Adaptive FEM for Electrical Impedance Tomography with a Piecewise Constant Conductivity

Adaptive Reconstruction for Electrical Impedance Tomography with a Piecewise Constant Conductivity

Bangti Jin* Yifeng Xu†

Abstract

In this work we propose and analyze a numerical method for electrical impedance tomography of recovering a piecewise constant conductivity from boundary voltage measurements. It is based on standard Tikhonov regularization with a Modica-Mortola penalty functional and adaptive mesh refinement using suitable *a posteriori* error estimators of residual type that involve the state, adjoint and variational inequality in the necessary optimality condition and a separate marking strategy. We prove the convergence of the adaptive algorithm in the following sense: the sequence of discrete solutions contains a subsequence convergent to a solution of the continuous necessary optimality system. Several numerical examples are presented to illustrate the convergence behavior of the algorithm.

Keywords: electrical impedance tomography, piecewise constant conductivity, Modica-Mortola functional, *a posteriori* error estimator, adaptive finite element method, convergence analysis

1 Introduction

Electrical impedance tomography (EIT) aims at recovering the electrical conductivity distribution of an object from voltage measurements on the boundary. It has attracted much interest in medical imaging, geophysical prospecting, nondestructive evaluation and pneumatic oil pipeline conveying etc. A large number of reconstruction algorithms have been proposed; see, e.g., [37, 36, 26, 35, 31, 30, 24, 15, 21, 40, 39, 18, 3, 34, 33, 57, 28, 54, 51] for a rather incomplete list. One prominent idea underpinning many imaging algorithms is regularization, especially Tikhonov regularization [29]. In practice, they are customarily implemented using the continuous piecewise linear finite element method (FEM) on quasi-uniform meshes, due to its flexibility in handling spatially variable coefficients and general domain geometry. The convergence analysis of finite element approximations was carried out in [22, 47, 27].

In several practical applications, the physical process is accurately described by the complete electrode model (CEM) [14, 50]. It employs nonstandard boundary conditions to capture characteristics of the experiment. In particular, around the electrode edges, the boundary condition changes from the Neumann to Robin type, which, according to classical elliptic regularity theory [23], induces weak solution singularity around the electrode edges; see, e.g., [45] for an early study. Further, the low-regularity of the sought-for conductivity distribution, especially that enforced by a nonsmooth penalty, e.g., total variation, can also induce weak interior singularities of the solution. Thus, a (quasi-)uniform triangulation of the domain can be inefficient for resolving these singularities, and the discretization errors around electrode edges and internal interfaces can potentially compromise the reconstruction accuracy. These observations motivate the use of an adaptive strategy to achieve the desired accuracy in order to enhance the overall computational efficiency.

For direct problems, the mathematical theory of AFEM, including *a posteriori* error estimation, convergence and computational complexity, has advanced greatly [1, 44, 53, 12]. A common adaptive FEM (AFEM) consists of the following successive loops:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \tag{1.1}$$

*Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK (b.jin@ucl.ac.uk, bangti.jin@gmail.com)

†Department of Mathematics and Scientific Computing Key Laboratory of Shanghai Universities, Shanghai Normal University, Shanghai 200234, China. (yfxu@shnu.edu.cn, mayfxu@gmail.com)

The module ESTIMATE employs the given problem data and computed solutions to provide computable quantities on the local errors, and distinguishes different adaptive algorithms.

In this work, we develop an adaptive EIT reconstruction algorithm with a piecewise constant conductivity. In practice, the piecewise constant nature is commonly enforced by a total variation penalty. However, it is challenging for AFEM treatment (see e.g., [5] for image denoising). Thus, we take an indirect approach based on a Modica-Mortola type functional:

$$\mathcal{F}_\varepsilon(\sigma) = \varepsilon \int_\Omega |\nabla \sigma|^2 \,\mathrm{d}x + \frac{1}{\varepsilon} \int_\Omega W(\sigma) \,\mathrm{d}x,$$

where the constant $\varepsilon > 0$ is small and $W(s) : \mathbb{R} \to \mathbb{R}$ is the double-well potential, i.e.,

$$W(s) = (s - c_0)^2 (s - c_1)^2, \tag{1.2}$$

with $c_0, c_1 > 0$ being two known values that the conductivity σ can take. The functional $\mathcal{F}_\varepsilon$ Γ-converges to the total variation semi-norm [42, 43, 2]. The corresponding regularized least-squares formulation reads

$$\inf_{\sigma \in \widetilde{\mathcal{A}}} \left\{ \mathcal{J}_\varepsilon(\sigma) = \tfrac{1}{2} \|U(\sigma) - U^\delta\|^2 + \tfrac{\widetilde{\alpha}}{2} \mathcal{F}_\varepsilon(\sigma) \right\}, \tag{1.3}$$

where $\tilde{\alpha} > 0$ is a regularization parameter; see Section 2 for further details. In this work, we propose *a posteriori* error estimators and an adaptive reconstruction algorithm of the form (1.1) for (1.3) based on a separate marking using three error indicators in the module MARK; see Algorithm 3.1. Further, we give a convergence analysis of the algorithm, in the sense that the sequence of state, adjoint and conductivity generated by the adaptive algorithm contains a convergent subsequence to a solution of the necessary optimality condition. The technical proof consists of two steps: Step 1 shows the subsequential convergence to a solution of a limiting problem, and Step 2 proves that the solution of the limiting problem satisfies the necessary optimality condition. The main technical challenges in the convergence analysis include the nonlinearity of the forward model, the nonconvexity of the double well potential and properly treating the variational inequality. The latter two are overcome by pointwise convergence of discrete minimizers and Lebesgue's dominated convergence theorem, and AFEM analysis techniques for elliptic obstacle problems, respectively. The adaptive algorithm and its convergence analysis are the main contributions of this work.

Last, we situate this work in the existing literature. In recent years, several adaptive techniques, including AFEM, have been applied to the numerical resolution of inverse problems. In a series of works [6, 7, 8, 9], Beilina et al studied the AFEM in a dual weighted residual framework for parameter identification. Feng et al [20] proposed a residual-based estimator for the state, adjoint and control by assuming convexity of the cost functional and high regularity on the control. Li et al [38] derived *a posteriori* error estimators for recovering the flux and proved their reliability; see [55] for a plain convergence analysis. Clason et al [17] studied functional *a posteriori* estimators for convex regularized formulations. Recently, Jin et al [32] proposed a first AFEM for Tikhonov functional for EIT with an $H^1(\Omega)$ penalty, and also provided a convergence analysis. This work extends the approach in [32] to the case of a piecewise constant conductivity. There are a number of major differences between this work and [32]. First, the $H^1(\Omega)$ penalty in [32] facilitates deriving the *a posteriori* estimator on the conductivity σ, by completing the squares and suitable approximation argument, which is not directly available for the Mordica-Mortola functional $\mathcal{F}_\varepsilon$. Second, we develop a sharper error indicator associated with the crucial variational inequality than that in [32], by a novel constraint preserving interpolation operator [13]; see the proof of Theorem 5.5, which represents the main technical novelty of this work. Third, Algorithm 3.1 employs a separate marking for the estimators, instead of a collective marking in [32], which automatically takes care of different scalings of the estimators.

The rest of this paper is organized as follows. In Section 2, we introduce the complete electrode model, and the regularized least-squares formulation. In Section 3, we give the AFEM algorithm. In Section 4, we present extensive numerical results to illustrate its convergence and efficiency. In Section 5, we present the lengthy technical convergence analysis. Throughout, $\langle \cdot, \cdot \rangle$ and $(\cdot, \cdot)$ denote the inner product on the Euclidean space and $(L^2(\Omega))^d$, respectively, by $\|\cdot\|$ the Euclidean norm, and occasionally abuse $\langle \cdot, \cdot \rangle$ for the duality pairing between the Hilbert space $\mathbb{H}$ and its dual space. The superscript t denotes the transpose of a vector. The notation c denotes a generic constant, which may differ at each occurrence, but it is always independent of the mesh size and other quantities of interest.

2 Regularized approach

This part describes the regularized approach for recovering piecewise constant conductivities.

2.1 Complete electrode model (CEM)

Let Ω be an open bounded domain in $\mathbb{R}^d$ $(d = 2, 3)$ with a polyhedral boundary $\partial\Omega$. We denote the set of electrodes by $\{e_l\}_{l=1}^L$, which are line segments/planar surfaces on $\partial\Omega$ and satisfy $\bar{e}_i \cap \bar{e}_k = \emptyset$ if $i \neq k$. The applied current on electrode e_l is denoted by I_l, and the vector $I = (I_1, \ldots, I_L)^{\mathrm{t}} \in \mathbb{R}^L$ satisfies $\sum_{l=1}^L I_l = 0$, i.e., $I \in \mathbb{R}^L_\diamond := \{V \in \mathbb{R}^L : \sum_{l=1}^L V_l = 0\}$. The electrode voltage $U = (U_1, \ldots, U_L)^{\mathrm{t}}$ is normalized, i.e., $U \in \mathbb{R}^L_\diamond$. Then the CEM reads: given the conductivity σ, positive contact impedances $\{z_l\}_{l=1}^L$ and input current $I \in \mathbb{R}^L_\diamond$, find $(u, U) \in H^1(\Omega) \otimes \mathbb{R}^L_\diamond$ such that [14, 50]

$$\begin{cases} -\nabla \cdot (\sigma \nabla u) = 0 & \text{in } \Omega, \\ u + z_l \sigma \frac{\partial u}{\partial n} = U_l & \text{on } e_l, l = 1, 2, \ldots, L, \\ \int_{e_l} \sigma \frac{\partial u}{\partial n} \mathrm{d}s = I_l & \text{for } l = 1, 2, \ldots, L, \\ \sigma \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \backslash \cup_{l=1}^L e_l. \end{cases} \tag{2.1}$$

The inverse problem is to recover the conductivity σ from a noisy version U^δ of the electrode voltage $U(\sigma^\dagger)$ (for the exact conductivity $\sigma^\dagger$) corresponding to one or multiple input currents.

Below the conductivity σ is assumed to be piecewise constant, i.e., in the admissible set

$$\mathcal{A} := \{\sigma \in L^\infty(\Omega) : \ \sigma = c_0 + (c_1 - c_0)\chi_{\Omega_1}\},$$

where the constants $c_1 > c_0 > 0$ are known, $\overline{\Omega}_1 \subset \Omega$ is an unknown open set with a Lipschitz boundary and χ_{Ω_1} denotes its characteristic function. We denote by $\mathbb{H}$ the space $H^1(\Omega) \otimes \mathbb{R}^L_\diamond$ with its norm given by

$$\|(u, U)\|^2_{\mathbb{H}} = \|u\|^2_{H^1(\Omega)} + \|U\|^2.$$

A convenient equivalent norm on the space $\mathbb{H}$ is given below.

Lemma 2.1. *On the space $\mathbb{H}$, the norm $\|\cdot\|_{\mathbb{H}}$ is equivalent to the norm $\|\cdot\|_{\mathbb{H},*}$ defined by*

$$\|(u, U)\|^2_{\mathbb{H},*} = \|\nabla u\|^2_{L^2(\Omega)} + \sum_{l=1}^L \|u - U_l\|^2_{L^2(e_l)}.$$

The weak formulation of the model (2.1) reads [50]: find $(u, U) \in \mathbb{H}$ such that

$$a(\sigma, (u, U), (v, V)) = \langle I, V \rangle \quad \forall (v, V) \in \mathbb{H}, \tag{2.2}$$

where the trilinear form $a(\sigma, (u, U), (v, V))$ on $\mathcal{A} \times \mathbb{H} \times \mathbb{H}$ is defined by

$$a(\sigma, (u, U), (v, V)) = (\sigma \nabla u, \nabla v) + \sum_{l=1}^L z_l^{-1} (u - U_l, v - V_l)_{L^2(e_l)},$$

where $(\cdot, \cdot)_{L^2(e_l)}$ denotes the $L^2(e_l)$ inner product. For any $\sigma \in \mathcal{A}$, $\{z_l\}_{l=1}^L$ and $I \in \Sigma^L_\diamond$, the existence and uniqueness of a solution $(u, U) \in \mathbb{H}$ to (2.2) follows from Lemma 2.1 and Lax-Milgram theorem.

2.2 Regularized reconstruction

For numerical reconstruction with a piecewise constant conductivity, the total variation (TV) penalty is popular. The conductivity σ is assumed to be in the space $\mathrm{BV}(\Omega)$ of bounded variation [4, 19], i.e.,

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

equipped with the norm $\|v\|_{\mathrm{BV}(\Omega)} = \|v\|_{L^1(\Omega)} + |v|_{\mathrm{TV}(\Omega)}$, where

$$|v|_{\mathrm{TV}(\Omega)} := \sup\left\{ \int_\Omega v\nabla\cdot\boldsymbol{\phi}\,\mathrm{d}x :\ \boldsymbol{\phi}\in (C_c^1(\Omega))^d,\ \|\boldsymbol{\phi}(x)\|\le 1\right\}.$$

Below we discuss only one dataset, since the case of multiple datasets is similar. Then Tikhonov regularization leads to the following minimization problem:

$$\min_{\sigma\in\mathcal{A}}\left\{\mathcal{J}(\sigma) = \tfrac{1}{2}\|U(\sigma)-U^\delta\|^2 + \alpha|\sigma|_{\mathrm{TV}(\Omega)}\right\}, \tag{2.3}$$

The scalar $\alpha>0$ is known as a regularization parameter. It has at least one minimizer [46, 22].

Since σ is piecewise constant, by Lebesgue decomposition theorem [4], the TV term $|\sigma|_{\mathrm{TV}(\Omega)}$ in (2.3) reduces to $\int_{S_\sigma}|[\sigma]|\,\mathrm{d}\mathcal{H}^{d-1}$, where S_σ is the jump set, $[\sigma]=\sigma^+-\sigma^-$ denotes the jump across S_σ and $\mathcal{H}^{d-1}$ refers to the $(d-1)$-dimensional Hausdorff measure. The numerical approximation of (2.3) requires simultaneously treating two sets of different Hausdorff dimensions (i.e., Ω and S_σ), which is very challenging. Thus, we replace the TV term $|\sigma|_{\mathrm{TV}(\Omega)}$ in (2.3) by a Modica–Mortola type functional [43]

$$\mathcal{F}_\varepsilon(z) := \begin{cases} \varepsilon\|\nabla z\|^2_{L^2(\Omega)} + \frac{1}{\varepsilon}\int_\Omega W(z)\,\mathrm{d}x & \text{if } z\in H^1(\Omega),\\ +\infty & \text{otherwise,}\end{cases}$$

where ε is a small positive constant controlling the width of the transition interface, and $W:\mathbb{R}\to\mathbb{R}$ is the double-well potential given in (1.2). The functional $\mathcal{F}_\varepsilon$ was first proposed to model phase transition of two immiscible fluids in [11]. It is connected with the TV semi-norm as follows [43, 42, 2]; see [10] for an introduction to Γ-convergence.

Theorem 2.1. *With $c_W = \int_{c_0}^{c_1}\sqrt{W(s)}\,\mathrm{d}s$, let*

$$\mathcal{F}(z) := \begin{cases} 2c_W|z|_{\mathrm{TV}(\Omega)} & \text{if } z\in \mathrm{BV}(\Omega)\cap\mathcal{A},\\ +\infty & \text{otherwise.}\end{cases}$$

Then $\mathcal{F}_\varepsilon$ Γ-converges to $\mathcal{F}$ in $L^1(\Omega)$ as $\varepsilon\to 0^+$. Let $\{\varepsilon_n\}_{n\ge1}$ and $\{v_n\}_{n\ge1}$ be given sequences such that $\varepsilon_n\to 0^+$ and $\{\mathcal{F}_{\varepsilon_n}(v_n)\}_{n\ge1}$ is bounded. Then v_n is precompact in $L^1(\Omega)$.

The proposed EIT reconstruction method reads

$$\inf_{\sigma\in\widetilde{\mathcal{A}}}\left\{\mathcal{J}_\varepsilon(\sigma) = \tfrac{1}{2}\|U(\sigma)-U^\delta\|^2 + \tfrac{\widetilde{\alpha}}{2}\mathcal{F}_\varepsilon(\sigma)\right\}, \tag{2.4}$$

where $\widetilde{\alpha}=\alpha/c_W$, and the admissible set $\widetilde{\mathcal{A}}$ is defined as

$$\widetilde{\mathcal{A}} := \left\{\sigma\in H^1(\Omega) : c_0\le\sigma(x)\le c_1 \text{ a.e. } x\in\Omega\right\}.$$

Now we recall a useful continuity result of the forward map [22, Lemma 2.2], which gives the continuity of the fidelity term in the functional $\mathcal{J}_\varepsilon$. See also [31, 18] for related continuity results.

Lemma 2.2. *Let $\{\sigma_n\}_{n\ge1}\subset\widetilde{\mathcal{A}}$ satisfy $\sigma_n\to\sigma^*$ in $L^1(\Omega)$. Then*

$$\lim_{n\to\infty}\|\,(u(\sigma_n)-u(\sigma^*), U(\sigma_n)-U(\sigma^*))\,\|_{\mathbb{H}} = 0. \tag{2.5}$$

Lemma 2.2 implies that $\{\mathcal{J}_\varepsilon\}_{\varepsilon>0}$ are continuous perturbations of $\mathcal{J}$ in $L^1(\Omega)$. Then the stability of Γ-convergence [2, Proposition 1(ii)] [10, Remark 1.4] and Theorem 2.1 indicate that $\mathcal{J}_\varepsilon$ Γ-converges to $\mathcal{J}$ with respect to $L^1(\Omega)$, and $\mathcal{J}_\varepsilon$ can (approximately) recover piecewise constant conductivities. Next we show the existence of a minimizer to $\mathcal{J}_\varepsilon$.

Theorem 2.2. *For each $\varepsilon>0$, there exists at least one minimizer to problem* (2.4).

Proof. Since $\mathcal{J}_\varepsilon$ is nonnegative, there exists a minimizing sequence $\{\sigma_n\}_{n\geq 1} \subset \widetilde{\mathcal{A}}$ such that $\mathcal{J}_\varepsilon(\sigma_n) \to m_\varepsilon := \inf_{\sigma\in\widetilde{\mathcal{A}}} \mathcal{J}_\varepsilon(\sigma)$. Thus, $\sup_n \|\nabla\sigma_n\|_{L^2(\Omega)} < \infty$, which, along with $c_0 \leq \sigma_n \leq c_1$, yields $\|\sigma_n\|_{H^1(\Omega)} \leq c$. Since $\widetilde{\mathcal{A}}$ is closed and convex, there exist a subsequence, relabeled as $\{\sigma_n\}_{n\geq 1}$, and some $\sigma^* \in \widetilde{\mathcal{A}}$ such that

$$\sigma_n \rightharpoonup \sigma^* \quad \text{weakly in } H^1(\Omega), \quad \sigma_n \to \sigma^* \quad \text{in } L^1(\Omega), \quad \sigma_n \to \sigma^* \quad \text{a.e. in } \Omega. \tag{2.6}$$

Since $W(s) \in C^2[c_0, c_1]$, $\{W(\sigma_n)\}_{n\geq 1}$ is uniformly bounded in $L^\infty(\Omega)$ and converges to $W(\sigma^*)$ almost everywhere in Ω. By Lebesgue's dominated convergence theorem [19, p. 28, Theorem 1.19], $\int_\Omega W(\sigma_n)\,\mathrm{d}x \to \int_\Omega W(\sigma^*)\,\mathrm{d}x$. By Lemma 2.2 and the weak lower semi-continuity of the $H^1(\Omega)$-seminorm, we obtain

$$\mathcal{J}_\varepsilon(\sigma^*) \leq \liminf_{n\to\infty} \mathcal{J}_\varepsilon(\sigma_n) \leq \lim_{n\to\infty} \mathcal{J}_\varepsilon(\sigma_n) = m_\varepsilon.$$

Thus σ^* is a global minimizer of the functional $\mathcal{J}_\varepsilon$. □

To obtain the necessary optimality system of (2.4), we use the standard adjoint technique. The adjoint problem for (2.2) reads: find $(p, P) \in \mathbb{H}$ such that

$$a(\sigma, (p, P), (v, V)) = \langle U(\sigma) - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}. \tag{2.7}$$

By straightforward computation, the Gâteaux derivative $\mathcal{J}'_\varepsilon(\sigma)[\mu]$ of the functional $\mathcal{J}_\varepsilon$ at $\sigma \in \widetilde{\mathcal{A}}$ in the direction $\mu \in H^1(\Omega)$ is given by

$$\mathcal{J}'_\varepsilon(\sigma)[\mu] = \widetilde{\alpha}\left[\varepsilon(\nabla\sigma, \nabla\mu) + \tfrac{1}{2\varepsilon}(W'(\sigma), \mu)\right] - (\mu\nabla u(\sigma), \nabla p(\sigma)).$$

Then the minimizer σ^* to problem (2.4) and the respective state (u^*, U^*) and adjoint (p^*, P^*) satisfy the following necessary optimality system:

$$\begin{cases} a(\sigma^*, (u^*, U^*), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}, \\ a(\sigma^*, (p^*, P^*), (v, V)) = \langle U^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}, \\ \widetilde{\alpha}\varepsilon(\nabla\sigma^*, \nabla(\mu - \sigma^*)) + \frac{\widetilde{\alpha}}{2\varepsilon}(W'(\sigma^*), \mu - \sigma^*) - ((\mu - \sigma^*)\nabla u^*, \nabla p^*) \geq 0 \quad \forall \mu \in \widetilde{\mathcal{A}}, \end{cases} \tag{2.8}$$

where the variational inequality at the last line is due to the box constraint in the admissible set $\widetilde{\mathcal{A}}$. The optimality system (2.8) forms the basis of the adaptive algorithm and its convergence analysis.

3 Adaptive algorithm

Now we develop an adaptive FEM for problem (2.4). Let $\mathcal{T}_0$ be a shape regular triangulation of $\overline{\Omega}$ into simplicial elements, each intersecting at most one electrode surface e_l, and $\mathbb{T}$ be the set of all possible conforming triangulations of $\overline{\Omega}$ obtained from $\mathcal{T}_0$ by the successive use of bisection. Then $\mathbb{T}$ is uniformly shape regular, i.e., the shape-regularity of any mesh $\mathcal{T} \in \mathbb{T}$ is bounded by a constant depending only on $\mathcal{T}_0$ [44, 52]. Over any $\mathcal{T} \in \mathbb{T}$, we define a continuous piecewise linear finite element space

$$V_\mathcal{T} = \left\{v \in C(\overline{\Omega}) : v|_T \in P_1(T)\ \forall T \in \mathcal{T}\right\},$$

where $P_1(T)$ consists of all linear functions on T. The space $V_\mathcal{T}$ is used for approximating the state u and adjoint p, and the discrete admissible set $\widetilde{\mathcal{A}}_\mathcal{T}$ for the conductivity is given by

$$\widetilde{\mathcal{A}}_\mathcal{T} := V_\mathcal{T} \cap \widetilde{\mathcal{A}}.$$

Given $\sigma_\mathcal{T} \in \widetilde{\mathcal{A}}_\mathcal{T}$, the discrete analogue of problem (2.2) is to find $(u_\mathcal{T}, U_\mathcal{T}) \in \mathbb{H}_\mathcal{T} \equiv V_\mathcal{T} \otimes \mathbb{R}^L_\diamond$ such that

$$a(\sigma_\mathcal{T}, (u_\mathcal{T}, U_\mathcal{T}), (v_\mathcal{T}, V)) = \langle I, V\rangle \quad \forall (v_\mathcal{T}, V) \in \mathbb{H}_\mathcal{T}. \tag{3.1}$$

Then we approximate problem (2.4) by minimizing the following functional over $\widetilde{\mathcal{A}}_\mathcal{T}$:

$$J_{\varepsilon,\mathcal{T}}(\sigma_\mathcal{T}) = \tfrac{1}{2}\|U_\mathcal{T}(\sigma_\mathcal{T}) - U^\delta\|^2 + \tfrac{\widetilde{\alpha}}{2}\mathcal{F}_\varepsilon(\sigma_\mathcal{T}). \tag{3.2}$$

Then, similar to Theorem 2.2, there exists at least one minimizer $\sigma_{\mathcal{T}}^*$ to (3.2), and the minimizer $\sigma_{\mathcal{T}}^*$ and the related state $(u_{\mathcal{T}}^*, U_{\mathcal{T}}^*) \in \mathbb{H}_{\mathcal{T}}$ and adjoint $(p_{\mathcal{T}}^*, P_{\mathcal{T}}^*) \in \mathbb{H}_{\mathcal{T}}$ satisfy

$$\begin{cases} a(\sigma_{\mathcal{T}}^*, (u_{\mathcal{T}}^*, U_{\mathcal{T}}^*), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}_{\mathcal{T}}, \\ a(\sigma_{\mathcal{T}}^*, (p_{\mathcal{T}}^*, P_{\mathcal{T}}^*), (v, V)) = \langle U_{\mathcal{T}}^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}_{\mathcal{T}}, \\ \widetilde{\alpha}\varepsilon(\nabla\sigma_{\mathcal{T}}^*, \nabla(\mu - \sigma_{\mathcal{T}}^*)) + \dfrac{\widetilde{\alpha}}{2\varepsilon}(W'(\sigma_{\mathcal{T}}^*), \mu - \sigma_{\mathcal{T}}^*) - ((\mu - \sigma_{\mathcal{T}}^*)\nabla u_{\mathcal{T}}^*, \nabla p_{\mathcal{T}}^*) \geq 0 \quad \forall \mu \in \widetilde{\mathcal{A}}_{\mathcal{T}}. \end{cases} \tag{3.3}$$

Further, $(u_{\mathcal{T}}^*, U_{\mathcal{T}}^*)$ and $(p_{\mathcal{T}}^*, P_{\mathcal{T}}^*)$ depend continuously on the problem data, i.e.,

$$\|(u_{\mathcal{T}}^*, U_{\mathcal{T}}^*)\|_{\mathbb{H}} + \|(p_{\mathcal{T}}^*, P_{\mathcal{T}}^*)\|_{\mathbb{H}} \leq c(\|I\| + \|U^\delta\|), \tag{3.4}$$

where the constant c can be made independent of α and ε.

To describe the error estimators, we first recall some useful notation. The collection of all faces (respectively all interior faces) in $\mathcal{T} \in \mathbb{T}$ is denoted by $\mathcal{F}_{\mathcal{T}}$ (respectively $\mathcal{F}_{\mathcal{T}}^i$) and its restriction to the electrode $\bar{e}_l$ and $\partial\Omega \backslash \cup_{l=1}^L e_l$ by $\mathcal{F}_{\mathcal{T}}^l$ and $\mathcal{F}_{\mathcal{T}}^c$, respectively. A face/edge F has a fixed normal unit vector $\boldsymbol{n}_F$ in $\overline{\Omega}$ with $\boldsymbol{n}_F = \boldsymbol{n}$ on $\partial\Omega$. The diameter of any $T \in \mathcal{T}$ and $F \in \mathcal{F}_{\mathcal{T}}$ is denoted by $h_T := |T|^{1/d}$ and $h_F := |F|^{1/(d-1)}$, respectively. For the solution $(\sigma_{\mathcal{T}}^*, (u_{\mathcal{T}}^*, U_{\mathcal{T}}^*), (p_{\mathcal{T}}^*, P_{\mathcal{T}}^*))$ to problem (3.3), we define two element residuals for each element $T \in \mathcal{T}$ and two face residuals for each face $F \in \mathcal{F}_{\mathcal{T}}$ by

$$\begin{aligned} R_{T,1}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*) &= \nabla \cdot (\sigma_{\mathcal{T}}^* \nabla u_{\mathcal{T}}^*), \\ R_{T,2}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, p_{\mathcal{T}}^*) &= \tfrac{\widetilde{\alpha}}{2\varepsilon} W'(\sigma_{\mathcal{T}}^*) - \nabla u_{\mathcal{T}}^* \cdot \nabla p_{\mathcal{T}}^*, \\ J_{F,1}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, U_{\mathcal{T}}^*) &= \begin{cases} [\sigma_{\mathcal{T}}^* \nabla u_{\mathcal{T}}^* \cdot \boldsymbol{n}_F] & \text{for } F \in \mathcal{F}_{\mathcal{T}}^i, \\ \sigma_{\mathcal{T}}^* \nabla u_{\mathcal{T}}^* \cdot \boldsymbol{n} + (u_{\mathcal{T}}^* - U_{\mathcal{T},l}^*)/z_l & \text{for } F \in \mathcal{F}_{\mathcal{T}}^l, \\ \sigma_{\mathcal{T}}^* \nabla u_{\mathcal{T}}^* \cdot \boldsymbol{n} & \text{for } F \in \mathcal{F}_{\mathcal{T}}^c, \end{cases} \\ J_{F,2}(\sigma_{\mathcal{T}}^*) &= \begin{cases} \widetilde{\alpha}\varepsilon[\nabla\sigma_{\mathcal{T}}^* \cdot \boldsymbol{n}_F] & \text{for } F \in \mathcal{F}_{\mathcal{T}}^i, \\ \widetilde{\alpha}\varepsilon \nabla\sigma_{\mathcal{T}}^* \cdot \boldsymbol{n} & \text{for } F \in \mathcal{F}_{\mathcal{T}}^l \cup \mathcal{F}_{\mathcal{T}}^c, \end{cases} \end{aligned}$$

where $[\cdot]$ denotes the jump across interior face F. Then for any element $T \in \mathcal{T}$, we define the following three error estimators

$$\begin{aligned} \eta_{\mathcal{T},1}^2(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, U_{\mathcal{T}}^*, T) &:= h_T^2 \|R_{T,1}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*)\|_{L^2(T)}^2 + \sum_{F \subset \partial T} h_F \|J_{F,1}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, U_{\mathcal{T}}^*)\|_{L^2(F)}^2, \\ \eta_{\mathcal{T},2}^2(\sigma_{\mathcal{T}}^*, p_{\mathcal{T}}^*, P_{\mathcal{T}}^*, T) &:= h_T^2 \|R_{T,1}(\sigma_{\mathcal{T}}^*, p_{\mathcal{T}}^*)\|_{L^2(T)}^2 + \sum_{F \subset \partial T} h_F \|J_{F,1}(\sigma_{\mathcal{T}}^*, p_{\mathcal{T}}^*, P_{\mathcal{T}}^*)\|_{L^2(F)}^2, \\ \eta_{\mathcal{T},3}^q(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, p_{\mathcal{T}}^*, T) &:= h_T^q \|R_{T,2}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, p_{\mathcal{T}}^*)\|_{L^q(T)}^q + \sum_{F \subset \partial T} h_F \|J_{F,2}(\sigma_{\mathcal{T}}^*)\|_{L^q(F)}^q \end{aligned}$$

with $q = d/(d-1)$. The estimator $\eta_{\mathcal{T},1}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, U_{\mathcal{T}}^*, T)$ is identical with the standard residual error indicator for the direct problem: find $(\tilde{u}, \tilde{U}) \in \mathbb{H}$ such that

$$a(\sigma_{\mathcal{T}}^*, (\tilde{u}, \tilde{U}), (v, V)) = \langle I, V\rangle, \quad \forall (v, V) \in \mathbb{H}.$$

It differs from the direct problem in (2.8) by replacing the conductivity σ^* with $\sigma_{\mathcal{T}}^*$ instead, and is a perturbation of the latter case. The perturbation is vanishingly small in the event of the conjectured (subsequential) convergence $\sigma_{\mathcal{T}}^* \to \sigma^*$. The estimator $\eta_{\mathcal{T},2}(\sigma_{\mathcal{T}}^*, p_{\mathcal{T}}^*, P_{\mathcal{T}}^*, T)$ admits a similar interpretation. These two estimators are essentially identical with that for the $H^1(\Omega)$ penalty in [32], and we refer to [32, Section 3.3] for a detailed heuristic derivation. The estimator $\eta_{\mathcal{T},3}(\sigma_{\mathcal{T}}^*, u_{\mathcal{T}}^*, p_{\mathcal{T}}^*, T)$ is related to the variational inequality in the necessary optimality condition (2.8), and roughly provides a quantitative measure how well it is satisfied. The estimator (including the exponent q) is motivated by the convergence analysis; see the proof of Theorem 5.5 and Remark 5.2 below. It represents the main new ingredient for problem (2.4), and differs from that for the $H^1(\Omega)$ penalty in [32].

Remark 3.1. *The estimator $\eta_{k,3}$ improves that in [32], i.e.,*

$$\eta^2_{\mathcal{T},3}(\sigma^*_{\mathcal{T}}, u^*_{\mathcal{T}}, p^*_{\mathcal{T}}, T) := h_T^4 \|R_{T,2}(\sigma^*_{\mathcal{T}}, u^*_{\mathcal{T}}, p^*_{\mathcal{T}})\|^2_{L^2(T)} + \sum_{F\subset\partial T} h_F^2 \|J_{F,2}(\sigma^*_{\mathcal{T}})\|^2_{L^2(F)},$$

in terms of the exponents on h_T and h_F. This improvement is achieved by a novel constraint preserving interpolation operator defined in (5.13) *below.*

Now we can formulate an adaptive algorithm for (2.4); see Algorithm 3.1. Below we indicate the dependence on the mesh $\mathcal{T}_k$ by the subscript k, e.g., $\mathcal{J}_{\varepsilon,k}$ for $\mathcal{J}_{\varepsilon,\mathcal{T}_k}$.

Algorithm 3.1 AFEM for EIT with a piecewise constant conductivity.

1: Specify an initial mesh $\mathcal{T}_0$, and set the maximum number K of refinements.
2: **for** $k = 0 : K-1$ **do**
3: (SOLVE) Solve problem (3.1)-(3.2) over $\mathcal{T}_k$ for $(\sigma^*_k, (u^*_k, U^*_k)) \in \widetilde{\mathcal{A}}_k \times \mathbb{H}_k$ and (3.3) for $(p^*_k, P^*_k) \in \mathbb{H}_k$.
4: (ESTIMATE) Compute error indicators $\eta^2_{k,1}(\sigma^*_k, u^*_k, U^*_k)$, $\eta^2_{k,2}(\sigma^*_k, p^*_k, P^*_k)$ and $\eta^q_{k,3}(\sigma^*_k, u^*_k, p^*_k)$.
5: (MARK) Mark three subsets $\mathcal{M}^i_k \subseteq \mathcal{T}_k$ $(i = 1, 2, 3)$ such that each $\mathcal{M}^i_k$ contains at least one element $\widetilde{T}^i_k \in \mathcal{T}_k$ $(i = 1, 2, 3)$ with the largest error indicator:

$$\eta_{k,i}(\widetilde{T}^i_k) = \max_{T\in\mathcal{T}_k} \eta_{k,i}. \tag{3.5}$$

Then $\mathcal{M}_k := \mathcal{M}^1_k \cup \mathcal{M}^2_k \cup \mathcal{M}^3_k$.
6: (REFINE) Refine each element T in $\mathcal{M}_k$ by bisection to get $\mathcal{T}_{k+1}$.
7: Check the stopping criterion.
8: **end for**
9: Output $(\sigma^*_k, (u^*_k, U^*_k), (p^*_k, P^*_k))$.

The MARK module selects a collection of elements in the mesh $\mathcal{T}_k$. The condition (3.5) covers several commonly used marking strategies, e.g., maximum, equidistribution, modified equidistribution, and Dörfler's strategy [49, pp. 962]. Compared with a collective marking in AFEM in [32], Algorithm 3.1 employs a separate marking to select more elements for refinement in each loop, which leads to fewer iterations of the adaptive process. The error estimators may also be used for coarsening, which is relevant if the recovered inclusions change dramatically during the iteration. However, the convergence analysis below does not carry over to coarsening, and it will not be further explored.

Last, we give the main theoretical result: for each fixed $\varepsilon > 0$, the sequence of discrete solutions $\{\sigma^*_k, (u^*_k, U^*_k), (p^*_k, P^*_k)\}_{k\geq 0}$ generated by Algorithm 3.1 contains a subsequence converging in $H^1(\Omega)\times\mathbb{H}\times\mathbb{H}$ to a solution of system (2.8). The proof is lengthy and technical, and thus deferred to Section 5.

Theorem 3.1. *The sequence of discrete solutions $\{\sigma^*_k, (u^*_k, U^*_k), (p^*_k, P^*_k)\}_{k\geq 0}$ by Algorithm 3.1 contains a subsequence $\{\sigma^*_{k_j}, (u^*_{k_j}, U^*_{k_j}), (p^*_{k_j}, P^*_{k_j})\}_{j\geq 0}$ convergent to a solution $(\sigma^*, (u^*, U^*), (p^*, P^*))$ of system* (2.8)*:*

$$\|\sigma^*_{k_j} - \sigma^*\|_{H^1(\Omega)},\ \|(u^*_{k_j} - u^*, U^*_{k_j} - U^*)\|_{\mathbb{H}},\ \|(p^*_{k_j} - p^*, P^*_{k_j} - P^*)\|_{\mathbb{H}} \to 0 \quad \text{as } j \to \infty.$$

4 Numerical experiments and discussions

Now we present numerical results to illustrate Algorithm 3.1 on a square domain $\Omega = (-1, 1)^2$. There are sixteen electrodes $\{e_l\}_{l=1}^L$ (with $L = 16$) evenly distributed along $\partial\Omega$, each of length $1/4$. The contact impedances $\{z_l\}_{l=1}^L$ are all set to unit. We take ten sinusoidal input currents, and for each voltage $U(\sigma^\dagger) \in \mathbb{R}^L_\diamond$, generate the noisy data U^δ by

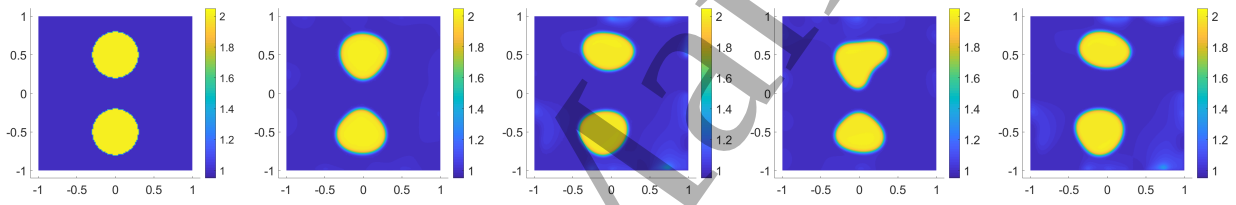
$$U^\delta_l = U_l(\sigma^\dagger) + \epsilon \max_l |U_l(\sigma^\dagger)|\xi_l, \quad l = 1, \ldots, L, \tag{4.1}$$

where ϵ is the (relative) noise level, and $\{\xi_l\}_{l=1}^L$ follow the standard normal distribution. Note that ϵ = 1e-2 refers to a relatively high noise level for EIT. The exact data $U(\sigma^\dagger)$ is computed using a much finer uniform mesh, to avoid the most obvious form of "inverse crime".

In the experiments, we fix K (the number of refinements) at 15, q (exponent in $\eta_{k,3}^q$) at 2, and ε (the functional $\mathcal{F}_\varepsilon$) at 1e-2. The marking strategy (3.5) in the module `MARK` selects a minimal refinement set $\mathcal{M}_k := \cup_{i=1}^3 \mathcal{M}_k^i \subseteq \mathcal{T}_k$ such that

$$\eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, \mathcal{M}_k^1) \geq \theta \eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*), \quad \eta_{k,2}^2(\sigma_k^*, p_k^*, P_k^*, \mathcal{M}_k^2) \geq \theta \eta_{k,2}^2(\sigma_k^*, p_k^*, P_k^*),$$
$$\eta_{k,3}^2(\sigma_k^*, u_k^*, p_k^*, \mathcal{M}_k^3) \geq \theta \eta_{k,3}^2(\sigma_k^*, u_k^*, p_k^*),$$

with a threshold $\theta = 0.7$. The refinement is performed with one popular refinement strategy, i.e., newest vertex bisection [41]. Specifically, it connects the midpoint x_T, as a newest vertex, of a reference edge F of an element $T \in \mathcal{T}_k$ to the opposite node of F, and employs two edges opposite to the midpoint x_T as reference edges of the two newly created triangles in $\mathcal{T}_{k+1}$. Problem (3.1)-(3.2) is solved by a Newton type method; see Appendix A for the detail. The conductivity on $\mathcal{T}_0$ is initialized to $\sigma_0 = c_0$, and then for $k = 1, 2, \ldots$, σ_{k-1}^* (defined on $\mathcal{T}_{k-1}$) is interpolated to $\mathcal{T}_k$ to warm start the optimization. The regularization parameter $\widetilde{\alpha}$ in (2.4) is determined in a trial-and-error manner. All computations are performed using `MATLAB` 2018a on a personal laptop with 8.00 GB RAM and 2.5 GHz CPU.



(a) true conductivity (b) adaptive (c) adaptive (d) uniform (e) uniform

Figure 1: The final recoveries by the adaptive and uniform refinements for Example 1(i). The results in (b) and (d) are for ϵ = 1e-3 and $\tilde{\alpha}$ = 2e-2, and (c) and (e) for ϵ = 1e-2 and $\tilde{\alpha}$ = 3e-2. d.o.f. in (b), (c), (d) and (e) are 15830, 18770, 16641 and 16641, respectively.

The first set of examples are concerned with two inclusions.

Example 1. *The background conductivity* $\sigma_0(x) = 1$.

(i) *The true conductivity* $\sigma^\dagger$ *is given by* $\sigma_0(x) + \chi_{B_1}(x) + \chi_{B_2}(x)$, *with* B_1 *and* B_2 *denote two circles centered at* $(0, 0.5)$ *and* $(0, -0.5)$, *respectively, both with a radius 0.3.*

(ii) *The true conductivity* $\sigma^\dagger$ *is given by* $\sigma_0(x) + 1 + 1.2e^{-\frac{25(x_1^2+(x_2-0.5)^2)}{2}} + 1.2e^{-\frac{25(x_1^2+(x_2+0.5)^2)}{2}}$, *i.e., two Gaussian bumps centered at* $(0, 0.5)$ *and* $(0, -0.5)$.

(iii) *The true conductivity* $\sigma^\dagger$ *is given by* $\sigma_0(x) + 5\chi_{B_1}(x) + 5\chi_{B_2}(x)$, *with* B_1 *and* B_2 *denote two circles centered at* $(0, 0.5)$ *and* $(0, -0.5)$, *respectively, both with a radius 0.3.*

The numerical results for Example 1(i) with ϵ = 1e-3 and ϵ = 1e-2 are shown in Figs. 1–5, where d.o.f. denotes the degree of freedom of the mesh. It is observed from Fig. 1 that with both uniform and adaptive refinements, the final recoveries have comparable accuracy and capture well the inclusion locations.

Next we examine the adaptive refinement process more closely. In Figs. 2 and 3, we show the meshes $\mathcal{T}_k$ during the iteration and the corresponding recoveries σ_k for Example 1(i) at two noise levels ϵ = 1e-3 and ϵ = 1e-2, respectively. On the coarse mesh $\mathcal{T}_0$, the recovery has very large errors and can only identify one component and thus fails to correctly identify the number of inclusions, due to the severe under-resolution of both state and conductivity. Nonetheless, Algorithm 3.1 can correctly recover the two components with reasonable accuracy after several adaptive loops, and accordingly, the support of the recovery is gradually refined with its accuracy improving steadily. In particular, the inclusion locations stabilize after several

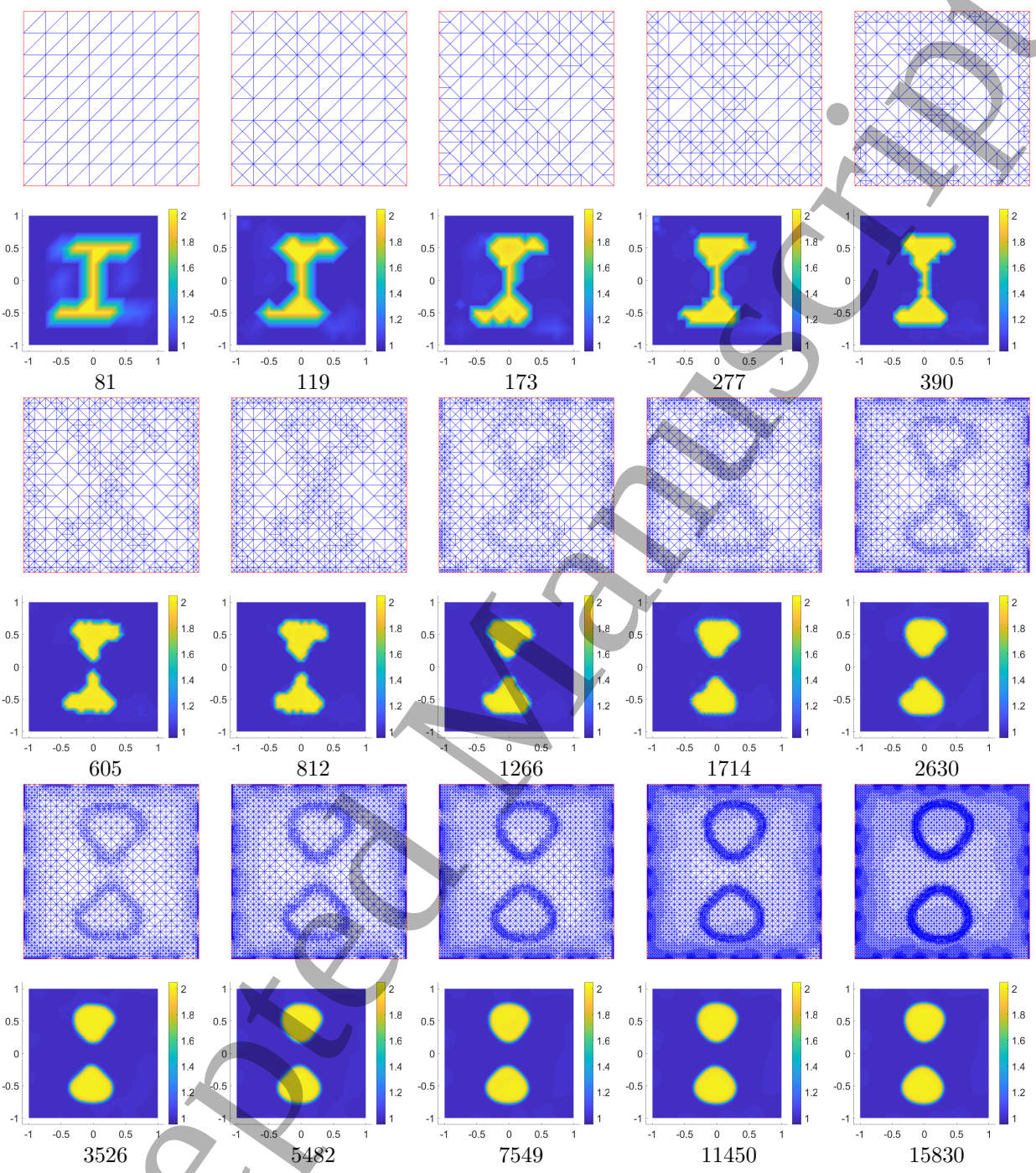


Figure 2: The meshes $\mathcal{T}_k$ and recoveries σ_k during the adaptive refinement, for Example 1(i) with $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 2e-2. The numbers refer to d.o.f.

loops, and thus coarsening of the mesh seems unnecessary. Throughout, the refinement occurs mainly in the regions around the electrode edges and internal interface, which is clearly observed for both noise levels. This is attributed to the separable marking strategy, which allows detecting different sources of singularities simultaneously. In Fig. 4, we display the evolution of the error indicators for Example 1(i) with $\epsilon =$ 1e-3. The estimators play different roles: $\eta^2_{k,1}$ and $\eta^2_{k,2}$ indicate the electrode edges during first iterations and then

also internal interface, whereas throughout $\eta_{k,3}^2$ concentrates on the internal interface. Thus, $\eta_{k,1}^2$ and $\eta_{k,2}^2$ are most effective for resolving the state and adjoint, whereas $\eta_{k,3}^2$ is effective for detecting internal jumps of the conductivity. The magnitude of $\eta_{k,2}^2$ is much smaller than $\eta_{k,1}^2$, since the boundary data $U^\delta - U(\sigma_k)$ for the adjoint is much smaller than the input current I for the state. Thus, a simple collective marking strategy (i.e., $\eta_k^2 = \eta_{k,1}^2 + \eta_{k,2}^2 + \eta_{k,3}^2$) may miss the correct singularity, due to their drastically different scalings. In contrast, the separate marking in (3.5) can take care of the scaling automatically.

In Fig. 5, we plot the $L^2(\Omega)$ and $L^1(\Omega)$ errors of the recoveries versus d.o.f. N, where the recovery on the corresponding finest mesh is taken as the reference (since the recoveries by the adaptive and uniform refinements are slightly different; see Fig. 1). Due to the discontinuity of the sought-for conductivity, the $L^1(\Omega)$ norm is especially suitable for measuring the convergence. The convergence of the algorithm is clearly observed for both adaptive and uniform refinements. Further, with a fixed d.o.f., AFEM gives more accurate results than the uniform one in both error metrics. These observations show the computational efficiency of the adaptive algorithm.

Examples 1(ii) and (iii) are variations of Example 1(i), and the results are presented in Figs. 6–9. The proposed approach assumes a piecewise constant conductivity with known lower and upper bounds. Example 1(ii) does not fulfill the assumption, since the true conductivity $\sigma^\dagger$ is not piecewise constant. Thus the algorithm can only produce a piecewise constant approximation to the exact one. Nonetheless, the inclusion support is reasonably identified. When the noise level ϵ increases from 1e-3 to 1e-2, the reconstruction accuracy deteriorates significantly; see Fig. 6. Example 1(iii) involves high contrast inclusions, which are well known to be numerically more challenging. This is clearly observed in Fig. 8, where the recovery accuracy is inferior, especially for the noise level ϵ = 1e-2. However, the adaptive refinement procedure works well similarly as the preceding examples: the refinement occurs mainly around the electrode edges and inclusion interface; see Figs. 7 and 9 for the details.

Now we consider one more challenging example with four inclusions.

Example 2. *The true conductivity $\sigma^\dagger$ is given by $\sigma_0(x) + \sum_{i=1}^4 \chi_{B_i}(x)$, with the circles B_i centered at $(0.6, \pm 0.6)$ and $(-0.6, \pm 0.6)$, respectively, all with a radius 0.2, and the background conductivity $\sigma_0(x) = 1$.*

The numerical results for Example 2 are given in Figs. 10–12. The results are in excellent agreement with the observations from Example 1: The algorithm converges steadily as the adaptive iteration proceeds, and with a low noise level, it can accurately recover all four inclusions, showing clearly the efficiency of the adaptive approach. The refinement is mainly around the electrode edge and interval interface.

5 Proof of Theorem 3.1

The lengthy and technical proof is divided into two steps: Step 1 shows the convergence to an auxiliary minimization problem over a limiting admissible set in Section 5.1, and Step 2 shows that the solution of the auxiliary problem satisfies the necessary optimality system (2.8) in section 5.2. The overall proof strategy is similar to [32], and hence we omit relevant arguments.

5.1 Auxiliary convergence

Since the two sequences $\{\mathbb{H}_k\}_{k\geq 0}$ and $\{\widetilde{\mathcal{A}}_k\}_{k\geq 0}$ generated by Algorithm 3.1 are nested, we may define

$$\mathbb{H}_\infty := \overline{\bigcup_{k\geq 0} \mathbb{H}_k} \text{ (in } \mathbb{H}\text{-norm)} \quad \text{and} \quad \widetilde{\mathcal{A}}_\infty := \overline{\bigcup_{k\geq 0} \widetilde{\mathcal{A}}_k} \text{ (in } H^1(\Omega)\text{-norm).}$$

Clearly $\mathbb{H}_\infty$ is a closed subspace of $\mathbb{H}$. For the set $\widetilde{\mathcal{A}}_\infty$, we have the following result [32, Lemma 4.1].

Lemma 5.1. *$\widetilde{\mathcal{A}}_\infty$ is a closed convex subset of $\widetilde{\mathcal{A}}$.*

Over the limiting set $\widetilde{\mathcal{A}}_\infty$, we define an auxiliary limiting minimization problem:

$$\min_{\sigma_\infty \in \widetilde{\mathcal{A}}_\infty} \left\{ \mathcal{J}_{\varepsilon,\infty}(\sigma_\infty) = \tfrac{1}{2}\|U_\infty(\sigma_\infty) - U^\delta\|^2 + \tfrac{\widetilde{\alpha}}{2}\mathcal{F}_\varepsilon(\sigma_\infty) \right\}, \tag{5.1}$$

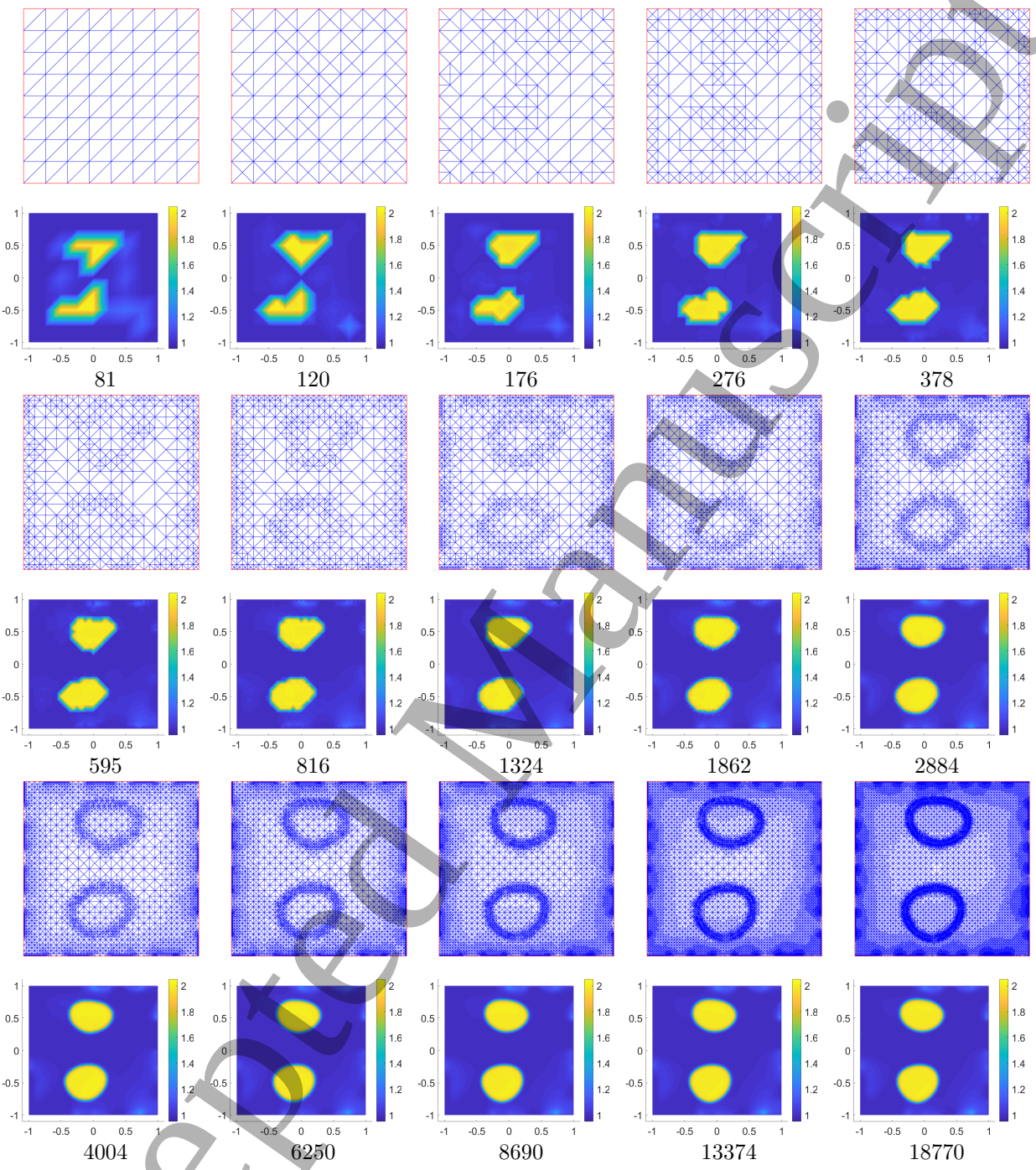


Figure 3: The meshes $\mathcal{T}_k$ and recoveries σ_k during the adaptive refinement, for Example 1(i) with $\epsilon =$ 1e-2 and $\tilde{\alpha} =$ 3e-2. The numbers refer to d.o.f.

where $(u_\infty, U_\infty) \in \mathbb{H}_\infty$ satisfies

$$a(\sigma_\infty, (u_\infty, U_\infty), (v, V)) = \langle I, V \rangle \quad \forall (v, V) \in \mathbb{H}_\infty. \tag{5.2}$$

By Lemma 2.1 and Lax-Milgram theorem, problem (5.2) is well-posed for any fixed $\sigma_\infty \in \mathcal{A}_\infty$. The next result gives the existence of a minimizer to (5.1)–(5.2).

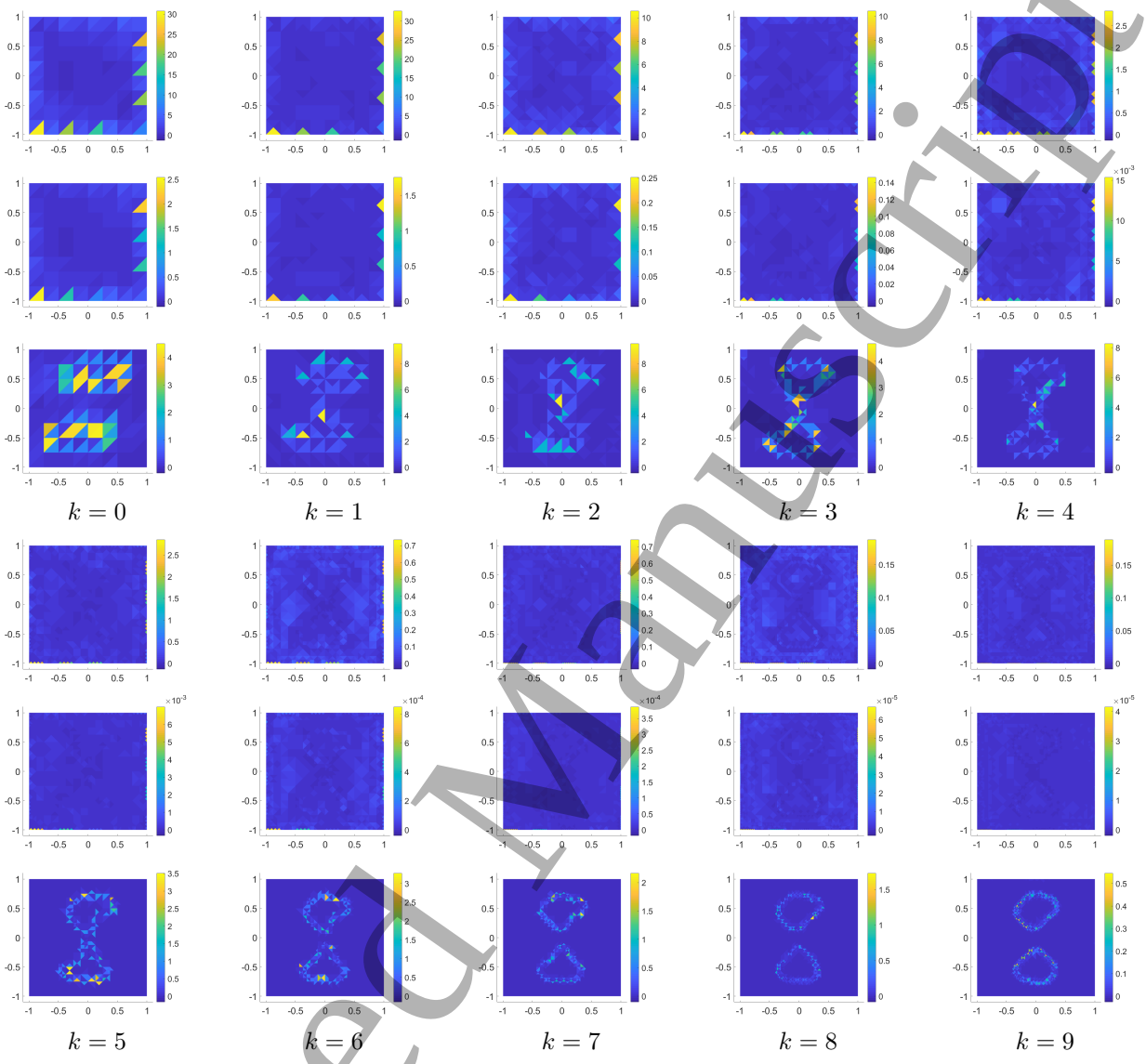


Figure 4: The evolution of the three error indicators $\eta_{k,i}^2$ for $k = 0, 1, \ldots, 9$, $i = 1$ (top), $i = 2$ (middle) and $i = 3$ (bottom), for Example 1(i) with $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 2e-2.

Theorem 5.1. *There exists at least one minimizer to problem* (5.1)–(5.2).

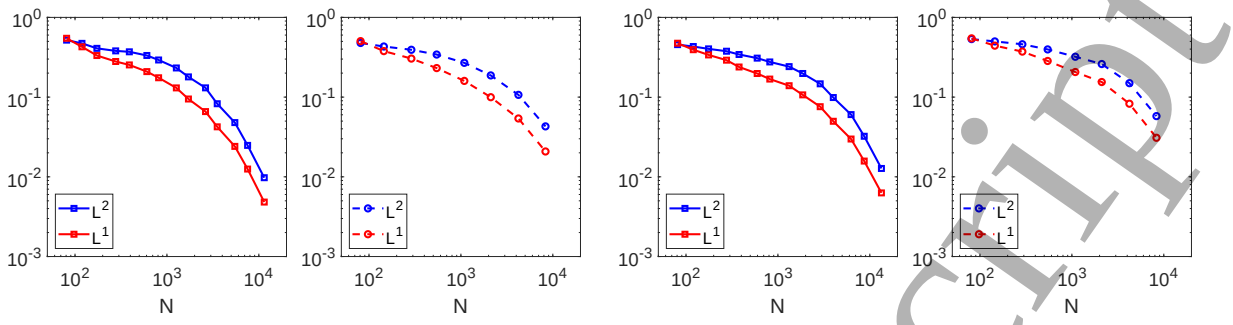
Proof. Let $\{\sigma_k^*, (u_k^*, U_k^*)\}_{k\geq 0}$ be the sequence of discrete solutions given by Algorithm 3.1. Since $c_1 \in \widetilde{\mathcal{A}}_k$ for all k, by (3.4), $\mathcal{J}_{\varepsilon,k}(\sigma_k^*) \leq \mathcal{J}_{\varepsilon,k}(c_1) \leq c$, and thus $\{\sigma_k^*\}_{k\geq 0}$ is uniformly bounded in $H^1(\Omega)$. By Lemma 5.1 and Sobolev embedding, there exist a subsequence, denoted by $\{\sigma_{k_j}^*\}_{j\geq 0}$, and some $\sigma^* \in \widetilde{\mathcal{A}}_\infty$ such that

$$\sigma_{k_j}^* \rightharpoonup \sigma^* \quad \text{weakly in } H^1(\Omega), \quad \sigma_{k_j}^* \to \sigma^* \quad \text{in } L^2(\Omega), \quad \sigma_{k_j}^* \to \sigma^* \quad \text{a.e. in } \Omega, \tag{5.3}$$

Next we introduce a discrete analogue of problem (5.2) with $\sigma_\infty = \sigma^*$: find $(u_{k_j}, U_{k_j}) \in \mathbb{H}_{k_j}$ such that

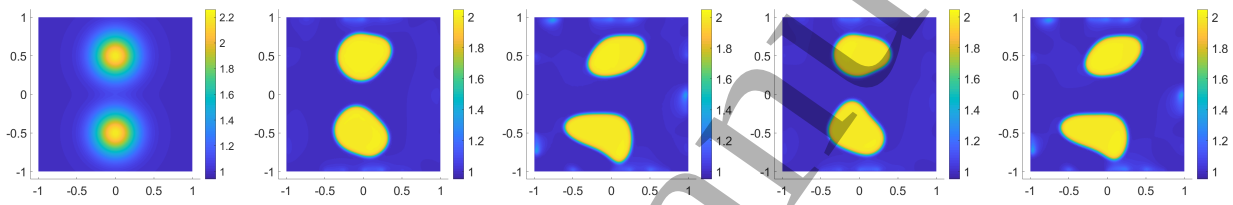
$$a(\sigma^*, (u_{k_j}, U_{k_j}), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}_{k_j}. \tag{5.4}$$

By Lemma 2.1, Cea's lemma and the construction of the space $\mathbb{H}_\infty$, the solution $(u_\infty^*, U_\infty^*) \in \mathbb{H}_\infty$ of (5.2)



(a) $\epsilon = $ 1e-3, $\tilde{\alpha} = $ 2e-2 (b) $\epsilon = $ 1e-2, $\tilde{\alpha} = $ 3e-2

Figure 5: The $L^2(\Omega)$ and $L^1(\Omega)$ errors versus d.o.f. N of the mesh, for Example 1(i), using the adaptive (solid) and uniform (dashed) refinement.



(a) true conductivity (b) adaptive (c) adaptive (d) uniform (e) uniform

Figure 6: The final recoveries by the adaptive and uniform refinements for Example 1(ii). The results in (b) and (d) are for $\epsilon = $ 1e-3 and $\tilde{\alpha} = $ 2e-2, and (c) and (e) for $\epsilon = $ 1e-2 and $\tilde{\alpha} = $ 5e-2. The d.o.f. of (b), (c), (d) and (e) are 17736, 20524, 16641 and 16641.

with $\sigma_\infty = \sigma^*$ satisfies

$$\|(u^*_\infty - u^*_{k_j}, U^*_\infty - U^*_{k_j})\|_{\mathbb{H}} \le c \inf_{(v,V)\in\mathbb{H}_{k_j}} \|(u^*_\infty - v, U^*_\infty - V)\|_{\mathbb{H}} \to 0. \tag{5.5}$$

Taking the test function $(v, V) = (u_{k_j} - u^*_{k_j}, U_{k_j} - U^*_{k_j}) \in \mathbb{H}_{k_j}$ in the first line of (3.3) and (5.4) and then applying the Cauchy-Schwarz inequality lead to

$$\begin{aligned}
&a(\sigma^*_{k_j}, (u_{k_j} - u^*_{k_j}, U_{k_j} - U^*_{k_j}), (u_{k_j} - u^*_{k_j}, U_{k_j} - U^*_{k_j}))\\
&= ((\sigma^*_{k_j} - \sigma^*)\nabla(u_{k_j} - u^*_\infty), \nabla(u_{k_j} - u^*_{k_j})) + ((\sigma^*_{k_j} - \sigma^*)\nabla u^*_\infty, \nabla(u_{k_j} - u^*_{k_j}))\\
&\le (\|(\sigma^*_{k_j} - \sigma^*)\nabla(u_{k_j} - u^*_\infty)\|_{L^2(\Omega)} + \|(\sigma^*_{k_j} - \sigma^*)\nabla u^*_\infty\|_{L^2(\Omega)})\|\nabla(u_{k_j} - u^*_{k_j})\|_{L^2(\Omega)}.
\end{aligned}$$

In view of (5.5), pointwise convergence in (5.3) and Lebesgue's dominated convergence theorem,

$$\|(\sigma^*_{k_j} - \sigma^*)\nabla(u_{k_j} - u^*_\infty)\|_{L^2(\Omega)} \le c_1\|\nabla(u_{k_j} - u^*_\infty)\|_{L^2(\Omega)} \to 0, \quad \|(\sigma^*_{k_j} - \sigma^*)\nabla u^*_\infty\|_{L^2(\Omega)} \to 0,$$

This and Lemma 2.1 imply $\|(u_{k_j} - u^*_{k_j}, U_{k_j} - U^*_{k_j})\|_{\mathbb{H}} \to 0$. Then, (5.5) and the triangle inequality imply

$$\|(u^*_{k_j} - u^*_\infty, U^*_{k_j} - U^*_\infty)\|_{\mathbb{H}} \to 0. \tag{5.6}$$

Meanwhile, repeating the argument of Theorem 2.2 gives

$$\int_\Omega W(\sigma^*_{k_j})\,\mathrm{d}x \to \int_\Omega W(\sigma^*)\,\mathrm{d}x. \tag{5.7}$$

next we apply a density argument. For any $\sigma_\infty \in \widetilde{\mathcal{A}}_\infty$, by the construction of the space $\mathbb{H}_\infty$, there exists a sequence $\{\sigma_k\}_{k\ge0} \subset \bigcup_{k\ge0}\widetilde{\mathcal{A}}_k$ such that $\sigma_k \to \sigma_\infty$ in $H^1(\Omega)$. Repeating the preceding argument

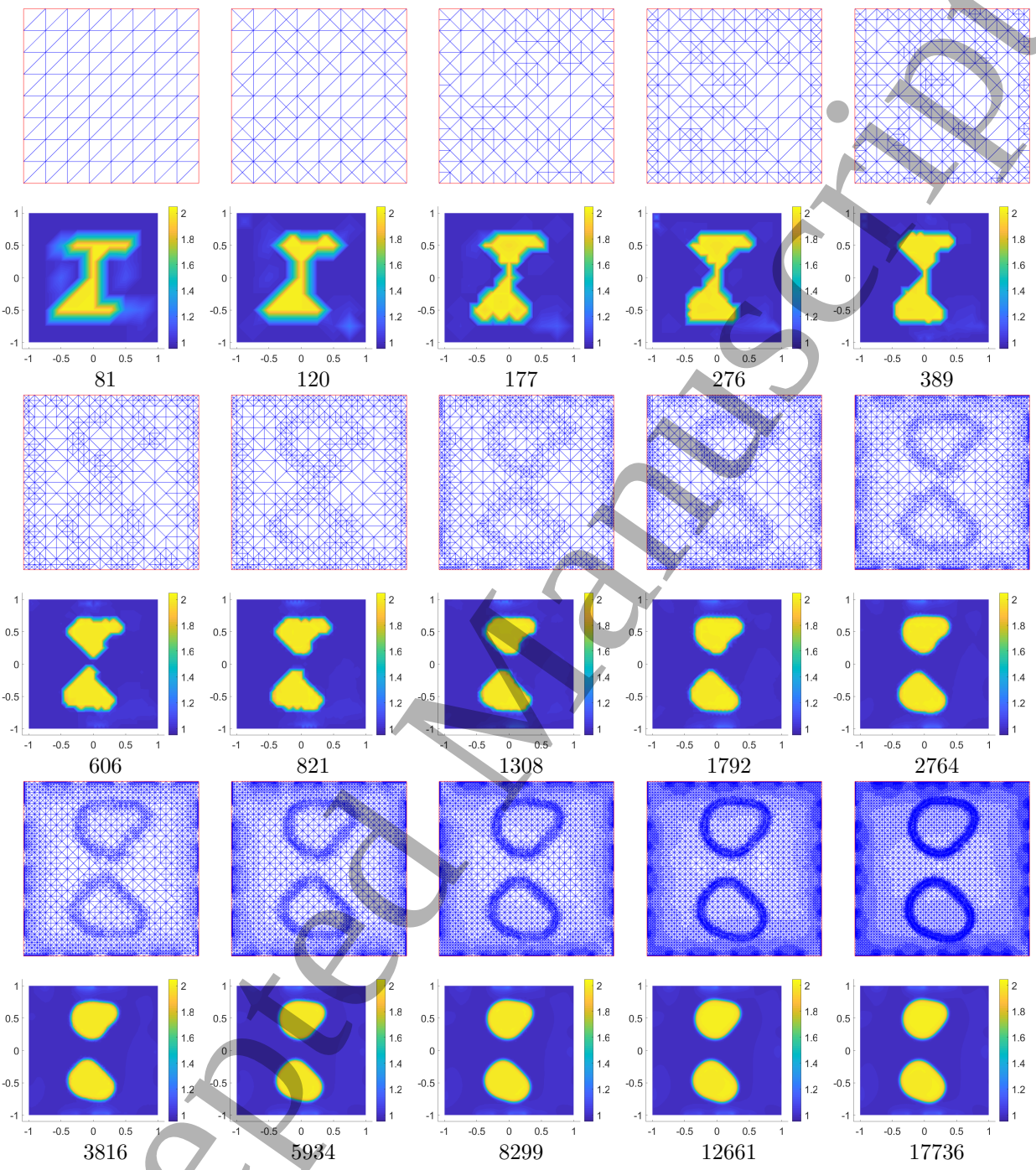
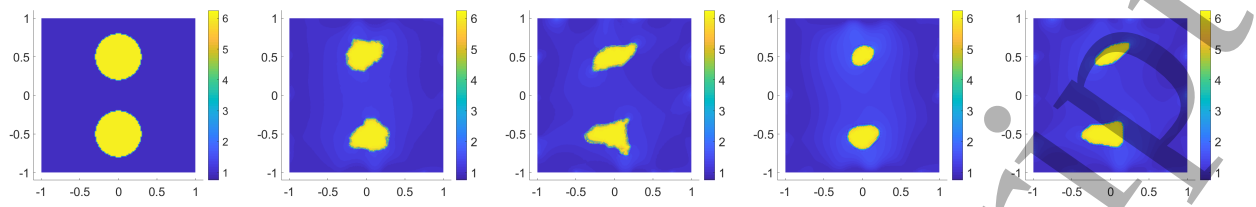


Figure 7: The meshes $\mathcal{T}_k$ and recovered conductivities σ_k during the adaptive iteration for Example 1(ii) with $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 2e-2. The number under each figure refers to d.o.f.

gives $\|U(\sigma_k) - U^\delta\|^2 \to \|U(\sigma_\infty) - U^\delta\|^2$ and $\int_\Omega W(\sigma_k)\,\mathrm{d}x \to \int_\Omega W(\sigma_\infty)\,\mathrm{d}x$. Now (5.6), the weak lower semicontinuity of the $H^1(\Omega)$-norm, (5.7) and the minimizing property of σ_k^* to $\mathcal{J}_{\varepsilon,k}$ over the set $\widetilde{\mathcal{A}}_k$ imply

$$\mathcal{J}_{\varepsilon,\infty}(\sigma^*) \le \liminf_{j\to\infty} \mathcal{J}_{\varepsilon,k_j}(\sigma_{k_j}^*) \le \limsup_{j\to\infty} \mathcal{J}_{\varepsilon,k_j}(\sigma_{k_j}^*)$$



(a) true conductivity (b) adaptive (c) adaptive (d) uniform (e) uniform

Figure 8: The final recoveries by the adaptive and uniform refinements for Example 1(iii). The results in (b) and (d) are for $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 1e-4, and (c) and (e) for $\epsilon =$ 1e-2 and $\tilde{\alpha} =$ 2e-4. The d.o.f. for (b), (c), (d) and (e) is 14620, 19355, 16641 and 16641 respectively.

$$\leq \limsup_{k\to\infty} \mathcal{J}_{\varepsilon,k}(\sigma_k^*) \leq \limsup_{k\to\infty} \mathcal{J}_{\varepsilon,k}(\sigma_k) = \mathcal{J}_{\varepsilon,\infty}(\sigma_\infty) \quad \forall \sigma_\infty \in \widetilde{\mathcal{A}}_\infty. \tag{5.8}$$

Since $\sigma^* \in \widetilde{\mathcal{A}}_\infty$, $\sigma_\infty^* := \sigma^*$ is a minimizer of $\mathcal{J}_{\varepsilon,\infty}$ over $\widetilde{\mathcal{A}}_\infty$. □

Further, we have the following auxiliary convergence.

Theorem 5.2. *The sequence of discrete solutions* $\{\sigma_k^*, (u_k^*, U_k^*)\}_{k\geq 0}$ *to problem* (3.2) *contains a subsequence* $\{\sigma_{k_j}^*, (u_{k_j}^*, U_{k_j}^*)\}_{j\geq 0}$ *convergent to a minimizer* $(\sigma_\infty^*, (u_\infty^*, U_\infty^*))$ *to problem* (5.1)–(5.2)*:*

$$\sigma_{k_j}^* \to \sigma_\infty^* \quad \text{in } H^1(\Omega), \quad \sigma_{k_j}^* \to \sigma_\infty^* \quad \text{a.e. in } \Omega, \quad (u_{k_j}^*, U_{k_j}^*) \to (u_\infty^*, U_\infty^*) \quad \text{in } \mathbb{H}.$$

Proof. The convergence of $(u_{k_j}^*, U_{k_j}^*)$ was already proved in Theorem 5.1. Taking $\sigma_\infty = \sigma_\infty^*$ in (5.8) gives $\lim_{j\to\infty} \mathcal{J}_{\varepsilon,k_j}(\sigma_{k_j}^*) = \mathcal{J}_{\varepsilon,\infty}(\sigma_\infty^*)$. By (5.6) and (5.7), we have $\|\nabla\sigma_{k_j}^*\|_{L^2(\Omega)}^2 \to \|\nabla\sigma_\infty^*\|_{L^2(\Omega)}^2$. Thus, the sequence $\{\sigma_{k_j}^*\}_{j\geq 0}$ converges to σ_∞^* in $H^1(\Omega)$. □

Next we consider the convergence of the sequence $\{(p_k^*, P_k^*)\}_{k\geq 0}$. With a minimizer $(\sigma_\infty^*, (u_\infty^*, U_\infty^*))$ to problem (5.1), we define a limiting adjoint problem: find $(p_\infty^*, P_\infty^*) \in \mathbb{H}_\infty$ such that

$$a(\sigma_\infty^*, (p_\infty^*, P_\infty^*), (v, V)) = \langle U_\infty^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}_\infty. \tag{5.9}$$

By Lemma 2.1 and Lax-Milgram theorem, (5.9) is uniquely solvable. We have the following convergence result for (p_∞^*, P_∞^*). The proof is identical with [32, Theorem 4.5], and hence omitted.

Theorem 5.3. *Under the condition of Theorem 5.2, the subsequence of adjoint solutions* $\{(p_{k_j}^*, P_{k_j}^*)\}_{j\geq 0}$ *generated by Algorithm 3.1 converges to the solution* (p_∞^*, P_∞^*) *of problem* (5.9)*:*

$$\lim_{j\to\infty} \|(p_{k_j}^* - p_\infty^*, P_{k_j}^* - P_\infty^*)\|_{\mathbb{H}} = 0.$$

5.2 Proof of Theorem 3.1

Theorem 3.1 follows directly by combining Theorems 5.2-5.3 in Section 5.1 and Theorems 5.4-5.5 below. The proof in this part relies on the marking condition (3.5). First, we show that the limit $(\sigma_\infty^*, (u_\infty^*, U_\infty^*), (p_\infty^*, P_\infty^*))$ solves the variational equations in (2.8).

Theorem 5.4. *The solutions* $(\sigma_\infty^*, u_\infty^*, U_\infty^*)$ *and* (p_∞^*, P_∞^*) *to problems* (5.1)-(5.2) *and* (5.9) *satisfy*

$$a(\sigma_\infty^*, (u_\infty^*, U_\infty^*), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H},$$
$$a(\sigma_\infty^*, (p_\infty^*, P_\infty^*), (v, V)) = \langle U_\infty^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}.$$

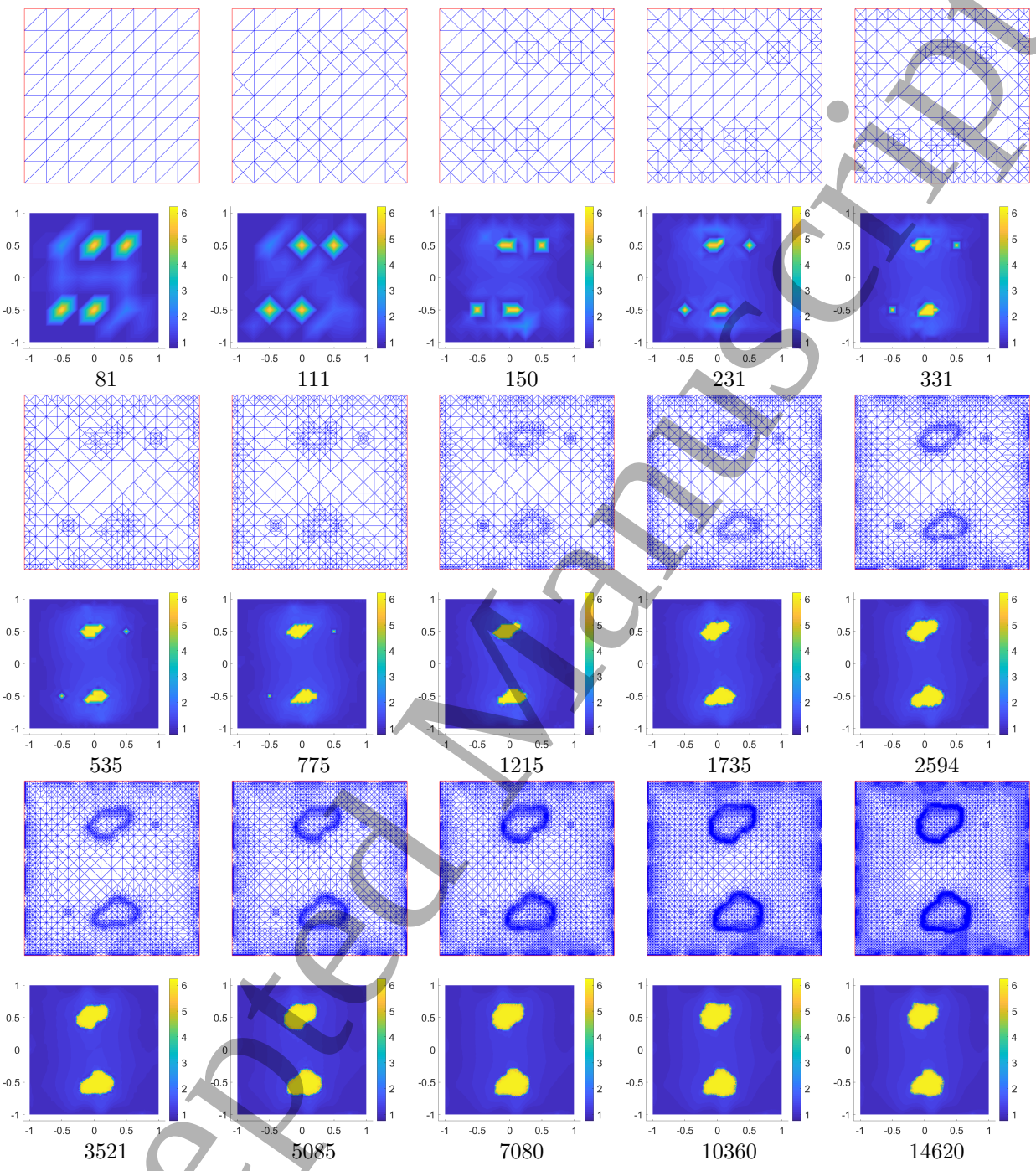
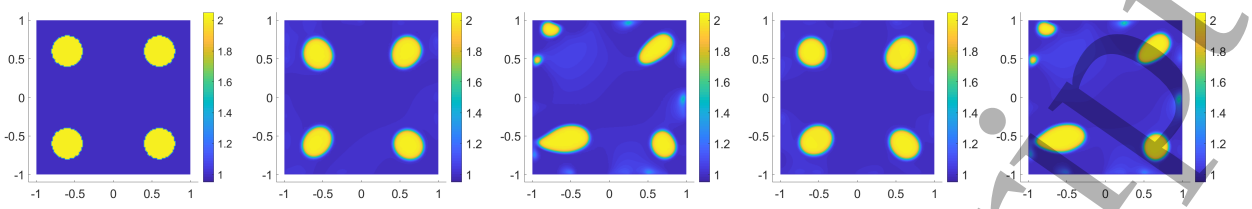


Figure 9: The meshes $\mathcal{T}_k$ and recovered conductivities σ_k during the adaptive iteration for Example 1(iii) with $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 1e-4. The number under each figure refers to d.o.f.

Proof. The proof is identical with [32, Lemma 4.8], using Theorems 5.2-5.3, and hence we only give a brief sketch. By [32, Lemma 3.5], for each $T \in \mathcal{T}_k$ with its face F (intersecting with e_l), there hold

$$\eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, T) \le c(\|\nabla u_k^*\|_{L^2(D_T)}^2 + h_F \|u_k^* - U_{k,l}^*\|_{L^2(F\cap e_l)}^2),$$
$$\eta_{k,2}^2(\sigma_k^*, p_k^*, P_k^*, T) \le c(\|\nabla p_k^*\|_{L^2(D_T)}^2 + h_F \|p_k^* - P_{k,l}^*\|_{L^2(F\cap e_l)}^2),$$



(a) exact (b) adaptive (c) adaptive (d) uniform (e) uniform

Figure 10: The final recoveries by the adaptive and uniform refinements for Example 2. The results in (b) and (d) are for $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 2e-2, and (c) and (e) for $\epsilon =$ 1e-2 and $\tilde{\alpha} =$ 3e-2. The d.o.f. of (b), (c), (d) and (e) is 18008, 21120 and 16641 and 16641, respectively.

where the notation D_T is defined below. Then by the marking condition (3.5), [32, Lemma 4.6] implies that for each convergent subsequence $\{\sigma^*_{k_j}, (u^*_{k_j}, U^*_{k_j}), (p^*_{k_j}, P^*_{k_j})\}_{j\geq 0}$ from Theorems 5.2 and 5.3, there hold

$$\lim_{j\to\infty} \max_{T\in\mathcal{M}^1_{k_j}} \eta_{k_j,1}(\sigma^*_{k_j}, u^*_{k_j}, U^*_{k_j}, T) = 0 \quad \text{and} \quad \lim_{j\to\infty} \max_{T\in\mathcal{M}^2_{k_j}} \eta_{k_j,2}(\sigma^*_{k_j}, p^*_{k_j}, P^*_{k_j}, T) = 0.$$

Last, by [32, Lemma 4.7] and Theorems 5.2-5.3, the argument of [32, Lemma 4.8] completes the proof. □

Remark 5.1. *The argument of Theorem 5.4 dates back to [49], and the main tools include the Galerkin orthogonality of the residual operator, the Lagrange and the Scott-Zhang interpolation operators [16, 48], the marking condition* (3.5) *and a density argument. Further, the error estimators* $\eta_{k,1}(\sigma^*_k, u^*_k, U^*_k)$ *and* $\eta_{k,2}(\sigma^*_k, p^*_k, P^*_k)$ *emerge in the proof and are then employed in the module* `ESTIMATE` *of Algorithm 3.1.*

Next we prove that the limit $(\sigma^*_\infty, (u^*_\infty, U^*_\infty), (p^*_\infty, P^*_\infty))$ satisfies the variational inequality in (2.8). The proof relies crucially on a constraint preserving interpolation operator. We denote by D_T the union of elements in $\mathcal{T}$ with a non-empty intersection with an element $T \in \mathcal{T}$, and by ω_F the union of elements in $\mathcal{T}$ sharing a common face/edge with $F \in \mathcal{F}_\mathcal{T}$. Let

$$\mathcal{T}^+_k := \bigcap_{l\geq k} \mathcal{T}_l, \quad \mathcal{T}^0_k := \mathcal{T}_k \setminus \mathcal{T}^+_k, \quad \Omega^+_k := \bigcup_{T\in\mathcal{T}^+_k} D_T, \quad \Omega^0_k := \bigcup_{T\in\mathcal{T}^0_k} D_T.$$

The set $\mathcal{T}^+_k$ consists of all elements not refined after the k-th iteration, and all elements in $\mathcal{T}^0_k$ are refined at least once after the k-th iteration. Clearly, $\mathcal{T}^+_l \subset \mathcal{T}^+_k$ for $l < k$. We also define a mesh-size function $h_k : \overline{\Omega} \to \mathbb{R}^+$ almost everywhere

$$h_k(x) = \begin{cases} h_T, & x \in T^i, \\ h_F, & x \in F^i, \end{cases}$$

where T^i denotes the interior of an element $T \in \mathcal{T}_k$, and F^i the relative interior of an edge $F \in \mathcal{F}_k$. It has the following property [49, Corollary 3.3]:

$$\lim_{k\to\infty} \|h_k \chi_{\Omega^0_k}\|_{L^\infty(\Omega)} = 0. \tag{5.10}$$

The next result gives the limiting behaviour of the maximal error indicator $\eta_{k,3}$.

Lemma 5.2. *Let* $\{(\sigma^*_k, (u^*_k, U^*_k), (p^*_k, P^*_k))\}_{k\geq 0}$ *be the sequence of discrete solutions generated by Algorithm 3.1. Then for each convergent subsequence* $\{\sigma^*_{k_j}, (u^*_{k_j}, U^*_{k_j}), (p^*_{k_j}, P^*_{k_j})\}_{j\geq 0}$*, there holds*

$$\lim_{j\to\infty} \max_{T\in\mathcal{M}^3_{k_j}} \eta_{k_j,3}(\sigma^*_{k_j}, u^*_{k_j}, p^*_{k_j}, T) = 0.$$

Proof. The inverse estimate and scaled trace theorem imply that for each $T \in \mathcal{T}_k$ (with its face F)

$$h^q_T \|\tfrac{\tilde{\alpha}}{2\varepsilon} W'(\sigma^*_k) - \nabla u^*_k \cdot \nabla p^*_k\|^q_{L^q(T)} \leq c h^q_T \|\nabla u^*_k \cdot \nabla p^*_k\|^q_{L^q(T)} + c h^q_T \|W'(\sigma^*_k)\|^q_{L^q(T)}$$

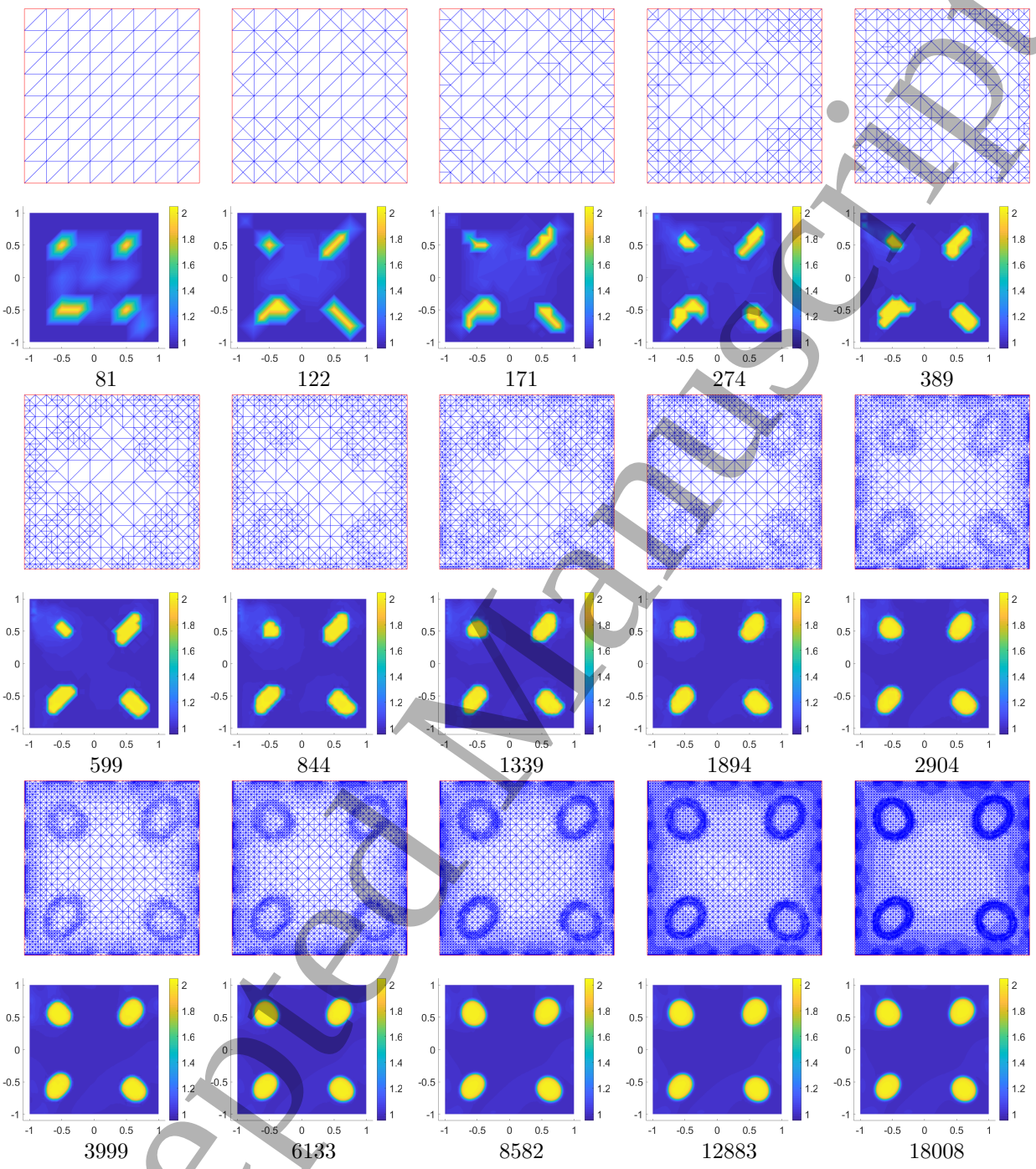
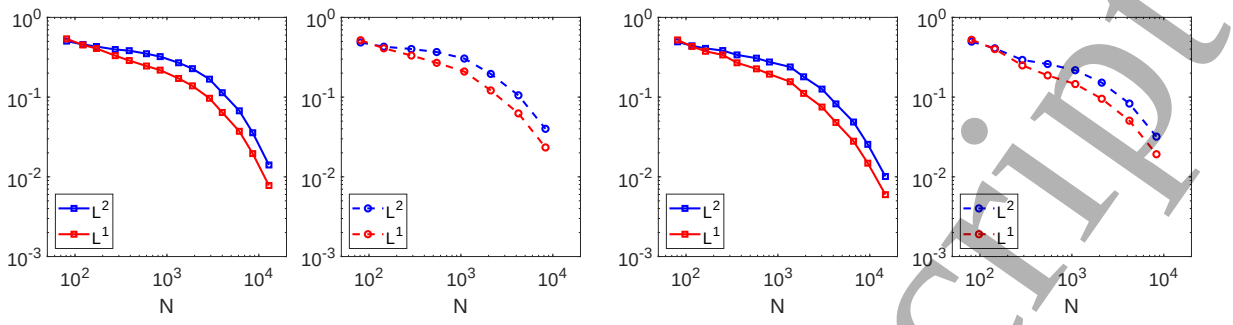


Figure 11: The meshes $\mathcal{T}_k$ and recovered conductivities σ_k during the adaptive refinement, for Example 2 with $\epsilon =$ 1e-3 and $\tilde{\alpha} =$ 2e-2. The number under each figure refers to d.o.f.

$$\leq ch_T^q h_T^{d-dq} \|\nabla u_k^* \cdot \nabla p_k^*\|_{L^1(T)}^q + ch_T^q \|W'(\sigma_k^*)\|_{L^q(T)}^q,$$

$$\sum_{F\subset\partial T} h_F \|J_{F,2}(\sigma_k^*)\|_{L^q(F)}^q \leq c \sum_{F\subset\partial T} h_F h_F^{-1} \|\nabla \sigma_k^*\|_{L^q(\omega_F)}^q.$$



(a) ϵ = 1e-3, $\tilde{\alpha}$ = 2e-2 (b) ϵ = 1e-2, $\tilde{\alpha}$ = 3e-2

Figure 12: The $L^2(\Omega)$ and $L^1(\Omega)$ errors versus the degree of freedom N of the mesh, for Example 2, using the adaptive (solid) and uniform (dashed) refinement.

With the choice $q = d/(d-1)$, combining these two estimates gives

$$\eta_{k,3}^q(\sigma_k^*, u_k^*, p_k^*, T) \le c(\|\nabla u_k^* \cdot \nabla p_k^*\|_{L^1(T)}^q + h_T^q \|W'(\sigma_k^*)\|_{L^q(T)}^q + \|\nabla \sigma_k^*\|_{L^q(D_T)}^q), \tag{5.11}$$

where c depends on $\widetilde{\alpha}$ and ε in $\mathcal{F}_\varepsilon$. Next, for the subsequence $\{\sigma_{k_j}^*, (u_{k_j}^*, U_{k_j}^*), (p_{k_j}^*, P_{k_j}^*)\}_{j\ge 0}$, let $\widetilde{T}_j^3 \in \mathcal{M}_{k_j}^3$ be the element with the largest error indicator $\eta_{k_j,3}(\sigma_{k_j}^*, u_{k_j}^*, p_{k_j}^*, T)$. Since $D_{\widetilde{T}_j^i} \subset \Omega_{k_j}^0$, (5.10) implies

$$|D_{\widetilde{T}_j^i}| \le c\|h_{k_j}\|_{L^\infty(\Omega_{k_j}^0)}^d \to 0 \quad \text{as } j \to \infty. \tag{5.12}$$

By (5.11), Cauchy-Schwarz inequality and triangle inequality, there holds

$$\begin{aligned}
\eta_{k_j,3}^q(\sigma_{k_j}^*, u_{k_j}^*, p_{k_j}^*, \widetilde{T}_j^3) &\le c(\|\nabla u_{k_j}^*\|_{L^2(\widetilde{T}_j^3)}^q \|\nabla p_{k_j}^*\|_{L^2(\widetilde{T}_j^3)}^q + h_{\widetilde{T}_j^3}^q \|W'(\sigma_k^*)\|_{L^q(\widetilde{T}_j^3)}^q + \|\nabla \sigma_{k_j}^*\|_{L^q(D_{\widetilde{T}_j^3})}^q) \\
&\le c\big((\|\nabla(u_{k_j}^* - u_\infty^*)\|_{L^2(\Omega)}^q + \|\nabla u_\infty^*\|_{L^2(\widetilde{T}_j^3)}^q)(\|\nabla(p_{k_j}^* - p_\infty^*)\|_{L^2(\Omega)}^q + \|\nabla p_\infty^*\|_{L^2(\widetilde{T}_j^3)}^q) \\
&\quad + h_{\widetilde{T}_j^3}^q(\|W'(\sigma_{k_j}^*) - W'(\sigma_\infty^*)\|_{L^q(\Omega)}^q + \|W'(\sigma_\infty^*)\|_{L^q(\widetilde{T}_j^3)}^q) \\
&\quad + (\|\nabla(\sigma_{k_j}^* - \sigma_\infty^*)\|_{L^q(\Omega)}^q + \|\nabla \sigma_\infty^*\|_{L^q(D_{\widetilde{T}_j^3})}^q)\big).
\end{aligned}$$

By Theorems 5.2 and 5.3, Lebesgue's dominated convergence theorem, the choice $q = d/(d-1) \le 2$ and Hölder inequality, we obtain $\|W'(\sigma_{k_j}^*) - W'(\sigma_\infty^*)\|_{L^q(\Omega)}^q \to 0$ and $\|\nabla(\sigma_{k_j}^* - \sigma_\infty^*)\|_{L^q(\Omega)}^q \to 0$. Then the absolute continuity of the norm $\|\cdot\|_{L^q(\Omega)}$ with respect to Lebesgue measure and (5.12) complete the proof. □

Due to a lack of Galerkin orthogonality for variational inequalities, we employ a local L^r-stable interpolation operator of Clément/Chen-Nochetto type. Let $\mathcal{N}_k$ be the set of all interior nodes of $\mathcal{T}_k$ and $\{\phi_x\}_{x\in\mathcal{N}_k}$ be the nodal basis functions in V_k. For each $x \in \mathcal{N}_k$, the support of ϕ_x is denoted by ω_x, i.e., the union of all elements in $\mathcal{T}_k$ with a non-empty intersection with x. Then we define $\Pi_k : L^1(\Omega) \to V_k$ by

$$\Pi_k v := \sum_{x\in\mathcal{N}_k} \frac{1}{|\omega_x|} \int_{\omega_x} v \,\mathrm{d}x \phi_x. \tag{5.13}$$

Clearly, $\Pi_k v \in \widetilde{\mathcal{A}}_k$ if $c_0 \le v \le c_1$ a.e. $x \in \Omega$. The definition is adapted from [13] (for elliptic obstacle problems) by replacing the maximal ball $\Delta_x \subset \omega_x$ centered at an interior node x by ω_x. $\Pi_k v$ satisfies following properties; see Appendix B for a proof.

Lemma 5.3. *For any $v \in W^{1,r}(\Omega)$, there hold for all $r \in [1, +\infty]$, any $T \in \mathcal{T}_k$ and any $F \subset \partial T$,*

$$\|\Pi_k v\|_{L^r(T)} \le c\|v\|_{L^r(D_T)}, \quad \|\nabla \Pi_k v\|_{L^r(T)} \le c\|\nabla v\|_{L^r(D_T)},$$
$$\|v - \Pi_k v\|_{L^r(T)} \le ch_T\|\nabla v\|_{L^r(D_T)}, \quad \|v - \Pi_k v\|_{L^r(F)} \le ch_F^{1-1/r}\|\nabla v\|_{L^r(D_T)}.$$

Last we show that the limit $(\sigma^*_\infty,(u^*_\infty,U^*_\infty),(p^*_\infty,P^*_\infty))$ satisfies the variational inequality in (2.8).

Theorem 5.5. *The solutions* $(\sigma^*_\infty,u^*_\infty,U^*_\infty)$ *and* (p^*_∞,P^*_∞) *to problems* (5.1)-(5.2) *and* (5.9) *satisfy*

$$\widetilde{\alpha}\varepsilon(\nabla\sigma^*_\infty,\nabla(\mu-\sigma^*_\infty))+\tfrac{\widetilde{\alpha}}{2\varepsilon}(W'(\sigma^*_\infty),\mu-\sigma^*_\infty)-(\nabla u^*_\infty,\nabla p^*_\infty(\mu-\sigma^*_\infty))\geq 0\quad\forall\mu\in\widetilde{\mathcal{A}}.$$

Proof. The proof is lengthy, and we break it into five steps.

Step i. Derive a preliminary variational inequality. We relabel the subsequence $\{\sigma^*_{k_j},(u^*_{k_j},U^*_{k_j}),(p^*_{k_j},P^*_{k_j})\}_{j\geq 0}$ in Theorems 5.2 and 5.3 as $\{\sigma^*_k,(u^*_k,U^*_k),(p^*_k,P^*_k)\}_{k\geq 0}$. Let I_k be the Lagrange interpolation operator on V_k, and let $\alpha'=\widetilde{\alpha}\varepsilon$ and $\alpha''=\frac{\widetilde{\alpha}}{2\varepsilon}$. For any $\mu\in\widetilde{\mathcal{A}}\cap C^\infty(\overline{\Omega})$, $I_k\mu\in\widetilde{\mathcal{A}}_k$ and let $\nu=\mu-I_k\mu$. Direct computation gives

$$\begin{aligned}
&\alpha'(\nabla\sigma^*_k,\nabla(\mu-\sigma^*_k))+\alpha''(W'(\sigma^*_k),\mu-\sigma^*_k)-((\mu-\sigma^*_k)\nabla u^*_k,\nabla p^*_k)\\
&\quad=\alpha'(\nabla\sigma^*_k,\nabla(\mu-I_k\mu))+\alpha''(W'(\sigma^*_k),\mu-I_k\mu)-((\mu-I_k\mu)\nabla u^*_k,\nabla p^*_k)\\
&\qquad+\alpha'(\nabla\sigma^*_k,\nabla(I_k\mu-\sigma^*_k))+\alpha''(W'(\sigma^*_k),I_k\mu-\sigma^*_k)-((I_k\mu-\sigma^*_k)\nabla u^*_k,\nabla p^*_k)\\
&\quad=\alpha'(\nabla\sigma^*_k,\nabla(\nu-\Pi_k\nu))+\alpha''(W'(\sigma^*_k),\nu-\Pi_k\nu)-((\nu-\Pi_k\nu)\nabla u^*_k,\nabla p^*_k)\\
&\qquad+\alpha'(\nabla\sigma^*_k,\nabla\Pi_k\nu)+\alpha''(W'(\sigma^*_k),\Pi_k\nu)-(\Pi_k\nu\nabla u^*_k,\nabla p^*_k)\\
&\qquad+\alpha'(\nabla\sigma^*_k,\nabla(I_k\mu-\sigma^*_k))+\alpha''(W'(\sigma^*_k),I_k\mu-\sigma^*_k)-((I_k\mu-\sigma^*_k)\nabla u^*_k,\nabla p^*_k)\\
&\quad\geq[\alpha'(\nabla\sigma^*_k,\nabla(\nu-\Pi_k\nu))+\alpha''(W'(\sigma^*_k),\nu-\Pi_k\nu)-((\nu-\Pi_k\nu)\nabla u^*_k,\nabla p^*_k)]\\
&\qquad+[\alpha'(\nabla\sigma^*_k,\nabla\Pi_k\nu)+\alpha''(W'(\sigma^*_k),\Pi_k\nu)-(\Pi_k\nu\nabla u^*_k,\nabla p^*_k)]:=\mathrm{I}+\mathrm{II},
\end{aligned}\tag{5.14}$$

where the last inequality is due to the variational inequality in (3.3) with $\mu_k=I_k\mu$.

Step ii. Bound the I. By elementwise integration by parts, Hölder inequality, the definition of the estimator $\eta_{k,3}$ and Lemma 5.3 with $r=q'$ (with q' being the conjugate exponent of q),

$$\begin{aligned}
|\mathrm{I}|&=\Big|\sum_{T\in\mathcal{T}_k}\int_T R_{T,2}(\sigma^*_k,u^*_k,p^*_k)(\nu-\Pi_k\nu)\,\mathrm{d}x+\sum_{F\in\mathcal{F}_k}\int_F J_{F,2}(\sigma^*_k)(\nu-\Pi_k\nu)\mathrm{d}s\Big|\\
&\leq\sum_{T\in\mathcal{T}_k}\Big(\|R_{T,2}(\sigma^*_k,u^*_k,p^*_k)\|_{L^q(T)}\|\nu-\Pi_k\nu\|_{L^{q'}(T)}+\sum_{F\subset\partial T}\|J_{F,2}(\sigma^*_k)\|_{L^q(F)}\|\nu-\Pi_k\nu\|_{L^{q'}(F)}\Big)\\
&\leq c\sum_{T\in\mathcal{T}_k}\Big(h_T\|R_{T,2}(\sigma^*_k,u^*_k,p^*_k)\|_{L^q(T)}+\sum_{F\subset\partial T}h_F^{1/q}\|J_{F,2}(\sigma^*_k)\|_{L^q(F)}\Big)\|\nabla\nu\|_{L^{q'}(D_T)}\\
&\leq c\sum_{T\in\mathcal{T}_k}\eta_{k,3}(\sigma^*_k,u^*_k,p^*_k,T)\|\nabla\nu\|_{L^{q'}(D_T)}.
\end{aligned}$$

Thus, for any $k>l$, by (discrete) Hölder's inequality and the finite overlapping property of the patches D_T, due to uniform shape regularity of the meshes $\mathcal{T}_k\in\mathbb{T}$, there holds

$$\begin{aligned}
|\mathrm{I}|&\leq c\Big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}(\sigma^*_k,u^*_k,p^*_k,T)\|\nabla\nu\|_{L^{q'}(D_T)}+\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}(\sigma^*_k,u^*_k,p^*_k,T)\|\nabla\nu\|_{L^{q'}(D_T)}\Big)\\
&\leq c\Big(\Big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta^q_{k,3}(\sigma^*_k,u^*_k,p^*_k,T)\Big)^{1/q}\|\nabla(\mu-I_k\mu)\|_{L^{q'}(\Omega_l^0)}\\
&\qquad+\Big(\sum_{T\in\mathcal{T}_l^+}\eta^q_{k,3}(\sigma^*_k,u^*_k,p^*_k,T)\Big)^{1/q}\|\nabla(\mu-I_k\mu)\|_{L^{q'}(\Omega_l^+)}\Big).
\end{aligned}$$

Since $W'(s)\in C^1[c_0,c_1]$, by the pointwise convergence of $\{\sigma^*_k\}_{k\geq 0}$ in Theorem 5.2 and Lebesgue's dominated convergence theorem, we deduce

$$W'(\sigma^*_k)\to W'(\sigma^*_\infty)\quad\text{in } L^2(\Omega).\tag{5.15}$$

Since $q=d/(d-1)\leq 2$, the sequence $\{W'(\sigma^*_k)\}_{k\geq 0}$ is uniformly bounded in $L^q(\Omega)$. By Theorems 5.2 and 5.3, the sequences $\{\sigma^*_k\}_{k\geq 0}$, $\{u^*_k\}_{k\geq 0}$ and $\{p^*_k\}_{k\geq 0}$ are uniformly bounded in $H^1(\Omega)$. Thus, (5.11) and (5.10),

and Hölder inequality give

$$\begin{aligned}
&\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\\
\le& c\Big(\|\nabla u_k^*\cdot\nabla p_k^*\|_{L^1(\Omega)}^{q-1}\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\|\nabla u_k^*\cdot\nabla p_k^*\|_{L^1(T)}+\|h_l\|_{L^\infty(\Omega_l^0)}^q\|W'(\sigma_k^*)\|_{L^q(\Omega)}^q+\|\nabla\sigma_k^*\|_{L^q(\Omega)}^q\Big)\\
\le& c\Big(\|\nabla u_k^*\|_{L^2(\Omega)}^q\|\nabla p_k^*\|_{L^2(\Omega)}^q+\|h_l\chi_{\Omega_l^0}\|_{L^\infty(\Omega)}^q\|W'(\sigma_k^*)\|_{L^q(\Omega)}^q+\|\nabla\sigma_k^*\|_{L^2(\Omega)}^q\Big)\le c. \qquad (5.16)
\end{aligned}$$

Then by the error estimate of I_k [16],

$$|\mathrm{I}|\le c\|h_l\chi_{\Omega_l^0}\|_{L^\infty(\Omega)}\|\mu\|_{W^{2,q'}(\Omega)}+c\big(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\|\mu\|_{W^{2,q'}(\Omega)}.$$

By (5.10), $c\|h_l\chi_{\Omega_l^0}\|_{L^\infty(\Omega)}\|\mu\|_{W^{2,q'}(\Omega)}\to 0$ as $l\to\infty$. Since $\mathcal{T}_l^+\subset\mathcal{T}_k$ for $k>l$, (3.5) implies

$$\big(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\le|\mathcal{T}_l^+|^{1/q}\max_{T\in\mathcal{T}_l^+}\eta_{k,3}(\sigma_k^*,u_k^*,p_k^*,T)\le|\mathcal{T}_l^+|^{1/q}\max_{T\in\mathcal{M}_k^3}\eta_{k,3}(\sigma_k^*,u_k^*,p_k^*,T).$$

By Lemma 5.2, for any small $\varepsilon>0$, we can choose $k_1>l_1$ for some large fixed l_1 such that whenever $k>k_1$,

$$c\big(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\|\mu\|_{W^{2,q'}(\Omega)}<\varepsilon.$$

Consequently,

$$\mathrm{I}\to 0\quad\forall\mu\in\widetilde{\mathcal{A}}\cap C^\infty(\overline{\Omega}). \qquad (5.17)$$

Step iii. Bound the term II. For the term II, elementwise integration and Hölder inequality yield

$$\begin{aligned}
|\mathrm{II}|&=\Big|\sum_{T\in\mathcal{T}_k}\int_T R_{T,2}(\sigma_k^*,u_k^*,p_k^*)\Pi_k\nu\,\mathrm{d}x+\sum_{F\in\mathcal{F}_k}\int_F J_{F,2}(\sigma_k^*)\Pi_k\nu\mathrm{d}s\Big|\\
&\le\sum_{T\in\mathcal{T}_k}\Big(\|R_{T,2}(\sigma_k^*,u_k^*,p_k^*)\|_{L^q(T)}\|\Pi_k\nu\|_{L^{q'}(T)}+\sum_{F\subset\partial T}\|J_{F,2}(\sigma_k^*)\|_{L^q(F)}\|\Pi_k\nu\|_{L^{q'}(F)}\Big)
\end{aligned}$$

By the scaled trace theorem, local inverse estimate, $L^{q'}$-stability of Π_k in Lemma 5.3, local quasi-uniformity and interpolation error estimate for I_k [16], we deduce that for $k>l$

$$\begin{aligned}
|\mathrm{II}|&\le c\sum_{T\in\mathcal{T}_k}\Big(h_T\|R_{T,2}(\sigma_k^*,u_k^*,p_k^*)\|_{L^q(T)}h_T^{-1}\|\Pi_k\nu\|_{L^{q'}(T)}+\sum_{F\subset\partial T}h_F^{1/q}\|J_{F,2}(\sigma_k^*)\|_{L^q(F)}h_F^{-1/q-1/q'}\|\Pi_k\nu\|_{L^{q'}(T)}\Big)\\
&\le c\sum_{T\in\mathcal{T}_k}\Big(h_T\|R_{T,2}(\sigma_k^*,u_k^*,p_k^*)\|_{L^q(T)}+\sum_{F\subset\partial T}h_F^{1/q}\|J_{F,2}(\sigma_k^*)\|_{L^q(F)}\Big)h_T^{-1}\|\nu\|_{L^{q'}(D_T)}\\
&\le c\sum_{T\in\mathcal{T}_k}\eta_{k,3}(\sigma_k^*,u_k^*,p_k^*,T)h_T^{-1}\|\mu-I_k\mu\|_{L^{q'}(D_T)}\\
&=c\big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}(\sigma_k^*,u_k^*,p_k^*,T)h_T\|\mu\|_{W^{2,q'}(D_T)}+\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}(\sigma_k^*,u_k^*,p_k^*,T)h_T\|\mu\|_{W^{2,q'}(D_T)}\big)\\
&\le c\|h_l\chi_{\Omega_l^0}\|_{L^\infty(\Omega)}\big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\|\mu\|_{W^{2,q'}(\Omega)}+c\big(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\|\mu\|_{W^{2,q'}(\Omega)}.
\end{aligned}$$

Since $\big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\le c$, cf. (5.16), there holds

$$|\mathrm{II}|\le c\|h_l\chi_{\Omega_l^0}\|_{L^\infty(\Omega)}\|\mu\|_{W^{2,q'}(\Omega)}+c\big(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^q(\sigma_k^*,u_k^*,p_k^*,T)\big)^{1/q}\|\mu\|_{W^{2,q'}(\Omega)}.$$

Now by repeating the argument for the term I, we obtain

$$\mathrm{II} \to 0 \quad \forall \mu \in \widetilde{\mathcal{A}} \cap C^\infty(\overline{\Omega}). \tag{5.18}$$

Step iv. Take limit in preliminary variational inequality. Using (5.15) and the $H^1(\Omega)$-convergence of $\{\sigma_k^*\}_{k\geq 0}$ in Theorem 5.2, we have for each $\mu \in \widetilde{\mathcal{A}} \cap C^\infty(\overline{\Omega})$

$$\alpha'(\nabla\sigma_k^*, \nabla(\mu - \sigma_k^*)) + \alpha''(W'(\sigma_k^*), \mu - \sigma_k^*) \to \alpha'(\nabla\sigma_\infty^*, \nabla(\mu - \sigma_\infty^*)) + \alpha''(W'(\sigma_\infty^*), \mu - \sigma_\infty^*). \tag{5.19}$$

Further, the uniform boundedness on $\{u_k^*\}_{k\geq 0}$ in $H^1(\Omega)$ and the convergence of $\{p_k^*\}_{k\geq 0}$ to p_∞^* in $H^1(\Omega)$ in Theorem 5.3 yield

$$|(\mu\nabla u_k^*, \nabla(p_k^* - p_\infty^*))| \leq c\|\nabla(p_k^* - p_\infty^*)\|_{L^2(\Omega)} \to 0.$$

This and Theorem 5.2 imply

$$(\mu\nabla u_k^*, \nabla p_k^*) = (\mu\nabla u_k^*, \nabla(p_k^* - p_\infty^*)) + (\mu\nabla u_k^*, \nabla p_\infty^*) \to (\mu\nabla u_\infty^*, \nabla p_\infty^*) \quad \forall \mu \in \widetilde{\mathcal{A}} \cap C^\infty(\overline{\Omega}). \tag{5.20}$$

In the splitting

$$\begin{aligned}(\sigma_k^*\nabla u_k^*, \nabla p_k^*) - (\sigma_\infty^*\nabla u_\infty^*, \nabla p_\infty^*) = (\sigma_k^*\nabla u_k^*, \nabla(p_k^* - p_\infty^*)) + ((\sigma_k^* - \sigma_\infty^*)\nabla u_k^*, \nabla p_\infty^*) \\ + (\sigma_\infty^*\nabla(u_k^* - u_\infty^*), \nabla p_\infty^*),\end{aligned}$$

the arguments for (5.20) directly yields

$$|(\sigma_k^*\nabla u_k^*, \nabla(p_k^* - p_\infty^*))| \to 0 \quad \text{and} \quad |(\sigma_\infty^*\nabla(u_k^* - u_\infty^*), \nabla p_\infty^*)| \to 0.$$

The boundedness on $\{u_k^*\}_{k\geq 0}$ in $H^1(\Omega)$, pointwise convergence of $\{\sigma_k^*\}_{k\geq 0}$ of Theorem 5.2 and Lebesgue's dominated convergence theorem imply

$$|((\sigma_k^* - \sigma_\infty^*)\nabla u_k^*, \nabla p_\infty^*)| \leq c\|(\sigma_k^* - \sigma_\infty^*)\nabla p_\infty^*\|_{L^2(\Omega)} \to 0.$$

Hence, there holds

$$(\sigma_k^*\nabla u_k^*, \nabla p_k^*) \to (\sigma_\infty^*\nabla u_\infty^*, \nabla p_\infty^*). \tag{5.21}$$

Now by passing both sides of (5.14) to the limit and combining the estimates (5.17)-(5.21), we obtain

$$\alpha'(\nabla\sigma_\infty^*, \nabla(\mu - \sigma_\infty^*)) + \alpha''(W'(\sigma_\infty^*), \mu - \sigma_\infty^*) - (\nabla u_\infty^*, \nabla p_\infty^*(\mu - \sigma_\infty^*))_{L^2(\Omega)} \geq 0 \quad \forall \mu \in \widetilde{\mathcal{A}} \cap C^\infty(\overline{\Omega}).$$

Step v. Density argument. By the density of $C^\infty(\overline{\Omega})$ in $H^1(\Omega)$ and the construction via a standard mollifier [19], for any $\mu \in \widetilde{\mathcal{A}}$ there exists a sequence $\{\mu_n\} \subset \widetilde{\mathcal{A}} \cap C^\infty(\overline{\Omega})$ such that $\|\mu_n - \mu\|_{H^1(\Omega)} \to 0$ as $n \to \infty$. Thus, $(\nabla\sigma_\infty^*, \nabla\mu_n) \to (\nabla\sigma_\infty^*, \nabla\mu)$, $(W'(\sigma_\infty^*), \mu_n) \to (W'(\sigma_\infty^*), \mu)$, and $(\mu_n\nabla u_\infty^*, \nabla p_\infty^*) \to (\mu\nabla u_\infty^*, \nabla p_\infty^*)$, after possibly passing to a subsequence. The desired result follows from the preceding two estimates. □

Remark 5.2. *The computable quantity $\eta_{k,3}(\sigma_k^*, u_k^*, p_k^*, T)$ emerges naturally from the proof, i.e., the upper bounds on* I *and* II*, which motivates its use as the a posteriori error estimator in Algorithm 3.1.*

Acknowledgements

The authors are grateful to an anonymous referee and the boarder member for the constructive comments, which have significantly improved the presentation of the paper. The work of Y. Xu was partially supported by National Natural Science Foundation of China (11201307), Ministry of Education of China through Specialized Research Fund for the Doctoral Program of Higher Education (20123127120001) and Natural Science Foundation of Shanghai (17ZR1420800).

A The solution of the variational inequality

Now we describe an iterative method for minimizing the energy functional

$$\frac{\tilde{\alpha}\varepsilon}{2}\|\nabla\sigma\|^2_{L^2(\Omega)} + \frac{\tilde{\alpha}}{2\varepsilon}\int_\Omega W(\sigma)\mathrm{d}x + \frac{1}{2}\|U(\sigma) - U^\delta\|^2.$$

Let $p(z) = (z - c_0)(z - c_1)$. Then one linearized approximation $p_L(z, z_k)$ reads (with $\delta z = z - z_k$)

$$\begin{aligned} p_L(z, z_k) &= p(z_k) + p'(z_k)(z - z_k) \\ &= (z_k^2 - (c_0 + c_1)z_k + c_0c_1) + (2z_k - c_0 - c_1)\delta z. \end{aligned}$$

Upon substituting the approximation $p_L(z, z_k)$ for $p(z)$ and linearizing the forward map $U(\sigma)$, we obtain the following surrogate energy functional (with $\delta\sigma = \sigma - \sigma_k$ being the increment and $\delta U = U^\delta - U(\sigma_k)$)

$$\tfrac{\tilde{\alpha}\varepsilon}{2}\|\nabla(\sigma_k + \delta\sigma)\|^2_{L^2(\Omega)} + \tfrac{\tilde{\alpha}}{2\varepsilon}\|p(\sigma_k) + p'(\sigma_k)\delta\sigma\|^2_{L^2(\Omega)} + \tfrac{1}{2}\|U'(\sigma_k)\delta\sigma - \delta U\|^2. \tag{A.1}$$

The treatment of the double well potential term $\int_\Omega W(\sigma)\mathrm{d}x$ is in the spirit of the classical majorization-minimization algorithm in the following sense (see [56] for a detailed derivation)

$$\int_\Omega W(\sigma_k)\mathrm{d}x = \int_\Omega p_L(\sigma_k, \sigma_k)^2\mathrm{d}x, \quad \nabla\int_\Omega W(\sigma_k)\mathrm{d}x = \nabla\int_\Omega p_L(\sigma_k, \sigma_k)^2\mathrm{d}x,$$
$$\text{and} \quad \nabla^2\int_\Omega W(\sigma_k)\mathrm{d}x \le \nabla^2\int_\Omega p_L(\sigma_k, \sigma_k)^2\mathrm{d}x.$$

This algorithm is known to have excellent numerical stability. Upon ignoring the box constraint on the conductivity σ, problem (A.1) is to find $\delta\sigma \in H^1(\Omega)$ such that

$$\begin{aligned} &(U'(\sigma_k)^*U'(\sigma_k)\delta\sigma, \phi) + \tilde{\alpha}\varepsilon(\nabla\delta\sigma, \nabla\phi) + \tfrac{\tilde{\alpha}}{\varepsilon}(p'(\sigma_k)^2\delta\sigma, \phi) \\ &\qquad = (U'(\sigma_k)^*\delta U, \phi) - \tfrac{\tilde{\alpha}}{\varepsilon}(p(\sigma_k)p'(\sigma_k), \phi) - \tilde{\alpha}\varepsilon(\nabla\sigma_k, \nabla\phi), \quad \forall\phi \in H^1(\Omega). \end{aligned}$$

This equation can be solved by an iterative method for the update $\delta\sigma$ (with the box constraint treated by a projection step). Note that $U'(\sigma_k)$ and $U'(\sigma_k)^*$ can be implemented in matrix-free manner using the standard adjoint technique. In our experiment, we employ the conjugate gradient method to solve the resulting linear systems, preconditioned by the sparse matrix corresponding to $\tilde{\alpha}\varepsilon(\nabla\delta\sigma, \nabla\phi) + \frac{\tilde{\alpha}}{\varepsilon}(p'(\sigma_k)^2\delta\sigma, \phi)$.

B Proof of Lemma 5.3

The proof follows that in [13, 25]. By Hölder inequality and $h_T^d \le |\omega_x|$ for each node $x \in T$,

$$\left|\frac{1}{|\omega_x|}\int_{\omega_x} v\,\mathrm{d}x\right| \le |\omega_x|^{-1/r}\|v\|_{L^r(\omega_x)} \le h_T^{-d/r}\|v\|_{L^r(\omega_x)}.$$

The desired L^r-stability follows from the estimate $\|\phi_x\|_{L^r(T)} \le ch_T^{d/r}$, by the local quasi-uniformity of the mesh. In view of the definition (5.13), $\Pi_k\zeta = \zeta$ for any $\zeta \in \mathbb{R}$. By local inverse estimate, the L^r-stability of Π_k, standard interpolation error estimate [16] and local quasi-uniformity,

$$\begin{aligned} \|\nabla\Pi_k v\|_{L^r(T)} &= \inf_{\zeta\in\mathbb{R}}\|\nabla\Pi_k(v - \zeta)\|_{L^r(T)} \le ch_T^{-1}\inf_{\zeta\in\mathbb{R}}\|\Pi_k(v - \zeta)\|_{L^r(T)} \\ &\le ch_T^{-1}\inf_{\zeta\in\mathbb{R}}\|v - \zeta\|_{L^r(D_T)} \le ch_T^{-1}\|v - \frac{1}{|D_T|}\int_{D_T} v\,\mathrm{d}x\|_{L^r(D_T)} \le c\|\nabla v\|_{L^r(D_T)}. \end{aligned} \tag{B.1}$$

Similarly,

$$\begin{aligned} \|v - \Pi_k v\|_{L^r(T)} &= \|v - \zeta - \Pi_k(v - \zeta)\|_{L^r(T)} \\ &\le c\inf_{\zeta\in\mathbb{R}}\|v - \zeta\|_{L^r(D_T)} \le ch_T\|\nabla v\|_{L^r(D_T)}. \end{aligned} \tag{B.2}$$

By the scaled trace theorem, for any $F \subset \partial T$, there holds

$$\|v - \Pi_k v\|_{L^r(F)} \le c(h_F^{-1/r}\|v - \Pi_k v\|_{L^r(T)} + h_F^{1-1/r}\|\nabla(v - \Pi_k v)\|_{L^r(T)}).$$

Then (B.1) and (B.2) complete the proof of the lemma.

References

[1] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. Wiley-Interscience, New York, 2000.

[2] G. Alberti. Variational models for phase transitionsan approach via Γ-convergence. In L. Ambrosio, N. Dancer, G. Buttazzo, A. Marino, and M. K. V. Murthy, editors, *Calculus of Variations and Partial Differential Equations*, pages 95–114. Springer, New York, 2000.

[3] G. S. Alberti, H. Ammari, B. Jin, J.-K. Seo, and W. Zhang. The linearized inverse problem in multi-frequency electrical impedance tomography. *SIAM J. Imaging Sci.*, 9(4):1525–1551, 2016.

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

[5] S. Bartels. Error control and adaptivity for a variational model problem defined on functions of bounded variation. *Math. Comp.*, 84(293):1217–1240, 2015.

[6] L. Beilina and C. Clason. An adaptive hybrid FEM/FDM method for an inverse scattering problem in scanning acoustic microscopy. *SIAM J. Sci. Comput.*, 28(1):382–402, 2006.

[7] L. Beilina and M. V. Klibanov. A posteriori error estimates for the adaptivity technique for the tikhonov functional and global convergence for a coefficient inverse problem. *Inverse Problems*, 26(4):045012, 27pp, 2010.

[8] L. Beilina and M. V. Klibanov. Reconstruction of dielectrics from experimental data via a hybrid globally convergent/adaptive algorithm. *Inverse Problems*, 26(12):125009, 30 pp, 2010.

[9] L. Beilina, M. V. Klibanov, and M. Y. Kokurin. Adaptivity with relaxation for ill-posed problems and global convergence for a coefficient inverse problem. *J. Math. Sci.*, 167(3):279–325, 2010.

[10] A. Braides. Γ-*convergenc for Beginners*. Oxford University Press, Oxford, UK, 2002.

[11] J. Cahn and J. Hilliard. Free energy of a non-uniform system I – Interfacial free energy. *J. Chem. Phys.*, 28:258–267, 1958.

[12] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[13] Z. Chen and R. Nochetto. Residual type a posteriori error estimates for elliptic obstacle problems. *Numer. Math.*, 84(4):527–548, 2000.

[14] K.-S. Cheng, D. Isaacson, J. C. Newell, and D. G. Gisser. Electrode models for electric current computed tomography. *IEEE Trans. Biomed. Eng.*, 36(9):918–924, 1989.

[15] Y. T. Chow, K. Ito, and J. Zou. A direct sampling method for electrical impedance tomography. *Inverse Problems*, 30(9):095003, 25 pp., 2014.

[16] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM, Philadelphia, PA, 2002.

[17] C. Clason, B. Kaltenbacher, and D. Wachsmuth. Functional error estimators for the adaptive discretization of inverse problems. *Inverse Problems*, 32(10):104004, 25 pp, 2016.

[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, 2015. revised edition.

[20] T. Feng, Y. Yan, and W. Liu. Adaptive finite element methods for the identification of distributed parameters in elliptic equation. *Adv. Comput. Math.*, 29(1):27–53, 2008.

[21] M. Gehre and B. Jin. Expectation propagation for nonlinear inverse problems with an application to electrical impedance tomography. *J. Comput. Phys.*, 259:513–535, 2014.

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

[23] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, MA, 1985.

[24] B. Harrach and M. Ullrich. Monotonicity-based shape reconstruction in electrical impedance tomography. *SIAM J. Math. Anal.*, 45(6):3382–3403, 2013.

[25] P. Hild and S. Nicaise. A posteriori error estimations of residual type for Signorini's problem. *Numer. Math.*, 101(3):523–549, 2005.

[26] M. Hintermüller and A. Laurain. Electrical impedance tomography: from topology to shape. *Control & Cybernetics*, 37:913–933, 2008.

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

[28] N. Hyvönen and L. Mustonen. Generalized linearization techniques in electrical impedance tomography. *Numer. Math.*, 140(1):95–120, 2018.

[29] K. Ito and B. Jin. *Inverse Problems: Tikhonov Theory and Algorithms*. World Scientific, Singapore, 2014.

[30] B. Jin, T. Khan, and P. Maass. A reconstruction algorithm for electrical impedance tomography based on sparsity regularization. *Internat. J. Numer. Methods Engrg.*, 89(3):337–353, 2012.

[31] 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.

[32] B. Jin, Y. Xu, and J. Zou. An adaptive finite element method for electrical impedance tomography. *IMA J. Numer. Anal.*, 37(3):1520–1550, 2017.

[33] M. Klibanov, J. Li, and W. Zhang. Convexification of electrical impedance tomography with restricted dirichlet-to-neumann map data. *Inverse Problems*, 35(3):035005, 33 pp., 2019.

[34] M. V. Klibanov. Convexification of restricted Dirichlet-to-Neumann map. *J. Inverse Ill-Posed Probl.*, 25(5):669–685, 2017.

[35] K. Knudsen, M. Lassas, J. L. Mueller, and S. Siltanen. Regularized D-bar method for the inverse conductivity problem. *Inverse Probl. Imaging*, 3(4):599–624, 2009.

[36] A. Lechleiter, N. Hyvönen, and H. Hakula. The factorization method applied to the complete electrode model of impedance tomography. *SIAM J. Appl. Math.*, 68(4):1097–1121, 2008.

[37] A. Lechleiter and A. Rieder. Newton regularizations for impedance tomography: a numerical study. *Inverse Problems*, 22(6):1967–1987, 2006.

[38] J. Li, J. Xie, and J. Zou. An adaptive finite element reconstruction of distributed fluxes. *Inverse Problems*, 27(7):075009, 25pp, 2011.

[39] D. Liu, V. Kolehmainen, S. Siltanen, and A. Seppänen. A nonlinear approach to difference imaging in EIT; assessment of the robustness in the presence of modelling errors. *Inverse Problems*, 31(3):035012, 25 pp., 2015.

[40] E. Malone, G. S. dos Santos, D. Holder, and S. Arridge. Multifrequency electrical impedance tomography using spectral constraints. *IEEE Trans. Med. Imag.*, 33(2):340–350, 2014.

[41] W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Trans. Math. Software*, 15(4):326–347 (1990), 1989.

[42] L. Modica. The gradient theory of phase transitions and the minimal interface criterion. *Arch. Rational Mech. Anal.*, 98:123–142, 1987.

[43] L. Modica and S. Mortola. Un esempio di Γ^--convergenza. *Boll. Un. Mat. Ital.*, 14-B:285–299, 1977.

[44] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In R. A. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*, pages 409–542. Springer, New York, 2009.

[45] M. K. Pidcock, S. Ciulli, and S. Ispas. Singularities of mixed boundary value problems in electrical impedance tomography. *Physiol. Meas.*, 16(3):A213–A218, 1995.

[46] L. Rondi. On the regularization of the inverse conductivity problem with discontinuous conductivities. *Inverse Probl. Imaging*, 2(3):397–409, 2008.

[47] L. Rondi. Discrete approximation and regularisation for the inverse conductivity problem. *Rend. Istit. Mat. Univ. Trieste*, 48:315–352, 2016.

[48] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[49] K. G. Siebert. A convergence proof for adaptive finite elements without lower bounds. *IMA J. Num. Anal.*, 31(3):947–970, 2011.

[50] E. Somersalo, M. Cheney, and D. Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.*, 52(4):1023–1040, 1992.

[51] C. Tan, S. Lv, F. Dong, and M. Takei. Image reconstruction based on convolutional neural network for electrical resistance tomography. *IEEE Sensors J.*, 19(1):196–204, 2019.

[52] C. Traxler. An algorithm for adaptive mesh refinement in n dimensions. *Computing*, 59:115–137, 1997.

[53] R. Verfürth. *A Posteriori Estimation Techniques for Finite Element Methods*. Oxford University Press, Oxford, 2013.

[54] J. Xiao, Z. Liu, P. Zhao, Y. Li, and J. Huo. Deep learning image reconstruction simulation for electromagnetic tomography. *IEEE Sensors J.*, 18(8):3290–3298, 2018.

[55] Y. Xu and J. Zou. Convergence of an adaptive finite element method for distributed flux reconstruction. *Math. Comp.*, 84(296):2645–2663, 2015.

[56] Q. Zhang, L. Chen, and Y. Xu. A minimization method for the double-well energy functional. Preprint, arXiv:1809.01839, 2018.

[57] L. Zhou, B. Harrach, and J. K. Seo. Monotonicity-based electrical impedance tomography for lung imaging. *Inverse Problems*, 34(4):045005, 25, 2018.