

# Abstract

The study of gravity and magnetism has a long history in Earth sciences and continues to play an important role in exploration geophysics. Fundamental knowledge about our planet has been gained from the processing of potential field data at all scales: from large tectonic processes, down to the size of mineral deposits. Attempting to model density and magnetization through the inverse process remains challenging however as it is a highly non-unique problem. Assumptions about physical property contrasts are needed in order to obtain a solution that is geologically meaningful.

A major limitation encountered with current inversion methodologies is in assessing the range of possible models within a deterministic framework. It is difficult to generate a suite of solutions that are sufficiently different while remaining geological plausible. Conventional unconstrained inversions generally yield either smooth models with poor edge definition or sparse and compact models that are too simplistic to represent subtitle features. Geological constraints can be incorporated in the inversion, but building and testing different scenarios are time-consuming and can rapidly become intractable. Thus, in current practise, geological interpretation is often made by using one, or a small number of realizations that represent a narrow portion of possible answers.

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

# Preface

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

The mixed norm algorithm presented in Chapter 3 and part of the learning algorithm presented in Chapter 6 resulted in the ~~published~~ article Fournier and Oldenburg (2019b) and presented at a conference (Fournier et al., 2016). The same algorithm has been used by other researchers and resulted in three research papers to which I am also co-author (Abedi et al., 2018a,b; Miller et al., 2017).

Improvements to the magnetic vector inversion in Chapter 4 will be featured in the accepted paper Fournier and Oldenburg (2019a), currently in the second round of revisions. I have co-authored a research paper investigating the geothermal resources at Mount Baker (Schermerhorn et al., 2017). Using the same methodology, I collaborated in the implementation of sparse vector inversion applied to self-potential problems to map hydrothermal circulation at Mount Tongariro, New Zealand (Miller et al., 2018).

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

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

optional argument to \caption  
allows specification of short  
versions to appear in the LoT &

## List of Tables

LoT

Table 2.1	Summary table for the forward modeling of a magnetized sphere using a Tensor and Octree discretization . . . . .	21
Table 3.1	Norm values and proportionality ratio obtained for the 1D solution presented in Figure 3.5(c). A proportionality ratio of $\lambda_\infty \approx 1$ indicates that the components of the regularization function are both contributing significantly to the final solution. . . . .	38
Table 3.2	IRLS algorithm in pseudo-code made of two stages: Stage 1 Initialization with convex least-squares inversion, Stage 2 IRLS updates with inner $\beta$ -search steps. . . . .	41
Table 3.3	Inversion summary after convergence of the S-IRLS for various cooling rates $\eta$ as presented in Figure 3.16. Each inversion trial was required to reach the target misfit $\phi_d^*$ and target $\varepsilon^* = 1e-6$ . . . . .	55
Table 7.1	Summary table grouping the various lithological units logged from bore hole. Expected density and magnetic susceptibility contrasts are derived from Figure 7.2. . . . .	116
Table 7.2	Intervals along boreholes KV200 and KV297 reporting significant remanent magnetization (Montonen, 2012) . . . . .	119

# Chapter 1

## Introduction

### 1.1 Motivation

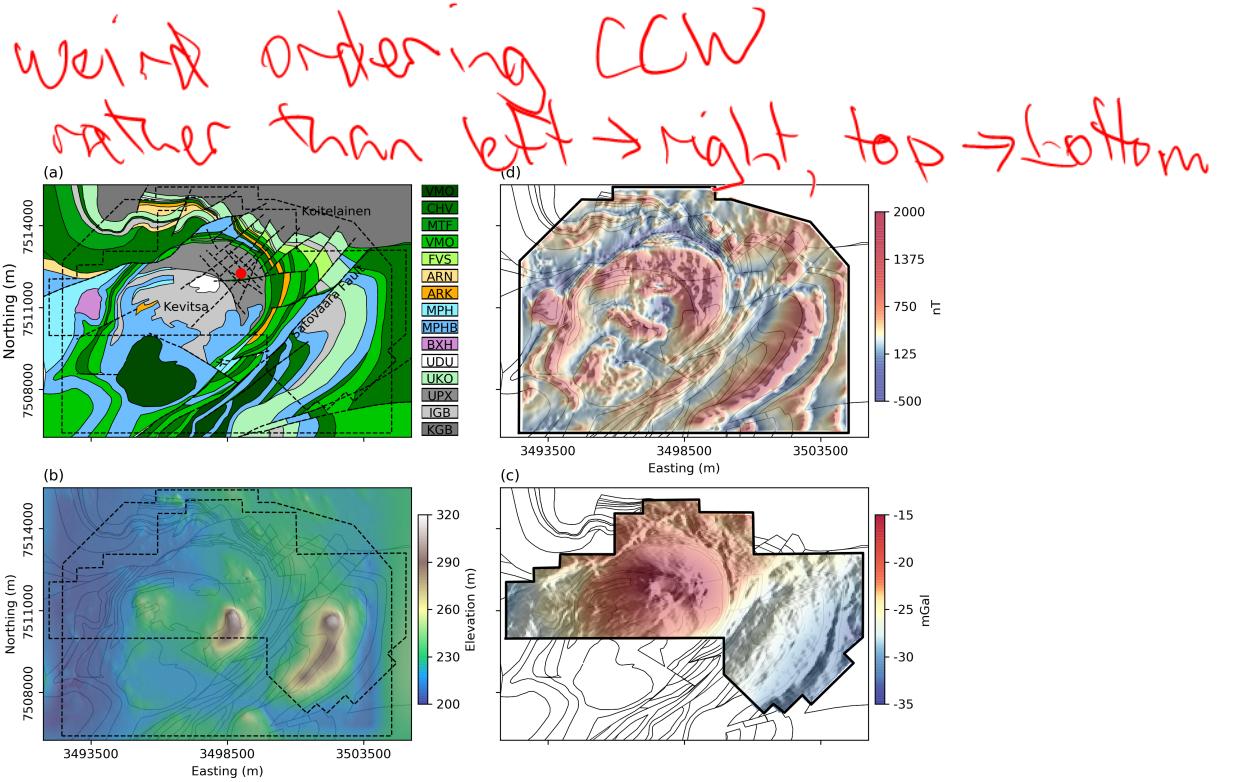
Density and magnetic permeability are important physical properties used in mineral exploration. Geologists routinely measure these properties on rock samples, along with textural and structural features, to target areas for further investigation. The vast majority of known mineral deposits have been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

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

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

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

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



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

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

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

### 1.1.3 Exploring the solution space

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

Sparsity assumptions have increased the flexibility of inversion algorithms for the modelling of compact targets with sharp edges. The usual strategy is to define global parameters applied over the entire model space. In most geological settings, however, we can expect to encounter a mix of features, either smooth and sparse in one region but elongated and blocky elsewhere. Determining exactly which assumption to use in a particular geological setting remains largely user-driven. Sun and Li (2014) have made inroads in imposing variable sparsity assumptions, either the  $\ell_2$  or  $\ell_1$ -norm measure of the model, to different regions of a 2D seismic tomography problem. They demonstrate that this choice could be automated based on information present in the data. Regions reacting favourably to the sparsity assumption are defined through a learning process.

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

#### 1.1.4 Dealing with remanence

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

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

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

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

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

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

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

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

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

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

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

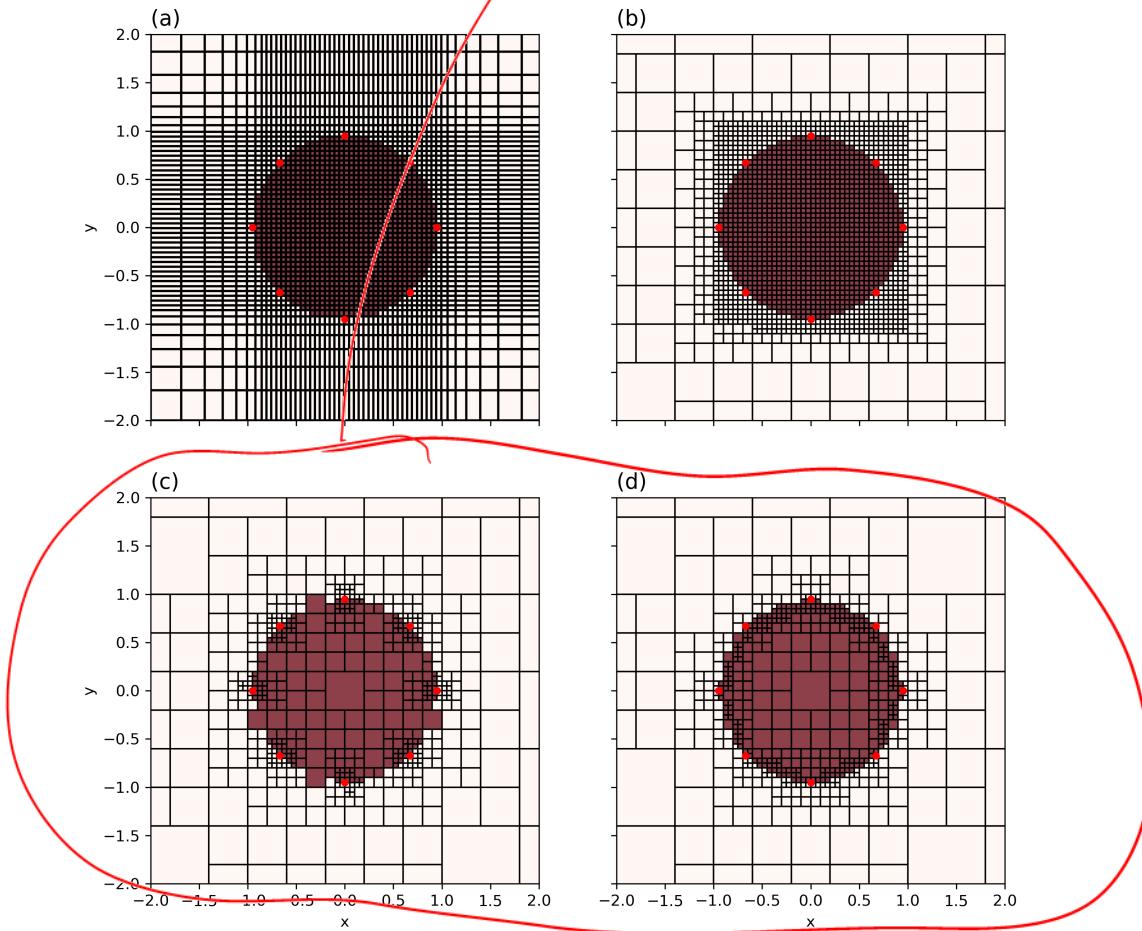
$$\mathbf{b} = \mathbf{T}\mathbf{m}, \quad (2.5)$$

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

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

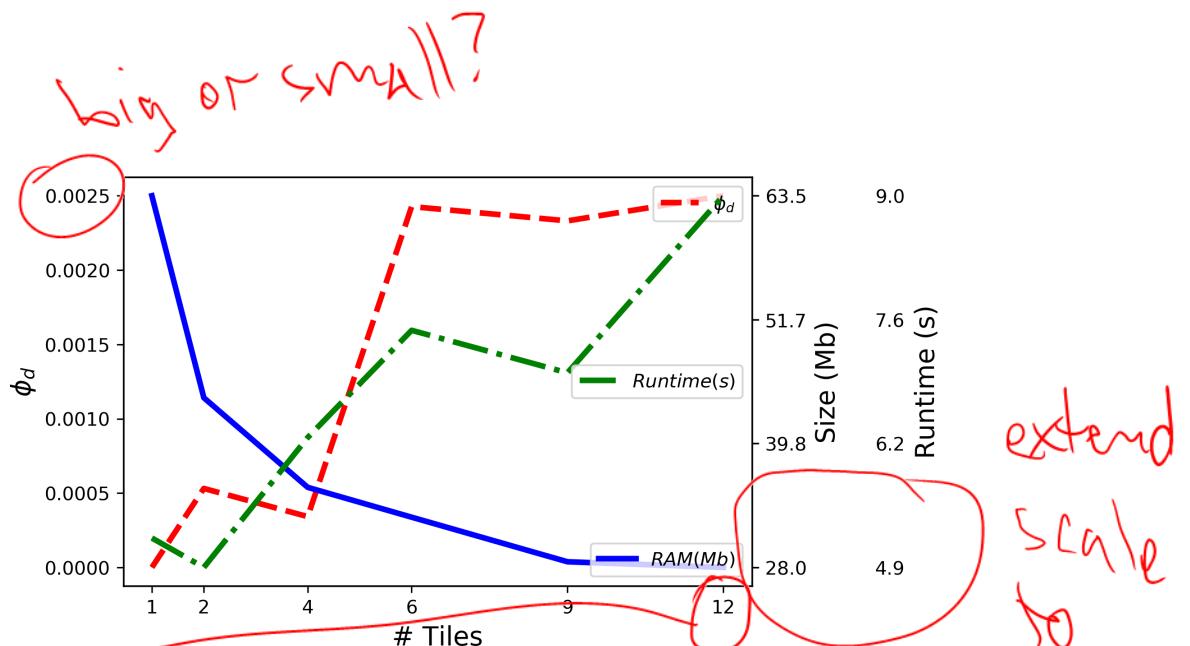
where  $\mu_0$  is the magnetic permeability of free-space. It is important to note that the tensor  $\mathbf{T}$  is a symmetric matrix with zero trace. Therefore only five of the nine

Why asymmetries?

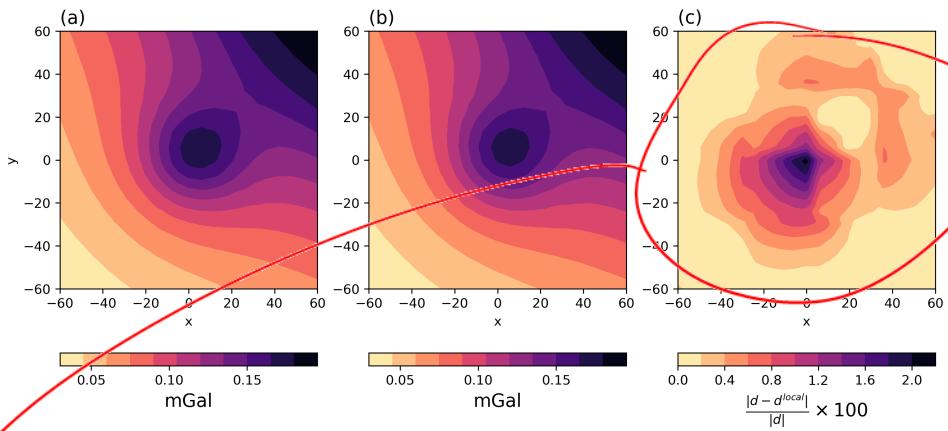


**Figure 2.4:** (a) Discretization of a sphere defined by discrete points (red) using the conventional Tensor mesh with a core region and padding cells. Octree meshes refined by (b) box, (c) radial and (d) surface methods from the SimPEG.discretize library.

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

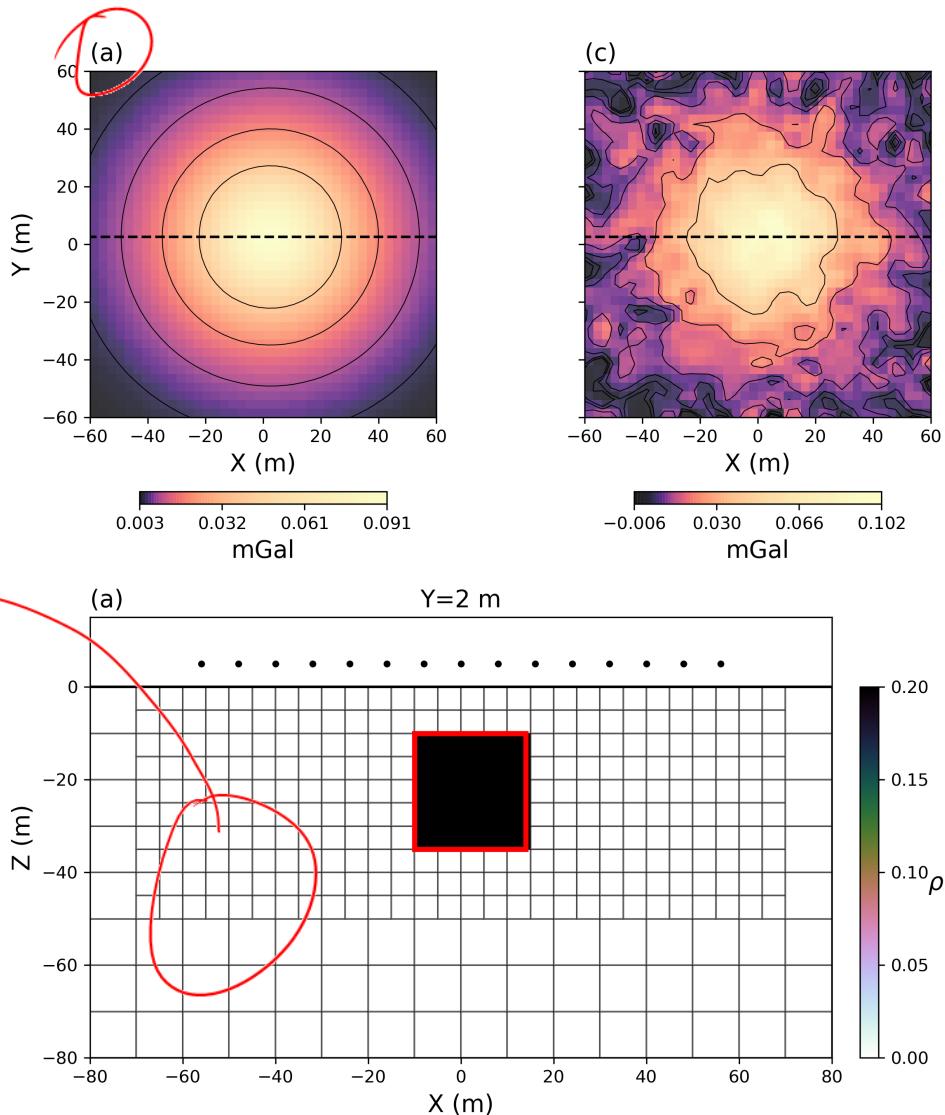


**Figure 2.7:** Trade-off curves between the size of the forward problem (blue) and data residual (red) over a range of number of tiles. The total size of the problem is calculated based on the sum of cells in all the local meshes times the number of data. The optimal number of tiles would be at the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



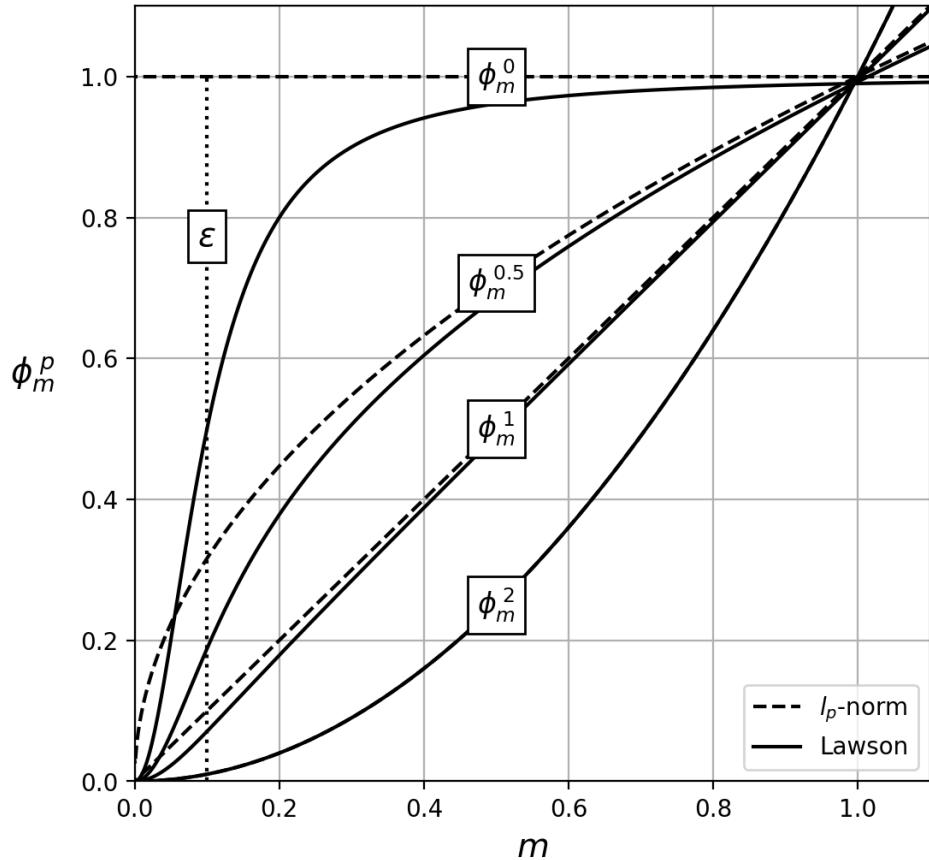
**Figure 2.8:** Simulated gravity data calculated from (a) the global model and (b) the 16 forward tiled calculations. (c) Data residuals show short wavelength discrepancies between adjacent tiles due to interpolation effects.

what makes this short not long?



**Figure 3.1:** (a) Vertical section through a 25 m cube with uniform density  $\rho = 0.2$  g/cc. (b) Simulated gravity data responses on a  $21 \times 21$  survey grid placed 15 m above the anomaly. (c) Gravity data with random Gaussian noise added,  $10^{-3}$  mGal standard deviation.

From the noisy data I will attempt to recover the block anomaly by the inverse

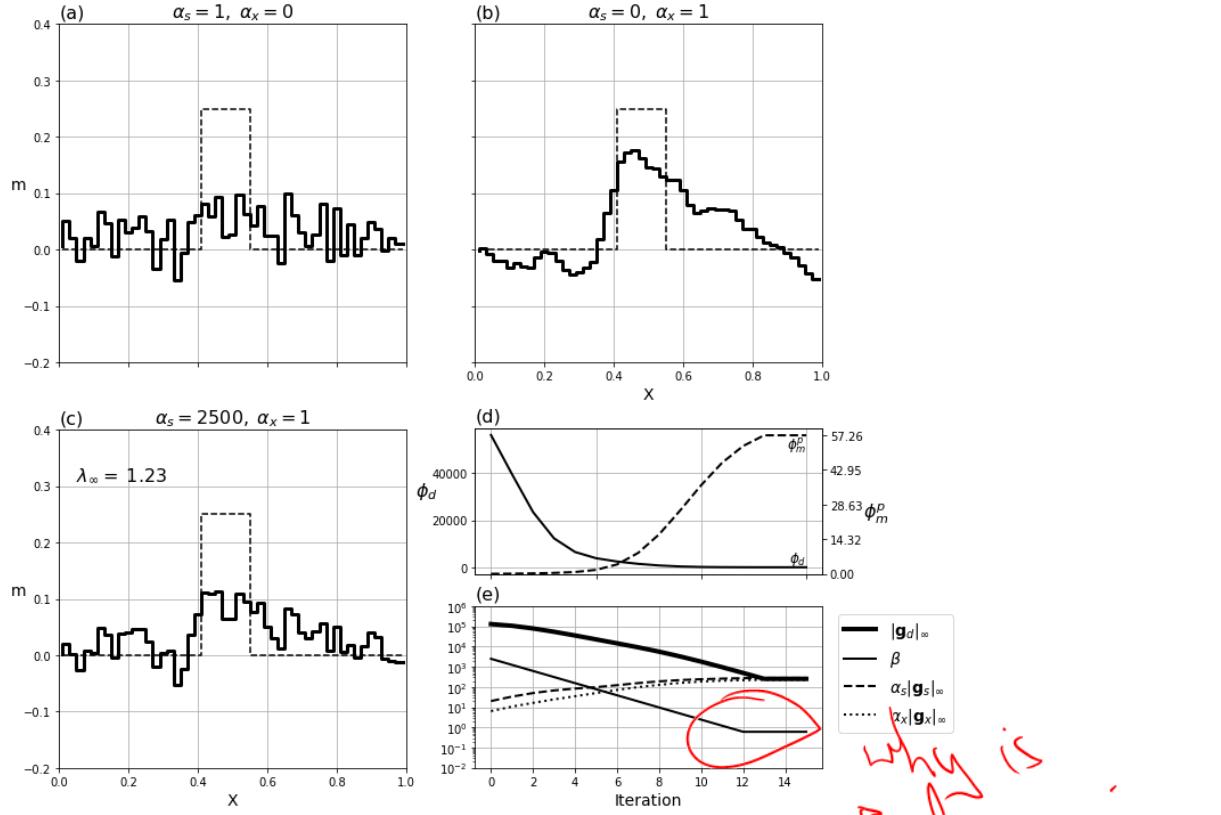


**Figure 3.3:** Approximated  $\ell_p$ -norm using the Lawson measure (Lawson, 1961) over a range of  $p$ -values and for a fixed threshold parameter  $\epsilon = 10^{-1}$ .

tials, sinusoidals), but my choice will serve to highlight the effects of various regularization functions. I generate 10 data, so  $\mathbf{F} \in \mathbb{R}^{N \times M}$  where  $M = 50$  and  $N = 10$ . Random Gaussian error  $\mathbf{e}$  ( $\sigma=0.025$ ) is added to simulate noise (Fig. 3.4(c)).

To begin my analysis, I invert my synthetic dataset with two simple regularization functions. Figure 3.5(a) presents the recovered model after reaching the target misfit ( $\phi_d^* = N$ ) using the smallness term alone ( $\alpha_x = 0$ ). The solution exhibits high variability similar to the stochastic kernel function, but model parameters remain near the zero reference value. Next, I invert the data using the model gradient term ( $\alpha_s = 0$ ); this yields the smoother model presented in 3.5(b). The solution shows

what is  
reference  
model?



**Figure 3.5:** Solution to the 1D inverse problem using (a) an  $\ell_2$ -norm on the model ( $\alpha_x = 0$ ), (b) the  $\ell_2$ -norm on model gradients ( $\alpha_s = 0$ ) and (c) combined regularization function ( $\alpha_s = 2500$ ,  $\alpha_x = 1$ ). (d) Convergence curve comparing the misfit ( $\phi_d$ ) and the regularization ( $\phi_m$ ) as a function of iterations. (e) Comparative plot for the relative contribution of the different components of the objective function measured in terms of maximum absolute gradient ( $\|\mathbf{g}_i\|_\infty$ )

gradients I use the infinity-norm

$$\|\mathbf{g}_r\|_\infty = \|\nabla_m \phi_t\|_\infty \quad (3.15)$$

The  $\|\mathbf{g}_r\|_\infty$  metric is appealing for a few reasons: (a) it is directly linked to the minimization process because I use gradient descent methods, (b) it does not depend on the dimension  $M$  of the parameter space as do other measures that involve a sum of components of the vector, (c) the theoretical maximum can be calculated

only  
 $\|\mathbf{g}_d\|_\infty$

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

$$\phi_m = \sum_{r=s,x} \int_V \frac{f_r(m)^2}{\left(f_r(m)^2 + \varepsilon_r^2\right)^{1-p_r/2}} dV, \quad (3.18)$$

This measure makes the inverse problem non-linear with respect to the model. The common strategy is to linearized the function through the Iterative Reweighted Least-Squares (IRLS) approach where the denominator is replaced by model parameters from the most recent iteration. The smallest model component can be written as:

$$\phi_s^{p_s} = \sum_i \frac{m_i^2}{((m_i^{(k-1)})^2 + \varepsilon_s^2)^{1-p_s/2}} V_i \quad (3.19)$$

where  $m_i^{(k-1)}$  are model parameters obtained at a previous iteration and  $V_i$  are volume terms connected with the discretization. In (3.19) I have explicitly written the objective function as  $\phi_s^{p_s}$  to indicate that I am evaluating a smallest model component with an  $\ell_p$ -norm with  $p = p_s$ . The approximated norm can be expressed in linear form as:

$$\phi_s^{p_s} = \|\mathbf{V}_s \mathbf{R}_s \mathbf{m}\|_2^2. \quad (3.20)$$

$$\mathbf{R}_s = \text{diag} \left[ \left( (\mathbf{m}^{(k-1)})^2 + \varepsilon_s^2 \right)^{p_s/2-1} \right]^{1/2}, \quad (3.21)$$

Carrying out the same IRLS operation on the derivative component yields

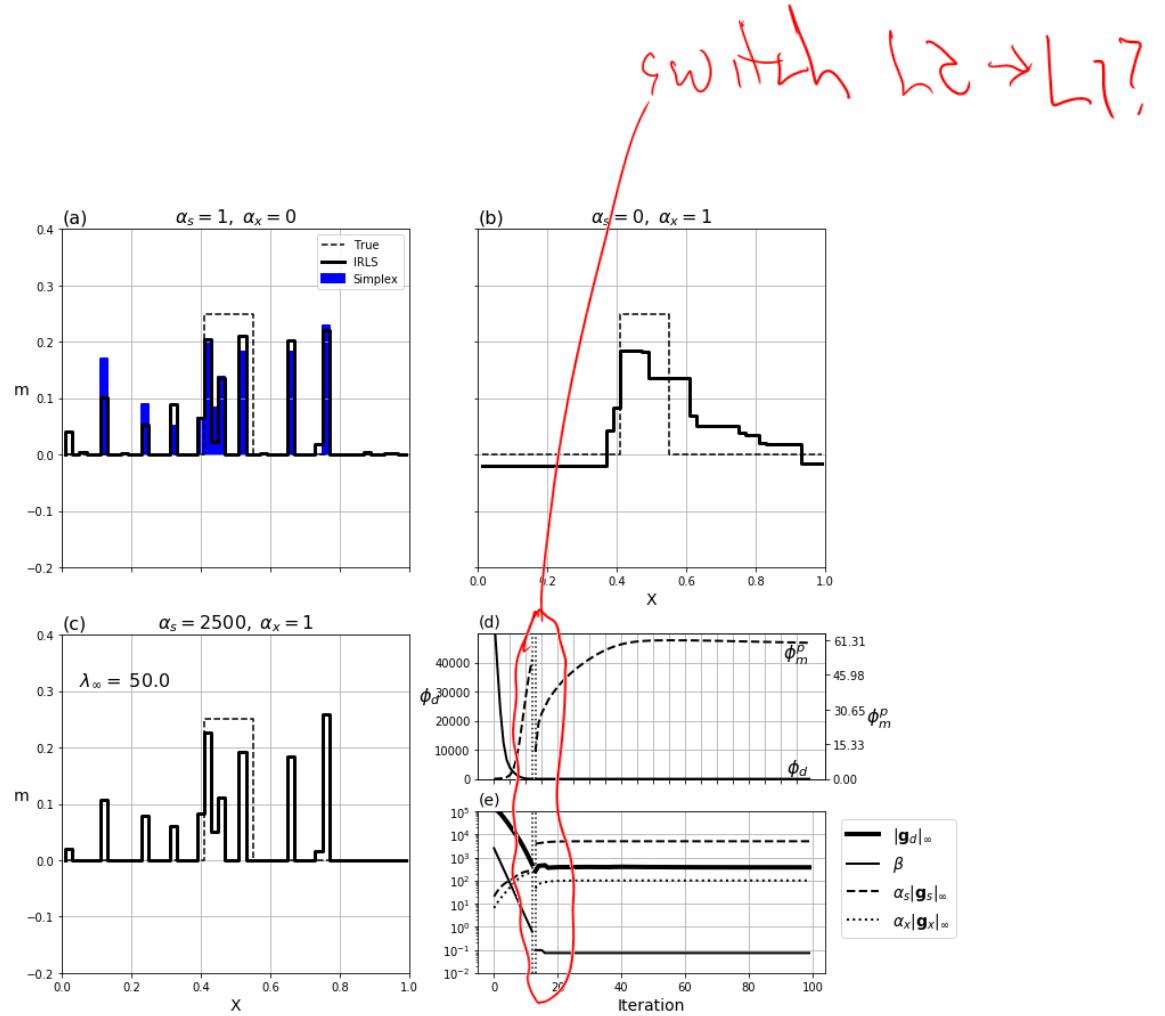
$$\phi_x^{p_x} = \|\mathbf{V}_x \mathbf{R}_x \mathbf{G}_x \mathbf{m}\|_2^2, \quad (3.22)$$

where the IRLS weights are calculated by:

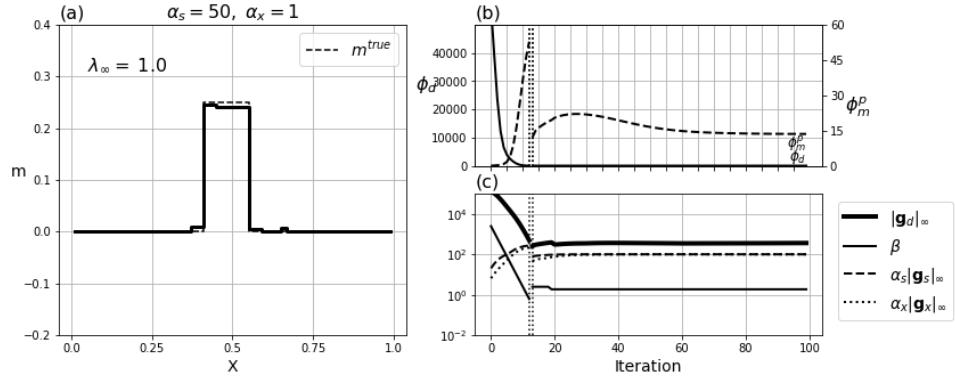
$$\mathbf{R}_x = \text{diag} \left[ \left( (\mathbf{G}_x \mathbf{m}^{(k-1)})^2 + \varepsilon_x^2 \right)^{p_x/2-1} \right]^{1/2}. \quad (3.23)$$

The final regularization function is thus

$$\phi_m^p = \alpha_s \|\mathbf{V}_s \mathbf{R}_s \mathbf{m}\|_2^2 + \alpha_x \|\mathbf{V}_x \mathbf{R}_x \mathbf{G}_x \mathbf{m}\|_2^2. \quad (3.24)$$



**Figure 3.6:** (a) Two solutions using an  $\ell_1$ -norm on the model: (blue) Simplex, and (black) IRLS method. (b) Solution obtained with the approximated  $\ell_1$ -norm (IRLS) penalty on model gradients alone and (c) with the combined penalty functions ( $\alpha_s = 2500, \alpha_x = 1$ ). The calculated proportionality ratio  $\lambda_\infty$  indicates that the combined penalties is dominated by the  $\phi_s^1$  term. (d) Convergence curve and (e) maximum partial derivatives associated with the components of the objective function as a function of iterations for the inversion in (c).



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

this, I replace the gradient term by a finite difference so that

$$\phi_x^{P_x} = \|\nabla_x \mathbf{R}_x \mathbf{D}_x \mathbf{m}\|_2^2 \quad (3.29)$$

where

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

over weight on  
small cells?

The IRLS weights are then defined as:

$$\mathbf{R}_x = \text{diag} \left[ \left( (\mathbf{D}_x \mathbf{m}^{(k-1)})^2 + \epsilon_x^2 \right)^{p_x/2-1} \right]^{1/2}, \quad (3.31)$$

Replacing the gradient with a finite difference makes  $\phi_m$  and  $\phi_x$  dimensionally equivalent and the default scaling values can be reset to  $\alpha_s = \alpha_x = 1$ . Re-inverting my 1D problem with this new strategy results in the solution and convergence curves presented in 3.8. For the uniform discretization in this example, both the  $\alpha_s$  scaling and the finite difference approach yields identical results (Figure 3.7, 3.8(a)). In both cases  $\phi_m^P$  changes rapidly between Stage 1 and 2, associated with

what is the gradient matrix if  
not finite difference?  
44

on a 2D seismic tomography problem. They demonstrated the benefits of dividing model space into regions with different  $\ell_p$ -norm penalties. The choice of norms was limited to be either  $l_1$  or  $l_2$ . Little has been published however on the independent mixing of model and gradient norms on the range  $p \in [0, 2]$ , although this problem was initially addressed in (Fournier, 2015). I now apply my algorithm to minimize

$$\phi_m^p = \alpha_s \phi_s^0 + \alpha_x \phi_x^2. \quad (3.32)$$

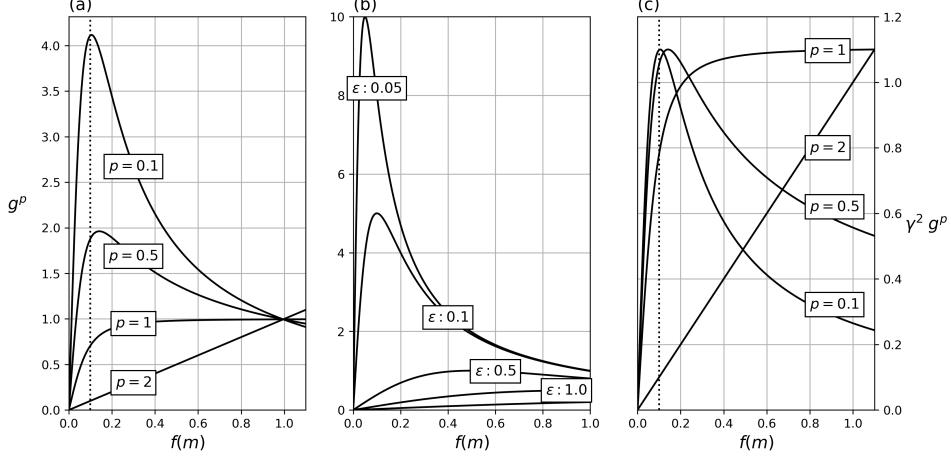
Based upon my previous work, I expect the solution to be sparse, in terms of non-zero model parameters, while smooth with respect to the model gradients. Unfortunately, following the current IRLS strategy, I recover the model presented in Figure 3.10(a). The anomaly is concentrated near the boxcar but appears to be dominated by  $\phi_s^0$ . There seems to be only marginal influence from  $\phi_x^2$ . Comparing the partial derivatives of the objective function confirms this. After convergence the calculated proportionality ratio is  $\lambda_\infty = 159$ . This is a significant change from the end of Stage 1 where  $\lambda_\infty \approx 1$ . Clearly, iteration 6, at Stage 2 of the IRLS, took the solution away from the proportionality condition (Fig 3.10(c)). I hypothesize that a more desirable solution could be obtained if proportionality was preserved among the components of the objective function throughout the IRLS process. In the following sections, I provide an important modification to the standard IRLS algorithm to achieve this goal.

### 3.2.1 Scaled-IRLS steps

Since the inverse problem is solved using gradients of the composite objective function  $\phi(\mathbf{m})$ , the relative magnitude of the individual gradients is a driving force in controlling the iteration step in (3.25). I want to ensure that each penalty term in the objective function is playing a significant role. Taking the partial derivatives of the linearized Lawson norm as prescribed in (3.19) yields:

$$g^p = \frac{\partial \phi^p}{\partial m} = \frac{f(m)}{(f(m^{(k-1)})^2 + \varepsilon^2)^{1-p/2}} V. \quad (3.33)$$

From Figure 3.11(a), I note that the magnitude of the derivatives increases rapidly for small  $p$  values as  $m_i \rightarrow 0$ . This trend is accentuated for small  $\varepsilon$  values as



**Figure 3.11:** Derivatives of the Lawson approximation over a range of model values for (a) a fixed threshold parameter  $\epsilon = 10^{-1}$  over a range of  $p$  values and for (b) a fixed  $p = 0$  over a range of  $\epsilon$  values. (c) Applying the  $\gamma$ -scaling to the gradients brings all maximums to be equal irrespective of  $p$  and  $\epsilon$ .

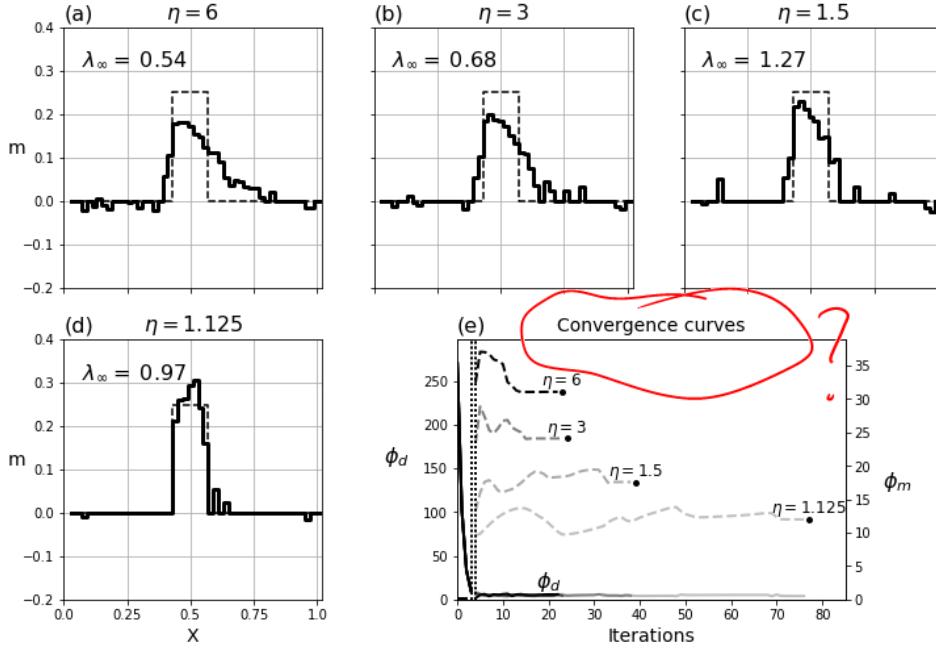
from which I can calculate  $\|g^p\|_\infty$  by substituting  $f(m)^*$  into (3.33). Figure 3.11(c) presents the derivatives for different approximated  $\ell_p$ -norms after applying the corresponding  $\gamma$ -scale. Note that the largest derivative of any norm is at most as large as the  $\ell_2$ -norm penalty for  $m \in \mathbb{R}$ . The scaling of the norm derivatives guarantees that two penalties can co-exist and impact the solution at every step of the IRLS, regardless of the chosen  $\{p, \epsilon\}$ -values.

I therefore define Scaled-IRLS weights such that (3.21) and (3.31) become:

$$\begin{aligned} \mathbf{R}_s &= \gamma_s \operatorname{diag} \left[ \left( (\mathbf{m}^{(k-1)})^2 + \epsilon_s^2 \right)^{p_s/2-1} \right]^{1/2} \\ \mathbf{R}_x &= \gamma_x \operatorname{diag} \left[ \left( (\mathbf{D}_x \mathbf{m}^{(k-1)})^2 + \epsilon_x^2 \right)^{p_x/2-1} \right]^{1/2}, \end{aligned} \quad (3.36)$$

where the scaling parameters  $\gamma_s$  and  $\gamma_x$  are:

$$\gamma_s = \left[ \frac{\|\mathbf{g}_s^2\|_\infty}{\|\mathbf{g}_s^{p_s}\|_\infty} \right]^{1/2}, \quad \gamma_x = \left[ \frac{\|\mathbf{g}_x^2\|_\infty}{\|\mathbf{g}_x^{p_x}\|_\infty} \right]^{1/2}, \quad (3.37)$$



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

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

**Table 3.3:** Inversion summary after convergence of the S-IRLS for various cooling rates  $\eta$  as presented in Figure 3.16. Each inversion trial was required to reach the target misfit  $\phi_d^*$  and target  $\varepsilon^* = 1e-6$ .

as  $\eta \rightarrow 1$ . I found experimentally that for  $\eta \approx 1.25$  generally yielded an optimal trade-off between computational time (number of iterations) and convergence to a suitable solution.

### 3.2.3 Summary

My goal is to solve an inverse problem where the regularization function is composed of multiple terms, each defined as an  $\ell_p$ -norm premultiplied by a scaling parameter. The scaling parameters,  $\alpha$ 's, are used to control how much each component contributes to the final solution. The relative influence of these components is quantified by evaluating the proportionality ratio  $\lambda_\infty$  (3.16). If two components contribute equally, then  $\lambda_\infty$  should be close to unity. Unfortunately, when the components of the regularization include model and gradient terms, the scaling is affected by the cell size chosen for discretization and by the  $\ell_p$ -values. To ameliorate this I remove the length scales from the measure of model gradients and replace gradients with a finite difference. This makes  $\phi_s^{p_s}$  and  $\phi_x^{p_x}$  dimensionally equivalent. The default values for obtaining equal contributions are thus  $\alpha_s = \alpha_x = 1$  for all combinations of  $\ell_p$ -norms on the gradients.

I solve the inverse problem by replacing the  $\ell_p$ -norms with their Lawson norm approximations. Thus I search for the model that minimizes

$$\begin{aligned} \phi(m) = & \|\mathbf{W}_d \Delta \mathbf{d}\|_2^2 + \\ & \beta \alpha_s \sum_{i=1}^M \frac{m_i^2}{(m_i^2 + \varepsilon_s^2)^{1-p_s/2}} V_{s_i} + \\ & \beta \alpha_x \sum_{i=1}^{M-1} \frac{(m_{i+1} - m_i)^2}{((m_{i+1} - m_i)^2 + \varepsilon_x^2)^{1-p_x/2}} V_{x_i}, \end{aligned} \quad (3.39)$$

for arbitrarily small  $\varepsilon_s, \varepsilon_x$  values. I solve the inverse problem using a two-stage approach. I first find a solution for the  $\ell_2$ -norm problem and then I change the objective functions to their final desired  $\ell_p$ -norms and solve the optimization problem using IRLS. To keep stability in the iterative process I successively rescale the

why should fit & smoothness  
matter the same amount?

IRLS weights  $\mathbf{R}$ . Thus at each iteration I solve a locally convex problem

$$\begin{aligned}\phi(m^{(k)}) = & \|\mathbf{W}_d \Delta \mathbf{d}\|_2^2 + \\ & \beta (\alpha_s \|\mathbf{V}_s \mathbf{R}_s \mathbf{m}\|_2^2 + \alpha_x \|\mathbf{V}_x \mathbf{R}_x \mathbf{D}_x \mathbf{m}\|_2^2) ,\end{aligned}\quad (3.40)$$

Although the local minimization problems involve scaled gradients, the final desired solution is that which minimizes (3.39). My ability to achieve this goal depends upon the value of  $\varepsilon$  and the chosen cooling rate. I find that the best (i.e. minimum norm) solution is obtained when  $\varepsilon$  is cooled slowly to a final small value. If the cooling is too fast then I obtain a substandard solution in which  $\lambda_\infty$  is not close to unity and my modelling objectives are not satisfied. Slower cooling and the frequent re-scaling of the gradients keeps the proportionality ratio near unity.

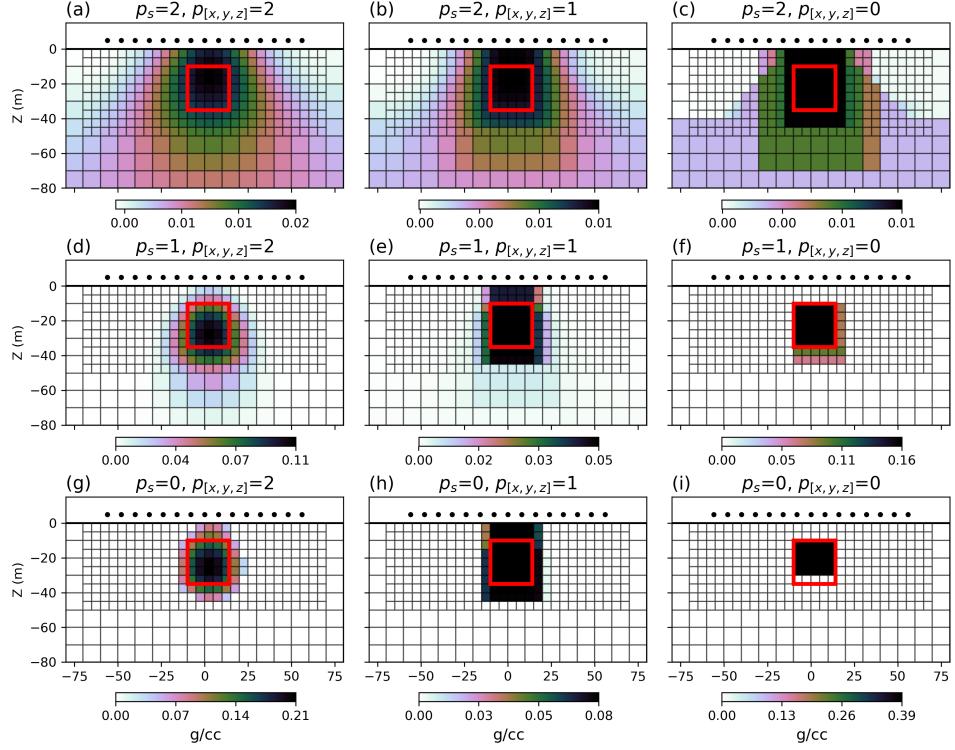
### 3.3 Exploring the model space

The smooth density model presented in Figure 3.2 was a poor approximation of a compact block, but it is one of many possible solutions. Now that I have developed an algorithm that can combine multiple regularization functions with different  $\ell_p$ -norm measure, I can explore the model space by generating a suite of solutions that have variable characteristics. I will demonstrate this on the synthetic gravity example shown in Figure 3.1. The function to be minimized for the 3D gravity problem becomes

$$\begin{aligned}\min_{\rho} \phi(\rho) = & \|\mathbf{G} \rho - \mathbf{d}^{obs}\|_2^2 + \beta \sum_{r=s,x,y,z} \alpha_r \|\mathbf{W}_r \mathbf{R}_r \mathbf{D}_r \rho\|_2^2 \\ \text{s.t. } & \phi_d \leq \phi_d^*\end{aligned}\quad (3.41)$$

I carry out eight additional inversions. I use a combination of norms on a range of  $p_s, p_{[x,y,z]}, \in [0, 1, 2]$  values. I set  $p_x = p_y = p_z$  in all cases. The solutions, nine models in total, are presented in Figure 3.17. All models have a final misfit  $\phi_d^* \approx 441$  and use the same  $\ell_2$ -norm solution to initiate the IRLS steps. I make the following general observations. There is a progressive transition from a smooth model (upper left) to a blocky solution (lower right) as  $p_s, p_x, p_y$  and  $p_z$  decrease.

The top of the density high is most often recovered at 10 m depth. Away from the anomalous region the velocity is relatively smooth and close to the background

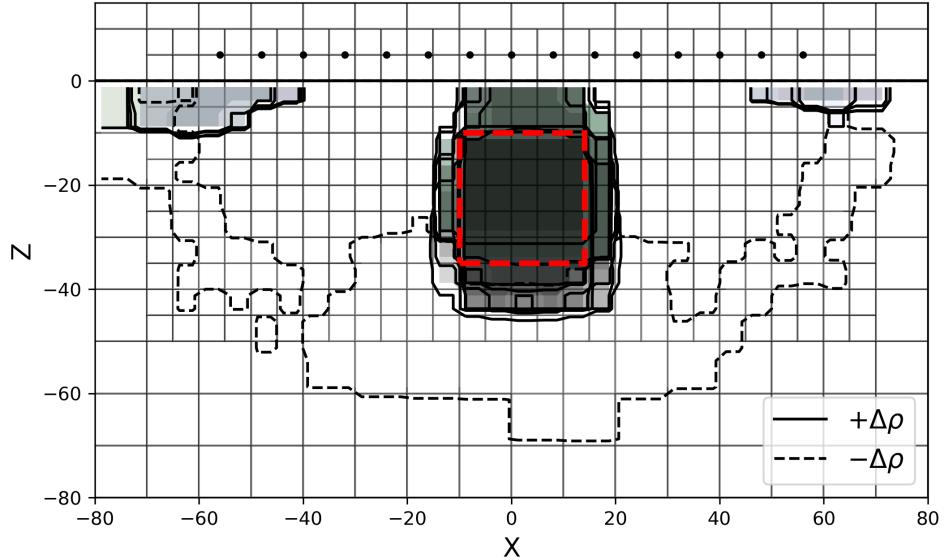


**Figure 3.17:** (a-i) Vertical section through a suite of density models recovered for varying  $\ell_p$ -norm penalties applied on the model and model gradients for  $p_s \in [0, 1, 2]$  and  $p_x = p_y = p_z \in [0, 1, 2]$ .

reference model of 0 g/cc. There is also a clear trend in the data misfit map such that the correlated residual decreases as  $p_s, p_x, p_y, p_z \rightarrow 0$ .

### 3.3.1 Interpretation

Accessing a range of solutions is important to assess the stability of different features and to avoid over-interpreting one specific realization. The next step requires to compare this ensemble of models and make a geological interpretation. In Chapter 6 I provide a more evolved methodology to extract local parameters, but for now, I will compare the solutions visually. Figure 3.18 presents an overlay for the 10th and 90th percentiles anomalous densities calculated from the suite of models shown in Figure 3.17. I can assess the robustness of features by comparing the



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

iso-contour lines of each model: tight clustering of the contours indicates that several models agree on the position of an edge, while a large spread indicates high variability. At the center of the model, I note that the top of the anomaly (solid) is highly correlated among models, but less so the bottom limit. This is expected as the resolution of the survey decreases with depth. Meanwhile, on the edges of the domain, the shape and sign of density contrasts vary substantially. From this simple analysis, I would assign high confidence on the horizontal and top of the positive density anomaly, and low confidence on features on either side of the inversion domain.

The normalized data residual maps for each inversion are shown in Figure 3.19. The decrease in correlated residual observed on the misfit maps (Figure 3.18) is also an important aspect to consider. While all the inversions have achieved the global target misfit, only after applying the proper constraints (blocky and compact)

of course the traditional assumption  
fails in this case

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

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

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

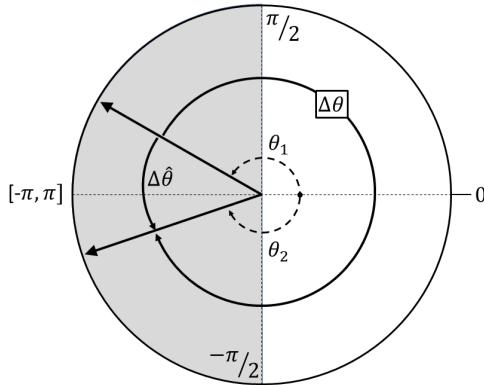
I begin with the conventional smooth assumptions ( $p_s, p_x, p_y, p_z = 2$ ) and attempt to recover the position and shape of the magnetic block. Since (4.2) is linear with respect to  $\kappa$ , I can use the same inversion methodology established in Chapter 3. The objective function to be minimized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

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

As I am dealing with strictly positive magnetic susceptibility  $\kappa$ , I impose bound constraints by the projected gradient method (Vogel, 2002). Model parameters  $\kappa_i$  that become negative are set to zero and ignored for the following Gauss-Newton step in (3.26). After reaching the target misfit criterion in equation (3.27), I recover the susceptibility model shown in Figure 4.2(a). I note that the position of the susceptibility anomaly is shifted to the side of the true block and appears to dip at 45° angle. This is due to the large negative data lobe introduced by the remanent component that I have purposefully ignored. Attempting to improve the solution by solving for a sparse model ( $p_s, p_x, p_y, p_z = 0$ ) yields the solution presented in Figure 4.2(c). The magnetic anomaly is imaged at the right depth and

just 1 step, or all <sup>62</sup> subsequent!

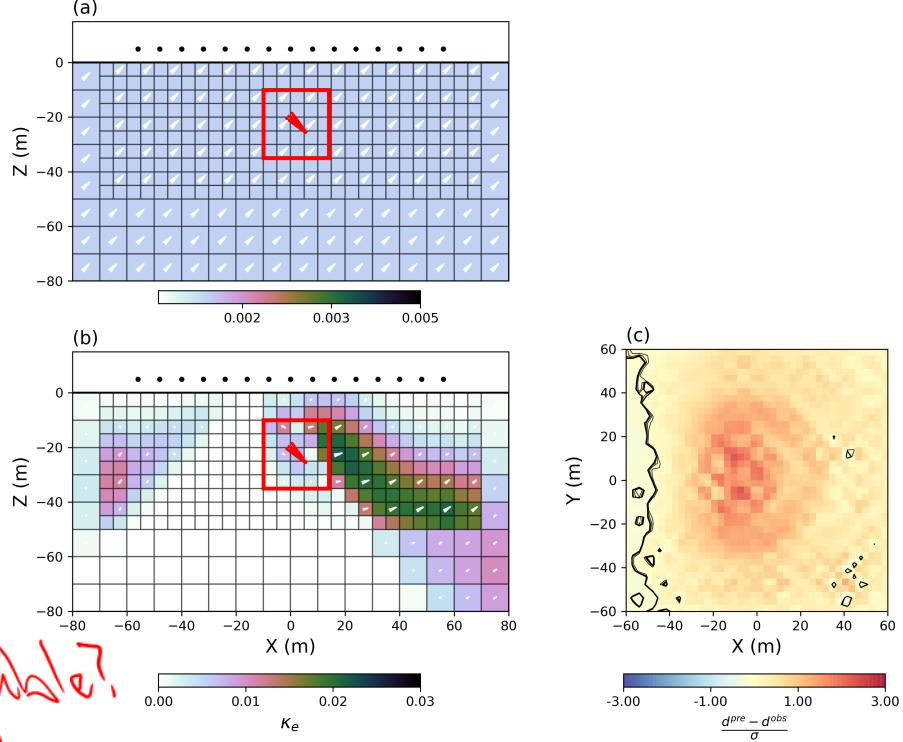
Thus the regularization only penalizes the change in angle between neighboring cells. Secondly, I perform an adjustment to the measure of angle differences to deal with the discontinuity at  $\theta = -\pi, \pi$  in order to prevent over penalizing angles that describe a similar orientation along the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



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

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

non-linear MVI-S formulation.



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

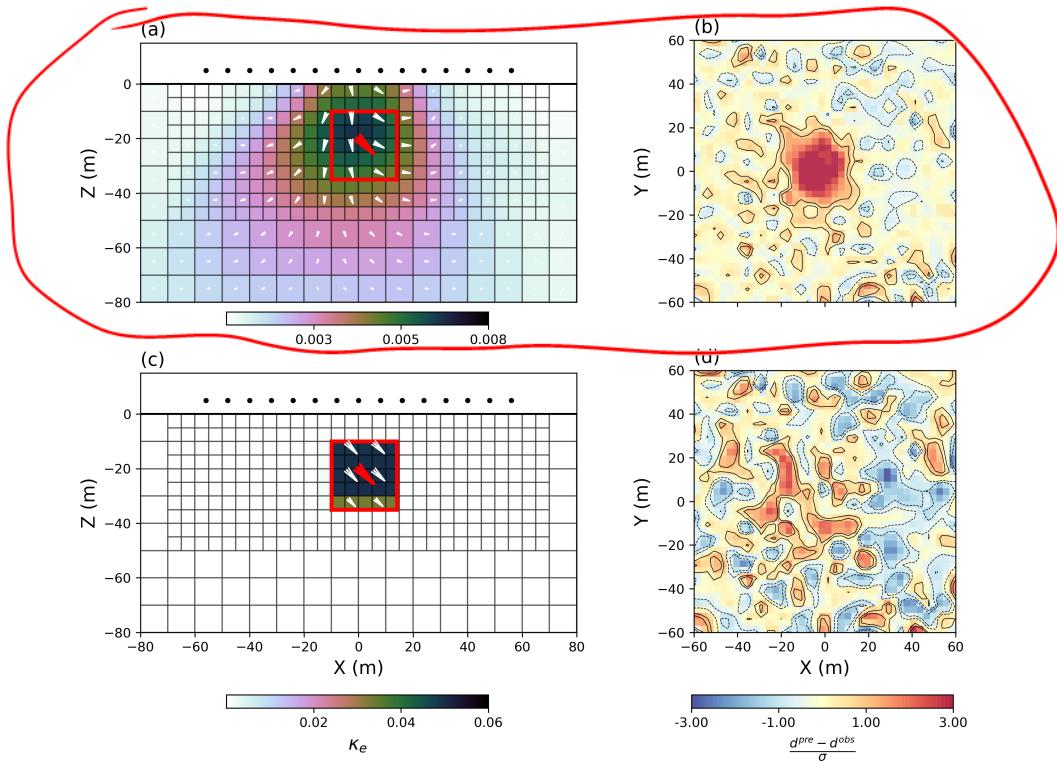
To gain some insight about the issues encountered with the MVI-S formulation, I consider a simpler two-parameter linear problem of the form

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

which I can express in matrix form as

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

*different from Fig 4.7 a&b how?*



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

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

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

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

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

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

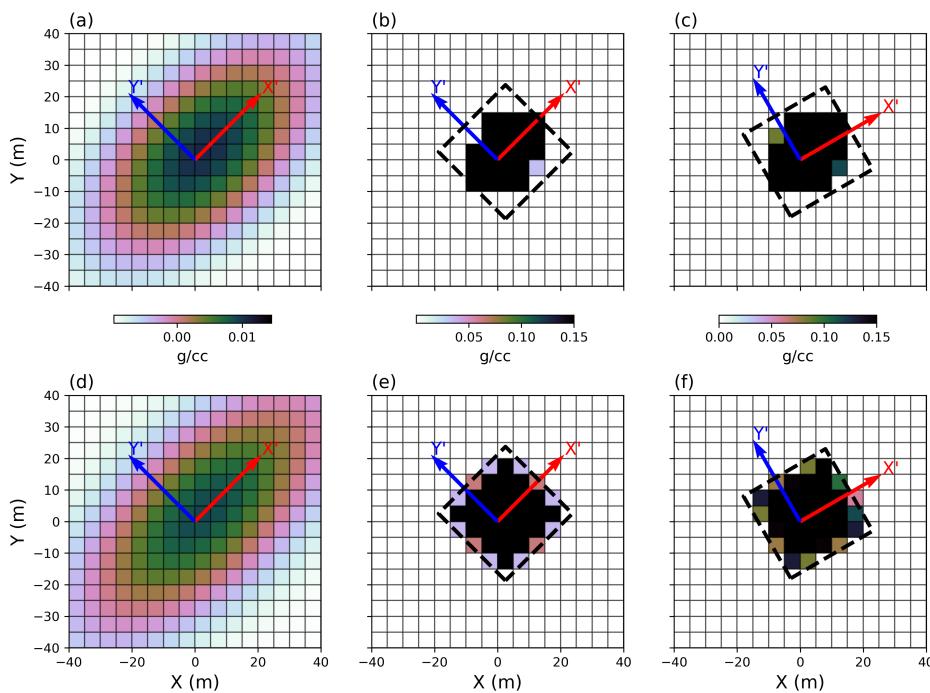
$$\begin{aligned}\phi_{m'} = & \mathbf{m}^\top \mathbf{W}_s^\top \mathbf{W}_s \mathbf{m} + \\ & \mathbf{m}^\top \frac{1}{8} \sum_{i=1}^8 \left( \mathbf{D}_{x,i}^\top \mathbf{W}_x^\top \mathbf{B}_{xx} \mathbf{W}_x \mathbf{D}_{x,i} + \right. \\ & \left. \mathbf{D}_{y,i}^\top \mathbf{W}_y^\top \mathbf{B}_{yy}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

*If you average  
forward & backward  
is it centered?*

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

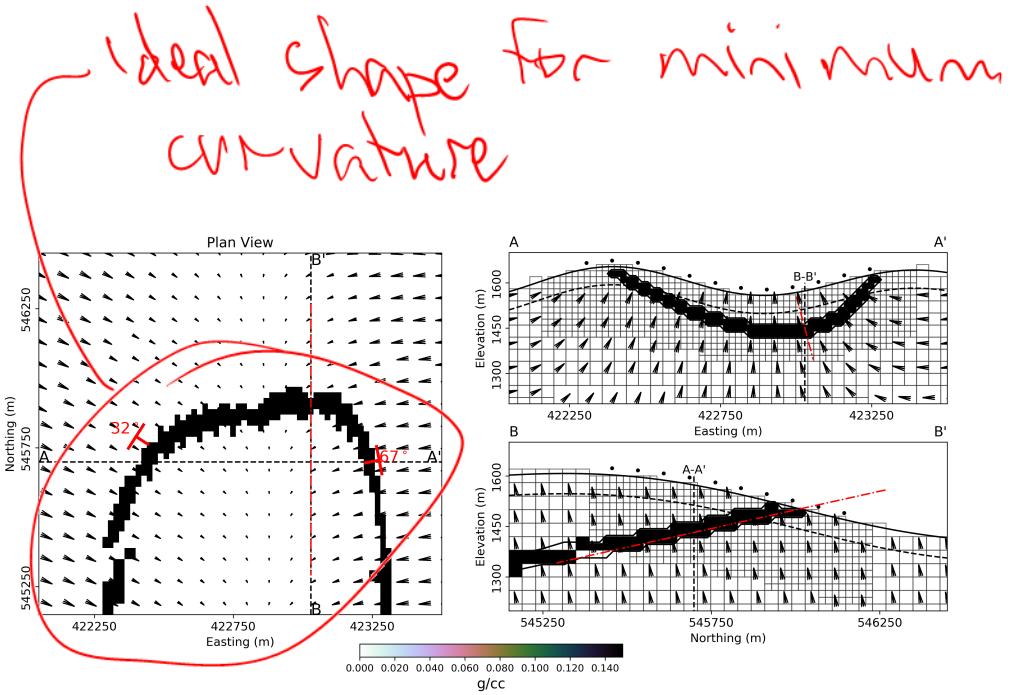
I showcase the effect of rotation with three inversions applied to the density

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



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

Highly artificial test case;  
what you really proved is that your  
sol'n method can be completely fooled  
by poor model parameters

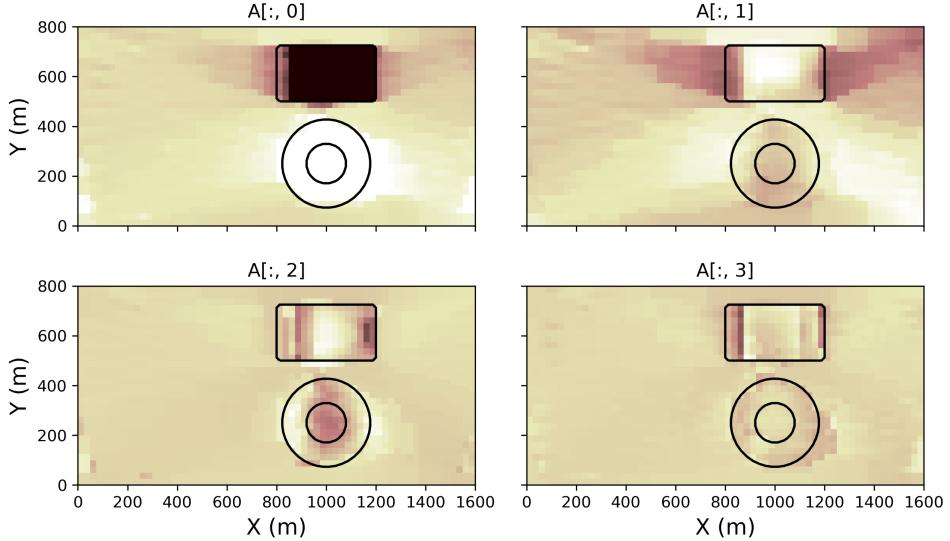


**Figure 5.9:** Interpolated dip and strike information used for the rotation of the objective function. The direction of the arrows define the estimated normal or the folded layer.

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

### 5.3 Summary

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



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

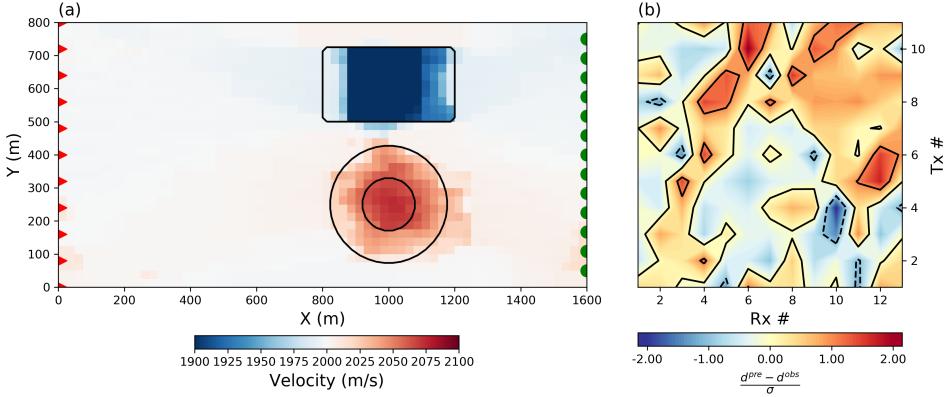
$\mathbb{R}^{nV \times nM}$ , also known as loadings, are scalars relating to how each principal component contributes to each model. I use the PCA algorithm from the open-source Python library `Scikit-Learn.decomposition.PCA` (Pedregosa et al., 2011). The number of principal components to be used in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

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

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

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

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



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

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

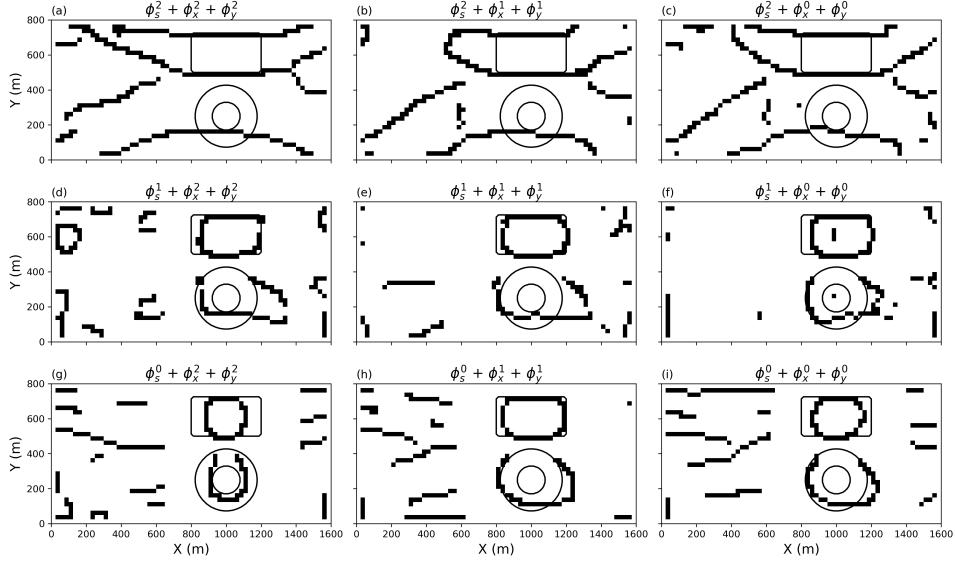
### 6.1.3 Parameter extraction

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

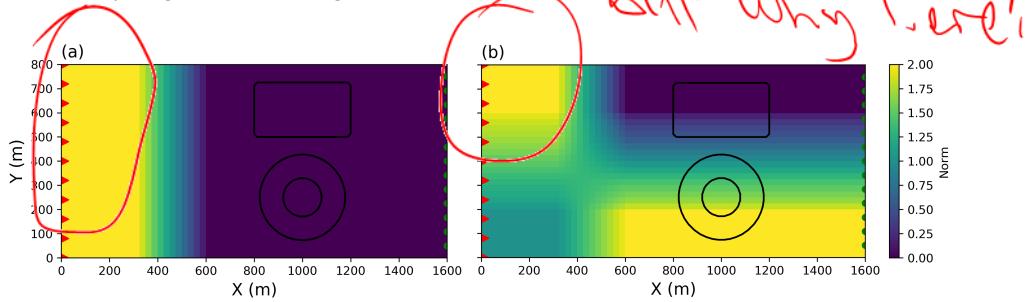
From this simplified representation of each model, I perform a moving window correlation  $r_{\mathbf{m}_i \mathbf{m}_P}$  between the average PCA model  $\mathbf{m}_A$  and each of the  $\mathbf{m}_i$  solutions:

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

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



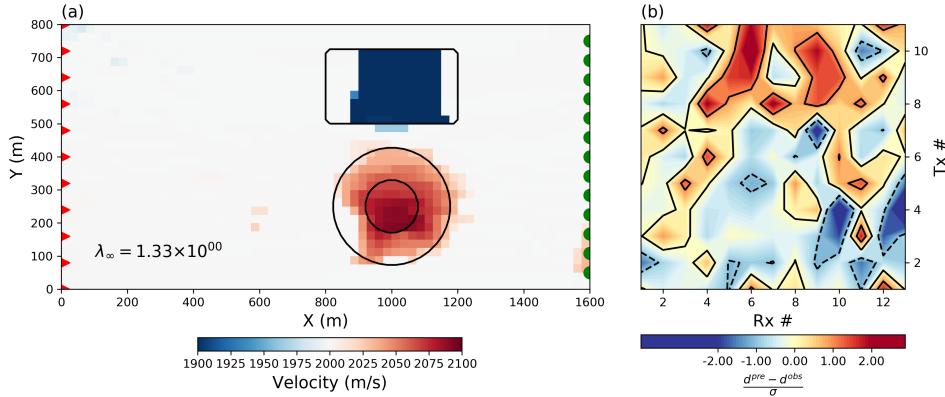
**Figure 6.6:** Parametric representation of the nine inverted models derived from the Canny edge detection algorithm.



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

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

*not obviously better than  
PCA result*

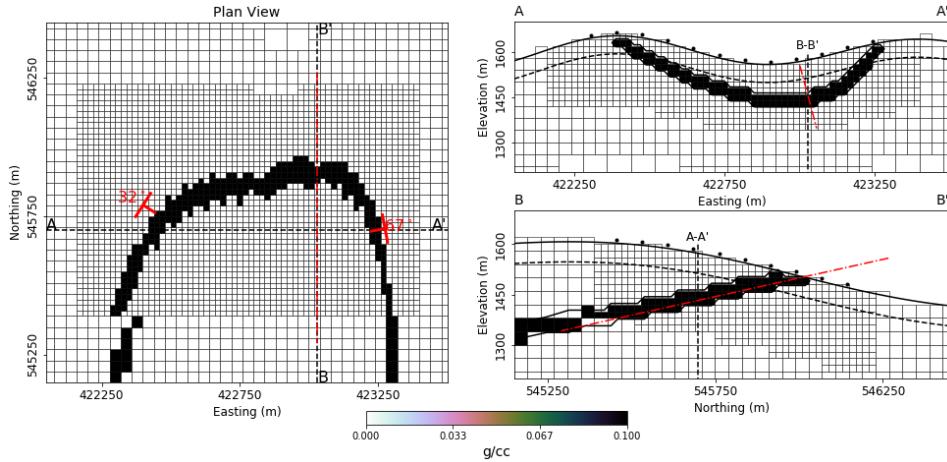


**Figure 6.8:** (a) The final SVMN inversion result and (b) normalized data residual map. Contour lines are shown in (a) for the true position and shape of the velocity anomalies.

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

## 6.2 Dip and strike estimation

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



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

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

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

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

# **Chapter 7**

## **Case Study - Kevitsa Ni-Cu-PGE deposit**

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

### **7.0.1 Geological setting**

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

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

As a first pass, I invert the data with the linear  $\ell_2$ -norm regularization. Figure 7.4(a) presents a section through the recovered density model at  $\approx 200$  m depth. I also extract a vertical section along the seismic profile E5 in Figure 7.4(b). I note that the bulk of the density anomaly is located inside the upper portion of the UPXO. As expected from smooth assumptions, the edges of the anomaly are poorly defined.

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

I then proceed with nine inversions over a range of mix-norm penalty functions for  $p_s, p_{y',z'} \in \{0, 1, 2\}$ ,  $p_{x'} = 2$ . Only the normal component of the geological contacts is fixed to the  $\ell_2$ -norm. Horizontal and vertical sections through the nine recovered density models are shown in Figure 7.6 and Figure 7.7 respectively. Residual data maps are shown in Figure 7.8. Correlated residuals indicate that some of the short ~~wave~~ wavelength information is under fitted.

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

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

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

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

To address issues posed by remanence, I proceed with the MVI-S algorithm. I perform a series of nine inversions with varying sparsity measures to assess the variability in the magnetization model. Starting from a common  $\ell_2$ -norm MVI-C model, I sequentially vary the combination of norms applied to the amplitude and its derivatives for ( $p_{\rho_s}, p_{\rho_{xyz}} \in [0, 2]$ ) and I apply the same rotation parameters as previously used for the gravity inversions. In this regard, the sparsity and rotation parameters tie the density and magnetization inversions together. While not jointly inverted, I expect some correlation between the two physical properties. Horizontal and vertical sections through the recovered nine magnetization models are shown in Figure 7.11 and 7.12 respectively. Residual data maps are shown in Figure 7.13 for all nine inversions.

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

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

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

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

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

## 7.4 Summary

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

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

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

# Chapter 8

## Conclusion

define

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

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

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

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

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

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

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

All the technology brought forward in this thesis were put to test in Chapter 7 on the Kevitsa Ni-Cu-PGE deposit. I inverted ground gravity and airborne mag-

**But you increased computational cost because you run more computations**

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

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

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