Inverse Problems

PAPER

Least-squares method for recovering multiple medium parameters

To cite this article: Kazufumi Ito *et al* 2022 *Inverse Problems* **38** 125004

IOP Publishing Inverse Problems

Inverse Problems **38** (2022) 125004 (23pp) https://doi.org/10.1088/1361-6420/ac999a

Least-squares method for recovering multiple medium parameters

Kazufumi Ito[1], **Ying Liang**[2] **and Jun Zou**[3,*]

[1] Department of Mathematics and Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC 27695, United States of America
[2] Department of Mathematics, Purdue University, West Lafayette, IN 47907, United States of America
[3] Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, People's Republic of China

E-mail: kito@unity.ncsu.edu, liang402@purdue.edu and zou@math.cuhk.edu.hk

Received 24 April 2022, revised 27 September 2022
Accepted for publication 12 October 2022
Published 28 October 2022

Abstract

We present a two-stage least-squares method for inverse medium problems of reconstructing multiple unknown coefficients simultaneously from noisy data. A direct sampling method is applied to detect the location of the inhomogeneity in the first stage, while a total least-squares method with a mixed regularization is used to recover the medium profile in the second stage. The total least-squares method is designed to minimize the residual of the model equation and the data fitting, along with an appropriate regularization, in an attempt to significantly improve the accuracy of the approximation obtained from the first stage. We shall also present an analysis on the well-posedness and convergence of this algorithm. Numerical experiments are carried out to verify the accuracies and robustness of this novel two-stage least-squares algorithm, with high tolerance of noise in the data.

Keywords: inverse medium problem, multiple parameters, least-squares method, mixed regularization

(Some figures may appear in colour only in the online journal)

*Author to whom any correspondence should be addressed.

[4] The work of this author was substantially supported by Hong Kong RGC General Research Fund (Projects 14306921 and 14306719).

1. Introduction

We are concerned in this work with recovering multiple medium parameters in a forward model of the general form

$$L(u, q) = g, \tag{1.1}$$

where L is a bilinear operator on $X \times Z$, $u \in X$ is the state variable, while $q \in Z$ represents one or multiple parameters in the model. The inverse problem of our interest is to recover the medium parameters q, under some measurement data of $f := Cu \in Y$. Here X, Y and Z are three Hilbert spaces, and C is an observation map from X to Y.

There are many applications where it is necessary to recover multiple parameters or coefficients simultaneously, for instance, the diffusive optical tomography (DOT) and the inverse electromagnetic medium problem. DOT aims at recovering the diffusion and absorption coefficients σ and μ from the governing equation [1, 2]:

$$-\nabla \cdot (\sigma(x)\nabla u) + \mu(x)\, u = 0 \quad \text{in } \Omega, \tag{1.2}$$

using the Cauchy data (f, h) collected at the boundary Γ of Ω:

$$f = u|_{\Gamma}, \qquad h = \frac{\partial u}{\partial \nu}|_{\Gamma}. \tag{1.3}$$

Another example would be the inverse electromagnetic medium problem to recover the unknown magnetic and electric coefficients μ and λ in the Maxwell's system [3, 4]:

$$\begin{aligned} \nabla \times \vec{H} + i\omega\mu(x)\,\vec{E} = 0 \quad \text{in } \Omega, \\ \nabla \times \vec{E} - i\omega\lambda(x)\,\vec{H} = 0 \quad \text{in } \Omega, \end{aligned}$$

using some measurement data of the electric or magnetic field $\vec{E}$ or $\vec{H}$.

The inverse reconstruction of multiple medium coefficients is generally much more technical and difficult than the single coefficient case. We shall propose a total least-squares formulation with appropriate regularization to transform the inverse problem into an optimization problem. The total least-squares strategy is actually not uncommon. One conventional approach for inverse medium problems is to seek an optimal parameter q from a feasible set $K \subset Z$ such that it minimizes an output least-squares functional j of the form

$$j(q) = \|Cu(q) - f\|_Y^2 + \alpha\, \psi(q), \tag{1.4}$$

where $u(q)$ solves the forward model (1.1) when q is given, ψ is a regularization term and $\alpha > 0$ is a regularization parameter. We refer the readers to [2, 5, 6] for more details about this traditional approach. A relaxed variational method of the least-squares form was proposed and studied in [7, 8] for the impedance computed tomography. The least-squares functional consists of a residual term associated with the governing equation while the measurement data is enforced at the feasible set. Different from the aforementioned approaches, we shall follow some basic principle of a total least-squares approach from [9] and treat the governing equation and the data fitting separately, along with a regularization. That is, we look for an optimal parameter q from Z and a state variable u from X together such that they minimize an extended functional of the form

$$J(u,q)=\|L(u,q)-g\|_X^2+\|Cu-f\|_Y^2+\alpha\,\psi(q). \tag{1.5}$$

This functional combines the residual of model equation, data fitting and constraints on parameters in the least-squares criterion. The combination in (1.5) results in a regularization effect to treat the model equation and allows a robustness as a quasi-reversibility method [9, 10]. Solving the minimization problem (1.4) with the Lagrangian approach is similar to our least-squares approach. Compared with the conventional approaches, the domain of $J(u,q)$ is much more regular and the semi-norm defined by the formulation is much stronger. More precisely, the total least-squares formulation aims to find a solution pair (u,q) simultaneously in a smooth class and is less sensitive to the noise and uncertainty in the inverse model. Another important feature of this formulation is that the functional $J(u,q)$ is quadratical and convex with respect to each variable u and q, if the regularization ψ is chosen to be quadratical and convex, while the traditional one $j(q)$ in (1.4) is highly nonlinear and non-convex in general. This special feature facilitates us naturally to minimize the functional $J(u,q)$ effectively by some natural alternating direction iterative (ADI) method so that only two quadratical and convex suboptimizations are needed to solve at each iteration in terms of the variables u and q respectively.

In addition to the functional (1.5) that uses the residual of the forward model (1.1), we will also address another least-squares functional that makes use of the equivalent first-order system of the forward model (1.1) and replaces the first term in (1.5) by the residuals of the corresponding first-order system. Using the first-order system has been the fundamental idea in modern least-squares methods in solving second-order PDEs [13–15]. The advantages of using first-order formulations are much more significant to the numerical solutions of inverse problems, especially when we aim at simultaneously reconstructing multiple coefficients as we do in this work. First, the multiple coefficients appear in separated first-order equations, hence are naturally decoupled. This would greatly reduce the nonlinearity and enhance the convexity of the resulting optimization systems. Second, the first-order formulation relaxes the regularity requirement of the solutions in the resulting analysis.

A crucial step to an effective reconstruction of multiple coefficients is to seek reasonable initial approximations to capture some general (possibly rather inaccurate) geometric and physical profiles of all the unknown multiple coefficients. This is a rather technical and difficult task in numerical reconstruction processes. Several effective approaches have been proposed recently to help find some reasonable initial approximations for solving the multidimensional hyperbolic coefficient inverse problems; see, e.g., [25–28]. In this work we shall propose to adopt the direct sampling-type methods (DSMs) that we have been developing in recent years (cf [16, 17, 23, 24]). Using the index functions provided by DSM, we shall determine a computational domain that is often much smaller than the original physical domain, then the restricted index functions on the computational domain serve as the initial guesses of the unknown coefficients. In this work, we will apply a newly developed DSM [18], where two groups of probing and index functions are constructed to identify and decouple the multiple inhomogeneous inclusions of different physical nature, which is principally different from the classical DSMs targeting the inhomogeneous inclusions from one single physical nature. As we shall see, DSMs turn out to be very effective, fast and reliable to provide some reasonable initial approximations.

The rest of the paper is structured as follows. In section 2, we justify the well-posedness of the least-squares formulation for the general inverse medium problems. In section 3, we propose an ADI method for solving the minimization problem and prove the convergence of the iterative method. We illustrate in section 4 how this total least-squares method applies to a concrete inverse problem, by taking the DOT problem as a benchmark problem. We present numerical results in section 5 for a couple of different types of inhomogeneous coefficients for the DOT problem to demonstrate the stability and effectiveness of this proposed method. Throughout the paper, c, c_0 and c_1 denote generic constants which may differ at each occurrence.

2. Well-posedness of the least-squares formulation for inverse medium problems

Recall that we proposed in section 1 the following least-squares formulation

$$\min_{u\in X,q\in Z} J(u,q) = \|L(u,q)-g\|_X^2 + \|Cu-f\|_Y^2 + \alpha\,\psi(q) \tag{2.1}$$

to solve the inverse medium problems modeled by (1.1). This section is devoted to the well-posedness of the total least-squares formulation (2.1), namely, the existence of a solution to (2.1) and the conditional stability of the reconstruction with respect to the measurement. To provide a mathematical justification of the well-posedness, we present several assumptions on the least-squares formulation. We will verify these assumptions in section 4 for a concrete example of such inverse medium problems.

Let us first introduce several notation. For simplicity, for a given $q\in Z$ (resp.$u\in X$), we will write L_q (resp.Φ_u) as

$$L_q u := L(u,q) \quad (\text{resp. } \Phi_u q := L(u,q)). \tag{2.2}$$

We denote the subdifferential of the regularization term ψ at q by $\partial\psi(q)$, and denote the inner products of the Hilbert spaces X, Y and Z by $(\cdot,\cdot)_X$, $(\cdot,\cdot)_Y$ and $\langle\cdot,\cdot\rangle$ respectively.

2.1. Existence of a minimizer

We present the following assumptions on the regularization term ψ and operators L and C in the forward model:

Assumption 1. The regularization term ψ is strictly convex and weakly lower semicontinuous. Furthermore, ψ is also coercive [6], i.e., $\psi(q)\geqslant c\|q\|_Z^2$.

This assumption implies that the level set $\{q\in Z : \psi(q)\leqslant c_0\}$ defines a bounded set in Z.

Assumption 2. Given a constant c_0, for q in the level set $\{q\in Z : \psi(q)\leqslant c_0\}$, $L_q : dom(L)\to X$, where $dom(L)\subset\{u\in X : L(u,q)\in X$ and $Cu\in Y$ for all $q\in Z$ satisfying $\psi(q)\leqslant c_0\}$, is a closed linear operator and is uniformly coercive, i.e., the graph norm $|u|_{W,q} := \|L(u,q)\|_X$ satisfies $|u|_{W,q}\geqslant c_1\|u\|_X$ uniformly in q for some constant $c_1>0$, and thus $L(u,q)=g\in X$ has a unique solution in $dom(L)$.

Under assumption 2, we can define the inverse operator $L_q^{-1} : X\to dom(L)$, which is uniformly bounded by the coercivity of L_q. We also need the following assumption on the sequentially closedness of operators L and C.

Assumption 3. The operators L and C are weakly sequentially closed, i.e., if a sequence $\{(u_n,q_n)\}_{n=1}^\infty$ converges to (u,q) weakly in $X\times Z$, then the sequence $\{L(u_n,q_n)\}_{n=1}^\infty$ converges to $L(u,q)$ weakly in X and the sequence $\{Cu_n\}_{n=1}^\infty$ converges to Cu weakly in Y.

Then we can verify the existence of the minimizers to the least-squares formulation (2.1).

Theorem 1. *Under assumptions 1–3, there exists a minimizer $(u^\star,q^\star)$ in $X\times Z$ of the least-squares formulation (2.1).*

Proof. Since X and Z are nonempty, there exists a minimizing sequence $\{(u_n, q_n)\}_{n=1}^{\infty}$ in $X \times Z$ such that

$$\lim_{n\to\infty} J(u_n, q_n) = \inf_{(u,q)\in X\times Z} J(u, q). \tag{2.3}$$

By assumptions 1 and 2, ψ is a coercive functional and the graph norm $|\cdot|_{W,q}$ is uniformly coercive, thus it follows from (1.5) that the sequence $\{(u_n, q_n)\}_{n=1}^{\infty}$ is uniformly bounded. Then there exists a subsequence of $\{(u_n, q_n)\}_{n=1}^{\infty}$, still denoted by $\{(u_n, q_n)\}_{n=1}^{\infty}$, and some $(u^\star, q^\star) \in X \times Z$ such that u_n converges to $u^\star$ weakly in X and q_n converges to $q^\star$ weakly in Z. As L and C are weakly sequentially closed by assumption 3, there hold that

$$\begin{aligned} &L(u_n, q_n) \text{ converges to } L(u^\star, q^\star) \text{ weakly in } X;\\ &\quad Cu_n \text{ converges to } Cu^\star \text{ weakly in } Y. \end{aligned} \tag{2.4}$$

From the weak lower semicontinuity of the norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, we have

$$\|L(u^\star, q^\star) - g\|_X^2 + \|Cu^\star - f\|_Y^2 \leqslant \lim_{n\to\infty} \inf\left(\|L(u_n, q_n) - g\|_X^2 + \|Cu_n - f\|_Y^2\right).$$

Together with the lower semicontinuity of the regularization term ψ, we can deduce that

$$J(u^\star, q^\star) \leqslant \lim_{n\to\infty} \inf J(u_n, q_n).$$

Hence it follows from (2.3) that $(u^\star, q^\star)$ is indeed a minimizer of the functional J in $X \times Z$. □

2.2. *Conditional stability*

In this subsection, we present some conditional stability estimates of the total least-squares formulation (2.1) for the general inverse medium problems. First we introduce two pairs $(\bar{u}, \bar{q})$ and $(\bar{g}, \bar{f})$ that satisfy

$$L(\bar{u}, \bar{q}) = \bar{g}, \quad C\bar{u} = \bar{f}. \tag{2.5}$$

Letting $(u^\star, q^\star)$ be the unique minimizer of (2.1) in a neighborhood of $(\bar{u}, \bar{q})$, we study the approximation error $(\delta u, \delta q) := (u^\star - \bar{u}, q^\star - \bar{q})$ to illustrate the stability of the least-squares formulation (2.1) with respect to the measurement f and also the term g in the governing equation (1.1). Denote the residual of the governing equation by $\epsilon_1 := L(u^\star, q^\star) - g$. As $(u^\star, q^\star)$ is the local minimizer of functional J in (1.5), we have $J(u^\star, q^\star) \leqslant J(\bar{u}, \bar{q})$. Therefore, by the definition of J we have the inequality that

$$\|\epsilon_1\|_X^2 + \|C(u^\star - \bar{u}) - (f - \bar{f})\|_Y^2 + \alpha\psi(q^\star) \leqslant \|g - \bar{g}\|_X^2 + \|f - \bar{f}\|_Y^2 + \alpha\psi(\bar{q}), \tag{2.6}$$

which directly leads to the following observation on $\psi(q^\star)$:

$$\alpha\psi(q^\star) \leqslant \|g - \bar{g}\|_X^2 + \|f - \bar{f}\|_Y^2 + \alpha\, \psi(\bar{q}). \tag{2.7}$$

If ψ is coercive, (2.7) provides a rough estimate of the reconstruction $q^\star$ with respect to the data noise.

We can further derive an estimate of the approximation error δq, under the following assumption on the operator L.

Assumption 4. There exists a norm $\|\cdot\|_W$ on Z and a subspace S of Z such that $\delta q \in S$ and it holds for any $q \in \bar{q} + S$ that

$$\|CL_q^{-1}\Phi_{\bar{u}}(q-\bar{q})\|_Y^2 \geqslant \|q-\bar{q}\|_W^2.$$

Assumption 4 is satisfied when S is a subspace of Z with finite rank. We can now deduce the following result of the approximation error δq.

Lemma 1. *Under assumptions 1–4, the approximation error δq is bounded in W-norm by the data noise and the regularization term, i.e.,*

$$\|\delta q\|_W^2 + \alpha\psi(q^\star) \leqslant c_0\big(\|g-\bar{g}\|_X^2 + \|f-\bar{f}\|_Y^2 + \alpha\,\psi(\bar{q})\big). \tag{2.8}$$

Proof. Using the bilinear property of L, one can rewrite the difference $L(u^\star, q^\star) - L(\bar{u}, \bar{q})$ as

$$L(u^\star, q^\star) - L(\bar{u}, \bar{q}) = L_{q^\star}(\delta u) + \Phi_{\bar{u}}(\delta q). \tag{2.9}$$

By assumption 2, L_q admits an inverse operator L_q^{-1} from X to $dom(L)$, which, together with (2.5) and (2.9) and the definition of ϵ_1, implies

$$\delta u = L_{q^\star}^{-1}(\epsilon_1 + g - \bar{g} - \Phi_{\bar{u}}(\delta q)). \tag{2.10}$$

Plugging (2.10) into (2.6) leads to an inequality:

$$\begin{aligned} &\|CL_{q^\star}^{-1}\Phi_{\bar{u}}\delta q - (\epsilon_1 + L_{q^\star}^{-1}(g-\bar{g})) + f - \bar{f}\|_Y^2 + \|\epsilon_1\|_X^2 + \alpha\,\psi(q^\star) \\ &\qquad \leqslant \|g-\bar{g}\|_X^2 + \|f-\bar{f}\|_Y^2 + \alpha\,\psi(\bar{q}). \end{aligned} \tag{2.11}$$

It follows from assumption 4 that there exists a norm $\|\cdot\|_W$ such that

$$\|CL_q^{-1}\Phi_{\bar{u}}\delta q\|_Y^2 \geqslant \|\delta q\|_W^2. \tag{2.12}$$

Then we can deduce from (2.11), the triangle inequality, the boundedness of L_q^{-1} and (2.12) that

$$\|\delta q\|_W^2 + \alpha\,\psi(q^\star) \leqslant c_0(\|g-\bar{g}\|_X^2 + \|f-\bar{f}\|_Y^2 + \alpha\,\psi(\bar{q})),$$

where c_0 is a constant, which completes the proof. □

The rest of this section is devoted to verifying the consistency of the least-squares formulation (2.1) as the noise level of measurement goes to zero, which is an essential property of a regularization scheme. If we choose an appropriate regularization parameter α according to the noise level of the data, we can deduce the convergence result of the reconstructed coefficients associated with the regularization parameter α. More precisely, given a set of exact data $(\bar{g}, \bar{f})$, we consider a parametric family $\{(g_\alpha, f_\alpha)\}$ such that $\|g_\alpha - \bar{g}\|_X^2 + \|f_\alpha - \bar{f}\|_Y^2 = o(\alpha)$. In the rest of this section, we denote the functional J in (1.5) with $g = g_\alpha$ and $f = f_\alpha$ by J_α, and the minimizer of J_α by (u_α, q_α). Then we justify the consistency of the least-squares formulation (2.1) by proving the convergence of the sequence of minimizers $\{q_\alpha\}$ to the minimum norm solution [6] of the system (2.5) as $\alpha \to 0$.

Theorem 2. *Let $\{\alpha_n\}_{n=1}^\infty \subset \mathbb{R}^+$ be a sequence converging to zero, and $\{(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^\infty$ be the corresponding sequence of minimizers of J_{α_n}. Then under assumptions 1–3, $\{(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^\infty$ has a subsequence that converges weakly to a minimum norm solution $(\hat{u}, \hat{q})$ of the system (2.5), i.e.,*

$$L(\hat{u}, \hat{q}) = \bar{g}, \quad C\hat{u} = \bar{f}.$$

$$\psi(\hat{q}) \leqslant \psi(\bar{q}) \text{ for all } (\bar{u}, \bar{q}) \text{ satisfying } (2.5).$$

Proof. As $(u_{\alpha_n}, q_{\alpha_n})$ is the minimizer of J_{α_n}, there holds that

$$J_{\alpha_n}(u_{\alpha_n}, q_{\alpha_n}) \leqslant J_{\alpha_n}(\bar{u}, \bar{q}). \tag{2.13}$$

By definition, $J_{\alpha_n}(\bar{u}, \bar{q}) = \|L(\bar{u}, \bar{q}) - g_{\alpha_n}\|_X^2 + \|C\bar{u} - f_{\alpha_n}\|_Y^2 + \alpha_n\,\psi(\bar{q})$, and thus $\{J_{\alpha_n}(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^{\infty}$ is uniformly bounded. Following the similar argument in the proof of theorem 1, there exists a subsequence of $\{(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^{\infty}$, still denoted as $\{(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^{\infty}$, and some $(\hat{u}, \hat{q})$ such that $\{(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^{\infty}$ converges to $(\hat{u}, \hat{q})$ weakly in $X \times Z$. By assumption 3, we have

$$L(u_{\alpha_n}, q_{\alpha_n})\,\text{converges to}\,L(\hat{u}, \hat{q})\,\text{weakly in}\,X;$$
$$Cu_{\alpha_n}\,\text{converges to}\,C\hat{u}\,\text{weakly in}\,Y.$$

From (2.13) one can also derive that

$$\begin{aligned}&\|L(u_{\alpha_n}, q_{\alpha_n}) - g_{\alpha_n}\|_X^2 + \|Cu_{\alpha_n} - f_{\alpha_n}\|_Y^2 + \alpha_n\,\psi(q_{\alpha_n})\\ &\quad\leqslant \|\bar{g} - g_{\alpha_n}\|_X^2 + \|\bar{f} - f_{\alpha_n}\|_Y^2 + \alpha_n\,\psi(\bar{q}),\end{aligned} \tag{2.14}$$

which implies

$$L(u_{\alpha_n}, q_{\alpha_n})\,\text{converges to}\,\bar{g}\,\text{strongly in}\,X;$$
$$Cu_{\alpha_n}\,\text{converges to}\,\bar{f}\,\text{strongly in}\,Y.$$

Therefore, as $\alpha_n \to 0$, $\{(u_{\alpha_n}, q_{\alpha_n})\}_{n=1}^{\infty}$ converges to $(\hat{u}, \hat{q})$ satisfying

$$L(\hat{u}, \hat{q}) = \bar{g}, \qquad C\hat{u} = \bar{f}. \tag{2.15}$$

Recall that one has an estimate of $\psi(q_{\alpha_n})$ from (2.14) that

$$\alpha_n\psi(q_{\alpha_n}) \leqslant \|g_{\alpha_n} - \bar{g}\|_X^2 + \|f_{\alpha_n} - \bar{f}\|_Y^2 + \alpha_n\,\psi(\bar{q}),$$

which leads to

$$\psi(q_{\alpha_n}) \leqslant \psi(\bar{q}) + o(1),$$

as $\alpha_n \to 0$. Using the lower semicontinuity of functional ψ, one obtain that

$$\psi(\hat{q}) \leqslant \psi(\bar{q}) \quad \text{for all } (\bar{u}, \bar{q}) \text{ satisfying (2.5)}.$$

Together with (2.15), we conclude that $(\hat{u}, \hat{q})$ is a minimum norm solution of (2.5). □

3. Alternating direction iterative method and its convergence

An important feature of the least-squares formulation (2.1) is that the functional $J(u, q)$ is quadratical and convex with respect to each variable u and q. This unique feature facilitates us naturally to minimize the functional J effectively by the ADI method [11, 12] so that only two quadratical and convex suboptimizations of one variable are required at each iteration. We shall carry out the convergence analysis of the ADI method in this section for general inverse medium problems.

Alternating direction iterative method for the minimization of (1.5).

Given an initial pair (u_0, q_0), find a sequence of pairs (u_k, q_k) for $k \geqslant 1$ as follows:

- Given q_k, find $u = u_{k+1} \in X$ by solving

$$\min_{u\in X} \|L(u, q_k) - g\|_X^2 + \|Cu - f\|_Y^2. \tag{3.1}$$

- Given u_{k+1}, find $q = q_{k+1} \in Z$ by solving

$$\min_{q\in Z} \|L(u_{k+1}, q) - g\|_X^2 + \alpha\psi(q). \tag{3.2}$$

We shall establish the convergence of the sequence $\{(u_k, q_k)\}_{k=1}^{\infty}$ generated by the above ADI method, under assumptions 1–3 on the least-squares formulation (2.1). For this purpose, we would like to introduce the Bregman distance [19] with respect to $\xi \in \partial\psi(p)$,

$$E(q, p) := \psi(q) - \psi(p) - \langle \xi, q - p\rangle \quad \forall\ q, p \in Z, \tag{3.3}$$

which is always nonnegative for convex function ψ. Now we are ready to present the convergence of the sequence $\{(u_k, q_k)\}_{k=1}^{\infty}$ generated by (3.1) and (3.2).

Lemma 2. *Under assumptions 1–3, the sequence $\{(u_k, q_k)\}_{k=1}^{\infty}$ generated by the ADI method (3.1) and (3.2) converges to a pair $(u^\star, q^\star)$ that satisfies the optimality condition of (2.1):*

$$(L_q)^*(L_q u - g) + C^*(Cu - f) = 0, -2(\Phi_u)^*(\Phi_u q - g) \in \partial(\alpha\psi(q)). \tag{3.4}$$

Proof. Using the optimality condition satisfied by the minimizer u_{k+1} of (3.1), we can deduce that

$$\begin{aligned} 0 &= \left(L_{q_k}^*(L_{q_k}u_{k+1} - g), u_{k+1} - u_k\right)_X + \left(C^*(Cu_{k+1} - f), u_{k+1} - u_k\right)_Y \\ &= \frac{1}{2}\Big(\|L_{q_k}u_{k+1} - g\|_X^2 + \|Cu_{k+1} - f\|_Y^2 - \left(\|L_{q_k}u_k - g\|_X^2 + \|Cu_k - f\|_Y^2\right) \\ &\quad + \|L_{q_k}(u_{k+1} - u_k)\|_X^2 + \|C(u_{k+1} - u_k)\|_Y^2\Big). \end{aligned} \tag{3.5}$$

Similarly, from the optimality condition satisfied by the minimizer q_{k+1} of (3.2), we can obtain that

$$-2\Phi_{u_{k+1}}^*(\Phi_{u_{k+1}}q_{k+1} - g) \in \partial\big(\alpha\psi(q_{k+1})\big).$$

On the other hand, we derive by taking $q = q_k$, $p = q_{k+1}$ and $\xi = -2\Phi_{u_{k+1}}^*(\Phi_{u_{k+1}}q_{k+1} - g)$ in (3.3) that

$$\langle -2\Phi_{u_{k+1}}^*(\Phi_{u_{k+1}}q_{k+1} - g), q_k - q_{k+1}\rangle + \alpha\psi(q_{k+1}) - \alpha\psi(q_k) + E(q_k, q_{k+1}) = 0. \tag{3.6}$$

We can readily rewrite the first term in (3.6) as

$$\begin{aligned} &\langle \Phi_{u_{k+1}}^*(\Phi_{u_{k+1}}q_{k+1} - g), q_{k+1} - q_k\rangle \\ &\quad = \langle \Phi_{u_{k+1}}q_{k+1} - g, \Phi_{u_{k+1}}(q_{k+1} - q_k)\rangle \\ &\quad = \frac{1}{2}\langle(\Phi_{u_{k+1}}q_{k+1} - g) + (\Phi_{u_{k+1}}q_k - g), \Phi_{u_{k+1}}q_{k+1} - g - (\Phi_{u_{k+1}}q_k - g)\rangle \\ &\qquad + \frac{1}{2}\langle(\Phi_{u_{k+1}}q_{k+1} - g) - (\Phi_{u_{k+1}}q_k - g), \Phi_{u_{k+1}}q_{k+1} - g - (\Phi_{u_{k+1}}q_k - g)\rangle \\ &\quad = \frac{1}{2}\left(\|\Phi_{u_{k+1}}q_{k+1} - g\|_X^2 - \|\Phi_{u_{k+1}}q_k - g\|_X^2 + \|\Phi_{u_{k+1}}(q_{k+1} - q_k)\|_X^2\right). \end{aligned} \tag{3.7}$$

Plugging this into (3.6), we can get that

$$\|\Phi_{u_{k+1}}q_{k+1} - g\|_X^2 + \alpha\psi(q_{k+1}) - \left(\|\Phi_{u_{k+1}}q_k - g\|_X^2 + \alpha\psi(q_k)\right)$$
$$+ \|\Phi_{u_{k+1}}(q_{k+1} - q_k)\|_X^2 + E(q_k, q_{k+1}) = 0. \tag{3.8}$$

As the sequence $\{(u_k, q_k)\}_{k=1}^{\infty}$ is generated by ADI method (3.1) and (3.2), the updated u_{k+1} (resp. q_{k+1}) minimizes the functional $J(u, q_k)$ (resp. $J(u_{k+1}, q)$) at each iteration, which would lead to

$$J(u_k, q_k) \geqslant J(u_{k+1}, q_k) \geqslant J(u_{k+1}, q_{k+1})$$

for all $k \geqslant 0$. Then we can further derive from (3.5) and (3.8) that for any $m \geqslant 1$, $J(u_m, q_m)$ satisfies

$$J(u_m, q_m) + \sum_{k=0}^{m-1} E(q_k, q_{k+1}) + \sum_{k=0}^{m-1} \left(\|L\big(u_{k+1}, q_k\big) - L(u_k, q_k)\|_X^2\right.$$
$$\left. + \|C\big(u_{k+1} - u_k\big)\|_Y^2 + \|L(u_{k+1}, q_{k+1}) - L(u_{k+1}, q_k)\|_X^2\right)$$
$$\leqslant J(u_0, q_0). \tag{3.9}$$

This implies that $\sum_{k=0}^{\infty}\|L(u_{k+1} - u_k, q_k)\|_X^2$ is bounded. Then we can conclude using assumption 2 that $\{u_k\}_{k=1}^{\infty}$ forms a Cauchy sequence and hence converges to some $u^\star \in$ *dom*(L). Since $\sum_{k=0}^{m-1} E(q_k, q_{k+1})$ is uniformly bounded for all m, we can derive that $\{q_k\}_{k=1}^{\infty}$ converges to some $q^\star \in Z$ from the strict convexity of ψ. As the sequence $\{J(u_k, q_k)\}_{k=1}^{\infty}$ is monotonely decreasing, there exists a limit $J^\star$. Following the similar argument in the proof of theorem 1, we conclude that $J^\star = J(u^\star, q^\star)$ by assumption 3. This completes the proof of the convergence. □

Remark 1. If (3.4) has a unique solution in a neighborhood of initial guess (u_0, q_0), then the solution is a local minimizer of the least-squares formulation (2.1), and we can apply the ADI method to generate a sequence that converges to this local minimizer as a plausible approximation of the exact coefficients.

4. Diffusive optical tomography

In this section, we take the DOT as a concrete example of inverse medium problems to illustrate the total least-squares approach we proposed and analysed in the previous sections. We will introduce an effective mixed regularization term, present the least-squares formulation of the first-order system of DOT, and then verify the assumptions in section 2 for the proposed formulation. We shall use the standard notation for Sobolev spaces. The objective of the DOT problem is to determine the unknown diffusion and absorption coefficients σ and μ simultaneously in a Lipschitz domain $\Omega \in \mathbb{R}^d$ $(d = 2, 3)$ from the model equation

$$-\nabla \cdot (\sigma(x)\nabla u) + \mu(x)\, u = 0 \quad \text{in } \Omega \tag{4.1}$$

with a pair of Cauchy data (f, h) on the boundary $\partial\Omega$, i.e.,

$$f = u|_{\partial\Omega}, \qquad h = \frac{\partial u}{\partial \nu}|_{\partial\Omega}.$$

Throughout this section, we shall use the notation $g := \delta_{\partial\Omega}h$ in the total least-squares formulation, where $\delta_{\partial\Omega}$ denotes the Neumann to source map. To define the Neumann to source map, we first introduce the boundary restriction mapping γ_0 on $\mathcal{D}(\overline{\Omega})$, i.e., $\gamma_0 u$ denotes the boundary value of $u \in \mathcal{D}(\overline{\Omega})$. Then we use T to denote the trace operator [20], which is formally defined to be the unique linear continuous extension of γ_0 as an operator from $L^2(\Omega)$ onto $H^{-1/2}(\partial\Omega)$. Using Riesz representation theorem, there exists a function in $L^2(\Omega)$, denoted by $\delta_{\partial\Omega}h$, such that for any $v \in L^2(\Omega)$,

$$(\delta_{\partial\Omega}h, v)_{L^2(\Omega)} = \langle h, Tv\rangle_{H^{1/2}(\partial\Omega), H^{-1/2}(\partial\Omega)}.$$

Then $\delta_{\partial\Omega}h$ will be viewed as g in the least-squares formulation.

4.1. Mixed regularization

In this subsection, we present an effective mixed regularization for solving the DOT problem. As the regularization term ψ in the least-squares formulation (2.1) shall encode the priori information, e.g., sparsity, continuity, lower or upper bound and other properties of the unknown coefficients, it is essential to choose an appropriate regularization term for a concrete inverse problem to ensure satisfactory reconstruction. In this work, we introduce a mixed L^1–H^1 regularization term ϕ for a coefficient $q : \Omega \to \mathbb{R}$:

$$\phi(q; \alpha, \beta, q_0, q_1) = \int_\Omega \frac{\alpha}{2}(|\nabla q|^2 + q^2)\mathrm{d}x + \int_\Omega \beta|q|\,\mathrm{d}x + \chi(q; q_0, q_1), \tag{4.2}$$

where $\chi(q;\ q_0, q_1)$ is given by

$$\chi(q; q_0, q_1) = \begin{cases} 0 & q_0 \leqslant q \leqslant q_1, \\ \infty & \text{otherwise}, \end{cases}$$

and q_0 and q_1 are some priori lower and upper bounds of the coefficient q respectively. The first term $\int_\Omega \frac{\alpha}{2}(|\nabla q|^2 + q^2)\mathrm{d}x$ in (4.2) is the H^1 regularization, the second term $\int_\Omega \beta|q|\mathrm{d}x$ corresponds to the L^1 regularization, and the third term $\chi(q;\ q_0, q_1)$ enforces the reconstruction to meet the constrains of the coefficient. Two constants α and β are used to balance the two regularization terms.

In practice, the L^1 regularization enhances the reconstruction and helps find geometrically sharp inclusions, but might generate spiky results. The H^1 regularization generates the reconstructions with overall clear structure, while the retrieved background may be blurry. Compared with other more conventional regularization methods, this mixed regularization technique in (4.2) combines two penalty terms and effectively promotes multiple features of the target solution, that is, it enhances the sparsity of the solution while it still preserves the overall structure at the same time.

4.2. First-order formulation of DOT

As we have emphasized in the Introduction, it may have some advantages to make use of the residuals of the first-order system of the model equation (4.1), instead of the residual of the original equation in the formulation (2.1), when we aim at recovering two unknown coefficients σ and μ simultaneously. Similarly to the formulation (2.1), we now have $q = (\sigma,\ \mu)$, and the operator L is given by

$$L(u, \mathbf{p}, \sigma, \mu) = \begin{pmatrix} -\nabla\cdot\mathbf{p} + \mu\, u \\ \mathbf{p} - \sigma\,\nabla u \end{pmatrix}, \tag{4.3}$$

where we have introduced an auxiliary vector flux $\mathbf{p}$ and each entry of L is of a first-order form such that two coefficients are separated naturally. Clearly, $L(v, q)$ is still bilinear with respect to the state variables $v = (u, \mathbf{p})$ and coefficients $q = (\sigma,\ \mu)$. Using the first-order system, we can then come to the following total least-squares functional:

$$\begin{aligned} J(u, \mathbf{p}, \sigma, \mu) = \| -\nabla \cdot \mathbf{p} + \mu(x)u - g\|^2_{L^2(\Omega)} + \|\mathbf{p} - \sigma(x)\,\nabla u\|^2_{(L^2(\Omega))^d} \\ + \|Cu - f\|^2_{L^2(\partial\Omega)} + \psi_1(\sigma) + \psi_2(\mu), \end{aligned} \tag{4.4}$$

where C is the trace operator, and $\psi_1(\sigma) = \phi(\sigma;\ \alpha_\sigma, \beta_\sigma, \sigma_0, \sigma_1)$ and $\psi_2(\mu) = \phi(\mu;\ \alpha_\mu, \beta_\mu,\ \mu_0,\ \mu_1)$ are the corresponding mixed regularization terms of σ and μ defined as in (4.2), μ_0, μ_1 are the lower and upper bounds of μ, and σ_0, σ_1 are the lower and upper bounds of σ. We shall minimize (4.4) over $(u, \mathbf{p}, \sigma, \mu) \in L^2(\Omega) \times (L^2(\Omega))^d \times L^2(\Omega) \times L^2(\Omega)$, that is, we have the spaces $X = L^2(\Omega) \times (L^2(\Omega))^d$ and $Z = L^2(\Omega) \times L^2(\Omega)$ at the current setting.

We will apply the ADI method to solve the least-squares formulation of $v = (u, \mathbf{p})$ and $q = (\sigma,\ \mu)$:

$$\begin{aligned} \min_{(v,q)\in X\times Z} J(u, \mathbf{p}, \sigma, \mu) = \| -\nabla \cdot \mathbf{p} + \mu(x)u - g\|^2_{L^2(\Omega)} \\ + \|\mathbf{p} - \sigma(x)\,\nabla u\|^2_{(L^2(\Omega))^2} + \|Cu - f\|^2_{L^2(\partial\Omega)} + \psi_1(\sigma) + \psi_2(\mu). \end{aligned} \tag{4.5}$$

Given an initial guess $(u_0, \mathbf{p}_0, \sigma_0,\ \mu_0)$, we find a sequence $(u_k, \mathbf{p}_k, \sigma_k,\ \mu_k)$ for $k \geqslant 1$ as below:

- Given σ_k, μ_k, find $u = u_{k+1}$, $\mathbf{p} = \mathbf{p}_{k+1}$ by solving

$$\begin{aligned} \min_{(u,\mathbf{p})\in X} J_1(u, \mathbf{p}) = \| -\nabla \cdot \mathbf{p} + \mu_k(x)u - g\|^2_{L^2(\Omega)} + \|\mathbf{p} - \sigma_k(x)\,\nabla u\|^2_{(L^2(\Omega))^d} \\ + \|Cu - f\|^2_{L^2(\partial\Omega)}, \end{aligned}$$

- Given u_{k+1}, $\mathbf{p}_{k+1}$, find $\sigma = \sigma_{k+1}$, $\mu = \mu_{k+1}$ by solving

$$\begin{aligned} \min_{(\sigma,\mu)\in Z} J_2(\sigma, \mu) = \| -\nabla \cdot \mathbf{p}_{k+1} + \mu(x)u_{k+1} - g\|^2_{L^2(\Omega)} + \|\mathbf{p}_{k+1} \\ - \sigma(x)\,\nabla u_{k+1}\|^2_{(L^2(\Omega))^d} + \psi_1(\sigma) + \psi_2(\mu). \end{aligned}$$

4.3. Well-posedness of the least-squares formulation for DOT

Recall that we have proved the well-posedness of the least-squares formulation in section 2 for general inverse medium problems. This subsection is devoted to the verification of assumptions in section 2 for the formulation (4.5) to ensure its well-posedness.

Firstly we consider assumption 1 on the regularization terms. It is observed from the formula (4.2) that each term of $\psi_1(\sigma)$ and $\psi_2(\mu)$ in (4.5) is convex and weakly lower semicontinuous. As the first term of (4.2) is the H^1 regularization term, $\psi_1(\sigma)$ and $\psi_2(\mu)$ are strictly convex by definition. One can also observe that there exists a positive constant c such that

$$\begin{aligned} \psi_1(\sigma) = \int_\Omega \frac{\alpha}{2}(|\nabla\sigma|^2 + \sigma^2)\mathrm{d}x + \int_\Omega \beta|\sigma|\,\mathrm{d}x + \chi(\sigma; \sigma_0, \sigma_1) \geqslant c\|\sigma\|^2_{H^1(\Omega)}, \\ \psi_2(\mu) = \int_\Omega \frac{\alpha}{2}(|\nabla\mu|^2 + \mu^2)\mathrm{d}x + \int_\Omega \beta|\mu|\,\mathrm{d}x + \chi(\mu; \mu_0, \mu_1) \geqslant c\|\mu\|^2_{H^1(\Omega)}, \end{aligned} \tag{4.6}$$

which imply that ψ_1 and ψ_2 are coercive in H^1-norm.

Next we verify assumption 2 on the closedness of L_q and the coercivity of its graph norm. For fixed $q = (\sigma,\ \mu)$, we denote the entries of the operator L_q in (4.3) by $L_{q,1}$, $L_{q,2}$, i.e., for $\delta v = (\delta u, \delta \mathbf{p}) \in dom(L)$,

$$L_{q,1}\delta v = -\nabla \cdot \delta \mathbf{p} + \mu \delta u, \qquad L_{q,2}\delta v = \delta \mathbf{p} - \sigma \nabla \delta u.$$

By means of the Hölder's inequality, we can readily have

$$\|L_q \delta v\|_X \leqslant \|\delta \mathbf{p}\|_{H(\mathrm{div},\Omega)} + \|\delta u\|_{H^1(\Omega)}\|\mu\|_{L^\infty(\Omega)} + \|\delta u\|_{H^1(\Omega)}\|\sigma\|_{L^\infty(\Omega)} \tag{4.7}$$

for q in the level set $\{q \in Z : \psi(q) = \psi_1(\sigma) + \psi_2(\mu) \leqslant c_0\}$ for some constant c_0. Then we can deduce from (4.7) that L_q is a closed linear operator. Next, we verify the coercivity of the graph norm. For simplicity, we consider the model problem with the homogeneous Neumann boundary condition $h = 0$ for the state variable, and set a side constraint for $dom(L)$ as $\mathbf{p} \cdot \nu = 0$ on $\partial\Omega$. Introduce the following notations

$$\sigma \nabla u - \mathbf{p} = \tilde{m}, \qquad -\nabla \cdot \mathbf{p} + \mu\, u = \tilde{g}. \tag{4.8}$$

From (4.8) and by integration by part, we can derive

$$\int_\Omega \sigma \nabla u \cdot \nabla u \mathrm{d}x + \int_\Omega \mu u \cdot u \mathrm{d}x = \int_\Omega \nabla u \cdot \tilde{m} \mathrm{d}x + \int_\Omega u \cdot \tilde{g} \mathrm{d}x,$$

which implies

$$\begin{aligned} &\|\nabla u\|^2_{(L^2(\Omega))^d} + \|u\|^2_{L^2(\Omega)} + \|\nabla \cdot \mathbf{p}\|^2_{L^2(\Omega)} + \|\mathbf{p}\|^2_{(L^2(\Omega))^d} \\ &\quad \leqslant c\,(\|\tilde{m}\|^2_{(L^2(\Omega))^d} + \|\tilde{g}\|^2_{L^2(\Omega)}) \end{aligned} \tag{4.9}$$

for some constant $c > 0$, when $q = (\sigma,\ \mu)$ satisfies $\psi_1(\sigma) \leqslant c_0$ and $\psi_2(\mu) \leqslant c_0$ for some constant $c_0 > 0$, as we did for (4.7). In this way we have verified that the graph norm corresponding to L_q defined as

$$|(u, \mathbf{p})|^2_{W,q} = \|\sigma \nabla u - \mathbf{p}\|^2_{(L^2(\Omega))^d} + \| - \nabla \cdot \mathbf{p} + \mu\, u\|^2_{L^2(\Omega)}$$

is uniformly coercive.

We next consider assumption 3 on the weakly sequentially closedness of operators L and C. It is noted that assumption 3 is applied in section 2 for general inverse medium problems to prove that the operator L maps a subsequence of a bounded sequence $\{(u_n, q_n)\}_{n=1}^\infty$ to a converging sequence $\{L(u_n, q_n)\}_{n=1}^\infty$ with its limit equal to $L(u, q)$, where (u, q) is the limit of $\{(u_n, q_n)\}_{n=1}^\infty$. In the analysis of the concrete DOT problem (4.1), the corresponding sequence $\{(u_n, \mathbf{p}_n, \sigma_n, \mu_n)\}_{n=1}^\infty$ is actually bounded in a stronger norm than $\| \cdot \|_X$ and $\| \cdot \|_Z$, as shown in (4.6) and (4.9). To prove the well-poseness of the least-squares formulation (4.5), it suffices to verify that for a sequence $\{(u_n, \mathbf{p}_n, \sigma_n, \mu_n)\}_{n=1}^\infty$ bounded in $H^1(\Omega) \times H(\mathrm{div},\ \Omega) \times H^1(\Omega) \times H^1(\Omega)$, there exists a subsequence weakly converging to $(u, \mathbf{p}, \sigma,\ \mu)$, and the operator L defined in (4.3) satisfies that $\{L(u_n, \mathbf{p}_n, \sigma_n, \mu_n)\}_{n=1}^\infty$ weakly converges to $L(u, \mathbf{p}, \sigma,\ \mu)$ in X.

For this verification, we first note that for a given bounded sequence $\{(u_n, \mathbf{p}_n, \sigma_n, \mu_n)\}_{n=1}^\infty$, there exists a subsequence, still denoted as $\{(u_n, \mathbf{p}_n, \sigma_n, \mu_n)\}_{n=1}^\infty$, weakly converging to $(v, q) := (u, \mathbf{p}, \sigma,\ \mu)$ in $H^1(\Omega) \times H(\mathrm{div},\ \Omega) \times H^1(\Omega) \times H^1(\Omega)$. Denoting $(u_n, \mathbf{p}_n)$ by v_n and $(\sigma_n,\ \mu_n)$ by q_n, there holds for any $V = (V_1, V_2) \in X$ that

$$
\begin{aligned}
&\lim_{k\to\infty}(L(v_n,q_n),V)_X-(L(v,q),V)_X\\
&\quad=\lim_{k\to\infty}\left(\left(L_{q_n}(v_n-v),V\right)_X+(L(v,q_n-q),V)_X\right)\\
&\quad=\lim_{k\to\infty}\big((-\nabla\cdot(\mathbf{p}_n-\mathbf{p})+\mu_n(u_n-u),V_1)_{L^2(\Omega)}\\
&\qquad+((\mathbf{p}_n-\mathbf{p})-\sigma_n\nabla(u_n-u),V_2)_{(L^2(\Omega))^d}\\
&\qquad+((\mu_n-\mu)u,V_1)_{L^2(\Omega)}-((\sigma_n-\sigma)\nabla u,V_2)_{(L^2(\Omega))^d}\big)=0,
\end{aligned}
$$

where we have used the weak convergence of $\{(u_n,\mathbf{p}_n,\sigma_n,\mu_n)\}_{n=1}^{\infty}$ and Hölder's inequality. Therefore, $\{L(u_n,\mathbf{p}_n,\sigma_n,\mu_n)\}_{n=1}^{\infty}$ weakly converges to $L(u,\mathbf{p},\sigma,\ \mu)$ in X. On the other hand, as C is the trace operator on $\partial\Omega$ in the DOT problem, it is linear and thus weakly sequentially closed, which completes our verification.

5. Numerical experiments

In this section, we carry out some numerical experiments for the DOT problem with different scenarios to illustrate the efficiency and robustness of the proposed two-stage algorithm in this work. Throughout these examples, we shall assume that we apply the Neumann boundary data h on $\partial\Omega$ and measure the corresponding Dirichlet data f to reconstruct the diffusion and absorption coefficients σ and μ simultaneously. The basic algorithm involves two stages: we apply the direct sampling method (DSM) in the first stage to get some initial approximations of the two unknown coefficients, and then adopt the total least-squares method to achieve more accurate reconstructions of the coefficients.

5.1. Direct sampling method for initialization

For all the numerical experiments, we shall use the DSM in the first stage of our algorithm, in an attempt to effectively locate the multiple inclusions inside the computational domain with limited measurement data. Here we give a brief description of DSM and refer the readers to [18] for more technical details about this DSM that can identify multiple coefficients. The DSM develops two separate families of probing functions, i.e., the monopole and dipole probing functions, for constructing separate index functions for multiple physical coefficients. The coefficients of the inhomogeneities can be approximated based on index functions due to the following two observations: the difference of scattered fields caused by inclusions can be approximated by the sum of Green's functions of the homogeneous medium and their gradients; and the two sets of probing functions have the mutually almost orthogonality property, i.e., they only interact closely with the Green's functions and their gradients respectively. Thus we can decouple the monopole and dipole effects and derive index functions for separate physical properties.

In practice, if the value of an index function ϕ for one physical coefficient at a sampling point x is close to 1, the sampling point is likely to stay in the support of inhomogeneity; whereas if $\phi(x)$ is close to 0, the sampling point x stays most probably outside the support. Hence the index functions give an image of the approximate support of the inhomogeneity, and we can determine a subdomain D to locate the support from index functions. The subdomain D could be chosen as $D=\{x\in\Omega:\phi(x)\geqslant\theta\}$ with θ being a suitable cut-off value, then we restrict the index function ϕ on D. We adopt this choice in the numerical experiments to remove the spurious oscillations in the homogeneous background. Once we have the restricted index function $\phi|_D$, we set the value of the approximation $\tilde{\phi}$ in D as $\tilde{\phi}|_D=c_\phi\phi|_D$, where the constant

c_ϕ is some priori estimate of the true coefficient, and set the value of $\tilde{\phi}$ outside of D as the background coefficient. In this way, we obtain $\tilde{\phi}$ as the initial guess for the total least-squares method for our further reconstruction in the second stage.

5.2. Examples

In all our numerical results with the proposed two-stage least-squares method for inverse medium problems, we discretize the objective functional J in (4.4) by a staggered finite difference scheme [21, 22]. The computational domain $\Omega = [0, 1] \times [0, 1]$ is divided into a uniform mesh consisting of small squares of width $h = 0.01$. The measurements u^δ of the true boundary data u are assumed to be of the noisy form:

$$u^\delta(x) = u(x) + \epsilon\eta \max_{x\in\partial\Omega} |u(x)|$$

where ϵ is the relative noise level and η follows the standard Gaussian distribution. The subdomain D is chosen by

$$D = \{x \in \Omega : \phi(x) \geqslant \theta\}, \tag{5.1}$$

where ϕ is the index function from DSM and the cutoff value θ is taken in the range (0.4, 0.7). The choice of the cut-off value θ will affect the size of the subdomain D, but it is insensitive to the reconstructed coefficients in the second stage. Without loss of generosity, we assume that the background diffusion and absorption coefficients $\underline{\sigma} = 1$ and $\underline{\mu} = 1$. The regularization parameters $(\alpha_\sigma, \beta_\sigma, \alpha_\mu, \beta_\mu)$ involved in the functional in (4.4) may be determined by various strategies for multiple-parameter regularizations, such as the balancing principle in [6]. But in this work, all the parameters are chosen in a trial and error manner, which are presented in more detail in table 1 for different examples. The numerical iterations are terminated either when the number of iterations reaches the maximum number K (it is set to be 50) or when the following stopping criteria are satisfied:

$$\frac{\|\sigma_{k+1} - \sigma_k\|_{L^2(\Omega)}}{\|\sigma_{k+1}\|_{L^2(\Omega)}} < 0.01 \quad \text{and} \quad \frac{\|\mu_{k+1} - \mu_k\|_{L^2(\Omega)}}{\|\mu_{k+1}\|_{L^2(\Omega)}} < 0.01.$$

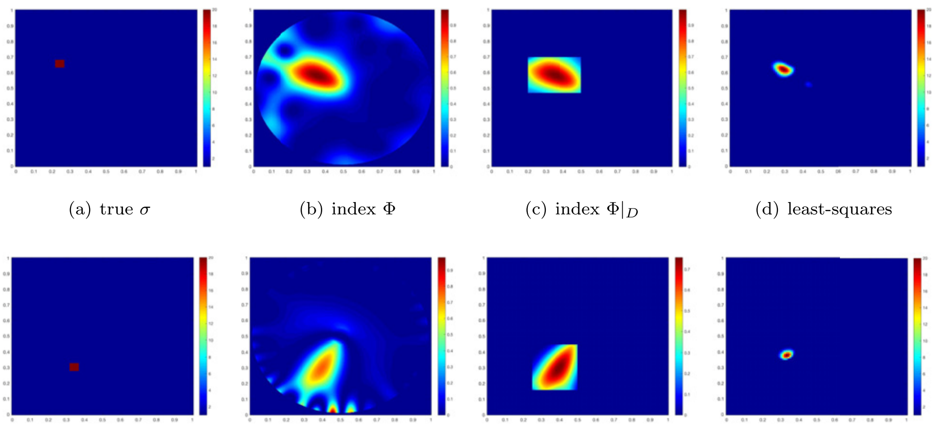
In all the experiments presented in this section, the above stopping criteria are achieved within about 20 iterations. All the computations were performed in MATLAB (R2018B) on a desktop computer.

Example 1. We consider the discontinuous diffusion and absorption coefficients σ and μ with one inclusion each. The inclusion of σ is of width 0.05 and centered at (0.25, 0.65) as shown in figure 1(a); the inclusion of μ is of width 0.05 and centered at (0.35, 0.3) as shown in figure 1(e). The magnitudes of the coefficients inside the inclusions are 20.

We shall use only one set of measurements to reconstruct the diffusion and absorption coefficients. It is shown in figures 1(b) and (f) that the index functions from the DSM separate inclusions of different physical nature well and give the initial approximate locations, while the exact locations of small inclusions are difficult to detect. If we simply take the maximal points of the index functions in figures 1(b) and (f) as the locations of the reconstructed inclusions, we may not be able to identify the true locations of inhomogeneity. Then we set the subdomain D using information from the first stage by (5.1), see figures 1(c) and (g), and set the value of approximation out of the subdomain to be the background coefficients. As in figures 1(d) and (h), this example illustrates that the least-squares formulation in the second stage works very well to improve the reconstruction and provides a much more accurate

Table 1. The regularization parameters (α, β) in each example for σ and μ without noise and with noise.

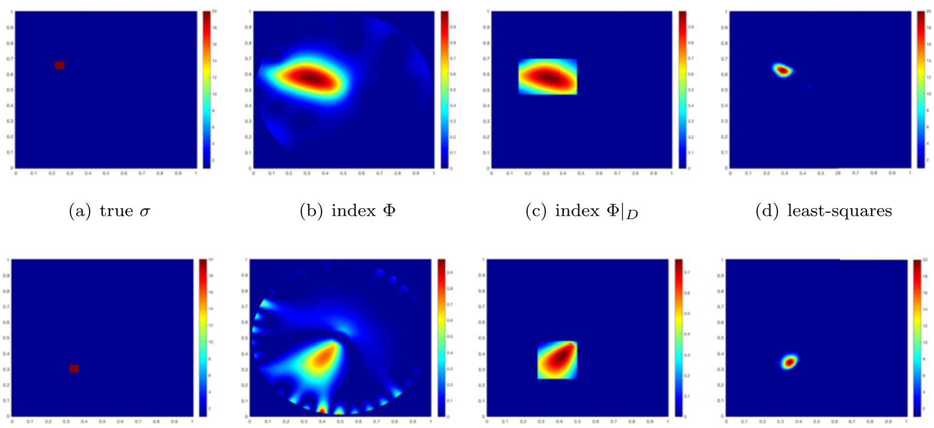
Noise level	$\epsilon = 0$		$\epsilon > 0$	
Example	$(\alpha_\sigma, \beta_\sigma)$	(α_μ, β_μ)	$(\alpha_\sigma, \beta_\sigma)$	(α_μ, β_μ)
1	$(1.0 \times 10^{-2}, 2.0 \times 10^{-2})$	$(5.0 \times 10^{-4}, 5.0 \times 10^{-4})$	$(1.0 \times 10^{-2}, 2.0 \times 10^{-2})$	$(5.0 \times 10^{-4}, 1.0 \times 10^{-3})$
2	$(1.0 \times 10^{-3}, 5.0 \times 10^{-3})$	N/A	$(1.0 \times 10^{-3}, 1.0 \times 10^{-2})$	N/A
3	$(1.0 \times 10^{-6}, 1.0 \times 10^{-3})$	N/A	$(1.0 \times 10^{-6}, 2.0 \times 10^{-3})$	N/A
4	$(1.0 \times 10^{-3}, 1.0 \times 10^{-2})$	$(1.0 \times 10^{-2}, 5.0 \times 10^{-3})$	$(1.0 \times 10^{-3}, 2.0 \times 10^{-2})$	$(1.0 \times 10^{-2}, 5.0 \times 10^{-3})$
5	$(1.0 \times 10^{-5}, 5.0 \times 10^{-4})$	N/A	$(1.0 \times 10^{-5}, 1.0 \times 10^{-3})$	N/A



(a) true σ (b) index Φ (c) index $\Phi|_D$ (d) least-squares

(e) true μ (f) index Φ (g) index $\Phi|_D$ (h) least-squares

Figure 1. Numerical results for example 1 with exact data: (a) true σ; (b) index Φ from DSM; (c) index $\Phi|_D$ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs for μ.

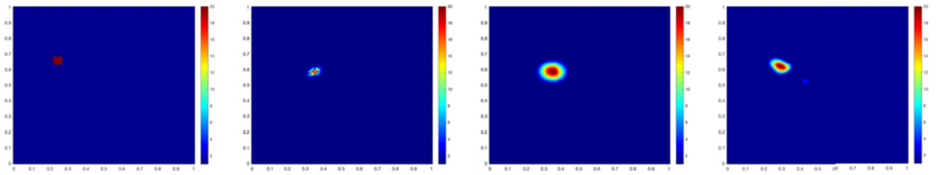


(a) true σ (b) index Φ (c) index $\Phi|_D$ (d) least-squares

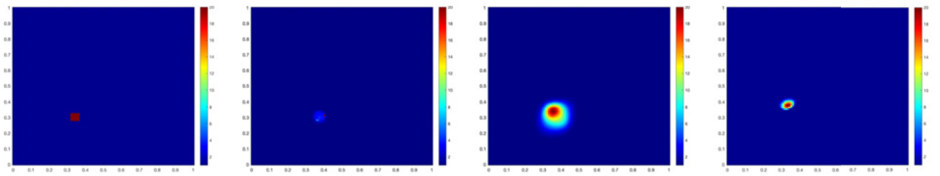
(e) true μ (f) index Φ (g) index $\Phi|_D$ (h) least-squares

Figure 2. Numerical results for example 1 using the data with 10% noise: (a) true σ; (b) index Φ from DSM; (c) index $\Phi|_D$ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs for μ.

location than that provided by DSM. This can be further evidenced by the centers of mass of the inhomogeneous part of the reconstructed $\sigma^\star$ and $\mu^\star$, which are located at (0.30, 0.62) and (0.34, 0, 37) respectively. With 10% noise in the measurements, the centers of mass of the inclusions in the reconstructed $\sigma^\star$ and $\mu^\star$ are located at (0.29, 0.62) and (0.35, 0, 34) respectively. We can observe that the reconstructions remain accurate in the noisy case, as shown in figure 2, which indicates that the two-stage algorithm gives quite stable reconstructions with respect to a big noise even though the DSM reconstructions become more blurry in figures 2(b) and (f). The regularization parameters $(\alpha_\sigma, \beta_\sigma, \alpha_\mu, \beta_\mu)$ are presented as in table 1.



(a) true σ (b) L^1 regularization (c) H^1 regularization (d) mixed regularization



(e) true μ (f) L^1 regularization (g) H^1 regularization (h) mixed regularization

Figure 3. Numerical results for example 1 with different regularization methods in the least-squares formulation: (a) true σ; (b) reconstruction with L^1 regularization; (c) reconstruction with H^1 regularization; (d) reconstruction with mixed regularization. (e)–(h) are the corresponding graphs for μ.

Compared with the reconstructions derived from the DSM, the improvement of the approximations for both σ and μ is significant: the recovered background is now mostly homogeneous, and the magnitude and size of the inhomogeneity approximate those of the true coefficients well. These results indicate clearly the significant potential of the proposed least-squares formulation with the mixed regularization for inverse medium problems.

From table 1 we obtain an insightful observation about the mixed regularization: the magnitude of parameter β is mostly larger than that of α. We can conclude that the L^1 penalty plays a predominant role in improving the performance of reconstruction for such inhomogeneous coefficients, whereas the H^1 penalty yields a locally smooth structure. This is why the mixed regularization has been applied for all the numerical examples in this section. We have also run the reconstructions with the single-parameter regularization and observed that the reconstruction with only the L^1 regularization often leads to a rather sparse and spiky result, while the reconstruction with only the H^1 regularization usually gives a more blurry and diffusive result; see figure 3 for a comparison.

Example 2. We implement the algorithm to reconstruct the discontinuous diffusion coefficient σ with two inclusions. We assume that μ is known, being the same as the background coefficient $\underline{\mu}$. One set of data is measured to locate two inclusions of σ, which are centered at (0.15, 0.5) and (0.5, 0.85) respectively, both of width 0.05. σ is taken to be 20 inside the regions as shown in figure 4(a).

The reconstructions for σ using exact measurements and measurements with 20% noise are shown in figure 4. In the first stage, the DSM gives the index function that separates these two inclusions well using only one set of data as in figure 4(b), while some inhomogeneity is observed in the background. This phenomenon comes from the ill-posedness of inverse medium problems and also the oscillation of fundamental solutions used in the DSM. Even so, the approximations in figures 4(b) and (f) still provide the basic modes of the inhomogeneity. We can identify that there are two inclusions, and capture the subdomain D for the second stage in figures 4(c) and (g). The least-squares formulation with mixed regularization

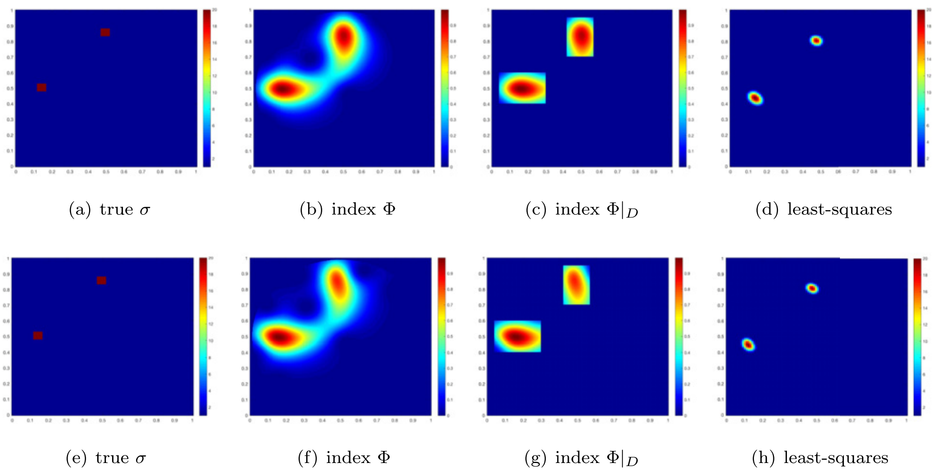


Figure 4. Numerical results for example 2 using the data with no noise (first row) and 20% noise (second row): (a) true σ; (b) index Φ from DSM; (c) index $\Phi|_D$ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs with 20% noise.

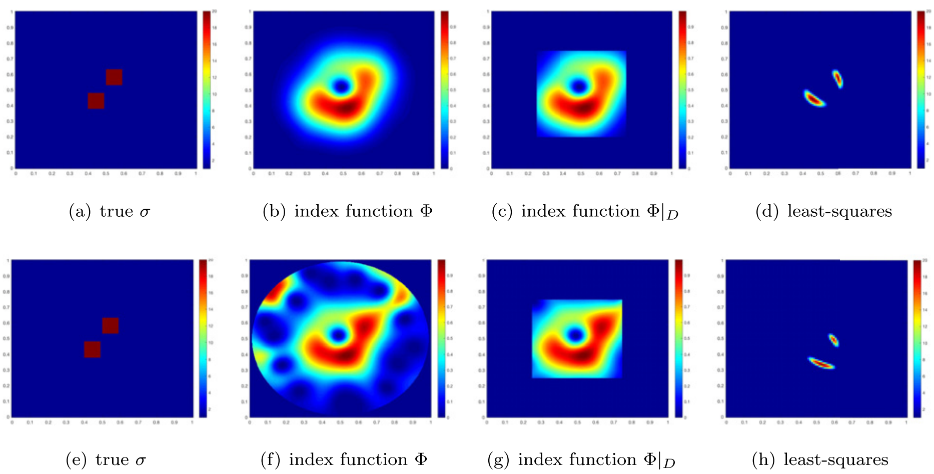
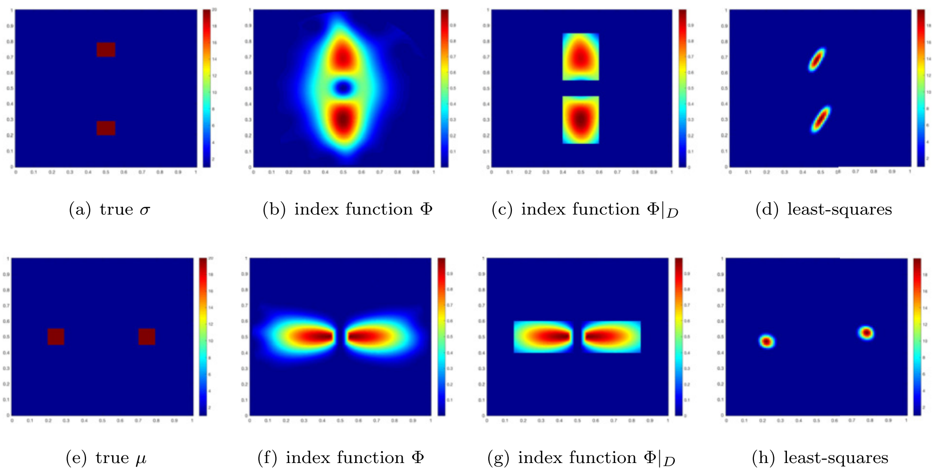


Figure 5. Numerical results for example 3 using the data without noise (first row) and with 2% noise: (a) true σ; (b) index Φ from DSM; (c) index $\Phi|_D$ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs using the data with 2% noise.

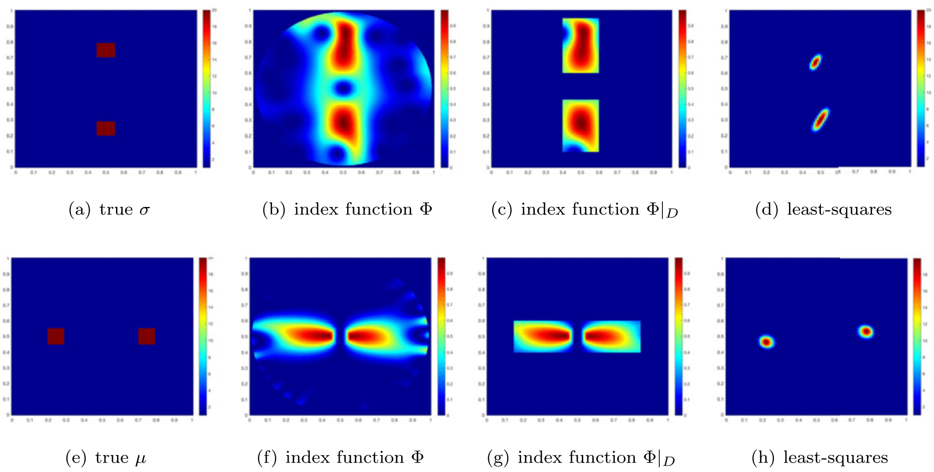
significantly improves the reconstruction: the locations of both inclusions are captured better with clear background and accurate size in figures 4(d) and (h), and the centers of mass of the two inclusions in the reconstructed $\sigma^\star$ (figure 4(d)) are located at (0.14, 0.43) and (0.48, 0, 81) respectively. Comparing figures 4(d) and (h), one can observe that the reconstruction deteriorates only slightly in that the left inclusion shrinks a little bit when the noise level ϵ increases from 0 to 20%, and the centers of mass of these two inclusions of inhomogeneity



(a) true σ (b) index function Φ (c) index function $\Phi|_D$ (d) least-squares

(e) true μ (f) index function Φ (g) index function $\Phi|_D$ (h) least-squares

Figure 6. Numerical results for example 4 using the exact data: (a) true σ; (b) index Φ for σ from DSM; (c) index $\Phi|_D$ for σ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs for μ.



(a) true σ (b) index function Φ (c) index function $\Phi|_D$ (d) least-squares

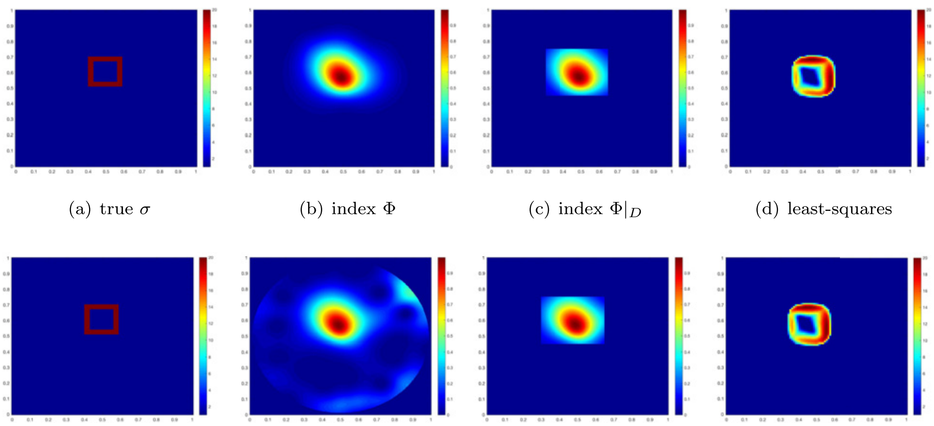
(e) true μ (f) index function Φ (g) index function $\Phi|_D$ (h) least-squares

Figure 7. Numerical results for example 4 using the data with 20% noise: (a) true σ; (b) index Φ for σ from DSM; (c) index $\Phi|_D$ for σ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs for μ.

are now located at (0.12, 0.44) and (0.48, 0, 81) respectively. This example verifies that the proposed two-stage algorithm is very robust with respect to the data noise.

Example 3. In this example we reconstruct two inclusions of diffusion coefficient σ that stay very close to each other. We assume that μ is the same as the background coefficient $\underline{\mu}$. The two inclusions of diffusion coefficient σ are centered at (0.45, 0.425) and (0.55, 0.575) respectively and of width 0.1 as shown in figure 5(a). The coefficient σ is 20 in both regions.

The two inclusions in example 2 are relatively far from each other, while example 3 considers the case when two inclusions are quite close to each other, which is more challenging as it



(a) true σ (b) index Φ (c) index $\Phi|_D$ (d) least-squares

(e) true σ (f) index Φ (g) index $\Phi|_D$ (h) least-squares

Figure 8. Numerical results for example 5 using the data without noise (first row) and with 20% noise (second row): (a) true σ; (b) index Φ from DSM; (c) index $\Phi|_D$ (constrained to the chosen subdomain D); (d) least-squares reconstruction. (e)–(h) are the corresponding graphs using data with 20% noise.

would be difficult to distinguish these two separated inclusions and reconstruct their locations and magnitudes precisely. As shown in figure 5(b), the index function from DSM presents limited information on the diffusion coefficient with only one set of data, and shows one connected inclusion. With 2% noise, the index function is blurred a lot as shown in figure 5(f). In both noisy and noiseless cases, only one subdomain can be detected from this index function from the first stage. The second stage still presents well separated reconstructions with the size and magnitude that match the exact diffusion coefficient well as shown in figure 5(d), and the centers of mass of the two inclusions in the reconstructed $\sigma^\star$ are located at (0.47, 0.43) and (0.61, 0, 57) respectively. For the case with 2% noise, this two-stage algorithm also gives satisfactory reconstruction (see figure 5(h)), and the centers of mass of the two inclusions are now located at (0.53, 0.32) and (0.60, 0, 48) respectively. Compared with figure 5(d), it is observable that the left inclusion moves towards x-axis and elongates a little bit, while we can still tell the sizes and locations of two inclusions from the reconstruction. This shows that the least-squares method provides much more details than DSM and is relatively robust with respect to the noise in the measurement.

Note that for the noise level higher than 2%, the DSM cannot provide a feasible initial guess for the least-squares method in the second stage with only one set of measurement data. But with an initial guess that can reflect some basic mode of the true coefficient, the least-squares method in the second stage has great tolerance of noise as shown in other examples.

Example 4. With this example, we reconstruct σ and μ simultaneously, with two inclusions for each coefficient. The inclusions are in the following scenario: the inclusions of diffusion coefficient σ are of width 0.1 and centered at (0.5, 0.25), (0.5, 0.75) respectively, and the magnitude inside the region is 20; the inclusions of absorption coefficient μ are of width 0.1 and centered at (0.25, 0.5), (0.75, 0.5) respectively, and the magnitude inside the region is 20, as shown in figures 6(a) and (e).

In example 1, we have only one inclusion for coefficients μ and σ each, and it is shown that this algorithm can reconstruct both medium coefficients well. Example 4 is more challenging than example 1, as the existence of two inclusions for each coefficient will influence the reconstruction of the other coefficient. As shown in figures 6(b) and (f), the index functions separate these inclusions well and give a rough approximation for their locations with only one set of data. However, it can be observed that the maximal points of index functions differ significantly from the exact coefficients. In the second stage of the proposed algorithm, the locations are improved significantly, as shown in figures 6(d) and (h), where the centers of mass of the two inclusions in the reconstruction $\sigma^{\star}$ are located at (0.51, 0.29) and (0.48, 0.68), while the centers of mass of the two inclusions in the reconstruction $\mu^{\star}$ are centered at (0.22, 0.46) and (0.79, 0.52). When we consider the case with 20% noise, DSM provides blurry approximations as shown in figures 7(b) and (f). In the second stage of reconstruction, figures 7(d) and (h) present results almost the same as figures 6(d) and (h), where the centers of mass of the two inclusions in figure 7(d) are located at (0.51, 0.29) and (0.48, 0.67), while the centers of mass of the two inclusions in figure 7(h) are centered at (0.22, 0.46) and (0.79, 0.53). This example demonstrates again the robust performance of the new method.

Example 5. In this example we assume that μ is the same as the background coefficient $\underline{\mu}$ and reconstruct the diffusion coefficient σ with ring-shaped square inclusion as shown in figure 8(a), with two sets of measurement data. The outer and inner side length of the ring are 0.2 and 0.15, and the rectangle ring is centered at (0.5, 0.6). The coefficient σ is taken to be 20 inside the region.

We apply the two-stage least-squares method with two sets of measurement data from different directions for this very challenging case. As we see from figure 8(b), the index function from the DSM can only reflect an approximated location of the inclusion, and it does not give any clue about the actual shape of the inclusion. But from figure 8(d), we can clearly see that the least-squares formulation can reconstruct the edges of the ring-shape inclusion. When 20% noise is present in the data, one has the reconstruction (see figure 8(h)) that is very similar to the results without noise, which shows the approximation is quite stable with respect to the noise.

6. Concluding remarks

We have presented a novel two-stage least-squares approach to reconstruct multiple unknown coefficients simultaneously for a class of inverse medium problems. A DSM is applied to provide an approximate initial support of the inhomogeneity in the first stage, then a total least-squares method with a mixed regularization is used to recover the medium profile in the second stage. The total least-squares formulation combines the residual of the model equation, the data fitting and the regularizations in the least-squares criterion. Analyses are provided to prove the well-posedness and convergence of the algorithm, and numerical results are provided to demonstrate the robustness and effectiveness of the algorithm. We are currently investigating this new methodology in several directions. Firstly, it is of great interest to explore how the formulations of the least L^1 criterion perform, which is expected to remove the potential outliers in the measurements and retain several good features of the total least-squares approach. Secondly, the robustness of the two-stage algorithm might be further improved with different DSMs in the first stage. Thirdly, it is very interesting and important to develop an effective deterministic strategy for providing desired regularization parameters for the total least-squares method. Although this paper concerns the inverse medium problem with multiple parameters, we believe that the proposed least-squares framework can be applied to solve other important parameter identification problems, and even more general stochastic inverse problems.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

Jun Zou https://orcid.org/0000-0002-4809-7724

References

[1] Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41
[2] Gibson A P, Hebden J C and Arridge S R 2005 Recent advances in diffuse optical imaging *Phys. Med. Biol.* **50** R1
[3] Dorn O, Bertete-Aguirre H, Berryman J G and Papanicolaou G C 1999 A nonlinear inversion method for 3d electromagnetic imaging using adjoint fields *Inverse Problems* **15** 1523
[4] Bao G and Li P 2005 Inverse medium scattering problems for electromagnetic waves *SIAM J. Appl. Math.* **65** 2049–66
[5] Cheney M, Isaacson D and Newell J C 1999 Electrical impedance tomography *SIAM Rev.* **41** 85–101
[6] Ito K and Jin B 2014 *Inverse Problems: Tikhonov Theory and Algorithms* (Singapore: World Scientific)
[7] Kohn R V and Vogelius M 1987 Relaxation of a variational method for impedance computed tomography *Commun. Pure Appl. Math.* **40** 745–77
[8] Wexler A, Fry B and Neuman M R 1985 Impedance-computed tomography algorithm and system *Appl. Opt.* **24** 3985–92
[9] Chung E, Ito K and Yamamoto M 2021 Least square formulation for ill-posed inverse problems and applications *Appl. Anal.* **100** 1–15
[10] Lattes R and Lions J L 1969 *The Method of Quasi-Reversibility. Applications to Partial Differential Equations* (Amsterdam: Elsevier)
[11] Tusnady G and Csiszar I 1984 Information geometry and alternating minimization procedures *Stat. Decisions* 205–37
[12] Byrne C L 2014 *Iterative Optimization in Inverse Problems* (Boca Raton, FL: CRC Press)
[13] Jespersen D C 1977 A least squares decomposition method for solving elliptic equations *Math. Comput.* **31** 873–80
[14] Bochev P B and Gunzburger M D 1998 Finite element methods of least-squares type *SIAM Rev.* **40** 789–837
[15] Pehlivanov A I, Carey G F and Lazarov R D 1994 Least-squares mixed finite elements for second-order elliptic problems *SIAM J. Numer. Anal.* **31** 1368–77
[16] Ito K, Jin B and Zou J 2012 A two-stage method for inverse medium scattering *J. Comput. Phys.* **237** 211–23
[17] Chow Y T, Ito K and Zou J 2014 A direct sampling method for electrical impedance tomography *Inverse Problems* **30** 095003
[18] Chow Y T, Han F and Zou J 2021 A direct sampling method for simultaneously recovering inhomogeneous inclusions of different nature *SIAM J. Sci. Comput.* **43** A2161–89
[19] Bregman L M 1967 The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming *USSR Comput. Math. Math. Phys.* **7** 200–17
[20] Girault V and Raviart P A 2012 *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms* vol 5 (Berlin: Springer)
[21] Virieux J 1984 SH-wave propagation in heterogeneous media: velocity-stress finite-difference method *Geophysics* **49** 1933–42
[22] Robertsson J O A, Blanch J O and Symes W W 1994 Viscoelastic finite-difference modeling *Geophysics* **59** 1444–56
[23] Chow Y T, Ito K, Liu K and Zou J 2015 Direct sampling method for diffusive optical tomography *SIAM J. Sci. Comput.* **37** 1658–A1684

[24] Chow Y T, Ito K and Zou J 2018 A time-dependent direct sampling method for recovering moving potentials in a heat equation *SIAM J. Sci. Comput.* **40** 2720–A2748

[25] Baudouin L, de Buhan M and Ervedoza S 2017 Convergent algorithm based on Carleman estimates for the recovery of a potential in the wave equation *SIAM J. Numer. Anal.* **55** 1578–613

[26] Klibanov M V, Kolesov A E and Nguyen D L 2019 Convexification method for an inverse scattering problem and its performance for experimental backscatter data for buried targets *SIAM J. Imaging Sci.* **12** 576–603

[27] Beilina L, Trung Thành N T, Klibanov M V and Malmberg J B 2014 Reconstruction of shapes and refractive indices from backscattering experimental data using the adaptivity *Inverse Problems* **30** 105007

[28] Lakhal A 2010 A decoupling-based imaging method for inverse medium scattering for Maxwell's equations *Inverse Problems* **26** 015007