



KANTOROVICH-RUBINSTEIN METRIC BASED LEVEL-SET METHODS FOR INVERTING MODULUS OF GRAVITY-FORCE DATA

WENBIN LI^{✉1} AND JIANLIANG QIAN^{✉*2,3}

¹School of Science, Harbin Institute of Technology, Shenzhen, Shenzhen 518055, China

²Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

³Department of CMSE, Michigan State University, East Lansing, MI 48824, USA

*In memory of Dr. Victor Isakov
for his many original contributions to inverse problems*

ABSTRACT. We propose to use the Kantorovich-Rubinstein (K-R) metric as a novel misfit function for the level-set based inverse gravity problems, where modulus of gravity-force data is used. By using the modulus data, we can satisfy the non-negativity requirement of distribution for the K-R metric naturally. Moreover, the K-R metric based level-set method can tolerate high level noise in the modulus data so that we can solve the domain inverse problem of gravimetry to high resolution. We develop the computational framework systematically. Numerical examples demonstrate the performance and effectiveness of the proposed algorithms.

1. Introduction. Inverse gravimetry as an inverse source problem has a long history [14], and in fact, prototypical inverse problems of potential theory were formulated by Laplace more than 200 years ago. However, such inverse problems are rather difficult to solve for two reasons: they are nonlinear and they are improperly posed. In 1943, to analyze stability of the inverse problem of potential theory, Tikhonov introduced certain important concepts of the theory of conditionally correct problems, which laid down the foundation for regularization theory of ill-posed problems. According to this theory, any conditionally correct problem can be studied numerically by means of regularization and the success of this solution process depends on the correctness class. Although existence theorems are of importance since they make sure that we do not use extra data, uniqueness questions are central in the theory of conditionally correct problems [14]. Over the years, a variety of uniqueness theorems for the inverse problem of potential theory have been established and most of these results are well summarized in the celebrated monograph by Dr. Victor Isakov [14]. Inspired by uniqueness results in [14], in [15, 16] Isakov and his collaborators developed a level-set framework to handle domain inverse problems which consist of one conditionally correct class of inverse gravimetry problems. Since such domain inverse problems have very poor stability

2020 *Mathematics Subject Classification.* Primary: 65N21, 65N06; Secondary: 65N80.

Key words and phrases. Kantorovich-Rubinstein metric, gravity, modulus data, inversion, level-set methods.

*Corresponding author: Jianliang Qian.

properties, such as logarithmic stability, and measured gravity anomaly data usually have very low precision, it is crucial to develop numerical inversion algorithms which can tolerate very high level of noise in measurements. Therefore, based on our recent works in [13, 12, 27], we further innovate the level-set framework for domain inverse problems by adopting the Kantorovich-Rubinstein (K-R) metric to measure data misfit, yielding an algorithm capable of tolerating very high level of noise in gravity-force data, where modulus of gravity force rather than gravity force itself is used in the inverse problem; this is the first time that the K-R metric is used for modulus of such data.

Since the gravitational force varies location-wise from that of a uniform sphere due to the earth's topographical features, such anomalous data, the so-called gravity anomaly, can be used to find subsurface anomalous volume mass distributions from the Newtonian gravitational force, which leads to the inverse problem of potential theory or the so-called inverse gravimetry. Consequently, we may solve such an inverse problem for many different applications at different scales, such as regional, explorational, or archaeological scales. Because of the equivalent source principle of gravitational potential, the inverse gravimetry is ill-posed in the Hadamard sense, where the Fredholm integral operator of the first kind defined by the Newtonian potential kernel has a nontrivial null space so that the inverse problem does not have a unique solution in general. Thus, to regain well-posedness from the inverse problem, we need to single out a conditionally correct class of problems that enjoys uniqueness, which is crucial for applying the Tikhonov regularization theory. One such conditionally correct class of problems is based on the equivalent source principle and is physically motivated by the following reasoning. Although subsurface anomalous mass can be distributed arbitrarily, the equivalent source principle indicates that there exists an average density-contrast function f and an associated domain D in R^n so that the anomalous measure $\mu = f\chi(D)$ can reproduce the given gravimetry data, where $\chi(D)$ is the characteristic function of domain D . However, without further assumption, one still cannot find density f and domain D simultaneously because there may exist infinitely many pairs $(f; D)$ that will reproduce the same gravimetry data. By making different assumptions on f and D , we can have various uniqueness results for different formulations of the inverse gravimetry problems, such as domain inverse problems, density inverse problems, and domain-density inverse problems [14]. In this paper, we will aim at the domain inverse problem, where the density-contrast function f is assumed to be a known constant, and the domain D is assumed to satisfy certain geometrical constraints so that the domain inverse problem has a unique solution which is logarithmically stable with respect to the given data in terms of modulus of gravity force.

Because a domain in the domain inverse problem may take a variety of shapes and may have multiple disconnected components, in [15, 16, 28] we have developed local level-set methods for the domain inverse problem, where the measurement data are gravity force. To do that, we start with a continuous function which is defined everywhere in the whole computational domain, and we further require that this function be positive inside a targeted domain and negative outside, which implies that the zero level-set where the function is zero describes exactly the boundary of the targeted domain, and this function is called the level-set function. A level-set implicit parametrization gives rise to many advantages, such as we have a globally defined function to manipulate, and the changes of geometry shape and connectivities can be automatically taken care of due to the underlying physical mechanism.

To tackle the so-called nil-zone problem induced by a varying density function of the anomalous mass distribution arising from the gravity inverse problem in geophysical prospecting [17, 18, 21, 1], where the density function should be understood as the “density contrast” function which may change its sign rather than as the “classical” density which is always non-negative, in [24] we have applied our level-set method to solve the corresponding domain inverse problem, where the given density function depends only on the x_3 -variable in the depth direction, and the measurement data are gravity force.

Although in a series of works [15, 16, 28, 29, 24, 26, 25, 12], we systematically developed level-set methods for solving this class of domain inverse problems, there are issues which are not well resolved, such as how to deal with sensitivity of L^2 -norm based misfit functions to noise and how to choose which set of gravity data to use. Advances in optimal transport theory [4, 9, 2, 3, 33, 31] in past several decades have motivated us to apply Wasserstein distances to measure the data-misfit in inverse gravity problems. In [13], we have developed a K-R metric based level-set method for inverting gradiometry (gravity-gradient) data in the conditionally correct class of domain inverse problems, which is capable of tolerating high-level noise; in [12] we proposed and analyzed a regularized Kantorovich-Rubinstein (K-R) metric to measure the distance between two given distributions, and we further applied this metric to solve the domain inverse problem of gravity so that we can recover the unknown support with high resolution, yielding an algorithm which is insensitive to Gaussian noise. On the other hand, to compute Wasserstein distances between two distributions, the two distributions are usually required to be non-negative, which is seldomly satisfied in many practical applications. To satisfy this non-negativity requirement, we propose to use the modulus of gravity-force data rather than the gravity force itself in the domain inverse problem so that we can apply the K-R metric in the level-set framework to measure data-misfit naturally, yielding an algorithm which is insensitive to high-level noise and using all the components of gravity force simultaneously rather than separately, where the modulus data enable us to avoid the issue of which set of data to use.

Other related works. We remark that in the literature the level-set method [30] has been widely used as a suitable and powerful tool for interfaces and shape-optimization problems mainly due to its ability in automatic interface merging and topological changes. In terms of non-geophysical inverse problems, the level-set method was first used for inverse obstacle problems in [32]; since then it has been applied to a variety of inverse problems. The level-set method was used for inverse scattering problems to reconstruct geometry of extended targets in [11] and [6]; see [5, 6] for surveys of related applications. In terms of geophysical inverse problems, the level-set method has also found its wide applications, such as in travel-time tomography problems in different settings [22, 36, 23].

Since contemporary gravity gradiometer measures various components of the gravity gradient, which is the second-order derivative of the exterior gravitational potential, in [29] we have developed fast local level-set methods to tackle the domain inverse problem, where the density function is assumed to be a known constant and the measurement data are gravity gradient; see [13, 19] for further developments in solving inverse problems using gravity gradiometry with different setups.

The paper is organized as the following. Section 2 formulates the inverse problem of modulus of gravity force. Section 3 develops the K-R metric based level-set

method, including implementation details. Section 4 shows some numerical examples to illustrate the performance and effectiveness of the proposed new algorithm. Section 5 provides a discussion about the choices of parameters in practice. And Section 6 draws the conclusion.

2. The inverse problem of modulus of gravity force. The gravity potential U generated by a mass distribution μ with $\text{supp } \mu \subset \Omega$ is defined as follows,

$$U(\mathbf{y}; \mu) = \gamma \int_{\Omega} K(\mathbf{y}, \mathbf{x}) d\mu(\mathbf{x}), \quad (1)$$

where $K(\mathbf{y}, \mathbf{x}) = K(|\mathbf{y} - \mathbf{x}|)$ is the fundamental solution of Laplace's equation,

$$K(\mathbf{y}, \mathbf{x}) = \begin{cases} -\frac{1}{2\pi} \ln |\mathbf{y} - \mathbf{x}| & , \quad \mathbf{x}, \mathbf{y} \in \mathbf{R}^2, \\ \frac{1}{4\pi |\mathbf{y} - \mathbf{x}|} & , \quad \mathbf{x}, \mathbf{y} \in \mathbf{R}^3, \end{cases} \quad (2)$$

and γ is a constant related to the universal gravitational constant. The vector $\nabla U(\mathbf{y}; \mu)$ represents the gravity force at $\mathbf{y}$ generated by the mass distribution μ , and $|\nabla U(\mathbf{y}; \mu)|$ is the modulus of gravity. Since $|\nabla U(\mathbf{y}; \mu)|$ is nonlinear stacking of multiple components of the gravity-force vector $\nabla U(\mathbf{y}; \mu)$, it is considered to be less sensitive to systematic measurement noises than $\nabla U(\mathbf{y}; \mu)$ which is directionally dependent. The inverse problem of modulus of gravity-force data is posed as follows: Given the modulus of gravity force $|\nabla U(\mathbf{y}; \mu)|$ on $\Sigma_0 \subset \mathbf{R}^n \setminus \Omega$, find the mass distribution μ with $\text{supp } \mu \subset \Omega$.

In this work, we will focus on a conditionally correct class of inverse problems as the following: the volume mass distribution takes the form of $\mu = f \chi_D$, where χ_D denotes the indicator function of the source domain D ,

$$\chi_D(x) = \begin{cases} 1 & , \quad x \in D \\ 0 & , \quad x \notin D \end{cases}, \quad (3)$$

and f is the density-contrast coefficient function. Starting from the work of Novikov (1938), a variety of well-posedness theorems have been proved for the gravity inverse problems with volume mass distribution [14]. We summarize some relevant uniqueness results in the following.

Let Ω be a bounded domain in $\mathbf{R}^n$ with connected $\mathbf{R}^n \setminus \bar{\Omega}$, and D_1, D_2 and D be open sets in Ω having piecewise smooth boundaries. Let Ω_0 be a convex domain with analytic (regular) boundary, $\Omega \subset \Omega_0$, and $\Sigma_0 \subset \partial\Omega_0$ be a nonempty hyper-surface. Denote the spatial coordinate $\mathbf{x} \in \mathbf{R}^n$ as $\mathbf{x} = (x_1, \dots, x_n)$.

We start with the following notion. The center of gravity of a positive measure μ is a point satisfying $\|\mu\|^{-1} \int \mathbf{x} d\mu(\mathbf{x})$, and accordingly, the center of gravity of a bounded open set Ω is the center of gravity of the measure $\chi(\Omega) d\mathbf{x}$. We have the following results.

Lemma 2.1 (Corollary 1.8.2 in [14]). *Consider the Laplacian operator defined by the Newtonian potential kernel. If Newtonian potentials of two positive measures with compact supports coincide outside a certain ball, then these measures have the same centers of gravity.*

Lemma 2.2 (Lemma 2.1.1 in [14]). $U_j = U(\cdot; \mu_j)$ with the measures $\mu_j \geq 0$, $1 \leq j \leq 2$. Then $|\nabla U_1| = |\nabla U_2|$ on $\Sigma_0 \implies U_1 = U_2$ on $\mathbf{R}^n \setminus \Omega$.

Theorem 2.1 (Theorem 2.2.1 in [14]). *Suppose that either (i) D_1 and D_2 are star-shaped with respect to their centers of gravity; or (ii) D_1 and D_2 are convex in the x_1 direction. Then $U(\cdot; \chi_{D_1}) = U(\cdot; \chi_{D_2})$ on $\mathbf{R}^n \setminus \Omega \implies D_1 = D_2$.*

Theorem 2.2 (Theorems 3.1.3 and 3.1.4 in [14]). *Suppose that either (i) D_1 and D_2 are convex, $f \in L_1(\Omega)$, and $0 < f$ on Ω ; or (ii) D_1 and D_2 are convex in x_1 , f does not depend on x_1 , $f \in C(\Omega)$, and $\Omega \subset \text{supp } f$. Then $U(\cdot; f \chi_{D_1}) = U(\cdot; f \chi_{D_2})$ on $\mathbf{R}^n \setminus \Omega \implies D_1 = D_2$.*

In the above theorems, a domain D in $\mathbf{R}^n$ is said to be star-shaped with respect to a point $\mathbf{x}_0$ if $\mathbf{x}_0 \in D$ and the intersection of any ray originated at $\mathbf{x}_0$ with D is an interval; an open set D is said to be convex in x_1 if the intersection of any straight line parallel to the x_1 -axis with D is an interval. Lemma 2.2 is a continuation result for the measurement gravity data. Theorems 2.1 and 2.2 are uniqueness theorems for recovering the unknown domain of the mass distribution. There are actually plenty of more well-posedness results provided in [14] including the uniqueness theorems for recovering the unknown density contrast and for simultaneously recovering the unknown domain and density contrast, some of which are also summarized in our previous work [27]. In this paper, we will focus on the recovery of unknown domain D in the form of volume mass distribution $\mu = f \chi_D$ by freezing the density-contrast coefficient function f . Combining Lemma 2.2, Theorem 2.1 and Theorem 2.2, it is immediate to get the following result [27].

Corollary 2.1. *Given data as the modulus of gravity, $d = |\nabla U|$, on Σ_0 and given $f \geq 0$ in Ω , the domain D of the volume mass distribution $\mu = f \chi_D$ can be uniquely determined if one of the following constraints is satisfied:*

- (i) D is star-shaped with respect to its center of gravity, and $f = C$, where C is a constant;
- (ii) D is convex in x_1 , and $f = C$, where C is a constant;
- (iii) D is convex in x_1 , f does not depend on x_1 , $f \in C(\Omega)$, and $\Omega \subset \text{supp } f$;
- (iv) D is convex, $f \in L_1(\Omega)$, and $0 < f$ on Ω .

The requirement $f \geq 0$ arises from Lemma 2.2, which ensures that the partial measurements of $|\nabla U|$ imply the full measurements of U . In practice, violating the condition $f \geq 0$ is admissible if there are sufficient and near complete measurements.

Applying Lemma 2.1 to the conditionally correct class of mass distribution, $\mu = f \chi_D$, we conclude that all those distributions will have the same center of gravity, no matter how f is chosen, large or small. Further, Theorem 2.1 says that D is unique if f is given. Therefore, even if we choose f unreasonably large or small, we will always be able to capture the same center of gravity of all those equivalent sources. Such reconstruction results will tell us a lot about the unknown source that we are interested in.

3. Level-set methods with Kantorovich-Rubinstein metrics.

3.1. Level-set formulation. Starting from the work [15], the level-set method is used to formulate the mass distribution in gravity inverse problems. Let ρ be the representation function of the mass distribution μ , viz. $\mu(\psi) = \int_{\Omega} \psi \rho dx$ for all $\psi \in C(\Omega)$ with $\text{supp } \psi \subset \Omega$. Since μ takes the form of volume mass distribution as $\mu = f \chi_D$, the representation function ρ can be naturally expressed as

$$\rho(\mathbf{x}) = f(\mathbf{x}) H(\phi(\mathbf{x})), \quad (4)$$

where ϕ is the level-set function,

$$\phi(\mathbf{x}) \begin{cases} > 0 & , \quad \mathbf{x} \in D \\ = 0 & , \quad \mathbf{x} \in \partial D \\ < 0 & , \quad \mathbf{x} \in \bar{D}^c \end{cases} , \quad (5)$$

and $H(\cdot)$ is the Heaviside function in the following form,

$$H(\phi) = \begin{cases} 1 & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases} . \quad (6)$$

Equation (4) is the level-set formulation for the representation function ρ . The zero level-set $\{\mathbf{x} \mid \phi(\mathbf{x}) = 0\}$ indicates the location of ∂D or the boundary of the domain D , and the Heaviside function $H(\phi)$ corresponds to the characteristic function χ_D in the volume mass distribution μ . To recover the unknown domain D with a given density-contrast coefficient f , one only needs to compute the level-set function $\phi(\mathbf{x})$. The derivative of the mass representation function ρ with respect to the level-set function ϕ is given as follows,

$$\frac{\partial \rho}{\partial \phi}(\mathbf{x}) = f(\mathbf{x}) \delta(\phi(\mathbf{x})), \quad (7)$$

where $\delta(\phi)$ is the Dirac delta function that arises from the derivative of the Heaviside function $H(\phi)$.

In numerical computation, continuous functions are used to evaluate $\phi(\mathbf{x})$, $H(\phi)$ and $\delta(\phi)$ so that we can avoid numerical instabilities in computation. Specifically, the level-set function ϕ is maintained to be a continuous signed distance function in the following form,

$$\phi(\mathbf{x}) = \begin{cases} \text{dist}(\mathbf{x}, \partial D) & , \quad \mathbf{x} \in D \\ -\text{dist}(\mathbf{x}, \partial D) & , \quad \mathbf{x} \in \bar{D}^c \end{cases} , \quad (8)$$

where $\text{dist}(\mathbf{x}, \partial D)$ denotes the distance between $\mathbf{x}$ and ∂D . The Heaviside function $H(\phi)$ is approximated by a smooth numerical Heaviside function $H_\tau(\phi)$,

$$H_\tau(\phi) = \begin{cases} 0 & , \quad \phi < -\tau \\ \frac{1}{2} + \frac{\phi}{2\tau} + \frac{1}{2\pi} \sin\left(\frac{\pi\phi}{\tau}\right) & , \quad -\tau \leq \phi \leq \tau \\ 1 & , \quad \phi > \tau \end{cases} , \quad (9)$$

where τ is a parameter that controls the thickness of the interface ∂D in numerical expression. The Dirac delta function $\delta(\phi)$ is evaluated by the following approximate delta function [35],

$$\delta_\epsilon(\phi) = \chi(T_\epsilon) |\nabla \phi|, \quad (10)$$

where $\chi(\cdot)$ is the indicator function, and T_ϵ denotes the region $T_\epsilon = \{\mathbf{x} : |\phi(\mathbf{x})| < \epsilon\}$ which is a neighborhood of the zero level set.

3.2. Kantorovich-Rubinstein metric for the modulus data fitting. The components of the gravity force $\mathbf{g} = \nabla U \in \mathbf{R}^n$ are given by

$$g_i(\mathbf{y}; \rho) = \gamma \int_{\Omega} K_i(\mathbf{y}, \mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}, \quad 1 \leq i \leq n, \quad (11)$$

where $\rho(\mathbf{x})$ is the representation function of the mass distribution μ , and the integral kernel $K_i(\mathbf{y}, \mathbf{x})$ has the following form,

$$K_i(\mathbf{y}, \mathbf{x}) = \frac{\partial K(\mathbf{y}, \mathbf{x})}{\partial y_i} = -\frac{1}{2^{n-1}\pi} \frac{y_i - x_i}{|\mathbf{y} - \mathbf{x}|^n}, \quad n = 2 \text{ or } 3. \quad (12)$$

The gravity modulus datum $d := |\nabla U|$ is then computed by

$$d(\mathbf{y}; \rho) = \left(\sum_{i=1}^n g_i^2(\mathbf{y}; \rho) \right)^{\frac{1}{2}}. \quad (13)$$

Denote the spatial coordinate $\mathbf{y} \in \mathbf{R}^n$ as $\mathbf{y} = (y_1, \dots, y_n)$, and suppose that the measurement data can be parameterized on an $(n-1)$ -dimensional plane

$$\Sigma_0 = \{\mathbf{s}(\mathbf{y}) : \mathbf{s}(\mathbf{y}) = (y_1, \dots, y_{n-1})\} \subset \mathbf{R}^{n-1}, \quad (14)$$

where we abuse the notation Σ_0 . We propose to use a Kantorovich-Rubinstein metric to measure the discrepancy between simulated modulus data and observations,

$$E_d = \|d(\mathbf{y}; \rho) - d^*(\mathbf{y})\|_{\text{KR},c} := \max_{v \in \text{BLip}_c} \int_{\Sigma_0} (d(\mathbf{y}; \rho) - d^*(\mathbf{y})) v(\mathbf{s}(\mathbf{y})) \, d\mathbf{s}(\mathbf{y}), \quad (15)$$

where d and d^* denote the simulated and observed gravity modulus data, respectively, and BLip_c denotes the bounded Lipschitz set defined as follows,

$$\text{BLip}_c = \left\{ v(\mathbf{s}(\mathbf{y})) : |v| \leq 1, \left| \frac{\partial v}{\partial y_i} \right| \leq c, i = 1, \dots, n-1 \right\}. \quad (16)$$

The Kantorovich-Rubinstein metric is closely related to the L^1 -Wasserstein distance [20, 33]. Some recent works have shown that using Wasserstein distance to measure data discrepancy has the advantage of mitigating non-convexity and local minima [7, 8], but directly applying the Wasserstein distance generally requires non-negativity of the data and the mass conservation between the simulated and observed data, which may cause difficulties to many applications. In our work, the gravity modulus data $|\nabla U|$ are naturally nonnegative, and the boundedness constraint $|v| \leq 1$ in the definition of BLip_c relaxes the requirement of data conservation. There are also some other slightly different formulations for the Kantorovich-Rubinstein metric [20, 12]. In this paper we choose to use the form defined by equations (15) and (16) for direct and simple computations of the metric in numerical implementation.

Let $\hat{v}$ denote the maximizer in equation (15),

$$\hat{v}(\mathbf{s}(\mathbf{y})) := \arg \max_{v \in \text{BLip}_c} \int_{\Sigma_0} (d(\mathbf{y}; \rho) - d^*(\mathbf{y})) v(\mathbf{s}(\mathbf{y})) \, d\mathbf{s}(\mathbf{y}). \quad (17)$$

The discrepancy function E_d can be computed as follows,

$$E_d = \|d(\mathbf{y}; \rho) - d^*(\mathbf{y})\|_{\text{KR},c} = \int_{\Sigma_0} (d(\mathbf{y}; \rho) - d^*(\mathbf{y})) \hat{v}(\mathbf{s}(\mathbf{y})) \, d\mathbf{s}(\mathbf{y}). \quad (18)$$

The Fréchet derivative of E_d with respect to d is given by

$$\frac{\partial E_d}{\partial d}(\mathbf{y}) = \hat{v}(\mathbf{s}(\mathbf{y})). \quad (19)$$

As a result, solving $\hat{v}$ in equation (15) is essential for the computations of the Kantorovich-Rubinstein-metric based data discrepancy.

Similar to the work [13], we propose to use an alternating-direction method of multipliers (ADMM) algorithm [10, 34] to efficiently compute the maximizer $\hat{v}(\mathbf{s}(\mathbf{y}))$. The details are as follows. Firstly, let

$$B = \{\mathbf{w} = (w_1, \dots, w_n)^T : |w_i(\mathbf{s})| \leq 1, 1 \leq i \leq n\}, \quad (20)$$

and introduce a linear operator $\mathcal{A}$ defined as

$$\mathcal{A}v(\mathbf{s}) = \left(v, \frac{1}{c} \frac{\partial v}{\partial y_1}, \dots, \frac{1}{c} \frac{\partial v}{\partial y_{n-1}} \right)^T. \quad (21)$$

The bounded Lipschitz set BLip_c can be re-formulated as

$$\text{BLip}_c = \{v(\mathbf{s}) : \mathcal{A}v(\mathbf{s}) \in B\}. \quad (22)$$

Denoting ι_B the zero-infinity indicator function,

$$\iota_B(\mathbf{w}) = \begin{cases} 0 & , \quad \mathbf{w} \in B \\ +\infty & , \quad \mathbf{w} \notin B \end{cases} \quad (23)$$

one can rewrite the optimization problem of (17) in the following equivalent form,

$$\widehat{v}(\mathbf{s}) = \arg \min_v -\langle d - d^*, v \rangle + \iota_B(\mathcal{A}v). \quad (24)$$

Introducing the auxiliary variable $\mathbf{w} = \mathcal{A}v$, the augmented Lagrangian for (24) is

$$L(v, \mathbf{w}; \lambda) = -\langle d - d^*, v \rangle + \iota_B(\mathbf{w}) + \frac{1}{\alpha} \langle \lambda, \mathcal{A}v - \mathbf{w} \rangle + \frac{1}{2\alpha} \|\mathcal{A}v - \mathbf{w}\|_2^2. \quad (25)$$

The optimizer $\widehat{v}(\mathbf{s})$ is achieved by solving the saddle-point problem of (25),

$$(\widehat{v}, \widehat{\mathbf{w}}; \widehat{\lambda}) = \sup_{\lambda} \inf_{v, \mathbf{w}} L(v, \mathbf{w}; \lambda). \quad (26)$$

We employ the ADMM algorithm to solve this saddle-point problem, which leads to the following alternating iterations,

$$\begin{cases} v_{k+1} &= \arg \min_v L(v, \mathbf{w}_k; \lambda_k), \\ \mathbf{w}_{k+1} &= \arg \min_{\mathbf{w}} L(v_{k+1}, \mathbf{w}; \lambda_k), \\ \lambda_{k+1} &= \lambda_k + \mathcal{A}v_{k+1} - \mathbf{w}_{k+1}. \end{cases} \quad (27)$$

In particular, we have

$$\begin{aligned} v_{k+1} &= \arg \min_v \frac{1}{2\alpha} \|\mathcal{A}v - \mathbf{w}_k + \lambda_k\|_2^2 - \langle d - d^*, v \rangle \\ &= (\mathcal{A}^* \mathcal{A})^{-1} (\mathcal{A}^* (\mathbf{w}_k - \lambda_k) + \alpha(d - d^*)) \end{aligned} \quad (28)$$

and

$$\begin{aligned} \mathbf{w}_{k+1} &= \arg \min_{\mathbf{w}} \frac{1}{2\alpha} \|\mathcal{A}v_{k+1} - \mathbf{w} + \lambda_k\|_2^2 + \iota_B(\mathbf{w}) \\ &= \max(-1, \min(\mathcal{A}v_{k+1} + \lambda_k, 1)), \end{aligned} \quad (29)$$

where $\mathcal{A}^*$ denotes the Hilbert adjoint of $\mathcal{A}$,

$$\mathcal{A}^* \mathbf{w}(\mathbf{s}) = \left(1, -\frac{1}{c} \frac{\partial}{\partial y_1}, \dots, -\frac{1}{c} \frac{\partial}{\partial y_{n-1}} \right) \mathbf{w}(\mathbf{s}) \quad \text{for } \mathbf{w}(\mathbf{s}) = (w_1(\mathbf{s}), \dots, w_n(\mathbf{s}))^T, \quad (30)$$

and so

$$\mathcal{A}^* \mathcal{A} = 1 - \frac{1}{c^2} \sum_{i=1}^{n-1} \frac{\partial^2}{\partial y_i^2}. \quad (31)$$

3.3. Optimization and level-set evolution. The level-set function $\phi(\mathbf{x})$ that models the unknown domain D of the mass distribution is reconstructed by solving the following optimization problem,

$$\phi := \arg \min_{\phi} E_d(\rho(\phi)), \quad (32)$$

where $E_d(\rho(\phi))$ denotes the data-discrepancy function defined by equation (15). In practice, some regularization terms can be added to the objective function, which help to provide smoothness for ϕ and avoid sharp oscillations on the interface; for example, see [27] for such details. Here we choose not to include regularization terms in the objective function, so that when we compare the Kantorovich-Rubinstein metric based solution with the L^1 -norm and L^2 -norm based ones, the effects of

regularization terms disappear and the differences of solutions are only caused by different choices of data discrepancies.

We will apply the gradient descent method to solve the optimization problem in equation (32). The gradient direction of E_d is evaluated in the following way. Firstly,

$$\begin{aligned}\frac{\partial E_d}{\partial \phi}(\mathbf{x}) &= \frac{\partial E_d}{\partial \rho}(\mathbf{x}) \frac{\partial \rho}{\partial \phi}(\mathbf{x}) = \frac{\partial \rho}{\partial \phi}(\mathbf{x}) \int_{\Sigma_0} \frac{\partial E_d}{\partial d}(\mathbf{y}) \frac{\partial d}{\partial \rho}(\mathbf{y}, \mathbf{x}) \, \mathrm{d}\mathbf{s}(\mathbf{y}) \\ &= \gamma f(\mathbf{x}) \delta(\phi(\mathbf{x})) \int_{\Sigma_0} \hat{v}(\mathbf{s}(\mathbf{y})) \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, \mathrm{d}\mathbf{s}(\mathbf{y}),\end{aligned}\quad (33)$$

where we have used equations (7), (11), (13), and (19) in the calculation. Replacing the delta function with the numerical delta function $\delta_\epsilon(\phi) = \chi(T_\epsilon) |\nabla \phi|$ as shown in equation (10), equation (33) is approximated by

$$\frac{\partial E_d}{\partial \phi}(\mathbf{x}) = \gamma f(\mathbf{x}) \chi(T_\epsilon) |\nabla \phi| \int_{\Sigma_0} \hat{v}(\mathbf{s}(\mathbf{y})) \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, \mathrm{d}\mathbf{s}(\mathbf{y}), \quad (34)$$

which can be denoted as

$$\frac{\partial E_d}{\partial \phi}(\mathbf{x}) = V_n(\mathbf{x}) |\nabla \phi(\mathbf{x})|, \quad (35)$$

with

$$V_n(\mathbf{x}) := \gamma f(\mathbf{x}) \chi(T_\epsilon) \int_{\Sigma_0} \hat{v}(\mathbf{s}(\mathbf{y})) \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, \mathrm{d}\mathbf{s}(\mathbf{y}). \quad (36)$$

The level-set function ϕ is then updated according to the negative gradient descent direction,

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E_d}{\partial \phi} = -V_n |\nabla \phi|, \quad (37)$$

where t denotes the artificial evolution time. The updating equation (37) can be viewed as a Hamilton-Jacobi equation. We propose to use a direct forward Euler scheme to discretize $\frac{\partial \phi}{\partial t}$, where the step-size Δt should satisfy the following Courant-Friedrichs-Lewy (CFL) condition to ensure the stability of evolution,

$$\Delta t \frac{\max_{\mathbf{x}} |V_n(\mathbf{x})|}{\min\{\Delta x_1, \dots, \Delta x_n\}} < 1. \quad (38)$$

In practice, we take the time step-size as

$$\Delta t = \beta \frac{\min\{\Delta x_1, \dots, \Delta x_n\}}{\max_{\mathbf{x}} |V_n(\mathbf{x})|}, \quad (39)$$

where $\beta \in (0, 1)$ is a constant, e.g. $\beta = 0.5$ in our numerical computation.

In addition, the level-set reinitialization is applied to ϕ after every update of ϕ , so that the level-set function ϕ is maintained to be roughly a continuous signed distance function as shown in equation (8). The details of level-set reinitialization can be found in our previous works [15, 27].

We summarize the overall algorithm as follows, which is in the framework of the Isakov-Leung-Qian algorithm [15, 28]:

-
1. Initialize the level-set function ϕ .
 2. Compute the simulated data according to (11) and (13).

3. Compute $\widehat{v}(\mathbf{s}(\mathbf{y}))$ according to (17) by the ADMM algorithm; compute the derivative term V_n according to (36).
4. Evolve the level set function ϕ according to (37).
5. Reinitialize the level set function to maintain the signed distance property.
6. Repeat 2-5 until it converges.

4. Numerical examples. We provide synthetic 2D examples to illustrate the efficacy of the inversion algorithm. The spatial coordinate $\mathbf{x} \in \mathbf{R}^2$ is denoted as $\mathbf{x} = (x, z)$, where x indicates the horizontal direction and z indicates the vertical direction. The synthetic measurement data are generated according to equations (11) and (13). We add Gaussian noises to the linear components of the gravity modulus data,

$$g_i^* = g_i (1 + \eta \cdot \mathcal{N}(0, 1)) , \quad i = 1, 2 \quad (40)$$

where g_i is generated according to equation (11), $\mathcal{N}(0, 1)$ denotes the Gaussian noises with zero mean and standard deviation 1, and η is the parameter prescribing the percentage of noises. Then the synthetic modulus data are generated according to equation (13):

$$d^* = \left((g_1^*)^2 + (g_2^*)^2 \right)^{\frac{1}{2}} .$$

To illustrate the effect of the Kantorovich-Rubinstein metric in the inversion algorithm, we will compare our solutions using the Kantorovich-Rubinstein metric as data fitting with the solutions using the L^1 -norm and L^2 -norm as data fitting, respectively. Here, the L^1 -norm and L^2 -norm data discrepancies are defined as follows,

$$\begin{aligned} E_{d,L^1} &:= \|d(\mathbf{y}; \rho) - d^*(\mathbf{y})\|_{L^1} = \int_{\Sigma_0} |d(\mathbf{y}; \rho) - d^*(\mathbf{y})| \, d\mathbf{s}(\mathbf{y}) , \\ E_{d,L^2} &:= \|d(\mathbf{y}; \rho) - d^*(\mathbf{y})\|_{L^2}^2 = \frac{1}{2} \int_{\Sigma_0} (d(\mathbf{y}; \rho) - d^*(\mathbf{y}))^2 \, d\mathbf{s}(\mathbf{y}) . \end{aligned}$$

The details of the corresponding inversion algorithms are included in Appendix A.

4.1. Example 1. Figure 1(a) shows the exact model, where the mass distribution includes two circles with a constant density contrast $f = 1$. The computational domain is $\Omega = (0, 1) \times (0, 0.5)$ km, and the mesh size is $\Delta x = \Delta z = 0.01$ km. The measurement data are uniformly collected along the surface

$$\Sigma_0 = \{(x, z) : -0.5 \leq x \leq 1.5, z = -0.05\} ,$$

where the sampling distance is taken as $\Delta x_m = 0.01$ km. To perform the inversion algorithm, we take the initial guess of the level-set function ϕ as follows,

$$\phi_0(x, z) = 0.1 - \sqrt{\left(\frac{x - 0.5}{2}\right)^2 + (z - 0.25)^2} ; \quad (41)$$

Figure 1(b) plots the initial structure of mass distribution.

Figures 2 and 3 show the inversion results, where we provide the solutions using the Kantorovich-Rubinstein metric, the L^1 norm and L^2 norm for data fitting, respectively. In this example, we take $c = 6$ for the Kantorovich-Rubinstein-metric based data discrepancy as shown in equation (15). We have performed inversions for the gravity modulus data with different rates of Gaussian noises, where the noises are added to the linear components of the modulus data as shown in equation (40).

We consider the data with 20%, 40%, 60% and 80% noise, respectively. The left column of Figure 2 provides the results for the modulus data with 20% Gaussian noise, where Figure 2(a) shows the noisy data, Figure 2(c) plots the solution using the Kantorovich-Rubinstein metric for data fitting, Figure 2(e) plots the solution using the L^1 norm for data fitting, and Figure 2(g) plots the solution using the L^2 norm for data fitting. Similarly, the right column of Figure 2 provides the results for the modulus data with 40% Gaussian noise. Figure 3 provides the inversion results for the modulus data with 60% and 80% Gaussian noise. In addition, Figure 4 plots the data discrepancy functions E_d , E_{d,L^1} and E_{d,L^2} in the above inversions. Basically we perform a large number of iterations and take the recovered solutions after full convergences. The data discrepancy function can be oscillating in the iterations partly due to the nonlinearity of level-set inversion and partly due to the noise contamination.

Under 20% or 40% Gaussian noise, inversions using the Kantorovich-Rubinstein metric, the L^1 norm, and the L^2 norm, all yield reasonable solutions. However, the effect of the Kantorovich-Rubinstein metric becomes significant as the level of noise contamination becomes high. As shown in Figure 3(c) and Figure 3(d), the K-R solutions adequately recover the true models even under the strong Gaussian noise of 60% or 80%, respectively. These results show that the K-R metric based level-set inversion algorithm using modulus of gravity-force data is insensitive to very high level of Gaussian noise.

4.2. Example 2. We consider a 2D salt model with varying density contrast in this example. Figure 5(a) shows the exact model. The salt contour is taken from a slice of the SEG/EAGE 3D salt model, and the varying density contrast is given as $f(x, z) = 0.2 \cdot (1.8 - z)$. The computational domain is $(0, 13.4) \times (0, 4)$ km and the mesh size is $\Delta x = \Delta z = 0.02$ km. The initial guess of the level-set function ϕ is taken as

$$\phi_0(x, z) = 1 - \sqrt{\left(\frac{x - 6.7}{2}\right)^2 + (z - 2)^2}; \quad (42)$$

Figure 5(b) shows the initial structure of mass distribution.

The measurement data are collected along

$$\Sigma_0 = \{(x, z) : -13 \leq x \leq 27, z = -0.1\}$$

with the sampling distance $\Delta x_m = 0.2$ km. We perform inversions for the modulus data with 20%, 40%, 60% and 80% Gaussian noise, respectively. Figures 6 and 7 show the inversion results, where we provide the solutions using the Kantorovich-Rubinstein metric, the L^1 norm, and the L^2 norm for data fitting, respectively. We take $c = 2$ in the formula of the Kantorovich-Rubinstein metric based data discrepancy E_d . To illustrate the convergence history, Figure 8 plots the data discrepancy functions E_d , E_{d,L^1} and E_{d,L^2} in the above inversions.

As shown in Figures 6 and 7, the solutions using the K-R metric based data discrepancy are generally more accurate than those using the L^1 norm and L^2 norm data discrepancy, which capture the top shape and the overall contour of the salt model more accurately. More importantly, the inversions using the K-R metric data discrepancy show strong capabilities of tolerating high-level noise. As shown in Figure 7(d), the solution remains stable even under 80% Gaussian noise. In contrast, the solutions with the L^1 norm or L^2 norm data discrepancies generate

plenty of artifacts due to the noise contamination as shown in Figures 6(f)-(h) and 7(e)-(h).

4.3. Example 3. We consider a 2D salt model with a constant density contrast in this example. Figure 9(a) shows the exact model. Similar to Example 2, the salt contour is taken from a slice of the SEG/EAGE 3D salt model, and the difference is that the density contrast is given as $f(x, z) = 0.2$. The computational domain is $(0, 13.4) \times (0, 4)$ km and the mesh size is $\Delta x = \Delta z = 0.02$ km. The initial guess of the level-set function ϕ is again taken as equation (42):

$$\phi_0(x, z) = 1 - \sqrt{\left(\frac{x - 6.7}{2}\right)^2 + (z - 2)^2};$$

Figure 9(b) shows the initial structure of mass distribution.

The measurement data are collected along

$$\Sigma_0 = \{(x, z) : -13 \leq x \leq 27, z = -0.1\}$$

with the sampling distance $\Delta x_m = 0.2$ km. We perform inversions for the modulus data without noise and with 5%, 10% and 20% Gaussian noise, respectively. Figures 10 and 11 show the inversion results, where we provide the solutions using the Kantorovich-Rubinstein metric, the L^1 norm, and the L^2 norm for data fitting, respectively. We take $c = 1.5$ in the formula of the Kantorovich-Rubinstein-metric based data discrepancy E_d . Figure 12 shows the convergence history of the data discrepancy functions E_d , E_{d,L^1} and E_{d,L^2} in the above inversions.

As shown in Figures 10 and 11, the K-R-metric based solutions are more reliable than the solutions using the L^1 norm and L^2 norm data discrepancy terms. By fitting the modulus data with the K-R metric, the solutions successfully recover the top shape and the overall contour of the salt model. The advantage of using the K-R metric becomes significant as the level of noises becomes high. As shown in the right column of Figure 11, 20% Gaussian noise with the L^1 norm or the L^2 norm data discrepancy generates artifacts polluting the solutions, while the solution using the K-R-metric based discrepancy remains stable and adequately recovers the salt model. We conclude that the proposed level-set inversion algorithm using gravity modulus data and the Kantorovich-Rubinstein metric has the capability of tolerating high-level random noises.

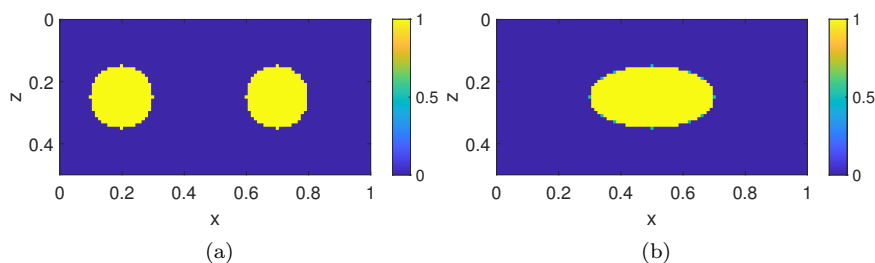


FIGURE 1. Example 1. Two circles with constant density contrast.

(a) Exact model of mass distribution; (b) initial guess.

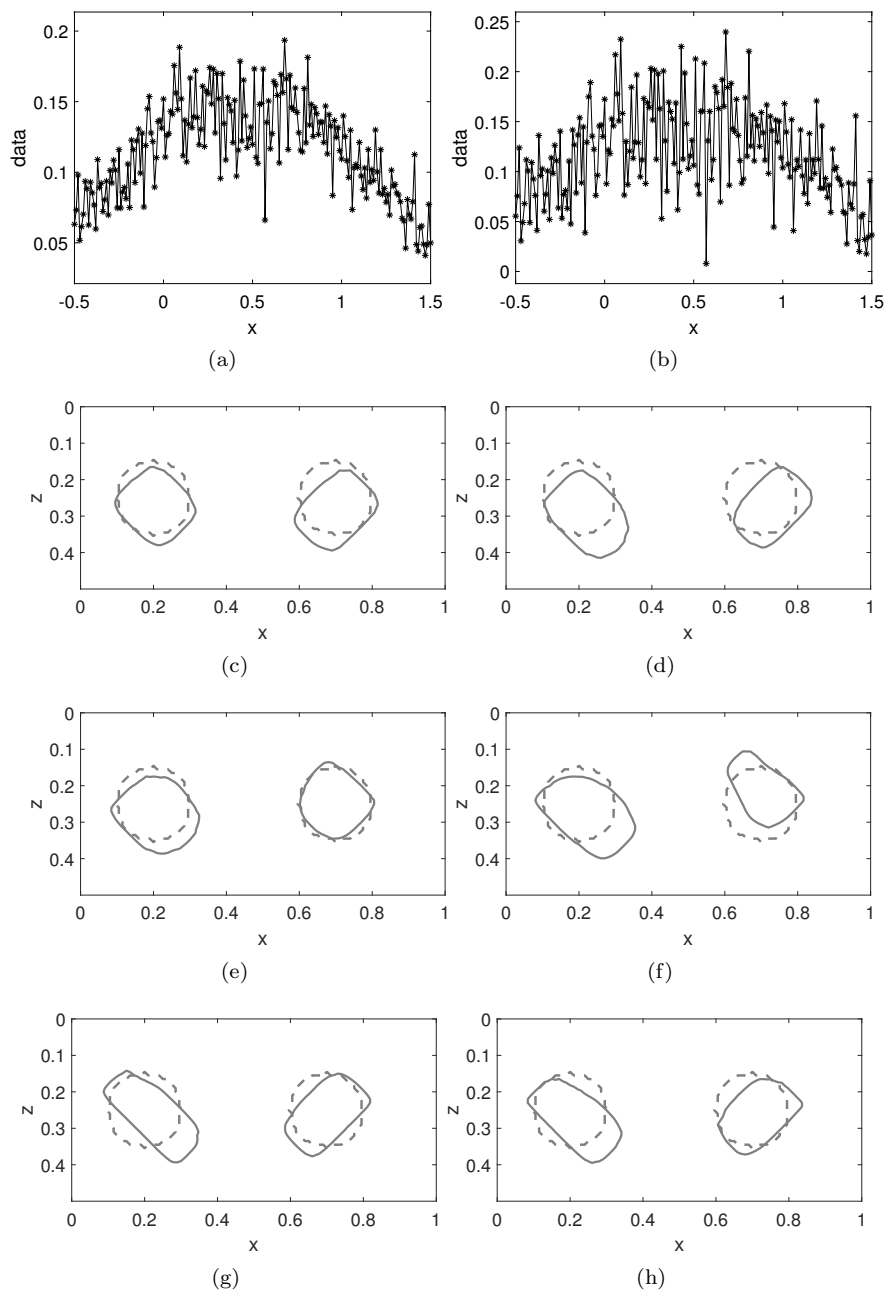


FIGURE 2. Example 1. Recovered solutions. The left column shows the results for the modulus data with 20% Gaussian noises added to the components; the right column shows the results for the modulus data with 40% Gaussian noises added to the components. (a), (b): modulus data with 20% noises and 40% noises, respectively; (c), (d): corresponding solutions using Kantorovich-Rubinstein metric for data fitting; (e), (f): corresponding solutions using L^1 norm for data fitting; (g), (h): corresponding solutions using L^2 norm for data fitting. In (c)-(h), the dashed line plots the exact model for reference.

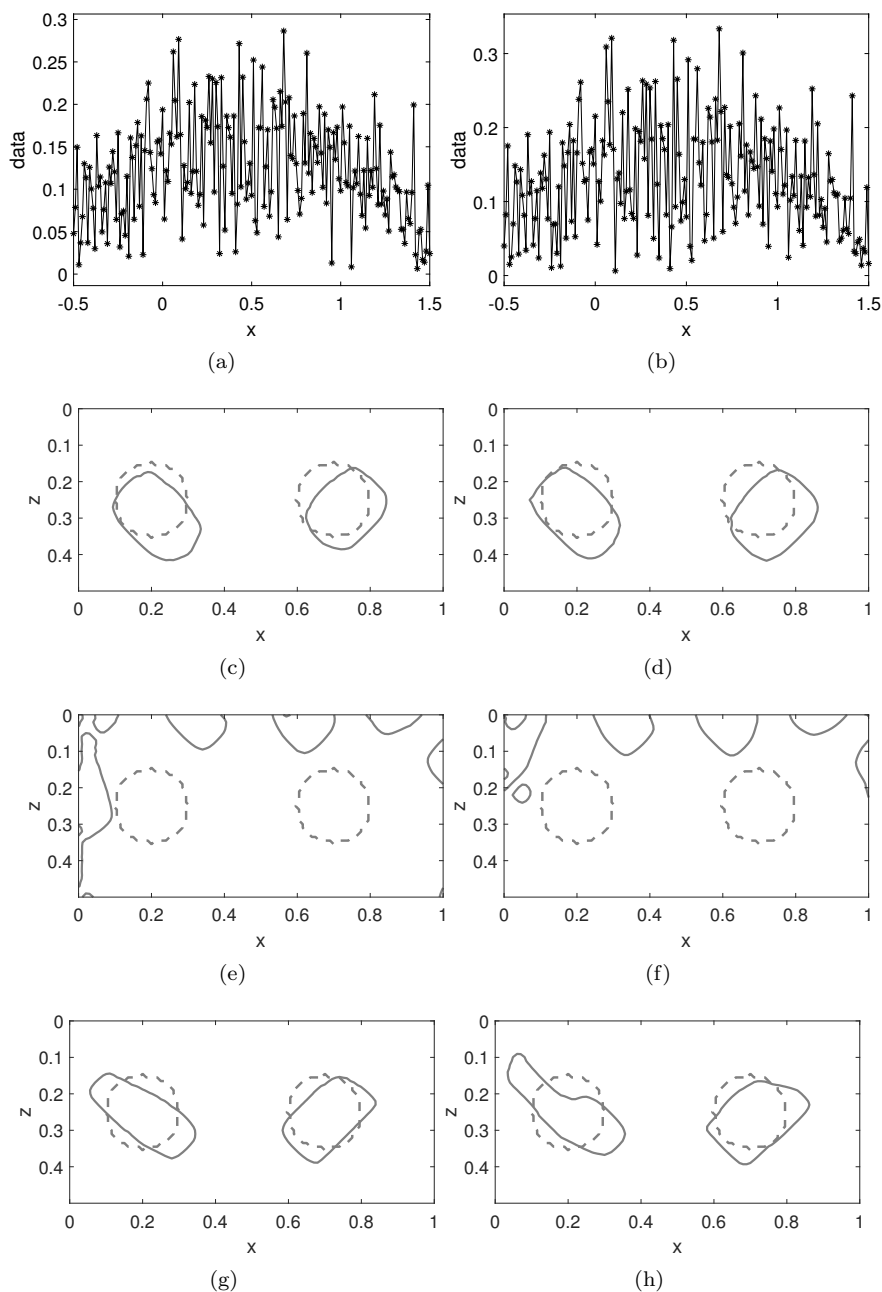


FIGURE 3. Example 1. Recovered solutions. The left column shows the results for the modulus data with 60% Gaussian noises added to the components; the right column shows the results for the modulus data with 80% Gaussian noises added to the components. (a), (b): modulus data with 60% noises and 80% noises, respectively; (c), (d): corresponding solutions using Kantorovich-Rubinstein metric for data fitting; (e), (f): corresponding solutions using L^1 norm for data fitting; (g), (h): corresponding solutions using L^2 norm for data fitting. In (c)-(h), the dashed line plots the exact model for reference.

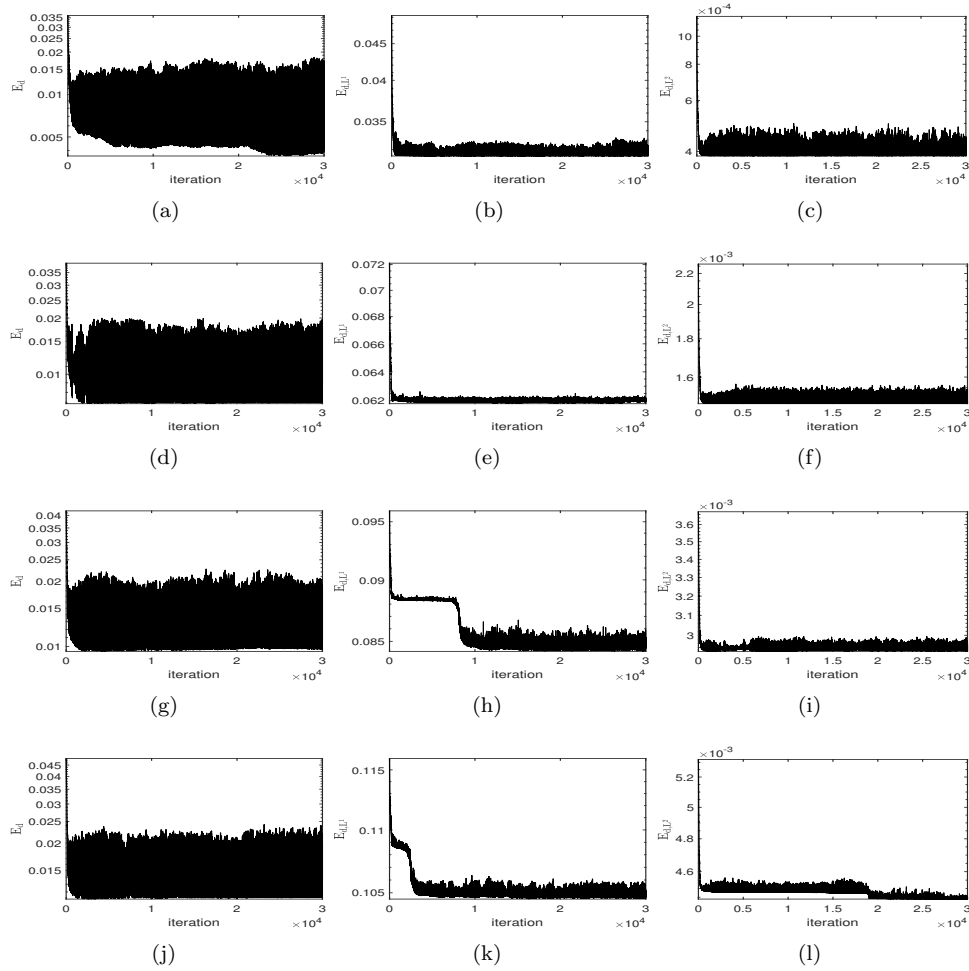


FIGURE 4. Example 1. Evolution of data-discrepancy functions. The first, second, and third column: the K-R metric E_d , the L^1 -norm E_{d,L^1} , and the L^2 -norm E_{d,L^2} , respectively. (a)-(c): the modulus data with 20% Gaussian noises added to the components; (d)-(f): 40% Gaussian noises; (g)-(i): 60% Gaussian noises; (j)-(l): 80% Gaussian noises.

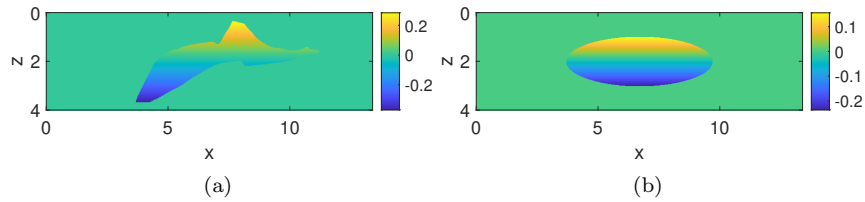


FIGURE 5. Example 2. Salt model with varying density contrast. (a) Exact model of mass distribution; (b) initial guess.

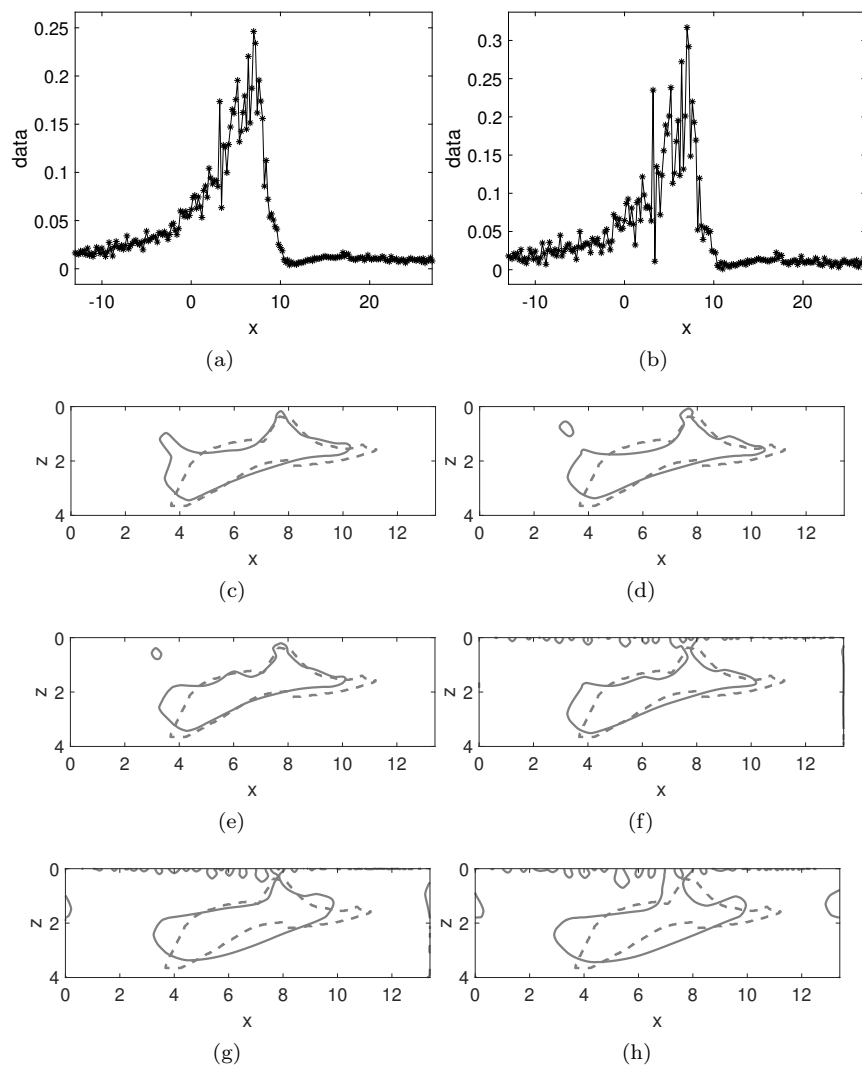


FIGURE 6. Example 2. Recovered solutions. The left column shows the results for the modulus data with 20% Gaussian noises added to the components; the right column shows the results for the modulus data with 40% Gaussian noises added to the components. (a),(b): modulus data with 20% noises and 40% noises, respectively; (c),(d): corresponding solutions using Kantorovich-Rubinstein metric for data fitting; (e),(f): corresponding solutions using L^1 norm for data fitting; (g),(h): corresponding solutions using L^2 norm for data fitting. In (c)-(h), the dashed line plots the exact model for reference.

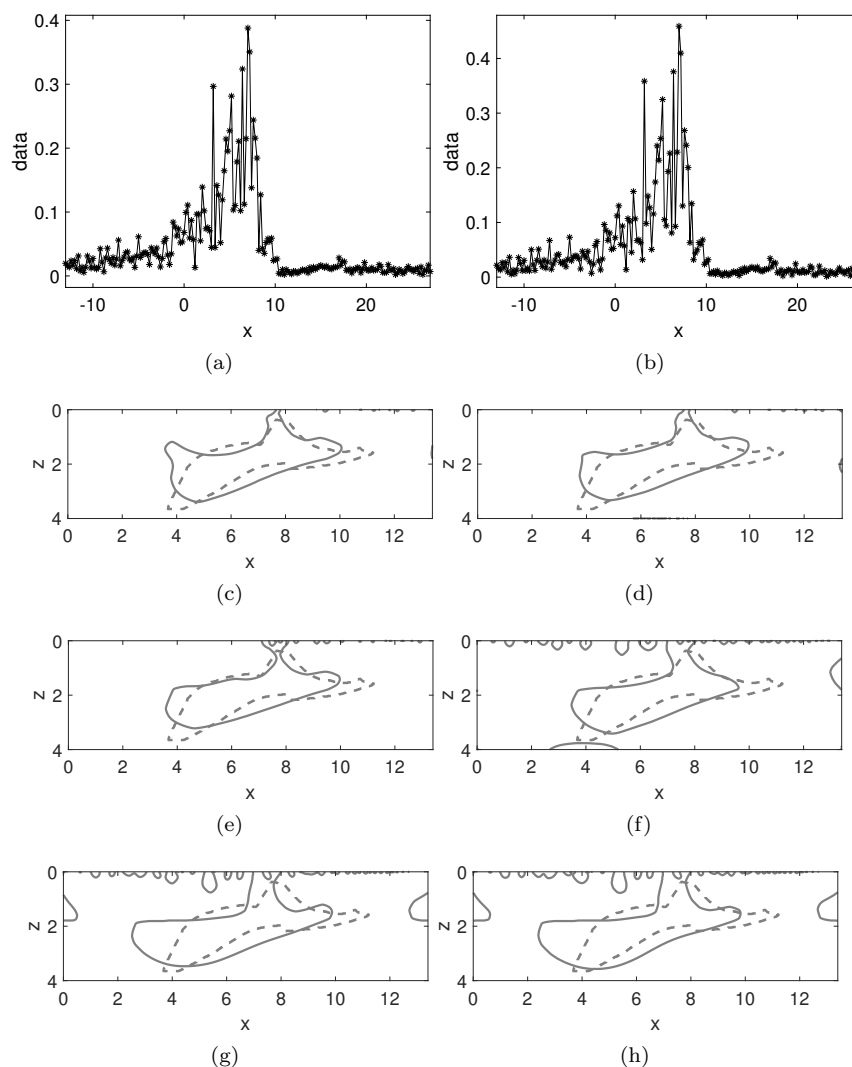


FIGURE 7. Example 2. Recovered solutions. The left column shows the results for the modulus data with 60% Gaussian noises added to the components; the right column shows the results for the modulus data with 80% Gaussian noises added to the components. (a),(b): modulus data with 60% noises and 80% noises, respectively; (c),(d): corresponding solutions using Kantorovich-Rubinstein metric for data fitting; (e),(f): corresponding solutions using L^1 norm for data fitting; (g),(h): corresponding solutions using L^2 norm for data fitting. In (c)-(h), the dashed line plots the exact model for reference.

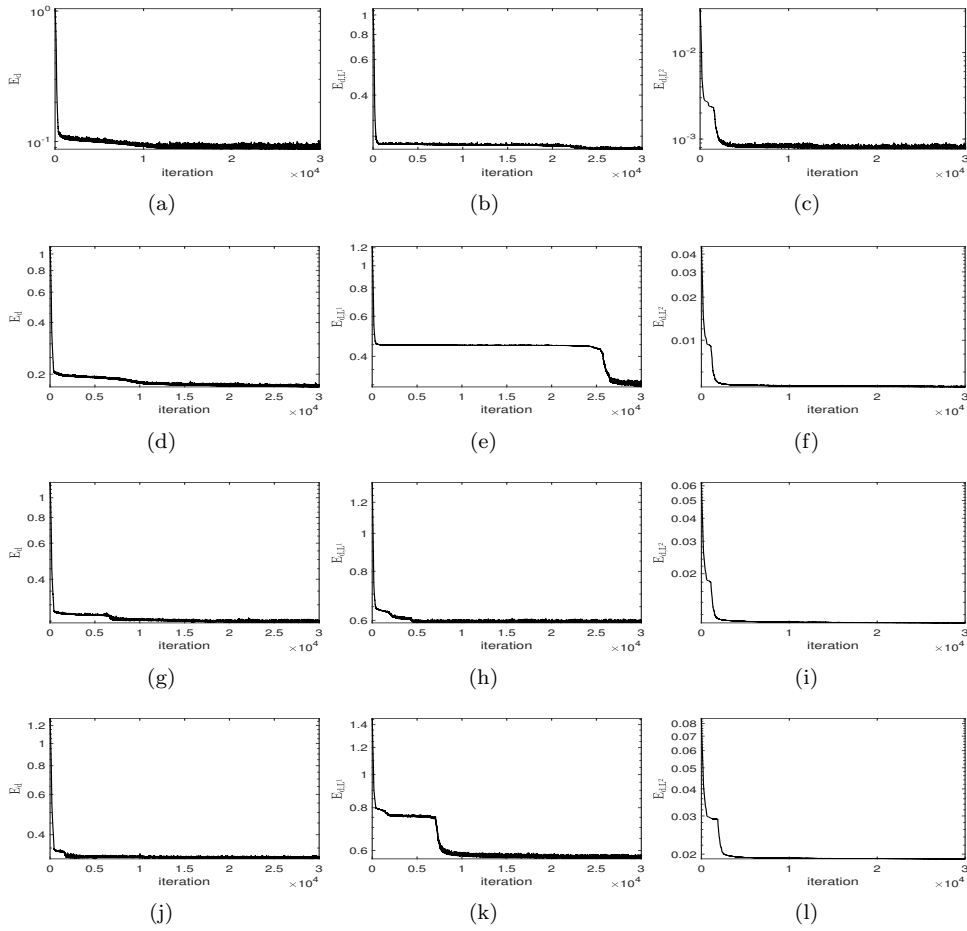


FIGURE 8. Example 2. Evolution of data-discrepancy functions. The first, second, and third column: the K-R metric E_d , the L^1 -norm E_{d,L^1} , and the L^2 -norm E_{d,L^2} , respectively. (a)-(c): the modulus data with 20% Gaussian noises added to the components; (d)-(f): 40% Gaussian noises; (g)-(i): 60% Gaussian noises; (j)-(l): 80% Gaussian noises.

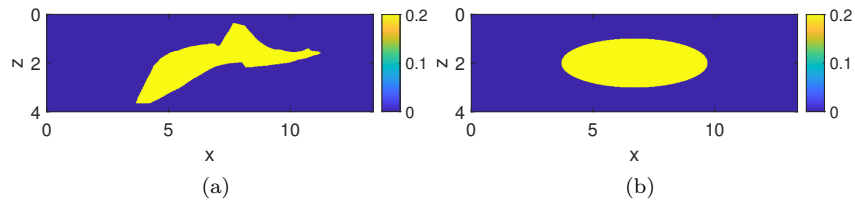


FIGURE 9. Example 3. Salt model with constant density contrast. (a) Exact model of mass distribution; (b) initial guess.

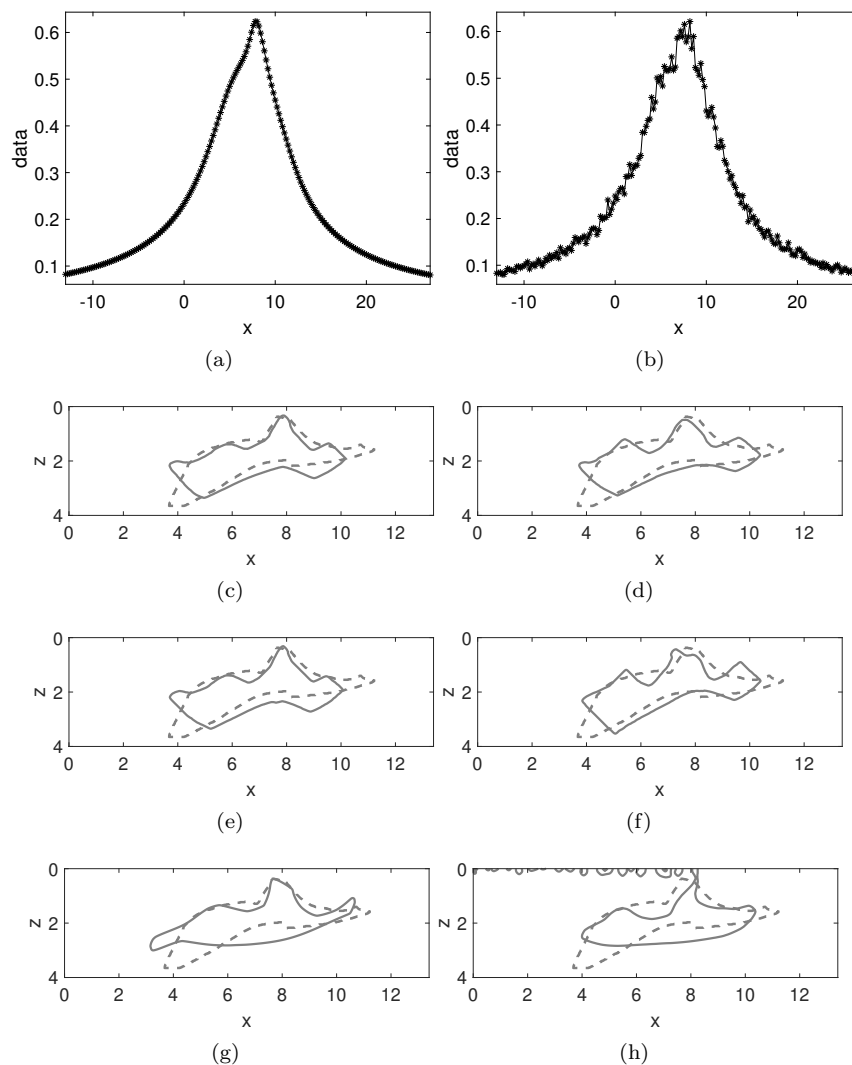


FIGURE 10. Example 3. Recovered solutions. The left column shows the results for the modulus data without noises; the right column shows the results for the modulus data with 5% Gaussian noises added to the components. (a), (b): modulus data without and with 5% noises, respectively; (c), (d): corresponding solutions using Kantorovich-Rubinstein metric for data fitting; (e), (f): corresponding solutions using L^1 norm for data fitting; (g), (h): corresponding solutions using L^2 norm for data fitting. In (c)-(h), the dashed line plots the exact model for reference.

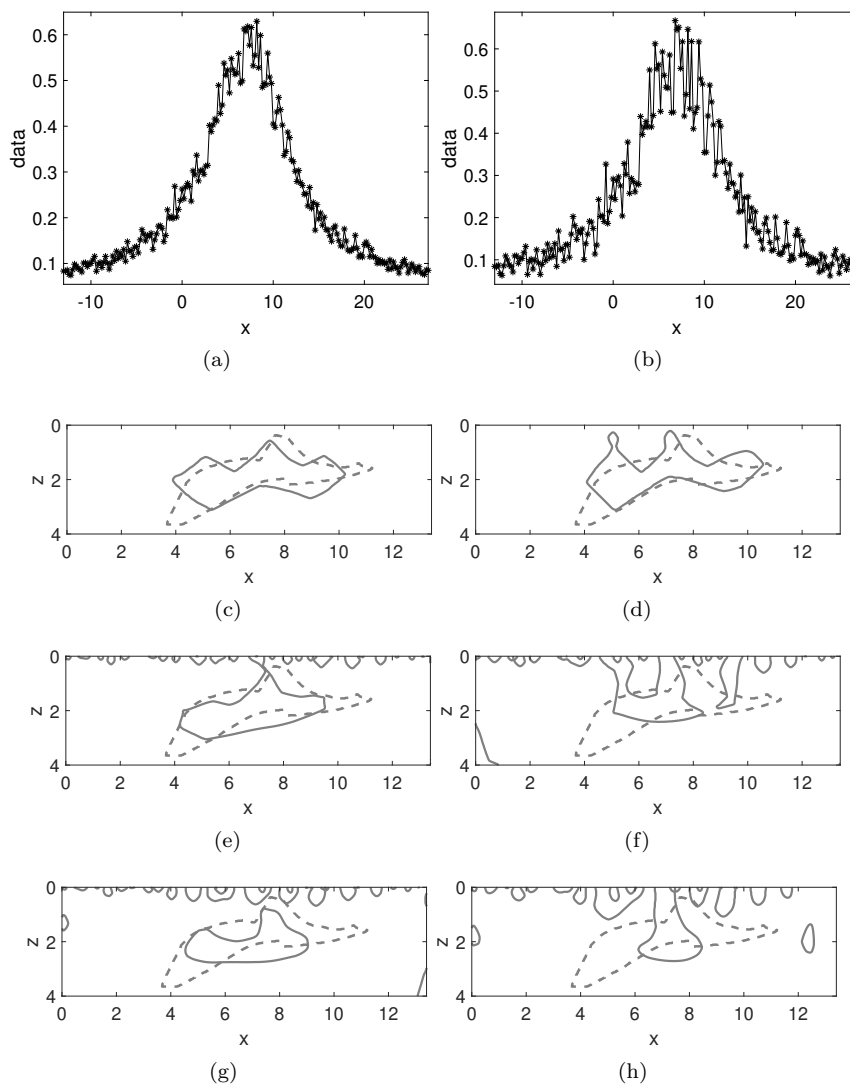


FIGURE 11. Example 3. Recovered solutions. The left column shows the results for the modulus data with 10% Gaussian noises added to the components; the right column shows the results for the modulus data with 20% Gaussian noises added to the components. (a), (b): modulus data with 10% noises and 20% noises, respectively; (c), (d): corresponding solutions using Kantorovich-Rubinstein metric for data fitting; (e), (f): corresponding solutions using L^1 norm for data fitting; (g), (h): corresponding solutions using L^2 norm for data fitting. In (c)-(h), the dashed line plots the exact model for reference.

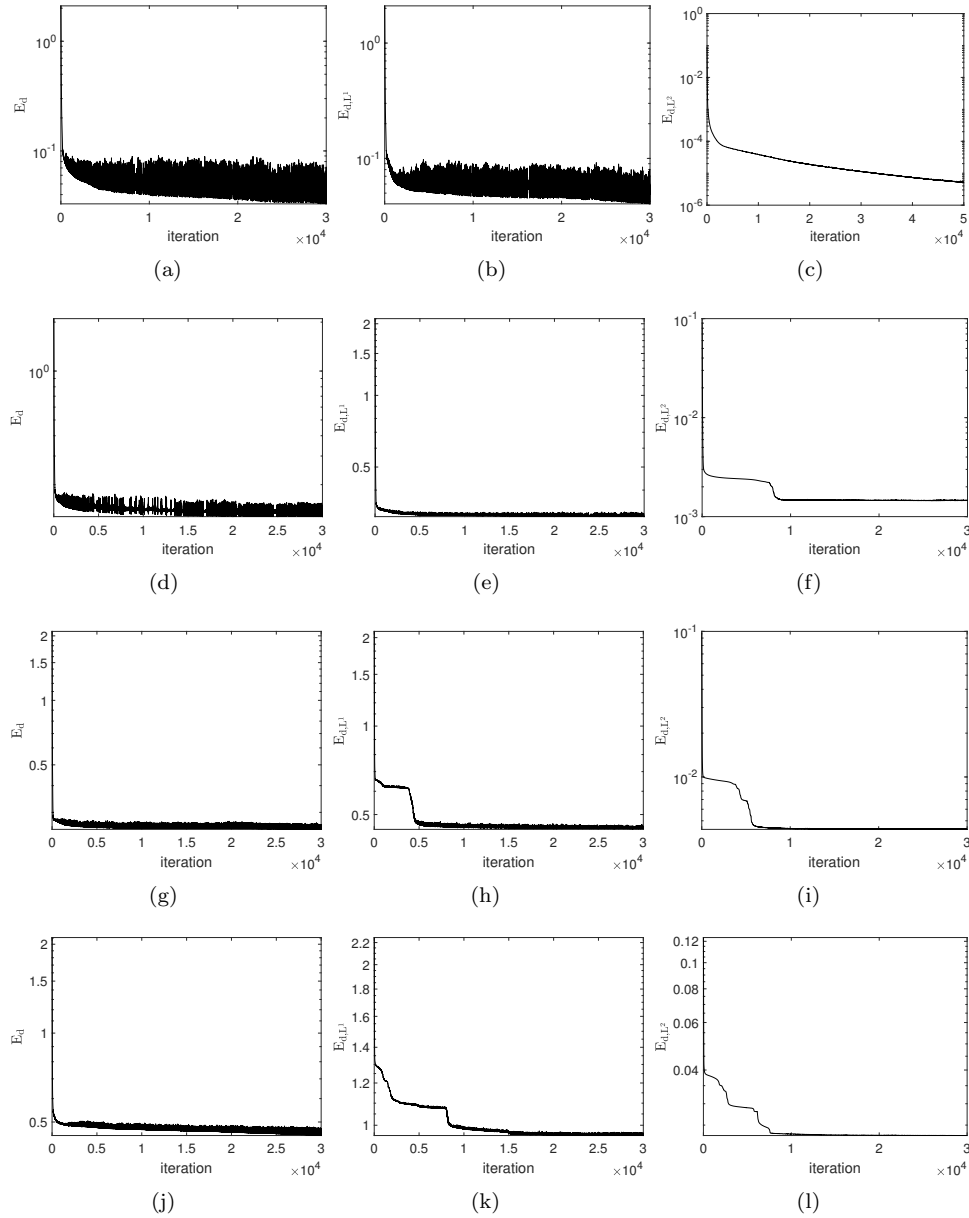


FIGURE 12. Example 3. Evolution of data-discrepancy functions. The first column plots the K-R metric data-discrepancy function E_d , the second column plots the L^1 -norm data-discrepancy function E_{d,L^1} , and the third column plots the L^2 -norm data-discrepancy function E_{d,L^2} . (a)-(c): Evolution of E_d , E_{d,L^1} and E_{d,L^2} for the modulus data without noises; (d)-(f): evolution of E_d , E_{d,L^1} and E_{d,L^2} for the data with 5% Gaussian noises; (g)-(i): evolution of E_d , E_{d,L^1} and E_{d,L^2} for the data with 10% Gaussian noises; (j)-(l): evolution of E_d , E_{d,L^1} and E_{d,L^2} for the data with 20% Gaussian noises.

5. Discussion. The parameter c plays an important role in constructing the K-R metric based data discrepancy function. In numerical computations, reliable inversion results rely on reasonable choices for the values of the parameter c . Here we provide some insights and discussions.

Let $r = d - d^*$ denote the data residual. Consider the situation that $r \in L^1(\mathbb{R})$, and suppose r has a compact support in $\mathbb{R}$. Similar to the discussions in [13], we have the following result,

$$\|r\|_{\text{KR},c} \sim \begin{cases} \|r\|_{L^1} & , \quad \text{as } c \rightarrow +\infty, \\ c \left\| \int_{-\infty}^x r(s) ds \right\|_{L^1} & , \quad \text{as } c \rightarrow 0. \end{cases} \quad (43)$$

Moreover, if the residual $r \in L^1(\mathbb{R})$ satisfies mass conservation, i.e. $\int_{\mathbb{R}} r(s) ds = 0$, we have

$$\|r\|_{\text{KR},c} \sim c \|r\|_{\text{KR},*} \quad \text{as } c \rightarrow 0, \quad (44)$$

where $\|\cdot\|_{\text{KR},*}$ denotes the K-R metric for measuring the distance between two distributions with equal masses. We omit the details on deriving equations (43) and (44), which can be found in [13].

The asymptotic behavior of $\|\cdot\|_{\text{KR},c}$ in $L^1(\mathbb{R})$ gives us some intuitions to select the parameter c . As $c \rightarrow +\infty$, $\|\cdot\|_{\text{KR},c}$ will be equivalent to the L^1 norm; as $c \rightarrow 0$, $\|\cdot\|_{\text{KR},c}$ behaves like the K-R metric for measuring the distance between two distributions with equal masses. It suggests that we should take a reasonably small value for c , so that $\|\cdot\|_{\text{KR},c}$ preserves the good properties of the K-R metric. Recalling the definition of BLip_c in equation (16), c is used to provide bound for ∇v , i.e. $\left| \frac{\partial v}{\partial y_i} \right| \leq c$. In discretization, $\left| \frac{\partial v}{\partial y_i} \right| \doteq \left| \frac{v(y_i + \Delta y_i) - v(y_i)}{\Delta y_i} \right| \leq \frac{2}{\Delta y_i}$, considering that $|v| \leq 1$ as $v \in \text{BLip}_c$. Therefore, taking a reasonably small value for c means that

$$c \ll \frac{2}{\Delta y_i}, \quad \forall i = 1, \dots, n-1, \quad (45)$$

so that c is effective to provide bound for ∇v in the bounded Lipschitz set BLip_c . On the other hand, the value of c can not be too small; otherwise we may require the property of mass conservation in the data residual. A possible strategy is to obtain a preliminary recovered solution and take it as the initial guess. Then hopefully the data residual can roughly satisfy the mass conservation in the inversion process, and we can choose a sufficiently small value for c . This point deserves a further study in the future work.

6. Conclusion. We have developed a Kantorovich-Rubinstein metric based level-set algorithm for the inverse problem of modulus of gravity force. Motivated by the well-posedness theorems of gravity inverse problems in [14], we propose to solve for the unknown domains of volume mass distributions from the measurements of modulus of gravity force. Following the framework of [15], we utilize the level-set method for the domain inverse problem, where the level-set function implicitly represents the boundary of the targeted domain. To develop robust inversion algorithms, we propose to use the Kantorovich-Rubinstein metric for data fitting of the modulus of gravity force. An ADMM algorithm is introduced to efficiently compute the K-R metric based data discrepancy and its Fréchet derivative. We implement the inversion algorithm in synthetic examples, where we compare the solutions using the K-R metric with those using the L^1 norm and L^2 norm for data fitting. The numerical examples demonstrate that the K-R metric based level-set algorithm is able to provide robust and stable solutions under extremely high-level random noises.

The proposed inversion algorithm shows the potential to process field measurement data with large noise contaminations in practical inverse gravimetry problems.

Acknowledgments. Wenbin Li is supported by Natural Science Foundation of Shenzhen (grant no. JCYJ20190806144005645) and NSFC (grant no. 41804096). Qian's research is partially supported by NSF (grants 2012046 and 2152011).

Appendix A. Level-set inversion algorithms using the L^1 norm and the L^2 norm as data discrepancies. The level-set function ϕ is recovered by solving the following optimization problem,

$$\phi := \arg \min_{\phi} E_{d,L^p}(\rho(\phi)), \quad p = 1, 2, \quad (\text{A-1})$$

and we will utilize the gradient descent method as well. Considering the formulas of E_{d,L^1} and E_{d,L^2} ,

$$E_{d,L^1} := \|d(\mathbf{y}; \rho) - d^*(\mathbf{y})\|_{L^1} = \int_{\Sigma_0} |d(\mathbf{y}; \rho) - d^*(\mathbf{y})| \, d\mathbf{s}(\mathbf{y}), \quad (\text{A-2})$$

$$E_{d,L^2} := \|d(\mathbf{y}; \rho) - d^*(\mathbf{y})\|_{L^2}^2 = \frac{1}{2} \int_{\Sigma_0} (d(\mathbf{y}; \rho) - d^*(\mathbf{y}))^2 \, d\mathbf{s}(\mathbf{y}), \quad (\text{A-3})$$

their gradient directions are evaluated in the following way,

$$\begin{aligned} \frac{\partial E_{d,L^1}}{\partial \phi}(\mathbf{x}) &= \frac{\partial E_{d,L^1}}{\partial \rho}(\mathbf{x}) \frac{\partial \rho}{\partial \phi}(\mathbf{x}) = \frac{\partial \rho}{\partial \phi}(\mathbf{x}) \int_{\Sigma_0} \frac{d(\mathbf{y}) - d^*(\mathbf{y})}{|d(\mathbf{y}) - d^*(\mathbf{y})|} \frac{\partial d}{\partial \rho}(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}) \\ &= \gamma f(\mathbf{x}) \delta(\phi(\mathbf{x})) \int_{\Sigma_0} \frac{d(\mathbf{y}) - d^*(\mathbf{y})}{|d(\mathbf{y}) - d^*(\mathbf{y})|} \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}), \end{aligned} \quad (\text{A-4})$$

$$\begin{aligned} \frac{\partial E_{d,L^2}}{\partial \phi}(\mathbf{x}) &= \frac{\partial E_{d,L^2}}{\partial \rho}(\mathbf{x}) \frac{\partial \rho}{\partial \phi}(\mathbf{x}) = \frac{\partial \rho}{\partial \phi}(\mathbf{x}) \int_{\Sigma_0} (d(\mathbf{y}) - d^*(\mathbf{y})) \frac{\partial d}{\partial \rho}(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}) \\ &= \gamma f(\mathbf{x}) \delta(\phi(\mathbf{x})) \int_{\Sigma_0} (d(\mathbf{y}) - d^*(\mathbf{y})) \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}), \end{aligned} \quad (\text{A-5})$$

where we have used equations (7), (11), and (13) in the calculation. Replacing the delta function with the numerical delta function $\delta_\epsilon(\phi) = \chi(T_\epsilon) |\nabla \phi|$ as shown in equation (10), the gradient directions are approximated by

$$\begin{aligned} \frac{\partial E_{d,L^1}}{\partial \phi}(\mathbf{x}) &= \gamma f(\mathbf{x}) \chi(T_\epsilon) |\nabla \phi| \int_{\Sigma_0} \frac{d(\mathbf{y}) - d^*(\mathbf{y})}{|d(\mathbf{y}) - d^*(\mathbf{y})|} \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}), \\ \frac{\partial E_{d,L^2}}{\partial \phi}(\mathbf{x}) &= \gamma f(\mathbf{x}) \chi(T_\epsilon) |\nabla \phi| \int_{\Sigma_0} (d(\mathbf{y}) - d^*(\mathbf{y})) \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}), \end{aligned}$$

which can be denoted as

$$\frac{\partial E_{d,L^p}}{\partial \phi}(\mathbf{x}) = V_{n,p}(\mathbf{x}) |\nabla \phi(\mathbf{x})|, \quad p = 1, 2, \quad (\text{A-6})$$

with

$$V_{n,1}(\mathbf{x}) := \gamma f(\mathbf{x}) \chi(T_\epsilon) \int_{\Sigma_0} \frac{d(\mathbf{y}) - d^*(\mathbf{y})}{|d(\mathbf{y}) - d^*(\mathbf{y})|} \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}), \quad (\text{A-7})$$

$$V_{n,2}(\mathbf{x}) := \gamma f(\mathbf{x}) \chi(T_\epsilon) \int_{\Sigma_0} (d(\mathbf{y}) - d^*(\mathbf{y})) \sum_{i=1}^n \frac{g_i(\mathbf{y})}{d(\mathbf{y})} K_i(\mathbf{y}, \mathbf{x}) \, d\mathbf{s}(\mathbf{y}). \quad (\text{A-8})$$

In equation (A-7), a small constant $\epsilon_0 > 0$ is added to the denominator $|d(\mathbf{y}) - d^*(\mathbf{y})|$ to avoid instability in the numerical computation, e.g. $\epsilon_0 = 10^{-8}$.

The level-set function ϕ is updated according to the negative gradient descent direction, which has the same formulation as equation (37),

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E_{d,p}}{\partial \phi} = -V_{n,p} |\nabla \phi|, \quad p = 1, 2. \quad (\text{A-9})$$

Correspondingly, the step-size Δt for the updating equation should satisfy the Courant-Friedrichs-Lewy (CFL) condition as shown in equation (38). And in practice, we take the step-size Δt according to equation (39),

$$\Delta t = \beta \frac{\min\{\Delta x_1, \dots, \Delta x_n\}}{\max_{\mathbf{x}} |V_{n,p}(\mathbf{x})|}, \quad p = 1, 2. \quad (\text{A-10})$$

where $\beta \in (0, 1)$ is a constant, e.g. $\beta = 0.5$ in our numerical computation. To be short, the same step-size rule is used to evolve the level-set function ϕ for all the three data-fitting strategies.

REFERENCES

- [1] J. E. Bain, T. R. Horscroft, J. Weyand, A. H. Saad and D. N. Bulling, Complex salt features resolved by integrating seismic, gravity and magnetics, *EAEG/EAPG Annual Meeting, Expanded Abstracts*, 1993.
- [2] J. D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, **84** (2000), 375-393.
- [3] J. D. Benamou and Y. Brenier, Mixed L^2 -Wasserstein optimal mapping between prescribed density functions, *J. Optim. Theory Appl.*, **111** (2001), 255-271.
- [4] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Commun. Pure Appl. Math.*, **44** (1991), 375-417.
- [5] M. Burger and S. Osher, A survey on level set methods for inverse problems and optimal design, *European J. Appl. Math.*, **16** (2005), 263-301.
- [6] O. Dorn and D. Lesselier, Level set methods for inverse scattering, *Inverse Problems*, **22** (2006), 67-131.
- [7] B. Engquist and B. D. Froese, Application of the Wasserstein metric to seismic signals, *Commun. Math. Sci.*, **12** (2014), 979-988.
- [8] B. Engquist, B. D. Froese and Y. Yang, Optimal transport for seismic full waveform inversion, *Commun. Math. Sci.*, **14** (2016), 2309-2330.
- [9] L. C. Evans, Partial differential equations and Monge-Kantorovich mass transfer, *Current Developments in Mathematics*. International Press, (1999), 65-126.
- [10] D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, *Computers & Mathematics with Applications*, **2** (1976), 17-40.
- [11] S. Hou, K. Solna and H.-K. Zhao, Imaging of location and geometry for extended targets using the response matrix, *J. Comput. Phys.*, **199** (2004), 317-338.
- [12] G. Huang and J. Qian, Analysis of Kantorovich-Rubinstein norm and its application to inverse gravity problem, *SIAM J. Imaging Sci.*, **12** (2019), 1528-1560.
- [13] G. Huang, X. Zhang and J. Qian, Kantorovich-Rubinstein misfit for inverting gravity-gradient data by the level-set method, *Geophysics*, **84** (2019), 55-73.
- [14] V. Isakov, *Inverse Source Problems*, American Mathematical Society, Providence, Rhode Island, 1990.
- [15] V. Isakov, S. Leung and J. Qian, A fast local level set method for inverse gravimetry, *Commun. Comput. Phys.*, **10** (2011), 1044-1070.
- [16] V. Isakov, S. Leung and J. Qian, A three-dimensional inverse gravimetry problem for ice with snow caps, *Inverse Probl. Imaging*, **7** (2013), 523-544.
- [17] G. J. Jorgensen and J. L. Kisabeth, Joint 3-D inversion of gravity magnetic and tensor gravity fields for imaging salt formations in the deepwater gulf of mexico, In *Expanded Abstracts*, (2000), 424-426.

- [18] R. A. Krahenbuhl and Y. Li, Inversion of gravity data using a binary formulation, *Geophysical Journal International*, **167** (2006), 543-556.
- [19] D. Leahy, *Inverse Problems Using Gravity Gradiometry*, PhD thesis, University of Manchester, England, 2019.
- [20] J. Lellmann, D. Lorenz, C. Schonlieb and T. Valkonen, [Imaging with Kantorovich-Rubinstein discrepancy](#), *SIAM J. Imaging Sci.*, **7** (2014), 2833-2859.
- [21] J. P. Leveille, I. F. Jones, Z. Z. Zhou, B. Wang and F. Liu, Subsalt imaging for exploration, production, and development: A review, *Geophysics*, **76** (2011), 3-20.
- [22] W. Li and S. Leung, A fast local level set based adjoint state method for first arrival transmission traveltimes tomography with discontinuous slowness, *Geophysical Journal International*, **195** (2013), 582-596.
- [23] W. Li, S. Leung and J. Qian, A level set-adjoint state method for crosswell transmission-reflection traveltimes tomography, *Geophys. J. Internat.*, **199** (2014), 348-367.
- [24] W. Li, W. Lu and J. Qian, A level set method for imaging salt structures using gravity data, *Geophysics*, **81** (2016), 35-51.
- [25] W. Li, W. Lu, J. Qian and Y. Li, A multiple level set method for three-dimensional inversion of magnetic data, *Geophysics*, **82** (2017), 61-81.
- [26] W. Li and J. Qian, Joint inversion of gravity and traveltimes data using a level-set based structural parameterization, *Geophysics*, **81** (2016), 107-119.
- [27] W. Li and J. Qian, [Simultaneously recovering both domain and varying density in inverse gravimetry by efficient level-set methods](#), *Inverse Probl. Imaging*, **15** (2021), 387-413.
- [28] W. Lu, S. Leung and J. Qian, [An improved fast local level set method for three-dimensional inverse gravimetry](#), *Inverse Probl. Imaging*, **9** (2015), 479-509.
- [29] W. Lu and J. Qian, A local level set method for three-dimensional inversion of gravity gradiometry data, *Geophysics*, **80** (2015), 35-51.
- [30] S. Osher and J. A. Sethian, [Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations](#), *J. Comput. Phys.*, **79** (1988), 12-49.
- [31] F. Santambrogio, *Optimal Transport for Applied Mathematics*, Birkhauser, 2015.
- [32] F. Santosa, [A level-set approach for inverse problems involving obstacles](#), *ESAIM Contrôle Optim. Calc. Var.*, **1** (1996), 17-33.
- [33] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003.
- [34] B. Wahlberg, S. Boyd, M. Annergren and Y. Wang, An ADMM algorithm for a class of total variation regularized estimation problems, *IFAC Proceedings*, **45** (2012), 83-88.
- [35] H.-K. Zhao, T. Chan, B. Merriman and S. J. Osher, [A variational level set approach for multiphase motion](#), *J. Comput. Phys.*, **127** (1996), 179-195.
- [36] P. Zhiglova, C. G. Farquharson and C. A. Hurich, 2-D reconstruction of boundaries with level set inversion of traveltimes, *Geophysical Journal International*, **192** (2013), 688-698.

Received December 2021; revised September 2022; early access October 2022.