and aspect ratios (e.g. an intrusive body should be disk-like or pipe-like).

An added consideration is that we should take into account the reliability of any type of information included in an inversion, be it geophysical data or geologic information. For example, direct observations should be considered more reliable than interpolations, which should be considered more reliable than inferences or hypotheses. If the information is qualitative (e.g. an expected shape of a target) it may be sound or speculative. Even if the information is quantitative (e.g. physical property measurements taken on rock samples) there may be a wide range of reliability reflected by the procedures taken to obtain that information. If a relative or absolute value can be placed on the reliability then this information should be incorporated alongside the related geologic information.

Rock type	Density (g/cc)	Slowness (s/km)
bedrock	0.0	0.0
cover	-0.5	0.5
block	0.5	-0.5

Table 2.1: The anomalous physical properties of each rock type for the synthetic example.

2.3 An illustrative synthetic scenario

We now introduce an illustrative 2D synthetic example in which we will incorporate much of the above geologic information while running through a hypothetical exploration scenario. The synthetic model is a simplified version of one used in Williams (2006), which was based on a nickel exploration scenario in the Eastern Goldfields of Yilgarn Craton in Western Australia. We choose to present a 2D example for ease of viewing but the methods are easily extended to 3D.

We consider two geophysical surveys: gravity and cross-well seismic tomography. These two surveys are complimentary in that they sense the earth in different ways. Each survey has different sensitivities to, and resolution capabilities in, different parts of the subsurface. Consequently, independent inversions of each data type are able to recover different parts of the subsurface to different degrees, as will become evident. These two surveys have been chosen for this example due to their computational simplicity but they could represent any two complimentary surveys.

The true model contains three rock types: bedrock, a cover layer and a dipping block. The anomalous physical properties (i.e. with background values subtracted) of each rock type are indicated in Table 2.1. From now on we will use the terms “density” and “slowness” to refer to the anomalous quantities. The discrete inversion mesh can be seen in Figure 2.1(b). The density and slowness models are shown in Figure 2.2(a) and 2.2(b) respectively.

Gravity and travel-time data are created for the models and a small amount of random noise is added before inverting. The gravity data are collected across the surface of the model and the cross-well seismic tomography ray-paths run from $x = -396\text{m}$ to $x = 396\text{m}$ every 50m in depth (z). The data are plotted in Figure 2.1.

2.4 Geologically unconstrained inversion

In our numerical inverse solutions, the Earth region of interest is divided into many cells within an orthogonal mesh with the physical property (or properties) of interest being constant across each cell. The inverse problem is formulated as an optimization that involves minimization of an objective function, Φ , that combines a data misfit measure, Φ_d , with a regularization measure, Φ_m :

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \beta \Phi_m(\mathbf{m}). \quad (2.1)$$

Here, \mathbf{m} is the model vector that holds the physical property values in each mesh cell of our discretised Earth and β is a trade-off parameter that controls the relative size of the misfit and regularization measure and allows us to tune the level of data fit as desired.

In the absence of more specific, constraining geologic information we choose to seek a smoothly varying model that does not contain unreasonably high values. We recognize that this “minimum structure” inversion is just one of several options, others including “compact” inversions (Last and Kubik, 1983), “focussed” inversions (Portniaguine and Zhdanov, 1999) and non-smooth inversions (Farquharson and Oldenburg, 1998). However, we consider the minimum structure option due to its computational simplicity. Our minimum structure regularization function is based on that of Li and Oldenburg (1996),

$$\begin{aligned} \Phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V w_x \left(\frac{\partial m}{\partial x} \right)^2 dv \\ & + \int_V w_y \left(\frac{\partial m}{\partial y} \right)^2 dv \\ & + \int_V w_z \left(\frac{\partial m}{\partial z} \right)^2 dv, \end{aligned} \quad (2.2)$$

which can be readily discretized on the inversion mesh. This regularization function makes the inverse problems tractable (it allows a single model to be recovered) and it allows incorporation of much geologic information into the inversion. The first term in equation (2.2) is a smallness or closeness term. In the case of a zero-valued reference model, m_{ref} , this term encourages the inversion to recover models with low values of the physical property. When a reference model is incorporated, the inversion attempts to match it as closely as possible while still fitting the data to the desired degree. The latter terms in equation (2.2) involving derivatives are smoothness terms that encourage recovery of spatially smooth models.

Default (unconstrained) inversion results for our synthetic example are shown in Figure 2.2(c) and 2.2(d). For all inversions we set $w_s = 0$ and $w_x = w_z = 1$ ($w_y = 0$ for the 2D problem). Without further geologic information, the models in Figure 2.2(c) and 2.2(d) would be our best guesses for the subsurface density and slowness distributions. However, the gravity inversion result fails to image the cover layer and dip of the central body, and the seismic tomography inversion fails to image the central body entirely.

2.5 Incorporating physical property value information

Physical property information is commonly available from measurements taken on surface outcrop or drill-core samples. This information can be incorporated into the inversions in a natural way using the reference models and weights in the regularization measure in equation (2.2) or through bound constraints added to the optimization problem:

$$\min_{\mathbf{m}} \quad \Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \beta \Phi_m(\mathbf{m}) \quad (2.3a)$$

$$s.t. \quad L_i \leq m_i \leq U_i \quad \text{for some or all } i \quad (2.3b)$$

where L_i and U_i are the lower and upper bounds for the i^{th} model cell. Li and Oldenburg (2003) discuss one possible method of solution to the bounded inverse problem, the logarithmic barrier approach. We employ a more robust and efficient gradient-projection method to solve the inverse problem: the gradient-projection-reduced-Newton (GPRN) method of Vogel (2002).

Returning to the synthetic scenario, assume that surface mapping has been performed in the next phase of exploration and physical property measurements have been taken on rock samples. We can now place bounds on the cells along the surface of the model. Reliability information (e.g. based on the statistics of the measurements taken on the surface samples) can be used to determine the spread on the bounds for any particular cell. Here, we set the bounds within ± 0.02 of the true values at the surface. The results with these bounds applied are shown in Figure 2.2(e) and 2.2(f). The results are immediately improved: the gravity inversion now indicates the cover layer and the dip of the central body, and the central body is now indicated in the tomography inversion.

2.6 Incorporating physical property trend information

Knowledge of the relative positions of rock units may allow us to specify physical property trends (i.e. physical property increase or decrease in particular directions). Assume now that in our synthetic scenario, knowledge of the geology leads us to expect a less dense and slower layer of cover that has a depth of no greater than 100m. Hence, away from the central body, the density should only increase with depth and the slowness should only decrease with depth in the top 100m. We may also specify that below a certain depth, say 150m for this example, we do not expect the density to increase or the slowness to decrease any further with depth.

We incorporate this physical property trend information through linear inequality constraints of the general form $\mathbf{A}\mathbf{m} \geq \mathbf{b}$. Lelièvre and Oldenburg (in press) (Chapter 4) provides detail on the implementation. For example, if density model element 1 (ρ_1) lies vertically above density model element 2 (ρ_2) then in the top 100m of the model we desire that $\rho_2 \geq \rho_1$. This leads to the inequality equation $\rho_2 - \rho_1 \geq 0$ which would define one row of \mathbf{A} and a zero element in \mathbf{b} . We use a logarithmic barrier method, as implemented by Li and Oldenburg (2003), to solve the resulting optimization problem

$$\min_{\mathbf{m}} \quad \Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \beta\Phi_m(\mathbf{m}) \quad (2.4a)$$

$$s.t. \quad \mathbf{A}\mathbf{m} \geq \mathbf{b}. \quad (2.4b)$$

Although there are alternatives to the logarithmic barrier method, it has proven to be a feasible solution method for large 3D geophysical inverse problems with simple bound constraints and it extends to linear inequalities without issue.

Bound constraints can be incorporated into the problem in equation (2.4) alongside the linear inequalities. If we wish to bound the i^{th} model parameter between lower and upper bounds L_i and U_i (i.e. $L_i \leq \rho_i \leq U_i$) then we would add two equations to the $\mathbf{A}\mathbf{m} \geq \mathbf{b}$ system:

$$\rho_i \geq L_i \quad (2.5)$$

$$-\rho_i \geq -U_i. \quad (2.6)$$

Incorporating the additional depth trend information into our synthetic inversions yields the improved results in Figure 2.2(g) and 2.2(h). Especially evident are the reduction of anomalous dense material at depth and an improved recovery of the cover unit in the density model.

2.7 Incorporating orientation and aspect ratio information

The smoothness terms in equation (2.2) measure model gradients in the three axial directions defined by the inversion mesh. By altering the relative values of the smoothness weights w_x , w_y and w_z in equation (2.2) one can cause the recovered models to become smoother (i.e. elongated) in some mesh-orthogonal direction(s) compared to the other(s). Aspect ratio information can be used to determine the relative values of the smoothness weights in each direction. A further generalization of the objective function by Li and Oldenburg (2000) (and see also Lelièvre and Oldenburg (in press); Chapter 4) allows the coordinate axes to be rotated such that elongations can be specified in any Cartesian coordinate frame. For a 2D problem the new regularization function is of the form

$$\begin{aligned}\Phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V w_{x'} \left(\cos \theta \frac{\partial m}{\partial x} + \sin \theta \frac{\partial m}{\partial z} \right)^2 dv \\ & + \int_V w_{z'} \left(-\sin \theta \frac{\partial m}{\partial x} + \cos \theta \frac{\partial m}{\partial z} \right)^2 dv\end{aligned}\quad (2.7)$$

where θ is the dip angle that rotates the original (mesh orthogonal) axes x and z into the new coordinate frame with axes x' and z' . In 3D there are three angles and we do not show the resulting complicated regularization measure here. The smoothness weights can be homogeneous across the entire mesh or can be set to different values in different regions. Orientations can thereby be specified globally or locally.

Returning to our synthetic scenario, the geologic information available and the previous inversion results lead us to expect that the region contains a horizontal cover layer interrupted by a central dipping body. Assume now that a drilling program spots a drill-hole as indicated by the white lines in Figure 2.3(a) and 2.3(b). If physical property measurements are taken on drill-core samples then we can place additional bounds on the cells along the drill-hole trace, producing the results in Figure 2.3(a) and 2.3(b). The surface mapping and drill-hole information also allow an interpretation of the dip of the central body, which we will now incorporate.

We expect that the cover unit is no thicker than 100m. Hence, down to that depth we set the w_z/w_x ratio below the default value of unity. This will encourage features that are elongated in the x -direction (horizontal) and may take larger jumps in the z -direction (vertical). To incorporate the structural orientation information regarding the dip of the central body we need

to rotate the coordinate system such that the new x' -axis lies along the expected dip direction. Hence, in a region immediately around the outcrop of the target unit we specify a dip of 45° and set $w_{z'}/w_{x'}$ below unity to encourage features that are elongated in the down-dip direction. Outside of this region we do not rotate the coordinate system or apply a non-unity $w_{z'}/w_{x'}$ ratio.

Reliability estimates and knowledge of the scale of geologic information can be used to guide to how far sparse information can be extrapolated to influence its surroundings and thereby populate a larger region of the inversion mesh. In the synthetic we set the $w_{z'}/w_{x'}$ ratio low ($= 0.01$) where this information is most reliable: cells at or adjacent to the surface and the drill-hole trace. We increase the ratio back to unity as the cells get further away from those. The inversion results with this information included are in Figure 2.3(c) and 2.3(d). These results show further improvement and nearly recover the correct depth extents of the cover and dipping block.

2.8 Recovering rock type models

In this final stage, we wish to better constrain the recovered density and slowness values to lie within three separated narrow ranges as is expected assuming the presence of three rock types. Figures 2.4 and 2.5 show histograms of the physical property values for all models presented. The best recovered models so far, in Figure 2.3(c) and 2.3(d), are each starting to bin into three narrow ranges, as seen in Figure 2.5(c) and 2.5(d), but we would like to do better.

2.8.1 Recovering sharp boundaries between rock units

The use of ℓ^2 -norms (sum-of-squares) in the discretized version of equation (2.2) results in recovered models that exhibit a smooth, smeared-out appearance. Unfortunately, smooth models do not generally fit with geologists' ideas about the subsurface, which can involve sharp interfaces (physical property discontinuities) between rock units (regions with nearly constant physical properties). Sharp boundaries can be generated using other norms, for example as implemented by Farquharson and Oldenburg (1998). However, moving away from ℓ^2 -norms complicates the optimization problem.

We have developed an iterative inversion procedure that allows for recovery of geologically

realistic models while using ℓ^2 -norms. Chapter 5 explains the strategy in detail. The essence of the procedure is to re-weight previous inversion results based on some measure of structure. The method does not require any alteration to the inverse problem and it can therefore be performed with pre-existing inversion algorithms.

The first step is to take the best model so far and calculate the magnitude of the spatial model gradient:

$$\|\vec{\nabla}m\| = \sqrt{(\nabla_x m)^2 + (\nabla_y m)^2 + (\nabla_z m)^2}. \quad (2.8)$$

Smoothness weights are then set low in the regions where $\|\vec{\nabla}m\|$ is high and vice versa. This encourages the subsequent inversion to make interfaces sharper (between rock units) and to make smoother areas constant (within rock units). This procedure is repeated in an iterative manner until convergence is reached.

2.8.2 Cooperative inversion

The iterative procedure mentioned above extends naturally to structural cooperative inversion. In the cooperative procedure, the gradient magnitude is calculated for two (or more) inversion results, each using a different type of geophysical data. Those gradient magnitudes are normalized and added together. That result is then used to set the smoothness weights for all of the inversions in the next iteration.

The synthetic scenario we present is ideal for a cooperative inversion. Each geophysical survey senses the subsurface in a different way and each adequately resolves different characteristics of the true model. For example, the gravity inversions resolve the lateral extents of the central body well but the tomography inversions do not, and the opposite is true for the depth extent of the cover layer. This is clear from comparing the geologically unconstrained results in Figure 2.2(c) and 2.2(d). By combining both data sets into a cooperative inversion we hope to overcome the lack of resolution in each geophysical survey.

The results in Figure 2.3(e) and 2.3(f) were obtained after using this cooperative strategy with only the surface bounds and depth trend information incorporated. For comparison, the results in Figure 2.3(g) and 2.3(h) were obtained through independent inversions incorporating that same geologic information and using approximate ℓ^1 -norms (an Ekblom measure with $p = 1.1$ as implemented by Farquharson and Oldenburg (1998); refer to Chapter 5 in equation (2.2) to create models with sharper interfaces.

The cooperative procedure is able to overcome the lack of resolution in each geophysical survey and provides an obvious improvement over the results in Figure 2.3(g) and 2.3(h). The three rock units are clearly defined in Figure 2.3(e) and 2.3(f) with sharp interfaces between them. The density and slowness values are better confined within three narrow ranges, as indicated by Figure 2.5(e) and 2.5(f), and are close to the true values. The true depth extents of the cover layer and the target unit have been located within one mesh cell. When more geologic information is placed into the cooperative procedure the results show further improvement.

2.9 Conclusion

When not constrained by geologic information, default inversions can generate reasonable results, recovering spatially simple physical property distributions that honour the survey data. However, such first-pass results may not honour the geologic information, much of which can now be incorporated to improve inversion results, as is evident from the examples above and in the work of Phillips (2001), Williams (2006) and Farquharson et. al. (2008).

Furthermore, different physical property models recovered independently from different geophysical data sets can be inconsistent with each other and inverting the data in a joint or cooperative fashion may help overcome the lack of resolution in each data set. We have developed a cooperative inversion strategy that provides such a tool.

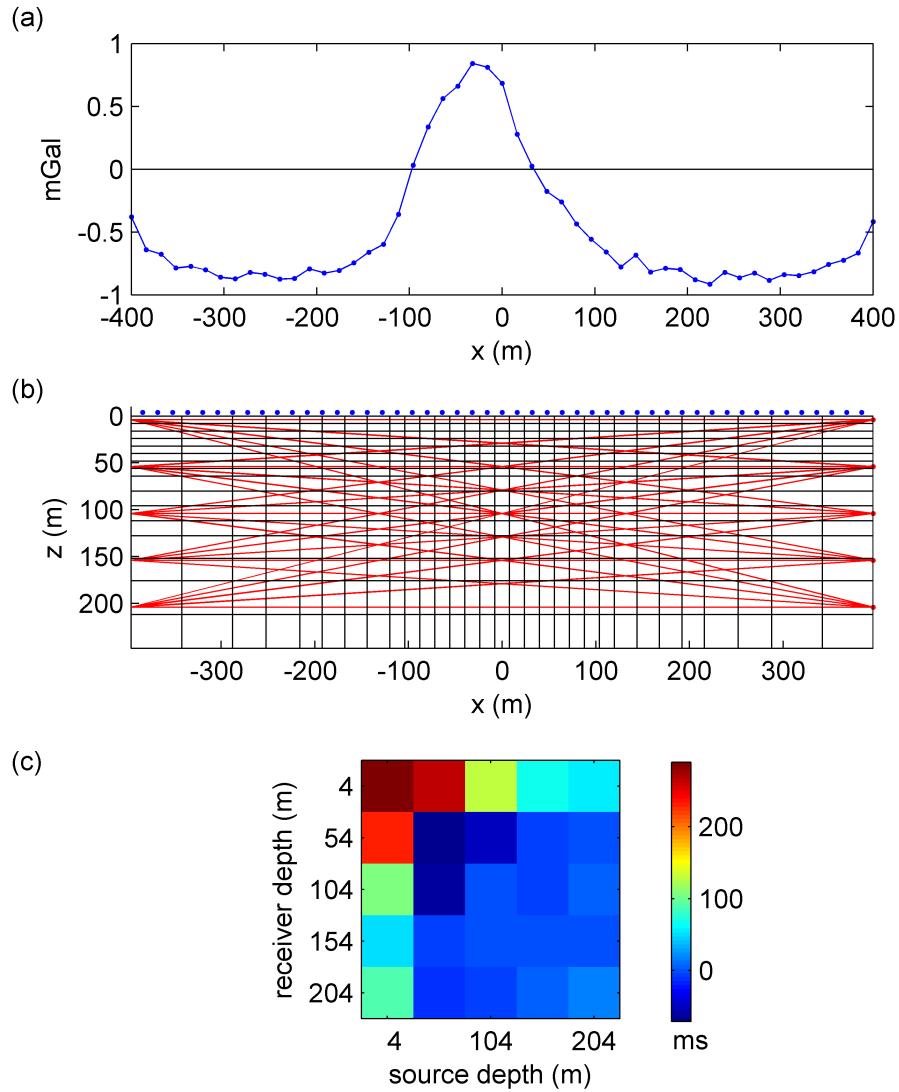


Figure 2.1: Image (a) shows the gravity data used in the inversions. Image (b) shows the inversion mesh in black, the locations of the gravity data are indicated with blue dots (above the surface), and the cross-well seismic tomography ray paths are indicated with red lines. Image (c) shows the seismic tomography travel-time data used in the inversions.

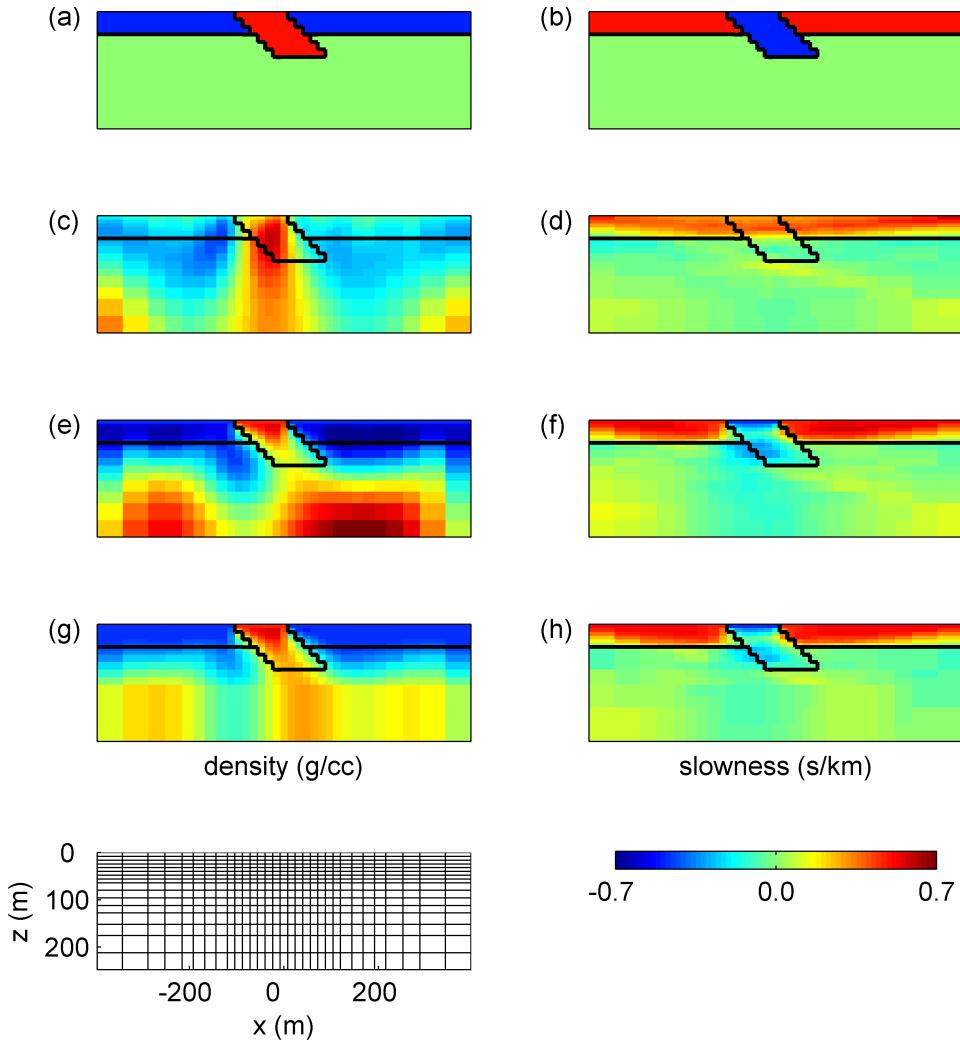


Figure 2.2: True and inverted models for the gravity and cross-well seismic tomography synthetic example. Density models are on left, slowness models on right. To remove clutter, we do not show axis labels on plots (a) through (h) but we provide the inversion mesh below (g), which is a duplication of that shown in Figure 2.1(b). The colour scale for all models is $[-0.7, 0.7]$, using the colour bar below (h). Superimposed black lines in (a) through (h) indicate the rock unit boundaries in the true model. From top to bottom the models are: (a,b) the true models; (c,d) geologically unconstrained results; (e,f) results with bounds applied along the surface; (g,h) results with additional linear inequality constraints applied to enforce expected depth trends. Models from subsequent inversions are shown in Figure 2.3.

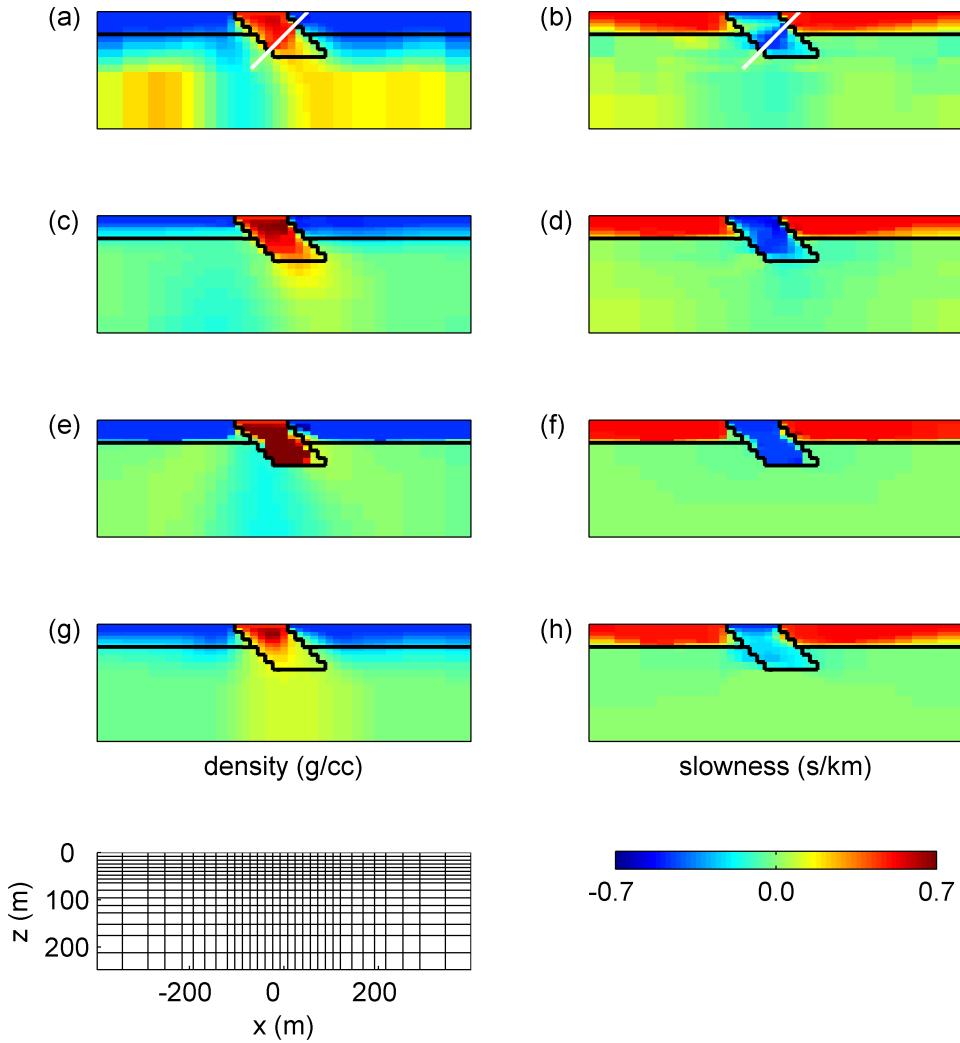


Figure 2.3: Inverted models for the gravity and cross-well seismic tomography synthetic example. Density models are on left, slowness models on right. To remove clutter, we do not show axis labels on plots (a) through (h) but we provide the inversion mesh below (g), which is a duplication of that shown in Figure 2.1(b). The colour scale for all models is $[-0.7, 0.7]$, using the colour bar below (h). Superimposed black lines in (a) through (h) indicate the rock unit boundaries in the true model. The location of a drill-hole is indicated with a white line in (a) and (b). From top to bottom the models are: (a,b) results with additional drill-hole bounds, (c,d) results with additional orientation information incorporated; (e,f) results of our structural cooperative inversion strategy after incorporating only surface bounds and expected depth trends; (g,h) independent results using approximate ℓ^1 -norms and the same geologic information as in (e,f). Models from previous inversions are shown in Figure 2.2.

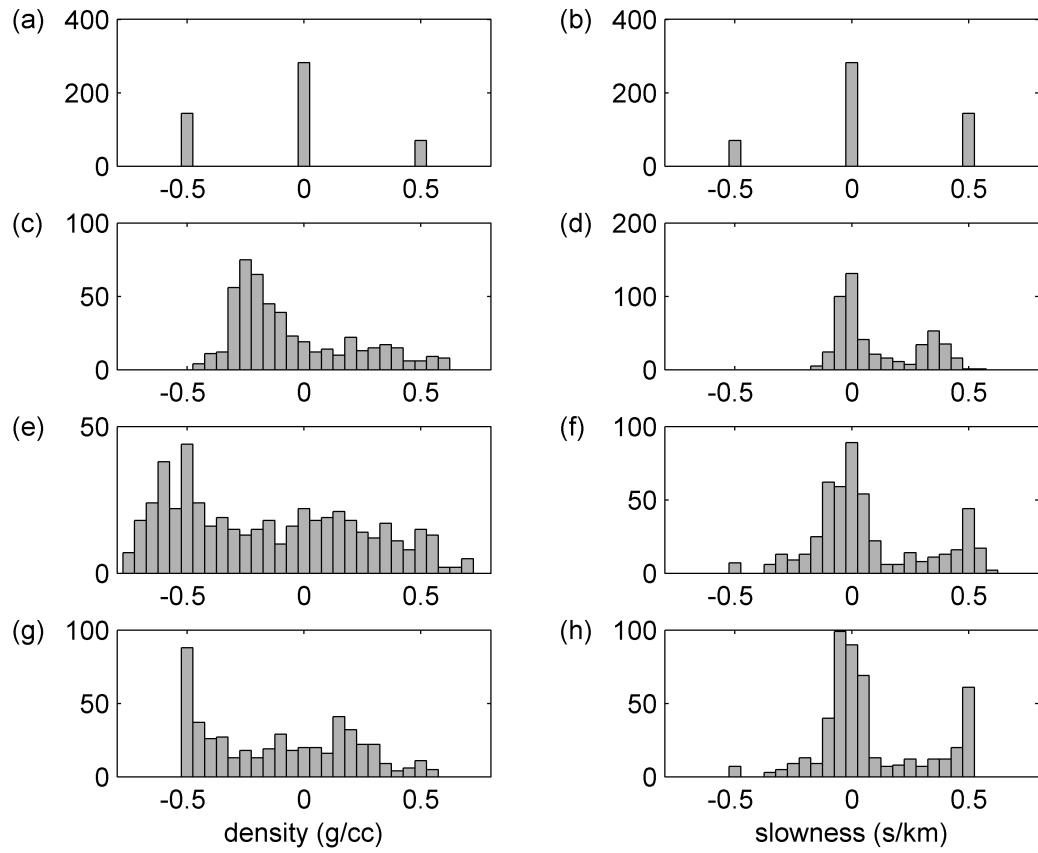


Figure 2.4: Histograms showing the distribution of physical property values in the corresponding models in Figure 2.2.

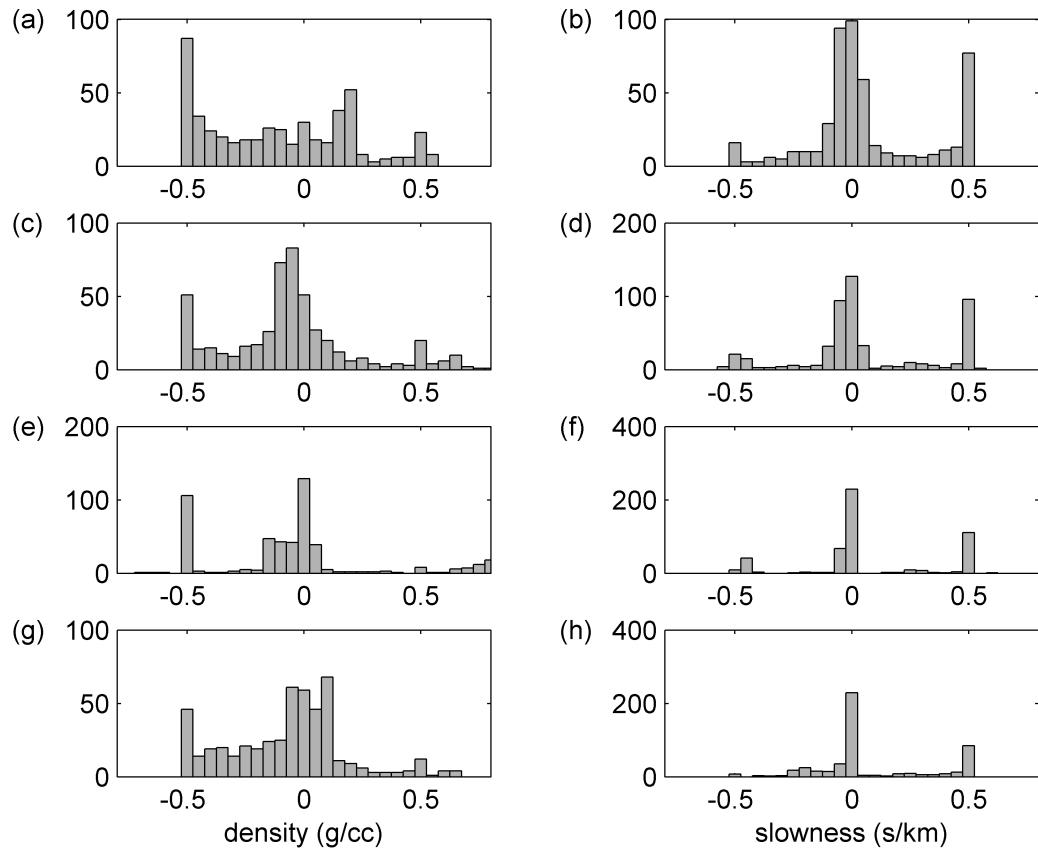


Figure 2.5: Histograms showing the distribution of physical property values in the corresponding models in Figure 2.3.

2.10 References

- Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure: *Geophysical Journal International*, **134**, 213–227.
- Farquharson, C. G., M. R. Ash, and H. G. Miller, 2008, Geologically constrained gravity inversion for the Voisey's Bay ovoid deposit: *The Leading Edge*, **27**, 64–69.
- Last, B. J., and K. Kubik, 1983, Compact gravity inversion: *Geophysics*, **48**, 713–721.
- Lelièvre, P. G, and D. W. Oldenburg, *in press*, A comprehensive study of including structural orientation information in geophysical inversions: *Geophysical Journal International*.
- Li, Y., and D. W. Oldenburg, 1996, 3D inversion of magnetic data: *Geophysics*, **61**, 394–408.
- Li, Y., and D. W. Oldenburg, 2000, Incorporating geologic dip information into geophysical inversions: *Geophysics*, **65**, 148–157.
- Li, Y., and D. W. Oldenburg, 2003, Fast inversion of large-scale magnetic data using wavelet transforms and logarithmic barrier method: *Geophysical Journal International*, **152**, 251–265.
- Phillips, N. D., 2001, Geophysical inversion in an integrated exploration program – examples from the San Nicolás deposit: Master's Thesis, University of British Columbia, Vancouver, Canada.
- Portniaguine, O., and M. S. Zhdanov, 1999, Focusing geophysical inversion images: *Geophysics*, **64**, 874–887.
- Vogel, C. R., 2002, Computational methods for inverse problems: SIAM.
- Williams, N. C., 2006, Applying UBC-GIF potential field inversions in greenfields or brownfields exploration: Australian Earth Sciences Convention, ASEG/GSA, Abstracts.

Chapter 3

A 3D total magnetization inversion applicable when significant, complicated remanence is present²

3.1 Introduction

3.1.1 The problem of remanence in magnetic interpretation

The total magnetization vector, \vec{J}_{tot} , within an isotropic body can be modelled as the vector sum

$$\vec{J}_{tot} = \chi \vec{H}_0 + \vec{J}_{rem} \quad (3.1)$$

where χ is magnetic susceptibility and \vec{H}_0 is the Earth's magnetic field (Bossavit, 1998). The first term in equation (3.1) is the induced component of the magnetization,

$$\vec{J}_{ind} = \chi \vec{H}_0, \quad (3.2)$$

and \vec{J}_{rem} is the remanent component. Remanent magnetization (or remanence) is a permanent magnetization obtained in the past that can often be oriented in a direction different from the Earth's field today. Hence, the induced and remanent components can be oriented in different directions.

Typical magnetic inversion routines such as that of Li and Oldenburg (1996) assume no remanent component exists. Hence, the magnetization is assumed to lie in the direction of the Earth's field and erroneous results can be obtained if this assumption is made incorrectly. To demonstrate

²A version of this chapter is in press: Lelièvre, P. G., and D. W. Oldenburg, A 3D total magnetization inversion applicable when significant, complicated remanence is present: *Geophysics*.

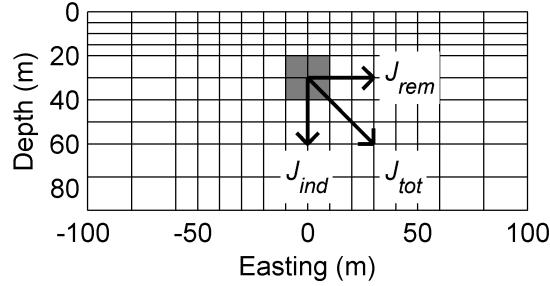


Figure 3.1: A vertical cross-section at northing = 0m through the true 3D synthetic model. The mesh cell spacing is identical in the easting and northing directions. A small magnetic body with $\chi = 0.1$ (plotted in grey here) resides in the centre of the mesh within a nonmagnetic background. The induced component of the magnetization is labelled J_{ind} , the remanent component J_{rem} and the total magnetization J_{tot} .

this, consider the following 3D synthetic example. The model mesh is shown in Figure 3.1 with an overlayed schematic showing the magnetization of the central body: the Earth's field has strength 40000nT, is vertical and downward (an inclination of 90°), and the remanent magnetization is horizontal towards the east (right in this diagram, a declination of 90°). The Königsberger ratio

$$Q = \frac{\|\vec{J}_{rem}\|}{\|\vec{J}_{ind}\|} = \frac{\|\vec{J}_{rem}\|}{\chi\|\vec{H}_0\|} \quad (3.3)$$

for the body is set to unity ($Q = 1$) so that the magnitudes of the induced and remanent magnetization components are equal; the resulting total magnetization is oriented with a 45° dip.

Figure 3.2 compares the different contributions to the TMI (total magnetic intensity) response (TMI measurements are $B = \mu_0 H$). To invert this data we use the methods of Li and Oldenburg (1996); Li and Oldenburg (2003). Figure 3.3(a) shows the susceptibility model recovered from inversion of the data in Figure 3.2(c) assuming no remanence exists (i.e. assuming the total magnetization is in the direction of the Earth's inducing field). The inversion had trouble fitting the data and the recovered model bears little resemblance to the true model, with significant susceptible material placed towards the mesh boundaries.

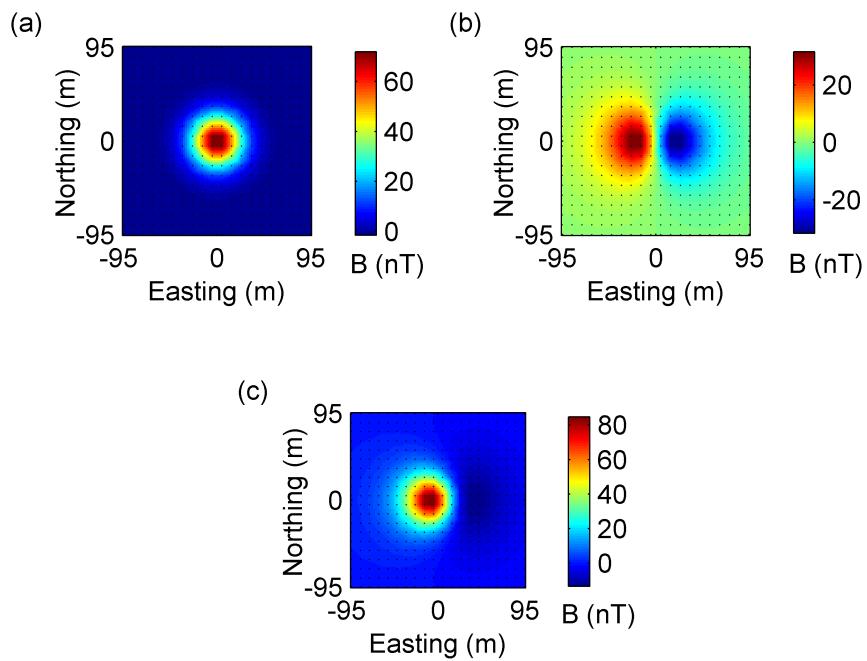


Figure 3.2: Map views of the (a) induced and (b) remanent components of the TMI response, and (c) total combined TMI response 10m above the mesh in Figure 3.1.

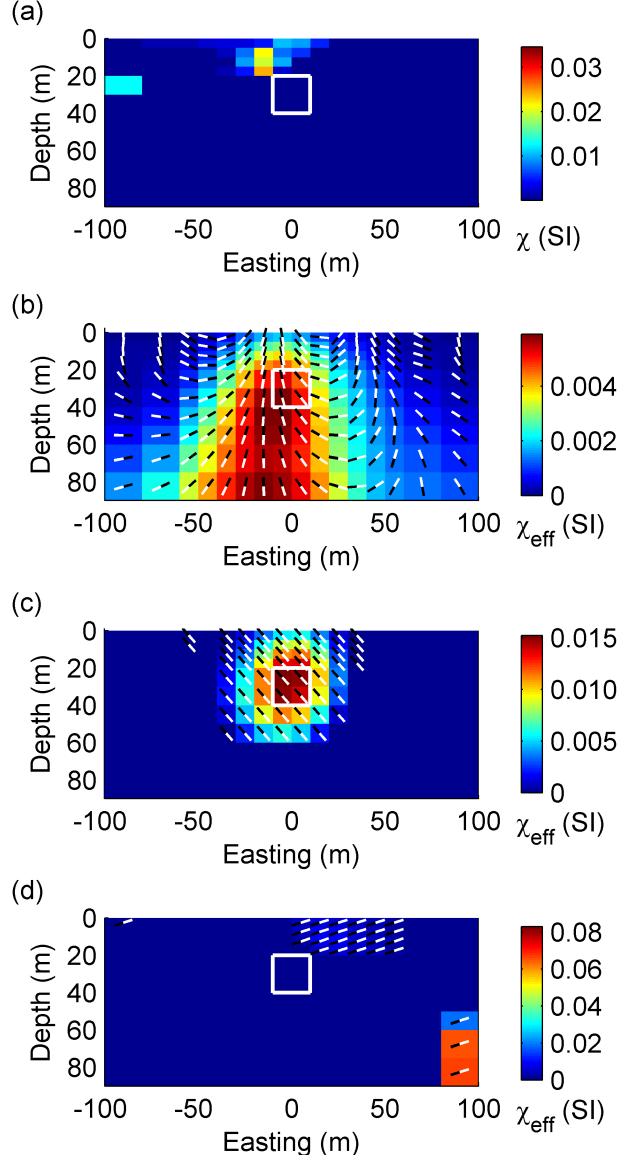


Figure 3.3: Vertical cross-sections at northing = 0m through 3D models recovered from inversion of the data in Figure 3.2(c): (a) susceptibility inversion with positivity imposed and assuming no remanence exists; (b) TMVC inversion with $\gamma = 1.0$ and no bounding applied; (c) TMVS inversion with $\theta = 90^\circ$ in the initial halfspace model; (d) TMVS inversion with $\theta = -45^\circ$ in the initial halfspace model. No reference models were included in these inversions. The colour scale used in (a) is SI susceptibility (unitless), in all others it is magnetization amplitude in units of effective susceptibility (i.e. normalized by the Earth's field strength). Total magnetization vectors are indicated by black and white lines with white at the head. Vectors with zero amplitude are not displayed. The white squares indicate the position of the block in the true model.

3.1.2 Previous approaches to dealing with remanence

Several authors have approached the problem of remanence by assuming simple causative bodies with uniform total magnetization directions. Li et al. (2004) provide an excellent summary of methods that can be used to estimate the total magnetization direction for such scenarios through analysis of the anomalies in magnetic data. The magnetization direction derived through these methods could be used in a subsequent inversion for an effective susceptibility, χ_{eff} , equal to the magnetization amplitude divided by the Earth's field strength:

$$\chi_{eff} = \frac{\|\vec{J}_{tot}\|}{\|\vec{H}_0\|} = \frac{\|\chi\vec{H}_0 + \vec{J}_{rem}\|}{\|\vec{H}_0\|}. \quad (3.4)$$

The methods include that of Phillips (2005), which makes use of Helbig's moment method (Helbig, 1963); the multiscale edge method of Haney and Li (2002), which makes use of a continuous wavelet transform; and the cross-correlation method of Dannemiller and Li (2006), which uses the reduction-to-pole process. Those methods rely on an adequate separation between the anomalies from different bodies and assume that for each body a constant net magnetization direction is sufficient for describing the whole body. This simplistic scenario is not generally applicable to the mineral exploration applications of interest in this paper where the subsurface magnetization can be complicated (e.g. varying remanence directions and amplitudes). The approaches discussed below are applicable to such complicated mineral exploration scenarios.

Another approach is to address properties of the measured magnetic field that are independent, or at least minimally dependent, on the direction of total magnetization. This removes the need to specify the direction of total magnetization exactly. Shearer and Li (2004) developed an algorithm that directly inverts a data quantity having a minimal dependence on the direction of the magnetization: they consider total gradient data (the amplitude of the anomalous magnetic field gradient). For 2D magnetic problems, this data quantity is independent of the magnetization direction but in 3D it is weakly dependent. The algorithm of Shearer and Li (2004) fully incorporates the nonlinear relationship between the total gradient data and the subsurface magnetization and they invert for the magnitude of the magnetization on a 3D mesh without knowing the direction of magnetization. The advantage of their approach is that it is applicable to situations with complicated magnetic bodies and multiple remanence directions. Their forward computation requires prescription of some constant magnetization direction across the volume; in general the prescribed direction will not correspond to the true direction and this can introduce some error into their inverse solutions and unacceptable artifacts (that would produce erroneous interpretations) can be recovered. Furthermore, if the total gradient data are not

measured directly but is instead created from total field data through linear transforms then further significant error can be introduced.

Whereas Shearer and Li (2004) invert for magnetization amplitude, in this paper our intention is to explicitly calculate the magnetization amplitude *and* direction of the source distribution. The method of Shearer and Li (2004) deals with data that have a low sensitivity to one of the important quantities driving the magnetic problem: the magnetization direction. In order to recover the magnetization direction it would be more appropriate to use data that are as sensitive to (i.e. contain enough information about) the magnetization as possible. Foss (2006) and Schmidt and Clark (2006) discuss how, compared to total magnetic intensity (TMI) data, three-component and gradient tensor magnetic data can potentially provide increased sensitivity to source magnetization direction. We go no further into those types of data here but mention that use of such data with our inversion methods that follow would be a fairly straightforward task.

3.2 Inversion for magnetization

To tackle the magnetic remanence problem we invert magnetic data for a three-component subsurface magnetization vector, as opposed to isotropic magnetic susceptibility (a scalar). We consider two approaches: a Cartesian and a spherical framework, discussed below. For abbreviation purposes we will refer to these methods as TMV (total magnetization vector) inversions, with TMVC denoting the Cartesian formulation and TMVS denoting the spherical formulation.

Wang et al. (2004) developed what they called a magnetization vector tomography imaging method. They recovered a three-component Cartesian magnetization model but their approach is more applicable to determining the total magnetization direction of separated, homogeneous bodies. As previously stated, we are interested in more complicated scenarios. The method of Parker et al. (1987) attempts to construct a three-component magnetization model that is the closest to a uniform solution (i.e. uniform magnetization direction). They use a procedure called seminorm minimization and allow for an upper bound on the magnetization amplitude. We are interested in methods with more wide-spread applicability (i.e. allowing for more complicated causative magnetization).

3.2.1 Discretization of the TMV formulations

To invert for magnetization we follow the methodology of Li and Oldenburg (1996) for inversion for isotropic susceptibility. In the susceptibility inversion, the model region is split into an orthogonal 3D mesh of M rectangular prismatic cells, each with constant susceptibility. If one assumes no remanence, the magnetization in the j^{th} cell is in the direction of the Earth's field and is

$$\vec{J}_j = \chi_j \vec{H}_0. \quad (3.5)$$

The N data predicted by the model (i.e. the response of the model), $\mathbf{d}^{pred} = [d_1, \dots, d_N]^T$, are calculated as

$$\mathbf{d}^{pred} = \mathbf{G} \boldsymbol{\chi} \quad (3.6)$$

where $\boldsymbol{\chi} = [\chi_1, \dots, \chi_M]^T$ is the model vector containing the susceptibilities in each cell and \mathbf{G} is the N -by- M full sensitivity matrix.

The Cartesian formulation

In the TMVC formulation, the magnetization is split into one component parallel to the Earth's field and two components perpendicular to the Earth's field. We define three orthogonal directions \hat{p} , \hat{s} and \hat{t} with \hat{p} in the direction of the Earth's field; \hat{s} and \hat{t} can be chosen in any convenient manner. The Earth model vector, \mathbf{m} , contains the three components of magnetization in each cell:

$$\mathbf{m} = [\mathbf{p}, \mathbf{s}, \mathbf{t}]^T \quad (3.7)$$

where $\mathbf{p} = [p_1, \dots, p_M]^T$ (similarly for \mathbf{s} and \mathbf{t}) and \mathbf{m} has length $3M$. We deal with units of effective susceptibility instead of magnetization strength by dividing the three magnetization components by the strength (magnitude) of the Earth's field, H_0 :

$$\vec{p} = \vec{J}_p / H_0 \quad (3.8a)$$

$$\vec{s} = \vec{J}_s / H_0 \quad (3.8b)$$

$$\vec{t} = \vec{J}_t / H_0. \quad (3.8c)$$

The magnetization in the j^{th} cell is then written

$$\vec{J}_j = H_0 (p_j \hat{p} + s_j \hat{s} + t_j \hat{t}) \quad (3.9)$$

and the predicted data are calculated as

$$\mathbf{d}^{pred} = \mathbf{G}_p \mathbf{p} + \mathbf{G}_s \mathbf{s} + \mathbf{G}_t \mathbf{t}. \quad (3.10)$$

The sensitivity matrices in equation (3.10) are related to that in equation (3.6). We simplify equation (3.10) as

$$\mathbf{d}^{pred} = \mathbf{G}_{pst} \mathbf{m} \quad (3.11)$$

by defining

$$\mathbf{G}_{pst} = [\mathbf{G}_p, \mathbf{G}_s, \mathbf{G}_t]. \quad (3.12)$$

The spherical formulation

In the TMVS formulation, the magnetization is represented by an amplitude, denoted a , a dip angle θ (positive down) and an azimuthal angle φ (positive east of north). The model vector now contains the amplitude and two angles in each cell:

$$\mathbf{m} = [\mathbf{a}, \theta, \varphi]^T. \quad (3.13)$$

Again, we deal with the amplitude as an effective susceptibility (magnetization amplitude divided by Earth's field strength H_0). We define Cartesian axes with $+x$ northing, $+y$ easting and $+z$ down and let

$$\vec{u} = \vec{J}_x / H_0 \quad (3.14a)$$

$$\vec{v} = \vec{J}_y / H_0 \quad (3.14b)$$

$$\vec{w} = \vec{J}_z / H_0. \quad (3.14c)$$

The Cartesian magnetization components in the j^{th} cell are determined through the following trigonometric expressions:

$$u_j = a_j \cos(\theta_j) \cos(\varphi_j) \quad (3.15a)$$

$$v_j = a_j \cos(\theta_j) \sin(\varphi_j) \quad (3.15b)$$

$$w_j = a_j \sin(\theta_j). \quad (3.15c)$$

The predicted data are calculated as

$$\mathbf{d}^{pred} = \mathbf{G}_u \mathbf{u} + \mathbf{G}_v \mathbf{v} + \mathbf{G}_w \mathbf{w}. \quad (3.16)$$

Again, the sensitivity matrices in equation (3.16) are related to that in equation (3.6) and we simplify equation (3.16) as

$$\mathbf{d}^{pred} = \mathbf{G}_{uvw}\mathbf{k} \quad (3.17)$$

by defining

$$\mathbf{k} = [\mathbf{u}, \mathbf{v}, \mathbf{w}]^T \quad (3.18)$$

and

$$\mathbf{G}_{uvw} = [\mathbf{G}_u, \mathbf{G}_v, \mathbf{G}_w]. \quad (3.19)$$

3.2.2 An underdetermined optimization approach

Following Li and Oldenburg (1996) we formulate the inversion as an underdetermined problem with Tikhonov regularization. The objective function to be minimized is

$$\Phi = \frac{1}{2}\Phi_d + \frac{\beta}{2}\Phi_m \quad (3.20)$$

where Φ_d is the data misfit and Φ_m is a parameterized model objective function that provides flexibility to generate models with different characteristics; β is a regularization parameter that allows us to control the data fit.

The data misfit is

$$\Phi_d = \|\mathbf{W}_d(\mathbf{d}^{pred} - \mathbf{d}^{obs})\|^2 \quad (3.21)$$

where \mathbf{d}^{obs} is the observed data and the weighting matrix \mathbf{W}_d contains measurement uncertainties as in Li and Oldenburg (1996). If the statistics of the noise in the data is Gaussian then the data misfit in equation (3.21) is a chi-squared variable with expected value equal to the number of data. Therefore, in our synthetic inversions we search for a value of β such that the misfit equals the number of data.

The Cartesian formulation

To simplify the mathematics we absorb \mathbf{W}_d into the observed data and the sensitivity matrices in equation (3.10) such that

$$\Phi_d = \|\tilde{\mathbf{G}}_{pst}\mathbf{m} - \tilde{\mathbf{d}}^{obs}\|^2. \quad (3.22)$$

Our model objective function is

$$\begin{aligned}\Phi_m = & \|\mathbf{W}_p(\mathbf{p} - \mathbf{p}_{ref})\|^2 + \gamma \|\mathbf{W}_s(\mathbf{s} - \mathbf{s}_{ref})\|^2 \\ & + \gamma \|\mathbf{W}_t(\mathbf{t} - \mathbf{t}_{ref})\|^2.\end{aligned}\quad (3.23)$$

The regularization functionals \mathbf{W}_p , \mathbf{W}_s and \mathbf{W}_t in equation (3.23) are equivalent to \mathbf{W}_m in Li and Oldenburg (1996). They include the desire for model smoothness and the depth weighting required for the magnetic problem. Each term in equation (3.23) contains three smoothness terms, each of which involves a finite difference operator for a different Cartesian direction. Reference models \mathbf{p}_{ref} , \mathbf{s}_{ref} and \mathbf{t}_{ref} can be included if desired. The factor $\gamma \in (0, \infty)$ controls the amount of magnetization that is tolerated in a direction away from the Earth's field: high values of γ will result in magnetization parallel to the Earth's field and low values of γ will result in magnetization perpendicular to the Earth's field.

To solve the inverse problem we set the gradient of the objective function to zero and obtain an equation of the form $\mathbf{A}\mathbf{m} = \mathbf{b}$ to solve for the magnetization model \mathbf{m} :

$$\mathbf{A} = \tilde{\mathbf{G}}_{pst}^T \tilde{\mathbf{G}}_{pst} + \beta \begin{bmatrix} \mathbf{W}_p^T \mathbf{W}_p & 0 & 0 \\ 0 & \gamma \mathbf{W}_s^T \mathbf{W}_s & 0 \\ 0 & 0 & \gamma \mathbf{W}_t^T \mathbf{W}_t \end{bmatrix} \quad (3.24)$$

and

$$\mathbf{b} = \tilde{\mathbf{G}}_{pst}^T \tilde{\mathbf{d}}^{obs} + \beta \begin{bmatrix} \mathbf{W}_p^T \mathbf{W}_p \mathbf{p}_{ref} \\ \gamma \mathbf{W}_s^T \mathbf{W}_s \mathbf{s}_{ref} \\ \gamma \mathbf{W}_t^T \mathbf{W}_t \mathbf{t}_{ref} \end{bmatrix}. \quad (3.25)$$

The spherical formulation

Again, we absorb \mathbf{W}_d into the observed data and sensitivity matrices in equation (3.16) such that

$$\Phi_d = \|\tilde{\mathbf{G}}_{uvw} \mathbf{k} - \tilde{\mathbf{d}}^{obs}\|^2. \quad (3.26)$$

Our model objective function is

$$\begin{aligned}\Phi_m = & \|\mathbf{W}_a(\mathbf{a} - \mathbf{a}_{ref})\|^2 + \gamma \|\mathbf{W}_\theta(\boldsymbol{\theta} - \boldsymbol{\theta}_{ref})\|^2 \\ & + \gamma \|\mathbf{W}_\varphi(\boldsymbol{\varphi} - \boldsymbol{\varphi}_{ref})\|^2.\end{aligned}\quad (3.27)$$

Again, the regularization functionals in equation (3.27) are equivalent to \mathbf{W}_m in Li and Oldenburg (1996), with each term in equation (3.27) containing three smoothness terms, each of which involving a finite difference operator for a different Cartesian direction. We keep the parameter γ in the TMVS formulation to allow us to deal with scaling issues; the amplitude values will generally lie across a different range than the angles and we can balance the terms in equation (3.27) by altering γ .

The forward modelling procedure in equation (3.26) is now a nonlinear operation due to the trigonometric conversion required in equation (3.15). We use an iterative Newton-type descent method to solve the resulting nonlinear inverse problem: we solve an equation of the form

$$\mathbf{H}\delta\mathbf{m} = -\mathbf{g} \quad (3.28)$$

with \mathbf{g} the gradient (first model derivative of the objective function Φ), \mathbf{H} the Hessian (second model derivative) and $\delta\mathbf{m}$ a model perturbation. After solving for $\delta\mathbf{m}$ a line search is performed for an appropriate step length α (the expected value is close to 1.0) and the model is updated as

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha\delta\mathbf{m} \quad (3.29)$$

(where $\mathbf{m}^{(k)}$ is the current iterate and $\mathbf{m}^{(k+1)}$ the updated model).

The gradient, \mathbf{g}_m , and Hessian, \mathbf{H}_m , of the model objective function Φ_m are

$$\mathbf{g}_m = \begin{bmatrix} \mathbf{W}_a^T \mathbf{W}_a (\mathbf{a} - \mathbf{a}_{ref}) \\ \gamma \mathbf{W}_\theta^T \mathbf{W}_\theta (\boldsymbol{\theta} - \boldsymbol{\theta}_{ref}) \\ \gamma \mathbf{W}_\varphi^T \mathbf{W}_\varphi (\boldsymbol{\varphi} - \boldsymbol{\varphi}_{ref}) \end{bmatrix} \quad (3.30)$$

and

$$\mathbf{H}_m = \begin{bmatrix} \mathbf{W}_a^T \mathbf{W}_a & 0 & 0 \\ 0 & \gamma \mathbf{W}_\theta^T \mathbf{W}_\theta & 0 \\ 0 & 0 & \gamma \mathbf{W}_\varphi^T \mathbf{W}_\varphi \end{bmatrix}. \quad (3.31)$$

For the data misfit term, Φ_d , we use the chain and product rules for matrix-vector equations.

The chain rule gives

$$\frac{d\Phi_d}{d\mathbf{m}} = \frac{\partial\Phi_d}{\partial\mathbf{k}} \frac{d\mathbf{k}}{d\mathbf{m}} \quad (3.32)$$

and the gradient, \mathbf{g}_d , of the misfit term is then

$$\mathbf{g}_d = \mathbf{S}^T \tilde{\mathbf{G}}_{uvw}^T \delta\mathbf{d} \quad (3.33)$$

where

$$\mathbf{S} = \frac{d\mathbf{k}}{d\mathbf{m}} = \begin{bmatrix} \mathbf{S}_{ua} & \mathbf{S}_{u\theta} & \mathbf{S}_{u\varphi} \\ \mathbf{S}_{va} & \mathbf{S}_{v\theta} & \mathbf{S}_{v\varphi} \\ \mathbf{S}_{wa} & \mathbf{S}_{w\theta} & \mathbf{S}_{w\varphi} \end{bmatrix} \quad (3.34)$$

$$\mathbf{S}_{ua} = \mathbf{S}_{ua}^T = \text{diag}(\mathbf{s}_{ua}) = \frac{d\mathbf{u}}{d\mathbf{a}} \quad (3.35)$$

(with similar definitions to equation (3.35) for any missing quantities). The *diag* function takes a length n vector and places it along the main diagonal of an n -by- n matrix.

Let f be some m -by- n matrix (or vector) and g be p -by- q , both being functions of some quantity x . The product rule for matrix-vector equations is then

$$\frac{df(x)g(x)}{dx} = \left(g^T \otimes \mathbf{I}_m \right) \frac{df}{dx} + \left(\mathbf{I}_q \otimes f \right) \frac{dg}{dx} \quad (3.36)$$

where \otimes is the Kronecker product and \mathbf{I}_m is an m -by- m identity matrix. The Hessian, \mathbf{H}_d , of the misfit term is then

$$\begin{aligned} \mathbf{H}_d &= \frac{d}{d\mathbf{m}} (\mathbf{S}^T \tilde{\mathbf{G}}_{uvw}^T \delta\mathbf{d}) \\ &= \left[(\delta\mathbf{d}^T \tilde{\mathbf{G}}_{uvw}) \otimes \mathbf{I}_M \right] \frac{d\mathbf{S}^T}{d\mathbf{m}} + \mathbf{S}^T \tilde{\mathbf{G}}_{uvw}^T \frac{d\delta\mathbf{d}}{d\mathbf{m}} \end{aligned} \quad (3.37)$$

which, after some nontrivial algebra, works out to

$$\mathbf{H}_d = \mathbf{T}_u \mathbf{D}_u + \mathbf{T}_v \mathbf{D}_v + \mathbf{T}_w \mathbf{D}_w + \mathbf{S}^T \tilde{\mathbf{G}}_{uvw}^T \tilde{\mathbf{G}}_{uvw} \mathbf{S} \quad (3.38)$$

where

$$\mathbf{T}_u = \begin{bmatrix} \mathbf{T}_{u,aa} & \mathbf{T}_{u,a\theta} & \mathbf{T}_{u,a\varphi} \\ \mathbf{T}_{u,\theta a} & \mathbf{T}_{u,\theta\theta} & \mathbf{T}_{u,\theta\varphi} \\ \mathbf{T}_{u,\varphi a} & \mathbf{T}_{u,\varphi\theta} & \mathbf{T}_{u,\varphi\varphi} \end{bmatrix} \quad (3.39)$$

$$\mathbf{T}_{u,a\theta} = \text{diag}(\mathbf{t}_{u,a\theta}) = \frac{d\mathbf{s}_{ua}}{d\theta} \quad (3.40)$$

$$\mathbf{D}_u = \text{diag} \left(\left[\tilde{\mathbf{G}}_u^T \delta\mathbf{d}, \tilde{\mathbf{G}}_u^T \delta\mathbf{d}, \tilde{\mathbf{G}}_u^T \delta\mathbf{d} \right]^T \right) \quad (3.41)$$

$$\delta\mathbf{d} = \tilde{\mathbf{G}}_{uvw} \mathbf{k} - \tilde{\mathbf{d}}^{obs} \quad (3.42)$$

with similar definitions for any missing quantities. We do not include the derivation of the expressions in equations (3.33) and (3.38) as they are quite involved (they are provided in

Appendix B of this thesis). The individual elements of the \mathbf{S} and \mathbf{T} quantities are simply first and second derivatives of the expressions in equation (3.15) with respect to a , θ and φ . Note that we are performing full Newton steps here. A Gauss-Newton approach would neglect the three $\mathbf{T}_\sim \mathbf{D}_\sim$ terms in equation (3.38) but investigation has shown that those terms give an important contribution to the Hessian and cannot be neglected.

3.2.3 Inversion functionality and important practical considerations

The susceptibility inversion of Li and Oldenburg (1996); Li and Oldenburg (2003) contains a high degree of nonuniqueness. To deal with this, Li and Oldenburg (2003) added depth weighting and allowed for further geologic information to be incorporated through weighting functions and physical property bounds, the latter enabling a positivity constraint on the susceptibility to maintain physical reality. In our magnetization inversions there are now three times as many model parameters and it is likely that additional information is required to obtain acceptable solutions. Therefore, we provide the same weighting and bounding functionality although, instead of following their logarithmic barrier approach for bound-constraints, we employ a more robust and efficient gradient-projection method (the gradient-projection-reduced-Newton (GPRN) method of Vogel (2002)).

In the TMVC (Cartesian) formulation, we can define a lower bound, p_L , that determines the value above which the p component of the magnetization must lie. If one is sure that no remanence exists then γ can be set to some high value and p_L set to zero such that all the recovered magnetizations must be in the direction of (i.e. parallel to) the Earth's field. In the limit $\gamma \rightarrow \infty$, such an inversion is identical to a susceptibility inversion assuming no remanence. If some remanence is expected then one can allow increasing amounts by decreasing γ towards zero and p_L below zero. With no prior information about remanence, γ should be set to unity and the p component not bounded.

An appropriate value for γ in the TMVS (spherical) formulation is difficult to estimate for any particular application. Hence, we suggest running several inversions with different values of γ . When γ is below a certain threshold the smoothness regularization on the angle terms is essentially removed and the results exhibit very scattered magnetization directions. Above a certain threshold the smoothness regularization on the angle terms becomes the primary focus and all magnetizations are in the same direction.

In the TMVS (spherical) formulation, depth weighting is only applied to the amplitude term in

the model objective function. Care must be taken to include smallness regularization on the angles only when reference models are provided. The TMVS objective function is nonconvex, resulting from the trigonometric identities required in the forward modelling. Hence, there is a chance that the objective function in the spherical formulation will suffer from negative curvatures and multiple minima. Therefore, our inversion algorithm contains checks for negative curvature, ensuring that all step directions $\delta\mathbf{m}$ are descent directions. If $\delta\mathbf{m}$ is not a descent direction then we multiply $\delta\mathbf{m}$ by -1.0 and perform a more careful line search in equation (3.29) (because now the expected value of $\alpha = 1.0$ is no longer appropriate). The following section demonstrates the existence of multiple minima in a simple example. A strategy for ameliorating this problem is also mentioned below.

3.2.4 A simple synthetic test

In mineral exploration applications the remanence can be significant and the subsurface magnetization complicated; there may be different Earth regions containing quite different remanence. In this section and sections that follow we investigate some illustrative synthetic problems to improve our understanding of our TMV inversion methods prior to applying them to real data. Here, the data in Figure 3.2(c) are inverted after adding a small amount of random noise (taken from normal distributions with zero mean and standard deviations equal to 2.0% of the absolute data values plus 1.0nT).

Results for the Cartesian formulation

Without any prior assumptions on the direction of remanence it is most appropriate to set $\gamma = 1.0$, and not enforce bounds on the magnetization or include reference models (here, resulting in the model in Figure 3.3(b)). From a target picking perspective, there is a dramatic improvement compared to the susceptibility inversion results (compare to Figure 3.3(a)).

Performing a TMVC inversion with no positivity on the p component is similar to performing a susceptibility inversion with no positivity (identical in the limit $\gamma \rightarrow \infty$). Our TMVC results mirror the well known effect often seen in susceptibility inversions: magnetic material is pushed deeper when no positivity is enforced. Hence, we suggest reducing the power of the depth weighting when performing TMVC inversions.

Results for the spherical formulation

Results for the spherical formulation inversion are shown in Figures 3.3(c) and 3.3(d). For the inversion in Figure 3.3(c) the initial model was a half-space with $a = 0.001$, $\theta = 90^\circ$ and $\varphi = 90^\circ$; i.e. the initial model specified a magnetization orientation in the direction of the Earth's field (which is not the true direction of the magnetization), and no reference models were included. We set $\gamma = 1.0$ such that the θ and φ terms dominated the model objective function resulting in a constant magnetization direction. The recovered model in Figure 3.3(c) shows further improvement over the Cartesian formulation results.

For the inversion in Figure 3.3(d), the initial halfspace model was altered to have $\theta = -45^\circ$. Again, no reference models were included. The inversion had trouble fitting the data (the model does not fit the data to the desired degree) and the result indicates the existence of multiple minima in the TMVS formulation. For this synthetic example, the results recovered with the TMVS formulation were essentially identical to that in Figure 3.3(c) for any initial orientations within 90° of the true magnetization.

3.3 Reducing the nonuniqueness in TMV inversions

Inverting for a vector magnetization model rather than scalar susceptibility greatly increases the nonuniqueness of the problem but inclusion of prior information can improve the results. Morris et al. (2007) provide a discussion of what types of remanence information could realistically be available for constraining a magnetic inversion. Below we discuss how that information could be incorporated into our TMV inversions.

3.3.1 Total magnetization known exactly

Obtaining thorough and reliable measurements of the remanence direction on oriented samples is not a common practice due to expense and difficulty. Morris et al. (2007) discuss methods for overcoming the difficulties involved. If oriented rock samples are available then χ (susceptibility), Q (Königsberger ratio) and the remanence direction can be measured. These can be combined with knowledge of the Earth's field to determine both the amplitude and direction of the total magnetization at the sample locations. This information can be easily incorporated into either

of our TMV inversions through the reference models and associated weights or through bounds on the model parameters.

3.3.2 Measurements of χ and Q

Obtaining measurements of χ and Q from rock samples is a fairly common practice. Assume that sampled measurements of χ and Q are available but the direction of remanence is not. To incorporate this information into the TMVC (Cartesian) magnetization inversion we start by writing the magnitudes of the induced, remanent and total magnetization as

$$\|\vec{J}_{ind}\| = \chi \|\vec{H}_0\| = \chi H_0 \quad (3.43a)$$

$$\|\vec{J}_{rem}\| = Q \|\vec{J}_{ind}\| = Q \chi H_0 \quad (3.43b)$$

$$\|\vec{J}_{tot}\| = \|\vec{J}_{ind} + \vec{J}_{rem}\|. \quad (3.43c)$$

The maximum magnitude of the total magnetization occurs when the remanent and induced components are parallel, and the minimum occurs when those components are antiparallel (opposing). This reasoning provides the following inequality statements:

$$|1 - Q| \chi H_0 \leq \|\vec{J}_{tot}\| \leq (1 + Q) \chi H_0. \quad (3.44)$$

In the TMVC formulation, the total magnetization is

$$\vec{J}_{tot} = H_0 (\vec{p} + \vec{s} + \vec{t}). \quad (3.45)$$

Hence, a constraint of the form

$$(1 - Q)^2 \chi^2 \leq p^2 + s^2 + t^2 \leq (1 + Q)^2 \chi^2 \quad (3.46)$$

can be included in the inverse problem. This is a fairly complicated nonlinear constraint and the minimization may be adversely affected. Much simpler bound constraints that follow similar reasoning are

$$(1 - Q)\chi \leq p \leq (1 + Q)\chi \quad (3.47a)$$

$$-Q\chi \leq s \leq Q\chi \quad (3.47b)$$

$$-Q\chi \leq t \leq Q\chi \quad (3.47c)$$

but these constraints provide more flexibility (are less constraining) than those in equation (3.46).

In contrast to the nonlinear constraints in equation (3.46), incorporating χ and Q information into the TMVS (spherical) magnetization inversion yields bounds on the magnetization amplitude:

$$|1 - Q|\chi \leq a \leq (1 + Q)\chi. \quad (3.48)$$

The method of Parker et al. (1987) also allows for an upper bound on the magnetization amplitude.

3.3.3 Known or assumed remanence direction

For most exploration drill-holes the core samples are only oriented well enough to determine the magnetic inclination, θ , but not the declination, φ . Morris et al. (2007) mention methods for approximating the declination using available geologic knowledge.

For the TMVC (Cartesian) formulation, if the direction of remanence can be assigned exactly then a linear equality constraint of the form

$$\mathbf{s} + \mathbf{At} + \mathbf{b} = 0 \quad (3.49)$$

can be created, which results from the requirement that

$$s_j \hat{s} + t_j \hat{t} = \vec{J}_{rem,\perp} \quad (3.50)$$

for all $j = 1 \dots M$ where $\vec{J}_{rem,\perp}$ is the component of the remanence perpendicular to the Earth's field. When the remanence direction can only be approximated, this information would result in inequality constraints related to the equality constraint in equation (3.49).

Sampled measurements of θ and/or φ can be incorporated into the TMVS (spherical) formulation through trivial use of the angle reference models or bounds. The direction of total magnetization lies somewhere between the remanence direction and the Earth's field, leading to bounds on θ and φ in the TMVS formulation.

3.3.4 Comparison of constraints in the Cartesian and spherical formulations

From the above discussion it is clear that the geologic information available can be incorporated into the spherical formulation much more easily than the Cartesian formulation. Assume that some combination of the true susceptibility, Königsberger ratio, the remanence direction and the total magnetization direction are known in the cells containing the true anomalous block in Figure 3.1. Calculating equations (3.46) and (3.47) for those cells shows that the recovered model in Figure 3.3(b) already satisfies those constraint equations. Hence, it is apparent that those constraints may not provide much utility in the TMV inversions and the full magnetization direction may be required.

3.4 A complicated synthetic scenario

We now move on to a more complicated 3D synthetic scenario that could not be approached with methods that assume a uniform magnetization direction. The mesh and model are shown in Figure 3.4. The Earth's field is oriented towards the north, inclined 30° above the horizontal, as indicated in Figure 3.4(b). The remanent magnetization is different for each of the two vertical tabular bodies, as indicated in Figure 3.4(c): the body to the west has a horizontal remanence directed 30° west of north; the body to the east has a horizontal remanence directed 30° east of north. The susceptibility of the bodies is 0.1 (SI units; within a zero background) and the Königsberger ratio is set to 3.0 for both bodies. This results in an effective susceptibility, χ_{eff} , of 0.38 for each body.

Figure 3.5 compares the different contributions to the TMI magnetic response of this model. The data in Figure 3.5(c) were first inverted for susceptibility assuming no remanence exists: we do not show the result but as expected, the inversion was not able to fit the data to the desired degree and the recovered model showed little resemblance to the true model. Figure 3.6(a) shows the result of the method of Shearer and Li (2004), which recovers a more acceptable central body but there is still significant magnetic material placed towards the mesh boundaries.

As previously mentioned, we suggest reducing the power of the depth weighting when performing TMVC inversions. Figures 3.6(b) and 3.7(a) show the result of a TMVC inversion with $\gamma = 1.0$, no reference models included and the depth weighting power reduced from 3.0 to 2.0. The TMVC result is comparable in shape to that using the method of Shearer and Li (2004). The TMVC inversion recovered lower values because overall it placed the magnetic material a few

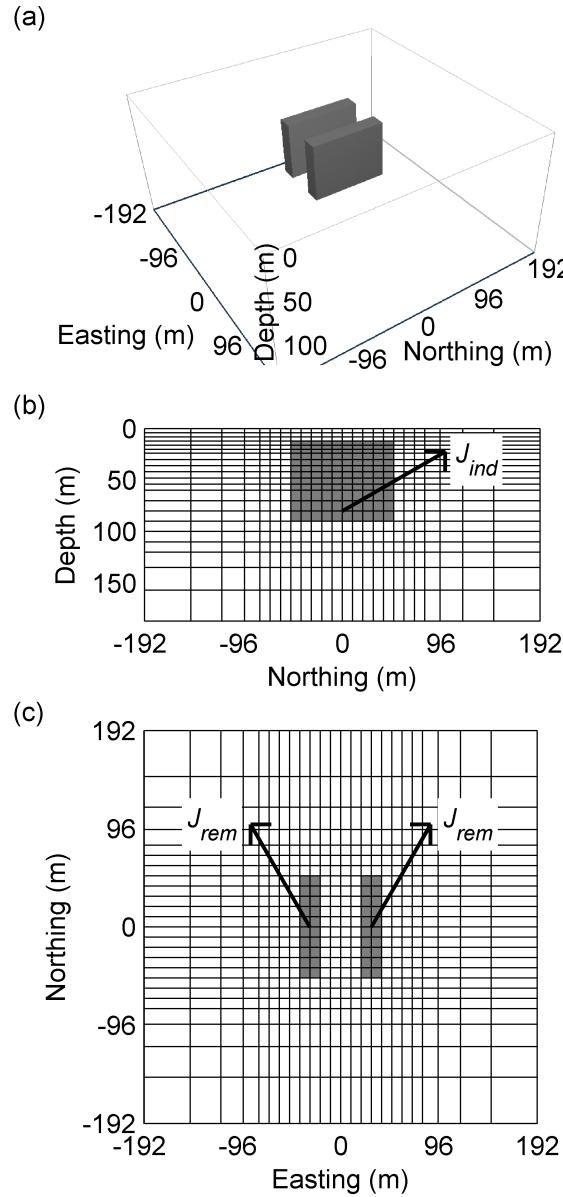


Figure 3.4: Views of the synthetic two body model: (a) shows a 3D perspective view, (b) shows a vertical cross-section at easting = 30m and (c) shows a horizontal cross-section at depth = 30m. Mesh cell spacing is identical in the easting and northing directions. The induced component of the magnetization is indicated in (b) and the remanent component in (c).

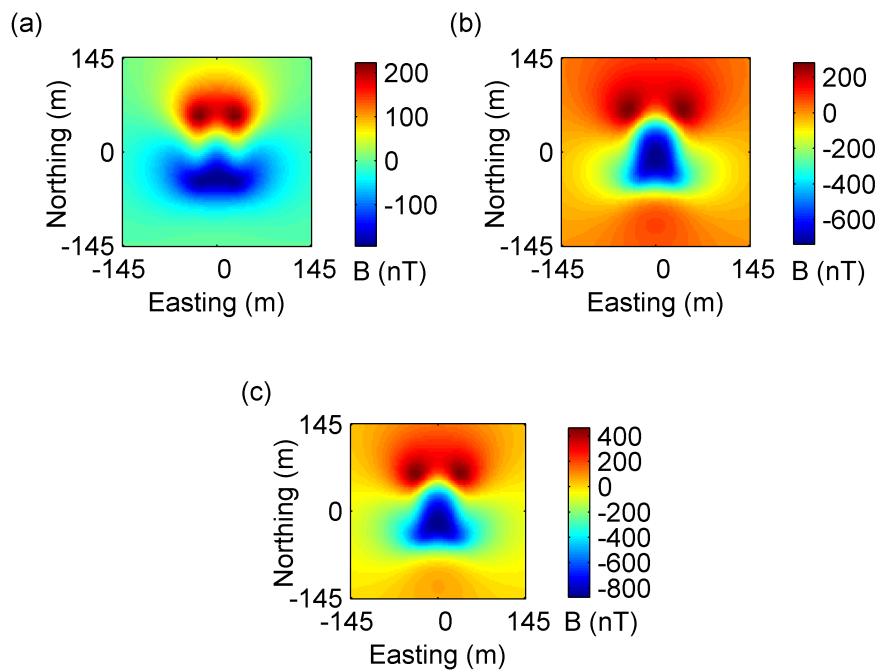


Figure 3.5: Map views of the (a) induced and (b) remanent components of the TMI response, and (c) total combined TMI response 10m above the mesh in Figure 3.4.

cells closer to the surface (a result of the altered depth weighting).

We now perform a TMVS inversion using the knowledge gained from the TMVC result to constrain the TMVS inversion such that it does not approach inappropriate local minima. First, the TMVC result allows us to reduce the size of the active mesh region (i.e. removing padding cells) to reduce the nonuniqueness of the problem. Second, the TMVC result indicates a split in the azimuth of the magnetizations: the highest magnetizations to the west have negative azimuths and those to the east have positive azimuths; this allows us to place bounds on φ in the TMVS inversion, with $-90 \leq \varphi \leq 0$ in the west and $0 \leq \varphi \leq 90$ in the east. No reference models were included. This TMVS inversion was able to fit the data and the result, seen in Figures 3.6(c) and 3.7(b), shows improvement over the TMVC result, now better indicating the presence of two vertical tabular bodies.

3.5 A real data example with multiple bodies and complicated remanence

Here we invert real survey data over a region containing multiple bodies with complicated remanence. The data, shown in Figure 3.8, have been base-station corrected only. Therefore, we design the inversion mesh to have padding cells extending beyond the data region to allow for any regional component that may exist in the data. However, the magnetization inversions place little to no magnetic material in the padding cells so we do not show those cells in the results. The region to the south contains several thin remanently magnetized units near the surface that have been tilted and faulted and, consequently, have remanent magnetizations in many different directions. This is evident in Figure 3.8 from the scattered small scale anomalies near the south. A deeper intrusive body lies further to the north of the remanently magnetized surface units, represented by a longer scale feature in the data.

Knowledge of the field survey characteristics lead us to assign uncertainties to the data that are composed of 2% of the absolute data values plus a floor of 50nT (which equals approximately 2% of the total data range). Unless otherwise mentioned, all inversions yield a chi-squared misfit value approximately equal to (within a few percent of) the number of data (= 7509). We found that the magnetization inversions were able to fit the data and lead to results with a reasonable level of data fit and a reasonable amount of structure. This indicates that the assigned uncertainties were appropriate. However, inversions for susceptibility have trouble fitting the data and the recovered models have geologically unrealistic features and unrealistically high susceptibility

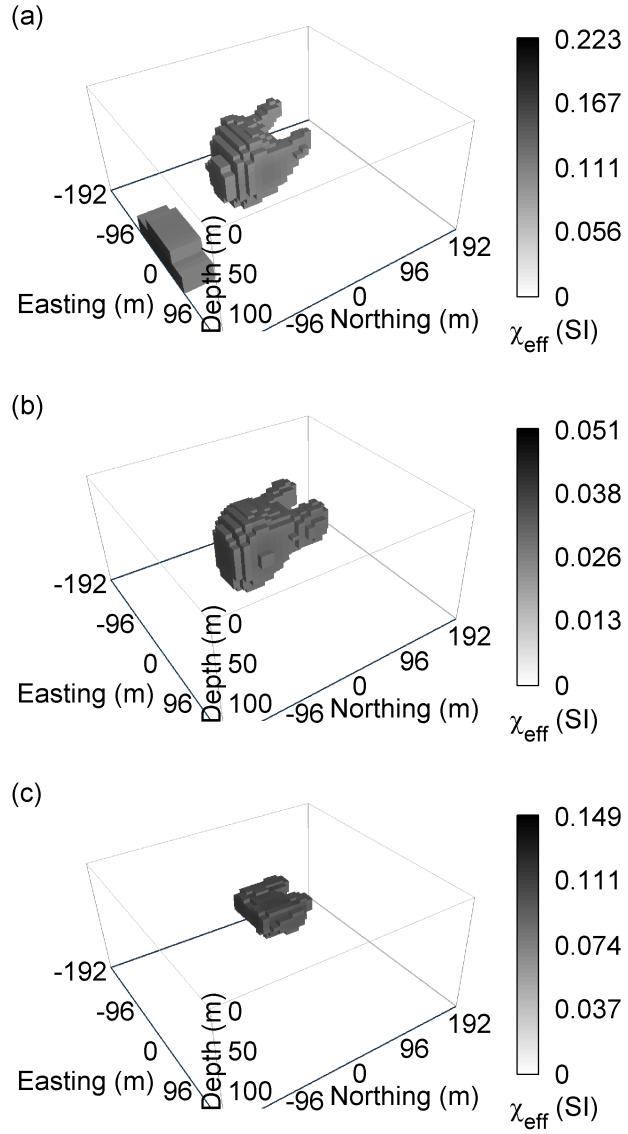


Figure 3.6: 3D perspective views of models recovered from inversion of the data in Figure 3.5(c). Plot (a) shows a $\chi_{eff} = 0.10$ isosurface of the model from a magnetization amplitude inversion (i.e. the method of Shearer and Li (2004)). Plot (b) shows a $\chi_{eff} = 0.032$ isosurface of the model from a TMVC inversion with $\gamma = 1.0$, no reference models included and a depth weighting power of 2.0. Plot (c) shows a $\chi_{eff} = 0.11$ isosurface of the model from a TMVS inversion with $\gamma = 1.0 \times 10^{-3}$, no reference models included and bounds applied.

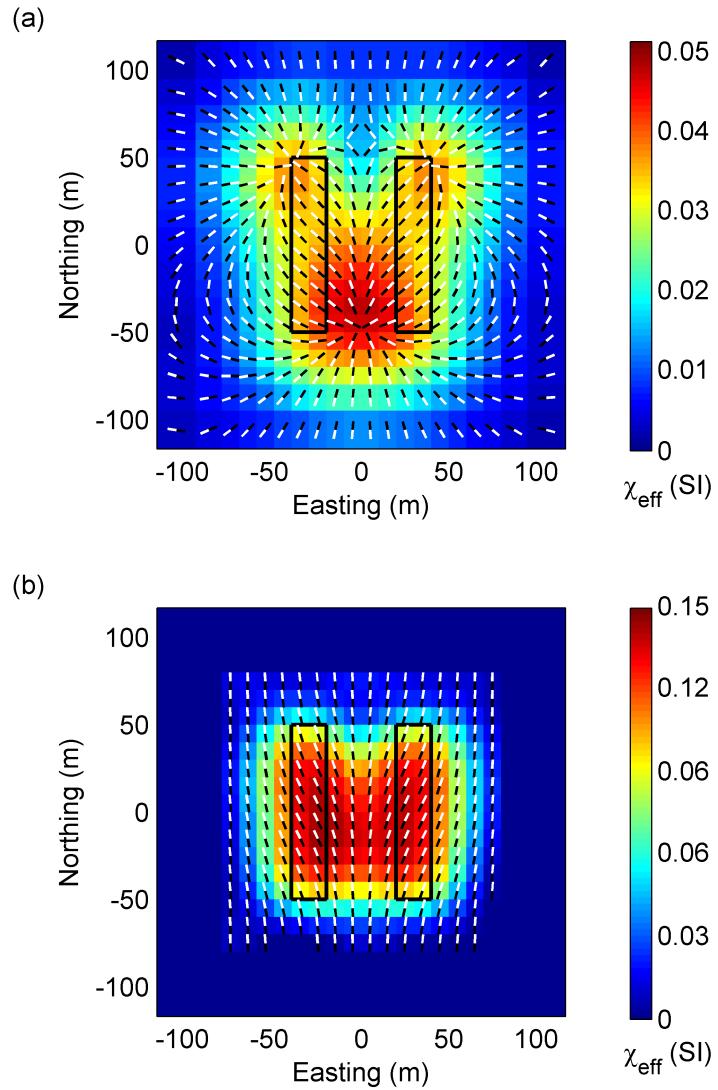


Figure 3.7: Plots (a) and (b) show horizontal cross-sections at depth = 30m through the models in Figures 3.6(b) and 3.6(c) respectively. Total magnetization vectors are indicated by black and white lines with white at the head. Vectors with zero amplitude are not displayed. The black rectangles indicate the positions of the blocks in the true model. Note that some padding cells have been removed in these plots.

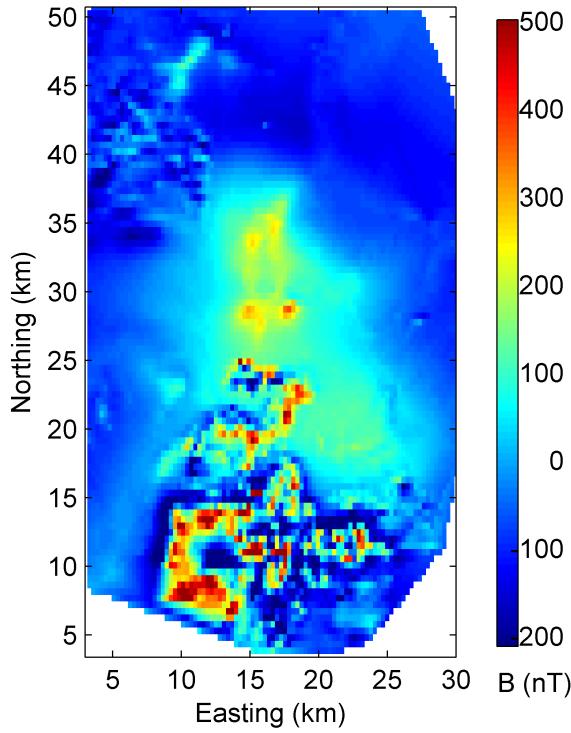


Figure 3.8: A map of the real TMI survey data. The data ranges from 726nT to 1062nT but we have altered the colour scale here to better represent the larger scale data features.

values. The conclusion is then that there is significant complicated remanence (high amplitude and many directions) in the subsurface that is affecting the susceptibility inversion results and inverting for susceptibility, and thereby assuming all magnetization in the region is parallel to the Earth's field, is not appropriate. We must then move to a magnetization inversion.

In the magnetization inversion we use the Cartesian formulation (TMVC). We set $\gamma = 1.0$ to allow magnetization in any direction, and we weight the smallness term to balance the values of the smoothness and smallness terms for the recovered models. No reference models are included. The result is shown in Figure 3.9: Figure 3.9(a) shows the p component recovered and Figure 3.9(b) shows the amplitude of the remaining magnetization, which we denote the st component ($= \sqrt{s^2 + t^2}$). These results clearly indicate significant remanence near the surface towards the south (where the st component is high) and a large susceptible body with little to no remanence at depth towards the north (where the st component is low and the p component is high). This is consistent with knowledge of the geology.

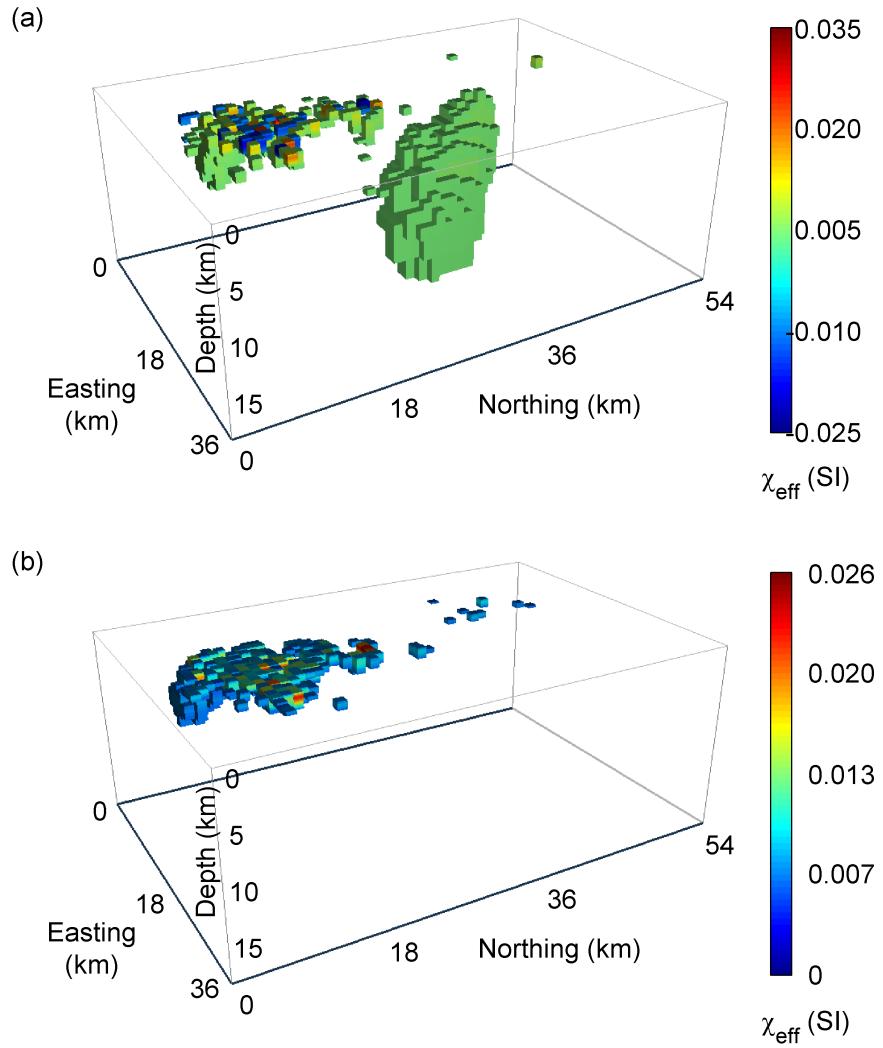


Figure 3.9: 3D perspective views from the southeast of the model recovered from a TMVC inversion of real survey data with $\gamma = 1.0$. Plot (a) shows $\chi_{\text{eff}} = -0.009$ and $\chi_{\text{eff}} = +0.005$ isosurfaces of the p component. Plot (b) shows a $\chi_{\text{eff}} = 0.006$ isosurface of the st component model ($= \sqrt{s^2 + t^2}$).

3.6 Application to regions containing high susceptibility

Before concluding, we mention another important use for our methods. Equation (3.2) is a simplification of the following:

$$\vec{J}_{ind} = \chi \vec{H} = \chi(\vec{H}_0 + \vec{H}_s), \quad (3.51)$$

where the secondary field \vec{H}_s includes any anomalous fields associated with magnetic material in the region. It is often assumed that the magnetic signal of any anomalous material is negligible compared to the Earth's field such that equation (3.51) reduces to equation (3.2). The induced magnetization is then parallel to the Earth's field. In most situations this assumption is acceptable, but when susceptibilities are high (e.g. above 0.1, SI units), this assumption can break down: the direction of the magnetic field (\vec{H}), and therefore the direction of the induced magnetization (\vec{J}_{ind}), at some location can be significantly affected by induced magnetization in the vicinity. This phenomenon is often referred to as demagnetization.

Hence, there are two situations for which typical magnetic inversion algorithms (i.e. those assuming that magnetization is parallel to the Earth's field) are not applicable: 1) in the presence of remanence, and 2) in the presence of high susceptibilities. Lelièvre and Oldenburg (2006) developed an algorithm for inverting magnetic data collected above regions of high magnetic susceptibility. However, that algorithm is not applicable when remanent magnetization is present. Our magnetization inversion methods are applicable to both situations.

Here we apply our methods to a synthetic example similar to one presented in Lelièvre and Oldenburg (2006). Although this example is on a small scale, it is applicable to larger scale mineral exploration problems where high susceptibilities can cause demagnetization effects (e.g. banded iron formations). The true model contains a highly susceptible ($\chi = 10.0$) prolate magnetic spheroid with a minor semi-axis of 3.0m and an eccentricity (long-to-short axis ratio) of 3.0. The situation is pictured in Figure 3.10(a). The long axis is horizontal and oriented in a west-east direction. The Earth's field is also towards the east and it dips 58.3° below horizontal. Demagnetization effects cause the spheroid's magnetization direction to rotate away from the Earth's field and towards the horizontal spheroid long axis, creating a magnetization that is inclined 31.7° below horizontal. Data are computed 8.0m above the centre of the spheroid (i.e. at an elevation of 4.0m) using an analytic method that takes demagnetization effects into account. A moderate amount of noise is added to this data and the resulting synthetic observed data are shown in Figure 3.11(a).

The unconstrained TMVC inversion result with $\gamma = 1.0$ and the depth weighting power lowered to 2.0 is shown in Figure 3.10(b). The unconstrained TMVS inversion result with $\gamma = 100$ and the Earth's field direction used in the initial model is shown in Figure 3.10(c). Similar to the results in Figure 3.3, where the true model was essentially a dipole source, the TMVC result generates a distribution of magnetization that is only slightly indicative of the true scenario. The TMVS inversion is successful in recovering a spheroidal distribution of high magnetization, with both the spheroid and the magnetization oriented correctly: the recovered magnetization declinations lie on $[89.4, 90.3]$ (degrees) and the recovered magnetization dips lie on $[33.2, 34.7]$, which are close to the true values of 90.0 and 31.7 respectively. Due to demagnetization effects, the effective susceptibility for the spheroid is lowered to approximately 3.0. The recovered values for the TMVS result are therefore reasonable, taking the smoothing of the recovered model into account. This example indicates that our magnetization inversion methods can provide reasonable results for scenarios involving demagnetization effects.

3.7 Conclusion

Inverting magnetic data for all three vector components of the subsurface magnetization is one possible route to interpreting magnetic data complicated by remanence. We have developed two three-component magnetization inversion routines (in Cartesian and spherical frameworks) that are appropriate for use on data that contain the response of material containing complicated magnetization (i.e. magnetization amplitude and direction varying throughout the region). The problem of remanence in magnetic inversions can now be dealt with through use of these two routines within an appropriate workflow, as indicated throughout this paper and discussed below.

Allowing a vector magnetization greatly increases the nonuniqueness of the magnetic inverse problem. We stress that the recovery of total magnetization is possible but it relies on the incorporation of reliable geologic information regarding remanence. Specifically, point measurements of the magnetization magnitude and direction obtained from oriented drill-core samples may be required to obtain acceptable magnetization models for more complicated scenarios. Therefore, we have designed our methods to allow incorporation of the types of information commonly collected.

The TMVC (Cartesian formulation) inversions tend to do a good job in the absence of additional information. In contrast, the TMVS (spherical formulation) inversions can perform poorly without an appropriate amount of constraints placed in them, but with those constraints the

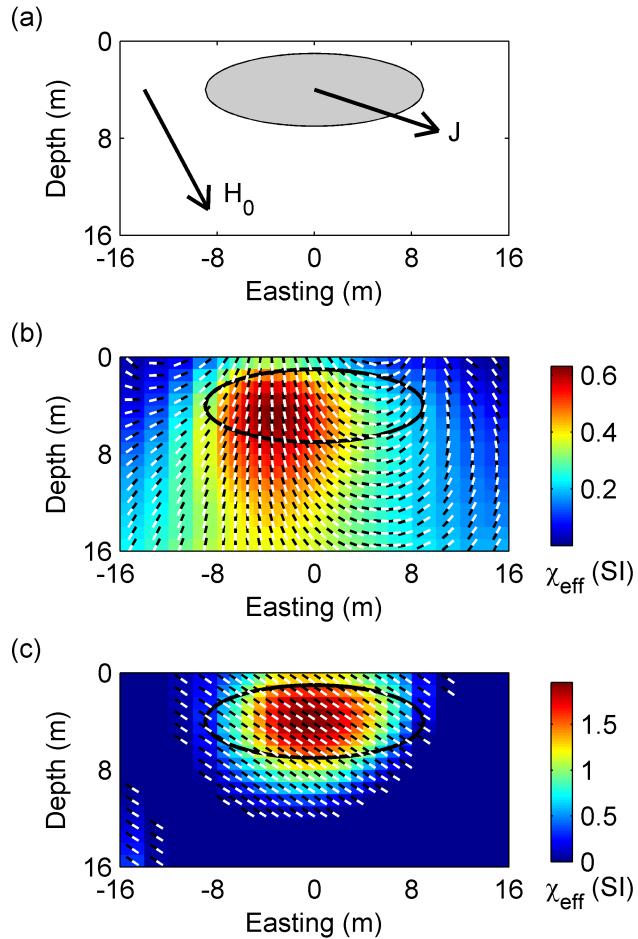


Figure 3.10: Vertical cross-sections at northing = 0m through 3D models for the high susceptibility example. The inversion mesh is identical in the northing and easting directions. This cross-section runs along the long axis of the true spheroid body, shown schematically in (a). The spheroid has a susceptibility of 10 (SI units). The TMVC inversion result with $\gamma = 1.0$ and a depth weighting power of 2.0 is shown in (b). The TMVS inversion result with $\gamma = 100$ and the Earth's field direction used in the initial model is shown in (c). The Earth's field is indicated in (a), labelled H_0 , and the resulting magnetization for the spheroid is labelled J . In (b) and (c), magnetization models are shown with the amplitude in units of effective susceptibility indicated on the colour bars. Total magnetization vector directions are indicated by black and white lines with white at the head. Vectors with zero amplitude are not displayed. The position of the spheroid is indicated in (b) and (c) with a black outline.

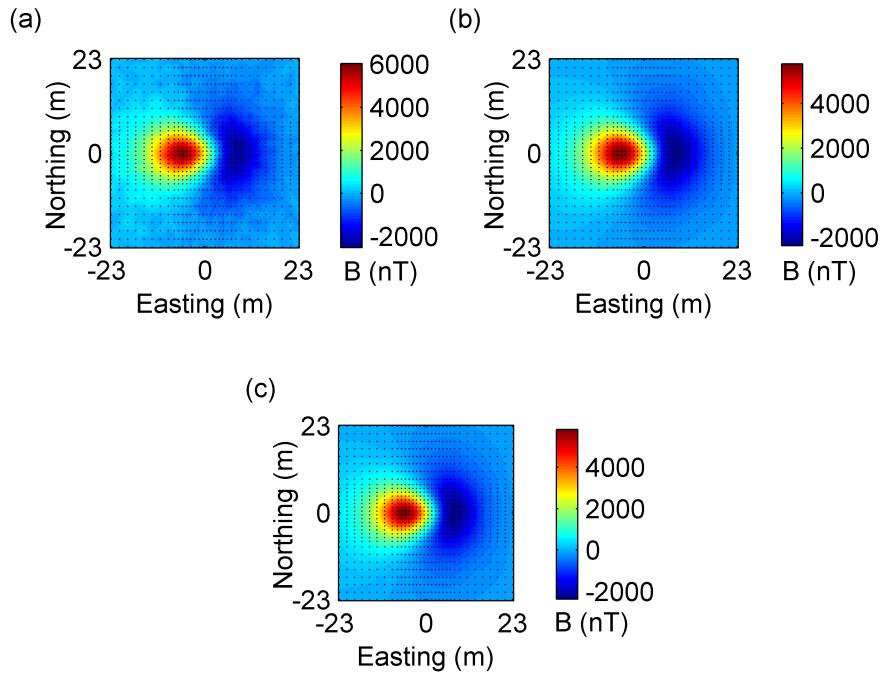


Figure 3.11: Total-field data maps calculated 4.0m above the surface for the spheroid model in Figure 3.10(a): (a) shows the observed data (forward modelled data with noise added); (b) shows the predicted data for the model in Figure 3.10(b); (c) shows the predicted data for the model in Figure 3.10(c). The locations of the data are indicated by black dots.

TMVS inversions can outperform the TMVC inversions. The information regarding remanence is more easily incorporated into the TMVS formulation. For these reasons, we suggest initial use of the TMVC formulation, followed by constrained TMVS inversions, as indicated in the following workflow:

1. invert with TMVC, including any geologic information through loose bound constraints;
2. visualize the resulting magnetization models (e.g. amplitude, inclination, declination, p and st components);
3. combine the TMVC result with geologic information to apply tighter bounds on a subsequent TMVS inversion.

With either routine, additional information incorporated as constraints can greatly improve the results, but due to the presence of multiple minima in the TMVS formulation it is vital to set bounds based on the TMVC result so that the TMVS inversion does not approach inappropriate local minima.

It is also important to note that our inversions recover the total magnetization, equal to the induced magnetization plus any existing remanence. The s and t components of the total magnetization are perpendicular to the inducing Earth's field and are, therefore, associated only with the remanence. However, it is not possible to disentangle the induced magnetization and the component of the remanence in the same direction, both of which are combined in the p component of the total magnetization. The process of viscous remanent magnetization can cause a remanent component in the direction of the current Earth's field. In that case, inversion methods assuming no remanence will behave well but return higher susceptibility values.

3.8 References

- Bossavit, A., 1998, Computational electromagnetism: Academic Press.
- Dannemiller, N., and Y. Li, 2006, A new method for determination of magnetization direction: *Geophysics*, **71**, no. 6, L69–L73.
- Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure, *Geophysical Journal International*: **134**, 213–227.
- Foss, C., 2006, Improvements in source resolution that can be expected from inversion of magnetic field tensor data: *The Leading Edge*, **25**, 81–84.
- Haney, M., and Y. Li, 2002, Total magnetization direction and dip from multiscale edges: 72nd International Annual Meeting, SEG, Expanded Abstracts, 735–738.
- Helbig, K., 1963, Some integrals of magnetic anomalies and their relation to the parameters of the disturbing body: *Zeitschrift für Geophysik*, **29**, 83–96.
- Lelièvre, P. G., and D. W. Oldenburg, 2006, Magnetic forward modelling and inversion for high susceptibility: *Geophysical Journal International*, **166**, 76–90.
- Li, Y., and D. W. Oldenburg, 1996, 3D inversion of magnetic data: *Geophysics*, **61**, 394–408.
- Li, Y., and D. W. Oldenburg, 2003, Fast inversion of large-scale magnetic data using wavelet transforms and logarithmic barrier method: *Geophysical Journal International*, **152**, 251–265.
- Li, Y., S. Shearer, M. Haney, and N. Dannemiller, 2004, Comprehensive approaches to the inversion of magnetic data with strong remanent magnetization: 74th Annual International Meeting, SEG, Expanded Abstracts, 1191–1194.
- Morris, B., H. Ugalde, and V. Thomson, 2007, Magnetic remanence constraints on magnetic inversion models: *The Leading Edge*, **26**, 960–964.
- Parker, R. L., L. Shure, and J. A. Hildebrand, 1987, The application of inverse theory to seamount magnetism: *Reviews of Geophysics*, **25**, 17–40.
- Phillips, J. D., 2005, Can we estimate total magnetization directions from aeromagnetic data using Helbig's integrals?: *Earth Planets Space*, **57**, 681–689.
- Schmidt, P. W., and D. A. Clark, 2006, The magnetic gradient tensor; its properties and uses in source characterization: *The Leading Edge*, **25**, 75–78.

Shearer, S., and Y. Li, 2004, 3D inversion of magnetic total gradient data in the presence of remanent magnetization: 74th International Annual Meeting, SEG, Expanded Abstracts, 774–777.

Vogel, C. R., 2002, Computational methods for inverse problems: SIAM.

Wang, M., Q. Di, K. Xu, and R. Wang, 2004, Magnetization vector inversion equations and forward [*sic*] and inverted [*sic*] 2-d model study: Chinese Journal of Geophysics, **47**, 601–609.

Chapter 4

A comprehensive study of including structural orientation information in geophysical inversions³

4.1 Introduction

To be reliable, Earth models used for mineral exploration should be consistent with all available geophysical and geologic information. Due to data uncertainty and other aspects inherent to the underdetermined geophysical inverse problem, there are an infinite number of models that can fit the geophysical data to the desired degree (i.e. the problem is nonunique). Further information is essential for a unique solution. Incorporating prior geologic knowledge can reduce ambiguity and enhance inversion results, leading to more reliable Earth models.

Structural orientation is an important form of available geologic information. This can involve the orientation of a body (i.e. the strike, dip and tilt of its major axes), aspect ratios (i.e. the relative lengths of a body's major axes) and physical property trends (i.e. increase, decrease or constant in a particular direction). The ability to specify such information becomes especially important for survey methods with limited depth resolution. The lack of resolution can lead to recovery of an object with an incorrect or distorted dip and by including orientation information the results can be dramatically improved at depth.

Many researchers have provided functionality for incorporating different types of geologic information into their particular inversion frameworks. In this paper we investigate how orientation information can be placed into our deterministic inversion framework in which a computation-

³A version of this chapter is in press: Lelièvre, P. G., and D. W. Oldenburg, A comprehensive study of including structural orientation and trend information in geophysical inversions: *Geophysical Journal International*.

ally well-behaved function is minimized subject to optional constraints. Before introducing our methods we provide an overview of some techniques used by other authors for comparison.

Bosch et al. (2001) and Guillen et al. (2008) work in a stochastic inversion framework that directly recovers rock type (i.e. a lithologic inversion) from a list of those assumed present. Prior information is placed into the problem through probability density functions and topology rules (relationships between rock units). The model space (i.e. all possible models) is investigated (sampled) through a random walk process, an approach proposed by Mosegaard and Tarantola (2002). This strategy provides not only model estimates but also statistical information regarding the model space. In contrast to the functions in our deterministic framework, their probability density functions and structural topology measures are not required to be differentiable and, hence, there is more flexibility in the types of geologic information that can be incorporated. However, their approach relies on random sampling methods that lead to much heavier computational costs than deterministic approaches.

Chasserau and Chouteau (2003) introduce regularization through a parameter covariance matrix, the elements of which are estimated statistically using variograms. The covariance matrix can be estimated from physical property data (e.g. measurements taken at surface or down drill-holes) or using some initial model(s) representing the best guess at the subsurface distribution. The variogram calculations involve three specified ellipsoid axes in any spatial directions and as such, their method allows construction of structures with different shapes and orientations. Estimation of the parameter covariance matrix requires significant memory and computation time.

Last and Kubik (1983) developed a compact (minimum volume) gravity inversion. Guillen and Menichetti (1984) extended the method to allow specification of a dip line along which the moment of inertia is minimized. Barbosa et al. (1994) extended the method further to allow specification of mass concentration information along several arbitrary axes. The compactness measures used lead to nonlinear functions to be minimized (in a deterministic framework) or sampled (in a probabilistic framework) which increases the computational burden above that of our methods.

Barbosa and Silva (2006) apply the method of Barbosa et al. (1994) within an interactive environment in which the interpreter can adjust the arbitrary axes as desired in order to aid geologic hypothesis testing. As such, their approach is similar to interactive forward modelling, where the interpreter investigates the model space, but differs in that the algorithm automatically fits the data.

Another novel approach is that of Wijns and Kowalczyk (2007) who, similarly to the approach of Barbosa and Silva (2006), allow for input from the interpreter to help ensure a geologically reasonable solution. Several inversions are performed with random values for several control parameters. The resulting suite of recovered models are then visually inspected by the interpreter and ranked by how geologically reasonable they are (based on the interpreter's prior knowledge). A genetic algorithm then takes that ranking into account and modifies the control parameter set to generate a new suite of models. This procedure progressively converges towards a reasonable set of solutions but requires a significant increase in the amount of inversions performed.

Orientation information can be incorporated in a natural way into all of the methods mentioned above. There is significant computational cost associated with most of those and we choose to work in a relatively computationally efficient deterministic framework.

Orientation information may come in different forms. Surface mapping can provide direct local measurements of structural orientation. Additional drilling may indicate approximate orientation information across larger volumes. If a geologic (rock) model can be created in the later stages of exploration then this can provide orientation information everywhere within the volume. Such information can be placed into an inversion as so-called "soft" or "hard" constraints, the former being a request and the latter being a guarantee. We begin by presenting soft and hard approaches for incorporating orientation information into our inversion framework. We then demonstrate the use of our methods on scenarios involving different types of orientation information.

4.2 Incorporating orientation information as soft constraints

4.2.1 Our deterministic inversion framework

In our numerical inverse solutions, the Earth region of interest is divided into many cells within an orthogonal mesh with the physical property of interest being constant across each cell. There are generally more cells than there are data and the resulting underdetermined inverse problem is formulated as an optimization that involves minimization of a (total) objective function, Φ , that combines a data misfit measure, Φ_d , with a regularization term, Φ_m , also called the model objective function:

$$\min_{\mathbf{m}} \quad \Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \beta \Phi_m(\mathbf{m}). \quad (4.1)$$

Here, \mathbf{m} is the model vector that holds the physical property values in each mesh cell of our discretised Earth and β is a trade-off parameter that controls the relative size of the Φ_d and Φ_m measures for the resulting model and allows us to tune the level of data fit as desired. The data misfit term controls how well we fit the data and the regularization term allows us to control the amount and type of structure in the recovered model.

The data misfit term measures the difference between the noisy observed data, \mathbf{d}^{obs} , and the data produced (predicted) by a candidate model, $\mathbf{d}^{pred} = F[\mathbf{m}]$. We define the data misfit as a sum-of-squares

$$\Phi_d = \sum_{i=1}^N \left(\frac{d_i^{pred} - d_i^{obs}}{\sigma_i} \right)^2 \quad (4.2)$$

where N is the number of data. Each data difference is normalized by an uncertainty, σ_i . These uncertainties are estimated errors in the observed data. The larger the uncertainty in an observed datum, the smaller its contribution to the misfit measure.

We use a model objective function that helps to recover smooth physical property models. Li and Oldenburg (1996) developed a model objective function that measured smoothness in three axial directions, with tunable parameters allowing specification of different elongations along those axes. Li and Oldenburg (2000) extended this formulation to allow the three axes to be arbitrarily rotated in 3D, thereby allowing inclusion of important orientation information (strike, dip and tilt) into the inversion in a soft manner. Below we provide a brief synopsis of the method of Li and Oldenburg (2000) and we demonstrate some drawbacks of, and our improvements to, their implementation.

4.2.2 Specifying preferred elongation directions and aspect ratios

Li and Oldenburg (1996) designed a model objective function of the following form:

$$\begin{aligned} \phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V w_x \left\{ \frac{\partial}{\partial x} (m - m_{ref}) \right\}^2 dv \\ & + \int_V w_y \left\{ \frac{\partial}{\partial y} (m - m_{ref}) \right\}^2 dv \\ & + \int_V w_z \left\{ \frac{\partial}{\partial z} (m - m_{ref}) \right\}^2 dv \end{aligned} \quad (4.3)$$

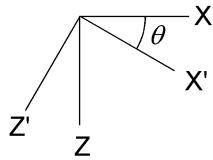


Figure 4.1: The mesh-orthogonal and rotated coordinate systems for a 2D problem.

(we use ϕ and m for continuous space and Φ and \mathbf{m} for the discrete case). By altering the relative values of the smoothness weights w_x , w_y and w_z in equation (4.3) we can cause the recovered models to become smoother (i.e. elongated) in some mesh-orthogonal direction(s) compared to the other(s), allowing specification of relative aspect ratios.

Specifying a preferred elongation in any (generally non-axial) direction is not possible using (4.3) because the derivatives are squared and directional information is thereby lost (Li and Oldenburg, 2000). In general, the geologic features will not be aligned with the mesh axes because the mesh is designed such that its horizontal axes are compatible with the survey grid over which the data were collected (with the remaining axis vertical).

4.2.3 A 2D dipping model objective function

A further generalization of the model objective function by Li and Oldenburg (2000) allows the coordinate axes to be rotated such that preferred elongations can be specified in any (perhaps non-axial) directions. Below we summarize their formulation for the 2D problem. The mesh-orthogonal axes are denoted x and z with the x -axis horizontal and z positive down. The rotated coordinates are x' and z' . The dip angle between the two coordinate systems is θ , measured downward from horizontal (i.e. from the x -axis towards the z -axis). This is represented graphically in Figure 4.1. The angle θ is used to align the rotated coordinates with the principal axes of the subsurface structures.

The new model objective function in 2D, with smoothness directions specified by the rotated

axes, is

$$\begin{aligned}\phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V w_{x'} \left(\frac{\partial m}{\partial x'} \right)^2 dv \\ & + \int_V w_{z'} \left(\frac{\partial m}{\partial z'} \right)^2 dv\end{aligned}\quad (4.4)$$

where we have simplified by removing the reference models in the smoothness terms. The rotation matrix between the two coordinate systems is

$$\mathbf{R} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (4.5)$$

which can be applied to the horizontal and vertical derivatives of the model so that orthogonal derivatives in an arbitrary coordinate system can be obtained:

$$\frac{\partial m}{\partial x'} = \cos \theta \frac{\partial m}{\partial x} + \sin \theta \frac{\partial m}{\partial z} \quad (4.6a)$$

$$\frac{\partial m}{\partial z'} = -\sin \theta \frac{\partial m}{\partial x} + \cos \theta \frac{\partial m}{\partial z}. \quad (4.6b)$$

Substitution of equation (4.6) into equation (4.4) provides the dipping model objective function

$$\begin{aligned}\phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V w_{x'} \left(\cos \theta \frac{\partial m}{\partial x} + \sin \theta \frac{\partial m}{\partial z} \right)^2 dv \\ & + \int_V w_{z'} \left(-\sin \theta \frac{\partial m}{\partial x} + \cos \theta \frac{\partial m}{\partial z} \right)^2 dv\end{aligned}\quad (4.7)$$

which rearranges to

$$\begin{aligned}\phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V (w_{x'} \cos^2 \theta + w_{z'} \sin^2 \theta) \left(\frac{\partial m}{\partial x} \right)^2 dv \\ & + \int_V (w_{x'} \sin^2 \theta + w_{z'} \cos^2 \theta) \left(\frac{\partial m}{\partial z} \right)^2 dv \\ & + \int_V 2 (w_{x'} - w_{z'}) \cos \theta \sin \theta \frac{\partial m}{\partial x} \frac{\partial m}{\partial z} dv\end{aligned}\quad (4.8)$$

and the discrete representation is

$$\begin{aligned}\Phi_m = & (\mathbf{m} - \mathbf{m}_{ref})^T \mathbf{W}_s^T \mathbf{W}_s (\mathbf{m} - \mathbf{m}_{ref}) \\ & + \mathbf{m}^T (\mathbf{D}_x^T \mathbf{B}_x \mathbf{D}_x + \mathbf{D}_z^T \mathbf{B}_z \mathbf{D}_z \\ & + \mathbf{D}_x^T \mathbf{B}_{xz} \mathbf{D}_z + \mathbf{D}_z^T \mathbf{B}_{xz} \mathbf{D}_x) \mathbf{m}\end{aligned}\quad (4.9)$$

where \mathbf{D}_x and \mathbf{D}_z are finite difference operators; \mathbf{W}_s is a diagonal matrix containing the smallness weights w_s ; \mathbf{B}_x , \mathbf{B}_z and \mathbf{B}_{xz} are diagonal matrices containing the trigonometric terms in equation (4.8); and the last term in equation (4.8) is represented by two cross-terms in equation (4.9) to promote symmetry.

Here we use ℓ^2 -norms (sum-of-squares) but note that general measures, such as those of Farquharson (1998), could be employed instead to encourage sharp interfaces. In that case, the objective function must remain in the form of equation (4.7) before discretisation because the form in equation (4.8) is only valid for ℓ^2 -norms.

4.2.4 Extension to 3D: specification of strike, dip and tilt

In the 3D coordinate system let $+x$ be in the northing direction, $+y$ easting and $+z$ down. Three angles are required to define the orientation of a 3D planar object (i.e. a plate): φ is the strike angle (the intersection of the plane with a horizontal surface) defined positive east of north; θ is the dip defined positive downward from horizontal; and ψ is the tilt. Li and Oldenburg (2000) used a tilt instead of a plunge because the strike direction as defined above is constant under arbitrary tilt angle. The graphic representation of the 3D scenario as depicted in Li and Oldenburg (2000) is shown in Figure 4.2. The tilt is the rotation of the object within its dipping plane i.e. around the y' axis in Figure 4.2).

The 3D rotation matrix is created via three sequential rotations. First, a rotation of φ is performed around the z axis. With $\varphi > 0$ this moves x towards y creating new axes x' and y' (refer to Figure 4.2). The first rotation matrix is

$$\mathbf{R}_z = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.10)$$

This is followed by a rotation of $\theta - 90$ degrees around the newly formed x' axis. With $0 < \theta < 90$

this moves z towards y' creating new axis z' . The second rotation matrix is

$$\begin{aligned}\mathbf{R}_x &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta - 90) & \sin(\theta - 90) \\ 0 & -\sin(\theta - 90) & \cos(\theta - 90) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin(\theta) & -\cos(\theta) \\ 0 & \cos(\theta) & \sin(\theta) \end{pmatrix}.\end{aligned}\tag{4.11}$$

Finally, a rotation of ψ occurs around the y' axis. With $\psi > 0$ this moves x' towards z' . Note that there is an inconsistency in sign between this last rotation and that shown in Figure 4.2 provided by Li and Oldenburg (2000). The third rotation matrix is

$$\begin{aligned}\mathbf{R}_y &= \begin{pmatrix} \cos(-\psi) & 0 & -\sin(-\psi) \\ 0 & 1 & 0 \\ \sin(-\psi) & 0 & \cos(-\psi) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\psi) & 0 & \sin(\psi) \\ 0 & 1 & 0 \\ -\sin(\psi) & 0 & \cos(\psi) \end{pmatrix}.\end{aligned}\tag{4.12}$$

The final 3D rotation matrix is created by combining those above:

$$\begin{aligned}\mathbf{R} &= \mathbf{R}_y \mathbf{R}_x \mathbf{R}_z = \begin{pmatrix} r_{xx} & r_{xy} & r_{xz} \\ r_{yx} & r_{yy} & r_{yz} \\ r_{zx} & r_{zy} & r_{zz} \end{pmatrix} \\ r_{xx} &= \cos \varphi \cos \psi - \sin \varphi \cos \theta \sin \psi \\ r_{xy} &= \sin \varphi \cos \psi + \cos \varphi \cos \theta \sin \psi \\ r_{xz} &= \sin \theta \sin \psi \\ r_{yx} &= -\sin \varphi \sin \theta \\ r_{yy} &= \cos \varphi \sin \theta \\ r_{yz} &= -\cos \theta \\ r_{zx} &= -\cos \varphi \sin \psi - \sin \varphi \cos \theta \cos \psi \\ r_{zy} &= -\sin \varphi \sin \psi + \cos \varphi \cos \theta \cos \psi \\ r_{zz} &= \sin \theta \cos \psi.\end{aligned}\tag{4.13}$$

Note that a typographical error in Li and Oldenburg (2000) provides an inconsistency in sign

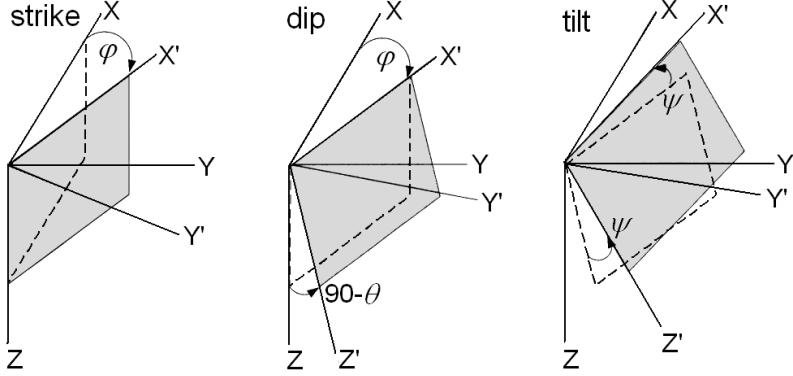


Figure 4.2: The mesh-orthogonal and rotated coordinate systems for a 3D problem (this figure was reproduced from Li and Oldenburg (2000)).

m_1	m_2	m_3	m_4
-------	-------	-------	-------

Figure 4.3: A 1D mesh with four cells.

between the r_{zy} quantity in equation (4.13) and that given in Li and Oldenburg (2000).

4.2.5 The choice of difference operators

The matrix multiplications and additions in equation (4.9) indicate that the matrices \mathbf{D}_x and \mathbf{D}_z must be square and of the same size. In the non-rotated (mesh-aligned) formulation, the traditional discrete differential operators calculate differences across cell-faces, with the x -direction gradients defined on vertical cell faces and the z -direction gradients defined on horizontal cell faces. Consequently, \mathbf{D}_x and \mathbf{D}_z are never square and are only the same size if the mesh contains the same number of cells in all directions. Li and Oldenburg (2000) defined the discrete x and z gradients at the centres of each cell so that the difference operators are the same size.

For a 1D problem with four model cells of unit dimensions (see Figure 4.3) the traditional \mathbf{D}_x operator would be

$$\mathbf{D}_x = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \quad (4.14)$$

m_1	m_2	m_3
m_4	m_5	m_6
m_7	m_8	m_9

Figure 4.4: A 3-by-3 2D mesh. The shaded cells indicate those involved in the finite differences defined at the centre of cell 5 when forward differences are employed.

which has three rows that define three differences operating across the three faces between the four cells. With differences defined at cell centres we need four differences (one for each cell) instead of three.

Cell-centred forward and backward differences

One option is to use forward differences for all cells and backward differences where necessary:

$$\mathbf{D}_x = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}. \quad (4.15)$$

Forward differences are used for cells 1 through 3 in Figure 4.3. A backward difference must be used for cell 4 which results in a difference operator with the lower two rows identical. This was the approach taken by Li and Oldenburg (2000). One may expect, as is shown below, that the use of backward differences instead may lead to different results.

Now consider a 3-by-3 package of cells in a 2D model, as shown in Figure 4.4. The x - and z -direction forward differences for the central cell 5 involve cells 6 and 8 respectively. Writing equation (4.7) for only cell 5 yields

$$\begin{aligned} \phi_m(m) = & w_s (m_5 - m_{ref,5})^2 dv \\ & + w_x' (\cos \theta (m_6 - m_5) + \sin \theta (m_8 - m_5))^2 dv \\ & + w_z' (-\sin \theta (m_6 - m_5) + \cos \theta (m_8 - m_5))^2 dv. \end{aligned} \quad (4.16)$$

With $\theta = +45^\circ$ this reduces to

$$\begin{aligned}\phi_m(m) = & w_s (m_5 - m_{ref,5})^2 dv \\ & + w_{x'} (-2m_5 + m_6 + m_8)^2 dv \\ & + w_{z'} (m_8 - m_6)^2 dv.\end{aligned}\quad (4.17)$$

(a constant in the smoothness terms equal to $\cos 45^\circ = \sin 45^\circ$ has been ignored here for simplification purposes). To recover a model elongated in the new x' direction (which dips at $+45^\circ$) we would set $w_{x'} \gg w_{z'}$ and effectively have

$$\begin{aligned}\phi_m(m) = & w_s (m_5 - m_{ref,5})^2 dv \\ & + w_{x'} (-2m_5 + m_6 + m_8)^2 dv.\end{aligned}\quad (4.18)$$

With $\theta = -45^\circ$ we would instead obtain

$$\begin{aligned}\phi_m(m) = & w_s (m_5 - m_{ref,5})^2 dv \\ & + w_{x'} (m_6 - m_8)^2 dv \\ & + w_{z'} (-2m_5 + m_6 + m_8)^2 dv.\end{aligned}\quad (4.19)$$

and setting $w_{x'} \gg w_{z'}$ effectively gives

$$\begin{aligned}\phi_m(m) = & w_s (m_5 - m_{ref,5})^2 dv \\ & + w_{x'} (m_6 - m_8)^2 dv.\end{aligned}\quad (4.20)$$

Note that there is an asymmetry indicated here: in equation (4.18) (for $\theta = +45^\circ$) the x' term considers values in three cells, whereas in equation (4.20) (for $\theta = -45^\circ$) the x' term contains only two cell values. This asymmetry of the forward differences leads to asymmetric results.

To demonstrate the asymmetric effects we present a small 2D gravity example. The true model is shown in Figure 4.5(a), for which gravity data are modelled and a small amount of noise added before inverting. The smallness and smoothness weights are constant across the model; w_s is set to 0.001 to balance the smallness and smoothness terms and the smoothness weights are either 1.0 or 0.001 (depending on the inversion). We set the reference model equal to the true model and apply this reference model in the smallness term only; we do so to make the asymmetric effects more apparent in the figures presented. Figure 4.5(b) shows the result of an inversion specifying no preferred elongation direction. The asymmetry of the forward differences is evident when comparing the result in Figure 4.6(a), in which we specify $\theta = +45^\circ$, to that in Figure 4.6(b), in which we specify $\theta = -45^\circ$. Ideally those two results should be symmetric

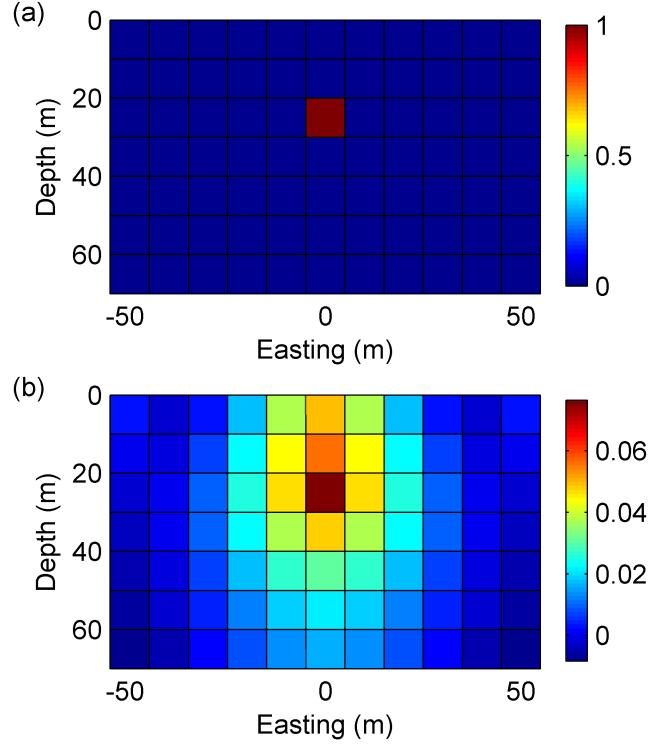


Figure 4.5: The true 2D density model is in (a). The recovered 2D density model with no preferred elongation direction specified ($w_{x'} = w_{z'} = 1.0$) is in (b).

across a vertical line bisecting the mesh (ignoring any asymmetry introduced by the random noise added to the data).

Cell-centred long differences

There are three options for promoting symmetry. The first is to use long differences across a cell. Consider again the 3-by-3 2D package of cells in Figure 4.4. Using long differences across cell 5 would provide the following contribution to the objective function in equation (4.7):

$$\begin{aligned} \phi_m(m) = & w_s (m_5 - m_{ref,5})^2 dv \\ & + w_{x'} (\cos \theta (m_6 - m_4) + \sin \theta (m_8 - m_2))^2 dv \\ & + w_{z'} (-\sin \theta (m_6 - m_4) + \cos \theta (m_8 - m_2))^2 dv. \end{aligned} \quad (4.21)$$

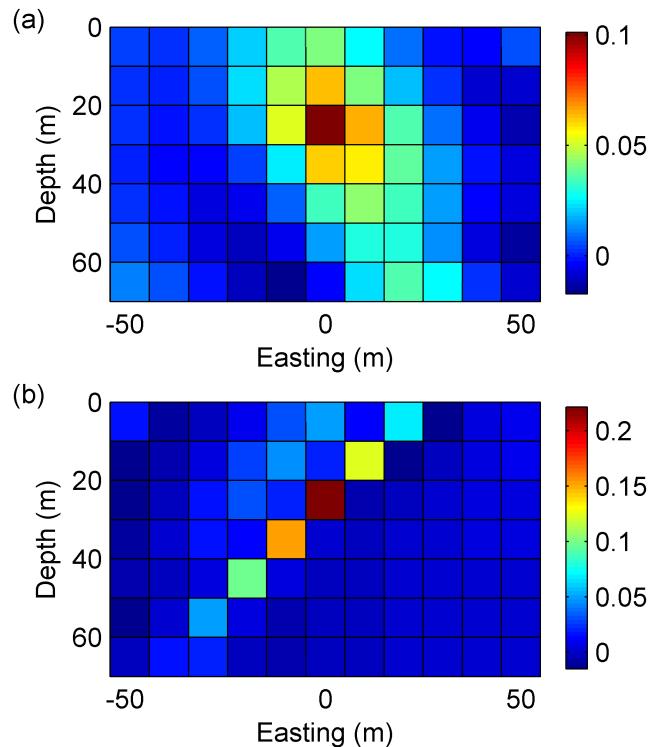


Figure 4.6: The recovered 2D density models using a dipping-objective function with forward differences used in the model objective function. The specified dips are (a) $+45^\circ$ and (b) -45° . The weights used were $w_{x'} = 1.0$ and $w_{z'} = 0.001$ across the entire mesh.

The result is that the even-numbered cells are all linked together through the finite-difference interactions in the model objective function, and similarly the odd-numbered cells are all linked together. However, no differences occur between even- and odd-numbered cells. In other words, each cell is linked through the difference operators to the cells diagonally adjacent to it, but there is no link to the cells across its faces. The mesh is thereby separated into two parts like a chess-board with a set of “black” cells and a set of “white” cells as in Figure 4.7(a): all the black cells are linked to each other, and the same is true for the white cells, but none of the black cells are linked to any of the white cells. Hence, smoothness can be maintained in the diagonal directions but not in the axial directions.

Mathematically, the chess-board pattern in Figure 4.7(a) is an annihilator for the discrete gradient operators formed using long differences. That is, given some model, we can arbitrarily add some value to all the “black” model cells, and arbitrarily add some other value to all the “white” model cells, and the value of the smoothness terms in the model objective function will not change (similarly, a constant value is also an annihilator for the smoothness terms). One might say that the chess-board pattern is invisible to the smoothness terms in the model objective function when long differences are used.

Node-centred integration

The second option for promoting symmetry is to move to a node-centred integration scheme. We would then use 2-by-2 packages of cells in 2D. Consider the top left 2-by-2 package of cells in Figure 4.4. A nodal scheme would integrate the model objective function around each node (rather than over each cell) such that the contribution to the model objective function for the node within the top left 2-by-2 package of cells would be

$$\begin{aligned}
 \phi_m(m) = & \text{(smallness terms)} \\
 & + w_x' \left(\cos \theta (m_2 - m_1 + m_5 - m_4) \right. \\
 & \quad \left. + \sin \theta (m_4 - m_1 + m_5 - m_2) \right)^2 dv \\
 & + w_z' \left(-\sin \theta (m_2 - m_1 + m_5 - m_4) \right. \\
 & \quad \left. + \cos \theta (m_4 - m_1 + m_5 - m_2) \right)^2 dv. \tag{4.22}
 \end{aligned}$$

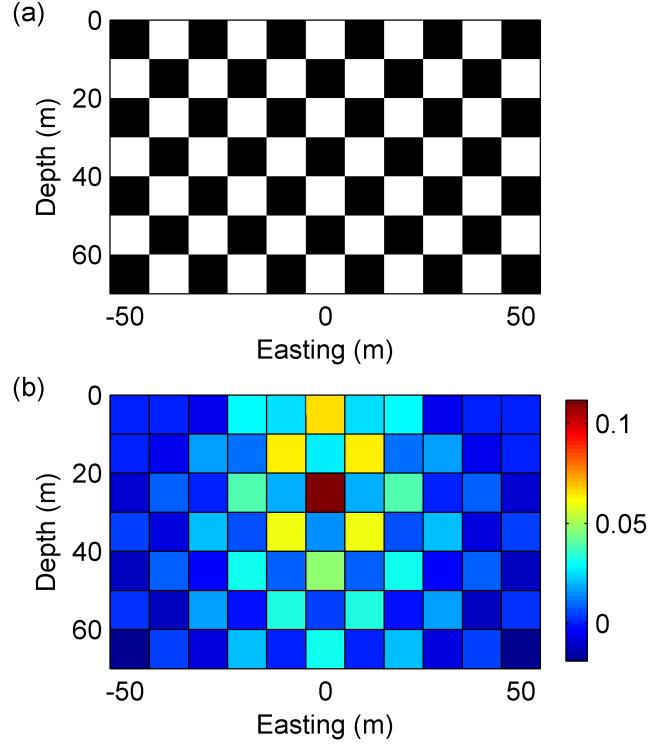


Figure 4.7: A chess-board pattern is in (a). The recovered 2D density model using node-centred finite-difference operators in the model objective function with $\theta = 45^\circ$ and $w_{x'} = w_{z'}$ is in (b).

With $\theta = 45^\circ$ this reduces to

$$\begin{aligned} \phi_m(m) = & \text{(smallness terms)} \\ & + w_{x'} (m_5 - m_1)^2 dv \\ & + w_{z'} (m_4 - m_2)^2 dv. \end{aligned} \quad (4.23)$$

At first glance this may seem appropriate, with diagonal differences as requested by setting $\theta = 45^\circ$. However, note that the chess-board annihilator again becomes an issue. In Figure 4.7(b) we show the result of inverting the example gravity data with a node-centered scheme, specifying $\theta = 45^\circ$ and with $w_{x'} = w_{z'}$. The chess-board annihilator pattern is clearly visible.

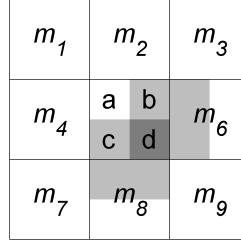


Figure 4.8: A 3-by-3 2D mesh. The central cell 5 is split into 4 quadrants. The shaded regions across cell faces indicate the regions across which x - and z -direction forward finite differences are calculated for cell 5 (those regions overlap in quadrant 5d).

set	\mathbf{D}_x	\mathbf{D}_z
1	backward	backward
2	forward	backward
3	backward	forward
4	forward	forward

Table 4.1: Difference options for the four sets of finite difference operators for the 2D dipping model objective function.

Four quadrant face-centred integration

Our solution for discretization avoids asymmetric results and the chess-board annihilator. We follow the traditional approach of calculating derivatives through finite differences across cell faces. In 2D, the elemental areas in which the x - and z -directional differences are defined are on different overlapping grids, as indicated in Figure 4.8. Hence, we must split each cell into four quadrants as indicated in Figure 4.8 and integrate over each quadrant separately. In quadrant 5d the x -direction difference is between the values in cells 6 and 5, and the z -direction difference is between the values in cells 8 and 5: this is much like using forward differences in quadrant 5d. In quadrant 5a we would essentially use backward differences for both the x - and z -directions; in quadrants 5b and 5c we would use a mixture of forward and backward differences. Hence, this approach is equivalent to using four sets of \mathbf{D}_x and \mathbf{D}_z operators, each set using a different option for the differences (forward or backward) as indicated in Table 4.1.

Let the i^{th} operators be denoted $\mathbf{D}_{x,i}$ and $\mathbf{D}_{z,i}$. The discrete representation of our 2D model

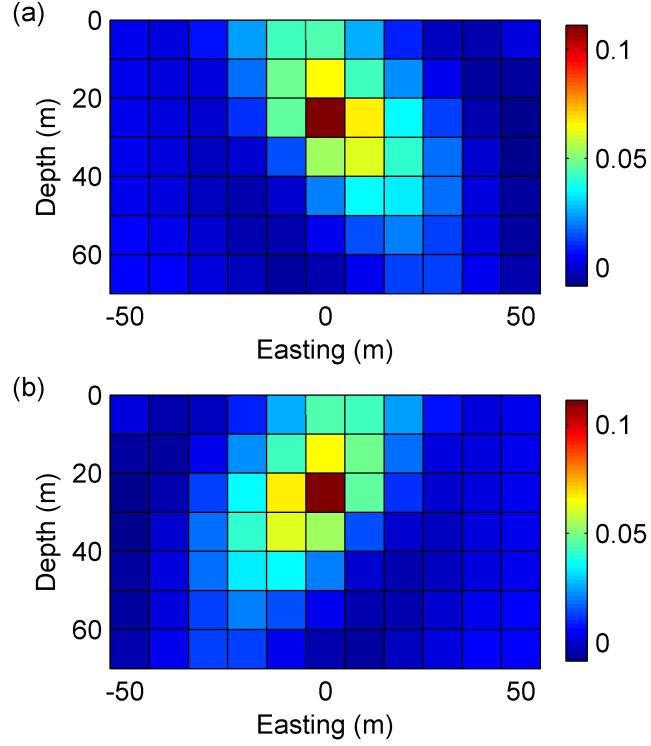


Figure 4.9: The recovered 2D density models using a dipping-objective function with four sets of differences used in the model objective function. The specified dips are (a) $+45^\circ$ and (b) -45° . The weights used were $w_{x'} = 1.0$ and $w_{z'} = 0.001$ across the entire mesh.

objective function is then

$$\begin{aligned}
 \Phi_m &= (\mathbf{m} - \mathbf{m}_{ref})^T \mathbf{W}_s^T \mathbf{W}_s (\mathbf{m} - \mathbf{m}_{ref}) \\
 &+ \mathbf{m}^T \frac{1}{4} \sum_{i=1}^4 (\mathbf{D}_{x,i}^T \mathbf{B}_x \mathbf{D}_{x,i} + \mathbf{D}_{z,i}^T \mathbf{B}_z \mathbf{D}_{z,i} \\
 &\quad + \mathbf{D}_{x,i}^T \mathbf{B}_{xz} \mathbf{D}_{z,i} + \mathbf{D}_{z,i}^T \mathbf{B}_{xz} \mathbf{D}_{x,i}) \mathbf{m} \\
 &= (\mathbf{m} - \mathbf{m}_{ref})^T \mathbf{W}_s^T \mathbf{W}_s (\mathbf{m} - \mathbf{m}_{ref}) + \mathbf{m}^T \mathbf{W}^T \mathbf{W} \mathbf{m}
 \end{aligned} \tag{4.24}$$

which yields a $\mathbf{W}^T \mathbf{W}$ operator that provides symmetric results (i.e. the results do not depend on the sign of θ) as is evident from comparing the results in Figures 4.9(a) and 4.9(b). Returning to the 2D example of Figure 4.4, the central cell 5 is now linked to all the cells across its faces, as indicated by the shaded cells in Figure 4.10. The chess-board issue is also ameliorated using this four set approach. In 3D we use eight sets of operators (all possible permutations).

m_1	m_2	m_3
m_4	m_5	m_6
m_7	m_8	m_9

Figure 4.10: A 3-by-3 2D mesh. The shaded cells indicate those involved in the finite differences defined at the centre of cell 5 when four sets of finite difference operators are employed.

4.2.6 Specifying orientations at point locations

The smoothness weights $w_{x'}$ and $w_{z'}$ can be homogeneous across the entire mesh or can be set to different values in different regions. It is thereby possible to specify orientations globally or locally. Li and Oldenburg (2000) elected to specify orientations on a regional basis. Instead, we extend the functionality such that different orientations can be specified at each mesh cell centre. Of course, orientations involve spatial gradients across finite distances so to say that we allow specification of an orientation at a point in space is somewhat misleading; in our formulation, each cell is linked to the cells adjacent to it so the cell-based orientations are thereby spread out to influence surrounding cells and the overlap inherent here acts as a smoothing mechanism.

4.2.7 A practical issue: depth and distance weighting

Depth weighting, or a more general distance weighting, must be incorporated into gravity and magnetic inversions due to the fall-off of the responses with distance. In Li and Oldenburg (1996, 1998), a depth weighting function $Z(z)$ is applied to the model such that the model objective function is of the form

$$\begin{aligned} \phi_m(m) = & \int_V w_s (Z(z)(m - m_{ref}))^2 dv \\ & + \int_V w_x \left(\frac{\partial Z(z)m}{\partial x} \right)^2 dv \\ & + \int_V w_y \left(\frac{\partial Z(z)m}{\partial y} \right)^2 dv \\ & + \int_V w_z \left(\frac{\partial Z(z)m}{\partial z} \right)^2 dv. \end{aligned} \tag{4.25}$$

Here, the depth weighting is applied inside the derivatives of the smoothness terms. The discrete representation would be

$$\begin{aligned}\Phi_m = & \|\mathbf{W}_s \mathbf{Z}(\mathbf{m} - \mathbf{m}_{ref})\|^2 \\ & + \|\mathbf{D}_x \mathbf{Zm}\|^2 + \|\mathbf{D}_y \mathbf{Zm}\|^2 + \|\mathbf{D}_z \mathbf{Zm}\|^2\end{aligned}\quad (4.26)$$

where we have absorbed the w_x , w_y and w_z weights into the finite difference operators. This approach had proven to work satisfactorily on a number of synthetic examples and became the default choice in the algorithm of Li and Oldenburg (1996).

There is an issue here however: with the depth weighting applied inside the derivatives, the z -smoothness term no longer provides an exact z -direction gradient because the model is altered by the depth weighting before the differential operator is applied. With depth weighting, this is only an issue for the z -direction since in the other smoothness terms the order of operations is irrelevant (i.e. Z is a function of z only and can therefore be taken outside of the derivatives with respect to x and y). With distance weighting, however, all smoothness terms are affected. It would be more appropriate to perform the depth or distance weighting outside the derivatives such that the model objective function is of the form

$$\begin{aligned}\phi_m(m) = & \int_V w_s Z_s(z)^2 (m - m_{ref})^2 dv \\ & + \int_V w_x Z_x(z)^2 \left(\frac{\partial m}{\partial x} \right)^2 dv \\ & + \int_V w_y Z_y(z)^2 \left(\frac{\partial m}{\partial y} \right)^2 dv \\ & + \int_V w_z Z_z(z)^2 \left(\frac{\partial m}{\partial z} \right)^2 dv.\end{aligned}\quad (4.27)$$

where we have added subscripts to the depth weighting matrices to indicate that they may be different sizes. The discrete representation would be

$$\begin{aligned}\Phi_m = & \|\mathbf{Z}_s \mathbf{W}_s(\mathbf{m} - \mathbf{m}_{ref})\|^2 \\ & + \|\mathbf{Z}_x \mathbf{D}_x \mathbf{m}\|^2 + \|\mathbf{Z}_y \mathbf{D}_y \mathbf{m}\|^2 + \|\mathbf{Z}_z \mathbf{D}_z \mathbf{m}\|^2.\end{aligned}\quad (4.28)$$

The issue is especially important when using a rotated objective function: if depth weighting is applied inside the z -derivative then the derivative calculated is not the exact z -gradient quantity required.

4.3 Incorporating orientation information as hard constraints

Specifying elongation information (i.e. axis directions and aspect ratios) through the smoothness measures is considered a soft constraint: we request of the inversion that the specified elongations be recovered but there is no guarantee. To obtain such a guarantee we can include the orientation information as hard constraints by bounding spatial model gradients and gradient ratios. We now consider two scenarios. In the first, we investigate how to bound the direction of the spatial gradient. In the second, we consider how to specify a physical property increase or decrease along a particular direction.

4.3.1 Bounding the spatial gradient direction

Let L and U specify lower and upper bounds on the dip angle θ : in 2D we can then write the trigonometric inequalities

$$\tan L \leq \left(\tan \theta = \frac{\nabla_z m}{\nabla_x m} \right) \leq \tan U. \quad (4.29)$$

After discretisation, equation (4.29) gives

$$\mathbf{L}\mathbf{D}_x\mathbf{m} \leq \mathbf{D}_z\mathbf{m} \leq \mathbf{U}\mathbf{D}_x\mathbf{m} \quad (4.30)$$

where \mathbf{L} and \mathbf{U} are diagonal matrices containing the bounding values on their main diagonals. To simplify we can split equation (4.30) into two inequalities

$$\mathbf{D}_z\mathbf{m} \geq \mathbf{L}\mathbf{D}_x\mathbf{m} \quad (4.31a)$$

$$\mathbf{U}\mathbf{D}_x\mathbf{m} \geq \mathbf{D}_z\mathbf{m} \quad (4.31b)$$

which rearrange to

$$(\mathbf{D}_z - \mathbf{L}\mathbf{D}_x)\mathbf{m} \geq \mathbf{0} \quad (4.32a)$$

$$(\mathbf{U}\mathbf{D}_x - \mathbf{D}_z)\mathbf{m} \geq \mathbf{0} \quad (4.32b)$$

and lead to linear constraints of the form

$$\mathbf{A}\mathbf{m} \geq \mathbf{b}. \quad (4.33)$$

Equation (4.33) is a system of equations, each of the form

$$a_1 m_1 + a_2 m_2 + \dots \geq b. \quad (4.34)$$

These constraints can be added to the inverse problem such that the resulting optimization problem is

$$\min_{\mathbf{m}} \quad \Phi_d(\mathbf{m}) + \beta \Phi_m(\mathbf{m}) \quad (4.35a)$$

$$s.t. \quad \mathbf{A}\mathbf{m} \geq \mathbf{b}. \quad (4.35b)$$

In 3D we could write more trigonometric inequalities for the three angles φ , θ and ψ . However, the resulting inequalities no longer reduce to linear constraints. For example, with $\psi = 0$ and $\varphi \neq 0$ the 3D equivalent to equation (4.29) is

$$\tan L \leq \left(\tan \theta = \frac{\nabla_z m}{\sqrt{(\nabla_x m)^2 + (\nabla_y m)^2}} \right) \leq \tan U \quad (4.36a)$$

and the squaring leads to nonlinear inequalities. Such constraints would increase the difficulty of the optimization problem and for this reason we do not consider them further.

4.3.2 Specifying directions of increase or decrease

We now re-pose the problem by specifying that a directional spatial derivative, in some specified location and direction \vec{v} , lies between some bounds. The directional derivative is the dot product of \vec{v} with the spatial model gradient and in 3D we obtain the inequality

$$L \leq (v_x \nabla_x + v_y \nabla_y + v_z \nabla_z) m \leq U \quad (4.37)$$

where the lower and upper bounds L and U can be used to specify that the physical property increases, decreases or remains constant along the direction \vec{v} . The discrete form is

$$\mathbf{L} \leq (\mathbf{V}_x \mathbf{D}_x + \mathbf{V}_y \mathbf{D}_y + \mathbf{V}_z \mathbf{D}_z) \mathbf{m} \leq \mathbf{U} \quad (4.38)$$

where matrices \mathbf{L} , \mathbf{U} , \mathbf{V}_x , \mathbf{V}_y and \mathbf{V}_z are diagonal. Again, equation (4.38) can be reduced to the simple form in equation (4.33).

4.3.3 Practical application of linear constraints

In equations (4.30) and (4.38) we again have a situation in which the difference operators must be the same size. We can take the same approach as in section 4.2.5 and use four sets of operators and therefore four sets of linear inequalities (or eight in 3D). This issue is removed if we simplify the problem to placing bounds on a particular gradient component, e.g.

$$\mathbf{L} \leq \mathbf{D}_z \mathbf{m} \leq \mathbf{U}, \quad (4.39)$$

or to specifying some other general relationship between model parameters, e.g.

$$a_1 m_1 + a_2 m_2 + \dots \geq b. \quad (4.40)$$

Consider a situation in which we can say with some certainty that one region of the subsurface is one rock type, another region is a second rock type, but we are not sure which of the two rock types falls between those two regions. Perhaps one rock type is an intrusion into the second, or perhaps the two units are separated by an offset fault at depth. Another possible scenario involves a cover unit of unknown thickness above another unit. If relative physical property values between rock types are known then model gradient values can be bounded in the unknown region. With less confident physical property information we may still be able to specify the sign of the model gradients in the unknown region (i.e. specifying a directional increase or decrease).

The linear constraints developed above allow specification of relative spatial relationships between rock units. They also allow incorporation of information regarding the direction and magnitude of alteration gradients. Another use is to incorporate poorly calibrated physical property measurements taken on rock samples, in which case the values could only be treated as relative (i.e. $L \leq m_1/m_2 \leq U$).

4.3.4 Linear inequality constrained inverse problems

Above we indicated how orientation information can be specified in a hard manner through the addition of linear constraints to the optimization problem. We now present a strategy for solving the inverse problem with linear inequality constraints. To minimize the problem in equation (4.35), we follow the approach of Li and Oldenburg (2003) and use a logarithmic

barrier method. Although there are modern alternatives to the logarithmic barrier method (refer to Gill et al. (1995)), it has proven to be a feasible solution method for large 3D geophysical inverse problems with simple bound constraints. The extension to linear constraints is as follows. The logarithmic barrier method adds a barrier term Φ_λ to the objective function and solves a sequence of unconstrained inversions

$$\min_{\mathbf{m}} \quad \Phi(\mathbf{m}) - \lambda \Phi_\lambda(\mathbf{m}) \quad (4.41)$$

while carefully cooling (reducing) the value of λ . The barrier term is

$$\Phi_\lambda = \sum_{i=1}^M \log(\mathbf{a}_i^T \mathbf{m} - b_i) \quad (4.42)$$

$$= \mathbf{e}^T \log(\mathbf{A}\mathbf{m} - \mathbf{b}) \quad (4.43)$$

where \mathbf{e} is a vector of ones, \mathbf{a}_i^T is the i^{th} row of \mathbf{A} and the log operation on a vector quantity is element-by-element.

We use a Newton-type descent method to solve the unconstrained subproblem and therefore we require the first and second order derivatives (the gradient and Hessian) of the logarithmic barrier term. The gradient is

$$\mathbf{g}_\lambda = \left(\frac{d\Phi_\lambda}{d\mathbf{m}} \right)^T \quad (4.44)$$

$$= \left(\mathbf{e}^T \frac{\partial \log(\mathbf{y})}{\partial \mathbf{y}} \frac{d\mathbf{y}}{d\mathbf{m}} \right)^T \Big|_{\mathbf{y}=\mathbf{A}\mathbf{m}-\mathbf{b}} \quad (4.45)$$

$$= (\mathbf{e}^T \text{diag}(\mathbf{y}^{-1}) \mathbf{A})^T \quad (4.46)$$

$$= \mathbf{A}^T \text{diag}(\mathbf{y}^{-1}) \mathbf{e} \quad (4.47)$$

$$= \mathbf{A}^T \mathbf{y}^{-1} \quad (4.48)$$

where the power operation on $\mathbf{y} = \mathbf{A}\mathbf{m} - \mathbf{b}$ is element-by-element. The Hessian is

$$\mathbf{H}_\lambda = \frac{d\mathbf{g}_\lambda}{d\mathbf{m}} \quad (4.49)$$

$$= \mathbf{A}^T \frac{\partial(\mathbf{y}^{-1})}{\partial \mathbf{y}} \frac{d\mathbf{y}}{d\mathbf{m}} \quad (4.50)$$

$$= -\mathbf{A}^T \text{diag}(\mathbf{y}^{-2}) \mathbf{A}. \quad (4.51)$$

Once a step direction $\delta\mathbf{m}$ is chosen we will update the current (k^{th}) model with

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha\delta\mathbf{m} \quad (4.52)$$

where the step length α is determined by performing a line search. We need to determine the maximum step length possible α_{max} without violating the constraints. For each linear constraint we have

$$\mathbf{a}_i^T \mathbf{m} \geq b_i \quad (4.53)$$

and we want to know the value α_i such that

$$\begin{aligned} \mathbf{a}_i^T (\mathbf{m} + \alpha_i \delta\mathbf{m}) &= b_i \\ \mathbf{a}_i^T \mathbf{m} + \alpha_i \mathbf{a}_i^T \delta\mathbf{m} &= b_i \\ \alpha_i &= (b_i - \mathbf{a}_i^T \mathbf{m}) / (\mathbf{a}_i^T \delta\mathbf{m}) . \end{aligned} \quad (4.54)$$

If the result for α_i above is negative then the search direction will not violate the i^{th} constraint regardless of the positive step length. Hence, we must calculate the α_i value for every linear constraint and take the smallest positive value of those as the limiting value α_{max} . We then set

$$\alpha = \min(\alpha_{max}, 1.0) . \quad (4.55)$$

Practical aspects of our algorithm follow Li and Oldenburg (2003). The step length is reduced slightly to ensure that we stay off the barrier:

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \gamma\alpha\delta\mathbf{m} , \quad \gamma = 0.925. \quad (4.56)$$

The barrier parameter is updated as

$$\lambda^{(k+1)} = [1 - \min(\alpha, \gamma)] \lambda^{(k)} \quad (4.57)$$

and the barrier iterations are continued until the barrier term has a negligible contribution to the total objective function.

The logarithmic barrier method is an interior-point method, meaning that the iterates remain feasible (the model always satisfies the constraints). This is required to avoid taking the logarithm of a negative number. The major practical difficulty is then that the initial model must be feasible and for complicated linear inequalities the creation of a feasible initial model may be a difficult task. We emphasize that these types of constrained optimization problems are at

the leading edge of optimization research and future advances in that field may provide solution methods that are more advantageous than the logarithmic barrier approach.

4.3.5 Applying hard orientation constraints to a 2D example

We now return to the small 2D gravity example in section 4.2.5 and add linear inequality constraints into the problem such that we specify a dip of $\theta = 45^\circ$ around the cell containing the anomalous density in the true model. The most thorough way to proceed would be to write equation (4.32) for that cell. This would be done for each four sets of permutations in Table 4.1 and would result in eight linear inequality constraints involving the cell of interest and its neighbours. Although that is possible, we continue with a more simple demonstration and create two pairs of linear constraints, each pair linking the cell of interest to one of its diagonally adjacent cells. If the cell of interest corresponds to cell 5 in Figure 4.4 then the linear constraints we use here specify

$$-0.001 \leq m_5 - m_1 \leq 0.001 \quad (4.58a)$$

$$-0.001 \leq m_5 - m_9 \leq 0.001 \quad (4.58b)$$

(i.e. cells 1 and 9 should contain values close to that in cell 5). We perform the same inversion that lead to the result in Figure 4.9(a) but apply the constraints in equation (4.58). The recovered model, shown in Figure 4.11, honours the linear inequality constraints in equation (4.58) and the $+45^\circ$ dip is clearly evident.

Alternatively, without some knowledge of the expected value of the cells being constrained, we may wish to specify that the relative change between two cells be less than some value, say 5%. This would lead to constraints of the form

$$\begin{aligned} -0.05 \leq \frac{m_5 - m_1}{m_1} \leq 0.05 \\ -0.05m_1 \leq m_5 - m_1 \leq 0.05m_1 \end{aligned} \quad (4.59)$$

which for this example would yield similar results.

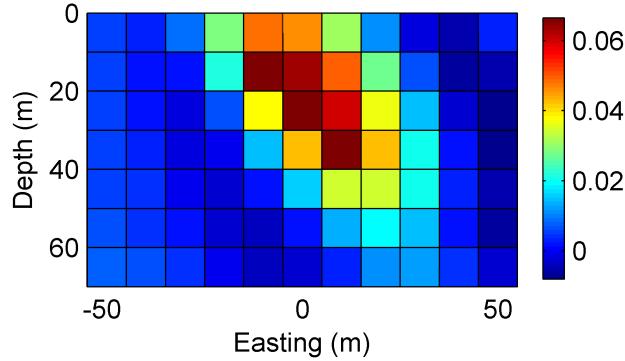


Figure 4.11: The recovered 2D density model from a similar inversion to that for Figure 4.9(a) but with additional linear inequality constraints in equation (4.58) added into the inversion.

4.4 Incorporating different forms of orientation information

As mentioned in the introduction, orientation information may come in different forms. We now demonstrate our methods for including different forms of orientation information on a more complicated 2D magnetics example. The true model, shown in Figure 4.12(a), represents a scenario in which a layered sequence of rock units has been folded into a syncline. The uniform discrete mesh is 47 cells by 23 cells. Magnetic data are calculated for the true model and a small amount of noise is added before inverting. We set the noise level below the amplitude of the signal from the lowest portion of the lower susceptible layer. The inversions are thus provided a chance to recover the lower susceptible layer but, as will become evident, the nonuniqueness of the problem makes it difficult to recover the layered scenario unless further geologic information (e.g. orientation information) is incorporated into the problem.

For all inversions mentioned below, the reference model is set to zero and smallness weights w_s are set to zero to emphasize the effects of incorporating orientation information. Figure 4.12(b) shows the inversion result with no preferred elongation direction specified ($w_x = w_z = 1.0$). This inversion fails to adequately resolve the two magnetic layers and places significant susceptible material where there is none in the true model.

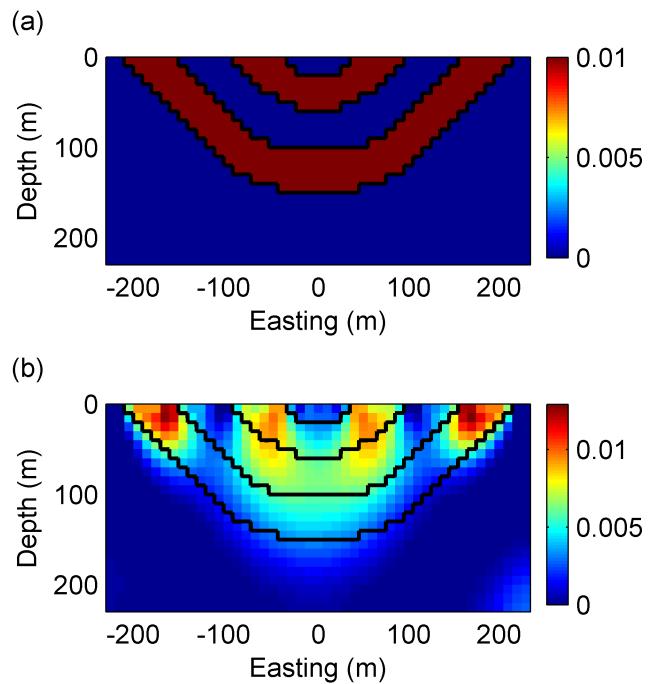


Figure 4.12: The true 2D susceptibility model is in (a). The recovered 2D susceptibility model with no preferred elongation direction specified ($w_x = w_z = 1.0$) is in (b). The location of the layers in the true model are indicated with a thick black line in (b).

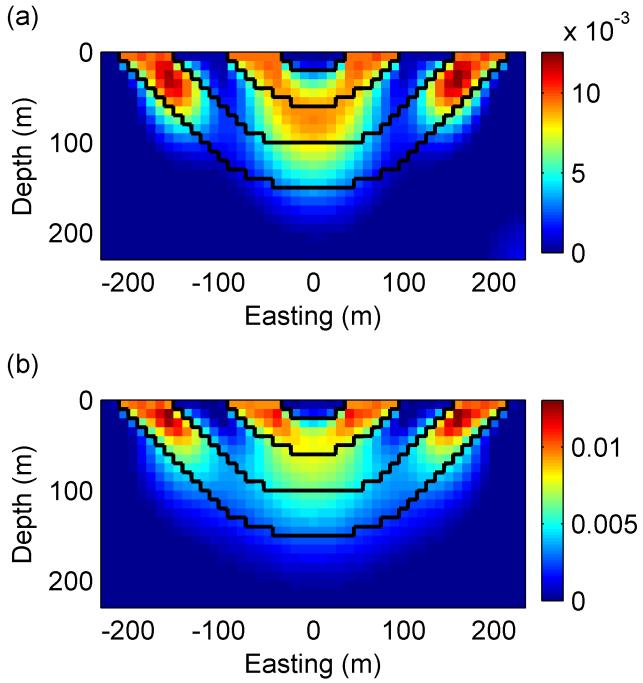


Figure 4.13: The recovered 2D susceptibility models with the following information incorporated into the inversions: (a) surface bounds; (b) surface bounds and surface orientations.

4.4.1 Incorporating surface orientation information

Surface mapping can provide direct local measurements of structural orientation. We now assume that surface mapping has provided physical property measurements and orientations at the surface which we can incorporate into the inversion. We set the susceptibility bounds across the surface equal to the true values $\pm 5.0 \times 10^{-4}$, which corresponds to 5% of the true value of the susceptible layers. The result with surface bounds included is in Figure 4.13(a).

In a real scenario we would need to make a decision on how deep to extend the orientation information. Here we take the approach of pushing the orientations to depth but weighting those orientations higher near the surface: this is done by setting $w_{x'}/w_{z'} = 100$ (where x' is in the along-dip direction) at the surface and decreasing that ratio to unity (i.e. no preferred elongation direction) at depths greater than five cells (50m). The result is shown in Figure 4.13(b). Incorporating the surface orientation information improves the result slightly. The upper susceptible layer is better recovered but the model still does not clearly indicate the presence of two magnetic layers at depth.

4.4.2 Incorporating interpreted volumetric orientation information

After surface mapping, subsequent drilling may allow interpretation of approximate orientations across larger volumes. Assuming that a drill-hole has been placed vertically through the centre of the model to a depth of 190m, we may then make an interpretation of a synclinal structure and develop a preliminary model of the subsurface dip as shown in Figure 4.14(a).

Without further investigation we would want to limit the weighting on those interpreted orientations as there is less confidence in them than for the measured orientations. We include the interpreted dip information in Figure 4.14(a) and set $w_{x'}/w_{z'} = 100$ near the surface and drill-hole location but decrease the ratio to 10.0 away from those locations to provide a lower weighting on the interpreted dips where we have less confidence. The $w_{x'}/w_{z'}$ ratios used across the volume are shown in Figure 4.14(b).

We also assume that the drill-core is logged with physical property measurements allowing us to include bounds in the cells along the drill-hole (we again set the bounds equal to the true susceptibility values $\pm 5.0 \times 10^{-4}$). Furthermore, we note that in previous inversions there has been a tendency to put higher susceptibility close to the surface. Hence, we limit the maximum susceptibility to be 0.0105 (the same as the largest upper bound for the surface and drill-hole cells).

The result with only the surface and drill-hole bounds (i.e. no orientation information) is in Figure 4.15(a). The result with the interpreted volumetric orientation information incorporated is in Figure 4.15(b). Incorporating this orientation information clearly improves the result and now indicates the presence of two distinct magnetic layers.

4.4.3 Incorporating orientation information from a rock model

Even without explicit physical property values attached, a geologic (rock) model still contains valuable orientation information. In this final iteration, we assume that a geologic model has been created in the later stages of exploration. Assuming that this model confidently locates the interfaces between the true rock units, we can calculate orientations associated with the interfaces and assign those orientations to the cells adjacent to the interfaces. We can then interpolate orientations in cells between the interfaces. The resulting orientations for our syncline example are shown in Figure 4.16(a).

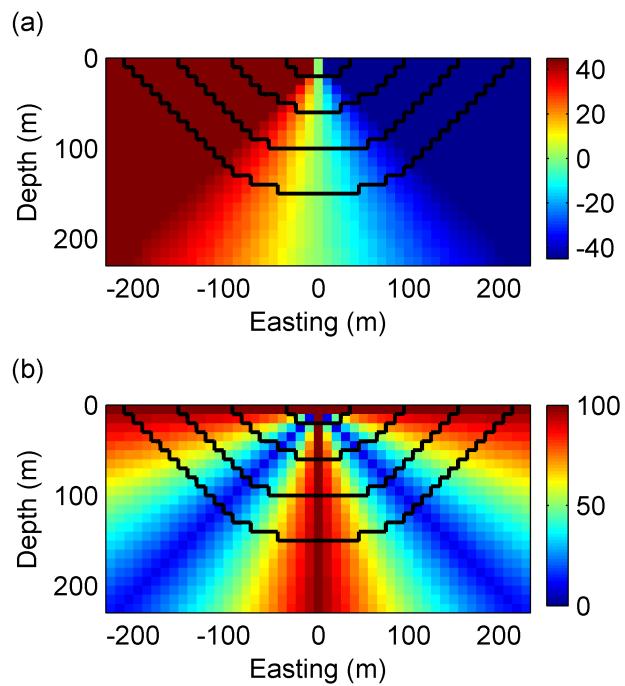


Figure 4.14: Interpreted dips (angles in degrees) for the 2D synclinal model are in (a). The $w_{x'}/w_{z'}$ ratios used across the volume for the result in Figure 4.15(b) are in (b).

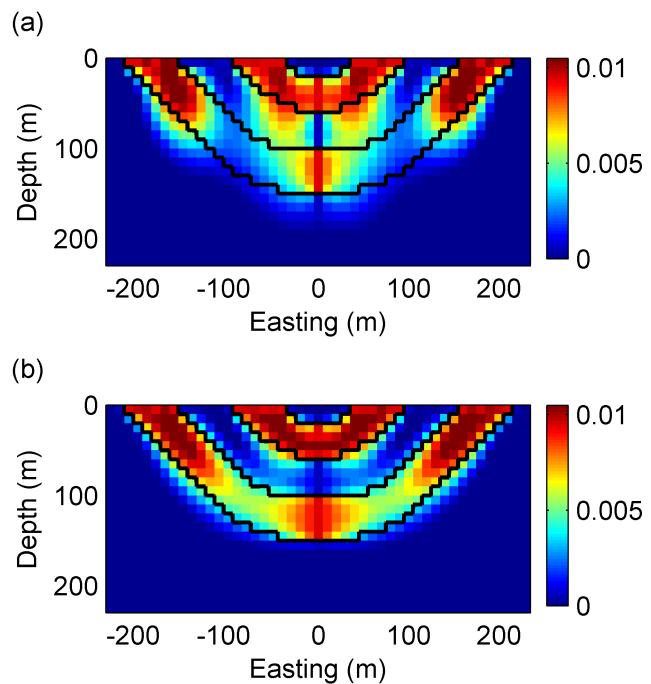


Figure 4.15: The recovered 2D susceptibility models with the following information incorporated into the inversions: (a) surface and drill-hole bounds; (b) surface and drill-hole bounds, and interpreted orientations at depth.

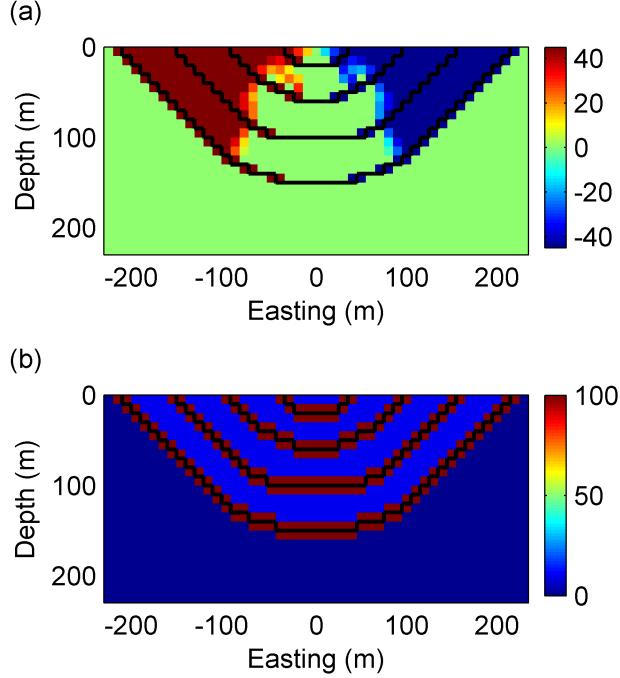


Figure 4.16: Calculated and interpolated dips (angles in degrees) for the 2D synclinal rock model are in (a). Cells where dip can not be calculated or interpolated (i.e. below the bottom-most susceptible layer) have been set to zero. The $w_{x'}/w_{z'}$ ratios used across the volume for the result in Figure 4.17(a) are in (b).

We set $w_{x'}/w_{z'}$ to a high value ($= 100$) in cells adjacent to the interfaces, to a lower value ($= 10.0$) where dips have been interpolated between interfaces, and to unity elsewhere. The $w_{x'}/w_{z'}$ ratios used across the volume are shown in Figure 4.16(b). If the geologic knowledge for a particular scenario leads us to believe that there is a more gradational physical property change (i.e. a less sharp interface) between rock units then this can be specified by decreasing the $w_{x'}/w_{z'}$ ratio closer to unity.

The inversion result with orientation information from the rock model incorporated is shown in Figure 4.17(a). The layers are well recovered, as expected since the rock model contains the true interface locations.

An alternative or additional strategy is to add linear constraints to the problem. We now add additional linear constraints of the form in equation (4.59):

$$-0.1m_j \leq m_i - m_j \leq 0.1m_j \quad (4.60)$$

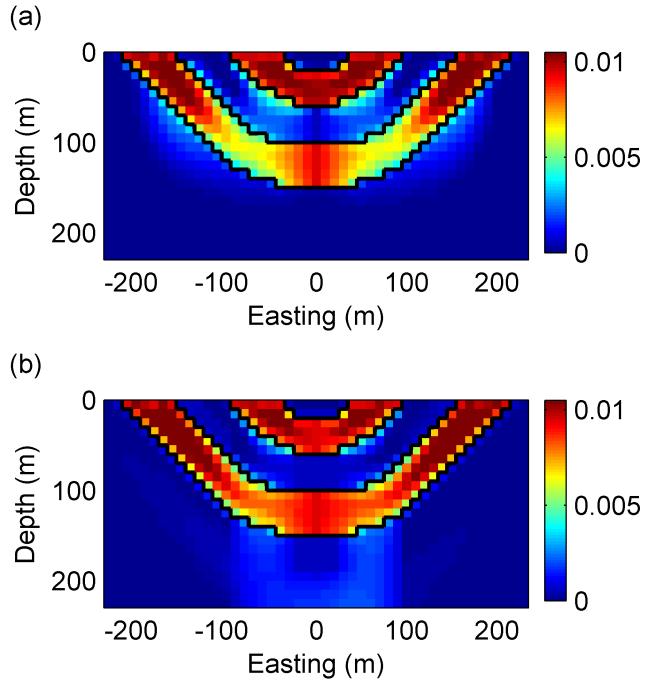


Figure 4.17: The recovered 2D susceptibility models with surface and drill-hole bounds, and interface location information incorporated into the inversions. The inversion for (a) had bound constraints only. Additional linear constraints were incorporated for (b).

which specifies a relative difference of 10% between the values in cells i and j . Between the easting coordinates -25m and $+25\text{m}$ we specify that cells horizontally adjacent to each other should obey equation (4.60) (i.e. such that the information along the drill-hole trace is extended out horizontally). To the west of -95m and to the east of $+95\text{m}$ we expect features that dip at $+45^\circ$ and -45° respectively, and we therefore specify that the diagonally adjacent cells along those dip directions should obey equation (4.60). The result with these linear constraints is shown in Figure 4.17(b), which provides further improvement.

4.5 Application to the San Nicolás deposit

Phillips (2001) performed considerable work on geophysical data from the San Nicolás massive sulphide copper-zinc deposit (Zacatecas, Mexico). Here we apply our methods to the gravity data therein. The inversion mesh is shown in Figure 4.18 and the data is plotted in Figures 4.18 and 4.19(a). There is a detailed geologic model available, interpreted from an extensive drilling

program, that we can use to constrain the inversions. Our first step is to create a density model from the geologic model and the physical property information available. That model is shown in Figure 4.20; the high density sulfide body is evident (red in those images). The deposit is bounded to the east by a southwest-dipping fault. Mineralization continues along the fault to depth to create a smaller keel structure that is evident on the left of Figure 4.20(a) and 4.20(b).

We next calculate the spatial model gradient in each cell for the model in Figure 4.20. High values of the gradient amplitude will occur where the interfaces in the geologic model exist (i.e. between rock units). Where the gradient amplitude is above some threshold we incorporate orientation information. The threshold value is determined by looking at isosurfaces and finding a threshold value for which the isosurface defines a fully connected set of interfaces for the main sulphide body and the keel. We use the orientation of the gradient vectors to specify rotation angles such that we can align one axis normal to, and two axes tangential to, the planar interfaces in the geologic model. If the x' axis is normal to an interface then we set $w_{x'} = 0.01$ and $w_{y'} = w_{z'} = 1.0$ to encourage the inversion to place sharp jumps in density across (normal to) the interface and maintain smoothness along (tangential to) the interface. Figure 4.21 indicates where the orientation information is applied (i.e. where the $w_{x'}$ smoothness weights are set low).

Figures 4.22 and 4.23 show cross-sections through recovered models obtained through inversions without and with orientation information incorporated respectively. The predicted data for those models are shown in Figure 4.19. The recovered model with orientation information incorporated better emphasizes the distinct high density sulphide body by placing sharp density jumps across the interfaces.

Due to corrections applied and regional components removed during data processing steps, it is difficult to compare the density values from the geologic model with those in the inversion results in an absolute sense. We can however compare the range of density values in the models (i.e. maximum value minus minimum value). Another issue is that forward modelling for the density model in Figure 4.20 (from the geologic model) creates a response with a range of approximately twice that of the observed data (see Figure 4.24). The density range for the geologic model is 1.70 (g/cc), but considering the forward modelling results, a value of 0.85 is a better value to compare with our results. The unconstrained result has a density range of 0.51 compared to 0.84 for the constrained result. Hence, the constrained result provides significantly improved density estimates.

In the unconstrained result, there is no indication of the keel. In the constrained result, there

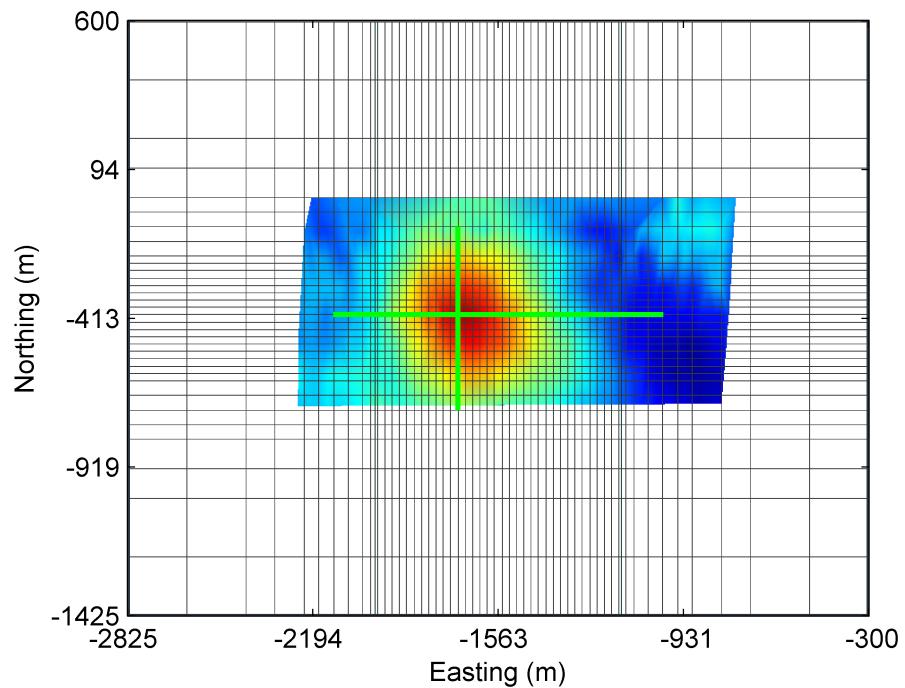


Figure 4.18: A top view of the mesh used to invert the San Nicolás gravity data, a map of which is overlayed. The mesh has 53-by-33-by-35 cells (easting-northing-depth). The locations of the cross-sections shown in the figures that follow are indicated with green lines.

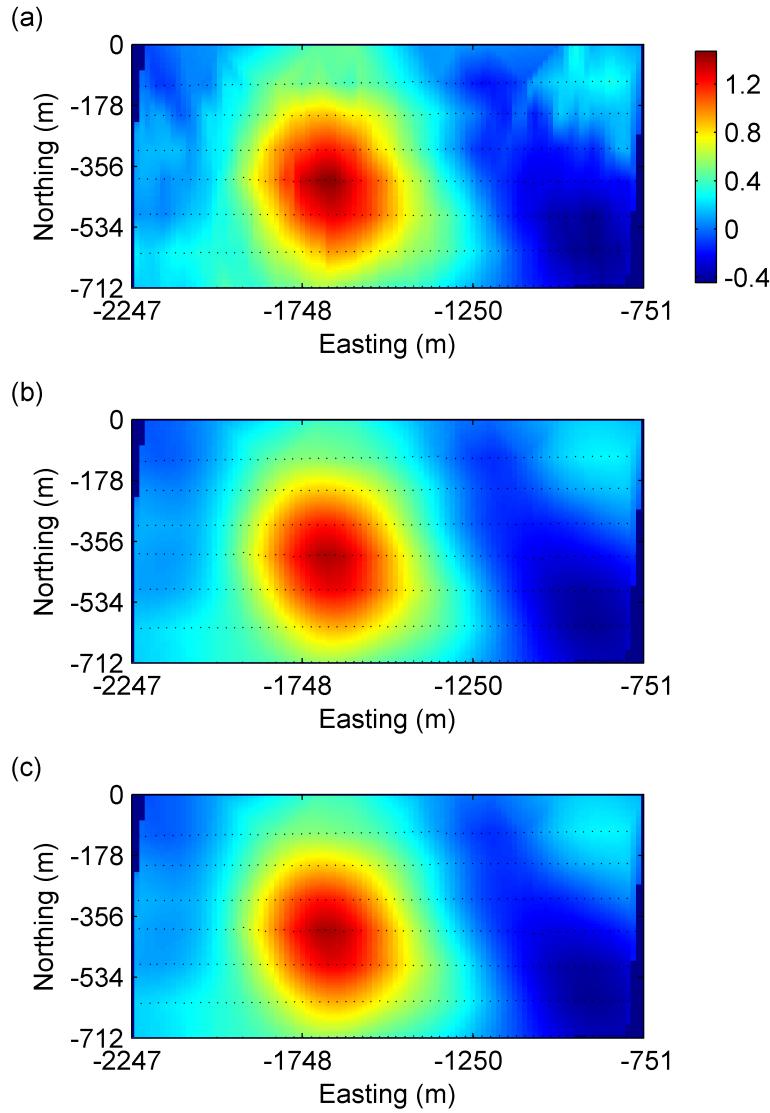


Figure 4.19: A map of the San Nicolás gravity data (units of mGal) is in (a); a regional component has been removed from the original data used in Phillips (2001). The predicted data for the model in Figure 4.22 is in (b). The predicted data for the model in Figure 4.23 is in (c). All three data maps are plotted on the same scale for comparison. The locations of the 422 data are indicated by black dots.

is a density structure recovered where the keel is expected to lie; this is most evident in Figure 4.25(b). The keel is difficult to model due to two factors. First, forward modelling experiments in Phillips (2001) showed that the gravity response of the keel is expected to lie only slightly above the estimated noise level for the data. Hence, there is only minimal data support for the keel and we can not expect the inversions to recover this deep structure well. Second, the discretization used results in mesh cells that are larger than some smaller spatial dimensions of the keel. If an inversion on this mesh happens to recover a structure indicative of the keel, we would expect upscaling of the keel structure onto the larger mesh cells to cause lower densities than expected (i.e. a small, high density structure becomes a larger, lower density structure once averaged onto a larger volume). This explains the lowered recovered density for the keel in the constrained result.

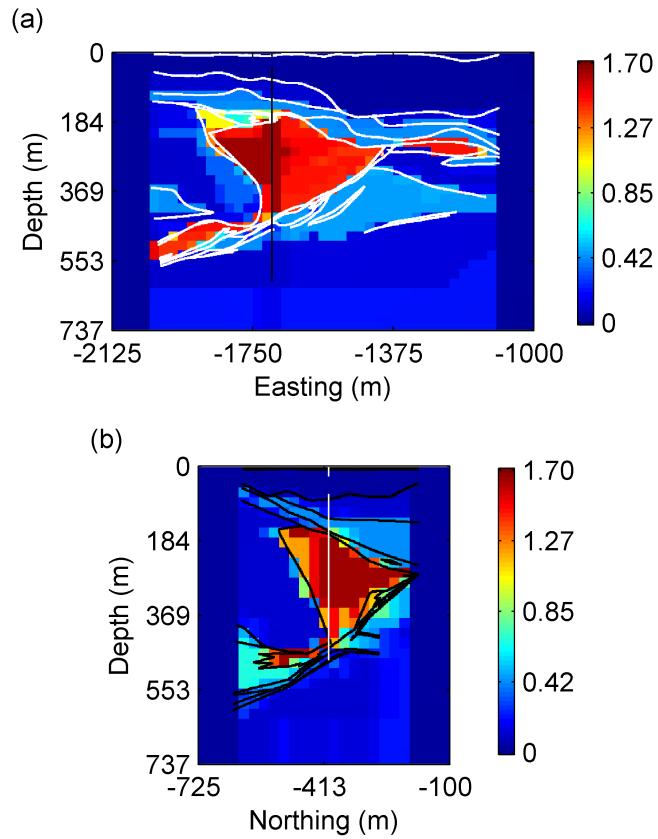


Figure 4.20: The anomalous density model (units of g/cc) created from the geologic model and physical property information for the San Nicolás deposit: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.

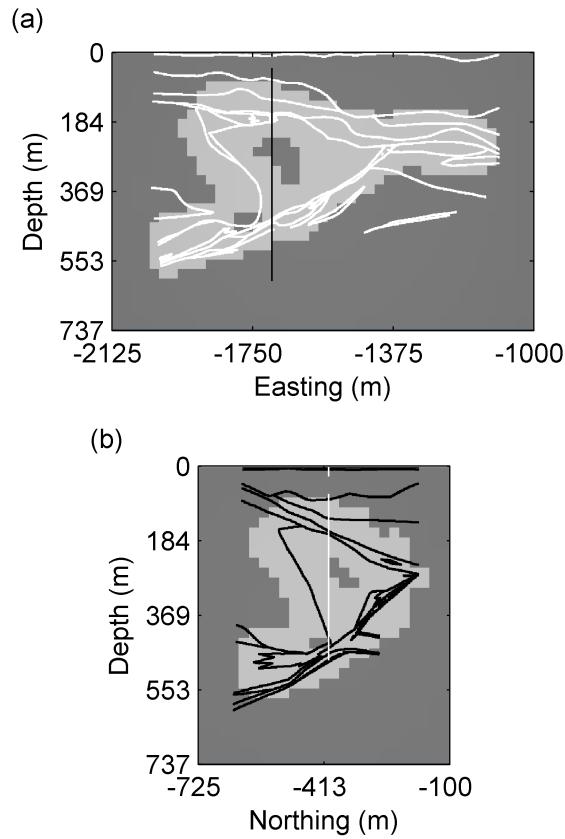


Figure 4.21: The $w_{x'}$ smoothness weights used in the inversion of the San Nicolás data with orientation information incorporated: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section. Weights of 1.0 are shown in dark grey and weights of 0.01 are shown in light grey.

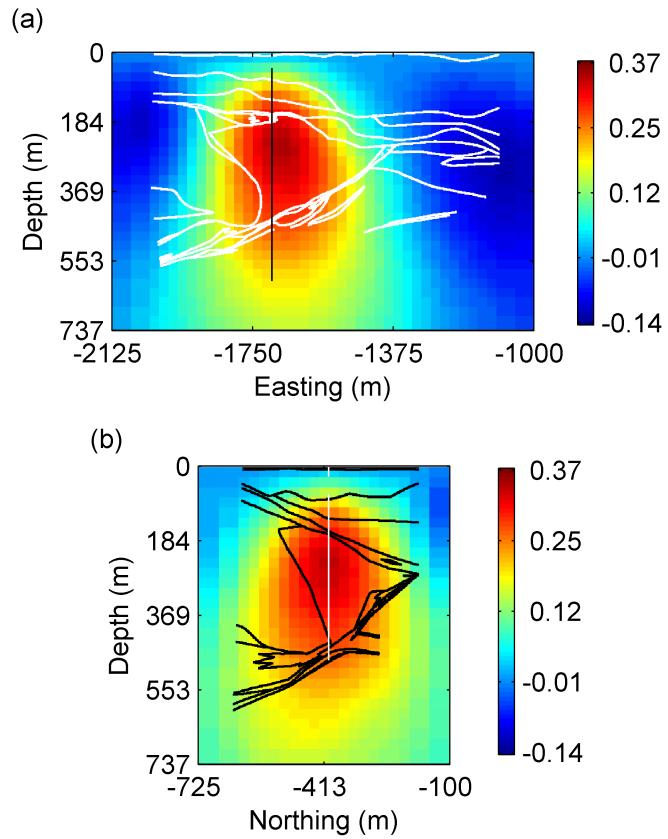


Figure 4.22: The recovered density model (units of g/cc) for the default, unconstrained inversion of the San Nicolás data: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.

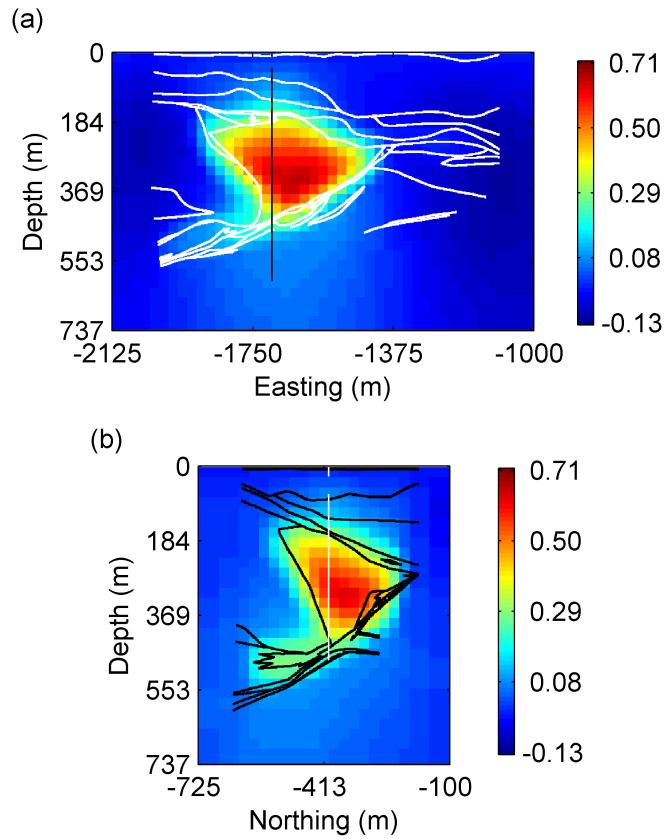


Figure 4.23: The recovered density model (units of g/cc) for the inversion of the San Nicolás data with orientation information incorporated: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.

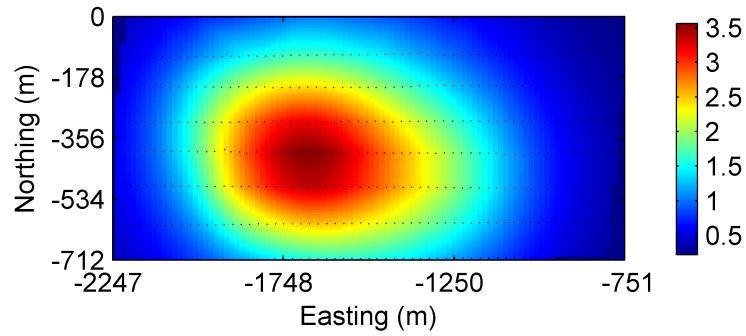


Figure 4.24: A map of the gravity data (units of mGal) forward modelled for the model in Figure 4.20. The locations of the 422 data are indicated by black dots.

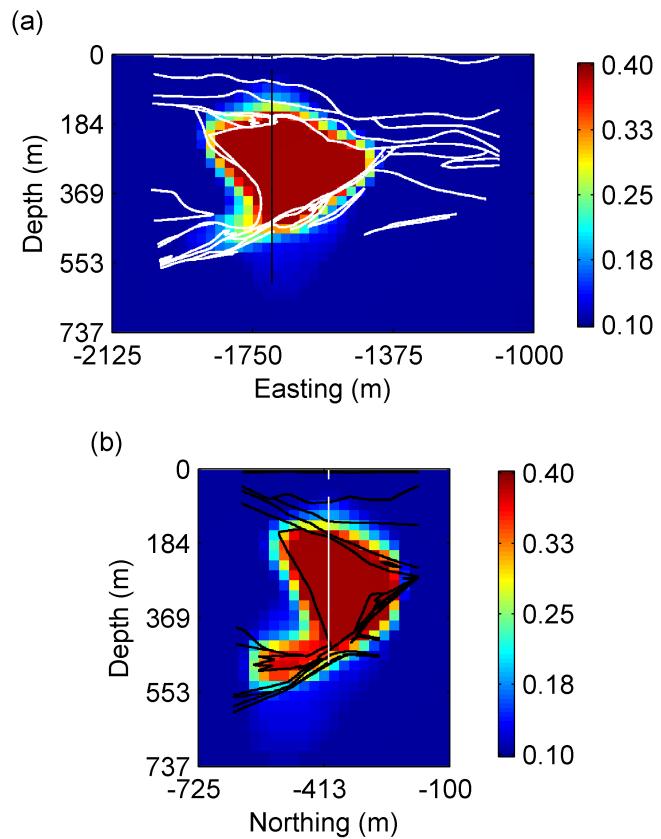


Figure 4.25: This figure shows the same information as in Figure 4.23 but the colour scale has been altered to better emphasize the keel structure at depth.

Region	Strike	Dip
1	115	80
2	71	90
3	115	90

Table 4.2: Strike and dip values (in degrees) assigned to each region indicated in Figure 4.26.

4.6 Application to the Hislop deposit

Mitchinson (2009) applied geophysical inversion to the Hislop gold deposit (eastern Timmins, Ontario, Canada) to help target Archean orogenic gold mineralization. Here we apply our methods to the magnetic data therein. The inversion mesh is shown in Figure 4.26 and the data is plotted in Figures 4.26 and 4.27(a). The Hislop deposit is hosted within a structurally complicated area, characterized by numerous faults and tight folds. Some larger faults define the edges of a high susceptibility region consisting of Fe-rich mafic and ultramafic volcanic rocks. Gold tends to occur in proximity to faults, making them important structural aspects of the subsurface.

Knowledge of the area has lead geologists to expect elongated tabular features that are steeply dipping and strike in different directions in different regions of the subsurface. Three regions are indicated in Figure 4.26 and Table 4.2 gives the strike and dip values assigned to each region. This information is used to rotate the smoothness directions across the inversion mesh. The smoothness weights are set to $w_{x'} = w_{z'} = 1$ and $w_{y'} = 0.01$ over the entire inversion volume to encourage tabular features that extend in the strike (x') and dip (z') directions. The recovered models without and with this information incorporated into the inversions are shown in Figure 4.28(a) and 4.28(b) respectively. The predicted data for those models are shown in Figure 4.27(b). The most obvious difference between the two recovered models is the change in shape of the highest susceptibility feature at centre, which becomes laterally more narrow once orientation information is incorporated. Another significant difference is the orientation of a lower susceptibility near-surface feature to the south of that body, which dips towards the south in Figure 4.28(a) but is nearly vertical in Figure 4.28(b).

Two large faults of interest are indicated in Figure 4.28, the depth traces of each having been interpreted from the recovered susceptibility models. Most gold deposits in the area are focused along the regional crustal-scale Porcupine-Destor Fault (shown in white in Figure 4.28); knowledge of its location and orientation is important for understanding the regional geology

4.6. Application to the Hislop deposit

and tectonics, and for focusing exploration programs. The local fault (which has no official name; shown in black in Figure 4.28) is where the majority of the gold is localized at the Hislop site; understanding its orientation and extent has implications on subsequent drill-hole spotting and mine planning.

The two interpretations for the depth trace of the local fault show different dip, especially closer to the surface. This indicates that more geologic information needs to be gathered (e.g. via drilling) to validate one interpretation or the other. One could also carry out more inversions with different local orientations and weights to assess what orientations of the local fault are reasonable from a data perspective. In contrast, the interpreted depth traces of the Porcupine-Destor Fault are similar, suggesting that this consistent depth trace is required by the data and providing confidence in the interpretation. The magnetic inversions have been helpful in mapping the Porcupine-Destor Fault location to depths below the limits of drilling.

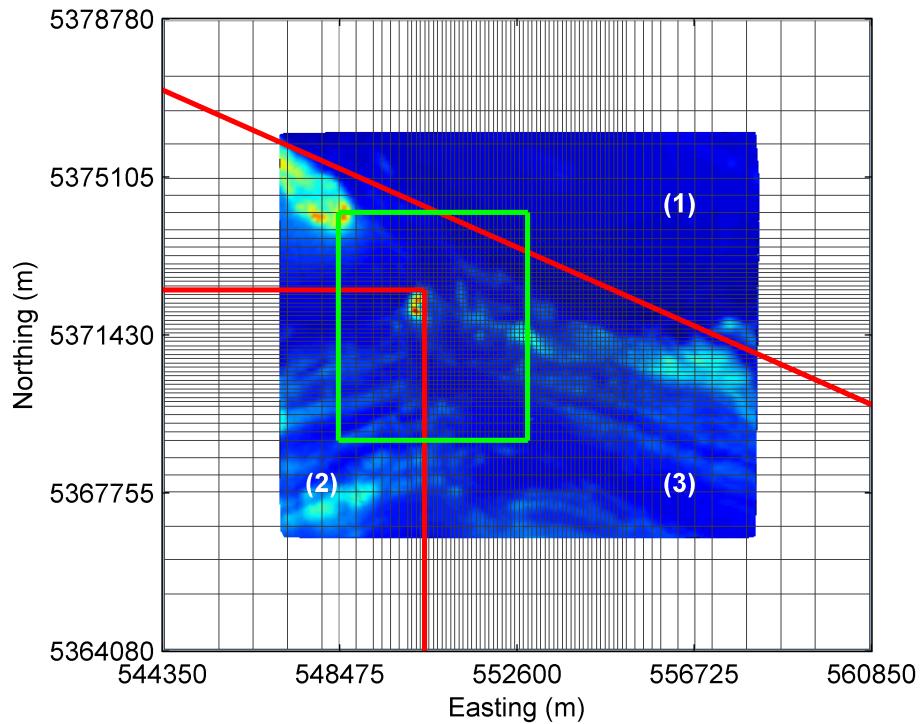


Figure 4.26: A top view of the mesh used to invert the Hislop magnetic data, a map of which is overlaid. The mesh has 73-by-55-by-40 cells (easting-northing-depth). Three regions are numbered and their boundaries indicated with red lines. The outline of the volumes shown in Figure 4.28 is indicated with a green rectangle.

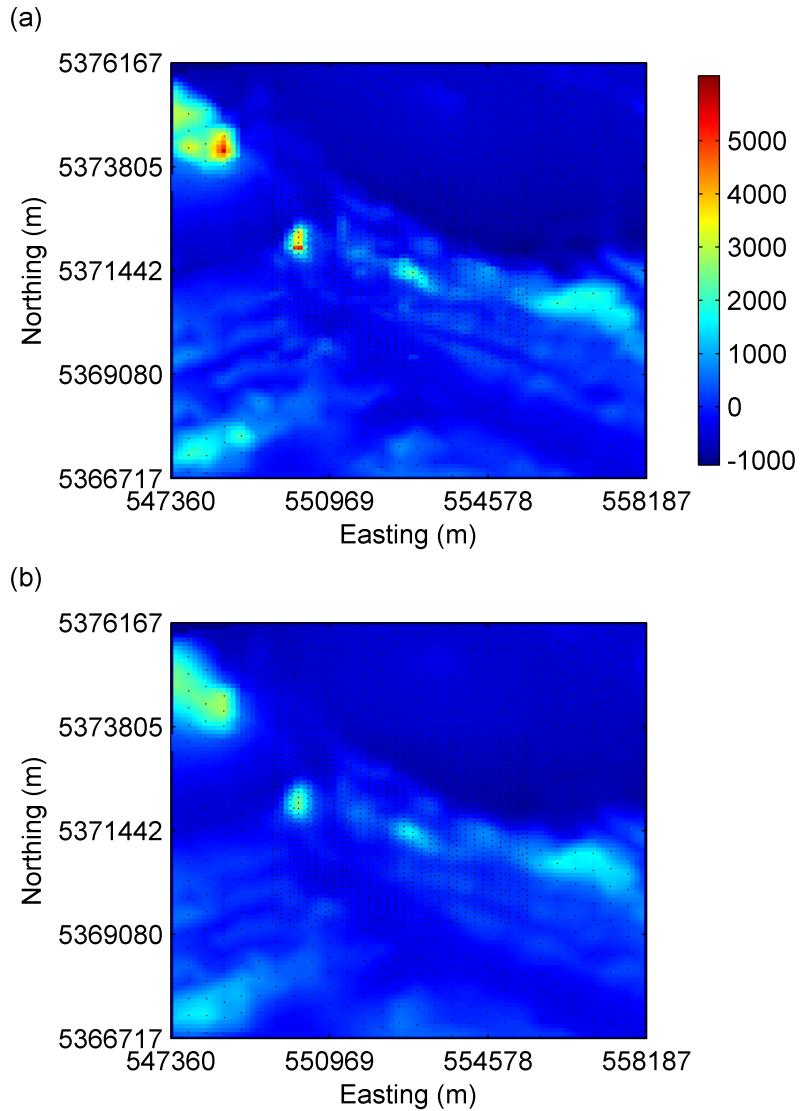


Figure 4.27: A map view of the Hislop magnetic data is shown in (a). The predicted data for the model in Figure 4.28(a) are in (b). The predicted data map for the model in Figure 4.28(b) is visibly indistinguishable from that in (b) so we do not show it. Both data maps are plotted on the same scale for comparison and the units are nT. The locations of the 1725 data are indicated by black dots.

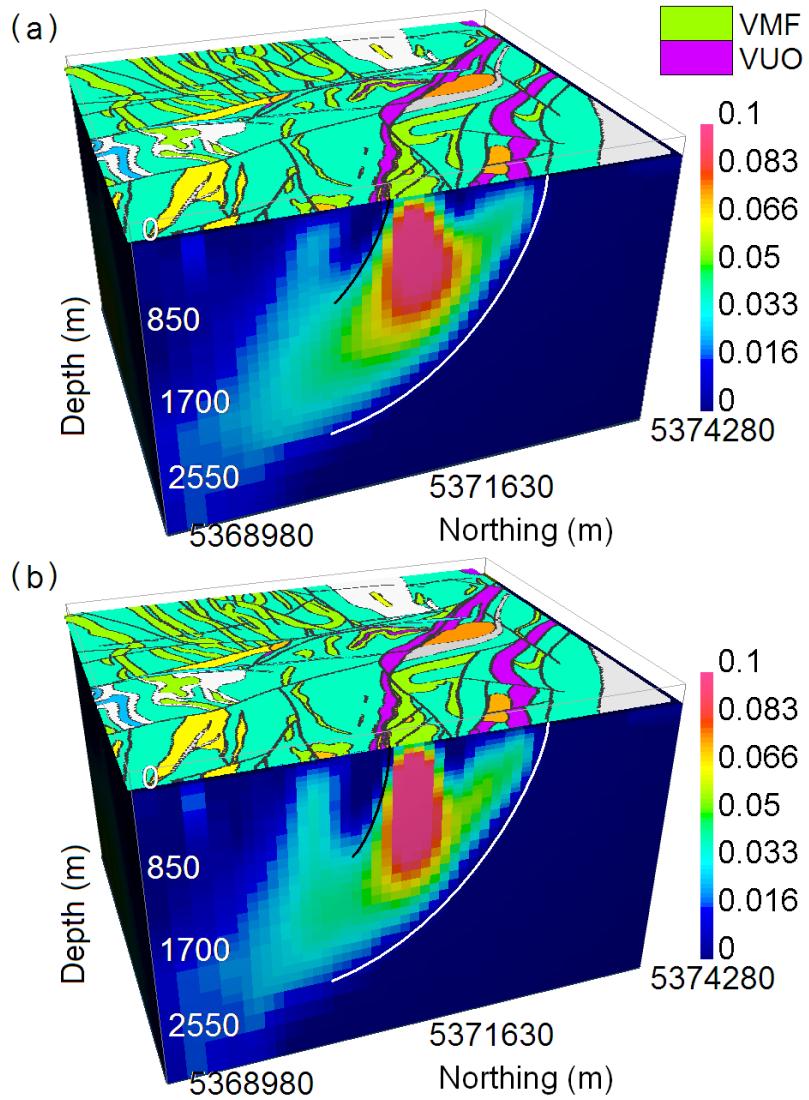


Figure 4.28: Vertical cross-sections at easting = 552850m through the susceptibility models (SI units) recovered from inversion of the Hislop data. The default, unconstrained inversion result is shown in (a). The result with orientation information incorporated is shown in (b). A geologic map, adapted from an interpretation by Power et al. (2004), has been overlaid on top of the mesh: the legend on top right in (a) indicates the colours of the magnetic mafic volcanic rocks (VMF) and ultramafic volcanic rocks (VUO). Interpreted depth traces for two faults are indicated with white and black lines: the Porcupine-Destor Fault is white and the local fault (no official name) black.

4.7 Conclusion

Incorporating orientation information into geophysical inversions can significantly improve the results, especially for gravity and magnetic problems which have poor resolution at depth, and we have provided a comprehensive look at the available methods for including this information. We have improved upon the work of Li and Oldenburg (2000) for including structural orientation information in geophysical inversions by developing a finite difference scheme for the numerical derivatives which ameliorates problems of asymmetry evident in the original implementation. We have also developed an approach that relies on additional linear constraints placed in the optimization problem.

4.8 References

- Barbosa, V. C. F., and J. B. C. Silva, 1994, Generalized compact gravity inversion: *Geophysics*, **59**, 57–68.
- Barbosa, V. C. F., and J. B. C. Silva, 2006, Interactive 2D magnetic inversion – a tool for aiding forward modeling and testing geologic hypotheses: *Geophysics*, **71**, L43–L50.
- Bosch, M., A. Guillen, and P. Ledru, 2001, Lithologic tomography – an application to geophysical data from the Cadomian belt of northern Brittany, France: *Tectonophysics*, **331**, 197–227.
- Chasserau, P., and M. Chouteau, 2003, 3D gravity inversion using a model of parameter covariance: *Journal of Applied Geophysics*, **52**, 59–74.
- Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure, *Geophysical Journal International*: **134**, 213–227.
- Gill, P., W. Murray, and M. Wright, 1995, *Practical Optimization*: Academic Press, London.
- Guillen, A., Ph. Calcagno, G. Courrioux, A. Joly, and P. Ledru, 2008, Geological modelling from field data and geological knowledge – Part II. Modelling validation using gravity and magnetic data inversion: *Physics of the Earth and Planetary Interiors*, **171**, 158–169.
- Guillen, A., and V. Menichetti, 1984, Gravity and magnetic inversion with minimization of a specific functional: *Geophysics*, **49**, 1354–1360.
- Last, B. J., and K. Kubik, 1983, Compact gravity inversion: *Geophysics*, **48**, 713–721.
- Li, Y., and D. W. Oldenburg, 1996, 3-D inversion of magnetic data: *Geophysics*, **61**, 394–408.
- Li, Y., and D. W. Oldenburg, 1998, 3D inversion of gravity data: *Geophysics*, **63**, 109–119.
- Li, Y., and D. W. Oldenburg, 2000, Incorporating geologic dip information into geophysical inversions: *Geophysics*, **65**, 148–157.
- Li, Y., and D. W. Oldenburg, 2003, Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method: *Geophysical Journal International*, **152**, 251–265.
- Mitchinson, D. E., 2009, Targeting Archean orogenic gold mineralization using physical properties and integrated geophysical methods, Doctoral Thesis, University of British Columbia, Canada.

- Mosegaard, K., and A. Tarantola, 2002, Probabilistic approach to inverse problems: International Geophysics, **81**, 237–265.
- Phillips, N. D., 2001, Geophysical inversion in an integrated exploration program – examples from the San Nicolás deposit, Master's Thesis, University of British Columbia, Canada.
- Power, W. L., D. Byrne, T. Worth, P. Wilson, L. Kirby, P. Gleeson, P. Stapleton, M. House, S. Robertson, N. Panizza, D. J. Holden, G. Cameron, R. Stuart, and N. J. Archibald, 2004, Geoinformatics evaluation of the eastward extension of the Timmins Gold Camp: Geoinformatics Exploration Inc., Unpublished report for St. Andrew Goldfields Ltd.
- Wijns, C., and P. Kowalczyk, 2007, Interactive geophysical inversion using qualitative geological constraints: Exploration Geophysics, **38**, 208–212.

Chapter 5

An iterative cooperative inversion strategy for obtaining geologically realistic models⁴

5.1 Introduction

Physical property models recovered through inversion of geophysical survey data aid geologists in interpreting the lithology of the subsurface. A common approach to the geophysical inverse problem is to design a function to minimize that combines terms which control the data fit and the amount of model structure (e.g. see Li and Oldenburg (1996)). Typical minimum-structure inversions regularize the model structure using using an ℓ^2 -norm (i.e. sum of squares) applied to the components of the spatial model gradient. This leads to the recovery of models characterized by smoothly varying physical property distributions. Smooth minimum-structure models are deemed appropriate when little geologic knowledge is available because the inherent smoothness emphasizes the uncertainty of the result and the lack of resolution inherent in the inverse problem. Also, use of ℓ^2 -norms is attractive as they produce computationally simple optimization problems.

Unfortunately, smooth models do not generally fit with geologists' ideas about the subsurface, which can involve sharp interfaces (physical property discontinuities) between rock units (regions with relatively constant physical properties). Hence, there is potential to be gained from performing inversions that can create models with such characteristics where appropriate. One approach is to perform an inversion involving discrete allowed values of physical properties (for

⁴A version of this chapter is in preparation to be submitted to a peer-reviewed journal: Lelièvre, P. G., D. W. Oldenburg and N. D. Phillips, An iterative cooperative inversion strategy for obtaining geologically realistic models.

example see Musil et al. (2003) and Krahenbuhl and Li (2004)). Another approach would be to directly seek the contacts between proposed rock units by parameterizing the model in terms of the contacts between rock units, either as functions or as tesselated surfaces within a complicated tetrahedral mesh.

Those options require significant alteration to the inverse problem and lead to computationally difficult optimization problems. Instead, we choose to work within a traditional deterministic framework in which the Earth is discretized into a 3D mesh of rectangular prismatic cells and we assign physical property values in each cell from a continuous range of allowed values. Our goal in this paper is then to develop an inversion strategy within that framework that, without adding complication to the inverse problem, can better emphasize distinct rock units and the sharp interfaces between them such that the recovered models are more geologically realistic.

Bosch et al. (2001) and Guillen et al. (2008) work in a stochastic inversion framework that directly recovers rock type (i.e. a lithologic inversion) from a list of those assumed present. Their approach allows a straight-forward incorporation of multiple geophysical data types and petrophysical information linking the rock types and the physical properties involved. Although their approach produces geologically realistic models, their algorithm relies on random sampling methods that lead to much heavier computational costs than deterministic approaches that minimize a function. To re-emphasize, we choose to work within a traditional deterministic framework and produce methods that do not cause a significant increase in computational cost.

We first develop an iterative inversion procedure that can help emphasize and delineate sharp interfaces between known rock units. We then extend that to allow cooperative use of multiple data types. We develop our methods on synthetic problems before applying them to data from the San Nicolás deposit.

5.2 An iterative approach for recovering geologically realistic models

5.2.1 Our deterministic inversion framework

In our deterministic inversion framework, we design an objective function, Φ , to minimize based on that of Li and Oldenburg (1996). A data misfit term, Φ_d , controls how well the data are fit and a regularization term, Φ_m , provides control over the amount and type of structure in the

recovered model, m :

$$\Phi(m) = \Phi_d(m) + \beta\Phi_m(m). \quad (5.1)$$

The trade-off parameter, β , allows us to fit the data to the desired degree (i.e. such that $\Phi_d(\mathbf{m})$ is equal to some target misfit Φ_d^*). The regularization term is of the form

$$\begin{aligned} \Phi_m(m) = & \int_V w_s (m - m_{ref})^2 dv \\ & + \int_V w_x \left(\frac{\partial m}{\partial x} \right)^2 dv \\ & + \int_V w_y \left(\frac{\partial m}{\partial y} \right)^2 dv \\ & + \int_V w_z \left(\frac{\partial m}{\partial z} \right)^2 dv. \end{aligned} \quad (5.2)$$

The first term measures the closeness of the model to the reference model m_{ref} (or in the absence of a reference model this term measures the amount of anomalous material in the subsurface). The latter three terms measure the smoothness of the model in each Cartesian direction.

The use of ℓ^2 -norms in such minimum-structure inversions results in smoothly varying models. Sharp boundaries can be generated using other measures on the individual smoothness terms, as implemented by Farquharson and Oldenburg (1998), or on the spatial model gradient $\vec{\nabla}m$ (e.g. the total variation approach developed by Rudin et al. (1992) and Vogel and Oman (1998)). One option for an alternate measure is the Ekblom measure

$$\rho(x) = \sum_i (x_i^2 + \epsilon^2)^{p/2} \quad (5.3)$$

where x is some vector of values and ϵ is some small value that ensures differentiability at zero (Ekblom, 1973; 1987). With $p \approx 1$ the Ekblom measure approximates an ℓ^1 -norm (i.e. sum of absolute values). With $p = 1$, the Ekblom measure applied to $\vec{\nabla}m$ yields the total variation measure. Use of such a measure complicates the optimization problem and we chose to explore how well sharp features can be encouraged using ℓ^2 -norms by employing an iterative strategy.

5.2.2 An overview of the iterative procedure

A diagram of our iterative procedure is shown in Figure 5.1. The procedure begins in step (1) with a model that is the best representation of the subsurface (e.g. after incorporating the available geophysical and geologic data through geologically constrained inversions). A measure of structure is then used to define a likelihood that sharp interfaces exist in particular regions in the model (steps (2) through (4)). The smoothness weighting is then altered in step (5) to encourage sharper interfaces in a subsequent inversion. The interface likelihood measure is used to guide how the smoothness weights are altered and the procedure is repeated iteratively. When a scenario exists in which sharp interfaces are only expected in some regions of the subsurface, the procedure can easily be performed on particular volumes of the subsurface model rather than on the entire volume. In such a case, it may even be advantageous to keep the rest of the subsurface model fixed.

Below we go into more detail on the individual steps of the procedure while running through an illustrative 2D synthetic cross-well seismic tomography example. The mesh for the example is a 32-by-18 grid of uniform equidimensional cells with 5.0m dimensions, as shown in Figure 5.2(a). The true model for this example is shown in Figure 5.2(a): the model contains a single body with an anomalous slowness (s/km) of 1.0 within a 0.0 background. Travel-time data are calculated for the true model with the ray-paths shown in Figure 5.2(b), which run from one side of the mesh to the other with sources and receivers located at depths every 15.0m from 7.5m to 82.5m. A small amount of random noise is added before inverting; that noise is taken from a normal distribution with zero mean and standard deviation equal to 2.0% of the range of the travel-time data. The default inversion result with $w_s = 0$ and ℓ^2 -norms on the smoothness measures is shown in Figure 5.2(c). The result with an Ekblom measure with $p = 1.1$ applied to the spatial model gradient $\vec{\nabla}m$ (which estimates a total variation measure) is shown in Figure 5.2(d).

Note that our iterative approach has similar elements to that of van Zon et al. (2007). They consider a situation in which there are only two units assumed present in the subsurface and their iterative strategy encourages those values to move towards the peaks in the prior bi-modal distribution obtained from petrophysical information. Because our strategy deals with interfaces (i.e. relative physical property changes) rather than absolute physical property values we can consider more than two rock units.

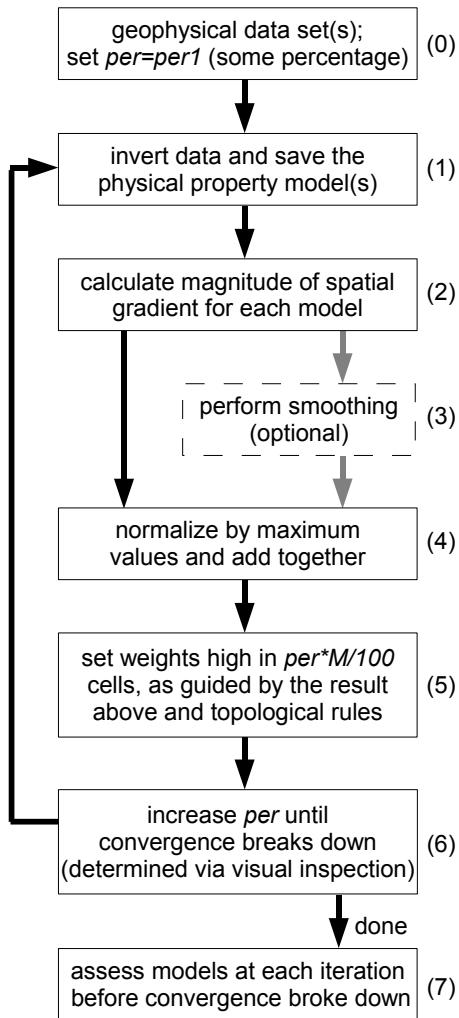


Figure 5.1: A flow diagram illustrating our iterative procedure. The number of mesh cells in the discretized Earth is denoted M .

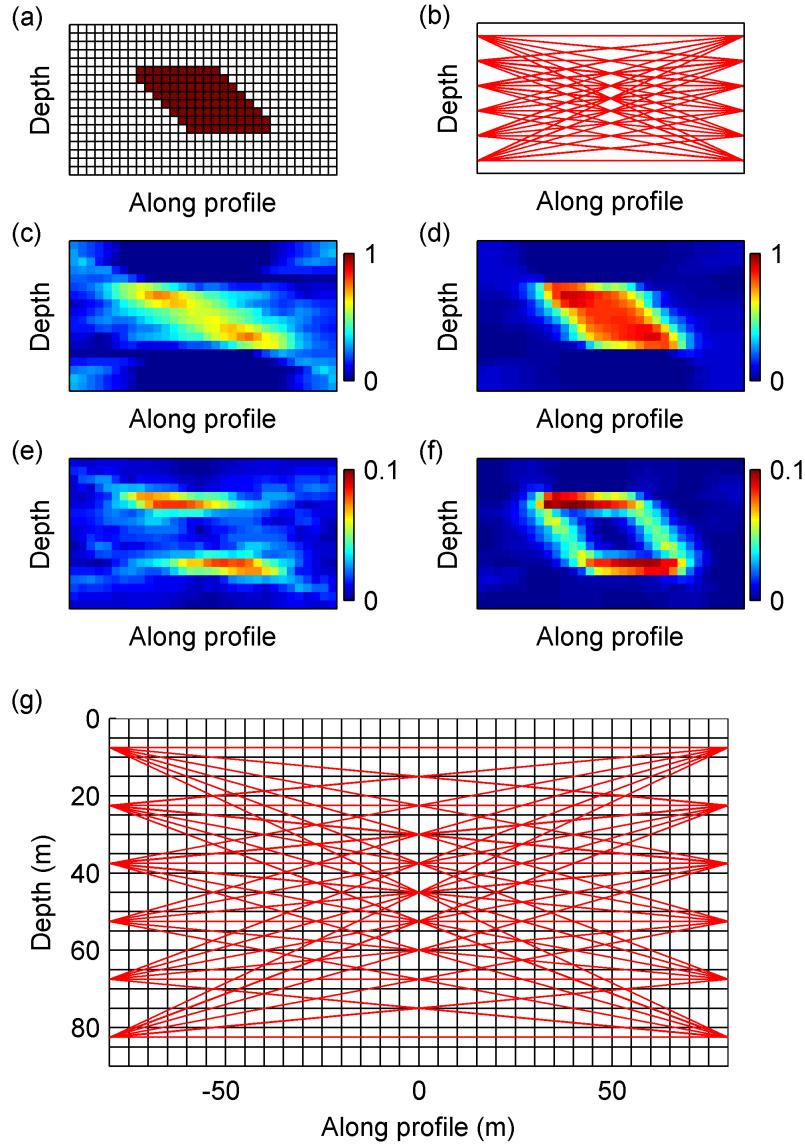


Figure 5.2: Plots for the first synthetic example: (a) shows the true slowness model (units of s/km) with the discrete mesh shown in black; (b) shows the ray-paths through the volume; (c) shows the inversion result with $w_s = 0$ and ℓ^2 -norms on the smoothness measures; (d) shows the inversion result with an approximate total variation measure (Ekblom with $p = 1.1$ applied to $\vec{\nabla}m$); (e) and (f) show the magnitude of the model gradient for the models in (c) and (d) respectively. To remove clutter, we do not show axis labels in (a) though (f) but we provide a labeled version of the inversion mesh in (g) with the ray-paths overlaid in red.

5.2.3 Measuring model structure

There are numerous possible definitions for structure. We define structure as the change in the model with position. The magnitude of the spatial model gradient is then a natural choice to use to provide a quantitative scalar measure of structure:

$$\|\vec{\nabla}m\| = \sqrt{(\nabla_x m)^2 + (\nabla_y m)^2 + (\nabla_z m)^2}. \quad (5.4)$$

Our iterative strategy then relies on the assumption that it is more likely that sharp interfaces between rock units lie in regions of the model in which $\|\vec{\nabla}m\|$ is high. This assumption will be valid provided the geophysical survey methods are sensitive enough to resolve the interfaces. Otherwise our procedure may force the model towards an invalid result. Improving the starting model by incorporating available geologic information into the procedure can help avoid that situation.

In step (2) of the procedure shown in Figure 5.1 the gradient magnitude in equation (5.4) is calculated on the discrete mesh. This is done using the same finite difference gradient operators used to discretize the regularization smoothness terms in equation (5.2). If those operators calculate gradients on cell faces (the standard approach) then interpolation is required to calculate values at cell centres:

$$\|\nabla \mathbf{m}\| = \sqrt{\mathbf{Q}_x((\mathbf{G}_x \mathbf{m})^2) + \mathbf{Q}_y((\mathbf{G}_y \mathbf{m})^2) + \mathbf{Q}_z((\mathbf{G}_z \mathbf{m})^2)} \quad (5.5)$$

where \mathbf{m} is the discretized model, \mathbf{G}_x is the x -direction gradient operator for cell faces, and \mathbf{Q}_x interpolates the x -direction gradients at cell centres (and similarly for the other terms). After smoothness weights are calculated in the next step of the process, those weights must again be interpolated back to the cell faces.

The gradient magnitude for the model in Figure 5.2(c) (the ℓ^2 -norm result) is shown in Figure 5.2(e). Note that, as expected, the gradient magnitude is high near the true interfaces; there is not one-to-one correspondence but we hope that the iterative procedure can improve the delineation of the true interfaces. For comparison, in Figure 5.2(f) we show the gradient magnitude for the model in Figure 5.2(d) (the approximate total variation result).

5.2.4 The iterative weighting procedure

The essence of our iterative procedure is to set the smoothness weights in equation (5.2) lower where contacts between rock units are expected. This will encourage the inversions to place sharper physical property jumps at those locations and maintain smoothness elsewhere (i.e. within the rock units). There are three questions that need to be answered: 1) where should the smoothness weights be set low, 2) over how much of the volume should the smoothness weights be set low, and 3) what should the relative weights (high and low) be?

The gradient magnitude measure in equation (5.4) can be used as a proxy for interface likelihood, which provides an idea of where the smoothness weights should be set low. However, we also need a way to determine over how much of the volume to set the smoothness weights low. Two options are to use a threshold (i.e. all cells with gradient magnitude higher than the threshold are weighted low) or use some volume percentile (i.e. some percentile of the cells with the highest gradient magnitude values are weighted low). At this stage one would have a rough idea of where the interfaces may lie in the volume. A hypothetical true Earth model (i.e. a geologic model) can be developed that contains contacts between constant rock units. This model can then be discretized on the inversion mesh and relative number of cells adjacent to the interfaces can be calculated. That provides an appropriate estimate for the volume percentile over which to set the smoothness weights low. It is more difficult to assign a threshold value since the scaling of the gradient magnitude measure will change as the iterations proceed. Therefore we take the volume percentile approach.

It is important that the iterative procedure be performed in an automated yet controlled fashion such that it can reliably delineate and emphasize the sharp interfaces. To do so we view our procedure as one which forces regions to be smooth and thereby forces other regions to alter as required (e.g. by becoming less smooth and exhibiting sharper interfaces). In the initialization step (0) shown in Figure 5.1, the procedure starts with a *low* volume percentile over which the smoothness weights are set *high*. Doing so will encourage only small regions of the model to become smoother (i.e. more constant). That volume percentile is then slowly increased (step (6) in Figure 5.1) such that larger regions of the model are encouraged to become smooth, and smaller regions of the model may contain sharp features. Performing the procedure in this controlled manner provides the inversions some flexibility to ignore the currently estimated interface delineation (at the current iteration) if required in order to fit the data. It is important to note that, generally, the result obtained after several iterations is significantly different from the result had only a single iteration been performed with the same final volume percentile. The

procedure creates a sequence of models, all of which fit the data but have smoother or sharper characteristics. This model suite can be subsequently compared and assessed in light of the available geologic knowledge.

Using a higher relative difference between the high and low weighting values results in sharper interfaces. We typically use a value of 1.0 for the high weights and 0.01 for the low weights. However, the exact values chosen will depend on the desired sharpness of those interfaces, which should be based on the geologic knowledge and the ability of the geophysical data to resolve them.

The procedure makes adjustments to the smoothness weights such that some regions of the model are weighted lower than other regions. Including a smallness term in the iteratively performed inversions can produce unwanted effects since an adjustment to the smoothness weight for a particular cell alters the relative importance of the smallness term for that cell. Hence, we set $w_s = 0$ when performing the iterative procedure. Although this removes the ability to include a reference model, we are still able to apply bound constraints, which can incorporate similar information into the inversions as a reference model can (e.g. physical property sampling of the true Earth).

5.2.5 Stopping criteria

If a reliable estimate is available of the volume percentile of mesh cells adjacent to the interfaces in the true Earth, then this provides a good idea of where to stop the procedure. When increasing the volume percentile value over which the weights are set high, it is obvious when the percentile value gets too high: above some critical value the sharp interfaces break and the models change rapidly. If the procedure is run up to a volume percentile of 100% the final model will be identical to the initial model. We typically run the procedure past the volume percentile estimate for the true Earth and save the models at each step, allowing us to select a model before the critical percentile value is reached (which can be easily determined via visual inspection). To re-emphasize, the sequence of models created all fit the data but have smoother or sharper characteristics, and the final model (before the critical iteration is reached) may not necessarily be the best as assessed by a geologist with knowledge of the region.

Moving again to the cross-well seismic tomography example in Figure 5.2, we perform the procedure such that the volume percentile over which the weights are set high increases from 46% to 94% over five iterations (i.e. an increase of 12% each iteration). This rate of increase is

larger than we would suggest for application to real data but it will suffice for demonstration purposes. Models and weights at successive iterations are shown in Figure 5.3, and a measure of the progression of the procedure is shown in Figure 5.4. Note that we use the term “progression” rather than “convergence” because, strictly speaking, there is no final result that we are converging towards: we merely adopt a model prior to break-down after the critical iteration. The break-down of the method after passing the critical volume percentile is evident at iteration 5. For this example, the critical percentile is expected to lie at around 88% as that is the ratio of mesh cells adjacent to the interfaces in the true model. The accepted final model is that before reaching the critical point (i.e. the model at iteration 4 where the percentile value is 82%). That model provides a much better representation of the true model than that in Figure 5.2(c): the two rock units (i.e. the body and background) are clearly defined and the outline of the body has been well recovered.

5.2.6 Applying topological rules

An advantage of the approach of Bosch et al. (2001) and Guillen et al. (2008) is that they can define topological rules that can be used to immediately reject any geologically unreasonable models created in their random sampling process. An example of such a rule might be that two particular rock units in their lithologic model should not contact each other.

There is some ability to apply topological rules regarding interfaces into our iterative procedure. In step (5) in Figure 5.1, these rules could be combined with the interface likelihood measure to guide which model cells will receive lower smoothness weights. For example, consider a situation in which one wishes to use our procedure to delineate a contact between two regions belonging to some known rock units: this scenario is illustrated in Figure 5.5. The procedure should begin with low weights between the known regions and shrink the low-weighted region inwards until the contact is located, effectively pushing the boundaries of those confident rock unit regions towards each other slowly such that the interface is eventually located. In such a case one may want to apply a rule that stipulates that no pockets of high weights may be created between the known regions, i.e. the situation in Figure 5.5(b) would be acceptable but that in Figure 5.5(c) would not.

Applying topological rules to the weighting procedure is, in effect, like applying regularization to the character of the weighting. The appropriate regularization will be dependent on the application, but we have found empirically that maintaining a smooth weighting function can help the procedure progress (i.e. convergence prior to the critical iteration) more smoothly.

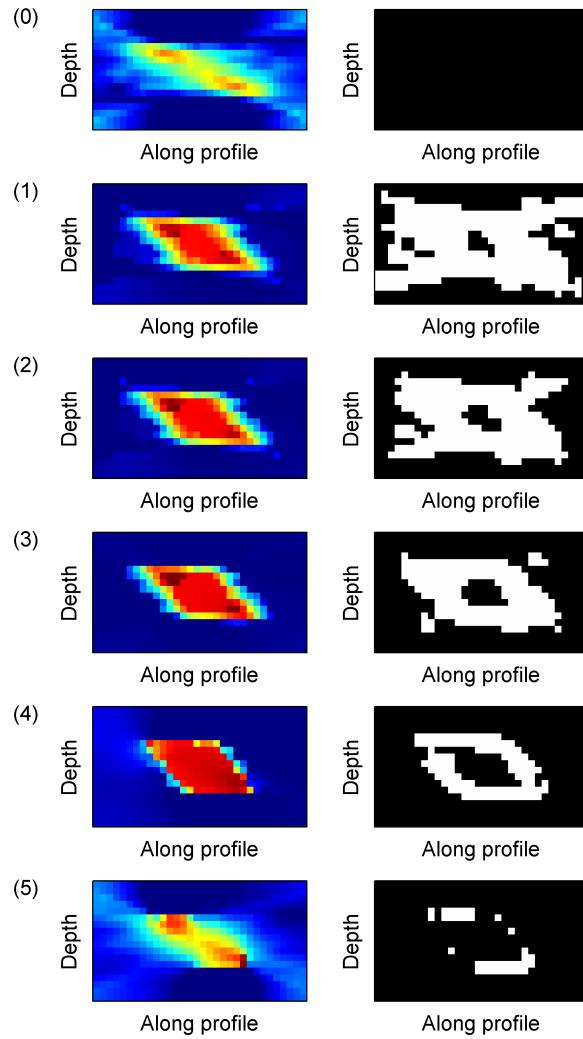


Figure 5.3: Results of our iterative inversion procedure for the first synthetic example with no smoothing applied. The recovered slowness models (units of s/km) are in the left column and weights applied in the inversions are in the right column. The numbers on the left indicate iteration number. The model at iteration 0 is the initial model recovered from an inversion with uniform weighting across the volume. The colour bar used is the same as that in Figure 5.2 (red high, blue low) and the colour scale for the models is on $[0, 1]$. The low weighting values are 0.01, indicated in white. To remove clutter, we do not show axis labels but we refer the reader to the labeled version of the inversion mesh in Figure 5.2(g).

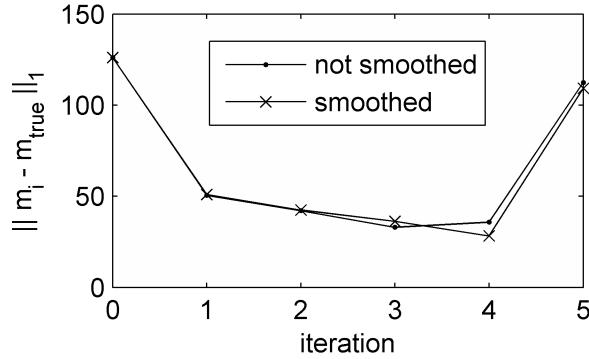


Figure 5.4: Progression curves for the iterative inversion procedure results in Figures 5.3 (black line with point markers) and 5.9 (grey line with cross markers). The progression measure is the ℓ^1 -norm of the difference between the model at each iteration and the true model.

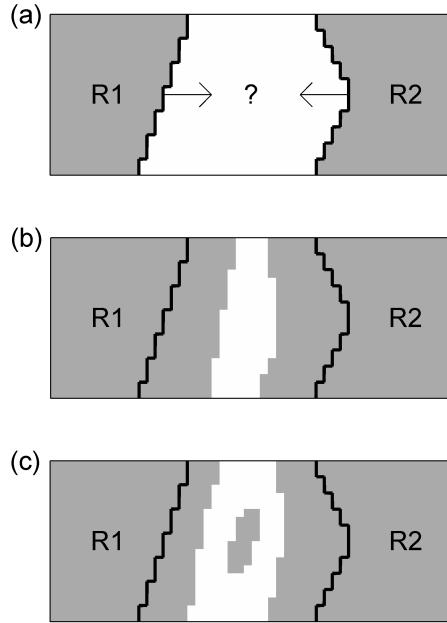


Figure 5.5: An illustrative schematic of our procedure. The volume of interest contains two rock units, R1 and R2. We are confident that the shaded regions in (a) contain the indicated rocks. Between those we are unsure and the essence of our procedure is to carefully expand those regions towards each other to obtain an improved estimate for the true interface location. In our inversions, the shaded regions would contain high smoothness weights and the white regions low weights. Two possible weighting scenarios with different topological characteristics are shown in (b) and (c).

Hence, we include the option to perform smoothing on the gradient magnitude values, calculated using equation (5.5), prior to using that information to set weights for the next inversion. This appears in step (3) in Figure 5.1.

This smoothing becomes important when the inverted models are expected to contain features at smaller scales than that of the interfaces of interest. As an example, consider again the 2D cross-well seismic tomography synthetic problem in Figure 5.2. The recovered model using ℓ^2 -norms, in Figure 5.2(c), contains such small scale features. Performing the iterative procedure with identical parameters but including the smoothing option yields the results in Figure 5.6. Visual comparison against the progression seen in Figure 5.3 indicates that the smoothing has a beneficial effect.

5.3 A more complicated example

We now present a more complicated synthetic example containing two regions, one containing a target body that we wish to remain smooth, and another containing a feature we wish to sharpen using our procedure. The true 2D density model is shown in Figure 5.7: the target body has a density of 1.0 and the sharper surface feather has a density of 0.2. All density values mentioned in this section have units of g/cc. Gravity data are collected across the top of the model at a height of 2.5m, every 5.0m along-profile from -200m to -200m. A small amount of random noise is added before inverting; that noise is taken from a normal distribution with zero mean and standard deviation equal to 2.0% of the range of the gravity data.

As mentioned previously, before applying our procedure it is important to improve the starting model by incorporating any available geologic information into the inversions. Geologic mapping can provide physical property bounds across the Earth's surface and knowledge of the rock types present in the subsurface can provide bounds over the entire volume. For the inversions leading to the results in Figure 5.7(b) and 5.7(c) (ℓ^2 and ℓ^1 -norm inversions respectively) we applied bounds in the cells at the surface of the mesh. Those surface bounds were equal to the true density values ± 0.02 (i.e. the bounds are [0.18, 0.22] on the left and [0.98, 1.02] on the right). In the rest of the volume we have set bounds on [0, 1.2].

Our iterative procedure is applied on the region to the left of the white line in Figure 5.7(a). The ℓ^2 -norm result in Figure 5.7(b) is used as the initial model, the gradients are smoothed during the re-weighting process, and the volume percentile (over which the weights are set high)

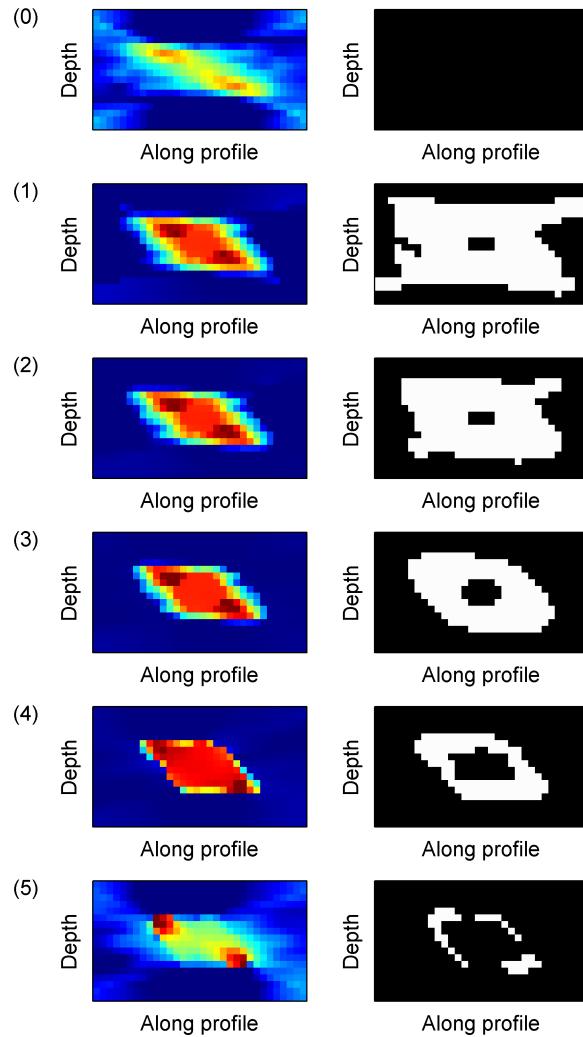


Figure 5.6: Results of our iterative inversion procedure for the first synthetic example with smoothing applied. The recovered slowness models (units of s/km) are in the left column and weights applied in the inversions are in the right column. The numbers on the left indicate iteration number. The model at iteration 0 is the initial model recovered from an inversion with uniform weighting across the volume. The colour bar used is the same as that in Figure 5.2 (red high, blue low) and the colour scale for the models is on $[0, 1]$. The low weighting values are 0.01, indicated in white. To remove clutter, we do not show axis labels but we refer the reader to the labeled version of the inversion mesh in Figure 5.2(g).

is increased from 24% to 84% over six iterations (i.e. an increase of 12% each iteration). The final model is shown in Figure 5.7(c). The sharp interface on the left of the mesh has been delineated within one mesh cell, and the density value of the unit above that interface is close to the true value of 0.2 (note that different colour-scales are used in Figure 5.7(a) and 5.7(c)). In doing so, we are left with a better recovery of the location of the body in the right side of the mesh (although that body has been further smoothed by the inversion and therefore displays a lowered amplitude).

5.4 Extension to allow cooperative use of multiple data types

5.4.1 Joint and cooperative inversion

The terms joint and cooperative inversion have been used in various ways in inversion literature. Here we use the term joint inversion to describe a method that simultaneously inverts multiple types of data over the same Earth region, each sensitive to a different physical property, while using some method to maintain similarity between the multiple physical property models involved. We use the term cooperative inversion to describe a method that inverts the multiple types of data individually but shares information between the inversion runs in order to improve the similarity of the multiple models.

By combining several different types of complimentary geophysical data through joint or cooperative inversion one hopes to reduce the ambiguity (due to nonuniqueness) of the inverse problem and enhance the inversion results. This becomes important whenever one type of data fails to adequately resolve the Earth, be it due to unpropitious survey aspects or lack of resolution inherent in the particular type of geophysical survey (e.g. the inherent lack of depth resolution for gravity and magnetics).

5.4.2 An overview of our cooperative strategy

Our iterative procedure for emphasizing and delineating sharp interfaces relies on the assumption that the gradient magnitude measure in equation (5.4) can be used as a proxy for interface likelihood. As previously mentioned, this assumption will be valid provided the geophysical survey methods are sensitive enough to resolve the interfaces. The philosophy of taking a cooperative approach is then that any individual set of data may provide reliable information

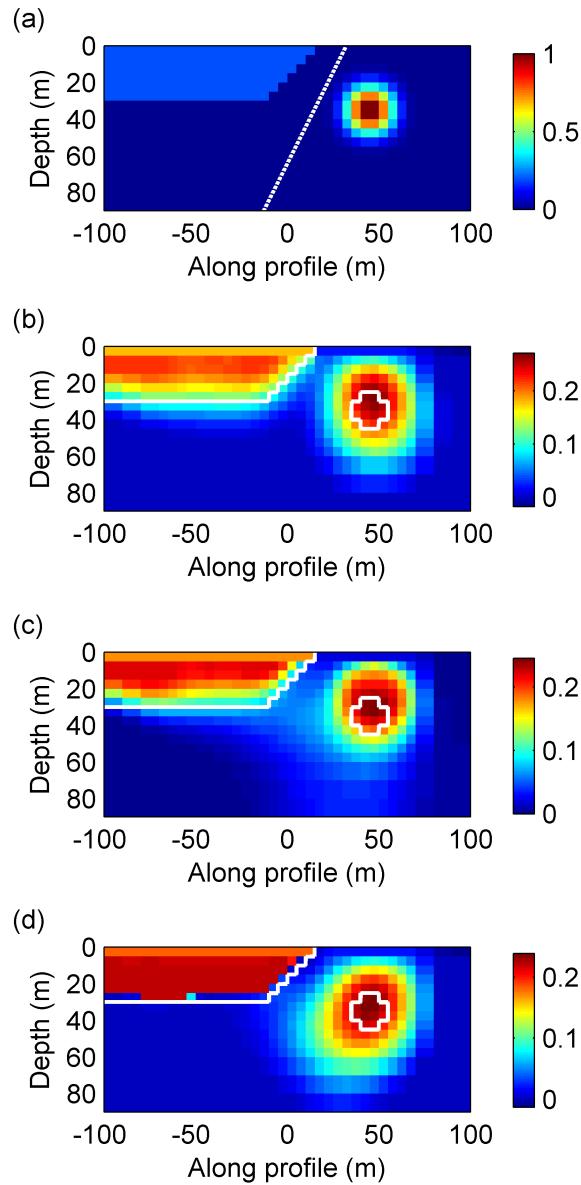


Figure 5.7: Plots for the second synthetic example: (a) shows the true density model (units of g/cc); (b) shows the inversion result with ℓ^2 -norms; (c) shows the inversion result with an approximate total variation measure (Ekblom with $p = 1.1$ applied to $\vec{\nabla}m$); (d) shows the result of applying our iterative procedure on the region to the left of the dotted white line in (a). Solid white lines in (b), (c) and (d) indicate the location of features in the true model.

about possible interfaces in some parts of the volume but not all, and the interfaces will be better resolved by combining all the data.

Our iterative procedure extends nicely to allow cooperative use of multiple data types. Several inversions are now performed in step (1) in Figure 5.1. In step (2), the gradient magnitude in equation (5.4) is calculated on each result. The gradient magnitudes for each model are then normalized (such that they lie on $[0, 1]$) and added together in step (4). The result is used to determine the smoothness weighting in step (5). With this iterative cooperative procedure it is simple to accommodate alteration differences between structural model units (e.g. one physical property varying more sharply across an interface than another physical property) by altering the low weighting values used in each inversion (i.e. a lower weight will result in a sharper interface).

A conventional cooperative practice is to invert one set of data independently and use the result to constrain a subsequent, independent inversion of the second set. However, the resulting model from this simple sequential cooperative inversion strategy can be biased towards the results of the first inversion (Gallardo and Meju, 2004). Instead of designing a procedure that works in a sequential (back-and-forth) manner, our iterative cooperative procedure is performed in a ladder style (i.e. simultaneously but independently). Hence, the procedure can be easily parallelized for computing efficiency.

5.4.3 Application of our cooperative strategy to a 2D synthetic example

Here we apply our procedure to an interesting 2D synthetic scenario. The synthetic model, shown in Figure 5.8(a) and 5.8(b), contains two blocks in a half-space. All density values mentioned in this section have units of g/cc and all slowness values have units of s/km. The block at surface has density 1.0 and slowness 1.0. The block at depth has density 1.0 and slowness -1.0 . We recognize that this is not a likely geologic scenario for density and slowness, but we emphasize that it is applicable for most other physical property pairs (e.g. magnetic susceptibility and slowness) so we continue with this example for application to more general scenarios.

The two geophysical methods for this example are gravity and cross-well seismic tomography. The gravity data are collected across the top of the model at a height of 2.5m, every 5.0m along-profile from -160m to -160m . The seismic sources and receivers are indicated in Figure 5.8(b) by dots on the left and right of the mesh: these are located at depths every 10.0m

from 2.5m to 82.5m. A small amount of random noise is added to both sets of data before inverting; for each type of data, that noise is taken from a normal distribution with zero mean and standard deviation equal to 2.0% of the range of the data.

We design this synthetic example such that one of the blocks is well resolved by only the gravity data and the other block is well resolved by only the tomography travel-time data. Gravity data are most sensitive to the cells close to the surface. Seismic tomography data are most sensitive to cells where there is a high density of ray-paths. Hence, we place one block at the surface (where gravity data are most sensitive and where there are fewer ray paths) and the second block is placed at depth near the centre of the mesh (where the gravity data are less sensitive and where there is a higher density of ray paths). This synthetic example thereby represents a scenario in which the geophysical data sets are complimentary and must be combined to adequately resolve the entire inversion volume.

The default inversion results with ℓ^2 -norms are shown in Figure 5.8(c) and 5.8(d). The results with an Ekblom measure with $p = 1.1$ applied to the spatial model gradient (which estimates a total variation measure) are shown in Figure 5.8(e) and 5.8(f). The results of performing our iterative cooperative procedure are presented in Figure 5.8(g) and 5.8(h). Models and weights at successive iterations are in Figure 5.9, and a measure of the progression of the procedure is in Figure 5.10. Again, the expected break-down of the method after passing a critical volume percentile is evident and the accepted final models at iteration (5) recover the true scenario well.

Because the gravity data have poor sensitivity to the body at depth, that body is not resolved in the ℓ^2 -norm inversion result in Figure 5.8(c). In contrast, our iterative procedure places an anomalous body at depth in the density model (Figure 5.8(g)), and the question arises here as to whether that body is being created 1) to fit the gravity data or 2) through a link to the body in the slowness model via the cooperative strategy. Either is advantageous in this scenario, but we would hope that if the body at depth did not exist in the density model, but did in the slowness model, that our cooperative procedure would not create an artifact at depth in the recovered density model. To test this, we run the synthetic example again but remove the body at depth from the true density model. The results of performing our iterative cooperative procedure for this new scenario are presented in Figure 5.8(i) and 5.8(j). We do not show the progression of the smoothness weighting but it is similar to that seen previously in Figure 5.9, the important aspect being that the cooperative strategy still sets the weights low around the location of the slowness anomaly at depth. As we hoped, an artifact is not created at depth in the density model in Figure 5.8(i), despite the smoothness weights for the inversion having

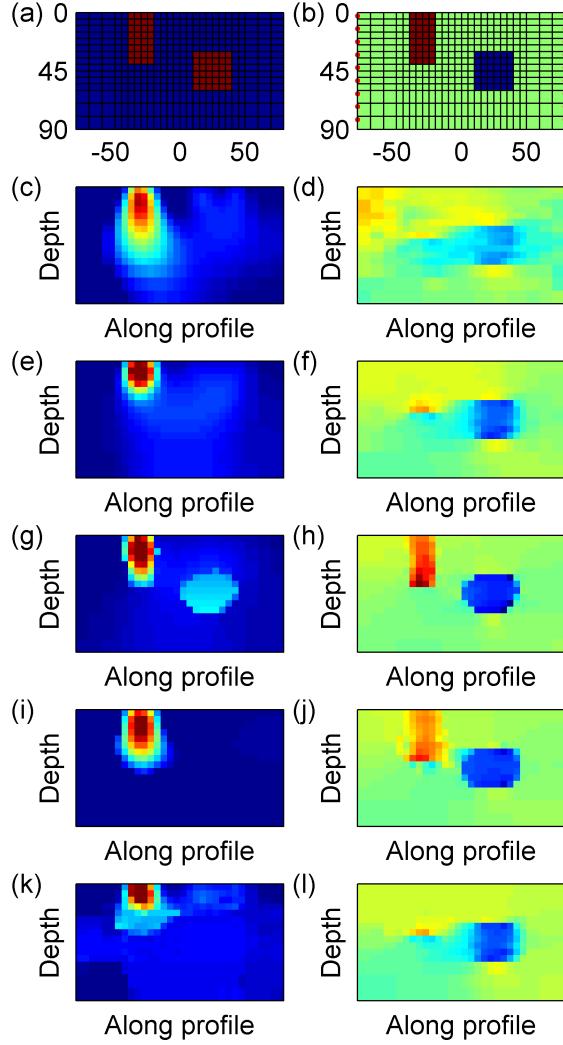


Figure 5.8: Plots for the third synthetic example: (a) shows the true density model (units of g/cc) and mesh; (b) shows the true slowness model (units of s/km) and mesh; (c) and (d) show the gravity and seismic tomography inversion results with ℓ^2 -norms; (e) and (f) show the inversion results with an approximate total variation measure (Ekblom with $p = 1.1$ applied to $\vec{\nabla}m$); (g) and (h) show the results for our iterative cooperative procedure; (i) and (j) show the results for our iterative cooperative procedure when the scenario is altered slightly to remove the body at depth in the density model only; (k) and (l) show the cross-gradient joint inversion results (again using the Ekblom measure with $p = 1.1$ applied to $\vec{\nabla}m$). The colour bar used is the same as that in Figure 5.2 (red high, blue low). The colour scales are on $[0, 1]$ for the density models and $[-1, 1]$ for the slowness models. The seismic sources and receivers are indicated in (b) by red dots on the left and right of the mesh. To remove clutter, we only show axis labels in (a) and (b): units are metres.

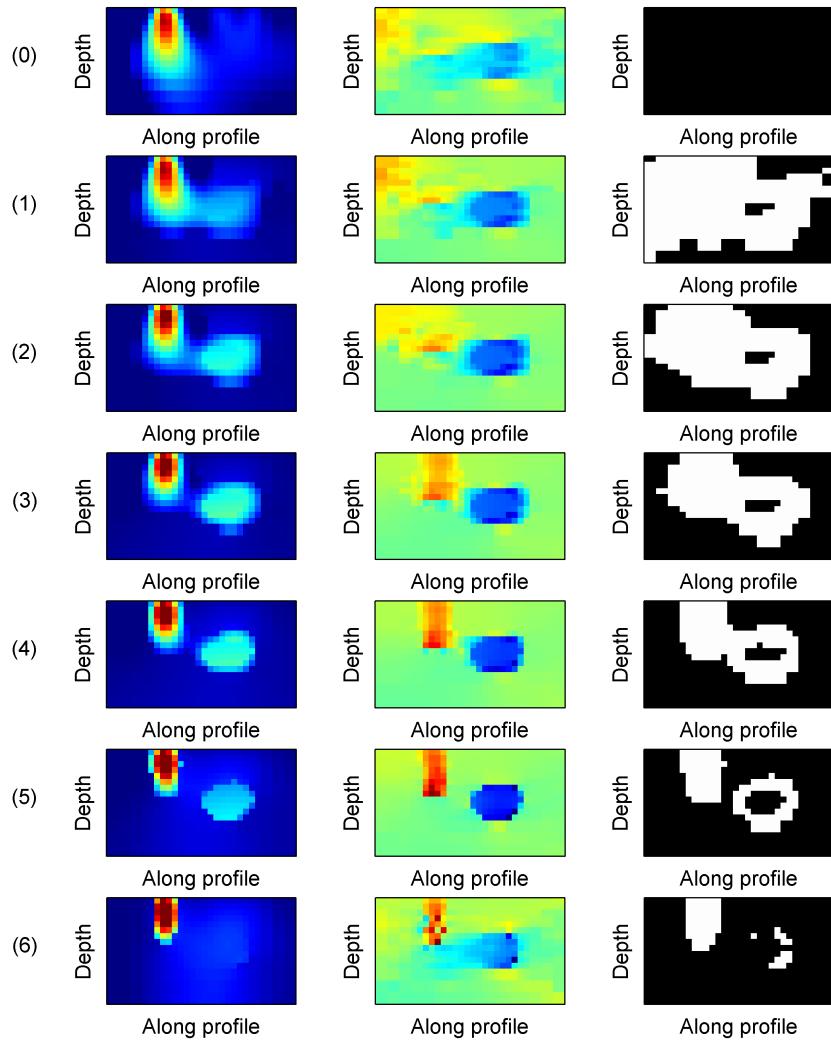


Figure 5.9: Results of the iterative cooperative procedure for the third synthetic example. The recovered density models (units of g/cc) are in the left column, the slowness models (units of s/km) are in the centre column, and the weights applied in the inversions are in the right column. The numbers on the left indicate iteration number. The models at iteration 0 are the initial models recovered from inversions with uniform weighting across the volume. The colour bar used is the same as that in Figure 5.2 (red high, blue low). The colour scales are on $[0, 1]$ for the density models and $[-1, 1]$ for the slowness models. The low weighting values are 0.01, indicated in white. To remove clutter, we do not show axis labels but we refer the reader to the labeled version of the inversion mesh in Figure 5.8(a) and 5.8(b).

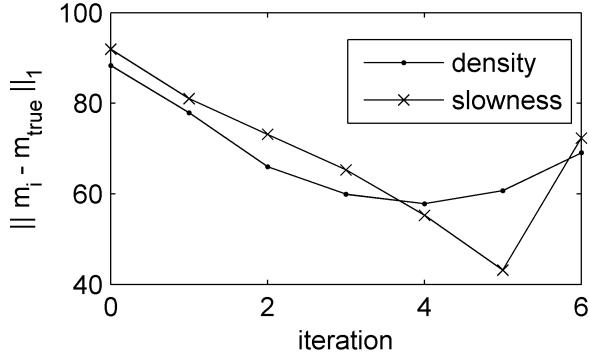


Figure 5.10: Progression curves for the iterative cooperative procedure results in Figure 5.9. The progression measure is the ℓ^1 -norm of the difference between the model at each iteration and the true model.

been set low in that region.

5.4.4 Comparison to cross-gradient joint inversion

Three notable cooperative inversion approaches are those of Lines et al. (1988), Oldenburg et al. (1997), and Paasche and Tronicke (2007), all of which assume some correlation between the physical properties involved but do not introduce an explicit coupling relationship. Our cooperative strategy differs in that rather than assuming some compositional correlation, we assume that the structural aspects of the multiple physical property distributions are spatially correlated (e.g. the models change in the same or similar places due to interfaces between rock units). In this way, our cooperative approach can be compared to the joint inversion approach of Gallardo and Meju (2004), who introduce a measure of structural similarity into the optimization constraints based on the cross-product of the spatial model gradients. The method of Gallardo and Meju (2004) has been well researched and applied to several realistic exploration scenarios and we therefore choose their method to compare against our own.

Our optimization problem with cross-gradient constraints is to minimize

$$\Phi = \Phi_{d1}(\mathbf{m}_1) + \Phi_{d2}(\mathbf{m}_2) + \beta(\Phi_{m1}(\mathbf{m}_1) + \Phi_{m2}(\mathbf{m}_2)) \quad (5.6)$$

subject to the equality constraint

$$(\mathbf{G}_x \mathbf{m}_1) \odot (\mathbf{G}_z \mathbf{m}_2) - (\mathbf{G}_z \mathbf{m}_1) \odot (\mathbf{G}_x \mathbf{m}_2) = \mathbf{0}. \quad (5.7)$$

Here, \mathbf{m}_1 and \mathbf{m}_2 are the two models involved, \mathbf{G}_x and \mathbf{G}_z are the two gradient operators for the 2D problem, and \odot denotes element-by-element multiplication. Note that there is only one trade-off parameter β with which to control the data misfit. It is then difficult (and most likely impossible) to fit both sets of data to the desired target levels such that $\Phi_{d1}(\mathbf{m}_1) = \Phi_{d1}^*$ and $\Phi_{d2}(\mathbf{m}_2) = \Phi_{d2}^*$. We search for a value of β such that

$$\Phi_{d1} + \Phi_{d2} = \Phi_{d1}^* + \Phi_{d2}^*. \quad (5.8)$$

This approach does a reasonable job of fitting both data sets to an appropriate level for our example: the target misfit values were 65.0 and 81.0 for the gravity and travel-time data respectively, and the final misfits recovered were 53.2 and 96.0.

To solve the joint inverse problem in equations (5.6) and (5.7) we use the nonlinear optimization software package KNITRO (Byrd et al., 2006) from Ziena Optimization Inc. Our iterative cooperative procedure creates sharp interfaces so for comparison purposes we use an approximate total variation regularization (an Ekblom measure with $p = 1.1$ applied to $\vec{\nabla}m$) to help the joint inversion produce sharp interfaces. The results of the cross-gradient joint inversion for the third synthetic example are shown in Figure 5.8(k) and 5.8(l).

Although our joint inversion is able to recover a result that honours the structural constraint in equation (5.7), it is unable to overcome the lack of resolution in each geophysical survey and recovers results similar to the independent results in Figure 5.8(e) and 5.8(f). To understand this, note that the nonconvex cross-gradient measure in equation (5.7) can become zero in two cases: 1) when the models contain structural changes in the same places (i.e. aligned model gradient vectors), or 2) when one or both models are constant (i.e. when one or both model gradient vectors are zero). These two scenarios represent local minima in the problem, with the latter scenario evident in the results in Figure 5.8(k) and 5.8(l): each body is recovered well in only one inversion and the other physical property model is more-or-less constant in the same region. In comparison, our iterative cooperative procedure (the results in Figure 5.8(g) and 5.8(h)) is able to overcome the lack of resolution in each geophysical survey and recovers both bodies well in both inversions.

We note that there are some differences between our joint inversion method and that used by Gallardo and Meju (2004). In Gallardo and Meju (2004), the Φ_{m1} and Φ_{m2} smoothness

components of the objective function are multiplied by factors that can be adjusted to control the amount of structure (smoothness) obtained in the models. Those factors provide the same control as can the w_x , w_y and w_z smoothness weights in equation (5.2). Their smoothness regularization factors are held fixed at a predetermined value, found via trial and error, while an auxiliary damping factor (equivalent to our trade-off parameter β) is cooled (decreased) iteratively in a manner similar to our methodology. As such, their regularization and cooling strategy are comparable to ours but there are details in their optimization strategy that Gallardo and Meju (2004) state are able to “reduc[e] the possibility of being trapped in local minima”. Given these differences, we acknowledge that Gallardo and Meju (2004) may be able to avoid the local minimum solution found in our synthetic scenario through judicious use of their strategy. We present the scenario as a warning that one should give sufficient consideration to the resolution of the geophysical surveys when inverting multiple data sets though a joint or cooperative strategy.

5.5 Application to the San Nicolás deposit

Phillips (2001) performed considerable work on geophysical data from the San Nicolás massive sulphide copper-zinc deposit (Zacatecas, Mexico). Here we apply our procedure to the gravity data therein. The inversion mesh is shown in Figure 5.11 and the data is plotted in Figures 5.11 and 5.12(a). There is a detailed geologic model available, interpreted from an extensive drilling program, that can be used to compare against our inversion results. Figure 5.13 shows cross-sections through the density model created from the geologic model and physical property information available; the high density sulfide body is evident (red in those images). The deposit is bounded to the east by a southwest-dipping fault. Mineralization continues along the fault to depth to create a smaller keel structure that is evident on the left of Figure 5.13(a) and 5.13(b).

The premise here is that detailed geologic information is *not* available. The hypothetical exploration question is: Given that we expect some blob of high density material in the subsurface (the sulphide body), what are some likely positions of the boundaries of that body? Our iterative inversion procedure can provide an answer to that question and we can use the detailed geologic information (i.e. the interfaces in the geologic model) to assess how well our procedure performs.

Figure 5.14 shows the same cross-sections through the density model recovered through default inversion (ℓ^2 -norms). The predicted data are shown in Figure 5.12(b). That model shows

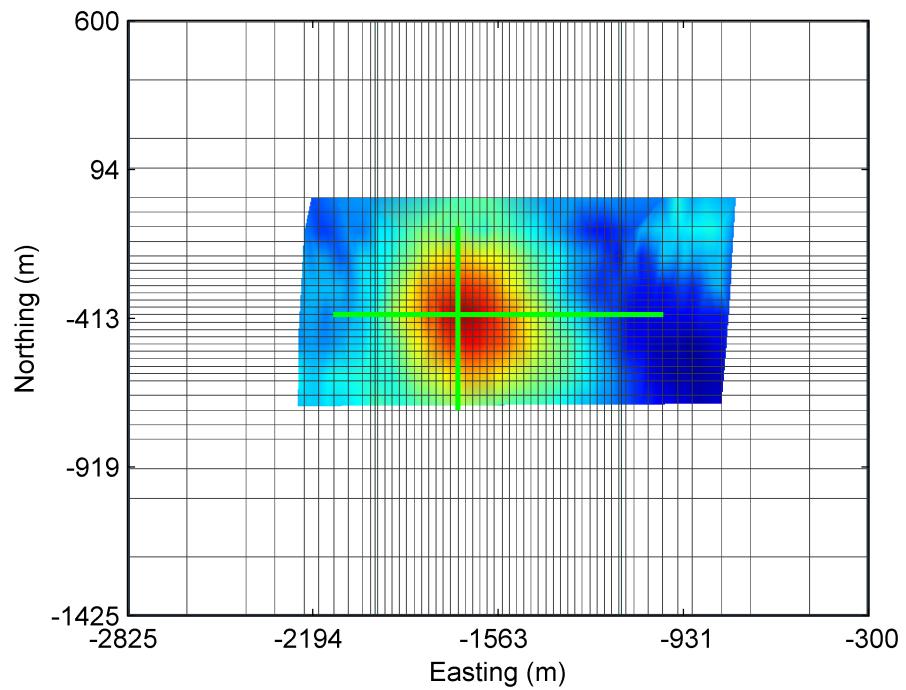


Figure 5.11: A top view of the mesh used to invert the San Nicolás gravity data, a map of which is overlayed. The mesh has 53-by-33-by-35 cells (easting-northing-depth). The locations of the cross-sections shown in the figures that follow are indicated with green lines.

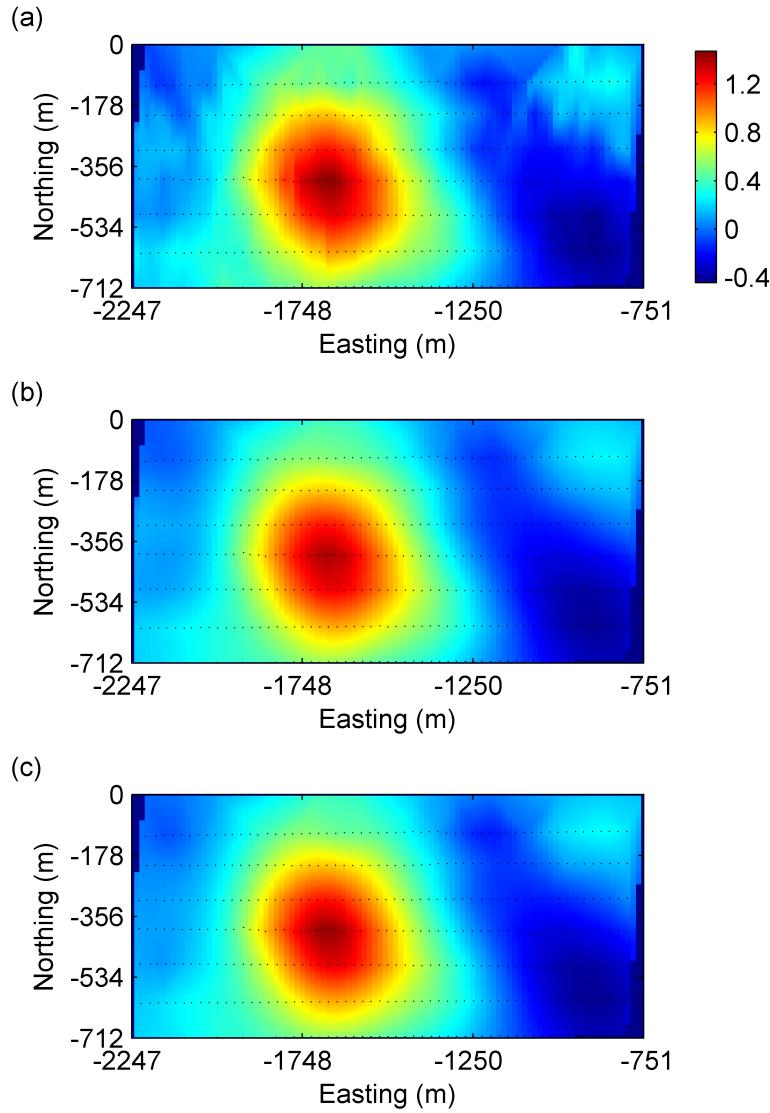


Figure 5.12: A map of the San Nicolás gravity data (units of mGal) is in (a); a regional component has been removed from the original data used in Phillips (2001). The predicted data for the model in Figure 5.14 is in (b). The predicted data for the model in Figure 5.15 is in (c). All three data maps are plotted on the same scale for comparison. The locations of the 422 data are indicated by black dots.

the smoothness inherent to minimum-structure inversions and it is difficult to interpret the boundaries of the sulphide body and estimate its overall size. Figure 5.15 shows the result after applying our iterative strategy, with the final smoothness weights shown in Figure 5.16. The predicted data are shown in Figure 5.12(c). A low weighting value of 0.01 was used and we did not apply smoothing within the procedure. As is evident from the overlayed interfaces from the geologic model, the result in Figure 5.15 now provides a better estimate of the size of the high density sulphide body. The depth extent of the sulphide body has been better recovered, as well as the northern extent. The southern extent of the body is less well resolved in Figure 5.15(b) but this may be a shallower reconstruction of the keel, the size of which is difficult to model with the discretization used.

Due to corrections applied and regional components removed during data processing steps, it is difficult to compare the density values from the geologic model with those in the inversion results in an absolute sense. We can however compare the range of density values in the models (i.e. maximum value minus minimum value). Another issue is that forward modelling for the density model in Figure 5.13 (from the geologic model) creates a response with a range of approximately twice that of the observed data (see Figure 5.17). The density range for the geologic model is 1.70 (g/cc), but considering the forward modelling results, a value of 0.85 is a better value to compare with our results. The unconstrained result has a density range of 0.51 compared to 0.93 for the result from our iterative procedure. Hence, the iterative procedure has provided significantly improved density estimates.

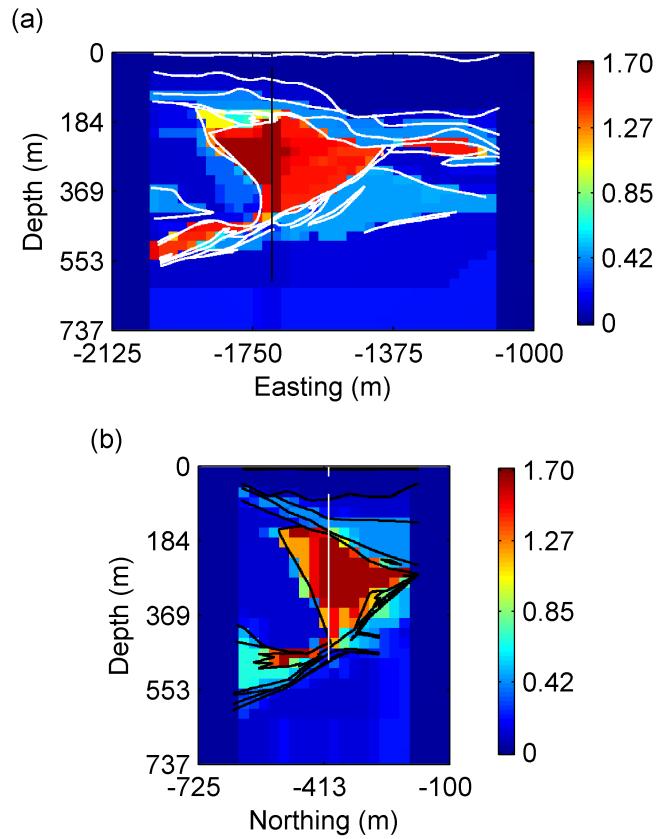


Figure 5.13: The anomalous density model (units of g/cc) created from the geologic model and physical property information for the San Nicolás deposit: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.

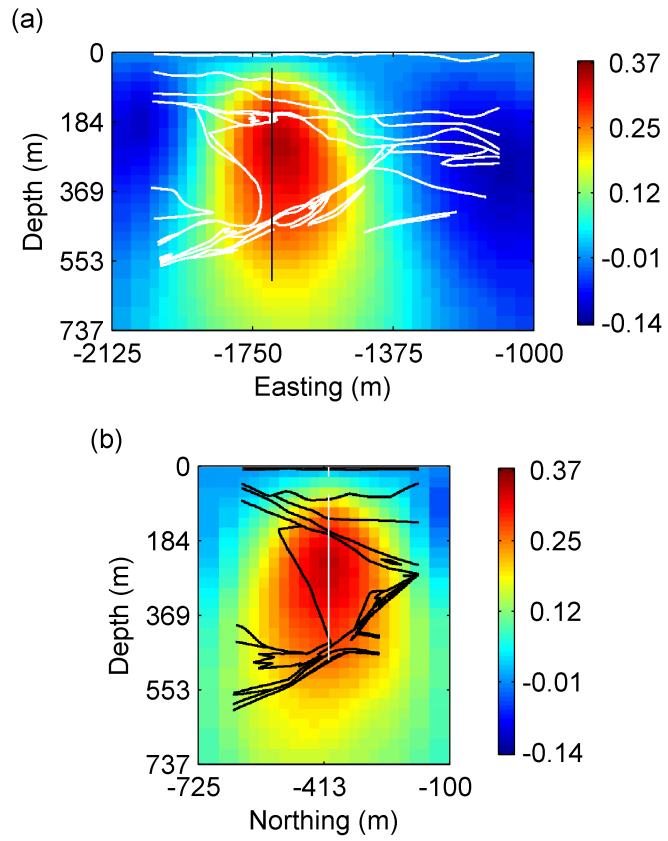


Figure 5.14: The recovered density model (units of g/cc) for the default, unconstrained inversion (ℓ^2 -norms) of the San Nicolás data: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.

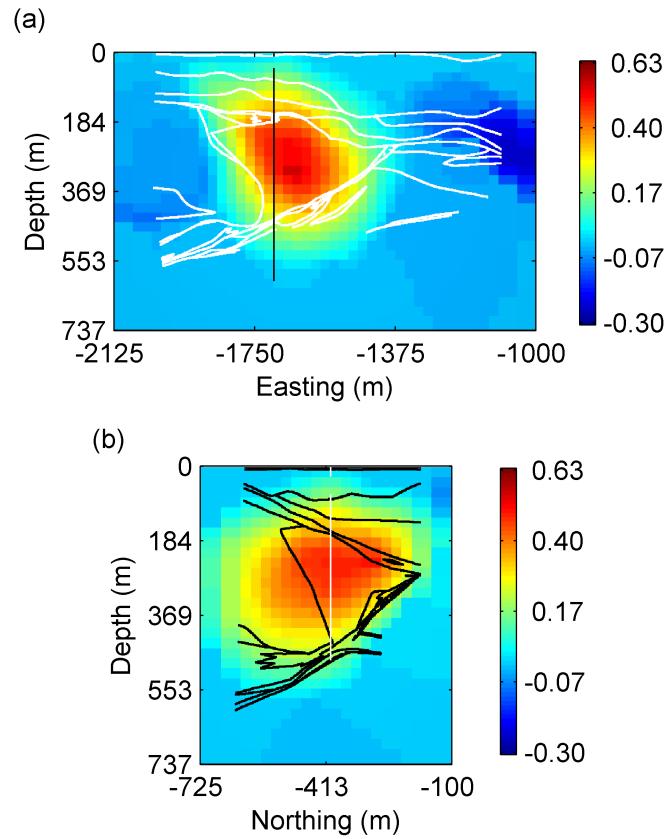


Figure 5.15: The recovered density model (units of g/cc) after applying our iterative procedure to the San Nicolás data: (a) shows a W-E cross-section at northing = -400m; (b) shows a S-N cross-section at easting = -1700m. Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section.

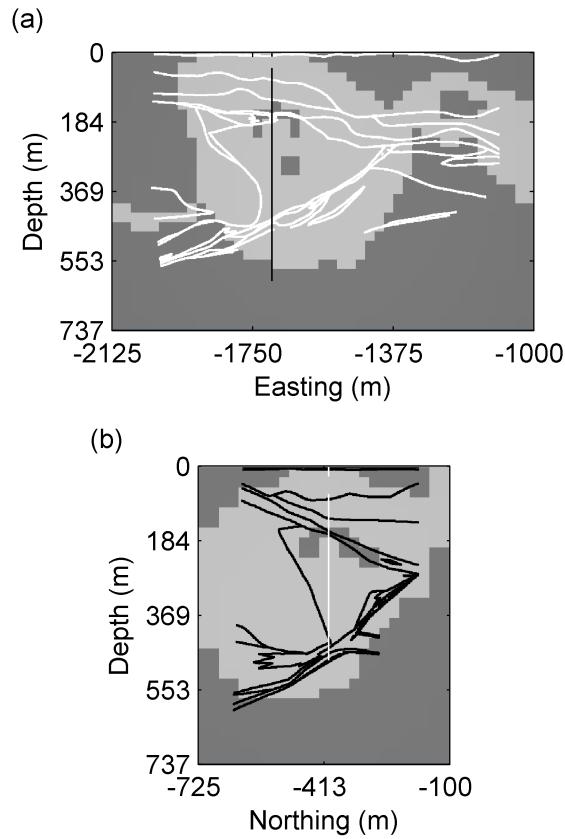


Figure 5.16: The final smoothness weights after applying our iterative procedure to the San Nicolás data: (a) shows a W-E cross-section at northing = -400m ; (b) shows a S-N cross-section at easting = -1700m . Interfaces in the geologic model are outlined in white for the W-E section and in black for the S-N section. Weights of 1.0 are shown in dark grey and weights of 0.01 are shown in light grey.

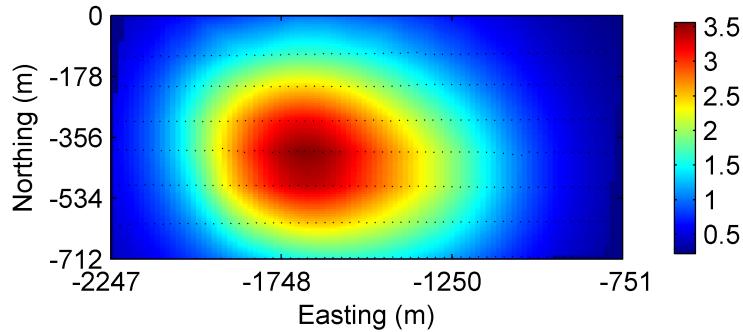


Figure 5.17: A map of the gravity data (units of mGal) forward modelled for the model in Figure 5.13. The locations of the 422 data are indicated by black dots.

5.6 Conclusion

We have developed an iterative inversion procedure that can obtain geologically realistic models by inverting one or more types of geophysical data in a cooperative fashion. The procedure can create models that involve sharp interfaces (physical property discontinuities) between rock units (regions with relatively constant physical properties) where appropriate in the subsurface. The procedure is fully automated, reliable and repeatable. It does not require any alteration to the inverse problem and can therefore be performed with pre-existing inversion algorithms.

Although our procedure is designed to allow recovery of sharp interfaces, we do not present it as a means to compete against methods like total variation regularization that can do the same. Instead, we present our procedure as one that could compliment such methods, i.e. we see no reason that our procedure could not be performed with non- ℓ^2 smoothness measures but such an alteration would require further investigation. Similarly, the procedure could be combined with the cross-gradient joint inversion method of Gallardo and Meju (2004).

We suggest application of our procedure to simple situations in which there are a few well-defined regions of differing rock types separated by sharp interfaces. However, the procedure can easily be used in more complicated scenarios by applying it over isolated volumes of the subsurface.

The procedure creates a sequence of models, all of which fit the data but have smoother or sharper characteristics. Ultimately, a geologist familiar with the region can best assess the resultant models and decide which features are realistic or not in light of the available geologic knowledge. New volumes can then be isolated for subsequent investigation using the iterative procedure. The method thereby allows for interaction between geologist and geophysicist.

5.7 References

- Bosch, M., A. Guillen, and P. Ledru, 2001, Lithologic tomography – an application to geophysical data from the Cadomian belt of northern Brittany, France: *Tectonophysics*, **331**, 197–227.
- Byrd, R. H., J. Nocedal, and R. A. Waltz, 2006, KNITRO – An integrated package for nonlinear optimization: *in* Large-Scale Nonlinear Optimization, G. di Pillo and M. Roma, eds, 35–59, Springer-Verlag.
- Ekblom, H., 1973, Calculation of linear best L_p -approximations: *BIT*, **13**, 292–300.
- Ekblom, H., 1987, The L_1 -estimate as limiting case of an L_p - or Huber-estimate: *in* Statistical Data Analysis Based on the L_1 -Norm and Related Methods, Y. Dodge, ed, 106–119, Elsevier.
- Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure: *Geophysical Journal International*, **134**, 213–227.
- Gallardo, L. A., and M. A. Meju, 2004, Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints: *Journal of Geophysical Research*, **109**, B03311.
- Guillen, A., P. Calcagno, G. Courrioux, A. Joly, and P. Ledru, 2008, Geological modelling from field data and geological knowledge – Part II. Modelling validation using gravity and magnetic data inversion: *Physics of the Earth and Planetary Interiors*, **171**, 158–169.
- Krahenbuhl, R. A., and Y. Li, 2004, Hybrid optimization for a binary inverse problem: 74th International Annual Meeting, SEG, Expanded Abstracts, 782–785.
- Li, Y., and D. W. Oldenburg, 1996, 3-D inversion of magnetic data: *Geophysics*, **61**, 394–408.
- Lines, L. R., A. K. Schultz, and S. Treitel, 1988, Cooperative inversion of geophysical data: *Geophysics*, **53**, 8–20.
- Musil, M., H. R. Maurer, and A. G. Green, 2003, Discrete tomography and joint inversion for loosely connected or unconnected physical properties – application to crosshole seismic and georadar data sets: *Geophysical Journal International*, **153**, 389–402.
- Oldenburg, D. W., Y. Li, and R. G. Ellis, 1997, Inversion of geophysical data over a copper gold porphyry deposit – a case history for Mt. Milligan: *Geophysics*, **62**, 1419–1431.
- Paasche, H. and J. Tronicke, 2007, Cooperative inversion of 2D geophysical data sets – a zonal approach based on fuzzy c-means cluster analysis: *Geophysics*, **72**, A35–A39.

- Phillips, N. D., 2001, Geophysical inversion in an integrated exploration program – examples from the San Nicolás deposit: Master's Thesis, University of British Columbia, Canada.
- Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms: *Physica D*, **60**, 259–268.
- van Zon, T., P. Ditmar, K. R. Chowdhury, and J. Mondt, 2007, Lithotype discrimination within a gridded model by inverting gravity data: *Geophysical Journal International*, **170**, 481–491.
- Vogel, C. R., and M. E. Oman, 1998, Fast total variation based reconstruction of noisy, blurred images: *IEEE Transactions on Image Processing*, **7**, 813–824.

Chapter 6

Conclusion

6.1 Summary and primary significance of results

My overall research objective was to develop geophysical inversion methods to help achieve the goal of more reliable Earth models. This involved development of geophysical inversion methods that 1) increase the kinds of geologic information that can be incorporated, and 2) can combine several complimentary types of geophysical data collected over the same Earth region.

It is now possible to incorporate all of the following geologic information into geophysical inversions in a deterministic framework:

- physical property measurements on rock samples;
- lithology observations (combined with petrophysical information);
- structural orientations;
- structural contacts between rock units;
- physical properties change sharply (e.g. across offset faults) or smoothly (e.g. across zones of alteration) between rock units;
- relative positions of rock units (e.g. a particular rock type is known to lie above another but the location of the contact is unknown);
- physical properties increase or decrease in particular directions (e.g. density often increases with depth);
- expected target shapes and aspect ratios (e.g. an intrusive body should be disk-like or pipe-like).

Below I provide further explanation of how those items can be dealt with using the methods I have developed.

I developed two magnetization vector inversion routines that are appropriate for use on data that contain the response of material containing complicated magnetization (i.e. magnetization amplitude and direction varying throughout the region) (Lelièvre and Oldenburg (in press, a); Chapter 3). The methods fill a previously empty niche between the method of Shearer and Li (2004) and those mentioned in Li et al. (2004). The problem of remanence in magnetic inversions can now be dealt with through use of those methods within an appropriate workflow in which geologic information regarding remanence is incorporated to improve the results. I purposefully designed my magnetization inversions to allow incorporation of the types of information commonly collected.

In Lelièvre and Oldenburg (in press, b) (Chapter 4) I developed methods for incorporating orientation information into geophysical inversions. I provided a comprehensive look at the available methods for including this information in both my inversion framework and those of other authors. I improved upon the work of Li and Oldenburg (2000) by developing a finite difference scheme for the numerical derivatives which ameliorates problems of asymmetry evident in the original implementation. I also developed an approach that relies on additional linear constraints placed in the optimization problem, which provides functionality for other types of information including being able to specify physical property trends (i.e. an increase or decrease in particular directions) and relative positions of rock units. These methods for incorporating structural orientation information can be used on large volumes to specify expected target shapes and aspect ratios. The methods can also be applied to incorporate orientations expected for located planar features such as contacts between rock units.

In Chapter 5 I developed an iterative inversion procedure that can obtain geologically realistic models by inverting one or more types of geophysical data in a cooperative fashion. Information regarding the relative positions of rock units can be used to define topological rules applied during the procedure. The procedure creates a sequence of models, all of which fit the data but have smoother or sharper characteristics. Ultimately, a geologist familiar with the region can best assess the resultant models and decide which features are realistic or not in light of the available geologic knowledge. New volumes of the subsurface can then be isolated for subsequent investigation using the iterative procedure. The method thereby allows for interaction between geologist and geophysicist.

None of the methods I developed involve significant complications to the inverse problem and they can be applied to current exploration problems without requiring additional computing resources.

6.2 A note on the San Nicolás examples

In sections 4.5 and 5.5, different methods were applied to the same gravity data. The premises of the inversion scenarios in those two sections are wholly different, as I clarify below in order to avoid any misunderstanding regarding the significance of the results.

In section 4.5, the premise is that detailed geologic information is available in the form of a geologic model. The hypothetical exploration question is: Given the detailed geologic model with interfaces identified from drilling, what is a likely density distribution that can honour those interfaces and the geophysical data? The subsequent inversion presented new knowledge by indicating density values that were significantly lower than those assigned to the geologic model. It is likely the case that density values measured on drill-core samples were used to assign density values to the geologic model. Those values are representative of the drill-core scale and can be different from representative bulk densities on the larger inversion mesh scale.

In section 5.5, the premise is that detailed geologic information is *not* available. The hypothetical exploration question is: Given that we expect some blob of high density material in the subsurface (the sulphide body), what are some likely positions of the boundaries of that body? Our iterative inversion procedure was able to provide an answer to that question and we used the detailed geologic information (i.e. the interfaces in the geologic model) to assess how well our procedure performed.

6.3 Advantages of increased regularization options

6.3.1 Feature appraisal

The research of this thesis is aimed at reducing model uncertainty. Considering the nonuniqueness of the inverse problem, more than one model is required to allow reliable conjecture regarding the subsurface. Once several models of the Earth have been recovered through inversion, one should assess the results before using them to guide further exploration (e.g. spotting holes for drilling programs). How far can we trust the existence of certain features in our model? What features are required by the data, which have been constructed by over-fitting the data (i.e. fitting errors), and which have been created in accordance with added regularization or model constraints (and therefore may be biased by the interpreter's prejudices)? To help answer

these questions, unbiased methods for assessing the reliability of the inversion results (i.e. feature appraisal) are necessary. It is possible that poor experimental design or poor data quality may lead to inconclusive inversion results once an appropriate assessment is undertaken.

Feature appraisal is a topic of ongoing research but an approach suggested by Oldenburg and Li (1999) is to perform several inversions, each with different regularization applied, and to look for common features in the models. One can have more confidence if model features of interest are not altered when the regularization is adjusted. A hypothesis-testing procedure that follows similar reasoning would be to alter the regularization in an attempt to actively remove a feature from the model. If this can be done then there is less confidence that the feature is required by the geophysical survey data. By expanding the possible regularization options through my methods, more parameters are available with which such feature appraisal methods can work.

6.3.2 Learned regularization

It is not always clear what type of regularization to use for a particular application and different choices may yield very different solutions, especially if the measures available are more specific to certain types of geologic structure. In learned regularization (for example, see Haber and Tenorio (2003) and Hewson (2004)), supervised learning techniques are applied to a training set of models that containing desired features. The techniques help determine what form of regularization (i.e what set of measures and what values of their associated parameters) is most appropriate such that a subsequent inversion using that regularization scheme will obtain models that are most consistent with the training set. In my work I have expanded the possible regularization options, which introduces more parameters for the learning process to include.

Wijns and Kowalczyk (2007) developed a novel procedure for incorporating user interaction into the inversion process. The procedure is related to the learned regularization approach mentioned above: it essentially determines the best form of the regularization to apply as guided by a human interpreter. Several inversions are preformed with random values for several control parameters (e.g. regularization parameters). The resulting suite of recovered models are then visually inspected by the interpreter and ranked by how geologically reasonable they are (based on the interpreter's prior knowledge). A genetic algorithm then takes that ranking into account and modifies the control parameter set to generate a new suite of models. This procedure progressively converges towards a reasonable set of solutions. Again, by expanding the possible regularization options I have introduced more parameters for use in such a method.

6.4 Critical assessment of methods and future research

6.4.1 Chapter 2

Lelièvre et al. (2008) and Lelièvre et al. (2009) (Chapter 2) combine the methods found in my thesis into a single demonstrative example. A similar workflow should be applied to an appropriately chosen real-world scenario to gain insight into the practical aspects of incorporating the various types of geologic information. Before the methods can be used by inversion practitioners on exploration scenarios, it may be necessary to produce utility programs that facilitate the connection between the inversion algorithms and geologic knowledge to help remove any practical barriers.

6.4.2 Chapter 3

My magnetization vector inversions involve a three-fold increase in parameters over the method of Shearer and Li (2004). This leads to larger problems that require more computing time and memory to solve. Due to the increase in nonuniqueness following the increase in parameters, it will likely require more geologic constraints in the magnetization inversions than in the amplitude inversions of Shearer and Li (2004) in order to provide a similar improvement in the results.

A serious issue is that the spherical formulation magnetization inversion involves minimization of a nonconvex function, which is problematic from an optimization viewpoint. We have suggested a two stage approach in which a Cartesian inversion is used to help constrain a spherical inversion in order to avoid local minima resulting from the nonconvex nature of the spherical formulation objective function. However, I would expect that strategy to only ensure desired results when applied to simpler scenarios such as those presented in Lelièvre and Oldenburg (in press, a) (Chapter 3). In more complicated scenarios the Cartesian inversion may prove to be the only applicable routine. It may be more appropriate to extend the Cartesian formulation by adding the nonlinear inequality constraints required to specify the same types of remanence information that can easily be incorporated through simple bound constraints in the spherical formulation. However, optimization algorithms that can solve large geophysical inverse problems with nonlinear constraints in a reasonable time-frame (for current computing power) have not yet been developed.

Due to the increase in nonuniqueness of the vector magnetization inversions, it may be useful

to consider inversion of data that are more sensitive to (i.e. contain more information about) the magnetization than the commonly collected total magnetization intensity (TMI) data. Foss (2006) and Schmidt and Clark (2006) discuss how, compared to TMI data, three-component and gradient tensor magnetic data can potentially provide increased sensitivity to source magnetization direction. Use of such data with the magnetization inversion methods would require a numerical solution to a new forward problem and may require alteration to the optimization strategies. Collection of three-component and gradient tensor data is not yet a common practice but could be encouraged if the magnetization inversion results show improvement when including those types of data.

The magnetization inversion algorithms were applied to a high susceptibility example involving demagnetization effects and they provided reasonable results. The magnetization inversion approach should be compared to the high susceptibility inversion method of Lelièvre and Oldenburg (2006) to gain an understanding of how they could be applied, either individually or in concert, in high susceptibility scenarios. Before doing so, the inversion algorithm of Lelièvre and Oldenburg (2006) needs to be updated to use the same optimization methods employed in the magnetization inversion codes (especially the gradient-projection-reduced-Newton approach for bound constraints).

6.4.3 Chapter 4

I developed approaches for including orientation information through the addition of linear inequality constraints to the optimization problem (Lelièvre and Oldenburg (in press, b); Chapter 4). To solve the resulting inverse problem I used the logarithmic barrier method, which has proven to be a feasible solution method for large 3D geophysical inverse problems with simple bound constraints (Li and Oldenburg, 2003). The major practical difficulty with this approach is that the initial model must be feasible (i.e. must obey the constraints) and for complicated linear inequalities the creation of a feasible initial model may be a difficult task. In the examples I presented, each linear constraint specified relationships between only two model cells and creating feasible initial models was still difficult to automate. Hence, development of methods for creating feasible initial models for more complicated linear constraints would be highly advantageous. Alternatively, there are other optimization approaches for solving linearly constrained inverse problems (e.g. refer to Gill et al. (1995)). Although many of these are not currently feasible for the large 3D geophysical inverse problems of interest to mineral exploration, future advancement in computing power may change this.

6.4.4 Chapter 5

Cooperative inversion strategies, such as the one I designed in Chapter 5, invert multiple types of data individually but share information between the inversion runs in order to improve the similarity of the multiple models. The alternative is to develop joint inversion methods in which the data are inverted simultaneously while using some method to maintain similarity between the multiple physical property models involved. My cooperative strategy has performed well on a challenging synthetic scenario and I feel that the method is robust, appropriate and has potential to provide helpful results for real-world applications. However, it needs to be compared to joint inversion approaches and should be applied to more real-world scenarios in order to gain improved understanding of the effects of the control parameters.

The suggested application of the procedure is to simple situations in which there are a few well-defined regions of differing rock types separated by sharp interfaces. It is unlikely that the procedure can deal with more complicated scenarios. However, the method can easily be used in more complicated scenarios by applying it over isolated volumes of the subsurface.

In the development of this procedure I limited the investigation to the use of ℓ_2 smoothness measures in the inversion objective functions. Also, I performed independent inversions at each iteration of the procedure. It is possible that the procedure may provide improved results if non- ℓ^2 smoothness measures are used or if joint inversions (such as that of Gallardo and Meju (2004)) were used at each iteration of the procedure. Such alterations would require further investigation.

6.4.5 Deterministic and stochastic methods

The methods of Bosch et al. (2001) and Guillen et al. (2008) allow for the incorporation of much geologic information within a stochastic inversion framework. Below I will refer to the commercial software GeoModeller that is based on their methods, developed by Intrepid Geophysics (Australia) and BRGM (Bureau de Recherches Géologiques et Minières, France). There are benefits and limitations associated with both their stochastic approach and the deterministic approach presented in this thesis. Although their approach allows incorporation of a wider variety of geologic information, GeoModeller requires significant computational burden relative to my deterministic inversions. It is difficult to make a case for one approach over the other as such a decision would depend greatly on the specific questions being asked in the particular

exploration scenario, but some general comments can be made.

Due to use of Monte Carlo sampling methods, GeoModeller requires an adequate initial estimate of the true Earth scenario in order to converge efficiently. Therefore, the deterministic inversion methods presented in this thesis are expected to provide more utility than GeoModeller inversions at earlier stages in the exploration of a site, during which time the geologic information available is typically too sparse to provide a full interpretation of the subsurface lithology. At later stages, models provided via deterministic inversions could be used to provide starting models for the GeoModeller inversion algorithm. At later stages in the exploration process, when more geologic knowledge is available, incorporation of more complicated types of information may be beyond the capabilities of deterministic inversions, at which point the stochastic method will prevail.

If used in concert through an appropriate workflow, perhaps as indicated above, these two approaches have the potential to provide a powerful set of tools for integrating geologic and geophysical data. Future work should investigate appropriate aspects of such a workflow.

6.5 The last word

To be reliable, Earth models used for mineral exploration should be consistent with all available geologic and geophysical information. Incorporation of geologic information into geophysical inversions is essential to obtain improved, geologically reasonable solutions. My research provides functional methods for applying geophysical inversion to exploration problems; the synthetic and real data examples in this thesis have demonstrated that this functionality shows promise for practical use. My research helps bring the inversion practitioner closer to the goal of more reliable Earth models, consistent with all information available.

6.6 References

- Bosch, M., A. Guillen, and P. Ledru, 2001, Lithologic tomography – an application to geophysical data from the Cadomian belt of northern Brittany, France: *Tectonophysics*, **331**, 197–227.
- Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure: *Geophysical Journal International*, **134**, 213–227.
- Foss, C., 2006, Improvements in source resolution that can be expected from inversion of magnetic field tensor data: *The Leading Edge*, **25**, 81–84.
- Gallardo, L. A., and M. A. Meju, 2004, Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints: *Journal of Geophysical Research*, **109**, B03311.
- Gill, P., W. Murray, and M. Wright, 1995, *Practical Optimization*: Academic Press, London.
- Guillen, A., P. Calcagno, G. Courrioux, A. Joly, and P. Ledru, 2008, Geological modelling from field data and geological knowledge – Part II. Modelling validation using gravity and magnetic data inversion: *Physics of the Earth and Planetary Interiors*, **171**, 158–169.
- Haber, E., and L. Tenorio, 2003, Learning regularization functionals – a supervised training approach: *Inverse Problems*, **19**, 611–626.
- Hewson, C., 2004, Developing Learned Regularization for Geophysical Inversions: Master's Thesis, The University of British Columbia, Canada.
- Lelièvre, P. G., and D. W. Oldenburg, 2006, Magnetic forward modelling and inversion for high susceptibility: *Geophysical Journal International*, **166**, 76–90.
- Lelièvre, P. G., D. W. Oldenburg, and N. C. Williams, 2008, Constraining geophysical inversions with geologic information: 78th International Annual Meeting, SEG, Extended Abstracts, 1223–1227.
- Lelièvre, P. G., D. W. Oldenburg and N. C. Williams, 2009, Integrating Geologic and Geophysical Data Through Advanced Constrained Inversions: Annual Meeting, ASEG, Extended Abstracts.
- Lelièvre, P. G., and D. W. Oldenburg, *in press (a)*, A 3D total magnetization inversion applicable when significant, complicated remanence is present: *Geophysics*.
- Lelièvre, P. G., and D. W. Oldenburg, *in press (b)*, A comprehensive study of including structural orientation information in geophysical inversions: *Geophysical Journal International*.

- Li, Y., and D. W. Oldenburg, 2000, Incorporating geologic dip information into geophysical inversions: *Geophysics*, **65**, 148–157.
- Li, Y., and D. W. Oldenburg, 2003, Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method: *Geophysical Journal International*, **152**, 251–265.
- Li, Y., S. Shearer, M. Haney, and N. Dannemiller, 2004, Comprehensive approaches to the inversion of magnetic data with strong remanent magnetization: 74th International Annual Meeting, SEG, Expanded Abstracts, 1191–1194.
- Oldenburg, D. W., and Y. Li, 1999, Estimating depth of investigation in dc resistivity and IP surveys: *Geophysics*, **64**, 403–416.
- Shearer, S., and Y. Li, 2004, 3D inversion of magnetic total gradient data in the presence of remanent magnetization: 74th International Annual Meeting, SEG, Expanded Abstracts, 774–777.
- Schmidt, P. W., and D. A. Clark, 2006, The magnetic gradient tensor – its properties and uses in source characterization: *The Leading Edge*, **25**, 75–78.
- Wijns, C., and P. Kowalczyk, 2007, Interactive geophysical inversion using qualitative geological constraints: *Exploration Geophysics*, **38**, 208–212.

Appendix A

Inversion primer

A.1 Introduction

Geophysical inversion seeks to recover models of the Earth's physical properties (e.g. density, conductivity) that can adequately reproduce anomalies in geophysical survey data (e.g. gravity, DC resistivity) while being consistent with geologic information. The physical properties are related to rock composition, structure and physical state. Hence, the physical property models recovered via inversion are an important source of information for understanding subsurface geology as it applies to mineral exploration.

As a simple example of inversion, consider the following gravity problem. The force of gravity at a particular location is determined by the mass in the surroundings. Let there be M spherical masses with densities ρ_i ($i = 1, \dots, M$), volumes v_i and locations $\vec{r}_i = (x_i, y_i, z_i)$. The vertical component of the gravity field, F_{gz} , due to these masses at a location \vec{r}_0 is determined from Newton's Law of Gravity and is written

$$F_{gz}(\vec{r}_0) = \gamma \sum_i^M \rho_i v_i \frac{z_i - z_0}{\|\vec{r}_i - \vec{r}_0\|^3} \quad (\text{A.1})$$

where γ is the gravitational constant. In a gravity survey, one measures the force of gravity across a region of interest; those measurements are the geophysical data. The task of inversion in this example is to reconstruct a scenario containing a number of masses at various locations (this is the model) that is consistent with the survey datum (the force of gravity at the measurement point). I will refer back to this example in the following discussion.

A.2 The forward problem

Associated with any inverse problem is a forward problem. A forward problem seeks to determine the data that would be measured by a particular geophysical survey given some model of the Earth (i.e. it calculates some physical response).

Returning to the simple gravity example introduced above, the forward problem is to determine the force of gravity for the given distribution of masses. A forward problem can be solved uniquely using the relevant physics, and is written

$$\mathbf{d}^{obs} = F[\mathbf{m}] + \epsilon \quad (\text{A.2})$$

where \mathbf{d}^{obs} is the observed (measured) data, F is the forward operator determined by the physics and geometry of the problem, and ϵ is some error that may be attributed to data collection or inaccuracies in the forward operator.

A.3 The inverse problem

The inverse problem is to determine \mathbf{m} (the model) from \mathbf{d}^{obs} (the data) in equation (A.2). To invert geophysical data, the Earth region of interest is discretized into many cells within an orthogonal mesh so that complicated structures can be modelled. There are many more model parameters than data and the problem is said to be underdetermined, meaning that there are not enough data to uniquely determine a model solution and there are therefore infinitely many appropriate models.

Returning again to our simple gravity example, given a single measurement of the force of gravity (at a specified location in a particular direction) one can create an infinite number of situations with different numbers of bodies and associated masses (i.e. densities and volumes) for which that force is obtained.

A.4 Fitting the data

A measure of misfit between the noisy observed data, \mathbf{d}^{obs} , and the data produced (predicted) by a candidate model, $\mathbf{d}^{pred} = F[\mathbf{m}]$, can be defined as a sum-of-squares

$$\Phi_d = \sum_{i=1}^N \left(\frac{d_i^{pred} - d_i^{obs}}{\sigma_i} \right)^2 \quad (\text{A.3})$$

where N is the number of data. Each data difference is normalized by an uncertainty, σ_i . These uncertainties are estimated errors in the observed data. The larger the uncertainty in an observed datum, the smaller its contribution to the misfit measure. As such, the inversion will put more effort into fitting the more trustworthy data.

Because there are errors in the survey data, we should not seek a model that fits the data exactly ($\Phi_d = 0$). Doing so would mean fitting the errors and would ensure an incorrect inverse solution. Normal practice is to assign a target misfit Φ_d^* that is determined from the error statistics. When creating synthetic inverse problems, in which the true Earth model is known, we can add noise to the forward modelled data from a known Gaussian distribution. In such case, the misfit defined in equation (A.3) is a chi-squared statistical variable and the expected value of the misfit is then $\Phi_d^* = N$.

A.5 Specifying desired model characteristics

As discussed above, there are an infinite number of models that can fit the data to the desired degree (i.e. such that $\Phi_d \simeq \Phi_d^*$), and further information is essential for a unique solution: the model is required to contain desired characteristics derived from the available geologic knowledge of the site. Administration of this requirement is called regularization. Regularization requires a measure of how well a candidate model conforms to the desired characteristics. This measure is sometimes called the model objective function.

Consider again our simple gravity example. If we specify that there is only a single body at a particular location then the inverse problem requires finding a density for that body that achieves the required force. This is trivial as there is now only one unknown to solve for in equation (A.1).

A.6 Optimization

The inverse problem can now be formulated as an optimization problem that involves minimization of a (total) objective function, Φ , that combines the misfit, Φ_d , and model objective function, Φ_m :

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \beta \Phi_m(\mathbf{m}). \quad (\text{A.4})$$

The trade-off parameter β in equation (A.4) controls the relative size of the Φ_d and Φ_m measures for the resulting model. Determining an appropriate value of β represents a trade-off: at high β values, recovered models will not fit the data well but will contain structure consistent with the applied regularization; at low β values, recovered models will fit the data well but may contain spurious structure resulting from fitting errors in the data. To obtain a model that achieves an appropriate balance, the minimization in equation (A.4) is performed for several values of β and the resulting models compared. One often seeks a value β^* that yields a misfit close to a target Φ_d^* .

The major computational bottleneck for inversion is the speed of the forward solution. Generally, this solution is required a great many times during the minimization of the objective function in equation (A.4). Compared to the time required to calculate Φ_d , that for Φ_m is relatively minor since it generally involves sparse operators.

Appendix B

Derivation of the Hessian for the spherical magnetization inversion formulation

Here I derive the expression in equation (3.38) for the Hessian of the misfit term for the spherical magnetization inversion formulation developed in Chapter 3. I begin with the expression in equation (3.33) for the gradient of the misfit term for the spherical formulation:

$$\mathbf{g}_d = \mathbf{S}^T \tilde{\mathbf{G}}_{uvw}^T \delta \mathbf{d}. \quad (\text{B.1})$$

This is expanded as

$$\mathbf{g}_d = \begin{bmatrix} \mathbf{g}_a \\ \mathbf{g}_\theta \\ \mathbf{g}_\phi \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{ua} & \mathbf{S}_{va} & \mathbf{S}_{wa} \\ \mathbf{S}_{u\theta} & \mathbf{S}_{v\theta} & \mathbf{S}_{w\theta} \\ \mathbf{S}_{u\phi} & \mathbf{S}_{v\phi} & \mathbf{S}_{w\phi} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}_u^T \\ \tilde{\mathbf{G}}_v^T \\ \tilde{\mathbf{G}}_w^T \end{bmatrix} \delta \mathbf{d}, \quad (\text{B.2})$$

and let $i, j \in \{a, \theta, \phi\}$ to define

$$\mathbf{g}_i = \mathbf{S}_{ui} \tilde{\mathbf{G}}_u^T \delta \mathbf{d} + \mathbf{S}_{vi} \tilde{\mathbf{G}}_v^T \delta \mathbf{d} + \mathbf{S}_{wi} \tilde{\mathbf{G}}_w^T \delta \mathbf{d} \quad (\text{B.3})$$

or more simply as

$$\mathbf{g}_i = \mathbf{S}_{ui} \mathbf{d}_u + \mathbf{S}_{vi} \mathbf{d}_v + \mathbf{S}_{wi} \mathbf{d}_w \quad (\text{B.4})$$

with the definitions

$$\mathbf{d}_u = \tilde{\mathbf{G}}_u^T \delta \mathbf{d} \quad (\text{B.5a})$$

$$\mathbf{d}_v = \tilde{\mathbf{G}}_v^T \delta \mathbf{d} \quad (\text{B.5b})$$

$$\mathbf{d}_w = \tilde{\mathbf{G}}_w^T \delta \mathbf{d}. \quad (\text{B.5c})$$

The nine Hessian terms, each an M -by- M matrix, are then given by

$$\begin{aligned}\mathbf{H}_{ij} &= \frac{d\mathbf{g}_i}{d\mathbf{j}} \\ &= \frac{d}{d\mathbf{j}} (\mathbf{S}_{ui}\mathbf{d}_u) + \frac{d}{d\mathbf{j}} (\mathbf{S}_{vi}\mathbf{d}_v) + \frac{d}{d\mathbf{j}} (\mathbf{S}_{wi}\mathbf{d}_w).\end{aligned}\quad (\text{B.6})$$

The product rule for matrix-vector equations is now required: if f is an m -by- n matrix (or vector) and g is p -by- q then

$$\frac{df(x)g(x)}{dx} = (g(x)^T \otimes \mathbf{I}_m) f'(x) + (\mathbf{I}_q \otimes f(x)) g'(x) \quad (\text{B.7})$$

where \otimes is the Kronecker product. This gives

$$\begin{aligned}\mathbf{H}_{ij} &= (\mathbf{d}_u^T \otimes \mathbf{I}) \frac{d\mathbf{S}_{ui}}{d\mathbf{j}} + \mathbf{S}_{ui} \tilde{\mathbf{G}}_u^T \frac{d\delta\mathbf{d}}{d\mathbf{j}} \\ &\quad + (\mathbf{d}_v^T \otimes \mathbf{I}) \frac{d\mathbf{S}_{vi}}{d\mathbf{j}} + \mathbf{S}_{vi} \tilde{\mathbf{G}}_v^T \frac{d\delta\mathbf{d}}{d\mathbf{j}} \\ &\quad + (\mathbf{d}_w^T \otimes \mathbf{I}) \frac{d\mathbf{S}_{wi}}{d\mathbf{j}} + \mathbf{S}_{wi} \tilde{\mathbf{G}}_w^T \frac{d\delta\mathbf{d}}{d\mathbf{j}} \\ &= (\mathbf{S}_{ui} \tilde{\mathbf{G}}_u^T + \mathbf{S}_{vi} \tilde{\mathbf{G}}_v^T + \mathbf{S}_{wi} \tilde{\mathbf{G}}_w^T) \frac{d\delta\mathbf{d}}{d\mathbf{j}} \\ &\quad + (\mathbf{d}_u^T \otimes \mathbf{I}) \frac{d\mathbf{S}_{ui}}{d\mathbf{j}} + (\mathbf{d}_v^T \otimes \mathbf{I}) \frac{d\mathbf{S}_{vi}}{d\mathbf{j}} + (\mathbf{d}_w^T \otimes \mathbf{I}) \frac{d\mathbf{S}_{wi}}{d\mathbf{j}}.\end{aligned}\quad (\text{B.8})$$

Equation (3.42) provides

$$\begin{aligned}\frac{d\delta\mathbf{d}}{d\mathbf{j}} &= \tilde{\mathbf{G}}_{uvw} \begin{bmatrix} \mathbf{S}_{uj} \\ \mathbf{S}_{vj} \\ \mathbf{S}_{wj} \end{bmatrix} \\ &= \tilde{\mathbf{G}}_u \mathbf{S}_{uj} + \tilde{\mathbf{G}}_v \mathbf{S}_{vj} + \tilde{\mathbf{G}}_w \mathbf{S}_{wj}.\end{aligned}\quad (\text{B.9})$$

Use of the identity

$$(\mathbf{a}^T \otimes \mathbf{I}) \frac{d}{d\mathbf{x}} \text{diag}(\mathbf{b}) = \text{diag}(\mathbf{a}) \frac{d\mathbf{b}}{d\mathbf{x}} \quad (\text{B.10})$$

and the definitions in equations (3.35) and (3.40) gives

$$\begin{aligned}((\delta\mathbf{d}^T \tilde{\mathbf{G}}_u) \otimes \mathbf{I}) \frac{d\mathbf{S}_{ui}}{d\mathbf{j}} &= \text{diag}(\tilde{\mathbf{G}}_u^T \delta\mathbf{d}) \frac{d\mathbf{s}_{ui}}{d\mathbf{j}} \\ &= \text{diag}(\tilde{\mathbf{G}}_u^T \delta\mathbf{d}) \mathbf{T}_{u,ij}.\end{aligned}\quad (\text{B.11})$$

I can now write

$$\mathbf{H}_{ij} = [\mathbf{S}_{ui}, \mathbf{S}_{vi}, \mathbf{S}_{wi}] \tilde{\mathbf{G}}_{uvw}^T \tilde{\mathbf{G}}_{uvw} \begin{bmatrix} \mathbf{S}_{uj} \\ \mathbf{S}_{vj} \\ \mathbf{S}_{wj} \end{bmatrix} + \text{diag}(\mathbf{d}_u) \mathbf{T}_{u,ij} + \text{diag}(\mathbf{d}_v) \mathbf{T}_{v,ij} + \text{diag}(\mathbf{d}_w) \mathbf{T}_{w,ij} \quad (\text{B.12})$$

which leads to the final expression

$$\mathbf{H}_d = \mathbf{S}^T \tilde{\mathbf{G}}_{uvw}^T \tilde{\mathbf{G}}_{uvw} \mathbf{S} + \mathbf{T}_u \mathbf{D}_u + \mathbf{T}_v \mathbf{D}_v + \mathbf{T}_w \mathbf{D}_w \quad (\text{B.13})$$

where

$$\mathbf{T}_u = \begin{bmatrix} \mathbf{T}_{u,aa} & \mathbf{T}_{u,a\theta} & \mathbf{T}_{u,a\phi} \\ \mathbf{T}_{u,\theta a} & \mathbf{T}_{u,\theta\theta} & \mathbf{T}_{u,\theta\phi} \\ \mathbf{T}_{u,\phi a} & \mathbf{T}_{u,\phi\theta} & \mathbf{T}_{u,\phi\phi} \end{bmatrix} \quad (\text{B.14})$$

$$\mathbf{D}_u = \begin{bmatrix} \text{diag}(\mathbf{d}_u) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \text{diag}(\mathbf{d}_u) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{diag}(\mathbf{d}_u) \end{bmatrix}. \quad (\text{B.15})$$