non-linear MVI-S formulation.

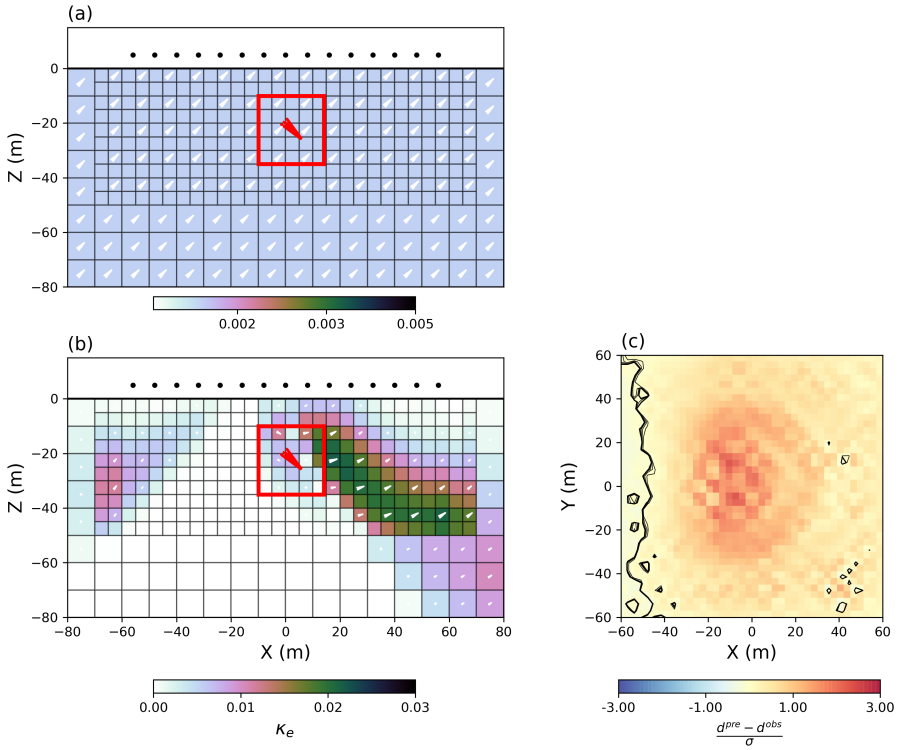

Figure 4.5: Vertical section through the (a) starting model and (b) recovered magnetization vector model in Spherical coordinates. The true position and magnetization orientation of the block are shown in red for reference. (c) Normalized data residuals show correlated signal. The inversion stopped after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

$$m_x + 2 * m_y = 1 \,, \tag{4.14}$$

which I can express in matrix form as

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

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

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

Figure 4.6(a) displays a contour map of the objective function along with its gradients. Following the same methodology as in (3.10), I find a solution such that the gradients of the objective function $\phi(m)$ vanish

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

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

As previously discussed for the gravity and magnetic problems, the simplest solution is often not satisfactory as it is strongly influenced by the physics of the experiment. From equation (3.7) in Chapter 3, I can introduce sensitivity based weights to counteract this bias towards large m_y value

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

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

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

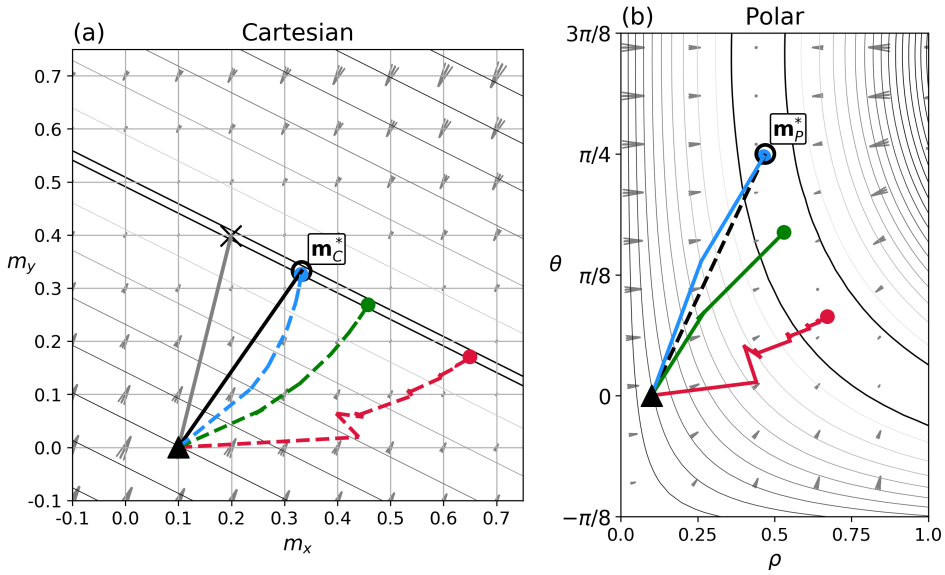

Figure 4.6: Contour map for two objective functions and their gradients (arrows) for a two-parameter inverse problem solved in Cartesian and polar coordinate systems. (a) The non-weighted (gray) problem yields a solution $\mathbf{m}_C = [0.2,\ 0.4]$ that reflects the size of the forward coefficients.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

and weighted gradients

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

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

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

tem under the transformation

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

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

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

The inverse problem is now non-linear with respect to the polar model so I solve it iteratively with the standard Gauss-Newton procedure described in equation (3.25). The partial derivatives of the forward mapping with respect to the polar coordinates are calculated by

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

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

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

The gradient of the objective function becomes

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

A trade-off parameter β_P is determined through the cooling schedule established in Chapter 3. The inversion is terminated once the data misfit and change in model norm falls below the tolerance η_{ϕ_d} and η_{ϕ_m} defined in equation (3.27) and (3.28) respectively.

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

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

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

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

assuming that I know the optimal trade-off parameter β_C^*. We already know that this gradient yields the target model $\mathbf{m}_C^{(0)}$. I can convert these gradients to polar coordinate by multiplying (4.26) with the matrix of partial derivatives $\mathbf{S}$ such that

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

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

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

I can simplify both sides of equation (4.29) by noting that $\mathbf{F}_C \mathbf{m}_C^{(0)} = \mathbb{F}[\mathbf{m}_P^{(0)}]$ which yields

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

I would like both sides to be roughly equal such that the gradient direction in polar space resemble the gradient direction calculated in the Cartesian space. First, I note from equation (4.29) that the transformation matrix $\mathbf{S}$ is missing from the right-hand side. This indicates that the gradient steps previously taken in polar

coordinates do not account for changes in sensitivity. I address this shortcoming by resorting to an iterative update of sensitivity weights

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

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

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

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

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

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

and scale the regularization as

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

My new scaled objective function becomes

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

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

4.2.1 Scaled MVI-S algorithm

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

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

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