An Efficient Alternating Algorithm for the L_p-Norm Cross-Gradient Joint Inversion of Gravity and Magnetic Data Using the 2-D Fast Fourier Transform

Saeed Vatankhah, Shuang Liu, Rosemary Anne Renaut, Xiangyun Hu, Jarom David Hogue, and Mostafa Gharloghi

***Abstract*— An efficient algorithm for the L_p-norm joint inversion of gravity and magnetic data using the cross-gradient constraint is presented. The presented framework incorporates stabilizers that use L_p-norms ($0 \leq p \leq 2$) of the model parameters, and/or the gradient of the model parameters. The formulation is developed from standard approaches for independent inversion of single data sets, and, thus, also facilitates the inclusion of necessary model and data weighting matrices, for example, depth weighting and hard constraint matrices. Using the block Toeplitz Toeplitz block structure of the underlying sensitivity matrices for gravity and magnetic models, when data are obtained on a uniform grid, the blocks for each layer of the depth are embedded in block circulant circulant block matrices. Then, all operations with these matrices are implemented efficiently using 2-D fast Fourier transforms, with a significant reduction in storage requirements. The nonlinear global objective function is minimized iteratively by imposing stationarity on the linear equation that results from applying linearization of the objective function about a starting model. To numerically solve the resulting linear system, at each iteration, the conjugate gradient algorithm is used. This is improved for large scale problems by the introduction of an algorithm in which updates for the magnetic and gravity parameter models are alternated at each iteration, further reducing total computational cost and storage requirements. Numerical results using a complicated 3-D synthetic model and real data sets obtained over the Galinge iron-ore deposit in the Qinghai province, north-west (NW) of China, demonstrate the efficiency of the presented algorithm.**

***Index Terms*— Block Toeplitz Toeplitz block (BTTB) structure, FFT, gravity, joint inversion, L_p-norm, magnetic.**

Manuscript received June 21, 2020; revised September 2, 2020; accepted October 13, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 41874122 and in part by the China Postdoctoral Science Foundation under Grant 2019M660191. The work of Rosemary Renaut was supported by the National Science Foundation (NSF) through the Approximate Singular Value Expansions and Solutions of Ill-Posed Problems under Grant DMS 1913136. *(Corresponding author: Shuang Liu.)*

Saeed Vatankhah, Shuang Liu, and Xiangyun Hu are with the Hubei Subsurface Multi-scale Imaging Key Laboratory, Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China (e-mail: lius@cug.edu.cn).

Rosemary Anne Renaut and Jarom David Hogue are with the School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-1804 USA.

Mostafa Gharloghi is with the Institute of Geophysics, University of Tehran, Tehran 14359-44411, Iran.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TGRS.2020.3033043.

Digital Object Identifier 10.1109/TGRS.2020.3033043

I. Introduction

POTENTIAL field surveys, gravity and magnetic, have been reported as effective strategies for delineating subsurface geological targets. They are applied in a wide range of studies including, for example, oil and gas exploration, mining applications, and mapping the basement topography [1], [2]. These surveys are relatively cheap, nondestructive passive remote sensing methods and can provide valuable information on the subsurface targets. Yet, they only require the measurement of variations in the Earth's natural fields that are caused by changes in the physical properties of the subsurface rocks. In the interpretation process, the acquired survey data can be used in an automatic inversion algorithm for the estimation of specific parameters of a subsurface target, for example, its geometry or physical properties. It is well known, however, that the potential field inversion problem is ill-posed. Thus, a stable and physically relevant solution is obtained by the application of suitable regularization strategies. An independent solution of the inverse problem for either gravity, or magnetic, data for the survey area will only provide information about the density or susceptibility, respectively, of the subsurface. On the other hand, a complementary solution of the inverse problem for both data sets can be used to reveal both density and magnetization variations present in a subsurface target. Thus, it is more appropriate to perform a simultaneous joint inversion that uses both data sets. Combined with regularization, this is an effective strategy for yielding a reliable subsurface geological model that simplifies the interpretation of the subsurface target(s). Thus, the development of efficient and stable joint inversion algorithms has received increased attention in the geophysical community.

Many different techniques have been developed for the simultaneous joint inversion of geophysical data sets. Generally, these techniques can be categorized into two main groups: 1) petrophysical and 2) structural approaches. Petrophysical techniques rely on a direct relationship between two or more physical properties of the subsurface target, for example, the assumption that the resistivity and the velocity are both functions of porosity and water saturation [3]. Although this strategy is attractive, it does depend on finding a reliable empirical relationship between physical properties. This is a

difficult task for general geological media because there is usually no simple or single relationship that approximates the whole range of effects [3]. Further discussion of the details of the application of petrophysical techniques for joint inversion is provided in the literature, including, for example, in [4]–[6]. Structural approaches use, instead, the model topology in order to enhance the structural similarity of reconstructed models [3], [7]–[10]. The main idea is that changes, at any point in the different models, should occur in the same or opposite spatial directions, or alternatively, changes will only occur in one of the models. Mathematically, this may be achieved by forcing the cross product of the gradient of the different model parameters to be zero everywhere [3], [8], [9]. Indeed, many successful results for simultaneous joint inversion with the inclusion of the cross-gradient constraint have been reported, [3], [8]–[16]. On the other hand, Zhdanov *et al.* [17] observed that the Gramian constraints can be used to enhance the correlation between different physical properties and/or their attributes. In this approach, the correlation is enhanced by minimizing the determinant of the Gram matrices of multimodel model parameters during the inversion process. The methodology has been used extensively in the joint inversion of different data sets, see, for example, [18]–[20]. In this study, we assume that the structure of the subsurface target(s) yields density and susceptibility model parameters over an approximately similar structure. Thus, our focus is on the use of the cross-gradient constraint within a general $L_{\mathbf{p}}$ formulation for efficient simultaneous joint inversion of gravity and magnetic data sets.

Different types of stabilizers have been adopted for the inversion of potential field data, dependent on the desired model features that are to be recovered. For example, it may be appropriate to reconstruct a model which represents only the large-scale features of the subsurface under the survey area without any arbitrary discontinuities. This is achieved with the maximum smoothness stabilizer which uses a L_2-norm[1] of the gradient of the model parameters, [21]–[24]. When it is anticipated that the subsurface structure exhibits discontinuities, stabilization can be achieved by imposing the L_1-norm, or L_0-norm, on the gradient of the model parameters [25]–[29]. If the assumed subsurface targets are localized and compact, it is more appropriate to apply the L_1 or L_0-norms directly on the model parameters [18], [26], [30]–[37]. In the potential field literature, stabilization by the application of the L_0-norm on the model parameters is usually referred to as the compactness constraint, whereas application of the L_1-norm, or L_0-norm, on the gradient is referred to as total variation (TV) and minimum gradient support (MGS), stabilization, respectively. A unifying approach for application of these constraints for single potential field inversion is presented in [38]. This approach also includes the modification of the stabilizers to account for additional model and data weighting matrices, such as required for imposition of depth weighting and hard constraint conditions, [22], [39]. Here, this unifying framework is developed for simultaneous joint inversion of gravity and magnetic data sets using the cross-gradient constraint.

The well-known and widely used formula of Fregoso and Gallardo [10] for the joint inversion of gravity and magnetic data with the cross-gradient constraint is based on the use of the generalized nonlinear least-squares framework developed by Tarantola and Valette [40]. Here, we adopt a deterministic viewpoint for parameter estimation and include deterministic constraints. The nonlinear objective function, that describes the inclusion of all stabilizing terms and the fit to data measurements, is minimized iteratively by imposing stationarity on the linear equation that results from applying linearization of the objective function about a starting model. To perform the inversion, the iteratively reweighted least-squares (IRLS) strategy is then used [41]. At each iteration, the conjugate gradient (CG) algorithm is applied to numerically solve the resulting linear system, and requires the efficient implementation of forward and transpose operations with the underlying sensitivity matrices.

Now, improved efficiency for the joint inversion of gravity and magnetic data is achieved in two ways. First, we take advantage of the underlying block Toeplitz Toeplitz block (BTTB) structure of the sensitivity matrices when the measurement data are obtained on a uniform grid, [42]–[44]. Then, all operations with the sensitivity matrices can be implemented using the 2-D fast Fourier transform (2-DFFT) and significant reductions in memory requirements are achieved. MATLAB software that both determines the relevant components of the matrices and implements the forward operations with the matrices using the 2-DFFT is available [45]. Moreover, the development and use of the 2-DFFT for the inversion of both gravity and magnetic data sets individually is carefully described in [46]. The application of these techniques within the joint inversion algorithm yields dramatic improvements in reduced computational costs and storage. Second, the inversion algorithm is improved, independent of the use of the structure of the matrices. A more efficient inversion algorithm in which the updates for the gravity and magnetic data are alternated, while maintaining coupling of the two data sets through the cross-gradient constraint, is introduced. This facilitates the inversion of large data sets in reasonable time and with reduced storage requirements.

This article is organized as follows. In Section II, the theoretical development of the algorithm is presented, along with the unifying framework that makes it possible to combine different types of stabilizers within the context of joint inversion. The novel alternating algorithm for the solution of the linear system of equations at each step of the CG algorithm is presented, as is the 2-DFFT approach. In Section III, the developed algorithm is validated on a large-scale synthetic example. Practical results are also presented for the inversion of gravity and magnetic data sets obtained over the Galinge iron-ore deposit in the Qinghai province, north-west (NW) of China. Section V is dedicated to a discussion of conclusions and future topics for research.

[1]The $L_{\mathbf{p}}$-norm of a vector $\mathbf{x} \in \mathcal{R}^n$ is defined as $\|\mathbf{x}\|_{\mathbf{p}} = (\sum_i^n |x_i|^{\mathbf{p}})^{1/\mathbf{p}}$, in which $1 \leq \mathbf{p} \leq \infty$. The L_0-norm of vector $\mathbf{x}$, $\|\mathbf{x}\|_0$, counts the number of nonzero entries in $\mathbf{x}$.

II. Joint Inversion Methodology

To formulate the problem, we use the well-known strategy for linear inversion of potential field data in which the subsurface is divided into a set of rectangular prisms with fixed size but unknown physical properties [22], [39]. Here, it is assumed that there is no remanent magnetization, and that self-demagnetization effects are also negligible. For ease of exposition, we first introduce some basic notation for the stacking of vectors (matrices) and generation of block-diagonal matrices. We use block_stack$(\cdot,\cdot)$ to indicate the stacking of vectors (or matrices) with the same number of columns in one vector (or matrix). Furthermore, block_diag(A, B) indicates a block diagonal matrix of size $(m_A+m_B)\times(n_A+n_B)$ when A and B are of sizes $(m_A\times n_A)$ and $(m_B\times n_B)$, respectively. Both definitions extend immediately for more than two entities.

We suppose that m measurements are taken for two sets of potential field data.[2] These are the vertical components of the gravity and total magnetic fields, and they are stacked in vectors $\mathbf{d}_1^{\text{obs}}$, and $\mathbf{d}_2^{\text{obs}}$, each of length m, respectively. The unknown physical parameters, the density, and the susceptibility, of n prisms, are also stacked in vectors $\mathbf{m}_1$ and $\mathbf{m}_2$, respectively. The data vectors and model parameters are then stacked consistently in vectors $\mathbf{d}^{\text{obs}} = \text{block\_stack}(\mathbf{d}_1^{\text{obs}}, \mathbf{d}_2^{\text{obs}}) \in \mathcal{R}^{2m}$, and $\mathbf{m} = \text{block\_stack}(\mathbf{m}_1, \mathbf{m}_2) \in \mathcal{R}^{2n}$. The measurements are connected to the model parameters via $G\mathbf{m} = \mathbf{d}^{\text{obs}}$ where $G = \text{block\_diag}(G_1, G_2) \in \mathcal{R}^{2m\times 2n}$, and G_1 and G_2 are the linear forward modeling operators for gravity and magnetic kernels, respectively. There are different alternative formulas which can be used to compute the entries of matrices G_1 and G_2. Here, we use the formulas developed by Haáz [47] for computing the vertical gravitational component, and Rao and Babu's [48], for the total magnetic field anomaly, of a right rectangular prism, respectively. Moreover, we assume that the measurements are taken at a uniform grid, on or near the surface, so that the resulting sensitivity matrices have BTTB structure for each depth layer of the volume, as described in [46]. Then all operations with these matrices use the 2-DFFT and require only the storage of the necessary transform data for each depth layer of the volume domain. This improves efficiency in reducing computational cost and memory required to implement operations with the relevant sensitivity matrices. Details, not replicated here, on how to use the 2-DFFT for operations using G_1 and G_2, and hence the block diagonal matrix G, are explained carefully in [46], and a software MATLAB package for generating the transform components that are used in place of G_1 and G_2 directly is available at [45].

The goal of the inversion is to find geologically plausible models $\mathbf{m}_1$ and $\mathbf{m}_2$ that predict $\mathbf{d}_1^{\text{obs}}$ and $\mathbf{d}_2^{\text{obs}}$, respectively, via a simultaneous joint algorithm that also facilitates the incorporation of relevant weighting matrices in the algorithm. We formulate the joint inversion for the determination of the model parameters $\mathbf{m}_1$ and $\mathbf{m}_2$ as the minimization of the global objective function, in which parameters α and λ are relative weighting parameters for the respective terms

$$P^{(\alpha,\lambda)}(\mathbf{m}) = \|W_{\mathbf{d}}(\mathbf{d}^{\text{obs}} - G\mathbf{m})\|_2^2 + \alpha^2\|WD(\mathbf{m}-\mathbf{m}^{\text{apr}})\|_2^2 + \lambda^2\ \|\mathbf{t}(\mathbf{m})\|_2^2. \quad (1)$$

The data misfit term, $\|W_{\mathbf{d}}(\mathbf{d}^{\text{obs}} - G\mathbf{m})\|_2^2$, measures how well the calculated data reproduce the observed data. Diagonal matrix $W_{\mathbf{d}} = \text{block\_diag}(W_{\mathbf{d}_1}, W_{\mathbf{d}_2}) \in \mathcal{R}^{2m\times 2m}$, where $W_{\mathbf{d}_1}$ and $W_{\mathbf{d}_2}$ are diagonal weighting matrices for the gravity and magnetic data, respectively. Here, we suppose that these diagonal elements are the inverses of the standard deviations of the independent, but potentially colored, noise in the data. The stabilizer, $\|WD(\mathbf{m}-\mathbf{m}^{\text{apr}})\|_2^2$, controls the growth of the solution with respect to the weighted norm and is especially significant as it determines the structural qualities of the desired solution. Here, this stabilizer is presented through a general L_2-norm formulation, but we will discuss how different choices of W and D lead to different $L_{\mathbf{p}}$-norm stabilizations. Furthermore, in (1) the vector $\mathbf{m}^{\text{apr}} = \text{block\_stack}(\mathbf{m}_1^{\text{apr}}, \mathbf{m}_2^{\text{apr}}) \in \mathcal{R}^{2n}$ is an initial starting model that may be known from previous investigations. It is also possible to set $\mathbf{m}^{\text{apr}} = \mathbf{0}$. The link between the gravity and magnetic models in this inversion algorithm is the cross-gradient function $\mathbf{t}(\mathbf{m}) \in \mathcal{R}^{3n}$. For this study, we assume that the model structures for $\mathbf{m}_1$ and $\mathbf{m}_2$ are approximately the same, and thus that it is important to measure the structural similarity using the cross-gradient constraint which will be approximately zero for models with similar structures. We, therefore, use

$$\mathbf{t}(\mathbf{m}) = \nabla\mathbf{m}_1\ (x,y,z) \times \nabla\mathbf{m}_2\ (x,y,z) \quad (2)$$

where ∇ indicates the gradient operator [3] and structural similarity is achieved when $\mathbf{t}(\mathbf{m}) = \mathbf{0}$, see Appendix for details. As noted already in Section I, this corresponds to the case in which the gradient vectors are in the same or opposite direction, or, alternatively, one of them is zero [3], [8]–[12], [14]. From a geological viewpoint, this means that if a boundary exists then it must be sensed by both methods in a common orientation regardless of how the amplitude of the physical property changes [3]. This means that information that is contained in one model is relevant to the other model and vice versa. Therefore, structures determined by one model can assist with the identification of structures in the other model, and, as a consequence, the structures of the two models can correct each other throughout the joint inversion process [3], [12]. On the other hand, while it is assumed that both models have similar structures at similar locations, it is also possible for one model to have a structure in a location where the other model has none [12]. Further details about characteristic properties of the cross-gradient constraint are provided in [3], [9], and [10].

The stabilizing term has a very significant impact on the solution that is obtained by minimizing (1). Depending on the type of the desired model to be recovered, there are many choices that can be considered for this stabilization, and that have been extensively adopted by the geophysical community. A stabilizer that is based on using the squared L_2-norm of the gradients of the model parameters produces smooth models with blurred boundaries. Then, any temptation to over-interpret

[2]We could assume different numbers of measurements for each field, m_1 and m_2 but for simplicity of the discussion we immediately assume $m_1 = m_2 = m$.

the data is reduced and features of the model that are not predicted by the data are avoided [21], [22]. Alternatively, if it is anticipated that a compact model with sharp interfaces would be an appropriate model, then a sparsity stabilizer based on either a L_1-norm or a L_0-norm, of the model parameters should be used. The L_1-norm stabilizer induces sparsity most efficiently among convex stabilizers, $\mathbf{p} \geq 1$, while the L_0-norm does not meet the mathematical requirement to be regarded as a norm. A detailed discussion of the different norms and sparsity regularization is provided in [49] and [50]. Here, we show how it is possible to use the given weighted L_2-norm regularizer in (1) to approximate different $L_\mathbf{p}$-norm stabilizers, $0 \leq \mathbf{p} \leq 2$, in a joint inversion algorithm.

Suppose D is the identity matrix, $D = I_{2n}$, and W is selected as $W = \text{block\_diag}(W_1, W_2) \in \mathcal{R}^{2n\times 2n}$, in which

$$W_i = (W_{\text{depth}})_i (W_h)_i (W_{L_\mathbf{p}})_i \in \mathcal{R}^{n\times n}, \quad i = 1, 2. \tag{3}$$

Here, the diagonal weighting matrix $(W_{L_\mathbf{p}})_i \in \mathcal{R}^{n\times n}$ is defined, assuming entries are calculated elementwise, by

$$(W_{L_\mathbf{p}})_i = \text{diag}\big(1/((\mathbf{m}_i - \mathbf{m}_i^{\text{apr}})^2 + \epsilon^2)^{(2-\mathbf{p})/4}\big), \quad i = 1, 2. \tag{4}$$

Taking $\mathbf{p} = 0$ or $\mathbf{p} = 1$, the stabilizing term in (1) is transformed to a L_0-norm or L_1-norm of the model parameters, respectively. Then, minimizing the objective function (1) leads to solutions which are compact. The choice $\mathbf{p} = 2$ provides the L_2-norm solution of the model parameters. The parameter ϵ is a small positive number, $0 < \epsilon \ll 1$, which is added to avoid the possibility of division by zero, and has an important effect on the solution. When ϵ is very small the solutions are sparse, while for large values the solutions are smooth. Further discussion on the impact of ϵ is given, for example, in [30], [25], [32], [33], [51], and [29]. On the other hand, it is also possible to choose D to provide an approximation to the gradient of the model parameters. Suppose, for example, that $D_1 = \text{block\_stack}\big(D_x, D_y, D_z\big) \in \mathcal{R}^{3n\times n}$, where D_x, D_y, and D_z are square and provide discrete approximations for derivative operators in x-, y-, and z-directions. Then, defining $\mathbf{0}_{3n\times n}$ to be the zero matrix of size $3n\times n$, and setting $D_2 = D_1$, we can use the matrix

$$D = \begin{pmatrix} D_1 & \mathbf{0}_{3n\times n} \\ \mathbf{0}_{3n\times n} & D_2 \end{pmatrix} \in \mathcal{R}^{6n\times 2n} \tag{5}$$

so that $D\mathbf{m}$ yields the approximate gradient of $\mathbf{m}$. More details about the structure of the matrices D_x, D_y, and D_z can be found in [52]–[54], and are also used in Appendix . Then, with this definition for D, and again using element-wise calculations, $(W_{L_\mathbf{p}})_i$ in (4) is replaced by

$$\begin{aligned}(W_{L_\mathbf{p}})_i &= \text{diag}\Bigg(1/\big((D_x(\mathbf{m} - \mathbf{m}_i^{\text{apr}}))^2 + (D_y(\mathbf{m} - \mathbf{m}_i^{\text{apr}}))^2 \\ &\qquad + (D_z(\mathbf{m} - \mathbf{m}_i^{\text{apr}}))^2 + \epsilon^2\big)^{(2-\mathbf{p})/4}\Bigg), \quad i = 1, 2.\end{aligned} \tag{6}$$

But now, for the multiplication in the stabilizing term to be dimensionally consistent, W is replaced by

$$W = \text{block\_diag}(W_1, W_1, W_1, W_2, W_2, W_2) \in \mathcal{R}^{6n\times 6n}. \tag{7}$$

In the definition of $(W_{L_\mathbf{p}})_i$, given in (6), picking $\mathbf{p} = 2$ produces a solution with minimum structure yielding a smooth model with blurred boundaries [21]–[23]. If we anticipate that there are true discontinuous jumps, then, it is possible to take $\mathbf{p} = 1$ or $\mathbf{p} = 0$ for a TV or MGS stabilization, respectively. In summary, all aforementioned definitions indicate how, dependent on the choices of $\mathbf{p}$ and D in W, it is possible to use the objective function (1) with a desired stabilizer. Well-known stabilizers, including TV, MGS, minimum structure, compactness, and L_1-norm can all be incorporated in a joint inversion methodology. Moreover, this unifying framework allows the use of additional $L_\mathbf{p}$-norms stabilizers, $0 \leq \mathbf{p} \leq 2$, which are not common in potential field inversion, simply by changing the choice of $\mathbf{p}$.

In (3), the diagonal depth weighting matrix, $(W_{\text{depth}})_i = \text{diag}(1/(z_j + z_0)^{\beta_i})$ is used to counteract the rapid decay of the kernel with depth [22], [23]. Here, z_j is the mean depth of prism j, z_0 depends both upon prism size and the height of the observed data. With the application of appropriate depth weighting, as determined by parameter β_i, all prisms participate with an approximately equal probability in the inversion process. The diagonal hard constraint matrix $(W_h)_i$, is generally an identity matrix. If geological information, or prior investigations in the survey area, can provide the values of the model parameter for some prisms, then, the information is included in $\mathbf{m}_i^{\text{apr}}$, and the corresponding diagonal entries of $(W_h)_i$ are set to a large value [34], [39], [55]. These known parameters are kept fixed during the iterative minimization of $P^{(\alpha,\lambda)}(\mathbf{m})$. Equivalently, the inversion algorithm searches only for unknown model parameters. As an important aside, note that all matrices D_x, D_y, D_z, W_{depth}, W_h, and $W_{L_\mathbf{p}}$, are sparse and can therefore be saved using a MATLAB `sparse` format, with very limited demand on the memory.

Now, both $W_{L_\mathbf{p}}$ and $\mathbf{t}(\mathbf{m})$ depend on the model parameters $\mathbf{m}$. Thus, the objective function $P^{(\alpha,\lambda)}(\mathbf{m})$ is nonlinear with respect to $\mathbf{m}$. We use a simple iterative strategy to convert $P^{(\alpha,\lambda)}(\mathbf{m})$ to a linear form, in which to linearize the cross-gradient constraint, the first-order Taylor expansion is used [3], [8]–[11]. First, suppose that the superscript ℓ applied to any variable indicates the value of that variable at iteration ℓ, so that $\mathbf{m}^{(\ell)}$ is the estimate of the model parameters at iteration ℓ. Then, we suppose that $\mathbf{m}^{(1)} = \mathbf{m}^{\text{apr}}$ and rewrite the objective function as

$$\begin{aligned}&P^{(\alpha,\lambda)}(\mathbf{m}) \\ &= \|W_\mathbf{d}(\mathbf{d}^{\text{obs}} - G\mathbf{m})\|_2^2 \\ &\quad + \alpha^2 \, \|W^{(\ell-1)} D(\mathbf{m} - \mathbf{m}^{(\ell-1)})\|_2^2 \\ &\quad + \lambda^2 \, \|\mathbf{t}^{(\ell-1)} + \mathbf{B}^{(\ell-1)}(\mathbf{m} - \mathbf{m}^{(\ell-1)})\|_2^2, \quad \ell = 2, 3, \ldots\end{aligned} \tag{8}$$

Note that $W^{(\ell-1)}$ indicates W estimated at iteration $\ell - 1$ through the nonlinear definition for $W_{L_\mathbf{p}}$ as given in (4) or (6). Here, $\mathbf{t}^{(\ell-1)}$ and $\mathbf{B}^{(\ell-1)} = (\nabla_\mathbf{m}\mathbf{t}^{(\ell-1)})$ are the cross-gradient and the Jacobian matrix of the discrete approximation for the cross-gradient function, respectively, evaluated at $\mathbf{m}^{(\ell-1)}$, consistent with the linear Taylor expansion for $\mathbf{t}$ around $\mathbf{m}^{(\ell-1)}$. The formulae used are given in Appendix . Taking $\nabla_\mathbf{m} P^{(\alpha,\lambda)}(\mathbf{m}) = \mathbf{0}$ defines the update $\mathbf{m}^{(\ell)}$ as the

solution of

$$-G^T W_{\mathbf{d}}^T W_{\mathbf{d}}(\mathbf{d}^{\text{obs}} - G\mathbf{m}) + \alpha^2 D^T (W^{(\ell-1)})^T W^{(\ell-1)} D(\mathbf{m} - \mathbf{m}^{(\ell-1)}) + \lambda^2 \left((\mathbf{B}^{(\ell-1)})^T \left\{ \mathbf{t}^{(\ell-1)} + \mathbf{B}^{(\ell-1)}(\mathbf{m} - \mathbf{m}^{(\ell-1)}) \right\} \right) = 0. \tag{9}$$

Equivalently, $\mathbf{m}^{(\ell)}$ solves

$$E^{(\ell)}\mathbf{m}^{(\ell)} = \mathbf{f}^{(\ell)} \tag{10}$$

where

$$E^{(\ell)} = \left(G^T W_{\mathbf{d}}^T W_{\mathbf{d}} G + \alpha^2 D^T (W^{(\ell-1)})^T W^{(\ell-1)} D + \lambda^2 \ (\mathbf{B}^{(\ell-1)})^T \mathbf{B}^{(\ell-1)} \right) \tag{11}$$

and

$$\mathbf{f}^{(\ell)} = \left(G^T W_{\mathbf{d}}^T W_{\mathbf{d}} \mathbf{d}^{\text{obs}} + \alpha^2 D^T (W^{(\ell-1)})^T W^{(\ell-1)} D\mathbf{m}^{(\ell-1)} + \lambda^2 (\mathbf{B}^{(\ell-1)})^T \left(\mathbf{B}^{(\ell-1)}\mathbf{m}^{(\ell-1)} - \mathbf{t}^{(\ell-1)} \right) \right). \tag{12}$$

Numerically the CG algorithm can be used to find $\mathbf{m}^{(\ell)}$. Here, we use the MATLAB function pcg. We should note that at each iteration of the algorithm lower and upper bounds on density and susceptibility are imposed. During the inversion process if an estimated physical property falls outside the specified bounds, it will be returned back to the nearest bound. Furthermore, to test the convergence of the solution at each iteration ℓ, we calculate a χ^2 measure for the respective data fit term at each iteration

$$(\chi_i^2)^{(\ell)} = \| W_{\mathbf{d}_i} (\mathbf{d}_i^{\text{obs}} - G_i \mathbf{m}_i^{(\ell)}) \|_2^2, \quad i = 1, 2. \tag{13}$$

The iteration will terminate at convergence only when $(\chi_i^2)^{(\ell)} \leq m + \sqrt{2m}$, for both $i = 1$, and $i = 2$. Otherwise, the iteration is allowed to proceed to a maximum number of iterations MAXIT. The steps of the joint inversion algorithm are summarized in Algorithm 1. Note that all operations using the block diagonal matrix G are implemented efficiently by using the 2-DFFT for each column block of the sensitivity matrices G_i, as described in [46].

In the objective function, the parameters α and λ are the important regularization parameters which give relative weights to the stabilizer and the cross-gradient term, respectively. To be more precise, we define α as block_diag$(\alpha_1 I_n, \alpha_2 I_n) \in \mathcal{R}^{2n \times 2n}$, where α_1 and α_2 are the relative weights for the gravity and magnetic terms, respectively. We should note that, although α is a diagonal matrix and can be used inside the stabilizer, we prefer to put α outside in order for the formulation to be consistent with the conventional Tikhonov objective functional. Weighting parameters α_1, α_2, and λ have an important effect on the estimated solution. Thus, they need to be determined carefully. But the application of an automatic parameter-choice method for determining α_1, α_2, and λ is difficult, or potentially impossible, and is outside the scope of this current study. Therefore, we adopt a simple but practical strategy for determining suitable values of these parameters. Previous investigations have demonstrated that it is efficient if the inversion starts with a large regularization parameter [34], [56]. We follow that strategy here and start the inversion with a large α_1 and α_2. In subsequent iterations, the parameters are reduced slowly dependent on parameters γ_1 and γ_2, respectively, using $\alpha_1^{(\ell)} = \alpha_1^{(\ell-1)}\gamma_1$ and $\alpha_2^{(\ell)} = \alpha_2^{(\ell-1)}\gamma_2$, where γ_1 and γ_2 are small numbers, $0 \ll \gamma_1, \gamma_2 < 1$. The process continues until the predicted data of one of the reconstructed models satisfies the observed data at the noise level. For that data set, the relevant parameter is then kept fixed during the following iterations. The parameter λ is held fixed in the implementation, although it is quite feasible that it is also iteration dependent. The amount of structural similarity obtained through the joint inversion algorithm can be adjusted using different choices of λ.

Algorithm 1 Generalized $L_{\mathbf{p}}$-Norm Joint Inversion of Gravity and Magnetic Data

Require: $\mathbf{d}^{\text{obs}}$, $\mathbf{m}^{\text{apr}}$, G, $W_{\mathbf{d}}$, $(W_h)_1$, $(W_h)_2$, D_x, D_y, D_z, ϵ, $\rho_{\min}$, $\rho_{\max}$, $\kappa_{\min}$, $\kappa_{\max}$, MAXIT, β_1, β_2, $\alpha_1^{(1)}$, $\alpha_2^{(1)}$, γ_1, γ_2 and λ.

1: Calculate $(W_{\text{depth}})_1$ and $(W_{\text{depth}})_2$ as determined by β_1 and β_2, respectively.
2: Set $W_1 = (W_{\text{depth}})_1 (W_h)_1$ and $W_2 = (W_{\text{depth}})_2 (W_h)_2$.
3: If D is the identity matrix, form $W = \text{block\_diag}(W_1, W_2)$. Otherwise set $W = \text{block\_diag}(W_1, W_1, W_1, W_2, W_2, W_2)$.
4: Initialize $\mathbf{m}^{(1)} = \mathbf{m}^{\text{apr}}$, $(W_{\text{L}_\mathbf{p}})_1^{(1)} = I$, $(W_{\text{L}_\mathbf{p}})_2^{(1)} = I$, and $\ell = 1$.
5: **while** Not converged, noise level not satisfied, and $\ell <$ MAXIT **do**
6: $\ell = \ell + 1$.
7: Compute $\mathbf{t}^{(\ell-1)}$ and $\mathbf{B}^{(\ell-1)}$, as given in Appendix .
8: Use CG to solve $E^{(\ell)}\mathbf{m}^{(\ell)} = \mathbf{f}^{(\ell)}$, defined by (11) and (12).
9: Impose constraints on $\mathbf{m}^{(\ell)}$ to force $\rho_{\min} \leq \mathbf{m}_1^{(\ell)} \leq \rho_{\max}$ and $\kappa_{\min} \leq \mathbf{m}_2^{(\ell)} \leq \kappa_{\max}$.
10: Test convergence criteria, (13), for χ_1^2 and χ_2^2. Exit loop if both satisfied.
11: Set $\alpha_1^{(\ell)} = \alpha_1^{(\ell-1)}\gamma_1$ and $\alpha_2^{(\ell)} = \alpha_2^{(\ell-1)}\gamma_2$. Update $\alpha^{(\ell)}$.
12: Determine $(W_{\text{L}_\mathbf{p}})_1^{(\ell)}$ and $(W_{\text{L}_\mathbf{p}})_2^{(\ell)}$, dependent on D and $\mathbf{p}$, (4) or (6).
13: Set $W^{(\ell)} = \text{block\_diag}(W_1^{(\ell)}, W_2^{(\ell)})$ when D is the identity. Otherwise set $W^{(\ell)} = \text{block\_diag}(W_1^{(\ell)}, W_1^{(\ell)}, W_1^{(\ell)}, W_2^{(\ell)}, W_2^{(\ell)}, W_2^{(\ell)})$
14: **end while**

Ensure: Solution $\rho = \mathbf{m}_1^{(\ell)}$, $\kappa = \mathbf{m}_2^{(\ell)}$, IT $= \ell$.

A. Alternating Algorithm for Joint Inversion

Algorithm 1 can still be expensive in terms of both storage and computational overhead, for large-scale problems, even when the multiplications with matrices G_i take advantage of the 2-DFFT. We now present a modification of Algorithm 1 in which the updates for $\mathbf{m}_1$ and $\mathbf{m}_2$ alternate. In this case, we use an approximate decoupling of the matrix $E^{(\ell)}$ for the

two parameter vectors. First, note that due to the block structure of G and $W_{\mathbf{d}}$, $W_{\mathbf{d}}G = \text{block\_diag}(W_{\mathbf{d}_1}G_1, W_{\mathbf{d}_2}G_2)$ and $G^T W_{\mathbf{d}}^T W_{\mathbf{d}} G = \text{block\_diag}(G_1^T W_{\mathbf{d}_1}^T W_{\mathbf{d}_1} G_1, G_2^T W_{\mathbf{d}_2}^T W_{\mathbf{d}_2} G_2)$. Also, $W^{(\ell)}D$ decouples as a block diagonal matrix $\text{block\_diag}(W_1^{(\ell)}D_1, W_2^{(\ell)}D_2)$. Based on the form for $\mathbf{B}$ given in (26) of Appendix

$$\mathbf{B}^T\mathbf{B} = \begin{pmatrix} \mathbf{B}_1^T\mathbf{B}_1 & \mathbf{B}_1^T\mathbf{B}_2 \\ \mathbf{B}_2^T\mathbf{B}_1 & \mathbf{B}_2^T\mathbf{B}_2 \end{pmatrix}. \tag{14}$$

This does not decouple into block diagonal form. On the other hand, we can approximate $\mathbf{B}^T\mathbf{B}$ using the block diagonal $\text{block\_diag}(\mathbf{B}_1^T\mathbf{B}_1, \mathbf{B}_2^T\mathbf{B}_2)$ which implies that the cross-terms due to $\mathbf{B}_1^T\mathbf{B}_2 = \mathbf{B}_2^T\mathbf{B}_1$ are ignored. Then we obtain the independent systems of equations to solve for $\mathbf{m}_1$ and $\mathbf{m}_2$, which are coupled through the updates for $\mathbf{B}$ and $\mathbf{t}$. In particular, we obtain

$$E_i^{(\ell)} = \Big(G_i^T W_{\mathbf{d}_i}^T W_{\mathbf{d}_i} G_i + \alpha^2 (W_i^{\ell-1} D_i)^T W_i^{\ell-1} D_i + \lambda^2 \ (\mathbf{B}_i^{(\ell-1)})^T \mathbf{B}_i^{(\ell-1)} \Big) \tag{15}$$

$$\mathbf{f}_i^{(\ell)} = \Big(G_i^T W_{\mathbf{d}_i}^T W_{\mathbf{d}_i} \mathbf{d}_i^{\text{obs}} + \alpha^2 (W_i^{(\ell-1)} D_i)^T W_i^{(\ell-1)} D_i \mathbf{m}_i^{(\ell-1)} + \lambda^2 (\mathbf{B}_i^{(\ell-1)})^T \left(\mathbf{B}_i^{(\ell-1)} \mathbf{m}_i^{(\ell-1)} - \mathbf{t}^{(\ell-1)} \right) \Big) \tag{16}$$

where $\mathbf{B}$ and $\mathbf{t}$ are updated between the solutions of the systems

$$E_i^{(\ell)} \mathbf{m}_i = \mathbf{f}_i^{(\ell)}, \quad i = 1, 2. \tag{17}$$

The essential steps replacing steps 7–9 in Algorithm 1 are provided in Algorithm 2.

Algorithm 2 Essential Steps of Alternating Generalized $L_{\mathbf{p}}$-Norm Joint Inversion of Gravity and Magnetic Data

Require: Proceed through steps 1-7 of Algorithm 1, replacing update for $\mathbf{B}^{(\ell-l)}$ by update of $\mathbf{B}_1^{(\ell-l)}$.

1: Solve $E_1^{(\ell)}\mathbf{m}_1^{(\ell)} = \mathbf{f}_1^{(\ell)}$, defined by (15) and (16).
2: Impose constraints $\rho_{\min} \le \mathbf{m}_1^{(\ell)} \le \rho_{\max}$.
3: Update $\mathbf{t}^{(\ell-1)}$ and $\mathbf{B}_2^{(\ell-1)}$.
4: Solve $E_2^{(\ell)}\mathbf{m}_2^{(\ell)} = \mathbf{f}_2^{(\ell)}$, defined by (15) and (16).
5: Impose constraints $\kappa_{\min} \le \mathbf{m}_2^{(\ell)} \le \kappa_{\max}$.
6: Move to step 10 of Algorithm 1.

We note that the ignored terms in this algorithm for updating $\mathbf{m}_1$ are $\lambda^2\mathbf{B}_1^T\mathbf{B}_2\mathbf{m}_2^{(\ell)}$ on the left and $\lambda^2\mathbf{B}_1^T\mathbf{B}_2\mathbf{m}_2^{(\ell-1)}$ on the right. Equivalently $\lambda^2\mathbf{B}_1^T\mathbf{B}_2(\mathbf{m}_2^{(\ell)} - \mathbf{m}_2^{(\ell-1)})$ is ignored in the update for $\mathbf{m}_1$, and, likewise, $\lambda^2\mathbf{B}_2^T\mathbf{B}_1(\mathbf{m}_1^{(\ell)} - \mathbf{m}_1^{(\ell-1)})$ is ignored in the update for $\mathbf{m}_2$. Ignoring these terms amounts to requiring that $\|\mathbf{B}_1^T\mathbf{B}_2(\mathbf{m}_2^{(\ell)} - \mathbf{m}_2^{(\ell-1)})\| \to 0$, $(\|\mathbf{B}_2^T\mathbf{B}_1(\mathbf{m}_1^{(\ell)} - \mathbf{m}_1^{(\ell-1)})\| \to 0)$ with increasing ℓ, which is not guaranteed, but is indeed desired for the convergence of the algorithm, for which not only $\|\mathbf{m}^{(\ell)} - \mathbf{m}^{(\ell-1)}\| \to 0$ but also $\|\mathbf{B}_1^T\mathbf{B}_2\| = \|\mathbf{B}_2^T\mathbf{B}_1\| \to 0$. The alternating Algorithm 2 amounts to applying a Quasi-Newton method for the minimization of (8).

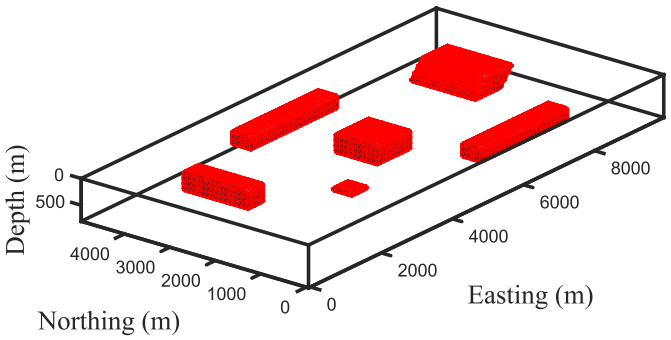

Fig. 1. Synthetic model consisting of six targets with different shapes, dimensions, and depths. The density contrast and the susceptibility of the targets are selected as $\rho = 0.6$ gr cm^{-3} and $\kappa = 0.06$ (SI unit), respectively, embedded in a homogeneous nonsusceptible background.

III. Simulations

The validity and effectiveness of the presented joint inversion algorithm is evaluated on a complicated model that consists of six subsurface targets with different shapes, dimensions, and depths. Fig. 1 presents a 3-D iso-surface view of the model. Furthermore, three plane-sections of the susceptibility distribution are also shown in the Fig. 2. The density contrast and the susceptibility of the targets are selected as $\rho = 0.6$ gr cm^{-3} and $\kappa = 0.06$ (SI unit), respectively, embedded in a homogeneous nonsusceptible background. To generate the total field anomaly, the intensity of the geomagnetic field, the inclination, and the declination are selected as 50 000 nT, 45° and 45°, respectively. The data on the surface are generated on a grid with $100 \times 50 = 5000$ points and grid spacing 100 m. Gaussian noise with zero mean and standard deviation $(\tau_1 \ |(\mathbf{d}_i^{\text{exact}})_j| + \tau_2 \ \max(|\mathbf{d}_i^{\text{exact}}|))$, $j = 1 \ldots n$, is added to each exact true measurement j, $(\mathbf{d}_i^{\text{exact}})_j$. The parameters pairs (τ_1, τ_2) are selected as $(0.01, 0.012)$ and $(0.01, 0.01)$ for gravity and magnetic data, respectively. These standard deviations are selected such that the signal to noise ratios (SNRs), as given by

$$\text{SNR}_i = 20 \log_{10}\left(\frac{\|\mathbf{d}_i^{\text{exact}}\|_2}{\|\mathbf{d}_i^{\text{obs}} - \mathbf{d}_i^{\text{exact}}\|_2} \right), \quad i = 1, 2 \tag{18}$$

are 25.47 and 25.13, respectively. The noise-contaminated gravity and magnetic data are illustrated in Fig. 3.

To perform the inversion, the subsurface volume is discretized into 40 000 prisms of sizes 100 m in each dimension. In all simulations, we use $\beta_1 = 0.8$ and $\beta_2 = 1.4$ in $(W_{\text{depth}})_1$ and $(W_{\text{depth}})_2$, respectively. The maximum number of iterations of the algorithm is selected as MAXIT = 100. For all inversions, the bound constraints $0 = \rho_{\min} \le \mathbf{m}_1 \le \rho_{\max} = 0.6$, gr cm^{-3}, and $0 = \kappa_{\min} \le \mathbf{m}_2 \le \kappa_{\max} = 0.06$, SI unit, are imposed. The initial regularization parameters are selected as $\alpha_1^{(1)} = 20\,000$ and $\alpha_2^{(1)} = 50\,000$. In Vatankhah *et al.* [57] we demonstrated that the gravity and magnetic sensitivity matrices, G_1 and G_2, have different spectral properties and, therefore, the regularization parameter should be much larger for the inversion of magnetic data as compared to that used for the inversion of gravity data. Thus, it is appropriate to control the speed of convergence for each model with parameters

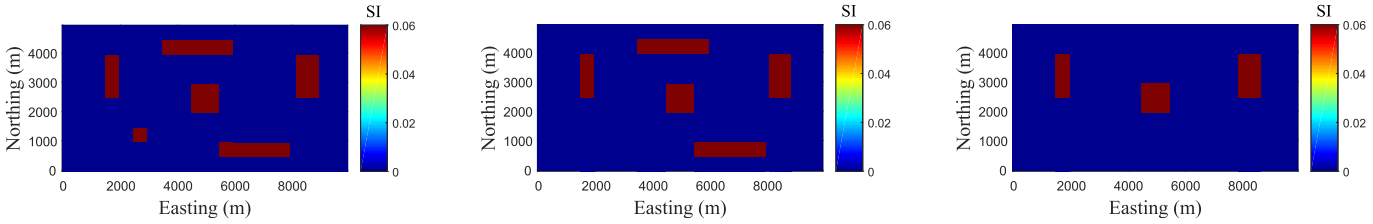

Fig. 2. Three plane-sections of the model shown in Fig. 1. Here, the susceptibility distribution of the model is presented at: (a) depth 100 m; (b) depth 200 m; and (c) depth 400 m.

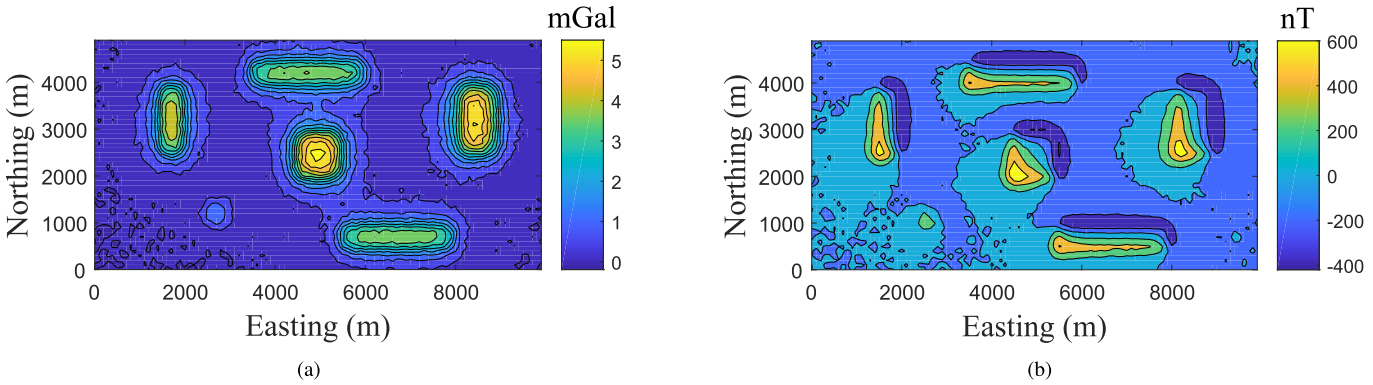

Fig. 3. Noise contaminated anomaly produced by the model shown in Fig. 1. (a) Vertical component of gravity. (b) Total magnetic field. The SNR for gravity and magnetic data, respectively, are 25.47 and 25.13.

$\gamma_1 = 0.9$ and $\gamma_2 = 0.95$, that are different. We present the results for the compact inversion using the L_1-norm of the model parameters imposed via $D = I$, $\mathbf{p} = 1$ and focusing parameter $\epsilon^2 = 1e^{-9}$ in (4). Further results for the application of a joint inversion algorithm on different models and with other stabilizers are presented as supplementary data in [58]. The alternating joint inversion algorithm is implemented using MATLAB release 2017b on a desktop computer with processor: Intel core i7 – 4790 CPU 3.6 GHz, and with 16.0 GB RAM.

We implement the inversion algorithm using the weight $\lambda = 10^7$ on the cross-gradient constraint. This selection is based on an analysis of entries of matrix **B**, which are very small. If we want to give enough weight to the cross-gradient term, it is necessary to use a large value for λ. On the other hand, if λ is selected too large, the results can be unsatisfactory. To show the effectiveness of applying the hard constraint matrix, when available information provide the values of some model parameters, we suppose that the physical properties of four prisms of the small cube are known. These prisms are located in the south-west corner of the cube in the second layer of the model which has a depth from 100 to 200m. Further, available information also suggests that the prisms above and below these known prisms, in the first and third layers, do not have any contrast with the background. The known density and susceptibility values are included in the initial model, with other model parameters are set as 0. The corresponding entries of matrices $(W_h)_1$ and $(W_h)_2$ are set to 100. After IT = 53 iterations the convergence criteria, χ_1^2 and χ_2^2, are satisfied and the inversion terminates. In this simulation, the χ_2^2 termination is reached at iteration 21. The susceptibility model is recovered more quickly than the density model, which requires additional iterations until both termination criteria are satisfied. The required time for performing the joint inversion is 2465 s. This compares with running Algorithm 1 without the alternating component, for which convergence is achieved in 59 iterations, requiring 4891s.

The 3-D visualizations of the reconstructed density and susceptibility models are illustrated in Fig. 4. Furthermore, three plane-sections of the reconstructed models are illustrated in Figs. 5 and 6, respectively. They are in good agreement with the original models and have nearly similar structures. The relative error of the reconstructed density and susceptibility models is 0.3794 and 0.3994, respectively. It is clear that sharp and focused images of the subsurface are obtained. Indeed, the results indicate that the algorithm is practical, can handle large data sets for the joint inversion in a reasonable time with a significant reduction in memory. Furthermore, it is clear that the known model parameters are kept fixed during the iterations, and, there is no incorrect extension of the small cube with known values. This demonstrates how the incorporation of available information for some model parameters can increase the reliability of the model obtained using the inverse algorithm. Finally, the gravity and magnetic data produced by the reconstructed density and susceptibility models are illustrated in Fig. 7.

In contrast, we also implement the algorithm for the case in which the cross-gradient constraint is not used in the inversion process. This is easy to do by selecting $\lambda = 0$. All other parameters of the implementation

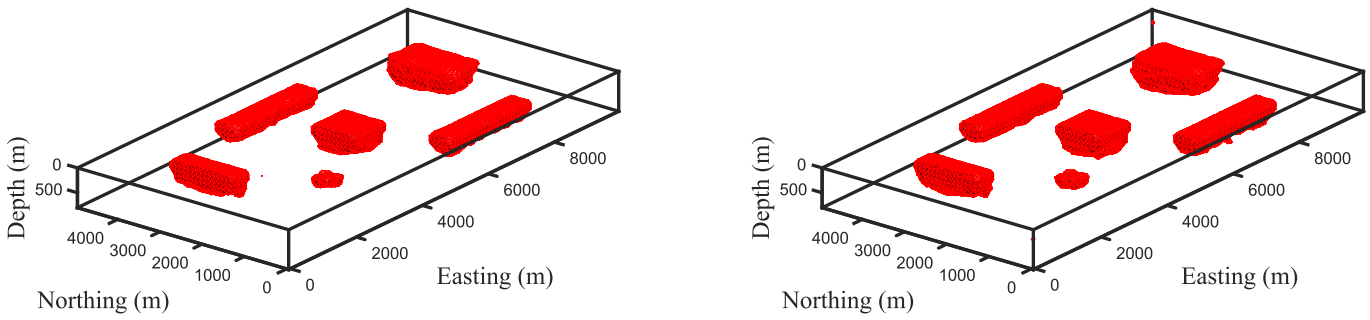

Fig. 4. 3-D view of the reconstructed models using the joint inversion algorithm with the L_1-norm of the model parameters as the stabilizer. (a) Density distribution. (b) Susceptibility distribution.

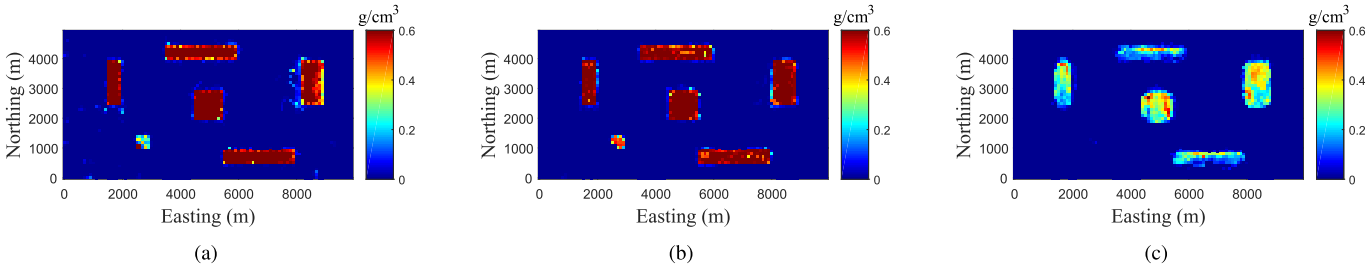

Fig. 5. Three plane-sections of the reconstructed density model using the joint inversion algorithm with the L_1-norm of the model parameters as the stabilizer. The plane-sections are presented at: (a) depth 100 m; (b) depth 200 m; and (c) depth 400 m.

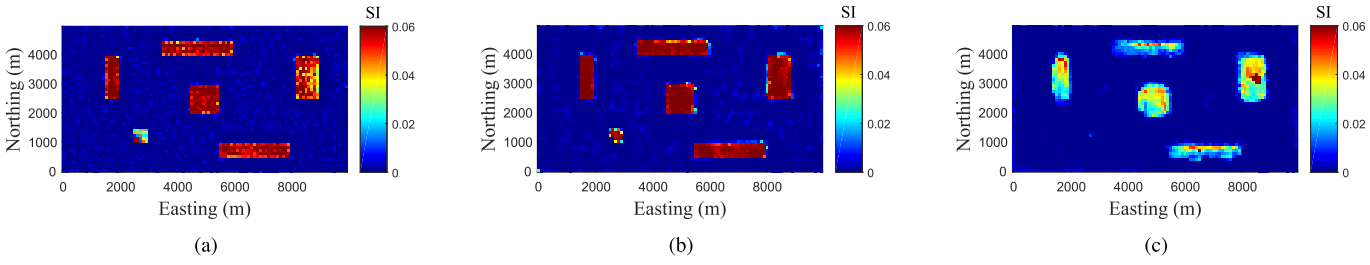

Fig. 6. Three plane-sections of the reconstructed susceptibility model using the joint inversion algorithm with the L_1-norm of the model parameters as the stabilizer. The plane-sections are presented at: (a) depth 100 m; (b) depth 200 m; and (c) depth 400 m.

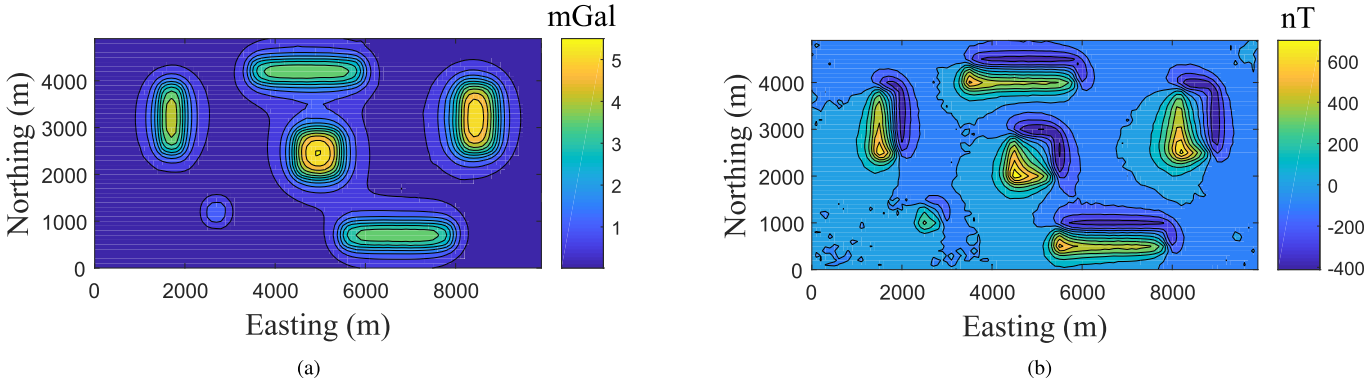

Fig. 7. Data produced by the models shown in Figs. 5 and 6. (a) Vertical component of gravity. (b) Total magnetic field.

are selected as before, with the same termination criteria, but without the cross-gradient term. The inversion again terminates after IT = 53 iterations, but now the relative error of the reconstructed density and susceptibility models are 0.4353 and 0.4317, respectively. These errors are larger than achieved when the cross-gradient is used. The reconstructed models are presented in Figs. 8 and 9. Although both models are close to the exact models, and

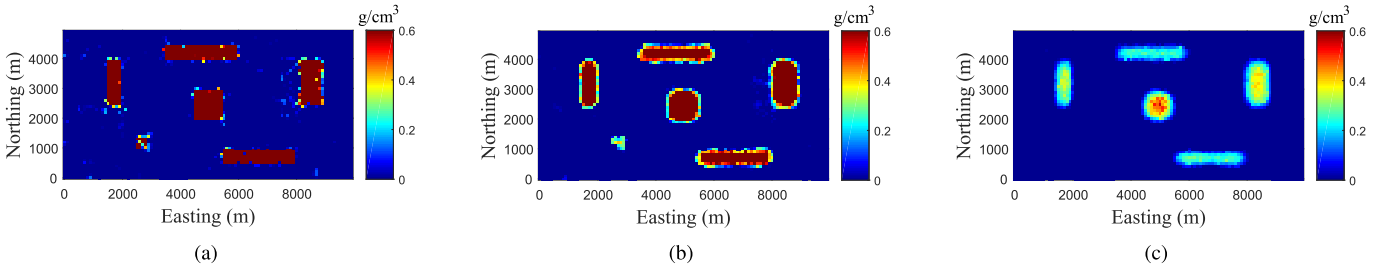

Fig. 8. Three plane-sections of the reconstructed density model using the inversion algorithm without using the cross-gradient constraint, i.e., for $\lambda = 0$. The plane-sections are presented at: (a) depth 100 m; (b) depth 200 m; and (c) depth 400 m.

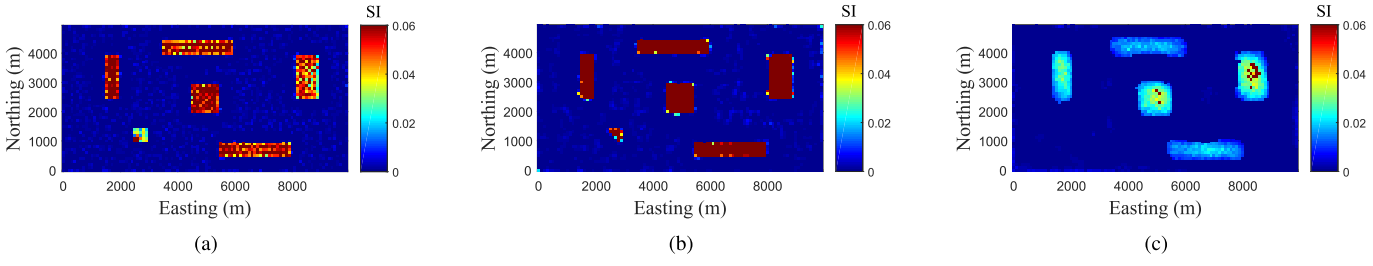

Fig. 9. Three plane-sections of the reconstructed susceptibility model using the inversion algorithm without using the cross-gradient constraint, i.e., for $\lambda = 0$. The plane-sections are presented at: (a) depth 100 m; (b) depth 200 m; and (c) depth 400 m.

thus geophysically acceptable, they are not as similar to each other as is the case when including the cross-gradient constraint.

We should note here that, without presenting the results, when parameter λ is selected too large, the inversion does not terminate at MAXIT. This means that it is not possible to satisfy the data misfit criteria with a large λ. The reconstructed models are not at all consistent with the original models, either with respect to the shape or to the maximum values of the physical properties. Although the main reconstructed bodies are quite similar to each other, as expected due to the strong requirement imposed by the use of the cross-gradient constraint, some additional unrealistic structures appear in the models. Clearly, the selection of λ is very important. But, this is not a difficult task. It is sufficient to consider the entries of $\mathbf{B}$, or to run the algorithm once, to determine a suitable λ.

IV. Real Data

The presented joint inversion strategy is now applied to the gravity and magnetic data that have been obtained over the Galinge iron-ore deposit in the Qinghai province, in NW China. The survey area is located in the center of the Qimantage metallogenic belt, one of the important skarn iron deposits in the province. In the given area, the bedrock is covered by a layer of Quaternary gravels with 117–210 m thickness [59]. Fig. 10 illustrates a geological map of the survey area. The defined stratum sequence, syn-genetic breccia of the horizons, and volcanic and subvolcanic rocks are the major factors determining the iron formations [60], [61]. The iron-ore deposit has probably been formed by volcanic exhalation and sedimentation as a reformed and superimposed deposit. The ore minerals mainly consist of magnetite and small amounts of hematite and siderite. There is a significant physical contrast between the orebodies and the surrounding rocks. Fig. 11 illustrates the gridded gravity and magnetic anomalies over the white rectangle of Fig. 10. Here, the data sets are gridded onto a grid of size 56×38, for a data set of size 2128 with 40 m spacing. The total field anomaly shows an isolated magnetic target elongated in the NW-SE direction. The situation for the gravity anomaly is not exactly the same, some additional anomalies appear in the observed data. This complicates the individual inversion and interpretation of the gravity data.

To perform the inversion, the subsurface is divided into $56 \times 38 \times 15 = 31\,920$ cubes of size 40 m in each dimension. We assume that the gravity and magnetic data are contaminated by Gaussian noise with standard deviation at each datum of $(\tau_1 \, |(\mathbf{d}_i^{\text{obs}})_j| + \tau_2 \, \max(|\mathbf{d}_i^{\text{obs}}|))$, $j = 1, \ldots, n$. Here, we take $(\tau_1, \tau_2) = (.01, .035)$ for the gravity data, and $(.01, .025)$ for the magnetic data. The intensity of the geomagnetic fields, the inclination and the declination are 53 800 nT, 56°, and −4°, respectively. We set MAXIT = 100, and impose bound constraints on the model parameters. In this case, these are $0 = \rho_{\min} \leq \mathbf{m}_1 \leq \rho_{\max} = 0.18$ gr cm^{-3}, and $0 = \kappa_{\min} \leq \mathbf{m}_2 \leq \kappa_{\max} = 0.2$, SI unit [59], [62]. The inversion algorithm terminates after IT = 60 iterations and requires 5477s. Note that the computational cost for this inversion is higher than the example used for the synthetic model. Although the respective data sets and a number of model parameters used for the real data are less than their counterparts for the synthetic data, $2128 < 5000$ and $31\,920 < 40\,000$, the number of layers in the model is increased from 8 to 15 for the real data. Thus, the 2-DFFT that is used per depth layer is smaller and the operations with the sensitivity matrices have to be accumulated over more depth layers. This increases the computational cost for each operation with the sensitivity matrices, and hence per iteration of the algorithm. Figs. 12 and 13 illustrate plane-sections of the reconstructed

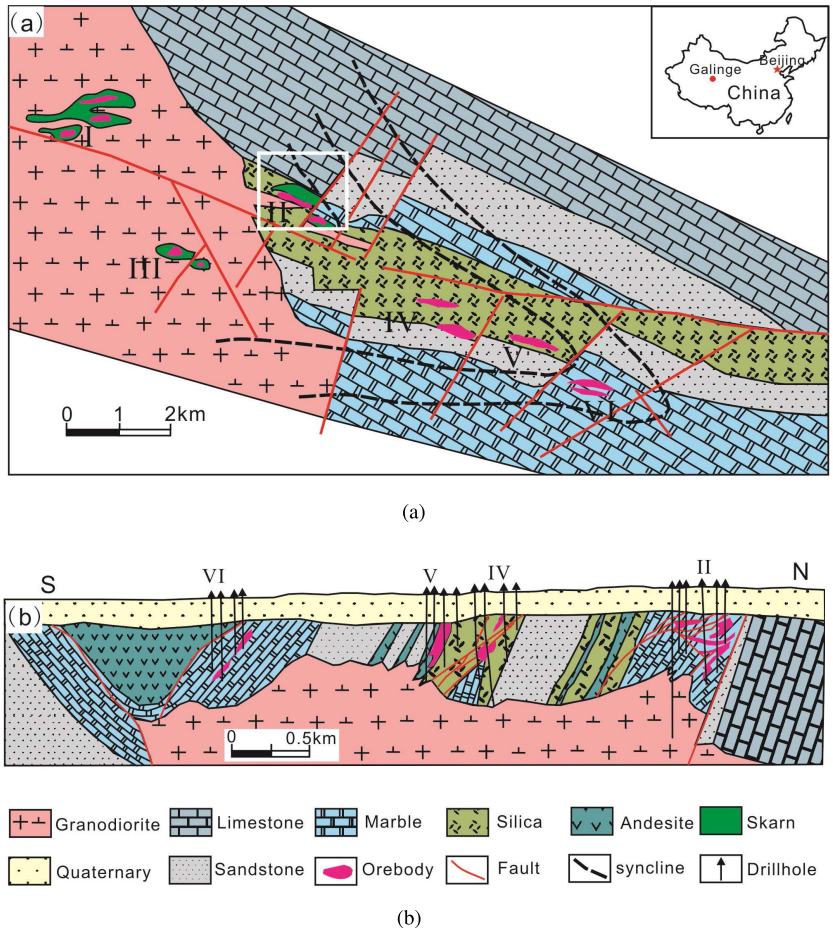

Fig. 10. (a) Geological map of the Galinge iron-ore deposit, NW China. The white rectangle indicates the geophysical survey area. (b) Cross section of the geological map shows bedrock, orebodies, and surrounding rocks [60].

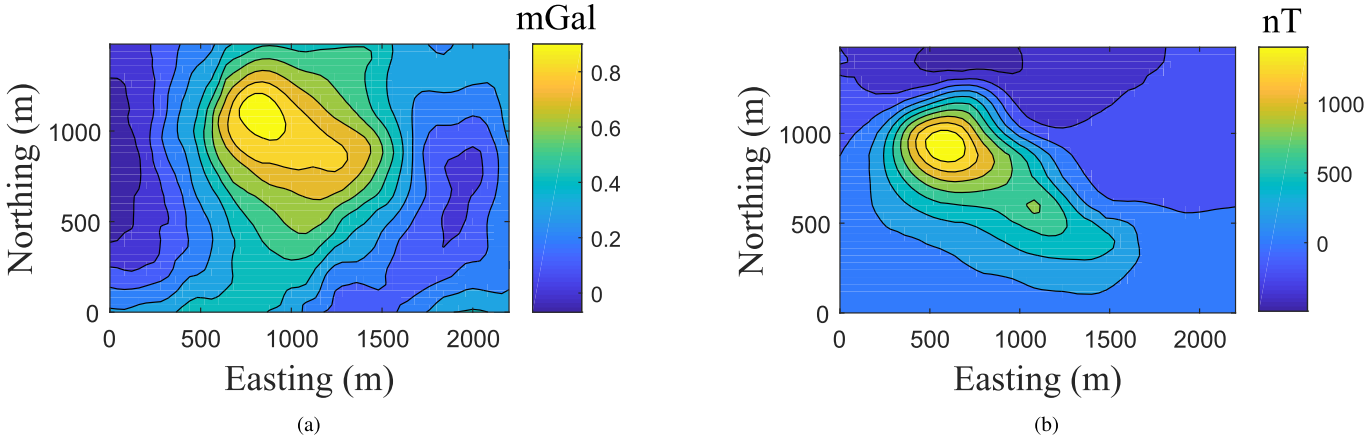

Fig. 11. Geophysical data over the Galinge iron-ore deposit, for the white rectangle in Fig. 10. (a) Gravity anomaly. (b) Magnetic total field anomaly.

density and susceptibility models, respectively. The 3-D view of the models is also shown in Fig. 14. Furthermore, the data produced by the reconstructed models are shown in Fig. 15. Generally, the target starts from depth about 160 m and continues to depth about 560 m, which is consistent with information from drill-holes and a previous study in the survey area [59]. The susceptibility model presents a single isolated subsurface target, but there are additional structures in the density distribution due to the additional gravity anomalies. As anticipated, the cross-gradient constraint only enforces structural similarity which is supported by the data; the additional structure is not imposed on the susceptibility due to the additional density structures. For comparison, we present the results of the inversion without the cross-gradient constraint

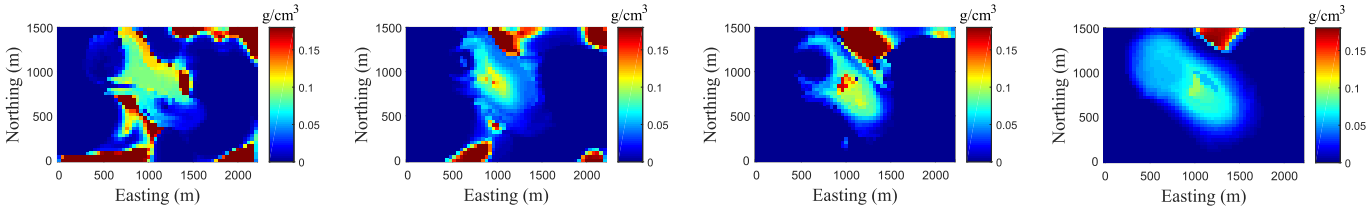

Fig. 12. Plane-sections of the reconstructed density model using the joint inversion algorithm for the data shown in Fig. 11. (a) Depth 160 m; (b) depth 300 m; (c) depth 440 m; and (d) depth 550.

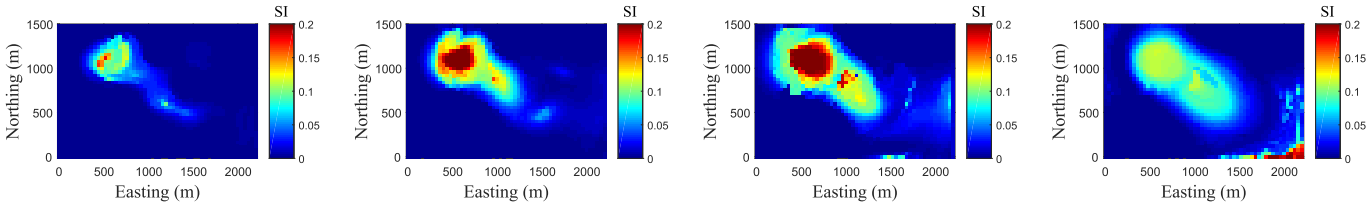

Fig. 13. Plane-sections of the reconstructed susceptibility model obtained using the joint inversion algorithm for the data shown in Fig. 11. (a) Depth 160 m; (b) depth 300 m; (c) depth 440 m; and (d) depth 550.

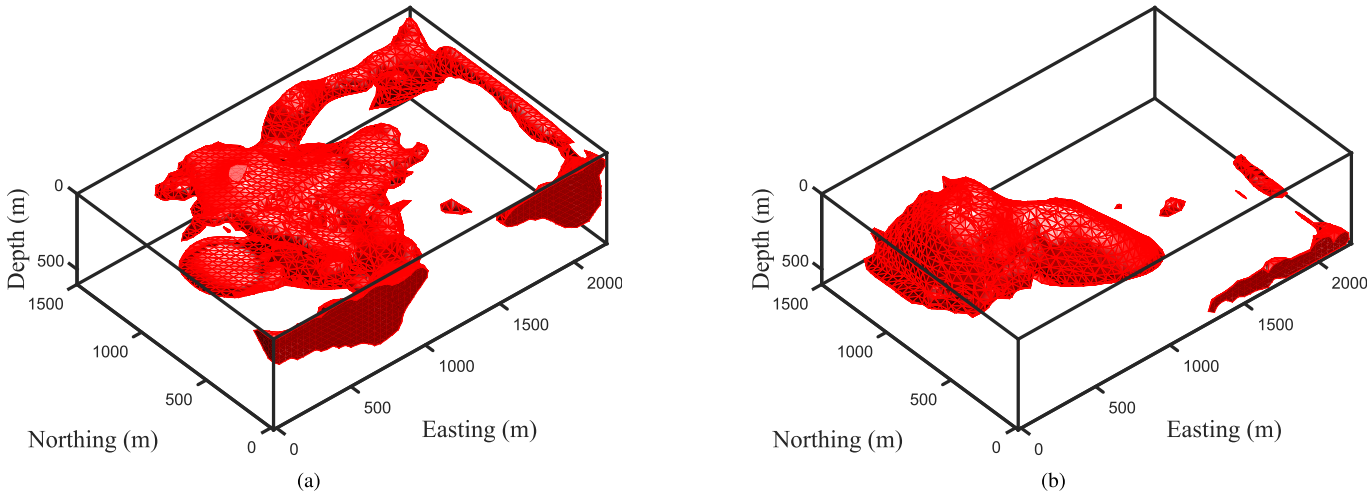

Fig. 14. 3-D view of the reconstructed models using the joint inversion algorithm for data shown in Fig. 11. (a) Density distribution. (b) Susceptibility distribution.

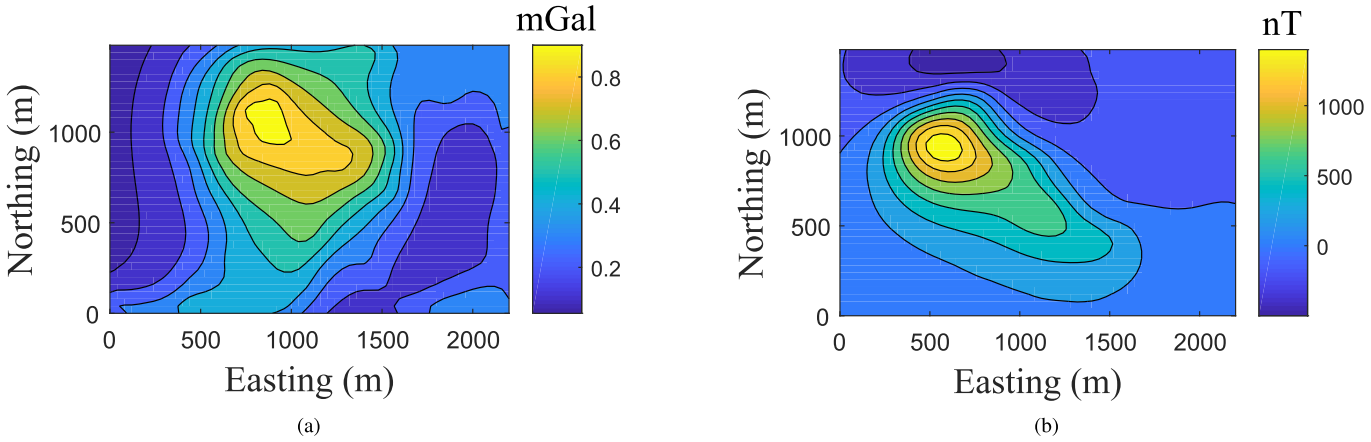

Fig. 15. Data predicted by the model shown in Fig. 14. (a) Gravity anomaly. (b) Magnetic total field anomaly.

in Figs. 16 and 17. The inversion terminates at iteration 84 and requires 6310 s. Clearly, there is no similarity between the reconstructed models. The density model is extended at shallower depths, which is not consistent with the drill-hole information. The results show how the joint inversion can improve the accuracy of the reconstructed models.

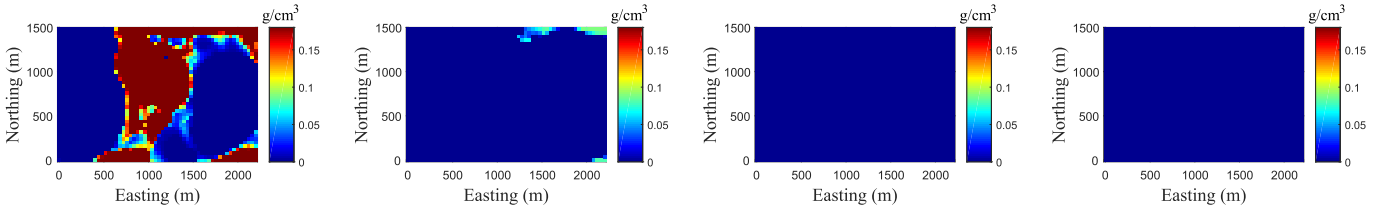

Fig. 16. Plane-sections of the reconstructed density model without using the cross-gradient constraint for the data shown in Fig. 11. (a) Depth 160 m. (b) Depth 300 m. (c) Depth 440 m. (d) Depth 550.

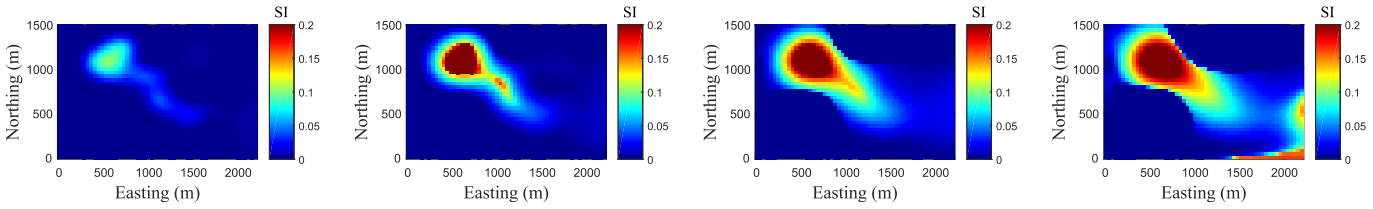

Fig. 17. Plane-sections of the reconstructed susceptibility model without using the cross-gradient constraint for the data shown in Fig. 11. (a) Depth 160 m. (b) Depth 300 m. (c) Depth 440 m. (d) Depth 550.

V. Conclusion

A framework for the incorporation of $L_{\mathbf{p}}$-norm stabilizers in an algorithm for joint inversion of gravity and magnetic data, in which the cross-gradient constraint provides the link between the two models, has been developed. This framework shows how it is possible to incorporate all well-known and widely used stabilizers, that are used for potential field inversion, within a joint inversion algorithm with the cross-gradient constraint. By suitable choices of the parameter $\mathbf{p}$ and the weighting matrix, that define the $L_{\mathbf{p}}$-norm constraint, it is possible to reconstruct a subsurface target exhibiting smooth, sparse, or blocky characteristics. The global objective function for the joint inversion consists of a data misfit term, a general form for the stabilizer, and the cross-gradient constraint. Their contributions to the global objective function are obtained using three different regularization parameters. A simple iterative strategy is used to convert the global nonlinear objective function to a linear form at each iteration, and the regularization weights can be adjusted at each iteration. Depth weighting and hard constraint matrices are also used in the presented inversion algorithm. These make it possible to weight prisms at depth and to include the known values of some prisms in the reconstructed model. Bound constraints on the model parameters may also be imposed at each iteration.

To make the algorithm effective for large-scale data sets, we take advantage of the underlying BTTB structure of the sensitivity matrices when the measurement data are obtained on a uniform grid. This allows an implementation in which all operations with the sensitivity matrices use the 2-DFFT. Moreover, the storage of the necessary transform matrices is significantly reduced as compared to the storage of the sensitivity matrices. Furthermore, we have demonstrated that it is possible to provide a more efficient inversion algorithm using alternating updates for the gravity and magnetic model parameters while maintaining the coupling of the two data sets through the cross-gradient constraint. This facilitates the inversion of large data sets in reasonable time and with reduced storage requirements. Results presented for a synthetic 3-D multiple model illustrate the performance of the developed algorithm. These results indicate that, when suitable regularization parameters can be estimated, the joint inversion algorithm yields suitable reconstructions of the subsurface structures. These reconstructions are improved in comparison with reconstructions obtained using independent gravity and magnetic inversions. The structures of the subsurface targets, for both density and susceptibility distributions, are similar and are close to the original models. A simple but practical strategy for the estimation of the regularization parameters is provided, by which large values are used at the initial step of the iteration, with a gradual decrease in subsequent iterations, dependent on selected scaling parameters for each of the imposed gravity and magnetic constraint terms. The weight on the cross-gradient linkage constraint is chosen to balance the three regularization terms. The results show that this strategy is effective, particularly given the lack of any known robust methods for automatically estimating these parameters.

Appendix

Although the expressions for the components of $\mathbf{t}$ and the Jacobian matrix $\mathbf{B}$ are found in the literature, [8], here we use a more compact description which corresponds to the implementation and is designed for computational and storage efficiency. As illustrated in Fig. 18, the subsurface is commonly divided into right rectangular prisms. Here, we suppose all prisms have same dimensions and that $\mathbf{m}_{ijk}$ represents the value of the current estimate for $\mathbf{m}$ at $(x_i, y_j, z_k) = (i\Delta x, j\Delta y, k\Delta z)$ where $0 \leq i \leq n_x - 1$, $0 \leq y \leq n_y - 1$ and $0 \leq k \leq n_z - 1$, for n_x, n_y and n_z blocks in x, y and z directions. The origin, $\mathbf{m}_{000}$, $(i = j = k = 0)$, is at the top left corner of the domain. All other parameters indexed in the same way

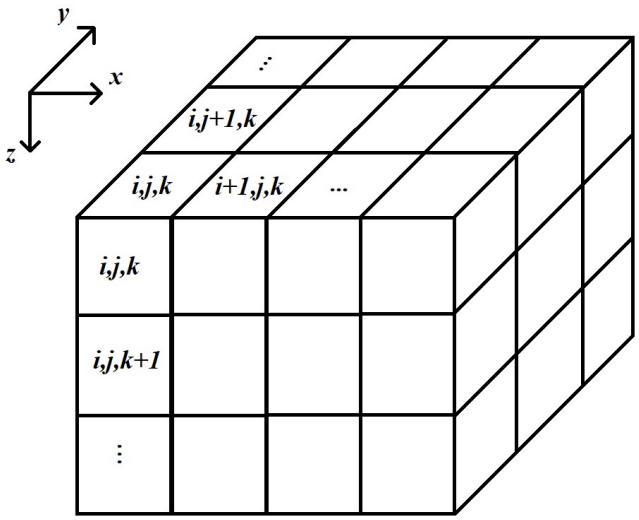

Fig. 18. Discretization of the subsurface into right rectangular prisms.

by ijk also correspond to the parameter at the given grid point (x_i, y_j, z_k) of the volume. We also use $\mathbf{t}^x(x, y, z)$ to indicate $\mathbf{t}^x(\mathbf{m}_1(x, y, z), \mathbf{m}_2(x, y, z))$ which is the x component of the cross-gradient function $\mathbf{t}$. Then, with this notation also applied for the y and z components, the components of the cross-gradient function (2) are given by

$$\mathbf{t}^x(x, y, z) = \frac{\partial \mathbf{m}_1}{\partial y}\frac{\partial \mathbf{m}_2}{\partial z} - \frac{\partial \mathbf{m}_1}{\partial z}\frac{\partial \mathbf{m}_2}{\partial y} \in \mathcal{R}^n \tag{19}$$

$$\mathbf{t}^y(x, y, z) = \frac{\partial \mathbf{m}_1}{\partial z}\frac{\partial \mathbf{m}_2}{\partial x} - \frac{\partial \mathbf{m}_1}{\partial x}\frac{\partial \mathbf{m}_2}{\partial z} \in \mathcal{R}^n \tag{20}$$

$$\mathbf{t}^z(x, y, z) = \frac{\partial \mathbf{m}_1}{\partial x}\frac{\partial \mathbf{m}_2}{\partial y} - \frac{\partial \mathbf{m}_1}{\partial y}\frac{\partial \mathbf{m}_2}{\partial x} \in \mathcal{R}^n. \tag{21}$$

This yields

$$\mathbf{t}(\mathbf{m}(x, y, z)) = \text{block\_stack}\left(\mathbf{t}^x, \mathbf{t}^y, \mathbf{t}^z\right) \in \mathcal{R}^{3n}.$$

We use forward difference operators to approximate the derivatives in $\mathbf{t}$, and for compact representation we introduce $D_x\mathbf{m}$ as the approximation for the x derivative of vector $\mathbf{m}$ at grid point ijk, with the y and z derivatives defined similarly. Specifically

$$D_x\mathbf{m} = \left(\frac{\mathbf{m}_{i+1,jk} - \mathbf{m}_{ijk}}{\Delta x}\right) \approx \left(\frac{\partial \mathbf{m}}{\partial x}\right)_{ijk}. \tag{22}$$

Using this notation, the approximate components of the cross-gradient are given by

$$\mathbf{t}^x_{ijk} = D_y\mathbf{m}_1 D_z\mathbf{m}_2 - D_z\mathbf{m}_1 D_y\mathbf{m}_2 \tag{23}$$

$$\mathbf{t}^y_{ijk} = D_z\mathbf{m}_1 D_x\mathbf{m}_2 - D_x\mathbf{m}_1 D_z\mathbf{m}_2 \tag{24}$$

$$\mathbf{t}^z_{ijk} = D_x\mathbf{m}_1 D_y\mathbf{m}_2 - D_y\mathbf{m}_1 D_x\mathbf{m}_2 \tag{25}$$

and we see that to calculate $\mathbf{t}$ we only need the calculation the approximate x, y, and z derivatives of $\mathbf{m}_1$ and $\mathbf{m}_2$ at each grid point (i, j, k).

The Jacobian matrix for the approximate cross-gradient function is given by

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{1x} & \mathbf{B}_{2x} \\ \mathbf{B}_{1y} & \mathbf{B}_{2y} \\ \mathbf{B}_{1z} & \mathbf{B}_{2z} \end{pmatrix} = (\mathbf{B}_1, \mathbf{B}_2) \in \mathcal{R}^{3n\times 2n} \tag{26}$$

Algorithm 3 Calculate Matrix $\mathbf{B}_1$

Require: Δx, Δy, Δz, n_x, n_y, n_z, m, n, $D_x\mathbf{m}$, $D_y\mathbf{m}$, $D_z\mathbf{m}$.

$Bx = \texttt{sparse}(\text{zeros}(n))$; $By = Bx$; $Bz = Bx$; $s = 1$; $d = n_x - 1$.
$dmx = \text{reshape}(D_x\mathbf{m})(n_x, n_y, n_z)$;
$dmy = \text{reshape}(D_y\mathbf{m})(n_x, n_y, n_z)$;
$dmz = \text{reshape}(D_z\mathbf{m})(n_x, n_y, n_z)$.
for $k = 1 : n_z$ **do**
  for $j = 1 : n_y$ **do**
    $I = s : s + d$; % Diagonal block and $i + 1$ block
    $Bx(I, I) = \texttt{sparse}(\text{diag}(dym(:, j, k)/\Delta z - dzm(:, j, k)/\Delta y))$.
    $By(I, I) = \texttt{sparse}(\text{diag}(dzm(:, j, k)/\Delta x - dxm(:, j, k)/\Delta z) - \text{diag}(dzm(1 : d, j, k)/\Delta x, 1))$.
    $Bz(I, I) = \texttt{sparse}(\text{diag}(dxm(:, j, k)/\Delta y - dym(:, j, k)/\Delta x) + \text{diag}(dym(1 : d, j, k)/\Delta x, 1))$.
    if $j < n_y$ **then**
      $J = s + d + 1 : s + 2d - 1$; % Block $j + 1$.
      $Bx(I, J) = \texttt{sparse}(\text{diag}(dzm(:, j, k)/\Delta y))$.
      $Bz(I, J) = \texttt{sparse}(-\text{diag}(dxm(:, j, k)/\Delta y))$.
    end if
    if $k < n_z$ **then**
      $J = s + m : s + m + d - 1$; % Block $k + 1$.
      $Bx(I, J) = \texttt{sparse}(-\text{diag}(dym(:, j, k)/\Delta z))$.
      $By(I, J) = \texttt{sparse}(\text{diag}(dxm(:, j, k)/\Delta z))$.
    end if
    $s = s + n_x$
  end for
end for
$B = [Bx; By; Bz]$;

Ensure: B.

where

$$\mathbf{B}_{1x} = \frac{\partial \mathbf{t}^x}{\partial \mathbf{m}_1}, \quad \mathbf{B}_{1y} = \frac{\partial \mathbf{t}^y}{\partial \mathbf{m}_1}, \quad \mathbf{B}_{1z} = \frac{\partial \mathbf{t}^z}{\partial \mathbf{m}_1} \tag{27}$$

with the equivalent definition for the components of $\mathbf{B}_2$ as partial derivatives with respect to $\mathbf{m}_2$. Now, using (22), we see that

$$\frac{\partial (D_x\mathbf{m})}{\partial \mathbf{m}_{pqr}} = \frac{1}{\Delta x}\begin{cases} 1 & p = i + 1 \; q = j \; r = k \\ -1 & p = i \; q = j \; r = k \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

Applying this for the discrete derivative of $\mathbf{t}^x$ with respect to $\mathbf{m}_1$ and $\mathbf{m}_2$, we obtain for the $p = i^{\text{th}}$ rows

$$(\mathbf{B}_{1\ x})_{ijk,pqr} = \begin{cases} \dfrac{D_y\mathbf{m}_2}{\Delta z} - \dfrac{D_z\mathbf{m}_2}{\Delta y} & q, r = j, k \\ \dfrac{D_z\mathbf{m}_2}{\Delta y} & q, r = j + 1, k \\ -\dfrac{D_y\mathbf{m}_2}{\Delta z} & q, r = j, k + 1 \end{cases}$$

$$(\mathbf{B}_{2\ x})_{ijk,pqr} = \begin{cases} \dfrac{D_z\mathbf{m}_1}{\Delta y} - \dfrac{D_y\mathbf{m}_1}{\Delta z} & q, r = j, k \\ -\dfrac{D_z\mathbf{m}_1}{\Delta y} & q, r = j + 1, k \\ \dfrac{D_y\mathbf{m}_1}{\Delta z} & q, r = j, k + 1. \end{cases} \tag{29}$$

Algorithm 4 Calculate $D_x\mathbf{m}$, $D_y\mathbf{m}$, and $D_z\mathbf{m}$

Require: Data $\mathbf{m}$, n_x, n_y, n_z, $\hat{D}_x$, $\hat{D}_y$, $\hat{D}_z$.

$\mathbf{m} = \text{reshape}(\mathbf{m}, n_x, n_y, n_z)$;
$mext = \text{zeros}(2n_x - 1, n_y, n_z)$;
$mext(1 : n_x, :, :) = \mathbf{m}$;
$dm = \text{real}(\texttt{ifft}(\hat{D}_x . * \texttt{fft}(mext)))$;
$dxm = dm(2 : n_x + 1, :, :)$;
$\mathbf{m} = \texttt{permute}(\mathbf{m}, [2, 1, 3])$;
$mext = \text{zeros}(2n_y - 1, n_x, n_z)$;
$mext(1 : n_y, :, :) = \mathbf{m}$;
$dm = \text{real}(\texttt{ifft}(\hat{D}_y . * \texttt{fft}(mext)))$;
$dym = \texttt{permute}(dm(2 : n_y + 1, :, :), [2, 1, 3])$;
$\mathbf{m} = \texttt{permute}(\mathbf{m}, [3, 1, 2])$;
$mext = \text{zeros}(2n_z - 1, n_y, n_x)$;
$mext(1 : n_z, :, :) = \mathbf{m}$;
$dm = \text{real}(\texttt{ifft}(\hat{D}_z . * \texttt{fft}(mext)))$;
$dzm = \texttt{permute}(dm(2 : n_z + 1, :, :), [3, 2, 1])$;

Ensure: dxm, dym, dzm.

Algorithm 5 Calculate $\hat{D}_x$, $\hat{D}_y$, and $\hat{D}_z$

Require: n_x, n_y, n_z, Δx, Δy, Δz.

$r = \text{zeros}(2 * n_x - 1, 1)$; $r(1 : 2) = [1, -1]/\Delta x$; $\hat{Dx} = \texttt{fft}(r)$;
$r = \text{zeros}(2 * n_y - 1, 1)$; $r(1 : 2) = [1, -1]/\Delta y$; $\hat{Dy} = \texttt{fft}(r)$;
$r = \text{zeros}(2 * n_z - 1, 1)$; $r(1 : 2) = [1, -1]/\Delta z$; $\hat{Dz} = \texttt{fft}(r)$;

Ensure: $\hat{D}_x$, $\hat{D}_y$ and $\hat{D}_z$.

Hence there are just six nonzero column entries in each row, which are calculated directly from the derivative vectors used already in (23). The matrices are sparse with a block structure, and the sum of the entries in any given row is 0, except for rows that correspond to the last row for a given block matrix, $i = n_x$, $j = n_y$ or $k = n_z$. The entries for B_{iy} and B_{iz} are calculated similarly. The compact algorithm for the calculation of $\mathbf{B}_1$ is given in Algorithm 3, using inputs $D_x\mathbf{m}_2$, $D_y\mathbf{m}_2$, and $D_z\mathbf{m}_2$, along with the problem specific scalars. A similar approach follows for $\mathbf{B}_2$ using $D_x\mathbf{m}_1$, $D_y\mathbf{m}_1$, $D_z\mathbf{m}_1$. Here, we note that the algorithm counts over the row index s which counts through the triples qr for each block of rows $s : s + n_x - 1$, but that it is possible to generate the same matrices by blocks, employing diag for the n_y and mth upper diagonals. Experiments demonstrated that this was less efficient. The required partial derivatives are calculated efficiently using Algorithm 4 which employs the Fourier coefficients for derivatives in x-, y- and z-directions obtained using Algorithm 5, and calculated only once before the start of the inversion.

ACKNOWLEDGMENT

The authors would like to thank the Editor and an Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and two anonymous reviewers. Their valuable comments helped us to significantly improve the quality of this article.

REFERENCES

[1] R. J. Blakely, *Potential Theory in Gravity & Magnetic Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
[2] M. N. Nabighian *et al.*, "Historical development of the gravity method in exploration," *Geophysics*, vol. 70, no. 6, pp. 63–89, 2005.
[3] L. A. Gallardo and M. A. Meju, "Characterization of heterogeneous near-surface materials by joint 2D inversion of DC resistivity and seismic data," *Geophys. Res. Lett.*, vol. 30, no. 13, p. 1658, Jul. 2003.
[4] L. Nielsen and B. H. Jacobsen, "Integrated gravity and wide-angle seismic inversion fortwo-dimensional crustal modelling," *Geophys. J. Int.*, vol. 140, no. 1, pp. 222–232, Jan. 2000.
[5] M. Moorkamp, B. Heincke, M. Jegen, A. W. Roberts, and R. W. Hobbs, "A framework for 3-D joint inversion of MT, gravity and seismic refraction data," *Geophys. J. Int.*, vol. 184, no. 1, pp. 477–493, 2011.
[6] M. Moorkamp, A. W. Roberts, M. Jegen, B. Heincke, and R. W. Hobbs, "Verification of velocity-resistivity relationships derived from structural joint inversion with borehole data," *Geophys. Res. Lett.*, vol. 40, no. 14, pp. 3596–3601, Jul. 2013.
[7] E. Haber and D. Oldenburg, "Joint inversion: A structural approach," *Inverse Problems*, vol. 13, no. 1, pp. 63–77, Feb. 1997.
[8] L. A. Gallardo, "Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints," *J. Geophys. Res.*, vol. 109, no. B3, 2004, Art. no. B03311.
[9] A. Tryggvason and N. Linde, "Local earthquake (LE) tomography with joint inversion for P-and S-wave velocities using structural constraints," *Geophys. Res. Lett.*, vol. 33, no. 7, 2006, Art. no. L07303.
[10] E. Fregoso and L. A. Gallardo, "Cross-gradients joint 3D inversion with applications to gravity and magnetic data," *Geophysics*, vol. 74, no. 4, pp. L31–L42, Jul. 2009.
[11] L. A. Gallardo, "Multiple cross-gradient joint inversion for geospectral imaging," *Geophys. Res. Lett.*, vol. 34, no. 19, 2007, Art. no. L19301.
[12] E. Haber and M. Holtzman Gazit, "Model fusion and joint inversion," *Surveys Geophys.*, vol. 34, no. 5, pp. 675–695, Sep. 2013.
[13] P. G. Lelièvre, C. G. Farquharson, and C. A. Hurich, "Joint inversion of seismic traveltimes and gravity data on unstructured grids with application to mineral exploration," *Geophysics*, vol. 77, no. 1, pp. K1–K15, Jan. 2012.
[14] E. Fregoso, L. A. Gallardo, and J. García-Abdeslem, "Structural joint inversion coupled with euler deconvolution of isolated gravity and magnetic anomalies," *Geophysics*, vol. 80, no. 2, pp. G67–G79, Mar. 2015.
[15] L. Gross, "Weighted cross-gradient function for joint inversion with the application to regional 3-D gravity and magnetic anomalies," *Geophys. J. Int.*, vol. 217, no. 3, pp. 2035–2046, Jun. 2019.
[16] Y. Z. A. Y. W. Sci, "Three-dimensional gravity-magnetic cross-gradient joint inversion based on structural coupling and a fast gradient method," *J. Comput. Math.*, vol. 37, no. 6, pp. 758–777, Jun. 2019.
[17] M. S. Zhdanov, A. Gribenko, and G. Wilson, "Generalized joint inversion of multimodal geophysical data using gramian constraints," *Geophys. Res. Lett.*, vol. 39, no. 9, pp. 1–7, May 2012, Art. no. L09301, doi: 10.1029/2012GL051233.
[18] M. S. Zhdanov, *Inverse Theory and Applications in Geophysics*. Amsterdam, The Netherlands: Elsevier, 2015.
[19] W. Lin and M. S. Zhdanov, "Joint multinary inversion of gravity and magnetic data using Gramian constraints," *Geophys. J. Int.*, vol. 215, no. 3, pp. 1540–1577, Aug. 2018.
[20] M. Jorgensen and M. S. Zhdanov, "Imaging yellowstone magmatic system by the joint Gramian inversion of gravity and magnetotelluric data," *Phys. Earth Planet. Interiors*, vol. 292, pp. 12–20, Jul. 2019.
[21] S. C. Constable, R. L. Parker, and C. G. Constable, "Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data," *Geophysics*, vol. 52, no. 3, pp. 289–300, Mar. 1987.
[22] Y. Li and D. W. Oldenburg, "3-D inversion of magnetic data," *Geophysics*, vol. 61, no. 2, pp. 394–408, 1996.
[23] M. Pilkington, "3-D magnetic imaging using conjugate gradients," *Geophysics*, vol. 62, no. 4, pp. 1132–1142, Jul. 1997.
[24] Y. Li and D. W. Oldenburg, "3-D inversion of gravity data," *Geophysics*, vol. 63, no. 1, pp. 109–119, Jan. 1998.
[25] C. G. Farquharson and D. W. Oldenburg, "Non-linear inversion using general measures of data misfit and model structure," *Geophys. J. Int.*, vol. 134, no. 1, pp. 213–227, 1998.

[26] O. Portniaguine and M. S. Zhdanov, "Focusing geophysical inversion images," *Geophysics*, vol. 64, no. 3, pp. 874–887, May 1999.

[27] H. Bertete-Aguirre, E. Cherkaev, and M. Oristaglio, "Non-smooth gravity problem with total variation penalization functional," *Geophys. J. Int.*, vol. 149, no. 2, pp. 499–507, May 2002.

[28] S. Vatankhah, R. A. Renaut, and V. E. Ardestani, "Total variation regularization of the 3-D gravity inverse problem using a randomized generalized singular value decomposition," *Geophys. J. Int.*, vol. 213, no. 1, pp. 695–705, Apr. 2018.

[29] D. Fournier and D. W. Oldenburg, "Inversion using spatially variable mixed ℓ_p norms," *Geophys. J. Int.*, vol. 218, no. 1, pp. 268–282, Jul. 2019.

[30] B. J. Last and K. Kubik, "Compact gravity inversion," *Geophysics*, vol. 48, no. 6, pp. 713–721, Jun. 1983.

[31] V. C. F. Barbosa and J. B. C. Silva, "Generalized compact gravity inversion," *Geophysics*, vol. 59, no. 1, pp. 57–68, Jan. 1994.

[32] M. Zhdanov and E. Tolstaya, "Minimum support nonlinear parametrization in the solution of a 3D magnetotelluric inverse problem," *Inverse Problems*, vol. 20, no. 3, pp. 937–952, Jun. 2004.

[33] J. B. Ajo-Franklin, B. J. Minsley, and T. M. Daley, "Applying compactness constraints to differential traveltime tomography," *Geophysics*, vol. 72, no. 4, pp. R67–R75, Jul. 2007.

[34] S. Vatankhah, V. E. Ardestani, and R. A. Renaut, "Automatic estimation of the regularization parameter in 2D focusing gravity inversion: Application of the method to the safo manganese mine in the northwest of iran," *J. Geophys. Eng.*, vol. 11, no. 4, Aug. 2014, Art. no. 045001.

[35] S. Vatankhah, R. A. Renaut, and V. E. Ardestani, "Regularization parameter estimation for underdetermined problems by the χ^2 principle with application to 2D focusing gravity inversion," *Inverse Problems*, vol. 30, no. 8, Aug. 2014, Art. no. 085002.

[36] J. Sun and Y. Li, "Adaptive lp inversion for simultaneous recovery of both blocky and smooth features in a geophysical model," *Geophys. J. Int.*, vol. 197, no. 2, pp. 882–899, May 2014.

[37] S. Vatankhah, R. A. Renaut, and V. E. Ardestani, "3-D projected l1 inversion of gravity data using truncated unbiased predictive risk estimator for regularization parameter estimation," *Geophys. J. Int.*, vol. 210, no. 3, pp. 1872–1887, Sep. 2017.

[38] S. Vatankhah, R. A. Renaut, and S. Liu, "Research note: A unifying framework for the widely used stabilization of potential field inverse problems," *Geophys. Prospecting*, vol. 68, no. 4, pp. 1416–1421, May 2020.

[39] O. Boulanger and M. Chouteau, "Constraint in 3D gravity inversion," *Geophys. Prospecting*, vol. 49, no. 2, pp. 265–280, 2001.

[40] A. Tarantola and B. Valette, "Generalized nonlinear inverse problems solved using the least squares criterion," *Rev. Geophys.*, vol. 20, no. 2, pp. 219–232, 1982.

[41] B. Wohlberg and P. Rodriguez, "An iteratively reweighted norm algorithm for minimization of total variation functionals," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 948–951, Dec. 2007.

[42] C. E. Bruun and T. B. Nielsen, "Algorithms and software for large-scale geophysical reconstructions," M.S. thesis, Tech. Univ. Denmark, Kongens Lyngby, Denmark, 2007.

[43] L. Chen and L. Liu, "Fast and accurate forward modelling of gravity field using prismatic grids," *Geophys. J. Int.*, vol. 216, no. 2, pp. 1062–1071, Feb. 2019.

[44] Y. Zhang and Y. S. Wong, "BTTB-based numerical schemes for three-dimensional gravity field inversion," *Geophys. J. Int.*, vol. 203, no. 1, pp. 243–256, Oct. 2015.

[45] J. D. Hogue, R. A. Renaut, and S. Vatankhah, "A tutorial and open source software for the efficient evaluation of gravity and magnetic kernels," *Comput. Geosci.*, vol. 144, Nov. 2020, Art. no. 104575.

[46] R. A. Renaut, J. D. Hogue, S. Vatankhah, and S. Liu, "A fast methodology for large-scale focusing inversion of gravity and magnetic data using the structured model matrix and the 2-D fast Fourier transform," *Geophys. J. Int.*, vol. 223, no. 2, pp. 1378–1397, Nov. 2020.

[47] I. B. Haáz, "Relations between the potential of the attraction of the mass contained in a finite rectangular prism and its first and second derivatives," *Geofizikai Kozlemenyek*, vol. 2, no. 7, pp. 57–66, 1953.

[48] D. B. Rao and N. R. Babu, "A rapid method for three-dimensional modeling of magnetic anomalies," *Geophysics*, vol. 56, no. 11, pp. 1729–1737, Nov. 1991.

[49] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[50] I. Selesnick, "Sparse regularization via convex analysis," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4481–4494, Sep. 2017.

[51] G. Fiandaca, J. Doetsch, G. Vignoli, and E. Auken, "Generalized focusing of time-lapse changes with applications to direct current and time-domain induced polarization inversions," *Geophys. J. Int.*, vol. 203, no. 2, pp. 1101–1112, Nov. 2015.

[52] Y. Li and D. W. Oldenburg, "Incorporating geological dip information into geophysical inversions," *Geophysics*, vol. 65, no. 1, pp. 148–157, Jan. 2000.

[53] P. G. Lelièvre and D. W. Oldenburg, "A comprehensive study of including structural orientation information in geophysical inversions," *Geophys. J. Int.*, vol. 178, no. 2, pp. 623–637, Aug. 2009.

[54] P. G. Lelièvre and C. G. Farquharson, "Gradient and smoothness regularization operators for geophysical inversion on unstructured meshes," *Geophys. J. Int.*, vol. 195, no. 1, pp. 330–341, Oct. 2013.

[55] S. Vatankhah, R. Anne Renaut, and V. E. Ardestani, "A fast algorithm for regularized focused 3D inversion of gravity data using randomized singular-value decomposition," *Geophysics*, vol. 83, no. 4, pp. G25–G34, Jul. 2018.

[56] C. G. Farquharson and D. W. Oldenburg, "A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems," *Geophys. J. Int.*, vol. 156, no. 3, pp. 411–425, Mar. 2004.

[57] S. Vatankhah, S. Liu, R. A. Renaut, X. Hu, and J. Baniamerian, "Improving the use of the randomized singular value decomposition for the inversion of gravity and magnetic data," *Geophysics*, vol. 85, no. 5, pp. G93–G107, Sep. 2020.

[58] J. D. Hogue, R. A. Renaut, and S. Vatankhah, "A tutorial and open source software for the efficient evaluation of gravity and magnetic kernels," 2019, *arXiv:1912.06976*. [Online]. Available: http://arxiv.org/abs/1912.06976

[59] S. Liu, X. Hu, Y. Xi, T. Liu, and S. Xu, "2D sequential inversion of total magnitude and total magnetic anomaly data affected by remanent magnetization," *Geophysics*, vol. 80, no. 3, pp. K1–K12, May 2015.

[60] M. Yu *et al.*, "40Ar-39Ar geochronology of the Galinge large skarn iron deposit in Qinghai province and geological significance," *Acta Geologica Sinica*, vol. 89, no. 3, pp. 510–521, 2015.

[61] M. Yu, "Geochemistry and zonation of the Galinge iron deposit, Qinghai province," M.S. thesis, School Earth Sci. Resour., China Univ. Geosci., Beijing, China, 2016.

[62] S. Liu, J. Baniamerian, and M. Fedi, "Imaging methods versus inverse methods: An option or an alternative?" *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3484–3494, May 2020.

Saeed Vatankhah received the M.Sc. and Ph.D. degrees in geophysics from the University of Tehran, Tehran, Iran, in 2009 and 2014, respectively. His Ph.D. dissertation focused on developing regularization parameter-choice rules for potential field inversion.

He has been an Assistant Professor of geophysics with the Institute of Geophysics, University of Tehran, since 2016. He is a Postdoctoral Research Fellow with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan, China. His main research interests include the inversion of gravity and magnetic potential field data using novel randomization techniques for large-scale problems, and the joint inversion of geophysical data sets. He has published several articles in major geophysics journals, such as *Geophysical Journal International*, *Geophysics*, and *Computers and Geosciences*.

Shuang Liu was born in Hubei, China, in 1985. He received the B.S., M.Sc., and Ph.D. degrees in geophysics from the China University of Geosciences, Wuhan, China, in 2008, 2011, and 2015, respectively.

He was a Postdoctoral Researcher with the Institute of Geology and Geophysics, Chinese Academic of Sciences, Beijing, China, and Napoli University of Federico II, Naples, Italy, from 2015 to 2018. He is an Associate Professor with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan. His research interest includes potential field data processing and inversion.

Rosemary Anne Renaut was born in Bristol, U.K., in 1958. She received the B.Sc. degree (Hons.) in mathematics from the University of Durham, Durham, U.K., in 1980, and the Part III of the Mathematical Tripos in applied mathematics and theoretical physics and the Ph.D. degree in applied mathematics from the University of Cambridge, Cambridge, U.K., in 1980 and 1985, respectively.

She has been holding a Faculty Position with Arizona State University, Tempe, AZ, USA, since 1987. She has held numerous Visiting Positions including with the Technical University of Kaiserslautern, Kaiserslautern, Germany, Technical University of Munich, Munich, Germany, and ETH Zürich, Zürich, Switzerland. She is a Chartered Mathematician in the U.K. and a fellow of the Institute of Mathematics and its Applications, Southend-on-Sea, U.K. She was also a Program Director with the National Science Foundation, Washington, DC, USA, from 2009 to 2012 and 2014 to 2017. She is a Full Professor with the Arizona State University. She has authored more than 100 research articles in areas of computational mathematics, and has general interests in the design of algorithms for the solution of practical inverse problems arising in geophysics and medical imaging.

Dr. Renaut was awarded the Royal Society Postdoctoral Fellowship for postdoctoral research at the Technical University of Aachen, Aachen, Germany.

Xiangyun Hu received the B.S. and M.Sc. degrees in geophysics from the Changchun College of Geology, Changchun, China, and the Ph.D. degree in mineral resource prospecting and exploration from the China University of Geosciences, Wuhan, China, in 2000.

Since 2005, he has been serving as a Professor with the Institute of Geophysics and Geomatics, China University of Geosciences. His research interests include 3-D MT/CSAMT/CSEM forward modeling and inverse problems, geoelectrical characteristics of the Tibetan Plateau, and integrated geophysical interpretation.

Jarom David Hogue was born in Orem, Utah, in 1983. He received the B.S. degree in mathematics and the master's and Ph.D. degrees in applied mathematics from Arizona State University, Tempe, AZ, USA, in 2014, 2016, and 2020, respectively.

He has authored articles in the area of computational mathematics, and has interests in modeling and efficient algorithms for the solution of inverse problems and dimensionality reduction techniques.

Mostafa Gharloghi was born in Tehran, Iran, in 1991. He received the B.S. degree in physics from Shahid Rajaee University, Tehran, in 2017. He is pursuing the M.Sc. degree in geophysics with the Institute of Geophysics, University of Tehran, Tehran.

His research interests include potential field data processing and the joint inversion of gravity and magnetic data sets.