Acousto-Electric Tomography with Total Variation Regularization

Bolaji James Adesokan[*]   Bjørn Jensen[*†]   Bangti Jin[‡]   Kim Knudsen[*]

October 16, 2018

Abstract

We study the numerical reconstruction problem in acousto-electric tomography of recovering the conductivity distribution in a bounded domain from interior power density data. We propose a numerical method for recovering discontinuous conductivity distributions, by reformulating it as an optimization problem with L^1 fitting and total variation penalty subject to PDE constraints. We establish continuity and differentiability results for the forward map, the well-posedness of the optimization problem, and present an easy-to-implement and robust numerical method based on successive linearization, smoothing and iterative reweighing. Extensive numerical experiments are presented to illustrate the feasibility of the proposed approach.

Keywords: acousto-electric tomography, reconstruction, total variation

1 Introduction

Acousto-electric tomography (AET) is one promising hybrid data imaging modality that has received increasing interest in the last decade [46, 5, 45]. It exploits the acousto-electric effect [32, 20, 31], i.e., the occurrence of small, localized changes in conductivities in the interior of a body due to a focused ultrasonic wave generated in the exterior. When the ultrasound is induced in combination with Electrical Impedance Tomography (EIT), one may reconstruct the interior power density data from EIT measurements [6], which can then be used for imaging the internal conductivity. The availability of internal data greatly improves the resolution and contrast in the reconstructions when compared with conventional EIT.

Mathematically, AET can be formulated as follows. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be an open bounded domain with a Lipschitz boundary Γ. The conductivity inside Ω is denoted by σ. Applying a current field f_j, $j = 1, \ldots, n$, to Γ induces an interior electric potential u_j given as the solution to

$$-\nabla \cdot (\sigma \nabla u_j) = 0 \text{ in } \Omega, \quad \text{with } \sigma \frac{\partial u_j}{\partial \nu} = f_j \text{ on } \Gamma.$$

Here ν denotes the unit outward normal vector on Γ. The AET forward problem is now to find the interior power density defined by $H_j(\sigma) = \sigma |\nabla u_j|^2$ from knowledge of σ, $j = 1, \ldots, n$. The corresponding inverse problem is, from knowledge of $H_j(\sigma)$ (and f_j), to recover the conductivity σ. The issues of unique recovery and stability have been extensively studied (see, e.g., [9, 3] and references therein). For example, in two spatial dimensions uniqueness is known for three properly chosen boundary conditions ensuring a non-vanishing Jacobian condition in the interior [3].

Various aspects of numerical reconstruction in AET have been considered [6, 22, 33, 27, 11]. An algorithm for recovering the conductivity from multiple power densities was proposed in [6]. It essentially relies on a perturbation approach and thus most suitable for small inclusions (relative to a known

background). Two optimal control formulations for reconstructing the conductivity were proposed in [15], and extensive numerical results to illustrate their effectiveness were also given there. The work [12] proposed a Levenberg-Marquardt iteration for numerical inversion, analyzed the convergence and regularizing properties of the algorithm in a Hilbert space setting (i.e., $H^s(\Omega)$ with $s > d/2$), under the assumptions that the derivative of the forward operator is injective and the data is noise-free. The case of limited angle boundary data was considered in [28]. The explicit formulation of the reconstruction problems as a regularized output least-squares problem was done in [2] and taken further to Perona-Malik type edge enhancing regularizers [40]. See also, e.g., [10, 38], for the case of anisotropic conductivities.

Total variation penalty is extremely popular for image processing. Since the seminal work [41] on image denoising, it has also been widely applied to solving inverse problems. There are several works on nonlinear parameter identifications for PDEs with total variation penalty, e.g., underground water flow [16], electrical impedance tomography [26] and quantitative photo-acoustic tomography [25], which have inspired the present work. Note that in these works, typically an $L^2(\Omega)$ fitting term is employed, which is well suited for theoretical considerations and numerical computations.

The main focus of this work is to reconstruct conductivities that are mostly piecewise constants using a PDE constrained optimal control formulation with a total variation penalty. Our main contributions are as follows. First, we provide a proper functional analytic setting for the reconstruction problem with discontinuous conductivity distributions. This is achieved by carefully analyzing the parameter-to-data map, e.g., continuity and differentiability in Theorem 3.2. The analysis relies crucially on the $W^{1,q}(\Omega)$ regularity of the state variable $u(\sigma)$ in Theorem 3.1. Second, we formulate the reconstruction problem as an optimization problem on an $L^1(\Omega)$ fitting term and total variation penalty term:

$$\min_{\sigma} \mathcal{J}_\beta(\sigma) = \sum_{j=1}^n \|H_j(\sigma) - z_j\|_{L^1(\Omega)} + \beta|\sigma|_{\mathrm{TV}},$$

over a suitable admissible set, and analyze the well-posedness of the formulation, e.g., existence and stability of minimizers in Theorems 4.1 and 4.2. Third, we describe an easy-to-implement numerical algorithm based on recursive linearization, smoothing and iterative reweighing, cf. Algorithm 1. Fourth, we present extensive numerical experiments with full and partial data to illustrate the effectiveness of the proposed approach. Further, we analyze the convergence of the discrete approximations by the Galerkin finite element method for both nonlinear and linearized models using suitable $W^{1,q}(\Omega)$ estimates on the finite element approximations in Lemmas A.1 and A.3.

The paper is organized as follows. In Section 2, we recall preliminary results on function spaces. Then in Section 3, we discuss mapping properties of the solution operator and parameter-to-data map, e.g., continuity and differentiability. In Section 4, we formulate the AET reconstruction into a PDE constrained optimization problem and analyze its analytic properties. In Section 5, we describe an algorithm for the numerical solution of the optimization problem. Last, in Section 6, we present extensive numerical results to illustrate the effectiveness of the reconstruction technique. In Appendix A, we give a finite element convergence analysis. Throughout, the notation C denotes a generic constant which may differ at each occurrence, but it is always independent of the mesh size h and other quantities under study.

2 Preliminaries on function spaces

This part reviews basic functional analytic tools and also fix the notation.

2.1 Sobolev spaces

First, we recall Sobolev spaces, which will be used extensively below. For any multi-index $\alpha \in \mathbb{N}^d$, $|\alpha|$ denotes the sum of all components. Given a domain $\Omega \subset \mathbb{R}^d$ with a Lipschitz continuous boundary Γ, for any $m \in \mathbb{N}$, $1 \leq p \leq \infty$, we follow [1] and define the Sobolev space $W^{m,p}(\Omega)$ by

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ for } 0 \leq |\alpha| \leq m\}.$$

It is equipped with the norm

$$\|u\|_{W^{m,p}(\Omega)} = \begin{cases} \left(\sum_{0 \le |\alpha| \le m} \|D^\alpha u\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}}, & \text{if } 1 \le p < \infty, \\ \max_{0 \le |\alpha| \le m} \|D^\alpha u\|_{L^\infty(\Omega)}, & \text{if } p = \infty. \end{cases}$$

The space $W_0^{m,p}(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $W^{m,p}(\Omega)$. Its dual space is denoted by $W^{-m,p'}(\Omega)$, with $1/p + 1/p' = 1$, i.e., p' is the conjugate exponent of p. Also we use $H^m(\Omega) = W^{m,2}(\Omega)$, and $H_0^m(\Omega) = W_0^{m,2}(\Omega)$. We denote by $(W^{m,p}(\Omega))'$ the dual space of $W^{m,p}(\Omega)$.

The bracket $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product, $(\cdot, \cdot)_{L^2(\Gamma)}$ the $L^2(\Gamma)$ inner product, and $\langle \cdot, \cdot \rangle$ dual pairing. The space $V \equiv H_\diamond^1(\Omega) \subset H^1(\Omega)$ consists of functions with zero mean on the boundary, that is

$$H_\diamond^1(\Omega) := \{u \in H^1(\Omega) : \int_\Gamma u \, ds = 0\} \simeq H^1(\Omega)/\mathbb{R}.$$

We denote by $H^{-\frac{1}{2}}(\Gamma)$ the dual space of $H^{\frac{1}{2}}(\Gamma)$, and $H_\diamond^{-\frac{1}{2}}(\Gamma) := \{v \in (H^{\frac{1}{2}}(\Gamma))' : \langle v, 1 \rangle = 0\}$. Functions in $H_\diamond^1(\Omega)$ satisfy the Poincaré type inequality

$$\|v\|_{H^1(\Omega)} \le C\|\nabla v\|_{L^2(\Omega)}, \quad \forall v \in H_\diamond^1(\Omega), \tag{2.1}$$

which can be derived by standard compactness arguments (see, e.g., [8, Section 5.4]).

2.2 Space of bounded variation

Below we shall use the total variation penalty in the regularized reconstruction, for which the proper function space is the space of bounded variation $BV(\Omega)$. We only describe some basic properties, and refer interested readers to [4, 8, 19] for details. The space $BV(\Omega)$ consists of functions $v \in L^1(\Omega)$ whose distributional derivative Dv is a Radon measure, i.e.,

$$BV(\Omega) = \{v \in L^1(\Omega) : |v|_{\mathrm{TV}} < \infty\},$$

where the total variation $|v|_{\mathrm{TV}}$ is defined by

$$|v|_{\mathrm{TV}} = \int_\Omega |Dv| = \sup\left\{\int_\Omega v \operatorname{div} \psi \, dx : \psi \in C_c^1(\Omega)^d, |\psi(x)| \le 1\right\},$$

and $|\cdot|$ denotes the Euclidean norm of vectors in $\mathbb{R}^d$. The space $BV(\Omega)$ is a Banach space when equipped with the norm

$$\|v\|_{BV(\Omega)} = \|v\|_{L^1(\Omega)} + |v|_{\mathrm{TV}}.$$

There are several different notions of convergence on the space $BV(\Omega)$. Besides the strong and the weak-$*$ topology on $BV(\Omega)$, there is the *intermediate* topology (also known as strict convergence), which is in between the two and characterized by strong $L^1(\Omega)$ convergence and convergence of $|\cdot|_{\mathrm{TV}}$ in $\mathbb{R}$, i.e. a sequence $\{v_n\} \subset BV(\Omega)$ converges in the intermediate sense to $v \in BV(\Omega)$ if $v_n \to v$ in $L^1(\Omega)$ and $|v_n|_{\mathrm{TV}} \to |v|_{\mathrm{TV}}$ as $n \to \infty$ [8, Definition 10.1.3, p. 374] [4, Definition 3.14, p. 125]. The intermediate convergence is very useful in practice, and we will equip $BV(\Omega)$ with this topology in the sequel.

The following results on $BV(\Omega)$ are very useful. The first assertion can be found at [8, Theorem 10.1.4, p. 378], the second at [19, Theorem 5.2, p. 199], and the third at [8, Theorem 10.1.2, p. 375]. For Assertion (i), the domain Ω has to satisfy suitable regularity, e.g., extension domain in the sense of [4, Definition 3.20, p. 130] (see, e.g., [4, Theorem 3.23, p. 132]). Any open set with a compact Lipschitz boundary is an extension domain [4, Proposition 3.21, p. 131].

Lemma 2.1. *The following properties hold on the space $BV(\Omega)$.*

(i) *The space $BV(\Omega)$ embeds compactly into $L^p(\Omega)$ for $p < \frac{d}{d-1}$.*

(ii) *The total variation is lower semi-continuous with respect to the convergence in $L^1(\Omega)$, i.e., if $\{v_n\} \subset BV(\Omega)$ and $v_n \to v$ in $L^1(\Omega)$, we have*

$$|v|_{\mathrm{TV}} \leq \liminf_{n \to \infty} |v_n|_{\mathrm{TV}}.$$

(iii) *The space $C^\infty(\overline{\Omega})$ is dense in $BV(\Omega)$ with respect to the convergence in the intermediate sense.*

3 Properties of the AET forward map

In this section we establish continuity and differentiability of the forward map in AET. For sufficiently regular conductivities, such results are well-known in the literature, but the issues become more delicate when non-smooth conductivities are considered.

For a fixed $\lambda \in (0, 1)$ we define the set

$$\mathcal{S} = \{\sigma : \lambda \leq \sigma \leq \lambda^{-1} \text{ a.e. in } \Omega\} \tag{A.1}$$

and we assume that $\sigma \in \mathcal{S}$. This assumption is physically reasonable and in particular it implies that $\sigma \in L^\infty(\Omega)$. We write $(\mathcal{S}, \tau)$ to denote the set endowed with a topology τ.

Given $f \in H_\diamond^{-\frac{1}{2}}(\Gamma)$ we consider the PDE problem

$$\begin{cases} -\nabla \cdot (\sigma \nabla u) = 0, & \text{in } \Omega, \\ \sigma \dfrac{\partial u}{\partial \nu} = f, & \text{in } \Gamma. \end{cases} \tag{3.1}$$

The weak formulation of problem (3.1) is to find $u \in V$ such that

$$(\sigma \nabla u, \nabla \psi) = \langle f, \psi \rangle \quad \forall \psi \in V.$$

Clearly, f must satisfy the compatibility condition $\langle f, 1 \rangle = 0$, which we shall assume throughout the rest of the paper independent of the space in which f is taken. By the Poincaré inequality (2.1), Lax-Milgram theorem implies that (3.1) has a unique weak solution $u = u(\sigma) \in V$ and

$$\|u(\sigma)\|_{H^1(\Omega)} \leq C\|f\|_{H^{-\frac{1}{2}}(\Gamma)},$$

where the constant C depends on λ and the domain Ω but is independent of σ. For a fixed f, we shall suppress the dependence of the solution u on f, and only indicate its dependence on σ by writing $u(\sigma)$. In passing, we note that as $\lambda \to 0^+$, problem (3.1) has a high-contrast conductivity coefficient, which is notoriously challenging to accurately resolve, and specialized numerical techniques are needed.

Next we recall a $W^{1,q}(\Omega)$ regularity result for elliptic problems; see [21] for a proof and also [13, 37] for closely related results. This result will play an important role in the analysis below.

Theorem 3.1. *Let $\sigma \in \mathcal{S}$ and suppose $g \in L^q(\Omega)$, $h \in L^q(\Omega)^d$ and $f \in (W^{1-\frac{1}{q}, q}(\Gamma))'$ with $\int_\Gamma f\, ds + \int_\Omega g\, dx = 0$. Then there exists a constant $Q > 2$, such that for any $q \in (2, Q)$, the problem*

$$\begin{cases} -\nabla \cdot (\sigma \nabla u) = g + \nabla \cdot h, & \text{in } \Omega, \\ \sigma \dfrac{\partial u}{\partial \nu} = f, & \text{on } \Gamma, \end{cases}$$

has a unique weak solution $u(\sigma) \in W^{1,q}(\Omega)$ satisfying

$$\|u\|_{W^{1,q}(\Omega)} \leq C(\|f\|_{(W^{1-\frac{1}{q}, q}(\Gamma))'} + \|g\|_{L^q(\Omega)} + \|h\|_{L^q(\Omega)^d}).$$

The constant $Q = Q(\lambda, d)$ depends only on the domain Ω, the spatial dimension d and the constant λ and the constant C depends only on Ω, λ, d and q.

Remark 3.1. *The parameter Q depends on the regularity of the domain Ω. If the domain Ω is of class C^1, then $Q(\lambda, d) \to \infty$ as $\lambda \to 1$ [24]. For a general Lipschitz domain, e.g., polyhedrons, there also always exists some $Q(\lambda, d) > 2$ for any $\lambda < 1$ [24, Section 5]. In the two-dimensional case, the optimal exponent Q was discussed in [7, 39].*

3.1 The continuity and differentiability of the solution map

In order to analyze the forward map $\sigma \mapsto H(\sigma)$, we first address the continuity of the solution operator $\sigma \mapsto u(\sigma)$. This map is identical with that for EIT and has been extensively studied in various function spaces [16, 30, 18]. Hence, we only sketch the proof for the convenience of readers.

Lemma 3.1. *Let $\{\sigma_k\} \subset \mathcal{S}$ satisfy $\sigma_k \to \sigma^*$ in $L^1(\Omega)$. Then the following statements hold.*

  (i) *If $f \in H_\diamond^{-\frac{1}{2}}(\Gamma)$, then there exists a subsequence of $\{u(\sigma_k)\}$ convergent to $u(\sigma^*)$ in $H^1(\Omega)$.*

  (ii) *If $f \in (W^{1-\frac{1}{r}, r}(\Gamma))'$ for some $r > 2$, then $u(\sigma_k) \to u(\sigma^*)$ in $W^{1,q}(\Omega)$ for any $q \in (2, \min(Q, r))$.*

Proof. Denote $u_k \equiv u(\sigma_k)$ and $u^* \equiv u(\sigma^*)$, and let $w_k = u_k - u^*$. By the weak formulations for u_k and u^*, we have

$$(\sigma_k \nabla w_k, \nabla \phi) = ((\sigma^* - \sigma_k)\nabla u^*, \nabla \phi) \quad \forall \phi \in V. \tag{3.2}$$

Now we discuss the two cases separately.

In case (i), letting $\phi = w_k \in V$ in (3.2) gives

$$\begin{aligned}(\sigma_k \nabla w_k, \nabla w_k) &= ((\sigma^* - \sigma_k)\nabla u^*, \nabla w_k) \\ &\leq \|\nabla w_k\|_{L^2(\Omega)} \|(\sigma_k - \sigma^*)\nabla u^*\|_{L^2(\Omega)}.\end{aligned}$$

By the standard measure theory, convergence in $L^p(\Omega)$, $p \geq 1$, implies almost everywhere convergence up to a subsequence [19, Theorem 1.21, p. 29]. Thus, one can extract a subsequence of $\{\sigma_k\}$, still denoted by $\{\sigma_k\}$, that converges almost everywhere in Ω. This, the trivial inequality $|\sigma_k - \sigma^*| \leq \lambda^{-1}$ a.e. for $\sigma_k, \sigma^* \in \mathcal{S}$ and Lesbesgue's dominated convergence theorem [19] imply

$$\lim_{k \to \infty} \|(\sigma_k - \sigma^*)\nabla u^*\|_{L^2(\Omega)} = 0.$$

This and the condition $\sigma_k \geq \lambda$ give the desired assertion in (i).

In case (ii), by Theorem 3.1, for any $q \in (2, \min(Q, r))$, there holds

$$\|\nabla w\|_{L^q(\Omega)} \leq C\|(\sigma_k - \sigma^*)\nabla u^*\|_{L^q(\Omega)}.$$

Since $q < \min(Q, r)$, we can choose $q' \in (q, \min(Q, r))$ and by Hölder inequality, we obtain

$$\|(\sigma_k - \sigma^*)\nabla u^*\|_{L^q(\Omega)} \leq \|\sigma_k - \sigma^*\|_{L^{p'}(\Omega)} \|\nabla u^*\|_{L^{q'}(\Omega)},$$

where the exponent p' satisfies $q'^{-1} + p'^{-1} = q^{-1}$. Since $\sigma_k, \sigma^* \in \mathcal{S}$, we have $|\sigma_k - \sigma^*| \leq \lambda^{-1}$ and

$$\|\sigma_k - \sigma^*\|_{L^{p'}(\Omega)} \leq \lambda^{\frac{1}{p'}-1} \|\sigma_k - \sigma^*\|_{L^1(\Omega)}^{\frac{1}{p'}},$$

from which it follows directly

$$\lim_{k \to \infty} \|\sigma_k - \sigma^*\|_{L^{p'}(\Omega)} = 0.$$

This and the uniform bound on $\|\nabla u^*\|_{L^{q'}(\Omega)}$ from Theorem 3.1 imply

$$\begin{aligned}0 \leq \lim_{k \to \infty} \|\nabla w_k\|_{L^q(\Omega)} &\leq \lim_{k \to \infty} \lambda^{-1}\|\sigma_k \nabla w_k\|_{L^q(\Omega)} \\ &\leq \lim_{k \to \infty} \lambda^{-1}\|(\sigma_k - \sigma^*)\nabla u^*\|_{L^q(\Omega)} = 0.\end{aligned}$$

This completes the proof of the lemma. □

Next we turn to the differentiability of the solution operator $\sigma \mapsto u(\sigma)$ for a fixed f. These properties are important for deriving necessary optimality conditions and developing numerical algorithms. However, one needs to be cautious: although the set $\mathcal{S}$ has an interior point with respect to the $L^\infty(\Omega)$ norm, it does not have any interior point with respect to the $L^p(\Omega)$ norm for any $1 \le p < \infty$. In the latter case, the results below have to be understood with respect to the relative topology. The next result gives the formula for the directional derivative $u'(\sigma)[\kappa]$ of $u(\sigma)$ at σ in the direction $\kappa \in L^\infty(\Omega)$.

Lemma 3.2. *For $\sigma \in \mathcal{S}$, the directional derivative $u'(\sigma)[\kappa]$ satisfies*

$$(\sigma \nabla u'(\sigma)[\kappa], \nabla \phi) = -(\kappa \nabla u(\sigma), \nabla \phi) \quad \forall \phi \in V. \tag{3.3}$$

If $f \in (W^{1-\frac{1}{r},r}(\Gamma))'$ for some $r > 2$, then for any $q \in (2, \min(Q,r))$ and $p > \frac{q\min(Q,r)}{\min(Q,r)-q}$, $u'(\sigma)[\cdot] : L^p(\Omega) \to W^{1,q}(\Omega)$ is continuous, and $u : (\mathcal{S}, L^p(\Omega)) \to W^{1,q}(\Omega)$ is Fréchet differentiable.

Proof. The expression of the directional derivative follows from direct computation. Specifically, let $h(t) = \frac{1}{t}(u(\sigma + t\kappa) - u(\sigma))$, $t > 0$. Assume that t is sufficiently small t so that $\lambda/2 < \lambda - t\|\kappa\|_{L^\infty}$. Now recalling the weak formulation for $u(\sigma + t\kappa)$ and $u(\sigma)$ appealing to the definition of $h(t)$ yield

$$(\sigma \nabla h(t), \nabla \phi) = -(\kappa \nabla u(\sigma + t\kappa), \nabla \phi) \quad \forall \phi \in V. \tag{3.4}$$

Letting $\phi = h(t)$ shows that $h(t)$ is uniformly bounded in $H^1(\Omega)$ as $t \to 0^+$. Further, we have $\|(\sigma + t\kappa) - \sigma\|_{L^p(\Omega)} = \lim_{t \to 0^+} t\|\kappa\|_{L^p(\Omega)} = 0$, for any $p > 0$. By Lemma 3.1, we deduce

$$\lim_{t \to 0^+} \|u(\sigma + t\kappa) - u(\sigma)\|_{H^1(\Omega)} = 0.$$

Thus letting $t \to 0^+$ in (3.4) yields the weak formulation for $u'(\sigma)[\kappa]$.

Now if $f \in (W^{1-\frac{1}{r},r}(\Gamma))'$, by Theorem 3.1, $u(\sigma) \in W^{1,s}(\Omega)$ for any $s \in [2, \min(Q,r))$. Thus, for any $q \in [2, \min(Q,r))$, we can choose any $q' \in (q, \min(Q,r))$, and $p'^{-1} + q'^{-1} = q^{-1}$, by Hölder's inequality, there holds

$$\|\kappa \nabla u(\sigma)\|_{L^q(\Omega)} \le \|\kappa\|_{L^{p'}(\Omega)} \|\nabla u(\sigma)\|_{L^{q'}(\Omega)}.$$

Then by Theorem 3.1, the solution $u'(\sigma)[\kappa]$ to (3.4) belongs to $W^{1,q}(\Omega)$ and

$$\|u'(\sigma)[\kappa]\|_{W^{1,q}(\Omega)} \le c\|\kappa\|_{L^{p'}(\Omega)} \|\nabla u(\sigma)\|_{L^{q'}(\Omega)}.$$

This shows the boundness of $u'(\sigma)[\kappa] : L^p(\Omega) \to W^{1,q}(\Omega)$. Since q' can be made arbitrarily close to $\min(Q,r)$, the desired continuity holds for any $p > \frac{q\min(Q,r)}{\min(Q,r)-q}$. Next we turn to Fréchet differentiability. Let $w(\sigma, \kappa) = u(\sigma + \kappa) - u(\sigma) - u'(\sigma)[\kappa]$. By the preceding argument, for any $q \in [2, \min(Q,r))$, the map $u'(\sigma)[\kappa] : L^p(\Omega) \to W^{1,q}(\Omega)$ is bounded, for any $p > \frac{q\min(Q,r)}{\min(Q,r)-q}$. It suffices to show

$$\lim_{\|\kappa\|_{L^p(\Omega)} \to 0} \frac{\|w(\sigma, \kappa)\|_{W^{1,q}(\Omega)}}{\|\kappa\|_{L^p(\Omega)}} = 0,$$

Next we take a sufficiently small κ (with $\sigma + \kappa \in \mathcal{S}$). Then the residual $w(\sigma, \kappa)$ satisfies

$$((\sigma + \kappa)\nabla w(\sigma, \kappa), \nabla \phi) = (\kappa \nabla u'(\sigma)[\kappa], \nabla \phi) \quad \forall \phi \in V. \tag{3.5}$$

By the preceding argument, we have

$$\|u'(\sigma)[\kappa]\|_{W^{1,q'}(\Omega)} \le c\|\kappa\|_{L^p(\Omega)},$$

where the exponent $q' \in (q, \min(Q,r))$ is sufficiently close to q. Therefore, by choosing p' large such that $p'^{-1} + q'^{-1} = q^{-1}$ and Hölder's inequality, there holds

$$\|\kappa \nabla u'(\sigma)[\kappa]\|_{L^q(\Omega)} \le \|\kappa\|_{L^{p'}(\Omega)} \|u'(\sigma)[\kappa]\|_{L^{q'}(\Omega)} \le c\|\kappa\|_{L^{p'}(\Omega)} \|\kappa\|_{L^p(\Omega)}.$$

This and Theorem 3.1 imply

$$\|w(\sigma,\kappa)\|_{W^{1,q}(\Omega)} \leq c\|\kappa\nabla u'(\sigma)[\kappa]\|_{L^q(\Omega)} \leq c\|\kappa\|_{L^{p'}(\Omega)}\|\kappa\|_{L^p(\Omega)}.$$

Now the desired assertion follows, in view of the inequality $\|\kappa\|_{L^{p'}(\Gamma)} \leq \lambda^{\frac{p}{p'}-1}\|\kappa\|_{L^p(\Omega)}^{\frac{p}{p'}}$. $\qquad\square$

Remark 3.2. *If $f \in H_\diamond^{-\frac{1}{2}}(\Gamma)$, the map $u(\sigma) : L^\infty(\Omega) \to H^1(\Omega)$ is holomorphic [17, p. 23]. Lemma 3.2 gives a slightly stronger result on the first derivative under the condition $f \in (W^{1-\frac{1}{r},r}(\Omega))'$. Our discussions have focused on the $L^p(\Omega)$ spaces, which is suitable for low-regularity penalties, e.g., total variation and $H^1(\Omega)$ penalty. More smoothing penalties, e.g., $H^2(\Omega)$, are also adopted [12]. Generally, if Ω has a C^{k+1}-boundary Γ, $\sigma, \kappa \in C^k(\overline{\Omega})$ and $f \in H^{k-\frac{1}{2}}(\Gamma)$ with $k > \frac{d}{2}$, then $u : H^k(\Omega) \to H^{k+1}(\Omega)$ is also Fréchet differentiable. This can be proved similarly [12].*

3.2 The continuity and differentiability of the forward map

The next result gives the continuity of the forward map $H(\sigma)$. It follows directly from Lemma 3.1.

Lemma 3.3. *Let $\{\sigma_k\} \subset \mathcal{S}$ be convergent to $\sigma^* \in \mathcal{S}$ in $L^1(\Omega)$. Then the following statements hold.*

(i) *If $f \in H_\diamond^{-\frac{1}{2}}(\Gamma)$, then there exists a subsequence of $\{H(\sigma_k)\}$ converging to $H(\sigma^*)$ in $L^1(\Omega)$.*

(ii) *If $f \in (W^{1-\frac{1}{r},r}(\Gamma))'$ for some $r > 2$, then $H(\sigma_k) \to H(\sigma^*)$ in $L^{\frac{q}{2}}(\Omega)$ for any $q \in [2, \min(Q, r))$.*

Proof. We split the difference $H(\sigma_k) - H(\sigma^*)$ into

$$H(\sigma_k) - H(\sigma^*) = (\sigma_k - \sigma^*)|\nabla u^*|^2 + \sigma_k(|\nabla u_k|^2 - |\nabla u^*|^2), \qquad (3.6)$$

and bound the two terms independently. We discuss case (i) and (ii) separately.

In case (i), by the $L^1(\Omega)$ convergence, we deduce that there exists a subsequence of $\{\sigma_k\}$, still relabelled as $\{\sigma_k\}$, that converges almost everywhere to σ^*. Since $|\sigma_k - \sigma^*| \leq \lambda^{-1}$, by Lebesgue's dominated convergence theorem and the condition $\sigma_k \to \sigma^*$ a.e., we deduce

$$\lim_{k\to\infty} \|(\sigma_k - \sigma^*)|\nabla u^*|^2\|_{L^1(\Omega)} = 0.$$

Meanwhile, by the triangle inequality and Hölder's inequality, there holds

$$\|\sigma_k(|\nabla u_k|^2 - |\nabla u^*|^2)\|_{L^1(\Omega)} \leq \lambda^{-1}\|(\nabla u_k - \nabla u^*)\cdot(\nabla u_k + \nabla u^*)\|_{L^1(\Omega)}$$
$$\leq \lambda^{-1}(\|\nabla u_k\|_{L^2(\Omega)} + \|\nabla u^*\|_{L^2(\Omega)})\|\nabla(u_k - u^*)\|_{L^2(\Omega)}.$$

The right hand side also tends to zero, in view of Lemma 3.1(i) and the fact that $\|\nabla u_k\|_{L^2(\Omega)}$ and $\|\nabla u^*\|_{L^2(\Omega)}$ are uniformly bounded independent of k.

The proof for case (ii) is similar. In the splitting (3.6), for any $q \in [2, \min(Q, r))$, the first term is now bounded using Hölder's inequality

$$\|(\sigma_k - \sigma^*)|\nabla u^*|^2\|_{L^{\frac{q}{2}}(\Omega)} \leq \|\sigma_k - \sigma^*\|_{L^{p'}(\Omega)}\|\nabla u^*\|_{L^{q'}(\Omega)}$$

where the exponent $q' \in (q, \min(Q, r))$ and the exponent $p' = \frac{qq'}{2(q'-q)} < \infty$. Now the assertion follows from $\|\sigma_k - \sigma^*\|_{L^{p'}(\Omega)} \leq \lambda^{-1+1/p'}\|\sigma_k - \sigma^*\|_{L^1(\Omega)}^{1/p'}$, cf. the proof of Lemma 3.1. Similarly, by the triangle inequality and Hölder's inequality, the second term can be bounded by

$$\|\sigma_k(|\nabla u_k|^2 - |\nabla u^*|^2)\|_{L^{\frac{q}{2}}(\Omega)} \leq \lambda^{-1}\|(\nabla u_k - \nabla u^*)\cdot(\nabla u_k + \nabla u^*)\|_{L^{\frac{q}{2}}(\Omega)}$$
$$\leq \lambda^{-1}(\|\nabla u_k\|_{L^q(\Omega)} + \|\nabla u^*\|_{L^q(\Omega)})\|\nabla(u_k - u^*)\|_{L^q(\Omega)}.$$

This and Lemma 3.1(ii) complete the proof of the lemma. $\qquad\square$

Last, we give the directional derivative $H'(\sigma)[\kappa]$ of the forward map $H(\sigma)$ and its adjoint $H'(\sigma)^*\xi$.

Theorem 3.2. *The directional derivative $H'(\sigma)[\kappa]$ is given by*

$$H'(\sigma)[\kappa] = \kappa|\nabla u(\sigma)|^2 + 2\sigma\nabla u(\sigma)\cdot\nabla u'(\sigma)[\kappa].$$

For any $q \in [2, Q)$ and $\kappa \in L^\infty(\Omega)$, if $f \in (W^{1-\frac{1}{q},q}(\Gamma))'$, then $H'(\sigma)[\kappa] \in L^{\frac{q}{2}}(\Omega)$, and the adjoint $H'(\sigma)^*\zeta$ is given by

$$H'(\sigma)^*\zeta = |\nabla u|^2\zeta + 2\nabla u\cdot\nabla v,$$

where $v = v(\zeta)$ solves the problem

$$(\sigma\nabla v, \nabla\phi) = -(\sigma\zeta\nabla u, \nabla\phi) \quad \forall\phi \in V.$$

Proof. The formula for $H'(\sigma)[\kappa]$ is direct to derive. By Theorem 3.1, $u(\sigma) \in W^{1,q}(\Omega)$. Thus, $\kappa\nabla u(\sigma) \in L^q(\Omega)$, and applying Theorem 3.1 again gives $u'(\sigma)[\kappa] \in W^{1,q}(\Omega)$. Then by Hölder's inequality, we obtain the first assertion.

To get the adjoint $H'(\sigma)^*$, we employ the weak formulations for $u'(\sigma)[\kappa]$ and v, i.e.,

$$(\sigma\nabla u'(\sigma)[\kappa], \nabla\phi) = -(\kappa\nabla u(\sigma), \nabla\phi) \quad \forall\psi \in V,$$
$$(\sigma\nabla v, \nabla\phi) = -(\zeta\sigma\nabla u(\sigma), \nabla\phi) \quad \forall\phi \in V.$$

Thus, by taking $\phi = v$ and $\phi = u'(\sigma)[\kappa]$, there holds $(\kappa\nabla u(\sigma), \nabla v) = (\zeta\sigma\nabla u(\sigma), \nabla u'(\sigma)[\kappa])$. Consequently,

$$
\begin{aligned}
(H'(\sigma)^*[\zeta], \kappa) &= (\zeta, H'(\sigma)[\kappa]) \\
&= (\zeta, \kappa|\nabla u(\sigma)|^2 + 2\sigma\nabla u(\sigma)\cdot\nabla u'(\sigma)[\kappa]) \\
&= (\zeta|\nabla u(\sigma)|^2, \kappa) + 2(\nabla u(\sigma)\cdot\nabla v, \kappa).
\end{aligned}
$$

By the definition of the adjoint, we obtain the formula for $H'(\sigma)^*[\zeta]$. $\qquad\square$

The next result gives the compactness of the forward map $H(\sigma)$. Hence, the nonlinear inverse problem is indeed ill-posed, and regularization is needed for stable reconstruction.

Theorem 3.3. *If $f \in (W^{1-\frac{1}{r},r}(\Gamma))'$ for some $r > 2$, then the map $H(\sigma) : (\mathcal{S}, BV) \to L^1(\Omega)$ is compact.*

Proof. By Lemma 2.1, the space $BV(\Omega)$ embeds compactly into $L^1(\Omega)$. Then by Lemma 3.3, the map $\sigma \mapsto H(\sigma)$ is continuous from $(\mathcal{S}, L^1(\Omega))$ to $L^1(\Omega)$, under the given regularity on f. Hence, $H(\sigma) : (\mathcal{S}, BV(\Omega)) \to L^1(\Omega)$ is compact. $\qquad\square$

4 Regularized problem

Now we discuss the well-posedness of the optimization problem arising in regularized reconstruction by means of a total variation penalty. Like before, we focus our discussion on one single dataset, i.e., $n = 1$, since the extension to multiple datasets is easy. Let $z = H(\sigma^*)$ be the power density data corresponding to the true conductivity σ^*, possibly corrupted by noise. The optimization problem reads:

$$\min_{\sigma\in\mathcal{A}} \left\{ \mathcal{J}_\beta(\sigma) = \|H(\sigma) - z\|_{L^1(\Omega)} + \beta|\sigma|_{\mathrm{TV}} \right\} \tag{4.1}$$

where $\beta > 0$ is a regularization parameter controlling the tradeoff between the two terms. The admissible set $\mathcal{A}$ is given by

$$\mathcal{A} := \{\sigma \in BV(\Omega) : \lambda \le \sigma \le \lambda^{-1} \text{ a.e. in } \Omega\} \equiv BV(\Omega) \cap \mathcal{S}.$$

In the model (4.1), the L^1 fidelity is motivated by analytical consideration: for any $\sigma \in \mathcal{A}$, the power density $H(\sigma)$ is only ensured to be $L^1(\Omega)$, but not the usual $L^2(\Omega)$, cf. Lemma 3.3. Statistically speaking, L^1 fitting is robust to outliers in the data, and popular for handling impulsive noise.

Now we prove that the functional $\mathcal{J}_\beta$ is (sequentially) continuous with respect to the intermediate convergence and that there exists a minimizer to problem (4.1).

Lemma 4.1. *Let $f \in (W^{1-\frac{1}{r},r}(\Gamma))'$ for some $r > 2$, then $\mathcal{J}_\beta$ is continuous on $\mathcal{A}$ in the intermediate topology of $BV(\Omega)$.*

Proof. Let $\{\sigma_k\} \subset \mathcal{A}$ be a sequence convergent in the intermediate topology. Then by definition it converges in $L^1(\Omega)$ and $|\sigma_k|_{\mathrm{TV}} \to |\sigma^*|_{\mathrm{TV}}$. Meanwhile, by Lemma 3.3, $H(\sigma_k) \to H(\sigma^*)$. Thus we obtain the desired continuity. □

Theorem 4.1. *For any $f \in H_\diamond^{-\frac{1}{2}}(\Gamma)$, problem (4.1) has at least one minimizer.*

Proof. Since $\mathcal{J}_\beta$ is bounded from below by zero, there exists a minimizing sequence $\{\sigma_k\} \subset \mathcal{A}$ such that $\lim_{k\to\infty} \mathcal{J}_\beta(\sigma_k) = \inf_{\sigma\in\mathcal{A}} \mathcal{J}_\beta(\sigma)$. This implies $|\sigma_k|_{\mathrm{TV}} \leq C$ for some $C > 0$. By the $L^\infty(\Omega)$ bound on the admissible set $\mathcal{A}$ and boundedness of Ω, we have $\|\sigma_k\|_{BV(\Omega)} \leq C$. By the uniform bound of σ_k in $BV(\Omega)$ and the compact embedding of $BV(\Omega)$ into $L^1(\Omega)$ in Lemma 2.1(i), there exists a subsequence, again relabeled as σ_k, convergent to some σ^* in $L^1(\Omega)$. Now by Lemma 3.3(i), $H(\sigma_k) \to H(\sigma^*)$ in $L^1(\Omega)$ (possibly by first passing to an a.e. pointwise convergent subsequence). The rest of the proof follows as Lemma 4.1, except that we only obtain $|\sigma^*|_{\mathrm{TV}} \leq \liminf_{k\to\infty} |\sigma_k|_{\mathrm{TV}}$. □

Remark 4.1. *Note that the existence of a minimizer requires only $f \in H_\diamond^{-\frac{1}{2}}(\Gamma)$, when compared with the sequential lower semi-continuity, since the proof does not require intermediately convergent subsequences in $BV(\Omega)$. Now we briefly mention the optimality conditions for problem (4.1). By means of the adjoint method, formally we derive the following optimality system: with $\zeta := \frac{\sigma|\nabla u|^2 - z}{|\sigma|\nabla u|^2 - z|}$ and u solving (3.1), the governing equation for σ reads*

$$\begin{cases} -\nabla \cdot \left(|\nabla\sigma|^{-1}\nabla\sigma\right) = -\beta^{-1}H'(\sigma)^*[\zeta] & in\ \Omega, \\ |\nabla\sigma|^{-1}\partial_\nu\sigma = 0 & on\ \Gamma. \end{cases}$$

This is a 1-Laplace equation for σ. It is known that a BV solution to such problems exist only under certain conditions on the source term and boundary condition [36].

In view of the lower semicontinuity of the functional, we have the following stability and consistency results; see [42, 43, 29] for a proof in the general case.

Theorem 4.2. *The following statements hold.*

(i) *Let the sequence $\{z_j\}$ be convergent to z in $L^1(\Omega)$, and $\{\sigma_j\}$ the corresponding minimizer to the functional $\mathcal{J}_\beta$ with z_j in place of z. Then the sequence $\{\sigma_j\}$ contains a subsequence convergent to a minimizer of $\mathcal{J}_\beta$ in the intermediate topology in $BV(\Omega)$.*

(ii) *Let $\{\delta_j\} \subset \mathbb{R}^+$ with $\delta_j \to 0$, $\{z_j\}$ be a sequence satisfying $\|z_j - z^*\|_{L^1(\Omega)} = \delta_j$ for some exact data z^*, and σ_j be a minimizer to the functional $\mathcal{J}_{\beta_j}$ with z_j in place of z^*. If β_js satisfy*

$$\lim_{j\to\infty} \beta_j = 0 \quad and \quad \lim_{j\to\infty} \frac{\delta_j}{\beta_j} = 0,$$

then the sequence $\{\sigma_j\}$ contains a subsequence convergent to an $|\cdot|_{\mathrm{TV}}$-minimizing solution $\sigma^\dagger$ in the intermediate topology $BV(\Omega)$.

Remark 4.2. *The functional $\mathcal{J}_\beta$ is generally not convex, similar to the quadratic fitting in [15]. Specifically, let $j(\sigma) = \int_\Omega |\sigma|\nabla u(\sigma)|^2 - z|\,dx$. Fix $\sigma \in \mathcal{A}$, z and f such that $\sigma|\nabla u(\sigma)|^2 < z$ a.e. in Ω. Since $u(\alpha\sigma) = \alpha^{-1}u(\sigma)$, we have $j(\alpha\sigma) = \int_\Omega (z - \alpha^{-1}\sigma|\nabla u(\sigma)|^2)\,dx$. Since $\alpha \mapsto \alpha^{-1}$ is convex on $\mathbb{R}_+$, the map $\alpha \mapsto j(\alpha\sigma)$ is concave, and thus, j is not convex in σ.*

5 Numerical algorithm

In this part, we describe an algorithm for problem (4.1). Problem (4.1) is numerically challenging to solve since it involves a nonlinear forward map $H(\sigma)$, and $\mathcal{J}_\beta$ is nonsmooth and nonconvex. We develop an easy-to-implement and robust algorithm using the following approximations at each outer iteration: recursively linearizing the operator $H(\sigma)$, smoothing the nonsmooth terms and approximating the L^1-norms by quadratic terms. The resulting intermediate constrained quadratic optimization problems are then solved by the conjugate gradient method.

5.1 Derivation

First, we tackle the nonconvexity with the fitting term $\|H(\sigma) - z\|_{L^1(\Omega)}$ by linearizing the forward map $H(\sigma)$. Consider a small change κ from a fixed σ (e.g., current approximation). Then by Theorem 3.2, $H(\sigma + \kappa) \approx H(\sigma) + H'(\sigma)\kappa$. Hence, for a fixed σ, we obtain a linearized functional $J_{\sigma,\beta}(\kappa)$ defined by

$$\mathcal{J}_\beta(\sigma + \kappa) \approx J_{\sigma,\beta}(\kappa) = \|H'(\sigma)\kappa - d_\sigma\|_{L^1(\Omega)} + \beta|\sigma + \kappa|_{\mathrm{TV}}, \tag{5.1}$$

with $d_\sigma = z - H(\sigma)$. Accordingly, we define a new admissible set $\mathcal{A}_\sigma$ by $\mathcal{A}_\sigma := \{\kappa \in BV(\Omega) : \sigma + \kappa \in \mathcal{A}\}$ to accommodate the box constraint on σ. This step gives rise to a nonsmooth but convex functional $J_{\sigma,\beta}(\kappa)$ (in the increment κ), since both TV seminorm and $L^1(\Omega)$ fitting are nondifferentiable. The nonsmoothness renders the numerical treatment inconvenient, and there are several possible strategies to handle this. Below, we employ the commonly used smoothing:

$$|\sigma|_{\mathrm{TV},\epsilon} = \int_\Omega |D\sigma|_\epsilon, \quad \text{and} \quad \|\cdot\|_\epsilon = \int_\Omega |\cdot|_\epsilon\, dx, \quad |\cdot|_\epsilon = \sqrt{|\cdot|^2 + \epsilon^2},$$

where $\epsilon > 0$ controls the amount of smoothing, and denote by $J_{\sigma,\beta,\epsilon}(\kappa)$ the smoothed functional:

$$J_{\sigma,\beta,\epsilon}(\kappa) = \|H'(\sigma)\kappa - d_\sigma\|_\epsilon + \beta|\sigma + \kappa|_{\mathrm{TV},\epsilon}.$$

The next result shows that a minimizer of $J_{\sigma,\beta,\epsilon}$ converges to a minimizer of $J_{\sigma,\beta}$ as ϵ goes to zero.

Theorem 5.1. *For any $\epsilon > 0$, there exists at least one minimizer κ_ϵ to the functional $J_{\sigma,\beta,\epsilon}$, and any accumulation point of the sequence of minimizers $\{\kappa_\epsilon\}_{\epsilon>0}$ is a minimizer of $J_{\sigma,\beta}$, as $\epsilon \to 0^+$.*

Proof. The existence of a minimizer to $J_{\sigma,\beta,\epsilon}$ follows easily as Theorem 4.1. Clearly, $J_{\sigma,\beta}(\kappa) \le J_{\sigma,\beta,\epsilon}(\kappa) \le J_{\sigma,\beta}(\kappa) + 2|\Omega|\epsilon$. Let κ^* minimize $J_{\sigma,\beta}$ and κ_ϵ minimize $J_{\sigma,\beta,\epsilon}$. Then by the minimizing properties of κ^* and κ_ϵ to $J_{\sigma,\beta}$ and $J_{\sigma,\beta,\epsilon}$, respectively, there hold

$$J_{\sigma,\beta}(\kappa^*) \le J_{\sigma,\beta}(\kappa_\epsilon) \le J_{\sigma,\beta,\epsilon}(\kappa_\epsilon) \le J_{\sigma,\beta,\epsilon}(\kappa^*) \le J_{\sigma,\beta}(\kappa^*) + 2|\Omega|\epsilon.$$

Thus any $\{\kappa_{\epsilon_k}\}_{k\in\mathbb{N}} \subset \{\kappa_\epsilon\}_{\epsilon>0}$, $\epsilon_k \to 0$ as $k \to \infty$, minimizes $J_{\sigma,\beta}$, and by repeating the arguments in the proof of Theorem 4.1, we deduce that every convergent subsequence converges to a global minimzer. $\square$

Next we derive a simple weighting scheme to facilitate the minimization of the functional $J_{\sigma,\beta,\epsilon}$. Differentiating $J_{\sigma,\beta,\epsilon}$ with respect to κ yields

$$J'_{\sigma,\beta,\epsilon}(\kappa)\eta = \int_\Omega \frac{H'(\sigma)\kappa - d_\sigma}{|H'(\sigma)\kappa - d_\sigma|_\epsilon} H'(\sigma)\eta\, dx + \beta \int_\Omega \frac{\nabla(\sigma + \kappa)}{|\nabla(\sigma + \kappa)|_\epsilon} \cdot \nabla\eta\, dx.$$

This is a highly nonlinear equation in κ (or more precisely variational inequality, under the constraint $\kappa \in \mathcal{A}_\sigma$). Then we freeze denominators of the integrands at some $\kappa' \in \mathcal{A}_\sigma$:

$$J'_{\sigma,\beta,\epsilon,\kappa'}(\kappa)\eta = \int_\Omega \frac{H'(\sigma)\kappa - d_\sigma}{|H'(\sigma)\kappa' - d_\sigma|_\epsilon} H'(\sigma)\eta\, dx + \beta \int_\Omega \frac{\nabla(\sigma + \kappa)}{|\nabla(\sigma + \kappa')|_\epsilon} \cdot \nabla\eta\, dx.$$

This corresponds to the gradient for the following weighted quadratic problem

$$J_{\sigma,\beta,\epsilon,\kappa'}(\kappa) = \frac{1}{2}\int_\Omega w(\sigma,\kappa')|H'(\sigma)\kappa - d_\sigma|^2\,dx + \frac{\beta}{2}\int w_0(\sigma,\kappa')|\nabla(\sigma+\kappa)|^2\,dx$$
$$= \tfrac{1}{2}\|H'(\sigma)\kappa - d_\sigma\|_{L^2_w(\Omega)}^2 + \tfrac{\beta}{2}\|\nabla(\sigma+\kappa)\|_{L^2_{w_0}(\Omega)}^2, \tag{5.2}$$

with the weight functions $w(\sigma,\kappa')$ and $w_0(\sigma,\kappa)$ given by

$$w(\sigma,\kappa') = |H'(\sigma)\kappa' - d_\sigma|_\epsilon^{-1} \quad \text{and} \quad w_0(\sigma,\kappa') = |\nabla(\sigma+\kappa')|_\epsilon^{-1}. \tag{5.3}$$

The reweighing step is repeated several times until a suitable stopping criterion is reached. This whole procedure for updating the increment κ is in the same spirit of lagged diffusivity for total variation denoising [44] or iteratively reweighed least-squares. Numerically, it is fairly robust to the initialization. Clearly, the procedure easily generalizes to multiple data sets.

5.2 Numerical algorithm

To compute the update κ from the functional $J_{\sigma,\beta,\epsilon,\kappa'}$, we employ the conjugate gradient method. To this end, we first derive the matrix representation. Let $\{\phi_i\}$, $\{\psi_i\}$ and $\{\xi_i\}$ be given bases for discretizing κ, σ and d_σ, respectively. Let

$$\kappa(x) = \sum_i \kappa_i \phi_i(x), \quad \sigma(x) = \sum_i \sigma_i \psi_i(x) \quad \text{and} \quad d_\sigma(x) = \sum_i (d_\sigma)_i \xi_i(x).$$

In our implementation, we take the standard P1 finite element basis as the basis for all three functions; and we refer to Appendix A for a convergence analysis of the finite element approximations. Next we define the corresponding stiffness and mass matrices respectively by

$$(\mathbf{K}_{w_0})_{ij} = \int_\Omega w_0 \nabla\psi_i \cdot \nabla\psi_j\,dx, \quad (\mathbf{L})_{ij} = \int_\Omega w H'_h(\sigma)\phi_i H'_h(\sigma)\phi_j\,dx, \quad \text{and} \quad (\mathbf{U})_{ij} = \int_\Omega w\xi_i H_h(\sigma)\phi_j\,dx.$$

Note that due to the locality of the standard P1 finite element basis functions ψ_i, the (stiffness) matrix $\mathbf{K}_{w_0}$ is sparse. However, the (discrete) operator $H'_h(\sigma)$ maps a localized basis ϕ_i to an almost everywhere nonzero function $H'_h(\sigma)\phi_i$, the matrices $\mathbf{L}$ is generally not sparse. Then upon expansion and with the vector representation of the variables, problem (5.2) reads

$$\min_{\boldsymbol{\kappa}} \boldsymbol{\kappa}^T(\mathbf{L} + \beta\mathbf{K}_{w_0})\boldsymbol{\kappa} - 2(\mathbf{d}_\sigma^T\mathbf{U} - \beta\boldsymbol{\sigma}^T\mathbf{K}_{w_0})\boldsymbol{\kappa} = \min_{\boldsymbol{\kappa}} \boldsymbol{\kappa}^T\mathbf{H}\boldsymbol{\kappa} - 2\mathbf{h}^T\boldsymbol{\kappa},$$

where $\mathbf{H} = \mathbf{L} + \beta\mathbf{K}_{w_0}$ and $\mathbf{h} = \mathbf{U}^T\mathbf{d}_\sigma - \beta\mathbf{K}_{w_0}\boldsymbol{\sigma}$, and the optimality condition simply becomes $\mathbf{H}\boldsymbol{\kappa} - \mathbf{h} = 0$. Since the matrix $\mathbf{H}$ is not sparse, it is costly to form and invert. Thus, in the implementation, rather than forming $\mathbf{H}$ explicitly, it is beneficial to access $\mathbf{L}$ only by matrix-vector product, and then to solve the linear system iteratively. The conjugage gradient method is a natural choice since $\mathbf{H}$ is symmetric and positive definite. To improve the numerical stability, one may add a small $\delta\mathbf{I}$ to $\mathbf{H}$, which allows to compensate the potential non-positivity of the matrix $\mathbf{H}$ due to numerical errors.

Now we can present the detailed procedure in Algorithm 1. There are two stopping criteria for the algorithm: one for the iteratively reweighed least-squares for updating the increment κ, and the other for the linearization at the outer iteration. Either can be based on the the relative change of the increment κ, and the second can be based on the magnitude of the derivative of $\mathcal{J}_\beta$ at σ in direction κ'. For the conjugate gradient iteration at Step 6, we initialize the iteration with κ', which can be fine tuned to be $\mu\kappa'$ for some $\mu > 0$, if needed, in a manner similar to the heavy ball method. In our numerical experiments, this choice is a good warm starting strategy.

6 Numerical results and discussions

Now we present numerical experiments to illustrate the proposed reconstruction algorithm.

Algorithm 1 Numerical algorithm for Problem (4.1).

1: Set initial guess σ_0, $\epsilon > 0$ and the maximum number of iterations I and K.
2: **for** $k = 1, \ldots, K$ **do**
3:   Set $\kappa' = 0$;
4:   **for** $i = 1, \ldots, I$ **do**
5:     Compute weights w and w_0 by (5.3);
6:     Solve (5.2) by the conjugate gradient method;
7:     Check stopping criterion;
8:     Set $\kappa' = \kappa$;
9:   **end for**
10:   Update $\sigma_{k+1} = \sigma_k + \kappa'$;
11:   Check stopping criterion;
12: **end for**

6.1 Experimental setting

The algorithm is implemented in Python using the *DOLFIN 2017.2.0* (FEniCS) package [35, 34]. Operators involving solving PDE's are formulated in the language of FEniCS and solved using standard solvers. The standard $P1$ finite elements are employed to approximate various functions. The meshes involved in generating simulated data and reconstruction are obtained from the public software package *gmsh 2.10.1*. All meshes are generated from a circle shape, but with various refinement levels given by the characteristic length h (i.e., mesh size). An overview of the statistics of the used meshes is given in Table 1. In the experiments, we consider the following boundary data:

$$f_1(x_1, x_2) = x_1, \quad f_2(x_1, x_2) = x_2, \quad f_3(x_1, x_2) = \tfrac{x_1 + x_2}{\sqrt{2}}, \quad f_4(x_1, x_2) = \tfrac{x_1 - x_2}{\sqrt{2}}.$$

The simulated power density data data $\mathcal{H} = (H_1, H_2, H_3, H_4)$ corresponding to the boundary fluxes $\mathcal{F} = (f_1, f_2, f_3, f_4)$ are generated using a finer mesh M_1, and the reconstructions are performed on the coarser mesh M_2, in order to mitigate the so-called inverse crime. Generally, the fine mesh should accurately resolve the forward problem far below the data noise, and the coarse mesh can deliver an accuracy commensurate with the data noise.

Table 1: Mesh statistics.

	Nodes	Triangles	h
M_1	41690	84010	0.01
M_2	6523	13296	0.025

We consider two different phantoms: the brain phantom in Fig. 1a and the geometrical shapes phantom in Fig. 1b. The corresponding interior data are shown in Fig. 2. When implementing the algorithm, the maximum number I of inner iterations is fixed at $I = 3$, and the number of conjugate gradient iterations is also fixed at 3. Our experiments indicate that increasing these numbers does not improve much the reconstruction quality, and hence we do not delve into fine tuning these parameters. Further, it is observed that the iterates stay nonnegative throughout the iteration, and hence the choice of the constant λ does not play a crucial role. Throughout, the regularization parameter β is determined in a trial-and-error way. The noisy data $\tilde{\mathbf{H}}$ is generated componentwise according to

$$\tilde{\mathbf{H}} = \mathbf{H} + \delta_e \frac{|\mathbf{H}|}{|\mathbf{e}|} \mathbf{e}$$

where $\mathbf{e}$ is a vector of appropriate size with each entry following a standard Gaussian distribution, δ_e the relative noise level, and $|\cdot|$ denotes the Euclidean norm of a vector. It is worth noting that in hybrid data tomography, the data $\tilde{\mathbf{H}}$ is often obtained from inverting the boundary data, and thus physically

Tissue/material	σ	Color
air	0.4	white
scalp	0.5232	green
skull	0.2983	blue
spinal fluid	1.0143	red
gray matter	0.55946	yellow
white matter	0.32404	cyan

(a) Head model phantom　　(b) Shape phantom　　(c) σ-values for different tissues.

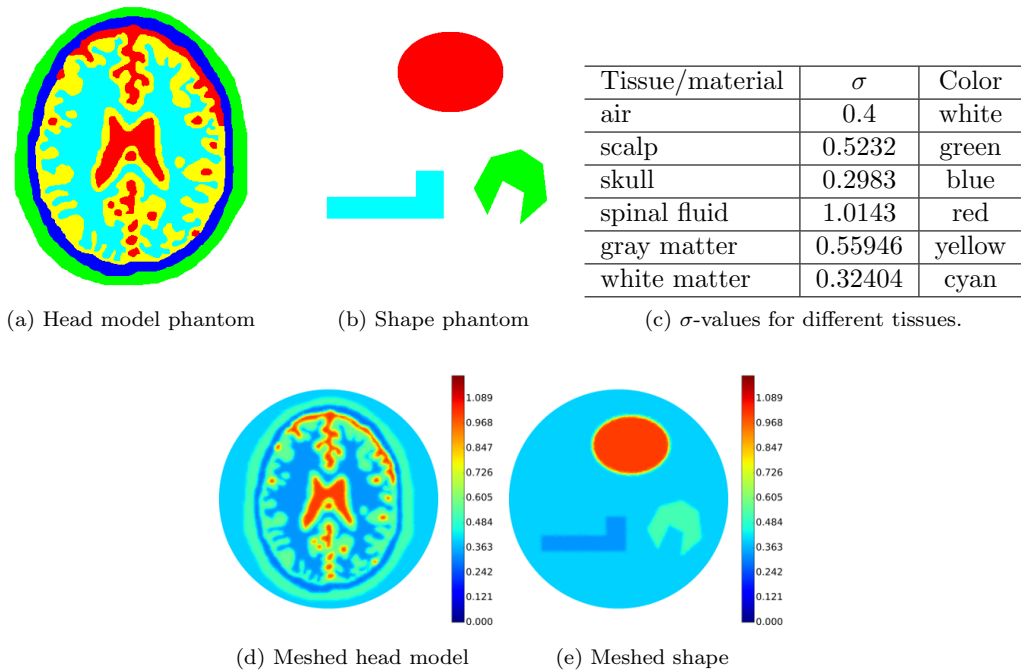(d) Meshed head model　　(e) Meshed shape

Figure 1: The two phantoms: (a) brain phantom and (b) geometric shapes, for simulated data and reconstruction, and (c) table of color-tissue correspondence.

realistic noise level is highly problem dependent. The choices of $\delta_e = 1\%$ and $\delta_e = 5\%$ below are typical for studying inversion algorithms.

6.2 Convergence of the algorithm

To gain insights into Algorithm 1, we present some numerical results on the convergence behavior in Fig. 3. In Figs. 3a and 3c, we show the evolution of the functional value $\mathcal{J}_\beta(\sigma_k)$ during the first 50 outer iterations, where the thick dashed line denotes the functional $\mathcal{J}_\beta(\sigma^*)$ evaluated at the true conductivity σ^* (interpolated at the reconstruction mesh). Due to regularization and discretization, it is not surprising that the true conductivity σ^* is generally not a global minimizer to $\mathcal{J}_\beta$. It is observed that the functional value decreases steadily as the iteration proceeds for both exact and noisy data, which shows clearly the robustness of the algorithm, and for exact data, eventually, the functional value falls below $\mathcal{J}_\beta(\sigma^*)$.

In Figs. 3b and 3d, we show the reconstruction errors in several metrics along the iteration. Given the choice of the space $BV(\Omega)$, we use the $L^1(\Omega)$-norm and total variation difference and a metric related to the intermediate topology of $BV(\Omega)$:

$$d_{BV}(\sigma, \eta) = \|\sigma - \eta\|_{L^1(\Omega)} + ||\sigma|_{\mathrm{TV}} - |\eta|_{\mathrm{TV}}|.$$

The peculiar jump in the plots at the beginning is related to the difference in total variation. This might be attributed to the fact that two functions with identical total variation can look anything alike. Thus it is mostly the tail of the plot, where the $L^1(\Omega)$-difference gets small, that is of significance. The error in total variation is dominating when compared with the $L^1(\Omega)$ error. The plot shows clearly that the convergence of the algorithm is quite steady for both exact and noisy data.
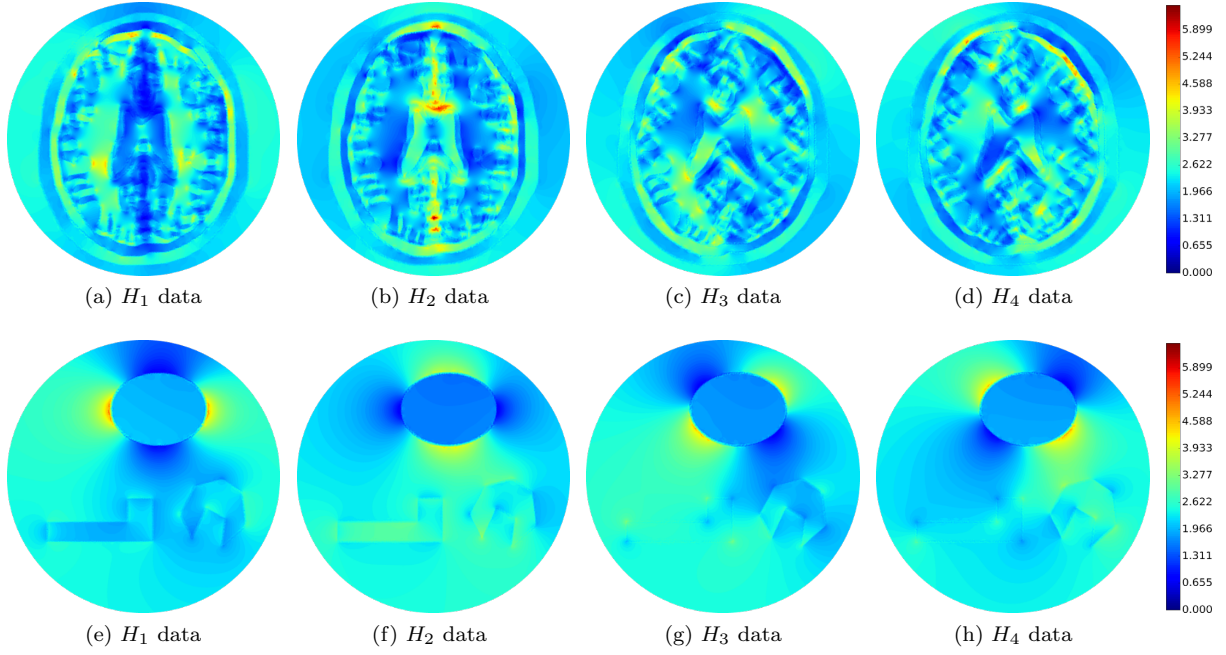
13

Figure 2: The power density data corresponding to f_i, $i = 1, \ldots, 4$ for each of the two phantoms.

(a) functional value $\mathcal{J}_\beta(\sigma_k)$

(b) errors

(c) functional value $\mathcal{J}_\beta(\sigma_k)$
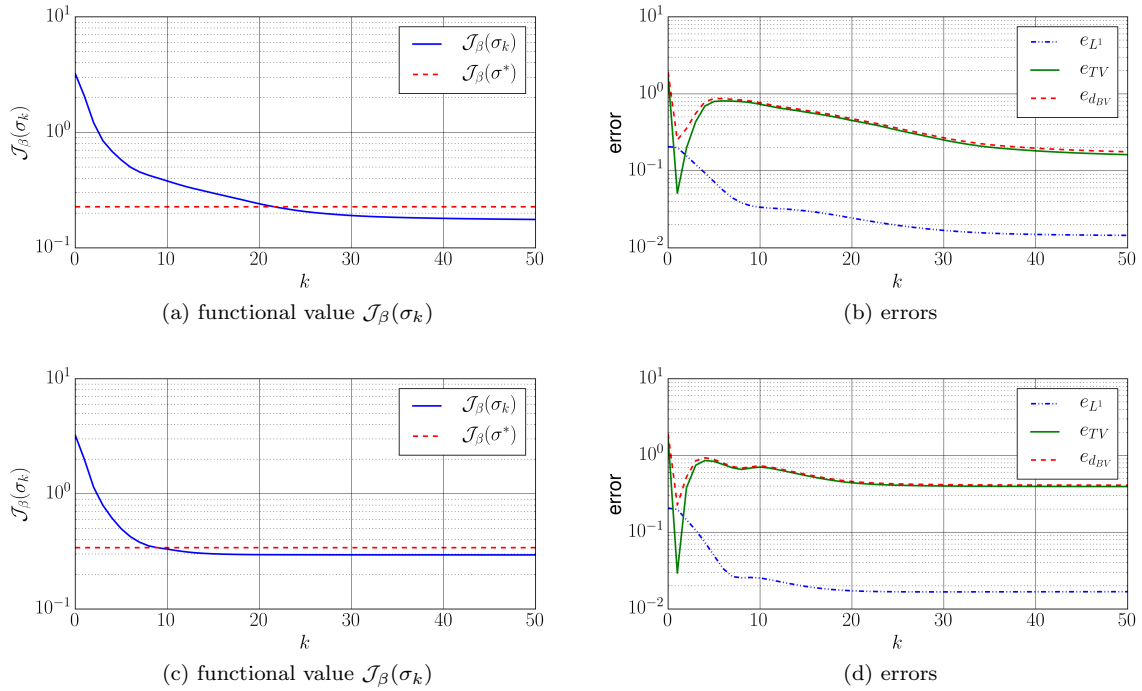
(d) errors

Figure 3: Convergence plots for the Shape phantom in Fig. 1b with exact data (top) and 1% noise (bottom). The parameters are taken to be $\beta = 3.5 \times 10^{-2}$ and $\epsilon = 10^{-4}$, $\lambda = 0.8$ and the data are H_1, H_2, H_3. The red line in (a) and (c) refers to $\mathcal{J}_\beta(\sigma^*)$. Figs. (b) and (d) show three metrics of the error $e_{L^1} := \|\sigma_k - \sigma^*\|_{L^1(\Omega)}$, $e_{TV} = \|\sigma_k|_{\mathrm{TV}} - |\sigma^*|_{\mathrm{TV}}|$ and $e_{d_{BV}} = d_{BV}(\sigma_k, \sigma^*)$.

14

6.3 Numerical reconstructions for full data

The reconstructions at two noise levels and different combinations of interior data are presented in Fig. 4, where the parameters for the reconstructions are shown in the captions. It is observed when the noise level increases, some fine details disappear in the reconstructions. Note that some details, e.g., the upper left and right part of the skull in Fig. 4f, can still be recovered by imposing less regularization, but at the cost of sacrificing the accuracy at the remaining part. The background in Fig. 4c appears slightly noisy, which is actually due to the discretization of the color-spectrum: 100 hues are used in the plots, but by changing it to 99 or 101, it is nearly completely uniform.

For the cases with only two boundary measurements, certain directional features are favored by the specified boundary data. The pairs (f_1, f_2) and (f_3, f_4) yields artifact that are different from each other, showing the influence of the choice of boundary data. For a microlocal analysis of the artifacts in the linearized model of AET, we refer interested readers to the recent work [11].
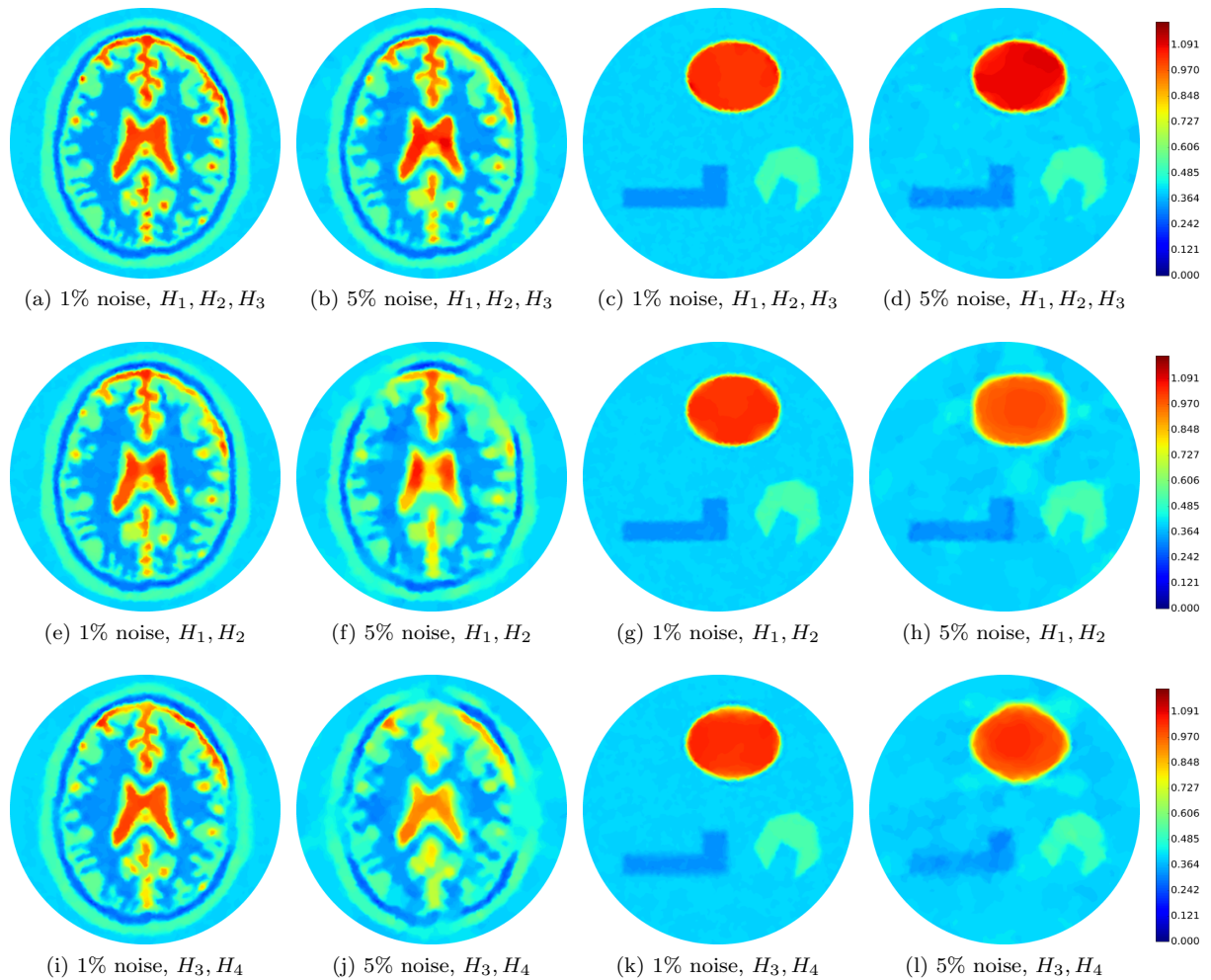

(a) 1% noise, H_1, H_2, H_3	(b) 5% noise, H_1, H_2, H_3	(c) 1% noise, H_1, H_2, H_3	(d) 5% noise, H_1, H_2, H_3
(e) 1% noise, H_1, H_2	(f) 5% noise, H_1, H_2	(g) 1% noise, H_1, H_2	(h) 5% noise, H_1, H_2
(i) 1% noise, H_3, H_4	(j) 5% noise, H_3, H_4	(k) 1% noise, H_3, H_4	(l) 5% noise, H_3, H_4

Figure 4: Reconstructions with different levels of noise, different amounts of data and different phantoms. Parameters: $\lambda = 0.8$, $\epsilon = 10^{-4}$, and $\beta = 3.5 \times 10^{-2}$ for the 1% noise case, $\beta = 0.7$ for the 5% noise case. The colorbar scale is $(0, 1.2)$.

6.4 Numerical reconstructions for partial interior data

The variational formulation in Section 4 extends straightforwardly to partial interior data, by restricting the integral in the fidelity term to a subdomain and with obvious modification, Algorithm 1 extends easily. To illustrate this flexibility, we consider two different subdomains for the available interior power density data: one small concentric disc and one half-disc. The corresponding numerical results are presented in Fig. 5, where the dashed red lines mark the boundary between subdomains with data and without. Notably, the reconstructions within the data-domains are fairly accurate, whereas the exterior has no significant updates during the iteration, and it is dominated completely by the initial guess σ_0.

Note that the scale of the colorbar in these plots is slightly larger compared to the other plots. This is due to an interesting effect happening near the boundary of the data-subdomain. Note that the inclusions in the phantom are reconstructed quite accurately near the center of the data-subdomain. However, as the boundary is approached, the values deviate from the background, seemingly in an attempt to compensate for the missing conductivity reconstruction outside the data-domain. However, the precise mechanism for the phenomenon is to be ascertained.

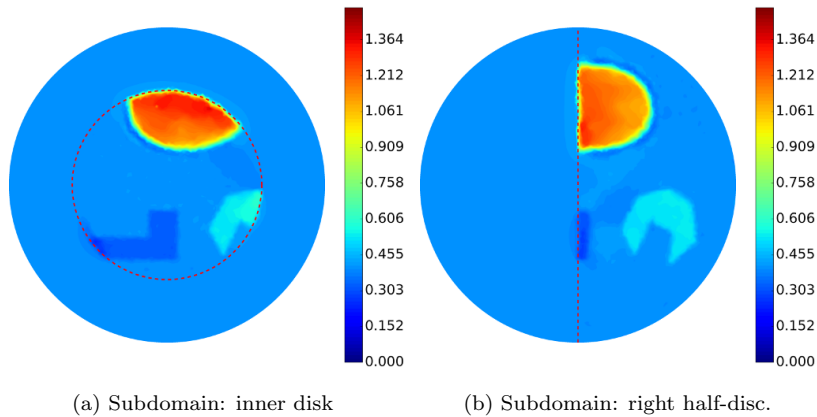(a) Subdomain: inner disk (b) Subdomain: right half-disc.

Figure 5: The reconstructions with data available on subdomains (the boundary is indicated by the dashed curves), obtained with 1% noise and three densities H_1, H_2 and H_3 data. Parameters: $\lambda = 0.8$, $\epsilon = 10^{-4}$, $\beta = 3.5 \times 10^{-2}$. The colorbar scale is $(0, 1.5)$.

7 Conclusion

In this paper we have studied the numerical reconstruction of acousto-electric tomography under very weak regularity assumptions on the conductivity. We have shown various continuity and differentiability results on the solutions of the elliptic PDE and the parameter-to-data map. We proposed a reconstruction algorithm based on total variation penalty, and showed the existence of a minimizer and its stability. Further, we proposed an algorithm based on recursively linearizing the forward map, smoothing and lagged diffusivity approximation, together with the conjugate gradient method for solving the resulting variational formulation. We have demonstrated the accuracy of the approach with extensive numerical experiments for both full and partial data.

8 Acknowledgement

We would like to thank Dr. Changyou Li for sharing with us the head phantom in Fig. 1a. The project was inspired by studies performed by Stratos Staboulis. We would like to thank Stratos for useful

A Convergence of finite element approximations

In this appendix, we discuss the convergence of the finite element approximation of the functional $\mathcal{J}_\beta$ and its linearization $J_{\beta,\sigma}$; see [23] for related analysis for electrical impedance tomography. Let $\mathcal{T}$ be a quasi-uniform partition of the domain Ω into simplicial elements, and consider the standard conforming piecewise linear finite element space:

$$X_h = \{\chi \in C(\overline{\Omega}) :\ \chi|_T \in \mathcal{P}_1\ \forall T \in \mathcal{T}_h\},$$

where $\mathcal{P}_1(T)$ denotes the space of linear functions on T, and $V_h := X_h \cap V$. The space V_h is used to discretize the state and adjoint variables, and X_h the conductivity σ. We denote by $I_h : C(\overline{\Omega}) \to X_h$ the standard Lagrangian nodal interpolation operator, and $R_h : V \to V_h$ the Ritz projection:

$$(\nabla R_h v, \nabla \phi) = (\nabla v, \nabla \phi) \quad \forall v \in V, \phi \in V_h.$$

The operators satisfy the following approximation properties:

$$
\begin{aligned}
\lim_{h \to 0} \|v - I_h v\|_{W^{1,p}(\Omega)} = 0 \quad &\forall v \in W^{1,p}(\Omega),\ \ p > d, \\
\lim_{h \to 0} \|v - R_h v\|_{H^1(\Omega)} = 0 \quad &\forall v \in V.
\end{aligned}
\tag{A.1}
$$

Below we discuss the discretization of the functionals $\mathcal{J}_\beta$ and $J_{\beta,\sigma}$ separately. Throughout the appendix, we assume $f \in (W^{1-\frac{1}{r},r}(\Gamma))'$, $r > 2$.

First we discuss the functional $\mathcal{J}_\beta$. For any $\sigma_h \in \mathcal{A}_h$, the discrete forward problem is to find $u_h \equiv u_h(\sigma_h) \in V_h$ such that

$$(\sigma_h \nabla u_h, \nabla \chi) = (f, \chi)_{L^2(\Gamma)} \quad \forall \chi \in V_h. \tag{A.2}$$

Then with $H_h(\sigma_h) = \sigma_h |\nabla u_h(\sigma_h)|^2$, the discrete analogue $\mathcal{J}_{\beta,h}$ of the functional $\mathcal{J}_\beta$ is given by

$$\mathcal{J}_{\beta,h}(\sigma_h) = \|H_h(\sigma_h) - z\|_{L^1(\Omega)} + \beta|\sigma_h|_{\mathrm{TV}},$$

which is to be minimized over the discrete admissible set $\mathcal{A}_h = \mathcal{A} \cap V_h$. It is easy to obtain the existence of a minimizer $\sigma_h^* \in \mathcal{A}_h$.

The following discrete analogue of Theorem 3.1 is useful.

Lemma A.1. *For any $\sigma_h \in \mathcal{A}_h$, there exist some $q \in (2, r)$ and some constant C independent h such that the solution $u_h(\sigma_h)$ to problem* (A.2) *satisfies*

$$\|u_h(\sigma_h)\|_{W^{1,q}(\Omega)} \leq C\|f\|_{(W^{1-\frac{1}{r},r}(\Gamma))'}.$$

Proof. By the repeating the argument of [14, Section 8.6] (the proof of Proposition 8.6.2), there exist some $C > 0$ and $q > 2$ sufficiently close to 2 such that for all $u_h \in V_h$, there holds

$$\|u_h\|_{W^{1,q}(\Omega)} \leq C \sup_{0 \neq v_h \in V_h} \frac{(\sigma_h \nabla u_h, \nabla v_h)}{\|v_h\|_{W^{1,q'}(\Omega)}}. \tag{A.3}$$

Taking $u_h = u_h(\sigma_h)$ in the inequality, (A.2) and Hölder's inequality yield the desired assertion. $\qquad\square$

We need a discrete analogue of Lemma 3.3 on the discrete forward map $\sigma_h \mapsto H_h(\sigma_h)$.

Lemma A.2. *Let the sequence $\{\sigma_h\}_{h>0} \subset \mathcal{A}_h \subset \mathcal{A}$ converge in $L^r(\Omega)$, $r \geq 1$, to some $\sigma \in \mathcal{A}$ as h tends to zero. Then $H_h(\sigma_h)$ converges to $H(\sigma)$ in $L^1(\Omega)$ as $h \to 0^+$.*

Proof. By Lax-Milgram theorem, u and u_h are uniformly bounded in $H^1(\Omega)$ independent of h. Setting the test function $\chi = R_h u - u_h \in V_h \subset V$ in the weak formulations and then subtracting them give

$$
\int_\Omega \sigma_h |\nabla(u - u_h)|^2 \, dx = - \int_\Omega (\sigma - \sigma_h) \nabla u \cdot \nabla(R_h u - u_h) \, dx
$$
$$
+ \int_\Omega \sigma_h \nabla(u - u_h) \cdot \nabla(u - R_h u) \, dx := \mathrm{I} + \mathrm{II}.
$$

It suffices to estimate the two terms I and II. For the term I, Hölder's inequality gives

$$
|\mathrm{I}| \leq \|\sigma - \sigma_h\|_{L^p(\Omega)} \|\nabla u\|_{L^q(\Omega)} \|\nabla(R_h u - u_h)\|_{L^2(\Omega)},
$$

where the exponent $q > 2$ is from Theorem 3.1, and $p^{-1} + q^{-1} = 2^{-1}$. Further, we note

$$
\|\nabla(R_h u - u_h)\|_{L^2(\Omega)} \leq C(\|u\|_{H^1(\Omega)} + \|u_h\|_{H^1(\Omega)}) \leq C.
$$

Thus, $\mathrm{I} \to 0$ as $h \to 0^+$, by the bound on σ_h and the convergence $\sigma_h \to \sigma^*$ in $L^r(\Omega)$. Further, by the bound on σ_h,

$$
|\mathrm{II}| \leq \|\sigma_h\|_{L^\infty(\Omega)} \|\nabla(u - u_h)\|_{L^2(\Omega)} \|\nabla(u - R_h u)\|_{L^2(\Omega)}
$$
$$
\leq \lambda^{-1} \|\nabla(u - u_h)\|_{L^2(\Omega)} \|\nabla(u - R_h u)\|_{L^2(\Omega)},
$$

which tends to zero, in view of (A.1). These two estimates imply $u_h \to u$ in $H^1(\Omega)$. Hence,

$$
\|H_h(\sigma_h) - H(\sigma)\|_{L^1(\Omega)} = \int_\Omega |\sigma_h |\nabla u_h(\sigma_h)|^2 - \sigma |\nabla u(\sigma)|^2| \, dx
$$
$$
\leq \int_\Omega |\sigma_h - \sigma| \, |\nabla u_h(\sigma_h)|^2 \, dx + \int_\Omega \sigma |(|\nabla u_h(\sigma_h)|^2 - |\nabla u(\sigma)|^2)| \, dx := \mathrm{III} + \mathrm{IV}
$$

By Lemma A.1 and the $L^\infty(\Omega)$ bound on $\mathcal{A}_h$, $\lim_{h\to 0^+} \mathrm{III} = 0$. The term IV also tends to zero due to $u_h \to u$ in $H^1(\Omega)$. This completes the proof. $\qquad \square$

The next result gives the convergence of the discrete minimizers σ_h^*.

Theorem A.1. *The sequence $\{\sigma_h^* \in X_h\}_{h>0}$ of minimizers to the discrete functionals $\mathcal{J}_{\beta,h}(\sigma_h)$ contains a subsequence converging in $L^1(\Omega)$ to a minimizer of $\mathcal{J}_\beta(\sigma)$ as $h \to 0^+$.*

Proof. Since $\sigma_h \equiv 1 \in \mathcal{A}_h$ for all h, the minimizing property of σ_h^* shows that the sequence $\{\mathcal{J}_{\beta,h}(\sigma_h^*)\}$ is uniformly bounded. Thus $\{|\sigma_h^*|_{\mathrm{TV}}\}$ is uniformly bounded, and there exists a subsequence, again denoted by $\{\sigma_h^*\}$, and some $\sigma^* \in \mathcal{A}$, such that $\sigma_h^* \to \sigma^*$ weak $*$ in $BV(\Omega)$. By Lemma 2.1(i), $\sigma_h^* \to \sigma^*$ in $L^1(\Omega)$. Lemma 2.1(ii) implies $|\sigma^*|_{\mathrm{TV}} \leq \liminf_{h\to 0} |\sigma_h^*|_{\mathrm{TV}}$. This and Lemma A.2 imply

$$
\mathcal{J}_\beta(\sigma^*) \leq \liminf_{h\to 0^+} \mathcal{J}_{\beta,h}(\sigma_h^*). \tag{A.4}
$$

For any $\sigma \in \mathcal{A}$, Lemma 2.1(iii) implies the existence of a sequence $\{\sigma^\epsilon\} \subset C^\infty(\overline{\Omega})$ such that $\int_\Omega |\sigma^\epsilon - \sigma| dx < \epsilon$ and $\left| \int_\Omega |\nabla \sigma^\epsilon| dx - \int_\Omega |D\sigma| \right| < \epsilon$. Let $\tilde{\sigma}_\epsilon = P_{[\lambda, \lambda^{-1}]}\sigma^\epsilon$, where $P_{[\lambda, \lambda^{-1}]}$ denotes pointwise projection. Since $\nabla \tilde{\sigma}^\epsilon = \nabla \sigma^\epsilon \chi_{\Omega_\epsilon}$ (with the set $\Omega_\epsilon = \{x \in \Omega : \lambda \leq \sigma^\epsilon \leq \lambda^{-1}\}$), which is uniformly bounded, and thus $\tilde{\sigma}^\epsilon \in \mathcal{A} \cap W^{1,\infty}(\Omega)$. The minimizing property of $\sigma_h^* \in \mathcal{A}_h$ gives $\mathcal{J}_h(\sigma_h^*) \leq \mathcal{J}_{\beta,h}(\mathcal{I}_h \tilde{\sigma}^\epsilon)$ for any $\epsilon > 0$. In view of (A.1), since $\tilde{\sigma}^\epsilon \in W^{1,\infty}(\Omega)$, we deduce $\lim_{h\to 0^+} \mathcal{I}_h \tilde{\sigma}^\epsilon = \tilde{\sigma}^\epsilon$ in $W^{1,1}(\Omega)$. Letting h to zero, and Lemma A.2 and (A.4) yield $\mathcal{J}_\beta(\sigma^*) \leq \mathcal{J}_\beta(\tilde{\sigma}^\epsilon)$. Then, by the contraction property of $P_{[\lambda, \lambda^{-1}]}$, there hold

$$
\int_\Omega |\nabla \tilde{\sigma}^\epsilon| \, dx = \int_{\Omega_\epsilon} |\nabla \sigma^\epsilon| \, dx \leq \int_\Omega |\nabla \sigma^\epsilon| \, dx \leq \int_\Omega |D\sigma| + \epsilon,
$$
$$
\int_\Omega |\tilde{\sigma}^\epsilon - \sigma| \, dx \leq \int_\Omega |\sigma^\epsilon - \sigma| \, dx < \epsilon.
$$

Letting ϵ to zero and Lemma 3.3 imply $\mathcal{J}_\beta(\sigma^*) \leq \mathcal{J}_\beta(\sigma)$ for any $\sigma \in \mathcal{A}$, completing the proof. $\qquad \square$

18

Next we discuss the discretization of the linearized problem (5.1) at some fixed $\sigma_0 \in \mathcal{A}$. With the approximation $u_h \equiv u_h(\sigma_0)$ defined by

$$(\sigma_0 \nabla u_h, \nabla \chi) = (f, \chi)_{L^2(\Gamma)} \quad \forall \chi \in V_h.$$

Then, for any $\kappa_h \in \mathcal{A}_{\sigma,h}$, find $v_h \equiv u_h'(\sigma_0)[\kappa_h] \in V_h$ such that

$$(\sigma_0 \nabla v_h, \nabla \chi) = -(\kappa_h \nabla u_h, \nabla \chi) \quad \forall \chi \in V_h.$$

Last, the discrete linearized functional $J_{\beta,h}$ of J_β (by omitting the subscript σ_0) reads

$$J_{\beta,h}(\kappa_h) = \|H_h'(\sigma_0)[\kappa_h] + H_h(\sigma_0) - z\|_{L^1(\Omega)} + \beta|\sigma_0 + \kappa_h|_{\mathrm{TV}},$$

where the linearized parameter-to-data map $H_h'(\sigma_0)[\kappa_h]$ is given by

$$H_h'(\sigma_0)[\kappa_h] = \kappa_h |\nabla u_h(\sigma_0)|^2 + 2\sigma_0 \nabla u_h(\sigma_0) \cdot \nabla u_h'(\sigma_0)[\kappa_h].$$

We need the following convergence result.

Lemma A.3. *There exists some $q > 2$ such that for any $\sigma \in \mathcal{A}$, $u_h(\sigma) \to u(\sigma)$ in $W^{1,q}(\Omega)$ as $h \to 0^+$.*

Proof. Let $u = u(\sigma)$ and $u_h = u_h(\sigma)$. Then for any $v_h \in V_h$, by the triangle inequality, there holds

$$\|u - u_h\|_{W^{1,q}(\Omega)} \le \|u - v_h\|_{W^{1,q}(\Omega)} + \|v_h - u_h\|_{W^{1,q}(\Omega)}.$$

It follows the inequality (A.3) and Galerkin orthogonality that with $q^{-1} + q'^{-1} = 1$

$$\begin{aligned}
\|u_h - v_h\|_{W^{1,q}(\Omega)} &\le C \sup_{0 \ne \chi \in V_h} \frac{(\sigma \nabla(u_h - v_h), \nabla \chi)}{\|\chi\|_{W^{1,q'}(\Omega)}} \\
&= C \sup_{0 \ne \chi \in V_h} \frac{(\sigma \nabla(u - v_h), \nabla \chi)}{\|\chi\|_{W^{1,q'}(\Omega)}} \le C\|u - v_h\|_{W^{1,q}(\Omega)},
\end{aligned}$$

where the last inequality is due to Hölder's inequality and the uniform bound on $\mathcal{A}$. Since the choice of $v_h \in V$ is arbitrary, combining the last two estimates gives

$$\|u - u_h\|_{W^{1,q}(\Omega)} \le C \inf_{v_h \in V_h} \|u - v_h\|_{W^{1,q}(\Omega)}.$$

Now the desired assertion follows by the density of V_h in $W^{1,q}(\Omega)$. $\square$

We have the following convergence for the finite element approximation.

Lemma A.4. *Let the sequence $\{\kappa_h\}_{h>0} \subset \mathcal{A}_{\sigma_0,h}$ converge in $L^r(\Omega)$, $r \ge 1$, to some $\kappa \in \mathcal{A}_\sigma$ as h tends to zero. Then $H_h'(\sigma_0)[\kappa_h] + H_h(\sigma_0) \to H'(\sigma_0)[\kappa] + H(\sigma_0)$ in $L^1(\Omega)$ as $h \to 0^+$.*

Proof. By Lemma A.3, $u_h \equiv u_h(\sigma_0) \to u(\sigma_0) := u$ in $W^{1,q}(\Omega)$ as $h \to 0^+$. Thus, $H_h(\sigma_0) = \sigma_0 |\nabla u_h|^2 \to \sigma_0 |\nabla u|^2 = H(\sigma_0)$ in $L^1(\Omega)$, and it suffices to show $H_h'(\sigma_0)[\kappa_h] \to H'(\sigma_0)[\kappa]$. By the triangle inequality and the $L^\infty(\Omega)$ bound on $\mathcal{A}_{\sigma_0}$ and $\mathcal{A}_{\sigma_0,h}$,

$$\begin{aligned}
\|H_h'(\sigma_0)[\kappa_h] - H'(\sigma_0)[\kappa]\|_{L^1(\Omega)} \le{}& \|\kappa_h |\nabla u_h(\sigma_0)|^2 - \kappa |\nabla u(\sigma_0)|^2\|_{L^1(\Omega)} \\
&+ C\|\nabla u_h(\sigma_0) \cdot \nabla u_h'(\sigma_0)[\kappa_h] - \nabla u(\sigma_0) \cdot \nabla u'(\sigma_0)[\kappa]\|_{L^2(\Omega)} := \mathrm{I} + \mathrm{II}.
\end{aligned}$$

By Hölder's inequality, the term I is bounded by (with $q > 2$ from Lemma A.1, and $p^{-1} + 2q^{-1} = 1$)

$$\mathrm{I} \le \|\kappa_h - \kappa\|_{L^p(\Omega)} \|\nabla u_h(\sigma_0)\|_{L^q(\Omega)}^2 + \|\kappa\|_{L^\infty(\Omega)} \||\nabla u_h(\sigma_0)|^2 - |\nabla u(\sigma_0)|^2\|_{L^1(\Omega)},$$

19

where both terms tend to zero, since $\kappa_h \to \kappa$ in $L^r(\Omega)$ and by Lemma A.3, $u_h(\sigma_0) \to u(\sigma_0)$ in $W^{1,q}(\Omega)$. Meanwhile, the term II is bounded by

$$
\begin{aligned}
\mathrm{II} &\le C\|\nabla(u_h(\sigma_0) - u(\sigma_0))\|_{L^2(\Omega)}\|\nabla u_h'(\sigma_0)[\kappa_h]\|_{L^2(\Omega)} \\
&+ C\|\nabla u(\sigma_0)\|_{L^2(\Omega)}\|\nabla(u_h'(\sigma_0)[\kappa_h] - u'(\sigma_0)[\kappa])\|_{L^2(\Omega)} := \mathrm{III} + \mathrm{IV}.
\end{aligned}
$$

By the uniform bound on κ_h, $\|u_h'(\sigma_0)[\kappa_h]\|_{L^2(\Omega)} \le C$ for some C independent of h, and thus the term $\mathrm{III} \to 0$ as $h \to 0^+$, in view of Lemma A.3. To bound the term IV, let $w_h \in V_h$ satisfy

$$
(\sigma_0 \nabla w_h, \nabla \chi) = (\kappa \nabla u(\sigma_0), \nabla \chi) \quad \forall \chi \in V_h.
$$

By Lemma A.3, there holds $\|\nabla w_h - \nabla u'(\sigma_0)[\kappa]\|_{L^2(\Omega)} \to 0$. Further, $v_h = w_h - u_h'(\sigma_0)[\kappa_h] \in V_h$ satisfies

$$
(\sigma_0 \nabla v_h, \nabla \chi) = (\kappa_h \nabla u_h(\sigma_0) - \kappa \nabla u(\sigma_0), \nabla \chi) \quad \forall \chi \in V_h.
$$

Letting $\chi = v_h$ and applying Cauchy-Schwarz inequality lead to

$$
\|\nabla v_h\|_{L^2(\Omega)} \le \lambda^{-1}\|\kappa_h \nabla u_h(\sigma_0) - \kappa \nabla u(\sigma_0)\|_{L^2(\Omega)}.
$$

Meanwhile, by Lemma A.3, the $L^\infty(\Omega)$ bound on κ_h and the fact $\kappa_h \to \kappa$ in $L^1(\Omega)$,

$$
\|\kappa_h \nabla u_h(\sigma_0) - \kappa \nabla u(\sigma_0)\|_{L^2(\Omega)} \to 0. \tag{A.5}
$$

Combining these estimates shows $\mathrm{IV} \to 0$ as $h \to 0^+$, which completes the proof of the lemma. $\qquad \square$

Last, we state the convergence of the discrete approximations to the linearized functional J_β. The proof is identical with that for Theorem A.1, but with Lemma A.4 in place of Lemma A.2.

Theorem A.2. *The sequence $\{\kappa_h^* \in \mathcal{A}_{\sigma_0,h}\}_{h>0}$ of minimizers to the the discrete functionals $J_{\sigma_0,\beta,h}(\kappa_h)$ contains a subsequence converging in $L^1(\Omega)$ to a minimizer of the functional $J_{\sigma_0,\beta}(\kappa)$ as $h \to 0^+$.*

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces.* Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] B. Adesokan, K. Knudsen, V. P. Krishnan, and S. Roy. A fully non-linear optimization approach to acousto-electric tomography. *Inverse Problems*, 34(10):104004, 16 pp., 2018.

[3] G. S. Alberti and Y. Capdeboscq. *Lectures on Elliptic Methods for Hybrid Inverse Problems.* Société Mathématique de France, Paris, 2018.

[4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* The Clarendon Press, Oxford University Press, New York, 2000.

[5] H. Ammari. *An Introduction to Mathematics of Emerging Biomedical Imaging.* Springer, Berlin, 2008.

[6] H. Ammari, E. Bonnetier, Y. Capdeboscq, M. Tanter, and M. Fink. Electrical impedance tomography by elastic deformation. *SIAM J. Appl. Math.*, 68(6):1557–1573, 2008.

[7] K. Astala, D. Faraco, and L. Székelyhidi, Jr. Convex integration and the L^p theory of elliptic equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 7(1):1–50, 2008.

[8] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces.* SIAM, Philadelphia, PA, 2006.

[9] G. Bal. Hybrid inverse problems and internal functionals. In *Inverse problems and applications: inside out. II*, volume 60 of *Math. Sci. Res. Inst. Publ.*, pages 325–368. Cambridge University Press, Cambridge, 2013.

[10] G. Bal, C. Guo, and F. Monard. Imaging of anisotropic conductivities from current densities in two dimensions. *SIAM J. Imaging Sci.*, 7(4):2538–2557, 2014.

[11] G. Bal, K. Hoffmann, and K. Knudsen. Propagation of singularities for linearised hybrid data impedance tomography. *Inverse Problems*, 34(2):024001, 19 pp., 2018.

[12] G. Bal, W. Naetar, O. Scherzer, and J. Schotland. The Levenberg-Marquardt iteration for numerical inversion of the power density operator. *J. Inverse Ill-Posed Probl.*, 21(2):265–280, 2013.

[13] B. V. Boyarskiĭ. Generalized solutions of a system of differential equations of first order and of elliptic type with discontinuous coefficients. *Mat. Sb. N.S.*, 43(85):451–503, 1957.

[14] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, third edition, 2008.

[15] Y. Capdeboscq, J. Fehrenbach, F. de Gournay, and O. Kavian. Imaging by modification: numerical reconstruction of local conductivities from corresponding power density measurements. *SIAM J. Imaging Sci.*, 2(4):1003–1030, 2009.

[16] Z. Chen and J. Zou. An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems. *SIAM J. Control Optim.*, 37(3):892–910, 1999.

[17] A. Cohen and R. DeVore. Approximation of high-dimensional parametric PDEs. *Acta Numer.*, 24:1–159, 2015.

[18] M. M. Dunlop and A. M. Stuart. The Bayesian formulation of EIT: analysis and algorithms. *Inverse Probl. Imaging*, 10(4):1007–1036, 2016.

[19] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, Boca Raton, FL, revised edition, 2015.

[20] F. E. Fox, K. F. Herzfeld, and G. D. Rock. The effect of ultrasonic waves on the conductivity of salt solutions. *Phys. Rev.*, 70:329–339, 1946.

[21] T. Gallouet and A. Monier. On the regularity of solutions to elliptic equations. *Rend. Mat. Appl. (7)*, 19(4):471–488, 1999.

[22] B. Gebauer and O. Scherzer. Impedance-acoustic tomography. *SIAM J. Appl. Math.*, 69(2):565–576, 2008.

[23] M. Gehre, B. Jin, and X. Lu. An analysis of finite element approximation in electrical impedance tomography. *Inverse Problems*, 30(4):045013, 24, 2014.

[24] K. Gröger. A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Math. Ann.*, 283(4):679–687, 1989.

[25] A. Hannukainen, N. Hyvönen, H. Majander, and T. Tarvainen. Efficient inclusion of total variation type priors in quantitative photoacoustic tomography. *SIAM J. Imaging Sci.*, 9(3):1132–1153, 2016.

[26] M. Hinze, B. Kaltenbacher, and T. N. T. Quyen. Identifying conductivity in electrical impedance tomography with total variation regularization. *Numer. Math.*, 138(3):723–765, 2018.

[27] K. Hoffmann and K. Knudsen. Iterative reconstruction methods for hybrid inverse problems in impedance tomography. *Sens. Imaging*, 15(1):art. id. 96, 2014.

[28] S. Hubmer, K. Knudsen, C. Li, and E. Sherina. Limited angle electrical impedance tomography with power density data. *Inv. Problems Sci. Eng.*, pages in press, `https://doi.org/10.1080/17415977.2018.1512983`, 2018.

[29] K. Ito and B. Jin. *Inverse Problems: Tikhonov Theory and Algorithms*. World Scientific, Hackensack, NJ, 2015.

[30] B. Jin and P. Maass. An analysis of electrical impedance tomography with applications to Tikhonov regularization. *ESAIM Control Optim. Calc. Var.*, 18(4):1027–1048, 2012.

[31] J. Jossinet, B. Lavandier, and D. Cathignol. The phenomenology of acousto-electric interaction signals in aqueous solutions of electrolytes. *Ultrasonics*, 36(1-5):607–613, 1998.

[32] F. Körber. Über den Einfluss des Druckes auf das elektrolytische Leitvermögen von Lösungen. *Z. Phys. Chem.*, 67(1):212–248, 1909.

[33] P. Kuchment and L. Kunyansky. 2D and 3D reconstructions in acousto-electric tomography. *Inverse Problems*, 27(5):055013, 21 pp., 2011.

[34] A. Logg and G. N. Wells. DOLFIN: automated finite element computing. *ACM Trans. Math. Software*, 37(2):Art. 20, 28 pp., 2010.

[35] A. Logg, G. N. Wells, T. F. Book, and G. N. Wells, editors. *Automated Solution of Differential Equations by the Finite Element Method*. Springer-Verlag, Berlin, 2012.

[36] A. Mercaldo, S. Segura de León, and C. Trombetti. On the solutions to 1-Laplacian equation with L1 data. *J. Funct. Anal.*, 256(8):2387–2416, 2009.

[37] N. G. Meyers. An L^pe-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 17:189–206, 1963.

[38] F. Monard and D. Rim. Imaging of isotropic and anisotropic conductivities from power densities in three dimensions. *Inverse Problems*, 34(7):075005, 26 pp., 2018.

[39] V. Nesi, M. Palombaro, and M. Ponsiglione. Gradient integrability and rigidity results for two-phase conductivities in two dimensions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 31(3):615–638, 2014.

[40] S. Roy and A. Borzì. A new optimization approach to sparse reconstruction of log-conductivity in acousto-electric tomography. *SIAM J. Imaging Sci.*, 11(2):1759–1784, 2018.

[41] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[42] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Springer, New York, 2009.

[43] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*. Walter de Gruyter GmbH & Co. KG, Berlin, 2012.

[44] C. R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17(1):227–238, 1996.

[45] T. Widlak and O. Scherzer. Hybrid tomography for conductivity imaging. *Inverse Problems*, 28(8):084008, 28 pp., 2012.

[46] H. Zhang and L. V. Wang. Acousto-electric tomography. In *Proc. SPIE 5320, Photons Plus Ultrasound: Imaging and Sensing*, pages doi: 10.1117/12.532610, 5 pp., 2004.