Kantorovich-Rubinstein misfit for inverting gravity-gradient data by the level-set method

Guanghui Huang¹, Xinming Zhang², and Jianliang Qian³

ABSTRACT

We have developed a novel Kantorovich-Rubinstein (KR) norm-based misfit function to measure the mismatch between gravity-gradient data for the inverse gradiometry problem. Under the assumption that an anomalous mass body has an unknown compact support with a prescribed constant value of density contrast, we implicitly parameterize the unknown mass body by a level-set function. Because the geometry of an underlying anomalous mass body may experience various changes during inversion in terms of level-set evolution, the classic least-squares ($L_2$-norm-based) and the $L_1$-norm-based misfit functions for governing the level-set evolution may potentially induce local minima if an initial guess of the level-set function is far from that of the target model. The KR norm from the optimal transport theory computes the data misfit by comparing the modeled data and the measured data in a global manner, leading to better resolution of the differences between the inverted model and the target model. Combining the KR norm with the level-set method yields a new effective methodology that is not only able to mitigate local minima but is also robust against random noise for the inverse gradiometry problem. Numerical experiments further demonstrate that the new KR norm-based misfit function is able to recover deep dipping flanks of SEG/EAGE salt models even at extremely low signal-to-noise ratios. The new methodology can be readily applied to gravity and magnetic data as well.

INTRODUCTION

The gravity gradiometer measures the gravity gradient tensor. Correct interpretation of gravity-gradient data provides essential guidelines in analyzing the composition of the earth and in delineating subsurface source bodies, such as mineral deposits. However, because gravity-gradient data are higher order derivatives of the potential, manual interpretation of such data is extremely challenging, whereas automatic interpretation calls for developing efficient inversion methods. In the past several decades, a lot of effort has been made to develop various inversion methods for gravity or gravity-gradient data (Last and Kubik, 1983; Li and Oldenburg, 1998; Condi and Talwani, 1999; Portniaguine and Zhdanov, 1999; Jorgensen and Kisabeth, 2000; Li, 2001a, 2001b, 2010; Routh et al., 2001; Zhdanov et al., 2004; Krahenbuhl and Li, 2006; Barnes et al., 2008; Martinez et al., 2010, 2013; Barnes and Barraud, 2012). In these methods, prior geologic constraints on density models are usually enforced so that the resulting solutions conform to realistic earth models (Last and Kubik, 1983; Li and Oldenburg, 1998; Condi and Talwani, 1999; Portniaguine and Zhdanov, 1999). Because most inversion techniques provide quantitative descriptions for subsurface structures, one has to extract the position of a source body from the resulting solutions, which is not an easy task.

The level-set method (Osher and Sethian, 1988) has been widely used as a suitable and powerful tool for shape-optimization problems mainly due to its abilities in automatic interface merging and splitting. Assuming that a homogeneous density-contrast distribution with the value of density contrast specified a priori was supported on an unknown bounded domain, we can convert the inverse gravity gradient problem into a domain inverse problem by implicitly parameterizing the unknown anomalous mass body with a level-set function, and we apply the steepest-descent method to minimize the least-squares objective so as to find the corresponding level-set function. Along this line, a series of works by Isakov et al. (2011, 2013),...
Lu et al. (2015), Lu and Qian (2015), Li et al. (2016, 2017), and Li and Qian (2016) have developed various level-set methods to delineate subsurface source bodies; see also Li and Qian (2016) and Zhelglova et al. (2017) for applying the level-set method to carry out the joint inversion of traveltime and gravity data. However, because the geometry of the underlying domains may evolve in a very complicated way, these level-set methods based on the classical least-squares objective are vulnerable to local minima induced by the misfit function and sensitive to random noise as shown by Lu et al. (2015), Lu and Qian (2015), and Li et al. (2016, 2017).

The optimal transport theory (Villani, 2003) has been applied in seismic inversion to build novel objective functions for mitigating the cycle-skipping-induced local minimum issues (Virieux and Operto, 2009), where the cycle skipping is caused by the mismatch of the traveltime shift in full-waveform inversion problems (Béna and Brenier, 2001; Engquist and Froese, 2014; Metivier et al., 2016). One of the important features of the optimal transport theory and sensitive to random noise as shown by Lu et al. (2015), Lu and Qian (2015), and Li et al. (2016, 2017).

Because it is hard to satisfy all of the requirements needed for applying the original optimal transport theory to inverse problems (Engquist and Froese, 2014; Metivier et al., 2016), such as the nonnegativity of data, we appeal to a variant of optimal transport distances, the so-called Kantorovich-Rubinstein (KR) norm (Villani, 2003; Lellmann et al., 2014), to measure the data misfit for inverse gravity-gradient problems in that the KR norm does not require nonnegativity of data. Although we choose to work on gravity-gradient problems here, the methodology will be naturally applicable to other potential data as well.

The rest of the paper is organized as follows. We present the methodology by reviewing the least-squares-based level-set method, developing the KR norm-based level-set method for gravity-gradient data, and further illustrating some properties of the new KR norm-based misfit function for gravity-gradient inversion. Numerical experiments demonstrate the performance and effectiveness of the new method for reconstructing complex salt bodies.

**METHODOLOGY**

**The level-set method for inverse gradiometry problem**

Gravity gradiometry measures the gradient of each component of the gravity field on a prescribed acquisition surface. To simplify the presentation, we only consider inverting the $zz$-component of the gravity gradient, which can be modeled by

$$G_{zz}(r) = \gamma \int_{\Gamma_r} K_{zz}(r, r') \rho(r') dr', \quad r \in \Gamma_r,$$  

(1)

where $\gamma = 6.67408 \times 10^{-8}$ cm$^3$ g$^{-1}$ s$^{-2}$ is the universal gravitational constant, $\rho(r)$ is the density-contrast function with compact support below the acquisition surface $\Gamma_r$, $r'$ denotes the coordinates $(x, z)$ in 2D or $(x, y, z)$ in 3D, and the kernel $K_{zz}(r, r')$ is given by

$$K_{zz}(r, r') = \begin{cases} \frac{1}{|r-r'|} (2|z-z'|^2 - |r-r'|^2) & \text{in } \mathbb{R}^2, \\ \frac{1}{|r-r'|^3} (3|z-z'|^2 - |r-r'|^2) & \text{in } \mathbb{R}^3. \end{cases}$$  

(2)

Figure 1. Two infinite horizontal cylinders. (a) The true model consists of two infinite horizontal cylinders (the solid) separated with a distance of 0.5 km and the test model (the dotted line) is a horizontal shift of the true model and (b) the synthesized data generated by the true model.

Figure 2. Behavior of the misfit functions based on the $L_2$ (the solid), $L_1$ (the circle line), and KR norm with $c = 1 \times 2 \Delta x^{-1}$ (the dotted line), $c = 0.08 \times 2 \Delta x^{-1}$ (the cross “×” line), $c = 0.02 \times 2 \Delta x^{-1}$ (the star “∗” line), and $c = 0.01 \times 2 \Delta x^{-1}$ (the diamond “○” line). The shift of cylinders in terms of $s$ ranges from −0.6 to 1. As expected, the $L_1$ objective function matches with the KR norm objective function with a large parameter $c$ and the KR norm-based objective becomes convex with a decrease of the parameter $c$. Each objective is normalized by its maximum value for the purpose of display.
The inverse problem can be stated as follows: Given the measured \(zz\)-component of gravity-gradient data \(d\), find a density-contrast function \(\rho(r)\) such that the modeling data \(u = Gzz\) using equation 1 match the measured data \(d\). It is well-known that the inverse gradiometry problem is exponentially ill-posed (Isakov, 1990), partially due to the nonuniqueness of the solution of the inverse problem (Jackson, 1972). The popular minimum-structure inversion (Li, 2001a) of the gravity-gradient data constructs the density-contrast function by minimizing a data-misfit function that is regularized by some depth-dependent model regularization function, and this approach can be further improved by imposing bound constraints on the density-contrast function as a priori. By assuming that the density-contrast \(\rho\) is of a constant value on an unknown support, the inverse gradiometry problem (for more about uniqueness see Jackson, 1972) is exponentially ill-posed (Isakov, 1990), partially due to the nonuniqueness of the solution of the inverse problem (Jackson, 1972). The popular minimum-structure inversion (Li, 2001a) of the gravity-gradient data constructs the density-contrast function by minimizing a data-misfit function that is regularized by some depth-dependent model regularization function, and this approach can be further improved by imposing bound constraints on the density-contrast function as a priori. By assuming that the density-contrast \(\rho\) is of a constant value on an unknown support, the inverse gradiometry problem (for more about uniqueness results, see Isakov, 1990, p. 36). We remark that the assumption of a constant density as a priori information on the unknown domain is reasonable in some practical situations, and multiple density-contrast values on multiple unknown domains can also be assumed (for applying multiple level-set methods to multiple domains, see Li et al., 2011).

Following Isakov et al. (2011, 2013), we represent the density-contrast function by

\[ \rho(r) = H(\phi(r)) \Delta \rho(r), \quad r \in \Omega, \quad (3) \]

where \(\Delta \rho(r)\) is a given function and \(H(\cdot)\) is the Heaviside function: \(H(\cdot) = 1\) if \(\cdot > 0\), \(H(\cdot) = 0.5\) if \(\cdot = 0\), and \(H(\cdot) = 0\) if \(\cdot < 0\); \(\phi(r)\) is the level-set function defining \(\Omega_\gamma\): \(\phi(r) > 0\) if \(r \in \Omega_\gamma\), \(\phi(r) = 0\) if \(r \in \partial \Omega_\gamma\), and \(\phi(r) < 0\) if \(r \in \Omega \setminus \Omega_\gamma\). Therefore, the data modeling operator using the level-set function can be written as

\[ G_{zz}[\phi](r) = \gamma \int_{\Omega_\gamma} K_{zz}(r, r') H(\phi(r')) \Delta \rho(r') \, dr', \quad r \in \Gamma_\gamma. \quad (4) \]

Isakov et al. (2011) propose a least-squares functional as an objective function to minimize the misfit between \(u = Gzz[\phi]\) and the measured data \(d\), namely,

\[ J_{LS}[\phi] = \frac{1}{2} \int_{\Gamma_\gamma} |u(r') - d(r')|^2 \, ds(r'). \quad (5) \]

The gradient of the least-squares objective functional is (Isakov et al., 2011)

\[ \nabla_{\phi} J_{LS}[\phi](r) = \gamma \delta(\phi(r)) \int_{\Gamma_\gamma} K_{zz}(r', r) (u(r') - d(r')) \, ds(r'), \quad (6) \]

where \(\delta(\cdot)\) is the Dirac-delta generalized function and the data residual \(u - d\) serves as the adjoint source that is backpropagated into the domain for updating the level-set function. Here, the data residual \(u - d\) being nonzero means that the data are not completely fitted yet so that we will backpropagate this data residual (which is defined on the boundary) into the computational domain by using this residual as a source; therefore, it is also called the adjoint source in the literature (Leung and Qian, 2006; Plessix, 2006).

Figure 3. Sensitivity against noise. (a, c, and e) Comparisons of noisy data (the solid line), clean data (the dotted line), and calculated data (the dotted dot). (b, d, and f) Comparisons of adjoint sources for the \(L_2\) (the solid line) and KR norm-based objective functions with two different relaxation parameters (the dotted line for the smaller one and the dotted dot for the larger one). Note that the data residual is normalized by its maximum absolute value for the purpose of illustration.
We introduce an artificial time parameter $t$ to drive the following evolution equation to steady state:

$$\frac{\partial}{\partial t} \phi(r, t) = -\nabla_{\phi} J_{L_1}[\phi](r),$$

(7)

where a Neumann boundary condition $\partial \phi / \partial n = 0$ and an initial condition $\phi(r, 0) = \phi_0(r)$ are imposed (Isakov et al., 2011). Then, the minimizer $\phi^*(r)$ of the least-squares objective 5 is given by $\phi^*(r) = \lim_{\alpha \to \infty} \phi_\alpha(r, t)$ and the desired boundary $\partial Q_\rho$ of the support of the density-contrast function $\rho(r)$ is $\partial Q_\rho = \{r : \phi^*(r) = 0, r \in \Omega\}$. Please refer to Isakov et al. (2011, 2013), Lu and Qian (2015), and Li et al. (2016) for more details; we also include a detailed implementation of the level-set method in Appendix A.

Similarly, we can replace the least squares with the $L_1$-norm as the misfit function, i.e., $J_{L_1}[\phi] = \int_{r_t} |u(r') - d(r')| ds(r')$. Following the strategy proposed by Isakov et al. (2011), we have the evolution equation for the level-set function:

$$\frac{\partial}{\partial t} \phi(r, t) = -\nabla_{\phi} J_{L_1}[\phi](r),$$

(8)

where the gradient $\nabla_{\phi} J_{L_1}[\phi](r)$ is given by

$$\nabla_{\phi} J_{L_1}[\phi](r) = \gamma \delta(\phi(r)) \Delta \rho(r) \times \int_{\Gamma_r} K_{zz}(r', r) \left( \frac{u(r') - d(r')}{|u(r') - d(r')|} \right) ds(r').$$

(9)

In our numerical examples, we replace the denominator $|u(r') - d(r')|$ in the above gradient with $\sqrt{|u(r') - d(r')|^2 + \epsilon^2}$ so as to avoid division by zero, where $\epsilon$ is a small positive parameter.

**KR norm-based misfit for inverse gradiometry**

Due to the nonlinear evolution of the level-set method, the level-set parameterization of the density-contrast function may induce

Figure 4. The 2D salt body. (a) A 2D slice (AA'-profile) of SEG/EAGE salt model, (b) the support of the salt body (the solid) overlapped with an ellipse (the initial guess, dotted line), and (c) comparison of recorded data (the solid) and simulated data (the dotted line) by initial guess.

Figure 5. The 2D salt body (5% noisy data with the maximum amplitude of 3.85 E). Comparisons of inverted results by (a) the $L_2$-norm, (b) the $L_1$-norm, and (c) the KR norm-based objective at the 20,000th iteration, (d) Comparison of the noisy data (the solid) and calculated data generated using inverted results by three objective functions (“-.” for $L_2$, “.”” for $L_1$, and “- -” for KR objectives).
nonconvexity in the least-squares objective 5 and the $L_1$-norm-based objective 8 so that minimizing these objective functions may potentially suffer from local minima; hence, they are sensitive to the initial guess. Therefore, designing a convex objective function that can mitigate local-minimum issues is quite important for successful applications of these level-set approaches. Compared with the classic $L_2$ and $L_1$ metric, an optimal-transport metric can easily resolve the spatial differences, such as shift, rotation, and dilation, between two probability distributions (Villani, 2003; Engquist et al., 2016), so that it is appealing to use this norm as a misfit function for better guiding the evolution of a level-set function. The difficulties for directly applying the optimal-transport metric to the inverse gradiometry problem are nonpositivity of the underlying functions and the lack of mass conservation between the modeled data and measured data. However, the KR norm as the dual formulation of an optimal-transport metric with the $L_1$ ground cost relaxes those two requirements and can be computed by solving a linear programming problem, so that it is attractive to use this norm in the inverse problem (Metivier et al., 2016).

Therefore, we first apply the KR norm as a discrepancy criterion to measure the misfit for inverting the 2D gravity-gradient data in that the corresponding data can be parameterized by a single variable on the real line $\mathbb{R}$, i.e.,

$$J_{\text{KR}}[\phi] = ||u - d||_{\text{KR}} \equiv \max_{v \in \text{BLip}_c} \int (u - d) \cdot v ds(r'),$$  \hspace{1cm} (10)$$

where $u = u(r)|_{\Gamma} = \gamma \int_{\Omega} K_{zz}(r', r') H(\phi(r')) \Delta \rho(r') dr'$ is the modeling data restricted to the given measurement surface $\Gamma \subset \mathbb{R}$ and the admissible Lipschitz set $\text{BLip}_c$ is defined by

$$\text{BLip}_c = \{ v(x) : |v(x)| \leq 1 \text{ and } |\partial x v(x)| \leq c, x \in \mathbb{R} \}. \hspace{1cm} (11)$$

We remark that the bound constraint on the function $v$ in the above admissible set can be viewed as a special case of the bound constraint introduced by Lellmann et al. (2014). Because the computation of the KR norm defined in equation 10 is equivalent to a linear programming problem with the constraint set 11, the KR norm can be efficiently computed by the alternating direction method of multipliers (ADMM) algorithm as presented in Appendix B (for a similar algorithm, see also Metivier et al., 2016).

We now present an example to demonstrate that the KR misfit induces convexity with respect to the horizontal shift of a mass anomaly and can effectively mitigate local minima in the level-set formulation of the inverse gradiometry problem in terms of the

Figure 6. The 2D salt body (25% noisy data with the maximum amplitude of 20.62 E). Comparisons of inverted results by (a) the $L_2$-norm at the 2100th iteration ($\chi_{L_2} = 0.572$), (b) the $L_1$-norm at the 3300th iteration ($\chi_{L_1} = 0.587$), and (c) the KR norm-based objective at the 11,000th iteration ($\chi_{KR} = 0.572$). (d) Comparison of the noisy data and calculated data generated using inverted results by three objective functions.

Figure 7. The 2D salt body (variable density). (a) A salt model with variable density decreasing with depth in the salt body, (b) The observed data contaminated by 5% random noise with the maximum amplitude of 3.79 E.
relaxation parameter $c$. Let $\phi(x, z) = \max(\psi_1, \psi_2)$ for $(x, z) \in [-1, 2] \times [0, 1]$ km be the level-set function for the true model as plotted in Figure 1a, in which the two level-set functions are $\psi_1(x, z) = 0.1 - \sqrt{(x-0.2)^2 + (z-0.3)^2}$ and $\psi_2(x, z) = 0.1 - \sqrt{(x-0.8)^2 + (z-0.3)^2}$. Letting the density-contrast function be $\Delta \rho = 0.1 \text{ g/cm}^3$, we generate the $G_{zz}$ data with the level-set function $\phi$ via the integral equation 4, which is the measured (synthetic) data $d$ as shown in Figure 1b.

To compare the behaviors of the $L_2, L_1$, and KR norm-based objectives, we use a shifted version of the true model as trial models that are parameterized by $s$ in terms of the level-set functions $\phi_s(x, z) = \max(\psi_{1,s}, \psi_{2,s})$, where $\psi_{1,s}$ and $\psi_{2,s}$ are given by

$$\psi_{1,s}(x, z) = 0.1 - \sqrt{(x-s)^2 + (z-0.3)^2}, \quad (12)$$

$$\psi_{2,s}(x, z) = 0.1 - \sqrt{(x-(s+0.6))^2 + (z-0.3)^2}, \quad (13)$$

where the shift parameter $s$ ranges from $s = -0.6$ to 1.

For each shift $s$, we can compute the $G_{zz}$ data with the level-set function $\phi_s$ by equation 4 again. Hence, the objective functions using the $L_2, L_1$, and KR norm only depend on the parameter $s$, and they are plotted in Figure 2. We can see that there are two local

Figure 8. The 2D salt body (variable density). Comparisons of inverted results with the density function (a, c, and e) $\Delta \rho = 0.2 \text{ g/cm}^3$ and (b, d, and f) $\Delta \rho = 0.24 \text{ g/cm}^3$ by the $L_2, L_1$, and KR norm-based objectives. Comparison of the measured data (the solid) and calculated data ("-" for $L_2$, "-" for $L_1$, and "-" for KR norm) generated using the inverted results by three objectives with the density function (g) $\Delta \rho = 0.2 \text{ g/cm}^3$ and (h) $\Delta \rho = 0.24 \text{ g/cm}^3$. 

Downloaded 09/06/19 to 172.5.43.192. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/
minima present in the least-squares and the $L_1$-norm-based objective functions, and the latter highly coincides with the KR norm-based objective function with a large relaxation parameter $c = 2\Delta x^{-1} (\Delta x = 0.01 \text{ km})$. The asymptotic relation between the KR norm and the $L_1$-norm for the large parameter $c$ is shown in Appendix C. The KR norm-based objective function becomes convex when the relaxation parameter is reduced to $c = 0.01\Delta x^{-1}$, and it only has a single minimum that is actually the global minimum. This simple example indicates that the KR norm-based objective function with a small relaxation parameter mitigates the local minima of the inverse gradiometry problem and tends to be convex so that it is amenable to a local gradient-based optimization method.

To illustrate the robustness of the KR norm-based nonlinear misfit functional $10$ with respect to $\phi$, analogous to the least-squares formulation $5$, we need to compute the gradient of the KR-based objective function $10$. Because the perturbation of $J_{\text{KR}}[\phi]$ is given by

$$\delta J_{\text{KR}}[\phi] = \langle \delta u, v^* \rangle_{\Gamma_r},$$

where $v^*$ is the maximizer of the right side in the objective function $10$ for a given $\phi$, we have

$$\nabla_{\phi} J_{\text{KR}}[\phi](\mathbf{r}) = \gamma \delta(\phi(\mathbf{r})) \rho(\mathbf{r}) K_{\text{zz}}(\mathbf{r}', \mathbf{r}) v^*(\mathbf{r}') ds(\mathbf{r}').$$

Comparing with the gradient in equation $6$ of the least-squares-based objective function, we realize that we only need to replace the adjoint source $u - d$ by the maximizer $v^*$ in the gradient $6$ to obtain the KR norm-based gradient. We further solve the following level-set evolution equation to steady state:

$$\frac{\partial}{\partial t} \phi(\mathbf{r}, t) = -\nabla_{\phi} J_{\text{KR}}[\phi](\mathbf{r}).$$

The KR norm-based objective function: Robustness to noise

To illustrate the robustness of the KR norm-based objective function against random noise, we compare the adjoint sources for the $L_2$ objective function and the KR norm-based objectives with small and large relaxation parameters $c$. The recorded data are the same as in the previous example consisting of two circles, and the modeling data are simulated with the level-set function $\phi_{\gamma \rightarrow 0.6}(x, z)$, as shown in Figure 3a. The adjoint sources with clean data for these three objective functions are plotted in Figure 3b, from which we can see that the adjoint source for the KR objective with a large parameter $c$ is the signum function as expected, whereas the adjoint source for a small parameter $c$ behaves like a smoothed version of the signum function.

In comparison with the objective function using clean data in Figure 3a and 3b, Figure 3c–3f indicates that the adjoint source for the KR norm-based objective with relaxation parameter $c = 0.01 \times 2(\Delta x)^{-1}$ is quite stable and does not change too much even if the data are contaminated with 10% or 15% random noise. On the other hand, for a large relaxation parameter such as $c = 2 \times (\Delta x)^{-1}$, the adjoint source of the KR norm-based objective is quite sensitive to random noise as the signs of the adjoint source vary a lot in comparison with the clean-data case, which may lead to wrong directions for updating the level-set function. This confirms that the KR norm-based misfit function with a small relaxation parameter is robust against random noise in data. In fact, we can actually show that the KR norm of Gaussian random noise decays of order $N^{-1/2}$ ($N$ is the total number of the measurement); hence, it is negligible so long as the sampling of the measurement is sufficiently dense (for a detailed proof, see Appendix D).

Extension to 3D inverse gradiometry problems

The KR norm-based objective function for the 3D inverse gradiometry problem reads as follows:

$$J_{\text{KR}}^{3D}[\phi] = \sup_{\mathbf{r} \in \text{BLip}^{3D}} \int_{\Gamma_r} (u - d) \cdot v ds(\mathbf{r}'),$$

where $u(x, y)$ and $d(x, y)$ are, respectively, the simulated and measured $zz$-component of the gravity-gradient data on the surface $\Gamma_r$, and the set $\text{BLip}^{3D}$ is given by

![Figure 9. Two balls. (a) The target source model is of grid size 1/32 km for generating the synthetic $G_{zz}$ data, (b) an ellipse as an initial level-set function, and (c) the noisy $G_{zz}$ data contaminated by 5% random noise with the maximum amplitude of 4.33 E.](image-url)
BLip_3D = \{ v(x, y) : |v(x, y)| \leq 1, |∂_x v(x, y)| \leq c, \\
and |∂_y v(x, y)| \leq c, x, y \in \mathbb{R} \},

which is a generalization of the 1D bounded Lipschitz set BLip_c in 3D.

The objective function 17 can be viewed as a direct generalization of the KR norm-based objective function for 2D problems, and it can be evaluated by the ADMM algorithm as well.

The first-order perturbation of the KR norm-based objective function \( J_{KR} \), with respect to \( u \) is given by

\[
\frac{∂}{∂u} J_{KR}[\phi] = v^*(x, y),
\]

where \( v^*(x, y) = \arg\max_{v \in \mathbb{R}^3} \{ v \cdot (r - d) \} \) and \( Av(x, y) = (v, c^{-1}∂_x v, c^{-1}∂_y v)^T \). Analogous to the 2D inverse gradiometry problem, \( v^*(x, y) \) is an adjoint source replacing the data residual \( u(x, y) - d(x, y) \) in the least-squares-based gradient equation 6.

**NUMERICAL EXPERIMENTS**

In all of the examples reported below, we assume that the measured data are only partially available on the surface. We apply the gradient-descent method to minimize objective functions by solving corresponding level-set equations, in which we set the Courant-Friedrichs-Lewy (CFL) constant as CFL = 0.8 in 2D examples and CFL = 0.6 in 3D examples for evolving level-set functions. To generate noisy data, we add the Gaussian random noise with mean zero and standard deviation \( \sigma = \max_{j = 1, \ldots, N} |d_j| \) to the clean measurement data \( d_j \), where \( \mu \) is the noise level used in this paper. To avoid over-fitting the data, we use a scaled \( \chi^2 \)-squared misfit as a criterion to terminate inversion (Li et al., 2017), where the scaled \( \chi^2 \)-squared misfit is computed by

\[
\chi^2 = \frac{1}{2N\sigma^2} \sum_{j=1}^{N} (G_{zz}(r_j) - d_{\text{noise}}(r_j))^2,
\]

where \( G_{zz}(r_j) (j = 1, 2, \ldots, N) \) are the simulated data sampled at receivers. The scaled \( \chi^2 \)-squared misfit is dimensionless with expectation 0.5, which is used as a stopping criterion by Li et al. (2017). However, here we assume that the scaled \( \chi^2 \)-squared misfit for the true model is exactly obtained for one random experiment, i.e.,

\[
\chi^2 = \frac{1}{2N\sigma^2} \sum_{j=1}^{N} (d(r_j) - d_{\text{noise}}(r_j))^2,
\]

which quantifies the strength of noise in a single random experiment. The motivation for this assumption is that we can only have one measurement at each receiver, which corresponds to a single random experiment and, thus, can be estimated in practice.

**2D examples**

*Example 1: 2D salt structure*

The target salt model as shown in Figure 4a is the 2D slice (A-A′ profile) of the SEG/EAGE salt model with dimensions of 13.6 km in distance and 4 km in depth (Aminzadeh et al., 1997). Although the original salt model is designed for testing seismic imaging methods, we capture the major salt structure shown in Figure 4b to test our method for inverse gradiometry problems. We assume that the density-contrast function is constant in the salt body as mentioned before, i.e., \( \Delta \rho = 0.2 \text{ g cm}^{-3} \). The computational domain is \( \Omega = [0, 13.6] \times [0, 4] \), which is uniformly discretized into 341 × 101 mesh points with a spacing of \( \Delta x = \Delta z = 0.04 \) km. A total of 341 measurements of the gravity-gradient \( G_{zz} \) are collected on the acquisition line \( \Gamma_r = \{ (x, z) : x \in [0, 13.6], z = -0.1 \} \) with a spacing of 0.04 km. The initial level-set function in Figure 4b (dotted line) is given by a circle, \( \phi_0(r) = 1 - \sqrt{(x-4)^2 + (z-2)^2} \).

*Case 1: Slightly noisy data*

To start with, we investigate the stability of the KR-based objective against random noise by

\[
\chi^2 = \frac{1}{2N\sigma^2} \sum_{j=1}^{N} (d(r_j) - d_{\text{noise}}(r_j))^2,
\]
Example 2: 2D salt body with variable density

We reconstruct salt models with density contrast varying in the depth direction. The reference salt model is the same as the previous example, but the target density-contrast function is \( \Delta \rho(r) = 0.2 + 0.04 \times (1.8 - z) \) g \( \cdot \) cm\(^{-3} \) inside the salt body, which decreases in depth as shown in Figure 7a. Figure 7b shows the measured data contaminated with 5% Gaussian random noise with the maximum amplitude of 3.79 E.

To investigate the robustness of the level-set method with respect to the density function, we compare inversion results by taking two different constant values of 0.2 and 0.24 g \( \cdot \) cm\(^{-3} \) (20% higher) in the unknown domain represented by the level-set function. These choices are apparently not consistent with the original measurement data generated by the variable density-contrast function.

We choose the relaxation parameter \( c = 0.002 \times (2/\Delta x) = 0.1 \) and the penalty parameter \( \alpha = 1 \), and we evaluate the KR norm using 400 ADMM iterations for both cases. The inversion is stopped when the \( \chi^2 \)-squared data misfit is close to the given \( \chi^2 \) misfit (\( \chi^2 = 0.483 \)) for all the methods.

For the first case with the density-contrast \( \Delta \rho = 0.2 \) g \( \cdot \) cm\(^{-3} \), Figure 8a, 8c, and 8e shows, respectively, the inverted results by the \( L_2 \)-norm-based objective at the 4750th iteration, the \( L_1 \)-norm-based objective at the 5400th iteration, and the KR norm-based objective at the 9700th iteration; the corresponding scaled \( \chi^2 \)-squared misfits are \( \chi_{L_2} = 0.481, \chi_{L_1} = 0.480, \) and \( \chi_{KR} = 0.511 \), respectively.

For the second case with the density-contrast \( \Delta \rho = 0.24 \) g \( \cdot \) cm\(^{-3} \), Figure 8b, 8d, and 8f shows, respectively, the inverted results by the \( L_2 \)-norm-based objective at the 15,100th iteration, the \( L_1 \)-norm-based objective at the 9100th iteration, and the KR norm-based objective at the 20,000th iteration; the corresponding scaled \( \chi^2 \)-squared misfits are almost equal: \( \chi_{L_2} = 0.492, \chi_{L_1} = 0.483, \) and \( \chi_{KR} = 0.498 \), respectively.

As we can see, the final reconstructed zero level-set functions from the KR norm-based objective match with the target salt boundary very well. Similar to the constant density-contrast example, we can observe that the top, base, and also the dipping flank of the salt body are well-recovered. Hence, we can conclude that the KR objective function is quite robust so long as the constant density
function is close to the target density function. Although the resi-
mulated data from the inverted results by the $L_2$ and $L_1$ objectives
are also consistent with the measured data as shown in Figure 8g
and 8h, the overall results indicate that the inverted result by the KR
objective converges to the target model in a more stable manner.

3D examples

Example 3: 3D two spheres

In this example, we assume that the target source model consists
of two spheres centered at $(0.6, 1, 0.25)$ and $(1.4, 1, 0.25)$ km with
the same radius of 0.15 km as shown in Figure 9a. The density-
contrast function is a constant of $0.2 \text{ g cm}^{-3}$ in the source bodies.

The computational domain $\Omega = [0, 2] \times [0, 2] \times [0, 0.5]$ km is dis-
cretized with the grid size of 1/32 km on each direction. A total
of $41 \times 41 = 1681$ synthesized data points are generated uniformly
on the surface $\Gamma_r = \{(x, y, z) : x \in [0, 2], y \in [0, 2], z = 0\}$ with a
spacing of 0.05 km in both directions. The measured data in
Figure 9c are contaminated by 5% Gaussian random noise with
maximum amplitude of 4.33 E.

We choose the relaxation parameter $c = 2$ and the penalty param-
eter $\alpha = 10$. We evaluate the KR norm with a maximum iteration of
200 and terminate the inversion when the $\chi$-squared data misfit is
close to the given $\chi$-squared misfit ($\chi_t = 0.485$) for all three meth-
ods. To illustrate the nonlinearity of the objective with respect to the
level-set function and possible local minima, we choose the initial
guess as a sphere centered at $(0.6, 1, 0.25)$ km with a radius of
0.15 km as shown in Figure 9b, which is the same as the first sphere.

Figure 12. Two balls (5% noisy data). Comparisons of (a, c, and e) the adjoint source and (b, d, and f) the 2D cross section of the gradient
(normalized by its maximum absolute value and overlapped with the target source model) at $y = 1$ km by the initial guess for the $L_2$, $L_1$, and
KR objectives. The marked square at $(0.75,0.25)$ km indicates the direction of the gradient of three objectives using the initial guess for
updating the level-set function.
of the target source model. The inverted results by the $L_2$ and $L_1$ objectives at the 3000th iteration and KR objective at the 650th iteration are plotted in Figure 10a, 10c, and 10e, respectively. The corresponding data residuals are displayed in Figure 10b, 10d, and 10f, which have the $\chi$-squared misfits of $\chi_{L_2} = 2.31, \chi_{L_1} = 2.32,$ and $\chi_{KR} = 0.486$, respectively. We can clearly see that the inversions based on the $L_2$ and $L_1$ objectives do not converge to the target model, which may be stuck in a local minimum. However, the inverted result by the KR objective converges successfully to the targeted one even if the initial guess is not close to the true model. This can also be confirmed by comparing the 2D cross sections at $y = 1$ km of the inverted results by these three objectives as shown in Figure 11a–11c.

To further understand why the $L_2$ and $L_1$ objectives are not able to update the level-set function in the correct direction, we compare the adjoint sources and gradients of the three objectives at the initial guess. Figure 12a displays the adjoint source of the $L_2$ objective with the initial guess, which is migrated into the domain by the adjoint operator $K_{zz}^T$. We can clearly see that the migrated image in Figure 12b is mainly concentrated below the adjoint source as expected and that the gradient of the $L_2$ objective in equation 6 defined by the restriction of the migrated image on the reference zero level-set function is, however, positive as marked in Figure 12b. This indicates that the level-set function may be updated in the negative direction after solving the linear programming problems 10 and 11.

$\phi_0(r) = 1 - \sqrt{ \frac{(x-7.8)^2}{2^2} + \frac{(y-5.2)^2}{2^2} + \frac{(z-2)^2}{1^2} }, \quad r \in \Omega, \quad (22)$

which is used to generate the simulated data as shown in Figure 13d that have more or less the same structure as the measured data.

For the inversion, we use a grid of 0.1 km in each direction to discretize the computational domain $\Omega$, which yields $137 \times 137 \times 41 = 769,529$ mesh points. The initial level-set function as plotted in Figure 13b is an ellipse given by

$$0.499 \times (x-7.8)^2 + (y-5.2)^2 + (z-2)^2 \leq 1,$$

$\Omega = \{ (x,y,z) : x \in [-4,18] \text{ km}, y \in [-4,18] \text{ km}, z = -0.1 \text{ km} \}$.

We consider to reconstruct the 3D SEG/EAGE salt structure (Aminzadeh et al., 1997). We extract the support of the salt body as shown in Figure 13a. The grid spacing is 0.04 km in each direction. The density contrast is 0.2 g · cm$^{-3}$. The computational domain $\Omega = [0,13.6] \times [0,13.6] \times [0,4] \text{ km}$ is discretized into $341 \times 341 \times 101 = 11,744,381$ mesh points. This grid is used to generate the measured data as plotted in Figure 13c. There are $276 \times 276 = 76,176$ data points in total with a spacing of 0.08 km in each direction, which are placed on the surface $\Gamma_r = \{ (x,y,z) : x \in [-4,18] \text{ km}, y \in [-4,18] \text{ km}, z = -0.1 \text{ km} \}$.

The inverted results are displayed in Figure 14a for the $L_2$ inversion at the 1800th iteration, in Figure 14c for the $L_1$ inversion at the 2600th iteration, and in Figure 14e for the KR inversion at the
2630th iteration, respectively, which roughly give the same scaled χ-squared misfit: \( \chi_{L_2} = 0.500 \), \( \chi_{L_1} = 0.500 \), and \( \chi_{KR} = 0.500 \), respectively. The evolution history of the objective function in Figure 15 indicates that the \( L_2 \) and \( L_1 \) objective functions are indeed convergent without further reduction of the objective value. The decrease of the KR objective value is more significant, although the oscillation is still present as is often occurring in the level-set inversion (Li et al., 2016). The corresponding data residuals are plotted in Figure 14b, 14d, and 14f for the three objectives, which do not show much structured information. These inversion results indicate that the \( L_2 \)- and \( L_1 \)-norm-based objectives are much more sensitive to random noise than the KR norm-based objective: The former two objectives are only able to image a portion of the top salt, whereas the KR norm-based objective function is able to recover crucial information of the dipping flank and the base of the salt. The 2D cross sections at \( x = 4 \) and 5.4 km in Figure 16a–16f.

Figure 14. The 3D SEG salt body (8% noisy data). Comparisons of the inverted results by the (a) \( L_2 \), (c) \( L_1 \), and (e) KR norm-based objective functions at the 2600th, the 3040th, and the 2090th iteration, respectively; comparisons of data residual using the (b) \( L_2 \), (d) \( L_1 \), and (f) KR norm inverted results, which have the χ-squared misfits of \( \chi_{L_2} = 0.500 \), \( \chi_{L_1} = 0.500 \), and \( \chi_{KR} = 0.500 \), respectively. The target χ-squared misfit is \( \chi_t = 0.499 \).
clearly demonstrate the advantages of the KR norm-based objective. Because the solution of the domain inverse problem investigated here is unique, it will correspond to the steady state of the level-set evolution equation 7. Although the inverted results produced by all three methods predict almost the same data fitting to the observed data, the overall results indicate that the result inverted by the KR norm-based objective is more reliable because the solution is more stably converging to the target model after a long time evolution.

To understand why the KR norm-based objective function is robust against random noise, we compare the adjoint sources of the three objective functions for the initial guess, which are plotted in Figure 17a–17c. We can observe that the $L_2$ adjoint source with noisy data is significantly impacted by random noise, and the same can be said about the $L_1$ adjoint source, but the KR adjoint source shows that the oscillations caused by random noise are smoothed out, which produces a much cleaner adjoint source for inversion.

**Example 5: 3D salt body (gravity inversion)**

We consider inverting gravity rather than gravity-gradient data to show that our new methodology is equally applicable. We use the same 3D SEG salt model to simulate gravity data (the $G_y$ component). The acquisition parameters are taken to be the same as the previous example. We also add $\mu = 8\%$ Gaussian random noise with the maximum amplitude of 1.79 mGal to the clean synthesized data to obtain noisy data as shown in Figure 18a, in which the scaled $\chi^2$-squared misfit is $\chi^2 = 0.499$. We choose the relaxation parameter $c = 10$ and the penalty parameter $\alpha = 1$. The KR norm is evaluated with the maximum iteration of 200 to guarantee the convergence of the ADMM algorithm. We use the same criterion as the previous example to terminate the inversions.

The inverted results are, respectively, plotted in Figure 19a for the $L_2$ inversion at the 550th iteration, in Figure 19c for the $L_1$ inversion at the 900th iteration, and in Figure 19e for the KR inversion at the 1500th iteration, which also give the same scaled $\chi^2$-squared misfit and produce the data residuals in Figure 18b–18d. As we can see, the KR norm-based objective function is able to recover crucial
information of the dipping flank and the base of the salt compared with the inverted results by the $L_2$ and $L_1$ objective functions. From the 2D cross sections at $x = 5.4$ km in Figure 19b, 19d, and 19f, we can also confirm the advantages of the KR norm-based objective when inverting the gravity data. Hence, the KR norm-based objective is effective not only for inverting gravity gradient but also for inverting gravity data.

Figure 17. The 3D SEG salt body (8% noisy data). Comparisons of the adjoint sources using the initial guess for the (a) $L_2$, (b) $L_1$, and (c) KR norm-based objectives.

Figure 18. The 3D SEG salt body (gravity data $G_z$). (a) The measured $G_z$ data contaminated by 8% random Gaussian noise with the maximum amplitude of 1.79 mGal and (b, c, and d) comparisons of the data residual generated by the inverted results using $L_2$, $L_1$, and KR norm-based objective functions.
DISCUSSION

By formulating the inverse gravity problem as a domain inverse problem, under reasonable assumptions such as an anomalous body being convex or \( z \)-convex (Isakov, 1990; Isakov et al., 2011), as shown by Isakov (1990, p. 36), the resulting domain inverse problem converges to the unique solution in terms of reconstructing the anomalous body. This unique solution has logarithmic stability in the sense that if the measurement data have accuracy \( \varepsilon \), then the inverted solution has accuracy of the order \( (\log(1/\varepsilon))^{-\beta} \) for some \( \beta > 0 \); this stability result is implied by Isakov (1990 p. 39). The level-set method is used here as a tool to parameterize the domain that we are interested in reconstructing, and one may use other approaches to parameterize the domain as well. Therefore, the level-set method itself does not change or introduce stability into the domain inverse problem.

The work by Isakov et al. (2011) is the first one that applies the level-set method to solve potential-field inverse problems; this method is shown to be an effective imaging tool when applied to field data (Li et al., 2017). It is well-known that the least-squares objective requires a good initial guess for the level-set-based method to be convergent to the desired solution (Dorn and Lesselier, 2006). The KR norm-based objective function investigated here can mitigate local minima that may arise in a least-squares formulation so that it can be more suitable for a local gradient-based optimization method as confirmed in our numerical examples. From the practical point of view, the KR norm-based objective function is quite stable and insensitive to random noise in measurements.

In terms of computational complexity, we will briefly discuss the 2D case only. The major computational cost for computing the KR norm is due to solving a tridiagonal matrix at each ADMM iteration, and this tridiagonal matrix can be solved in \( O(N) \) operations. Our numerical experiences indicate that it usually requires several dozens of iterations for ADMM to converge at each update step. Comparing with the \( L_2 \) and \( L_1 \)-norm-based objectives, the computational cost of the KR norm-based objective only increases a little bit, but this extra cost pays as shown in numerical experiments. Analogous analysis holds for 3D cases as well.

Our theory shows that for a large relaxation parameter the KR norm tends to the \( L_1 \)-norm, whereas for a small relaxation parameter the KR norm behaves like the \( L_1 \)-norm of an indefinite integral of the underlying function. This criterion allows us to differentiate the KR norm from the \( L_1 \)-norm, but it does not guide us on how to choose a small relaxation parameter for each model if the datum is contaminated with random noise, which is crucial for the KR inversion. This issue will be pursued in our future work.

CONCLUSION

We propose to use a novel KR norm-based misfit function for inverting gravity-gradient data by a level-set formulation of the inverse gradiometry problem. In comparison to the classic \( L_2 \) and \( L_1 \) metric, the KR norm leads to evolution of the level-set function by resolving spatial differences between modeled and measured data. Our numerical examples show that the KR norm-based objective

Figure 19. The 3D SEG salt body (8% noisy gravity data \( G_z \)). Comparisons of the inverted results by the (a) \( L_2 \), (c) \( L_1 \), and (e) KR norm-based objective functions at the 550th, the 900th, and the 1500th iterations, respectively; comparisons of 2D cross sections of the inverted results at (b, d, and f) \( x = 5.4 \ km \) by the \( L_2, L_1 \), and KR objectives.
function with a proper choice of relaxation parameter is able to provide a stable and robust reconstruction even if starting with a poor initial guess when the measured data are contaminated with strong random noise. The reconstructions of 2D and 3D SEG/EAGE complex salt structure further demonstrate the capabilities and potential of the KR norm-based misfit and the level-set formulation for large-scale practical inverse gradiometry problems.

ACKNOWLEDGMENTS

We would like to acknowledge Michigan State University HPCC for providing the computational resources that have contributed the research results reported here. Zhang’s research is supported by the Natural Science Foundation of Guangdong Province of China. Qian’s research is partially supported by the National Science Foundation of USA.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

APPENDIX A

THE LEVEL-SET METHOD FOR INVERSE GRADIOMETRY PROBLEM

We summarize the level-set method for the inverse gradiometry problem. To avoid numerical instabilities, the Heaviside function is approximated by the $\varepsilon$-Heaviside function as

$$H_{\varepsilon}(\phi) = \begin{cases} 0, & \text{if } \phi < -\varepsilon, \\ \frac{1}{2} + \frac{1}{2\varepsilon} \sin \left( \frac{\pi \phi}{\varepsilon} \right), & \text{if } |\phi| \leq \varepsilon, \\ 1, & \text{if } \phi > \varepsilon. \end{cases} \quad (A-1)$$

Algorithm 1. The level-set method for the gravity-gradient data.

Step 1. Given the measured gravity gradient ($\Delta \rho$) data $d$, an a priori constant-density $\Delta \rho$, the data misfit level $\varepsilon$ and the maximum evolution step $N_{\text{max}}$, choose an initial level-set function $\phi_0(r)$, the CFL constant $cfl$, and set the evolution step $j = 0$.

Step 2. Generate the gravity gradient data $u_j = G_{zz}[\phi_j]$ via the integral equation (4).

Step 3. Compute the data residual $r_j = u_j - d$ as the adjoint source.

Step 4. Back-propagate the adjoint source $r_j$ into the domain to find the gradient flow $g_j = \nabla J_{LS}[\phi_j]$ via the equation (6), and update the temporary level-set function by

$$\bar{\phi}_{j+1} = \phi_j - \Delta t g_j, \quad (A-2)$$

where the step size $\Delta t$ is given by $\Delta t = cfl \min \left[ \frac{\Delta r}{\max |\phi_j|} \right]$ and $V_j$ satisfies the relation $g_j = V_j / |\nabla \phi_j|$.

Step 5. Reinitialize the level-set function $\phi_{j+1}$ to maintain the signed distance property by solving the following equation:

$$\frac{\partial \Phi}{\partial \xi} + \text{sign}(\phi_{j+1})(|\nabla \Phi| - 1) = 0, \quad \frac{\partial \Phi}{\partial n} \bigg|_{\partial \Omega} = 0, \quad (A-3)$$

starting with the initial condition $\Phi(r, \xi = 0) = \phi_{j+1}(r)$.

Step 6. Set the updated level-set function as $\phi_{j+1}(r) = \Phi^\infty(r)$.

Step 7. If $j+1 < N_{\text{max}}$ and the data misfit level is above $\varepsilon$, continue to Step 2. Otherwise, output the final level-set function $\phi_{j+1}(r)$.

The delta function $\delta(\phi)$ is approximated by $\delta(\phi) = \chi_T |\nabla \phi|$, where $\chi_T$ denotes the support of the set $T_r = \{ r \in \Omega : |\phi(r)\xi| \leq \varepsilon \}$.

We remark that the reinitialization of the level-set function $\bar{\phi}_{j+1}$ to maintain the signed-distance property is necessary, which serves as a kind of regularization on the model space. The function $\Phi^\infty(r)$ is the steady-state solution of the evolution equation given in step 5, which is, however, approximated by several $\Delta \xi$ steps because we are only interested in the solution near the zero level set (for further discussion, see Isakov et al., 2011).

APPENDIX B

COMPUTATION OF THE KR NORM $\| \cdot \|_{KR}$

To compute the KR norm, we need to solve the linear programming problem (10) with the inequality constraint (11). We first introduce an indicator function $t_c$ of the bounded set $C = \{ w = (w_1, w_2)^T : |w_1(x)| \leq 1 \text{ and } |w_2(x)| \leq 1, x \in \mathbb{R} \}$,

$$t_c(w) = \begin{cases} 0 & \text{if } w \in C, \\ +\infty & \text{if } w \notin C. \end{cases} \quad (B-1)$$

Then, computing the KR norm of a single-variable function $r \in L_1(\mathbb{R})$ is equivalent to solving the following nonconstrained minimization problem:

$$\min_v \{ -\langle r, v \rangle + t_c(Av) \}, \quad (B-2)$$

where the operator $A$ is defined as $Av(x) = (v, c^{-1} \partial_x v)'$.

Introducing the auxiliary variable $w = Av$, the augmented Lagrangian method for problem B-2 reads (Glowinski, 2015)

$$L(v, w; \lambda) = -\langle r, v \rangle + t_c(w) + \frac{1}{\alpha} \langle \lambda, Av - w \rangle + \frac{1}{2\alpha} \| Av - w \|_{L_2}^2, \quad (B-3)$$

where $\alpha > 0$ is an augmentation parameter, $\lambda$ is a Lagrange multiplier, and $\langle \cdot, \cdot \rangle$ denotes the $L_2$ inner product.
Applying the ADMMs to the augmented Lagrangian B-3 yields

\[
\begin{align*}
    v_{k+1} &= \arg\min_v L(v, w_k; \lambda_k), \\
    w_{k+1} &= \arg\min_w L(v_{k+1}, w; \lambda_k), \\
    \lambda_{k+1} &= \lambda_k + Av_{k+1} - w_{k+1},
\end{align*}
\]  

where the notation argmin refers to the minimizer that attains the minimum of its corresponding problem.

By equation B-3, we have

\[
\begin{align*}
v_{k+1} &= \arg\min_v -\langle r, v \rangle + \frac{1}{\alpha} \langle \lambda_k, Av - w_k \rangle + \frac{1}{2\alpha} \|Av - w_k\|_2^2 \\
    &= \arg\min_v \frac{1}{2\alpha} \|Av - w_k + \lambda_k\|_2^2 - \langle r, v \rangle \\
    &= (A^*A)^{-1}(A^*(w_k - \lambda_k) + ar)
\end{align*}
\]  

and

\[
\begin{align*}
w_{k+1} &= \arg\min_w t_C(w) + \frac{1}{\alpha} \langle \lambda_k, Av_{k+1} - w \rangle + \frac{1}{2\alpha} \|Av_{k+1} - w\|_2^2 \\
    &= \arg\min_w \frac{1}{2\alpha} \|Av_{k+1} - w + \lambda_k\|_2^2 \\
    &= \max(-1, \min(Av_{k+1} + \lambda_k, 1)),
\end{align*}
\]

where the adjoint operator $A^*\lambda = \lambda_1 - c^{-1}\partial_1\lambda_2^2$ with $\lambda$ being a vector function with two components $\lambda_1$ and $\lambda_2$, and $A^*A = I - c^{-2}\partial_\alpha$ is the Laplacian operator with a homogeneous Neumann boundary condition, i.e., $\partial_\alpha v_{k+1} = 0$ on the boundary. The Laplace equation defined by $A^*A$ can be solved efficiently in $O(N)$ operations by a tridiagonal matrix solver based on finite-difference or finite-element discretization, where $N$ is the number of discretized grid points in the $x$ variable. We summarize the above details in Algorithm 2.

The above algorithm can be easily extended to compute the KR norm needed in evaluating the functional defined in equation 17.

Algorithm 2. The ADMM for evaluation of the KR norm.

1. Let $u_0 = 0, w_0 = 0$ and set the user-defined maximum iteration number MAXIT.
2. Fork $k = 0$: MAXIT
3. $v_{k+1} = (A^*A)^{-1}(A^*(w_k - \lambda_k) + ar)$.
4. $s_{k+1} = Av_{k+1} + \lambda_k$.
5. $w_{k+1} = \max(-1, \min(s_{k+1}, 1))$.
6. $\lambda_{k+1} = s_{k+1} - w_{k+1}$.
7. End

APPENDIX C

ASYMPTOTIC BEHAVIOR OF THE KR NORM

Let us first recall the definition of the KR norm for a function $r$ with zero mean,

\[\|r\|_{KR}^* = \max_{r \in \text{LiP}_1} \langle r, v \rangle,\]

where $\text{LiP}_1 = \{ v : |\nabla v|_{L^\infty} \leq 1 \}$ denotes the 1-Lipschitz set.

For $d = 1$, we have an explicit expression for the KR norm (Villani, 2003, p. 75),

\[\|r\|_{KR} = \left\| \int_{-\infty}^{x} r(s) ds \right\|_{L_1([R])}.\]

Letting the residual $r = u - d$ be of mean zero, we can obtain the following asymptotic behavior of the KR norm given in equation 10 with the bounded Lipschitz set 11,

\[\|r\|_{KR} = \begin{cases} \|r\|_{L_1([R])}, & \text{if } c \to +\infty, \\ c \|r\|_{KR}, & \text{if } c \to 0 \ (c \neq 0). \end{cases}\]

By a dual formulation of the KR norm, we have for any $r \in C_1([R])$ (Lellmann et al., 2014, Theorem 3.4)

\[\|r\|_{KR} = \min_{\psi \in C_1([R])} \|\partial_x \psi - r\|_{L_1([R])} + c \|\psi\|_{L_1([R])}.\]

Because the functions 0 and $\int_{-\infty}^{x} r(s) ds$ are both differentiable, we have

\[\|r\|_{KR} \leq \|r\|_{L_1([R])} \text{ and } \|r\|_{KR} \leq c \left\| \int_{-\infty}^{x} r(s) ds \right\|_{L_1([R])} = c \|r\|_{KR}.\]

Then, the minimizer $\psi^*$ of the problem in equation C-4 (for given $c$) satisfies $c \|\psi^*\|_{L_1([R])} \leq \|r\|_{L_1([R])}$ by the first inequality in formula C-5. Taking $c \to +\infty$, we have $\psi^* \to 0$ so that we can obtain

\[\|r\|_{KR} \to \|r\|_{L_1} \text{ as } c \to \infty,\]

which implies that the KR norm tends to the well-known $L_1$ norm.

Similarly, the minimizer $\psi^*$ satisfies $|\partial_x \psi^* - r|_{L_1} \leq c \|r\|_{KR}$ by the second inequality in formula C-5. Taking $c \to 0$, we have $\partial_x \psi^* \to r$, which implies that $\psi \to \int_{-\infty}^{x} r(s) ds$. Consequently, we have

\[\frac{1}{c} \|r\|_{KR} \to \left\| \int_{-\infty}^{x} r(s) ds \right\|_{L_1([R])} \text{ as } c \to 0,\]

which shows that the KR norm of a function behaves like the $L_1$-norm of its indefinite integral.

As the behavior of the KR norm discussed above changes quite differently for different relaxation $c$, a question arises naturally: How do we determine whether the relaxation parameter $c$ is small or not? Here, we suggest a heuristic way to achieve this goal that is effective in all of our numerical examples.
For the discretization of the admissible set BLip₂ = \{v_j; |v_j| \leq 1, |v_{j+1} - v_j| \leq c \Delta x, j = 0, 1, 2, \ldots, N\} (\Delta x is the grid size and N is the total number of the discretized points), the maximum value of the first-order finite difference should be bounded by \(2(\Delta x)^{-1}\) as

\[
\frac{|v_{j+1} - v_j|}{\Delta x} \leq \frac{1}{\Delta x} (|v_{j+1}| + |v_j|) \leq \frac{2}{\Delta x}, \quad j = 0, 1, \ldots, N.
\]

(C-8)

We have that the second constraint in the set BLip₂ will not be active so long as \(c > 2(\Delta x)^{-1}\). Therefore, we use the critical value \(c^* = 2(\Delta x)^{-1}\) as a threshold to choose the relaxation parameter, and we then can expect that the minimizer \(v^*\) should be a good approximation of the signum function for \(c = c^*\), as shown in the first asymptotic relation C-3 for parameter \(c\) large.

**APPENDIX D**

**ROBUSTNESS OF THE KR NORM AGAINST RANDOM NOISE**

In this appendix, we show that the 1D Gaussian random noise measured by the KR norm is negligible if the total number of measurement points is sufficiently large. Letting \(\zeta\) be a Gaussian random variable with mean zero and variance \(\sigma\), which is further approximated by the constant \(\zeta_j\) in each subinterval \((j-1)N^{-1}, jN^{-1} \}\ (j = 0, 1, 2, \ldots, N)\). Thus, we have by the explicit expression C-2:

\[
\|\zeta\|^2_{KR} = \int_0^1 \int_0^x \zeta^2 dt dx \approx N^{-2} \sum_{j=1}^{N} \sum_{k=1}^{j} \zeta_k^2.
\]

(D-1)

Define \(\nu_j = \sum_{k=1}^{j} \zeta_k\). By the fact that any linear combination of Gaussian random distributions is still a Gaussian random distribution, we have that each \(\nu_j\) is an independent Gaussian distribution with mean zero and variance \(j^2 \sigma^2\) \((j = 1, 2, \ldots, N)\). A direct computation of the expectation of Gaussian distributions \(\nu_j\) leads to

\[
\mathbb{E}[\nu_j] = \frac{1}{\sigma^2 \sqrt{2\pi}} \int_\mathbb{R} t e^{-\frac{t^2}{2\sigma^2}} dt = \frac{2\sqrt{\pi} \sigma}{\sqrt{2\pi}} \int_0^{\infty} s e^{-s^2/2} ds = \sqrt{\frac{2j}{\pi}},
\]

where we make a change of variable \(s = \sqrt{j} t\) in the second equation.

Hence, we have

\[
\mathbb{E}(\|\zeta\|^2_{KR}) \approx N^{-2} \sum_{j=1}^{N} \sqrt{\frac{2j}{\pi}} \approx O(N^{-1/2}),
\]

which implies that the KR norm for Gaussian random noise decays in the order \(N^{-1/2}\), hence, the KR norm is robust against Gaussian random noise.

**REFERENCES**

Aminzadeh, F., B. Jean, and T. Kunz, 1997, 3-D salt and overthrust models: SEG.


Engquist, B., B. D. Froese, and Y. Yang, 2016, Optimal transport for seismic full waveform inversion: Communications in Mathematical Sciences, 14, 2399–2415.

Glowinski, R., 2015, Variational methods for the numerical solution of nonlinear elliptic problems: SIAM.


Martinez, C., Y. Li, R. Krahenbuhl, and M. A. Braga, 2013, 3D inversion of airborne gravity gradiometry data in mineral exploration: A case study in


