An Adaptive Finite Element Method and Its Convergence for Electrical Impedance Tomography

Bangti Jin* Yifeng Xu† Jun Zou‡

Abstract

In this work we develop and analyze a novel adaptive finite element method for efficiently solving electrical impedance tomography – a severely ill-posed nonlinear inverse problem for recovering the conductivity from boundary voltage measurements. The reconstruction technique is based on Tikhonov regularization with a smoothness penalty and discretizing the forward model using continuous piecewise linear finite elements. We derive an adaptive algorithm based on a novel a posteriori error estimator in terms of the concerned state and adjoint variables and the conductivity. The convergence of the algorithm is established, in the sense that the sequence of discrete solutions contains a convergent subsequence to a solution of the optimality system for the continuous formulation. Numerical results are presented to verify the convergence and efficiency of the algorithm.

Keywords: electrical impedance tomography, a posteriori error estimator, adaptive finite element method, convergence analysis

1 Introduction

Electrical impedance tomography (EIT) is a popular diffusive imaging modality for probing internal structures of the concerned object, by recovering its electrical conductivity/permittivity distribution from voltage measurements on the boundary. One typical experimental setup is as follows. One first attaches a set of metallic electrodes to the surface of the object, then injects an input current into the object through these electrodes, which induces an electromagnetic field inside the object. Finally, one measures the induced electric voltages on the electrodes. The procedure is often repeated several times with different input currents in order to yield sufficient information on the sought-for conductivity distribution. In many applications, the physical process can be most accurately described by the complete electrode model [12, 35], but occasionally the simpler continuum model is also employed. The imaging modality has attracted considerable interest in practical applications, e.g., medical imaging, geophysical prospecting, nondestructive evaluation and pneumatic oil pipeline conveying.

A large number of reconstruction algorithms have been developed for the EIT inverse problem (see [1, 24, 25, 28, 37, 27, 20, 14] and references therein). One prominent idea underlying many existing imaging algorithms is Tikhonov regularization with either a smoothness or sparsity type (including total variation) penalty, and they have demonstrated very encouraging reconstructions with experimental data. In practice, they are numerically realized using the continuous piecewise linear finite element method, due to its flexibility in handling variable coefficients and solid theoretical underpinnings. It has been customarily adopted in practice. Despite its popularity, it was only rigorously justified very recently in [17] for the complete electrode model for both polygonal and smooth convex domains.

*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, Scientific Computing Key Laboratory of Shanghai Universities and E-Institute for Computational Science of Shanghai Universities, Shanghai Normal University, Shanghai 200234, China. (yfxuma@aliyun.com)

‡Department of Mathematics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. (zou@math.cuhk.edu.hk)

The accuracy of the complete electrode model relies essentially on nonstandard boundary conditions for capturing important characteristics of the physical experiment, notably the contact impedance effect. In the model, around the boundary of the electrodes, the boundary condition changes from Neumann to Robin type, which induces weak singularity of the forward solution around the interface [19]. The potential discontinuity of the sought-for conductivity field, enforced by low Sobolev smoothness penalty, will also induce weak solution singularities. With a quasi-uniform partition of the domain, the solution singularities are not effectively resolved and the errors around electrode edges and discontinuity interfaces are dominant, and can significantly compromise the reconstruction accuracy. This naturally motivates the use of an adaptive strategy to achieve the desired accuracy with reduced computational complexity. In this work, we shall develop a novel adaptive finite element method (AFEM) for the EIT inverse problem and analyze its convergence behavior.

The AFEM generates a sequence of nested triangulations and discrete solutions by a successive loop:

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

The key ingredient in the procedure is the module `ESTIMATE`, which consists of computing a posteriori error estimators, i.e., computable quantities from the discrete solution, the local mesh size and other given data. This has been thoroughly studied for forward problems (see e.g. the surveys [2, 36]). Over the past few decades, there are also many important works on the a posteriori error analysis of PDE-constrained optimal control problems (see [21, 22, 30, 31] for a partial list). However, the behavior of nonlinear inverse problems such as EIT is quite different from that of optimal control problems due to the ill-posed nature, the existence of the noise in the data and inherent high-degree nonlinearity.

The adaptive idea, including the AFEM, has started to attract some attention in the context of inverse problems in recent years. In [3, 4, 5, 6] the AFEM using a dual weighted residual framework was studied for parameter identification problems, and high order terms in relevant Lagrangian functionals were ignored. Feng et al. [16] proposed a residual-based estimator for state, costate (adjoint) and parameter by assuming convexity of the cost functional and high regularity on the parameter. Li et al. [29] derived rigorous a posteriori error estimators for reconstructing the distributed flux under a practical regularity assumption, in the sense that like for forward problems, the errors of the state variable, the adjoint variable and the flux are bounded from above and below by the estimators up to some multiplicative constants. In a series of interesting works [7, 8, 10], Beilina et al. adopted the AFEM for hyperbolic coefficient inverse problems; see the comprehensive monograph [9] for in-depth discussions. Kaltenbacher et al. [18, 26] described and analyzed adaptive strategies for choosing the regularization parameter in Tikhonov regularization and iterative regularization techniques, e.g., Gauss-Newton methods. Unlike the AFEM for forward problems, for which the convergence and computational complexity have been systematically studied (cf. the survey papers [11, 32]), the theoretical analysis of the AFEM for inverse problems is still in its infancy. Recently, the second and third authors established the convergence of the AFEM for recovering the flux and the Robin coefficient [39, 38].

In this paper, we shall develop a novel AFEM for the EIT and analyze its convergence. The proposed AFEM is of the standard form (1.1): it does not require a separate marking for data oscillation in the module `MARK` and the interior node property in the module `REFINE`, and hence it is easy to implement. The derivation of the a posteriori error estimators is constructive: it lends itself to a route for convergence analysis. The analysis relies on a limiting output least-squares problem defined on the closure of the adaptively generated finite element spaces, and it consists of two steps. First, the sequence of discrete minimizers is shown in Section 4.1 to contain a subsequence converging to a solution of the limiting problem, and then the limiting minimizer and related state and adjoint variables are proved in Section 4.2 to solve the necessary optimality system of the continuous Tikhonov functional.

This work is a continuation of our prior work [17] on the FEM discretization of EIT, but differs considerably in several aspects. The major effort of [17] was to justify the convergence of the finite element approximation of the Tikhonov formulation of the EIT, and no any a posteriori error estimate and adaptive method were studied, which is the main goal of the present work. The convergence analysis in [17] was provided only for *quasi-uniform* finite element meshes, and relies crucially on the $W^{1,q}(\Omega)$ ($q > 2$) regularity of the forward solution and the density of the finite element spaces V_h in $H^1(\Omega)$. The

density does not hold generally for adaptively generated finite element spaces. Hence, the analysis in [17] does not carry over to the AFEM algorithm. In this work, we shall adopt a strategy developed in [38] for recovering the Robin coefficient to overcome these technical difficulties. Nonetheless, there are major differences in the analysis due to higher degree of nonlinearity and ill-posedness of the EIT problem. In [38], the continuity of the parameter-to-state map from $L^2(\Gamma_i)$ to $L^2(\Gamma_c)$ plays a crucial role. For the EIT, only the $H^1(\Omega)$ weak continuity of the map holds (cf. Lemma 4.2), and we shall exploit the pointwise convergence of discrete minimizers and Lebesgue's dominated convergence theorem (cf. proofs of Theorems 4.1-4.3). In particular, it allows us to establish the $H^1(\Omega)$ convergence of discrete state variables (cf. proof of Theorem 4.2), and thus enables us to verify that the limiting solution also satisfies the optimality system of the continuous Tikhonov functional (Lemmas 4.5 and 4.6).

The rest of this paper is organized as follows. In Section 2, we describe the complete electrode model, regularized least-squares formulation and its optimality system. The finite element discretization is described, and a novel adaptive algorithm for the EIT is proposed in Section 3. The convergence analysis of the algorithm is given in Section 4. Some numerical results are presented in Section 5 to illustrate its convergence and efficiency. For notational simplicity, we use $\langle\cdot,\cdot\rangle$ and $(\cdot,\cdot)$ to 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 space $\mathbb{H}$ and its dual space. Throughout this work, the notation c is used for a generic constant, which may differ at each occurrence, but it is always independent of the mesh size and other quantities of interest.

2 Preliminary

We shall recall in this section the mathematical model for the EIT problem and describe the reconstruction technique based on Tikhonov regularization and its necessary optimality system.

2.1 Complete electrode model

Let Ω be an open bounded domain in $\mathbb{R}^d$ ($d = 2, 3$) with a polyhedral boundary Γ. We denote the set of electrodes by $\{e_l\}_{l=1}^L$, which are line segments/planar surfaces on Γ and disjoint from each other, i.e., $\bar{e}_i \cap \bar{e}_k = \emptyset$ if $i \neq k$. The applied current on the lth electrode e_l is denoted by I_l, and the current vector $I = (I_1, \ldots, I_L)^{\mathrm{t}}$ satisfies $\sum_{l=1}^L I_l = 0$ by the law of charge conservation. Let the space $\mathbb{R}^L_\diamond$ be the subspace of the vector space $\mathbb{R}^L$ with zero mean, i.e., $I \in \mathbb{R}^L_\diamond$. The electrode voltage $U = (U_1, \ldots, U_L)^{\mathrm{t}}$ is also normalized such that $U \in \mathbb{R}^L_\diamond$. Then the mathematical formulation of the complete electrode model (CEM) reads: given the electrical conductivity σ, positive contact impedances $\{z_l\}_{l=1}^L$ and an input current pattern $I \in \mathbb{R}^L_\diamond$, find the potential $u \in H^1(\Omega)$ and electrode voltage $U \in \mathbb{R}^L_\diamond$ such that

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

The physical motivation behind the model (2.1) is as follows. The governing equation is derived under a quasi-static low frequency assumption on the electromagnetic process [35]. The second line describes the contact impedance effect: When injecting electrical currents into the object, a highly resistive thin layer forms at the electrode-electrolyte interface (due to certain electrochemical processes), which causes potential drops across the electrode-electrolyte interface. The potential drop is described by Ohm's law, with the positive constants $\{z_l\}_{l=1}^L$ being the proportionality factors. It takes into account the fact that metallic electrodes are perfect conductors, and hence the voltage is constant on each electrode. The third line reflects the fact that the current I_l injected through the electrode e_l is completely confined to e_l itself. The nonstandard boundary conditions capture important features of the physical process, and thus the model (2.1) is capable of reproducing experimental data within the measurement precision

[12, 35]. However, the change of the boundary condition from Neumann to Robin type along the the boundary of the electrodes induces inherent singularity in the solution [19].

Due to physical constraint, the conductivity distribution σ is naturally bounded both from below and above by positive constants. Hence we introduce the following admissible set

$$\mathcal{A} = \{\lambda : \lambda \le \sigma(x) \le \lambda^{-1} \text{ a.e. } x \in \Omega\},$$

for some $\lambda \in (0, 1)$. Further, we denote by $\mathbb{H}$ the product space $H^1(\Omega) \otimes \mathbb{R}^L_\diamond$ with its norm defined 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 [35].

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 [35]: 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 fixed $\sigma \in \mathcal{A}$ and given contact impedances $\{z_l\}_{l=1}^L$ and current $I \in \Sigma^L_\diamond$, the existence and uniqueness of a solution $(u, U) \equiv (u(\sigma), U(\sigma)) \in \mathbb{H}$ to (2.2) follows from Lemma 2.1 and Lax-Milgram theorem, and further, it depends continuously on the input current pattern I [35]; see also [25, 17, 14] for the continuity of (u, U) on the conductivity σ.

2.2 Tikhonov regularization

The inverse problem is to reconstruct the conductivity σ from noisy measurements U^δ of the electrode voltage $U(\sigma^\dagger)$, corresponding to one or multiple input currents. It is severely ill-posed in the sense that small errors in the data can lead to very large deviations in the reconstructions. Hence, some sort of regularization is beneficial, and it is incorporated into imaging algorithms, either implicitly or explicitly, in order to yield physically meaningful images. One prominent idea behind numerous existing imaging algorithms is Tikhonov regularization [23], which minimizes the following functional

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

and then takes the minimizer as an approximation to the true conductivity $\sigma^\dagger$. The first term in the functional J integrates the information in the data U^δ. For notational simplicity, we consider only one dataset in the discussion, and the adaptation to multiple datasets is straightforward. The second term imposes a priori regularity assumption (smoothness) on the expected conductivity σ. The scalar $\alpha > 0$ is known as a regularization parameter, and controls the tradeoff between the two terms [23]. Problem (2.3) has at least one minimizer, and it depends continuously on the data perturbation [25].

Next we introduce the following adjoint problem for (2.1): find $(p, P) \equiv (p(\sigma), P(\sigma)) \in \mathbb{H}$ such that

$$a(\sigma, (p, P), (v, V)) = \langle U(\sigma) - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}. \tag{2.4}$$

The Gâteaux derivative of the functional $J(\sigma)$ at $\sigma \in \mathcal{A}$ in the direction μ is given by

$$J'(\sigma)[\mu] = (\alpha \nabla \sigma, \nabla \mu) - (\mu \nabla u(\sigma), \nabla p(\sigma)).$$

Then the minimizer σ^* to problem (2.3) and the respective forward solution (u^*, U^*) and the adjoint solution (p^*, P^*) satisfies the following necessary optimality system:

$$\begin{aligned} a(\sigma^*, (u^*, U^*), (v, V)) &= \langle I, V \rangle \quad \forall (v, V) \in \mathbb{H}, \\ a(\sigma^*, (p^*, P^*), (v, V)) &= \langle U^* - U^\delta, V \rangle \quad \forall (v, V) \in \mathbb{H}, \\ \alpha(\nabla \sigma^*, \nabla(\mu - \sigma^*))_{L^2(\Omega)} - ((\mu - \sigma^*) \nabla u^*, \nabla p^*)_{L^2(\Omega)} &\geq 0 \quad \forall \mu \in \mathcal{A}, \end{aligned} \tag{2.5}$$

where the variational inequality corresponds to the box constraint in the admissible set $\mathcal{A}$.

3 Adaptive finite element method

Now we describe the finite element method for discretizing problem (2.3), derive the a posteriori error estimator and develop a novel adaptive algorithm, which uses a general marking strategy and thus is easy to implement. The convergence analysis of the algorithm will be presented in Section 4.

3.1 Finite element discretization

To discretize the problem, we first triangulate the domain Ω. Let $\mathcal{T}$ be a shape regular triangulation of the polyhedral domain $\overline{\Omega}$ consisting of closed simplicial elements, with a local mesh size $h_T := |T|^{1/d}$ for each element $T \in \mathcal{T}$, which is assumed to intersect at most one electrode surface e_l. On the triangulation $\mathcal{T}$, 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 the space $P_1(T)$ consists of all linear functions on the element T. The space $V_{\mathcal{T}}$ is also used for approximating the potential u and the conductivity σ. Nonetheless, we observe that in practice, it is possible to employ different meshes for the potential and the conductivity. The use of piecewise linear finite elements is popular since the problem data, e.g., boundary conditions, have only limited regularity.

Now we can describe the finite element approximation. First, we approximate the forward map $(u(\sigma), U(\sigma)) \in \mathbb{H}$ by $(u_{\mathcal{T}}, U_{\mathcal{T}}) \equiv (u_{\mathcal{T}}(\sigma_{\mathcal{T}}), U_{\mathcal{T}}(\sigma_{\mathcal{T}})) \in \mathbb{H}_{\mathcal{T}} \equiv V_{\mathcal{T}} \otimes \mathbb{R}^L_\diamond$ defined by

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

where the (discretized) conductivity $\sigma_{\mathcal{T}}$ lies in the discrete admissible set

$$\mathcal{A}_{\mathcal{T}} = \{ \sigma_{\mathcal{T}} \in V_{\mathcal{T}} : \lambda \leq \sigma_{\mathcal{T}} \leq \lambda^{-1} \ \text{ a.e. } \Omega \} = \mathcal{A} \cap V_{\mathcal{T}}.$$

Then the discrete optimization problem reads

$$\min_{\sigma_{\mathcal{T}} \in \mathcal{A}_{\mathcal{T}}} \left\{ J_{\mathcal{T}}(\sigma_{\mathcal{T}}) = \tfrac{1}{2} \|U_{\mathcal{T}}(\sigma_{\mathcal{T}}) - U^\delta\|^2 + \tfrac{\alpha}{2} \|\nabla \sigma_{\mathcal{T}}\|^2_{L^2(\Omega)} \right\}. \tag{3.2}$$

Due to the compactness of the space $\mathcal{A}_{\mathcal{T}}$, there exists at least one minimizer $\sigma^*_{\mathcal{T}}$ to problem (3.1)-(3.2) [17]. The minimizer $\sigma^*_{\mathcal{T}}$ and the related forward solution $(u^*_{\mathcal{T}}, U^*_{\mathcal{T}}) \equiv (u^*_{\mathcal{T}}(\sigma^*_{\mathcal{T}}), U^*_{\mathcal{T}}(\sigma^*_{\mathcal{T}})) \in \mathbb{H}_{\mathcal{T}}$ and adjoint solution $(p^*_{\mathcal{T}}, P^*_{\mathcal{T}}) \equiv (p^*_{\mathcal{T}}(\sigma^*_{\mathcal{T}}), P^*_{\mathcal{T}}(\sigma^*_{\mathcal{T}})) \in \mathbb{H}_{\mathcal{T}}$ satisfies the following necessary optimality system

$$\begin{aligned} a(\sigma^*_{\mathcal{T}}, (u^*_{\mathcal{T}}, U^*_{\mathcal{T}}), (v_{\mathcal{T}}, V)) &= \langle I, V \rangle \quad \forall (v_{\mathcal{T}}, V) \in \mathbb{H}_{\mathcal{T}}, \\ a(\sigma^*_{\mathcal{T}}, (p^*_{\mathcal{T}}, P^*_{\mathcal{T}}), (v_{\mathcal{T}}, V)) &= \langle U^*_{\mathcal{T}} - U^\delta, V \rangle \quad \forall (v_{\mathcal{T}}, V) \in \mathbb{H}_{\mathcal{T}}, \\ \alpha(\nabla \sigma^*_{\mathcal{T}}, \nabla(\mu_{\mathcal{T}} - \sigma^*_{\mathcal{T}}))_{L^2(\Omega)} - ((\mu_{\mathcal{T}} - \sigma^*_{\mathcal{T}}) \nabla u^*_{\mathcal{T}}, \nabla p^*_{\mathcal{T}})_{L^2(\Omega)} &\geq 0 \quad \forall \mu_{\mathcal{T}} \in \mathcal{A}_{\mathcal{T}}, \end{aligned} \tag{3.3}$$

which is the discrete analogue of (2.5). Like in the continuous case, it is straightforward to verify that the discrete solutions $(u^*_{\mathcal{T}}, U^*_{\mathcal{T}})$ and $(p^*_{\mathcal{T}}, P^*_{\mathcal{T}})$ depend continuously on the input current pattern I, i.e.,

$$\|(u^*_{\mathcal{T}}, U^*_{\mathcal{T}})\|_{\mathbb{H},*} + \|(p^*_{\mathcal{T}}, U^*_{\mathcal{T}})\|_{\mathbb{H},*} \leq c(\|I\| + \|U^\delta\|). \tag{3.4}$$

3.2 Adaptive algorithm

Now we can present a novel AFEM for problem (2.2)-(2.3). First we introduce some notation. Let $\mathbb{T}$ be the set of all possible conforming triangulations of the domain $\overline{\Omega}$ obtained from some shape-regular initial mesh by the successive use of bisection. We call $\mathcal{T}' \in \mathbb{T}$ a refinement of $\mathcal{T} \in \mathbb{T}$ if $\mathcal{T}'$ can be obtained from $\mathcal{T}$ by a finite number of bisections. 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 on the electrode $\bar{e}_l$ and $\Gamma \setminus \cup_{l=1}^L e_l$ by $\mathcal{F}_\mathcal{T}^l$ and $\mathcal{F}_\mathcal{T}^c$, respectively. The scalar $h_F := |F|^{1/(d-1)}$ denotes the diameter of a face $F \in \mathcal{F}_\mathcal{T}$, which is associated with a fixed normal unit vector $\boldsymbol{n}_F$ in $\overline{\Omega}$ with $\boldsymbol{n}_F = \boldsymbol{n}$ on the boundary Γ. Further, we denote by D_T (respectively D_F) the union of all elements in $\mathcal{T}$ with non-empty intersection with an element $T \in \mathcal{T}$ (respectively $F \in \mathcal{F}_\mathcal{T}$).

Remark 3.1. *The family $\mathbb{T}$ is uniformly shape regular during the refinement process [32], i.e., the shape regularity of any $\mathcal{T} \in \mathbb{T}$ is uniformly bounded by a constant depending only on the initial mesh. Thus all constants only depend on the initial mesh and the given data but not on any subsequent mesh.*

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

$$
\begin{aligned}
R_{T,1}(\sigma_\mathcal{T}^*, u_\mathcal{T}^*) &= \nabla \cdot (\sigma_\mathcal{T}^* \nabla u_\mathcal{T}^*),\\
R_{T,2}(u_\mathcal{T}^*, p_\mathcal{T}^*) &= \nabla u_\mathcal{T}^* \cdot \nabla p_\mathcal{T}^*,\\
J_{F,1}(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*) &= \begin{cases} [\sigma_\mathcal{T}^* \nabla u_\mathcal{T}^* \cdot \boldsymbol{n}_F] & \text{for } F \in \mathcal{F}_\mathcal{T}^i,\\ \sigma_\mathcal{T}^* \nabla u_\mathcal{T}^* \cdot \boldsymbol{n} + (u_\mathcal{T}^* - U_{\mathcal{T},l}^*)/z_l & \text{for } F \in \mathcal{F}_\mathcal{T}^l,\\ \sigma_\mathcal{T}^* \nabla u_\mathcal{T}^* \cdot \boldsymbol{n} & \text{for } F \in \mathcal{F}_\mathcal{T}^c, \end{cases}\\
J_{F,2}(\sigma_\mathcal{T}^*) &= \begin{cases} [\alpha \nabla \sigma_\mathcal{T}^* \cdot \boldsymbol{n}_F] & \text{for } F \in \mathcal{F}_\mathcal{T}^i,\\ \alpha \nabla \sigma_\mathcal{T}^* \cdot \boldsymbol{n} & \text{for } F \in \mathcal{F}_\mathcal{T}^l \cup \mathcal{F}_\mathcal{T}^c, \end{cases}
\end{aligned}
$$

where $[\cdot]$ denotes the jumps across interior faces F. Then for any collection of elements $\mathcal{M}_\mathcal{T} \subseteq \mathcal{T}$, we introduce the following error estimator

$$
\begin{aligned}
\eta_\mathcal{T}^2(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*, \mathcal{M}_\mathcal{T}) &:= \sum_{T \in \mathcal{M}_\mathcal{T}} \eta_\mathcal{T}^2(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*, T)\\
&:= \sum_{T \in \mathcal{M}_\mathcal{T}} \eta_{\mathcal{T},1}^2(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*, T) + \eta_{\mathcal{T},2}^2(\sigma_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*, T) + \eta_{\mathcal{T},3}^2(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, p_\mathcal{T}^*, T),
\end{aligned} \tag{3.5}
$$

where the three components $\eta_{\mathcal{T},i}^2$, $i = 1, 2, 3$, are defined by

$$
\begin{aligned}
\eta_{\mathcal{T},1}^2(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*, T) &:= h_T^2 \|R_{T,1}(\sigma_\mathcal{T}^*, u_\mathcal{T}^*)\|_{L^2(T)}^2 + \sum_{F \subset \partial T} h_F \|J_{F,1}(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*)\|_{L^2(F)}^2,\\
\eta_{\mathcal{T},2}^2(\sigma_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*, T) &:= h_T^2 \|R_{T,1}(\sigma_\mathcal{T}^*, p_\mathcal{T}^*)\|_{L^2(T)}^2 + \sum_{F \subset \partial T} h_F \|J_{F,1}(\sigma_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*)\|_{L^2(F)}^2,\\
\eta_{\mathcal{T},3}^2(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, p_\mathcal{T}^*, T) &:= h_T^4 \|R_{T,2}(u_\mathcal{T}^*, p_\mathcal{T}^*)\|_{L^2(T)}^2 + \sum_{F \subset \partial T} h_F^3 \|J_{F,2}(\sigma_\mathcal{T}^*)\|_{L^2(F)}^2.
\end{aligned}
$$

We defer the derivation of the a posteriori error estimator $\eta_\mathcal{T}(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*, \mathcal{M}_\mathcal{T})$ to Section 3.3 below. The notation $\mathcal{M}_\mathcal{T}$ will be omitted whenever $\mathcal{M}_\mathcal{T} = \mathcal{T}$. Note that the estimator $\eta_\mathcal{T}$ depends only on the discrete solutions $(\sigma_\mathcal{T}^*, u_\mathcal{T}^*, U_\mathcal{T}^*, p_\mathcal{T}^*, P_\mathcal{T}^*)$ and the given problem data (e.g., impedance coefficients $\{z_l\}_{l=1}^L$), and all the quantities involved in $\eta_\mathcal{T}$ are completely computable. As it is shown in Section 4.2, this error estimator is sufficient for the convergence of the resulting adaptive algorithm.

Now we can formulate an adaptive algorithm for the EIT inverse problem, cf. Algorithm 1. Below we indicate the dependence on the triangulation $\mathcal{T}_k$ by the iteration number k in the subscript.

Algorithm 1 Adaptive finite element method for EIT

1: Specify a shape regular initial mesh $\mathcal{T}_0$, and set $k := 0$.
2: (`SOLVE`) Solve problem (3.1)-(3.2) over $\mathcal{T}_k$ for the minimizer $(\sigma_k^*, u_k^*, U_k^*) \in \mathcal{A}_k \times \mathbb{H}_k$ and the adjoint solution $(p_k^*, P_k^*) \in \mathbb{H}_k$; see (3.3).
3: (`ESTIMATE`) Compute the error estimator $\eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*)$ by (3.5).
4: (`MARK`) Mark a subset $\mathcal{M}_k \subseteq \mathcal{T}_k$ with at least one element $\widetilde{T} \in \mathcal{T}_k$ with the largest error indicator:

$$\eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*, \widetilde{T}) = \max_{T \in \mathcal{T}_k} \eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*, T). \tag{3.6}$$

5: (`REFINE`) Refine each element T in $\mathcal{M}_k$ by bisection to get $\mathcal{T}_{k+1}$.
6: Set $k = k + 1$, and return to Step 2, until a certain stopping criterion is fulfilled.

Remark 3.2. *Assumption* (3.6) *in the module* `MARK` *is fairly general, and it covers several commonly used marking strategies, e.g., maximum strategy, equidistribution, modified equidistribution strategy, and Dörfler's strategy [34, pp. 962]. Our convergence analysis in Section 4 covers all these marking strategies.*

Remark 3.3. *The solver in the module* `SOLVE` *can be either a gradient descent method or a Gauss-Newton method, each equipped with a suitable step size selection rule.*

We end this section with a geometric observation on the mesh sequence $\{\mathcal{T}_k\}$ and a stability result on error indicators $\eta_{k,1}(\sigma_k^*, u_k^*, U_k^*)$, $\eta_{k,2}(\sigma_k^*, p_k^*, P_k^*)$ and $\eta_{k,3}(\sigma_k^*, u_k^*, p_k^*)$ given in Algorithm 1. For the purpose, we introduce

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

That is, the set $\mathcal{T}_k^+$ consists of all elements not refined after the k-th iteration while all elements in $\mathcal{T}_k^0$ 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 by $h_k(x) = h_T$ for x in the interior of an element $T \in \mathcal{T}_k$ and $h_k(x) = h_F$ for x in the relative interior of an edge $F \in \mathcal{F}_k$. It has the following important property in the region of Ω involving marked elements (see [34, Corollary 3.3] for the proof):

Lemma 3.1. *Let χ_k^0 be the characteristic function of Ω_k^0. Then*

$$\lim_{k \to \infty} \|h_k \chi_k^0\|_{L^\infty(\Omega)} = 0.$$

The next result gives preliminary bounds on the a posteriori error estimators.

Lemma 3.2. *Let the sequence of discrete solutions $\{(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*)\}$ be generated by Algorithm 1. Then for each $T \in \mathcal{T}_k$ with its face F, there hold that*

$$\begin{aligned}
\eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, T) &\leq c(\|\nabla u_k^*\|_{L^2(D_T)}^2 + h_F \|u_k^* - U_{k,l}^*\|_{L^2(F \cap e_l)}^2), \\
\eta_{k,2}^2(\sigma_k^*, p_k^*, P_k^*, T) &\leq c(\|\nabla p_k^*\|_{L^2(D_T)}^2 + h_F \|p_k^* - P_{k,l}^*\|_{L^2(F \cap e_l)}^2), \\
\eta_{k,3}^2(\sigma_k^*, u_k^*, p_k^*, T) &\leq c(h_T^{4-d} \|\nabla u_k^*\|_{L^2(T)}^2 \|\nabla p_k^*\|_{L^2(T)}^2 + h_T^2 \|\nabla \sigma_k^*\|_{L^2(D_T)}^2).
\end{aligned}$$

Proof. We only prove the third estimate, and the first two follow analogously. By the inverse estimates and the trace theorem, the local quasi-uniformity of $\mathcal{T}_k$ yields

$$h_T^4 \|\nabla u_k^* \cdot \nabla p_k^*\|_{L^2(T)}^2 \leq c h_T^{4-d} \|\nabla u_k^* \cdot \nabla p_k^*\|_{L^1(T)}^2 \leq c h_T^{4-d} \|\nabla u_k^*\|_{L^2(T)}^2 \|\nabla p_k^*\|_{L^2(T)}^2,$$

$$\sum_{F \subset \partial T} h_F^3 \|J_{F,2}(\sigma_k^*)\|_{L^2(F)}^2 \leq c h_T^2 \|\nabla \sigma_k^*\|_{L^2(D_T)}^2.$$

□

3.3 Derivation of a posteriori error estimators

Now we motivate the estimator $\eta_{\mathcal{T}}$ defined in (3.5) which serves as an error indicator in the module `ESTIMATE` of Algorithm 1. The algorithm generates a sequence of discrete solutions $\{(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*)\}$ in a sequence of finite element spaces $\{V_k\}$ and discrete admissible sets $\{\mathcal{A}_k\}$ over a sequence of meshes $\{\mathcal{T}_k\}$. Since the discrete forward problem (3.1) needs to be solved repeatedly in Algorithm 1, some arguments in the a posteriori error estimation for direct problems will naturally be employed. We shall need the following error estimates on the Lagrange interpolation operator $I_k : H^2(\Omega) \to V_k$ [13] and the Scott-Zhang interpolation operator $I_k^{sz} : H^1(\Omega) \to V_k$ [33] over the triangulation $\mathcal{T}_k$.

Lemma 3.3. *For any $T \in \mathcal{T}_k$ and any $F \in \mathcal{F}_k$,*

$$\|v - I_k v\|_{L^2(T)} \le c h_T^2 \|v\|_{H^2(T)}, \quad \|v - I_k v\|_{L^2(F)} \le c h_T^{3/2} \|v\|_{H^2(\omega_F)},$$
$$\|v - I_k^{sz} v\|_{L^2(T)} \le c h_T \|v\|_{H^1(D_T)}, \quad \|v - I_k^{sz} v\|_{L^2(F)} \le c h_T^{1/2} \|v\|_{H^1(D_F)},$$

where ω_F is the union of elements with F as a face.

To motivate the error estimator $\eta_{\mathcal{T}}$, we begin with two auxiliary problems: find $(\widetilde{u}(\sigma_k^*), \widetilde{U}(\sigma_k^*)) \in \mathbb{H}$ and $(\widetilde{p}(\sigma_k^*), \widetilde{P}(\sigma_k^*)) \in \mathbb{H}$ such that

$$a(\sigma_k^*, (\widetilde{u}, \widetilde{U}), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}, \tag{3.7}$$

$$a(\sigma_k^*, (\widetilde{p}, \widetilde{P}), (v, V)) = \langle \widetilde{U} - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}. \tag{3.8}$$

Note that the first equation in (3.3) is actually the finite element scheme of the auxiliary problem (3.7) over $\mathcal{T}_k$. Hence, the standard a posteriori error analysis for forward problems can be applied. In particular, by setting $v_k = I_k^{sz} v \in V_k$ in the first equation in (3.3) for any $(v, V) \in \mathbb{H}$, applying elementwise integration by parts and appealing to Lemma 3.3, there hold

$$\begin{aligned} a(\sigma_k^*, (\widetilde{u} - u_k^*, \widetilde{U} - U_k^*), (v, V)) &= \langle I, V\rangle - (\sigma_k^* \nabla u_k^*, \nabla v) - \sum_{l=1}^{L} z_l^{-1} (u_k^* - U_{k,l}^*, v - V_l)_{L^2(e_l)} \\ &= (\sigma_k^* \nabla u_k^*, \nabla (I_k^{sz} v - v)) + \sum_{l=1}^{L} z_l^{-1} (u_k^* - U_{k,l}^*, I_k^{sz} v - v)_{L^2(e_l)} \\ &\le c \Big(\sum_{T \in \mathcal{T}_k} \eta_{k,1}^2 (\sigma_k^*, u_k^*, U_k^*, T) \Big)^{1/2} \|v\|_{H^1(\Omega)}. \end{aligned}$$

Taking $(v, V) = (\widetilde{u} - u_k^*, \widetilde{U} - U_k^*) \in \mathbb{H}$ and using Lemma 2.1 yield

$$\|(\widetilde{u} - u_k^*, \widetilde{U} - U_k^*)\|_{\mathbb{H},*} \le c \Big(\sum_{T \in \mathcal{T}_k} \eta_{k,1}^2 (\sigma_k^*, u_k^*, U_k^*, T) \Big)^{1/2}. \tag{3.9}$$

Further, from the first equation in (2.5) and (3.7) we find for any $(v, V) \in \mathbb{H}$

$$a(\sigma_k^*, (u^* - \widetilde{u}, U^* - \widetilde{U}), (v, V)) = ((\sigma_k^* - \sigma^*) \nabla u^*, \nabla v) \le \|(\sigma^* - \sigma_k^*) \nabla u^*\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}.$$

Consequently,

$$\|(u^* - \widetilde{u}, U^* - \widetilde{U})\|_{\mathbb{H},*} \le c \|(\sigma^* - \sigma_k^*) \nabla u^*\|_{L^2(\Omega)}. \tag{3.10}$$

Likewise, for $(p^* - p_k^*, P^* - P_k^*)$, we appeal to the second equation in the discrete optimality system (3.3)

and the auxiliary problem (3.8) to deduce

$$\begin{aligned}
a(\sigma_k^*,(\widetilde{p}-p_k^*,\widetilde{P}-P_k^*),(v,V)) &= \langle \widetilde{U}-U^\delta,V\rangle - a(\sigma_k^*,(p_k^*,P_k^*),(v,V))\\
&= \langle \widetilde{U}-U_k^*,V\rangle + \langle U_k^*-U^\delta,V\rangle - a(\sigma_k^*,(p_k^*,P_k^*),(v,V))\\
&= \langle \widetilde{U}-U_k^*,V\rangle + (\sigma_k^*\nabla p_k^*,\nabla(I_k^{sz}v-v)) + \sum_{l=1}^{L} z_l^{-1}(p_k^*-P_{k,l}^*,I_k^{sz}v-v)_{L^2(e_l)}\\
&\le c\Big(\big(\sum_{T\in\mathcal{T}_k}\eta_{k,2}^2(\sigma_k^*,p_k^*,P_k^*,T)\big)^{1/2} + \|\widetilde{U}-U_k^*\|\Big)\|(v,V)\|_{\mathbb{H},*},
\end{aligned}$$

and

$$\begin{aligned}
a(\sigma_k^*,(p^*-\widetilde{p},P^*-\widetilde{P}),(v,V)) &= ((\sigma_k^*-\sigma^*)\nabla p^*,\nabla v) + \langle U^*-\widetilde{U},V\rangle\\
&\le \Big(\|(\sigma^*-\sigma_k^*)\nabla p^*\|_{L^2(\Omega)} + \|U^*-\widetilde{U}\|\Big)\,\|(v,V)\|_{\mathbb{H},*},
\end{aligned}$$

which, together with (3.9) and (3.10) and Lemma 2.1, implies

$$\begin{aligned}
\|(p^*-p_k^*,P^*-P_k^*)\|_{\mathbb{H},*} \le c\Big(\big(\sum_{T\in\mathcal{T}_k}\eta_{k,1}^2(\sigma_k^*,u_k^*,U_k^*,T)+\eta_{k,2}^2(\sigma_k^*,p_k^*,P_k^*,T)\big)^{1/2}\\
+\|(\sigma^*-\sigma_k^*)\nabla u^*\|_{L^2(\Omega)} + \|(\sigma^*-\sigma_k^*)\nabla p^*\|_{L^2(\Omega)}\Big).
\end{aligned} \tag{3.11}$$

In view of (3.9)-(3.11), $\eta_{k,1}$ and $\eta_{k,2}$ can bound $(u^*-u_k^*,U^*-U_k^*)$ and $(p^*-p_k^*,P^*-P_k^*)$ from above up to the terms $\|(\sigma^*-\sigma_k^*)\nabla u^*\|_{L^2(\Omega)}$ and $\|(\sigma^*-\sigma_k^*)\nabla p^*\|_{L^2(\Omega)}$. This motivates our choice of a computable upper bound for $\sigma^*-\sigma_k^*$, upon discarding the uncomputable terms. Upon appealing to the variational inequalities in (2.5) and (3.3), and $I_k\mu\in\mathcal{A}_k$ for any $\mu\in\mathcal{A}\cap C^\infty(\overline{\Omega})$, we deduce

$$\begin{aligned}
\alpha\|\nabla(\sigma^*-\sigma_k^*)\|_{L^2(\Omega)}^2 &\le \alpha(\nabla\sigma_k^*,\nabla(\sigma_k^*-\sigma^*)) - ((\sigma_k^*-\sigma^*)\nabla u^*,\nabla p^*)\\
&= \alpha(\nabla\sigma_k^*,\nabla(\sigma_k^*-\sigma^*)) - ((\sigma_k^*-\sigma^*)\nabla u_k^*,\nabla p_k^*)\\
&\quad + (\nabla u_k^*\cdot\nabla p_k^* - \nabla u^*\cdot\nabla p^*,\sigma_k^*-\sigma^*)\\
&\le \alpha(\nabla\sigma_k^*,\nabla(I_k\mu-\sigma^*)) - ((I_k\mu-\sigma^*)\nabla u_k^*,\nabla p_k^*)\\
&\quad + (\nabla u_k^*\cdot\nabla p_k^* - \nabla u^*\cdot\nabla p^*,\sigma_k^*-\sigma^*)\\
&= \alpha(\nabla\sigma_k^*,\nabla(I_k\mu-\mu)) - (\nabla u_k^*,\nabla p_k^*(I_k\mu-\mu))\\
&\quad + (\nabla u_k^*\cdot\nabla p_k^* - \nabla u^*\cdot\nabla p^*,\sigma_k^*-\sigma^*)\\
&\quad + \alpha(\nabla\sigma_k^*,\nabla(\mu-\sigma^*)) - ((\mu-\sigma^*)\nabla u_k^*,\nabla p_k^*) := \mathrm{I}+\mathrm{II}+\mathrm{III}.
\end{aligned}$$

Now Lemma 3.3 and elementwise integration by parts yield

$$|\mathrm{I}| \le c(\sum_{T\in\mathcal{T}_k}\eta_{k,3}^2(\sigma_k^*,u_k^*,p_k^*,T))^{1/2}\|\mu\|_{H^2(\Omega)} \quad \forall\mu\in\mathcal{A}\cap C^\infty(\overline{\Omega}). \tag{3.12}$$

By the minimizing property of $\{\sigma_k^*\}$ for $J_k(\cdot)$, $\|\nabla\sigma_k^*\|_{L^2(\Omega)}$ is bounded. Then the estimate (3.4) and the density of $\mathcal{A}\cap C^\infty(\overline{\Omega})$ in $\mathcal{A}$ ensure that III can be made arbitrarily small. For the term II, we have

$$\begin{aligned}
|\mathrm{II}| &= |(\nabla u_k^*\cdot\nabla p_k^* - \nabla u_k^*\cdot\nabla p^* + \nabla u_k^*\cdot\nabla p^* - \nabla u^*\cdot\nabla p^*,\sigma^k-\sigma^*)|\\
&\le \|\nabla(p_k^*-p^*)\|_{L^2(\Omega)}\|(\sigma_k^*-\sigma^*)\nabla u_k^*\|_{L^2(\Omega)} + \|\nabla(u_k^*-u^*)\|_{L^2(\Omega)}\|(\sigma_k^*-\sigma^*)\nabla p^*\|_{L^2(\Omega)},
\end{aligned}$$

which is expected to be higher order terms. Upon discarding the terms $\|(\sigma^*-\sigma_k^*)\nabla u^*\|_{L^2(\Omega)}$ and $\|(\sigma^*-\sigma_k^*)\nabla p^*\|_{L^2(\Omega)}$ in (3.10)-(3.11) and the nonlinear term II, we get all computable quantities in (3.9), (3.11) and (3.12), which are exactly the a posteriori error estimator η_k defined in (3.5). Thus we may view it as a reliable upper bound for the norm error and employ it in the module `ESTIMATE` to drive the adaptive refinement process. Moreover the derivation of (3.12) suggests a natural way to deal with the variational inequality in (2.5) in the convergence analysis of Algorithm 1.

4 Convergence analysis

In this section, we shall establish the main theoretical result of this work, the convergence of Algorithm 1, namely the sequence of discrete solutions $\{(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*)\}$ to (3.3) generated by Algorithm 1, contains a subsequence converging in $H^1(\Omega) \times \mathbb{H} \times \mathbb{H}$ to a solution to the optimality system (2.5). The main technical difficulty lies in the lack of density of the adaptively generated finite element space V_k in the space $H^1(\Omega)$. To overcome the challenge, the proof is carried out in two steps. In the first step (Section 4.1), we analyze a "limiting" optimization problem posed over a limiting set induced by $\{\mathcal{A}_k\}$, and show that the sequence of discrete minimizers contains a convergent subsequence to a minimizer to the limiting problem. In the second step (Section 4.2), we show that the solution to the optimality system for the limiting problem actually solves the optimality system (2.5). It is worth noting that all the proofs in Section 4.1 only depends on the nestedness of finite element spaces and discrete admissible sets, and the error estimator (3.5) and the marking assumption (3.6) are used only in Section 4.2.

4.1 Limiting optimization problem

For the sequences $\{\mathbb{H}_k\}$ and $\{\mathcal{A}_k\}$ generated by Algorithm 1, we define a limiting finite element space $\mathbb{H}_\infty$ and a limiting admissible set $\mathcal{A}_\infty$ respectively by

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

It is easy to see that $\mathbb{H}_\infty$ is a closed subspace of $\mathbb{H}$. For the set $\mathcal{A}_\infty$, we have the following lemma.

Lemma 4.1. *$\mathcal{A}_\infty$ is a closed convex subset of $\mathcal{A}$.*

Proof. The definition of $\mathcal{A}_\infty$ implies its strong closedness. For any μ and ν in $\mathcal{A}_\infty$, there exist two sequences $\{\mu_k\}$ and $\{\nu_k\} \subset \bigcup_{k\geq 0} \mathcal{A}_k$ such that $\mu_k \to \mu$ and $\nu_k \to \nu$ in $H^1(\Omega)$. The convexity of the set $\mathcal{A}_k$ implies $\{t\mu_k + (1-t)\nu_k\} \subset \bigcup_{k\geq 0} \mathcal{A}_k$ for any $t \in (0,1)$. Then $t\mu_k + (1-t)\nu_k \to t\mu + (1-t)\nu$ in $H^1(\Omega)$, i.e. $t\mu + (1-t)\nu \in \mathcal{A}_\infty$ for any $t \in (0,1)$. Hence $\mathcal{A}_\infty$ is convex. Moreover, we have $\mu_k \to \mu$ a.e. in Ω after (possibly) passing to a subsequence, which, along with the constraint $\lambda \leq \mu_k \leq \lambda^{-1}$ a.e. in Ω, indicates that $\lambda \leq \mu \leq \lambda^{-1}$ a.e. in Ω. Lastly, the fact that $\mathcal{A}_\infty \subset H^1(\Omega)$ concludes $\mathcal{A}_\infty \subset \mathcal{A}$. □

Over the limiting set $\mathcal{A}_\infty$, we introduce a limiting minimization problem:

$$\min_{\sigma_\infty \in \mathcal{A}_\infty} \left\{ J_\infty(\sigma_\infty) = \tfrac{1}{2}\|U_\infty(\sigma_\infty) - U^\delta\|^2 + \tfrac{\alpha}{2}\|\nabla \sigma_\infty\|^2_{L^2(\Omega)} \right\}, \tag{4.1}$$

where $(u_\infty, U_\infty) \equiv (u_\infty(\sigma_\infty), U_\infty(\sigma_\infty)) \in \mathbb{H}_\infty$ satisfies the variational problem:

$$a(\sigma_\infty, (u_\infty, U_\infty), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}_\infty. \tag{4.2}$$

The next result shows the existence of a minimizer to the limiting problem (4.1)-(4.2).

Theorem 4.1. *There exists at least one minimizer to problem* (4.1)-(4.2).

Proof. It is clear that $\inf J_\infty(\sigma)$ is finite over $\mathcal{A}_\infty$, so there exists a minimizing sequence $\{\sigma^n\} \subset \mathcal{A}_\infty$, i.e., $J_\infty(\sigma^n) \to \inf_{\sigma\in\mathcal{A}_\infty} J_\infty(\sigma)$. Thus, the sequence $\{\sigma^n\}$ is uniformly bounded in $H^1(\Omega)$, and by Sobolev embedding theorem [15] and the fact that $\mathcal{A}_\infty$ is a closed convex subset of $H^1(\Omega)$, cf. Lemma 4.1, there exists a subsequence, still denoted by $\{\sigma^n\}$, and some $\sigma^* \in \mathcal{A}_\infty$ such that $\sigma^n \to \sigma^*$ weakly in $H^1(\Omega)$, $\sigma^n \to \sigma^*$ a.e. in Ω. By taking $\sigma_\infty = \sigma^n \in \mathcal{A}_\infty$ in (4.2), then $(u^n, U^n) \equiv (u^n(\sigma^n), U^n(\sigma^n))$ satisfies

$$a(\sigma^n, (u^n, U^n), (v, V)) = \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}_\infty. \tag{4.3}$$

Then by Lemma 2.1, $\{(u^n, U^n)\}$ is uniformly bounded in $\mathbb{H}$, which upon passing to a subsequence yields that there exists a subsequence, also denoted by $\{(u^n, U^n)\}$, and some $(u^*, U^*) \in \mathbb{H}_\infty$ such that

$$(u^n, U^n) \to (u^*, U^*) \quad \text{weakly in } \mathbb{H} \quad \text{and} \quad u^n \to u^* \quad \text{in } L^2(\Gamma). \tag{4.4}$$

We claim that $(u^*, U^*) = (u^*(\sigma^*), U^*(\sigma^*)) \in \mathbb{H}_\infty$. To this end, first we observe the splitting

$$(\sigma^n \nabla u^n, \nabla v) = ((\sigma^n - \sigma^*)\nabla u^n, \nabla v) + (\sigma^* \nabla u^n, \nabla v).$$

The pointwise convergence of the sequence $\{\sigma^n\}$, Lebesgue's dominated convergence theorem [15] and uniform boundedness of $\{u^n\}$ in $H^1(\Omega)$ imply that

$$|((\sigma^n - \sigma^*)\nabla u^n, \nabla v)| \le \|\nabla u^n\|_{L^2(\Omega)} \|(\sigma^n - \sigma^*)\nabla v\|_{L^2(\Omega)} \to 0.$$

This and the weak convergence of $\{u^n\}$ in $H^1(\Omega)$ give $(\sigma^n \nabla u^n, \nabla v)_{L^2(\Omega)} \to (\sigma^* \nabla u^*, \nabla v)_{L^2(\Omega)}$. Then by (4.4), we obtain

$$(u^n - U_l^n, v - V_l)_{L^2(e_l)} \to (u^* - U_l^*, v - V_l)_{L^2(e_l)}.$$

Upon taking into account these relations, we deduce

$$a(\sigma^*, (u^*, U^*), (v, V)) = \langle I, V \rangle \quad \forall (v, V) \in \mathbb{H}_\infty,$$

i.e., the desired claim $(u^*, U^*) = (u^*(\sigma^*), U^*(\sigma^*)) \in \mathbb{H}_\infty$. This and the weak lower semicontinuity of the norm imply that σ^* is a minimizer of $J_\infty(\cdot)$ over $\mathcal{A}_\infty$. □

As a corollary of the above proof, we have the following weak continuity result.

Lemma 4.2. *Let the sequence $\{\sigma_k\} \subset \bigcup_{k\ge 0} \mathcal{A}_k$ converge to some $\sigma^* \in \mathcal{A}_\infty$ weakly in $H^1(\Omega)$ and a.e. in Ω. Then there exist a subsequence $\{\sigma_{k_m}\}$ and the solution $(u(\sigma^*), U(\sigma^*)) \in \mathbb{H}_\infty$ to (4.2) such that the sequence of solutions $\{(u_{k_m}(\sigma_{k_m}), U_{k_m}(\sigma_{k_m}))\} \subset \bigcup_{k\ge 0} \mathbb{H}_k$ to (3.1) over $\mathcal{T}_{k_m}$ satisfies*

$$(u_{k_m}(\sigma_{k_m}), U_{k_m}(\sigma_{k_m})) \to (u(\sigma^*), U(\sigma^*)) \quad \textit{weakly in } \mathbb{H}.$$

Now we analyze the limiting behavior of the sequence $\{(\sigma_k^*, u_k^*, U_k^*)\}$ generated by Algorithm 1: It contains a subsequence converging in $H^1(\Omega) \times \mathbb{H}$ to a minimizer of the limiting problem (4.1)-(4.2). This result will play a crucial role in the convergence analysis in Section 4.2.

Theorem 4.2. *Let $\{\mathcal{A}_k \times \mathbb{H}_k\}$ be a sequence of discrete admissible sets and finite element spaces generated by Algorithm 1. Then the sequence of discrete solutions $\{(\sigma_k^*, u_k^*, U_k^*)\}$ to problem (3.2) has a subsequence $\{(\sigma_{k_m}^*, u_{k_m}^*, U_{k_m}^*)\}$ converging to a minimizer $(\sigma_\infty^*, u_\infty^*, U_\infty^*)$ to problem (4.1)-(4.2) in the sense that*

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

Proof. Since the constant function $1 \in \mathcal{A}_k$ for all k and $J_k(\sigma_k^*)$ attains its minimum at $\sigma_k^* \in \mathcal{A}_k$, the sequence $\{\sigma_k^*\}$ is uniformly bounded in $H^1(\Omega)$. Hence, by Sobolev embedding theorem [15], there exists a subsequence $\{\sigma_{k_m}^*\}$ and some $\sigma_\infty^* \in \mathcal{A}_\infty$ such that $\sigma_{k_m}^* \to \sigma_\infty^*$ weakly in $H^1(\Omega)$, $\sigma_{k_m}^* \to \sigma_\infty^*$ a.e. in Ω. By Lemma 4.2, there exists a subsubsequence of $\{(u_{k_m}^*, U_{k_m}^*)\}$, still denoted by $\{(u_{k_m}^*, U_{k_m}^*)\}$, such that

$$(u_{k_m}^*, U_{k_m}^*) \to (u_\infty^*(\sigma_\infty^*), U_\infty^*(\sigma_\infty^*)) \quad \text{weakly in } \mathbb{H},$$

where $(u_\infty^*(\sigma_\infty^*), U_\infty^*(\sigma_\infty^*))$ satisfies problem (4.2) with $\sigma_\infty = \sigma_\infty^*$. First we prove that σ_∞^* is a minimizer to J_∞ over $\mathcal{A}_\infty$. For any $\sigma \in \mathcal{A}_\infty$, the definition of $\mathcal{A}_\infty$ ensures the existence of a sequence $\{\sigma_l\} \subset \bigcup_{k\ge 0} \mathcal{A}_k$ such that $\sigma_l \to \sigma$ in $H^1(\Omega)$. By Lemma 4.2, there exists a subsequence $\{\sigma_{l_n}\}$ such that the solution $(u_{l_n}(\sigma_{l_n}), U_{l_n}(\sigma_{l_n}))$ to problem (3.1) over $\mathcal{T}_{l_n}$ and $(u_\infty(\sigma), U_\infty(\sigma)) \in \mathbb{H}$ satisfies

$$(u_{l_n}(\sigma_{l_n}), U_{l_n}(\sigma_{l_n})) \to (u_\infty(\sigma), U_\infty(\sigma)) \quad \text{weakly in } \mathbb{H}.$$

By the minimizing property of $\sigma_{k_m}^*$ to the functional J_{k_m} over $\mathcal{A}_{k_m}$, for $k_m \ge l_n$ (and large l_n) and the nestedness of the space $\mathcal{A}_{k_m} \supset \mathcal{A}_{l_n}$, there holds $J_{k_m}(\sigma_{k_m}^*) \le J_{k_m}(\sigma_{l_n}) = J_{l_n}(\sigma_{l_n})$. Consequently,

$$J_\infty(\sigma_\infty^*) \le \liminf_{m\to\infty} J_{k_m}(\sigma_{k_m}^*) \le \limsup_{n\to\infty} J_{l_n}(\sigma_{l_n}) = J_\infty(\sigma) \quad \forall \sigma \in \mathcal{A}_\infty.$$

Further, by taking $\sigma = \sigma_\infty^*$, we derive $\lim_{m\to\infty} J_{k_m}(\sigma_{k_m}^*) = J_\infty(\sigma_\infty^*)$, and thus $\lim_{m\to\infty} \|\nabla\sigma_{k_m}^*\|_{L^2(\Omega)}^2 = \|\nabla\sigma_\infty^*\|_{L^2(\Omega)}^2$. This and the weak convergence of $\sigma_{k_m}^*$ in $H^1(\Omega)$ shows the first assertion.

It remains to show the convergence of $\{u_{k_m}^*\}$ in $H^1(\Omega)$, which in view of its weak convergence in $H^1(\Omega)$ follows directly from the identity $\|\nabla(u_{k_m}^* - u_\infty^*)\|_{L^2(\Omega)} \to 0$. Using the discrete problem (3.1) over $\mathcal{T}_{k_m}$, the convergence of $\{U_{k_m}^*\}$ and the limiting problem (4.2) imply

$$a(\sigma_{k_m}, (u_{k_m}^*, U_{k_m}^*), (u_{k_m}^*, U_{k_m}^*)) = \langle I, U_{k_m}^*\rangle \to \langle I, U_\infty^*\rangle = a(\sigma_\infty, (u_\infty^*, U_\infty^*), (u_\infty^*, U_\infty^*)),$$

By the compact embedding from $H^1(\Omega)$ into $L^2(\Gamma)$ [15], the sequence $\{u_{k_m}^*\}$ converges to u_∞^* in $L^2(\Gamma)$, and the convergence of $\{U_{k_m}^*\}$ yield $(\sigma_{k_m}^*\nabla u_{k_m}^*, \nabla u_{k_m}^*) \to (\sigma_\infty^*\nabla u_\infty^*, \nabla u_\infty^*)$. Now it follows from the elementary identity

$$\|\sqrt{\sigma_{k_m}^*}\nabla(u_{k_m}^* - u_\infty^*)\|_{L^2(\Omega)}^2 = \|\sqrt{\sigma_{k_m}^*}\nabla u_{k_m}^*\|_{L^2(\Omega)}^2 - 2(\sigma_{k_m}^*\nabla u_{k_m}^*, \nabla u_\infty^*) + \|\sqrt{\sigma_{k_m}^*}\nabla u_\infty^*\|_{L^2(\Omega)}^2$$

and the triangle inequality that

$$\begin{aligned}\|\nabla(u_{k_m}^* - u_\infty^*)\|_{L^2(\Omega)}^2 \le c(&|(\sigma_{k_m}^*\nabla u_{k_m}^*, \nabla u_{k_m}^*) - (\sigma_\infty^*\nabla u_\infty^*, \nabla u_\infty^*)| + |(\sigma_{k_m}^* - \sigma_\infty^*, |\nabla u_\infty^*|^2)| \\ &+ |(\sigma_{k_m}^*\nabla u_{k_m}^* - \sigma_\infty^*\nabla u_\infty^*, \nabla u_\infty^*)|) := \mathrm{I} + \mathrm{II} + \mathrm{III}\end{aligned}$$

The second term II tends to zero by the pointwise convergence of the sequence $\{\sigma_{k_m}^*\}$ and Lebesgue's dominated convergence theorem [15]. For the third term III, there holds

$$\begin{aligned}\mathrm{III} &\le |((\sigma_{k_m}^* - \sigma_\infty^*)\nabla u_{k_m}^*, \nabla u_\infty^*)| + |(\sigma_\infty^*\nabla(u_{k_m}^* - u_\infty^*), \nabla u_\infty^*)| \\ &\le \|\nabla u_{k_m}^*\|_{L^2(\Omega)}\|(\sigma_{k_m}^* - \sigma_\infty^*)\nabla u_\infty^*\|_{L^2(\Omega)} + |(\sigma_\infty^*\nabla(u_{k_m}^* - u_\infty^*), \nabla u_\infty^*)| \to 0\end{aligned}$$

by the weak convergence of $\{u_{k_m}^*\}$ in $H^1(\Omega)$ and the pointwise convergence of $\{\sigma_{k_m}^*\}$. The preceding three estimates together complete the proof of the theorem. □

Next we turn to the optimality system of problem (4.1). Like in the continuous case, the optimality condition for the minimizer $(\sigma_\infty^*, u_\infty^*, U_\infty^*)$ and the adjoint solution $(p_\infty^*, P_\infty^*) \in \mathbb{H}_\infty$ is given by

$$\begin{aligned} a(\sigma_\infty^*, (u_\infty^*, U_\infty^*), (v, V)) &= \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H}_\infty, \\ a(\sigma_\infty^*, (p_\infty^*, P_\infty^*), (v, V)) &= \langle U_\infty^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}_\infty, \\ \alpha(\nabla\sigma_\infty^*, \nabla(\mu - \sigma_\infty^*)) - (\nabla u_\infty^*, \nabla p_\infty^*(\mu - \sigma_\infty^*)) &\ge 0 \quad \forall \mu \in \mathcal{A}_\infty. \end{aligned} \tag{4.5}$$

The next result shows the convergence of the sequence of adjoint solutions.

Theorem 4.3. *Under the condition of Theorem 4.2, the subsequence of adjoint solutions $\{(p_{k_m}^*, P_{k_m}^*)\}$ generated by Algorithm 1 converges to the solution (p_∞^*, P_∞^*) to the limiting adjoint problem in* (4.5)*:*

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

Proof. The discrete version of the limiting adjoint problem (4.5) reads: find $(\widetilde{p}_{k_m}, \widetilde{P}_{k_m}) \in \mathbb{H}_{k_m}$ such that

$$a(\sigma_\infty^*, (\widetilde{p}_{k_m}, \widetilde{P}_{k_m}), (v, V)) = \langle U_\infty^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}_{k_m}. \tag{4.6}$$

By Cea's lemma and the construction of the space $\mathbb{H}_\infty$, we deduce

$$\|(p_\infty^* - \widetilde{p}_{k_m}, P_\infty^* - \widetilde{P}_{k_m})\|_{\mathbb{H},*} \le c \inf_{(v,V)\in\mathbb{H}_{k_m}} \|(p_\infty^* - v, P_\infty^* - V)\|_{\mathbb{H},*} \to 0. \tag{4.7}$$

By taking $(v_{k_m}, V_{k_m}) = (\widetilde{p}_{k_m} - p_{k_m}^*, \widetilde{P}_{k_m} - P_{k_m}^*)$ in the second equation of (3.3) and $(v, V) = (\widetilde{p}_{k_m} - p_{k_m}^*, \widetilde{P}_{k_m} - P_{k_m}^*)$ in (4.6), we obtain

$$\begin{aligned}&\|\sqrt{\sigma_{k_m}}\nabla(\widetilde{p}_{k_m} - p_{k_m}^*)\|_{L^2(\Omega)}^2 + \sum_{l=1}^{L} z_l^{-1}\|\widetilde{p}_{k_m} - p_{k_m}^* - \widetilde{P}_{k_m,l} + P_{k_m,l}^*\|_{L^2(e_l)}^2 \\ &\quad= \langle U_\infty^* - U_{k_m}^*, \widetilde{P}_{k_m} - P_{k_m}^*\rangle + ((\sigma_{k_m}^* - \sigma_\infty^*)\nabla(\widetilde{p}_{k_m} - p_\infty^*), \nabla(\widetilde{p}_{k_m} - p_{k_m}^*)) \\ &\qquad+ ((\sigma_{k_m}^* - \sigma_\infty^*)\nabla p_\infty^*, \nabla(\widetilde{p}_{k_m} - p_{k_m}^*)) := \mathrm{I} + \mathrm{II} + \mathrm{III}.\end{aligned}$$

The Cauchy-Schwarz inequality and the box constraints on $\sigma^*_{k_m}$ and σ^*_∞ give

$$\begin{aligned}
|\mathrm{I}| &\le \|U^*_\infty - U^*_{k_m}\|_{\mathbb{R}^L} \|\widetilde{P}_{k_m} - P^*_{k_m}\|_{\mathbb{R}^L},\\
|\mathrm{II}| &\le c\|\nabla(\widetilde{p}_{k_m} - p^*_\infty)\|_{L^2(\Omega)} \|\nabla(\widetilde{p}_{k_m} - p^*_{k_m})\|_{L^2(\Omega)},\\
|\mathrm{III}| &\le \|(\sigma^*_{k_m} - \sigma^*_\infty)\nabla p^*_\infty\|_{L^2(\Omega)} \|\nabla(\widetilde{p}_{k_m} - p^*_{k_m})\|_{L^2(\Omega)},
\end{aligned}$$

which, together with Lemma 2.1, implies

$$\|(\widetilde{p}_{k_m} - p^*_{k_m}, \widetilde{P}_{k_m} - P^*_{k_m})\|_{\mathbb{H},*} \le c(\|U^*_\infty - U^*_{k_m}\|_{\mathbb{R}^L} + \|\nabla(\widetilde{p}_{k_m} - p^*_\infty)\|_{L^2(\Omega)} + \|(\sigma^*_{k_m} - \sigma^*_\infty)\nabla p^*_\infty\|_{L^2(\Omega)}).$$

Thanks to the convergence of $\{U^*_{k_m}\}$, the pointwise convergence of $\{\sigma^*_{k_m}\}$ in Theorem 4.2 and (4.7), the right-hand side tends to zero. Now the desired assertion follows from the triangle inequality and (4.7). □

4.2 Convergence of AFEM

Now we establish the main theoretical result of this work: the sequence of discrete solutions generated by Algorithm 1 contains a convergent subsequence $\{(\sigma^*_{k_m}, u^*_{k_m}, U^*_{k_m}, p^*_{k_m}, P^*_{k_m})\}$, whose limit solves the necessary optimality system (2.5). By Theorems 4.2 and 4.3, it suffices to show that the limiting solution $\{(\sigma^*_\infty, u^*_\infty, U^*_\infty, p^*_\infty, P^*_\infty)\}$ also solves the necessary optimality system (2.5). Our arguments begin with the observation that the maximal error indicator over marked elements has a vanishing limit; see Lemma 4.3 below. Then two sequences of residuals with respect to $(u^*_{k_m}, U^*_{k_m})$ and $(p^*_{k_m}, P^*_{k_m})$ are proved to converge to zero weakly in Lemma 4.4 below. This together with Theorems 4.2 and 4.3, implies the claim for the first two equations in (2.5), cf. Lemma 4.5 below. The variational inequality in (2.5) can be verified analogously, cf. Lemma 4.6 below.

First we show that the maximal error indicator over the marked elements has a vanishing limit.

Lemma 4.3. *Let $\{\mathcal{T}_k, \mathcal{A}_k \times \mathbb{H}_k, (\sigma^*_k, u^*_k, U^*_k, p^*_k, P^*_k)\}$ be the sequence of meshes, discrete admissible sets, finite element spaces and discrete solutions generated by Algorithm 1 and $\mathcal{M}_k$ the set of marked elements by* (3.6)*. Then for each convergent subsequence $\{(\sigma^*_{k_m}, u^*_{k_m}, U^*_{k_m}, p^*_{k_m}, P^*_{k_m})\}$, there holds*

$$\lim_{m\to\infty} \max_{T\in\mathcal{M}_{k_m}} \eta_{k_m}(\sigma^*_{k_m}, u^*_{k_m}, U^*_{k_m}, p^*_{k_m}, P^*_{k_m}, T) = 0.$$

Proof. We denote by $\widetilde{T}$ the element with the largest error indicator in $\mathcal{M}_{k_m}$. Since the set $D_{\widetilde{T}} \subset \Omega^0_{k_m}$, it follows from Lemma 3.1 that

$$|D_{\widetilde{T}}| \le c\|h_{k_m}\|^d_{L^\infty(\Omega^0_{k_m})} \to 0, \quad |\partial\widetilde{T} \cap e_l| \le c\|h_{k_m}\|^{d-1}_{L^\infty(\Omega^0_{k_m})} \to 0 \quad \text{as } m\to\infty. \tag{4.8}$$

By Lemma 3.2, the local quasi-uniformity of $\mathcal{T}_{k_m}$, inverse estimates, trace theorem [15] and the triangle inequality, we have

$$\begin{aligned}
\eta^2_{k_m,1}(\sigma^*_{k_m}, u^*_{k_m}, U^*_{k_m}, \widetilde{T}) &\le c(\|\nabla u^*_{k_m}\|^2_{L^2(D_{\widetilde{T}})} + h_{\widetilde{T}}\|u^*_{k_m,l} - U^*_{k_m,l}\|^2_{L^2(\partial\widetilde{T}\cap e_l)})\\
&\le c(\|(u^*_{k_m} - u^*_\infty, U^*_{k_m} - U^*_\infty)\|^2_{\mathbb{H},*} + \|\nabla u^*_\infty\|^2_{L^2(D_{\widetilde{T}})} + \|u^*_{\infty,l} - U^*_{\infty,l}\|^2_{L^2(\partial\widetilde{T}\cap e_l)}),\\
\eta^2_{k_m,2}(\sigma^*_{k_m}, p^*_{k_m}, P^*_{k_m}, \widetilde{T}) &\le c(\|\nabla p^*_{k_m}\|^2_{L^2(D_{\widetilde{T}})} + h_{\widetilde{T}}\|p^*_{k_m,l} - P^*_{k_m,l}\|^2_{L^2(\partial\widetilde{T}\cap e_l)})\\
&\le c(\|(p^*_{k_m} - p^*_\infty, P^*_{k_m} - P^*_\infty)\|^2_{\mathbb{H},*} + \|\nabla p^*_\infty\|^2_{L^2(D_{\widetilde{T}})} + \|p^*_{\infty,l} - P^*_{\infty,l}\|^2_{L^2(\partial\widetilde{T}\cap e_l)}),
\end{aligned}$$

$$\begin{aligned}
\eta^2_{k_m,3}(\sigma^*_{k_m}, u^*_{k_m}, p^*_{k_m}, \widetilde{T}) &\le c(h^{4-d}_{\widetilde{T}}\|\nabla u^*_{k_m}\|^2_{L^2(\widetilde{T})}\|\nabla p^*_{k_m}\|^2_{L^2(\widetilde{T})} + h^2_{\widetilde{T}}\|\nabla\sigma^*_{k_m}\|^2_{L^2(D_{\widetilde{T}})})\\
&\le c|D_{\widetilde{T}}|^{4/d-1}\big((\|\nabla(u^*_{k_m} - u^*_\infty)\|^2_{L^2(\widetilde{T})} + \|\nabla u^*_\infty\|^2_{L^2(\widetilde{T})})(\|\nabla(p^*_{k_m} - p^*_\infty)\|^2_{L^2(\widetilde{T})} + \|\nabla p^*_\infty\|^2_{L^2(\widetilde{T})})\\
&\quad + (\|\nabla(\sigma^*_{k_m} - \sigma^*_\infty)\|^2_{L^2(D_{\widetilde{T}})} + \|\nabla\sigma^*_\infty\|^2_{L^2(D_{\widetilde{T}})})\big).
\end{aligned}$$

Now the desired result follows from the convergence of $\{\sigma^*_{k_m}\}$ and $\{(u^*_{k_m}, U^*_{k_m})\}$ in Theorem 4.2 and $\{(p^*_{k_m}, P^*_{k_m})\}$ from Theorem 4.3, (4.8), and the absolute continuity of the norms $\|\cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_{L^2(\Gamma)}$ with respect to the Lebesgue measure. □

Now we define two residuals with respect to (u_k^*, U_k^*) and (p_k^*, P_k^*) as

$$\begin{aligned}\langle \mathcal{R}(u_k^*, U_k^*), (v, V)\rangle &:= a(\sigma_k^*, (u_k^*, U_k^*), (v, V)) - \langle I, V\rangle \quad \forall (v, V) \in \mathbb{H},\\ \langle \mathcal{R}(p_k^*, P_k^*), (v, V)\rangle &:= a(\sigma_k^*, (p_k^*, P_k^*), (v, V)) - \langle U_k^* - U^\delta, V\rangle \quad \forall (v, V) \in \mathbb{H}.\end{aligned}$$

Clearly, by definition, we have the Galerkin orthogonality

$$\begin{aligned}\langle \mathcal{R}(p_k^*, P_k^*), (v, V)\rangle &= 0 \quad \forall (v, V) \in \mathbb{H}_k,\\ \langle \mathcal{R}(p_k^*, P_k^*), (v, V)\rangle &= 0 \quad \forall (v, V) \in \mathbb{H}_k.\end{aligned} \tag{4.9}$$

To relate the limit $\{(\sigma_\infty^*, u_\infty^*, U_\infty^*, p_\infty^*, P_\infty^*)\}$ to the optimality system (2.5), we exploit the marking assumption (3.6) in Algorithm 1. The next result gives the weak convergence of the residuals to zero.

Lemma 4.4. *For the convergent subsequence $\{(\sigma_{k_m}^*, u_{k_m}^*, U_{k_m}^*, p_{k_m}^*, P_{k_m}^*)\}$ given in Theorems 4.2 and 4.3, there hold*

$$\begin{aligned}\lim_{m\to\infty} \langle \mathcal{R}(u_{k_m}^*, U_{k_m}^*), (v, V)\rangle &= 0 \quad \forall (v, V) \in \mathbb{H},\\ \lim_{m\to\infty} \langle \mathcal{R}(p_{k_m}^*, P_{k_m}^*), (v, V)\rangle &= 0 \quad \forall (v, V) \in \mathbb{H}.\end{aligned}$$

Proof. We only prove the first assertion since the second follows analogously. For notational simplicity, we relabel the index k_m by k. Let I_k and I_k^{sz} be the Lagrange and Scott-Zhang interpolation operators respectively associated with the space V_k. Then by the Galerkin orthogonality (4.9), elementwise integration by parts and the error estimate for I_k^{sz}, cf. Lemma 3.3, we deduce for $k > l$ and any $(\psi, V) \in C^\infty(\overline{\Omega}) \times \mathbb{R}_\diamond^L$

$$\begin{aligned}\big|\langle \mathcal{R}(u_k^*, U_k^*), (\psi, V)\rangle\big| &= \big|\langle \mathcal{R}(u_k^*, U_k^*), (\psi - I_k\psi, 0)\rangle\big| = \big|\langle \mathcal{R}(u_k^*, U_k^*), (w - I_k^{sz} w, 0)\rangle\big|\\ &= \Big|(\sigma_k^* \nabla u_k^*, \nabla(w - I_k^{sz} w)) + \sum_{l=1}^L z_l^{-1}((u_k^* - U_{k,l}^*), (w - I_k^{sz} w))_{L^2(e_l)}\Big|\\ &\le c \sum_{T\in\mathcal{T}_k} \eta_{k,1}(\sigma_k^*, u_k^*, U_k^*, T)\|w\|_{H^1(D_T)}\\ &= c\Big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+} \eta_{k,1}(\sigma_k^*, u_k^*, U_k^*, T)\|w\|_{H^1(D_T)} + \sum_{T\in\mathcal{T}_l^+} \eta_{k,1}(\sigma_k^*, u_k^*, U_k^*, T)\|w\|_{H^1(D_T)}\Big).\end{aligned}$$

where $w = \psi - I_k\psi$. By appealing to Lemma 3.2 and (3.4), we deduce

$$\Big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+} \eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, T)\Big)^{1/2} \le c$$

and further by the error estimate of the interpolation operator I_k from Lemma 3.3, we arrive at

$$\big|\langle \mathcal{R}(u_k^*, U_k^*), (\psi, V)\rangle\big| \le c_1\|h_l\|_{L^\infty(\Omega_l^0)}\|\psi\|_{H^2(\Omega)} + c_2\Big(\sum_{T\in\mathcal{T}_l^+} \eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, T)\Big)^{1/2}\|\psi\|_{H^2(\Omega)}.$$

By Lemma 3.1, $c_1\|h_l\|_{L^\infty(\Omega_l^0)}\|\psi\|_2 \to 0$ as $l \to \infty$. From $\mathcal{T}_l^+ \subset \mathcal{T}_k^+ \subset \mathcal{T}_k \subset \mathcal{M}_k$ for $k > l$ and the marking condition (3.6), we deduce

$$\begin{aligned}\Big(\sum_{T\in\mathcal{T}_l^+} \eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, T)\Big)^{1/2} &\le \sqrt{|\mathcal{T}_l^+|}\max_{T\in\mathcal{T}_l^+} \eta_{k,1}(\sigma_k^*, u_k^*, U_k^*, T) \le \sqrt{|\mathcal{T}_l^+|}\max_{T\in\mathcal{T}_k^+} \eta_{k,1}(\sigma_k^*, u_k^*, U_k^*, T)\\ &\le \sqrt{|\mathcal{T}_l^+|}\max_{T\in\mathcal{M}_k} \eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*, T).\end{aligned}$$

Now Lemma 4.3 implies that for any fixed large l_1, we can choose some $k_1 > l_1$ such that

$$c_2\Big(\sum_{T\in\mathcal{T}_l^+} \eta_{k,1}^2(\sigma_k^*, u_k^*, U_k^*, T)\Big)^{1/2}\|\psi\|_2 < \varepsilon$$

for any positive small number ε and $k > k_1$. Thus, we arrive at

$$\lim_{m\to\infty}\langle \mathcal{R}(u_{k_m}^*, U_{k_m}^*), (v, V)\rangle = 0 \quad \forall (v, V) \in C^\infty(\overline{\Omega}) \times \mathbb{R}_\diamond^L,$$

which, together with the density of $C^\infty(\overline{\Omega})$ in $H^1(\Omega)$, gives the desired assertion. □

Next we show that the limit $(\sigma_\infty^*, u_\infty^*, U_\infty^*, p_\infty^*, P_\infty^*)$ actually solves the variational equations in (2.5).

Lemma 4.5. *The solution to problem* (4.5) *solves the two variational equations in* (2.5)*, i.e.,*

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

Proof. We prove only the first assertion, since the proof of the second is analogous. Given the convergent subsequence $\{(\sigma_{k_m}^*, u_{k_m}^*, U_{k_m}^*, p_{k_m}^*, P_{k_m}^*)\}$ in Theorems 4.2 and 4.3, for any $(v, V) \in \mathbb{H}$, there holds

$$\begin{aligned} \Big|a(\sigma_\infty^*, (u_\infty^*, U_\infty^*), (v, V)) - \langle I, V\rangle\Big| &\le \sum_{l=1}^L z_l^{-1}\Big|(u_\infty^* - U_{\infty,l}^* - u_{k_m}^* + U_{k_m,l}^*, v - V_l)_{L^2(e_l)}\Big| \\ &\quad + \Big|((\sigma_\infty^*\nabla u_\infty^* - \sigma_{k_m}^*\nabla u_{k_m}^*), \nabla v)_{L^2(\Omega)}\Big| + \big|\langle \mathcal{R}(u_{k_m}^*, U_{k_m}^*), (v, V)\rangle\big|. \end{aligned}$$

In view of Theorem 4.2 and Lemma 4.4, the first and third terms tend to zero. For the second term,

$$\begin{aligned} |((\sigma_\infty^*\nabla u_\infty^* - \sigma_{k_m}^* u_{k_m}^*), \nabla v)| &\le |(\sigma_\infty^*\nabla(u_\infty^* - u_{k_m}^*), \nabla v)| + |((\sigma_\infty^* - \sigma_{k_m}^*)\nabla u_{k_m}^*, \nabla v)| \\ &\le |(\sigma_\infty^*\nabla(u_\infty^* - u_{k_m}^*), \nabla v)| + \|\nabla u_{k_m}^*\|_{L^2(\Omega)}\|(\sigma_\infty^* - \sigma_{k_m}^*)\nabla v\|_{L^2(\Omega)} \to 0, \end{aligned}$$

by the convergence of $\{u_{k_m}^*\}$, and the pointwise convergence of $\{\sigma_{k_m}^*\}$ in Theorem 4.2 and Lebesgue's dominated convergence theorem [15]. □

Now we turn to the variational inequality in (2.5). We resort again to a density argument like in the proof of Lemma 4.4. Specifically, we first show that the variational inequality holds for $(\sigma_\infty^*, u_\infty^*, p_\infty^*)$ over the smooth subset, and then extend the assertion to $\mathcal{A}$ by a density argument.

Lemma 4.6. *The solution to the variational inequality of problem* (4.5) *satisfies*

$$\alpha(\nabla\sigma_\infty^*, \nabla(\mu - \sigma_\infty^*)) - (\nabla u_\infty^*, \nabla p_\infty^*(\mu - \sigma_\infty^*)) \ge 0 \quad \forall \mu \in \mathcal{A}.$$

Proof. Like before, we relabel the convergent subsequence $\{(\sigma_{k_m}^*, u_{k_m}^*, U_{k_m}^*, p_{k_m}^*, P_{k_m}^*)\}$ in Theorems 4.2-4.3 by $\{(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*)\}$, and let I_k be the Lagrange interpolation operator associated with V_k. Then for any $\mu \in \widetilde{\mathcal{A}} := \mathcal{A} \cap C^\infty(\overline{\Omega})$, $I_k\mu \in \mathcal{A}_k$ and further the discrete variational inequality in (3.3) yields

$$\begin{aligned} &\alpha(\nabla\sigma_k^*, \nabla(\mu - \sigma_k^*)) - ((\mu - \sigma_k^*)\nabla u_k^*, \nabla p_k^*) \\ &= \alpha(\nabla\sigma_k^*, \nabla(\mu - I_k\mu)) - ((\mu - I_k\mu)\nabla u_k^*, \nabla p_k^*) \\ &\quad + \alpha(\nabla\sigma_k^*, \nabla(I_k\mu - \sigma_k^*)) - ((I_k\mu - \sigma_k^*)\nabla u_k^*, \nabla p_k^*) \\ &\ge \alpha(\nabla\sigma_k^*, \nabla(\mu - I_k\mu)) - ((\mu - I_k\mu)\nabla u_k^*, \nabla p_k^*). \end{aligned} \tag{4.10}$$

Using elementwise integration by parts, the definition of $\eta_{k,3}$ and error estimates for I_k, cf. Lemma 3.3, we deduce that for $k > l$, there holds

$$\begin{aligned} \Big|\alpha(\nabla\sigma_k^*, \nabla(\mu - I_k\mu)) - ((\mu - I_k\mu)\nabla u_k^*, \nabla p_k^*)\Big| &\le c\sum_{T\in\mathcal{T}_k}\eta_{k,3}(\sigma_k^*, u_k^*, p_k^*, T)\|\mu\|_{H^2(T)} \\ &\le c_3\Big(\Big(\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}^2(\sigma_k^*, u_k^*, p_k^*, T)\Big)^{1/2} + \Big(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^2(\sigma_k^*, u_k^*, p_k^*, T)\Big)^{1/2}\Big)\|\mu\|_{H^2(\Omega)}. \end{aligned}$$

The Lemma 3.2, (3.4) and the $H^1(\Omega)$-convergence of the sequence $\{\sigma_k^*\}$ from Theorem 4.2 and Lemma 3.1 give

$$\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\eta_{k,3}^2(\sigma_k^*,u_k^*,p_k^*,T)\le c(\|h_l\|_{L^\infty(\Omega_0^l)}^{4-d}\|\nabla p_k\|_{L^2(\Omega)}^2\sum_{T\in\mathcal{T}_k\setminus\mathcal{T}_l^+}\|\nabla u_k^*\|_{L^2(T)}^2+\|h_l\|_{L^\infty(\Omega_0^l)}^2\|\nabla\sigma_k^*\|_{L^2(\Omega)}^2)$$
$$\le c(\|h_l\|_{L^\infty(\Omega_0^l)}^{4-d}+\|h_l\|_{L^\infty(\Omega_0^l)}^2)\le c\|h_l\|_{L^\infty(\Omega_0^l)}^{4-d}\to 0.$$

Upon noting the inclusion $\mathcal{T}_l^+\subset\mathcal{T}_k$ for $k>l$, we deduce from the marking condition (3.6)

$$(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^2(\sigma_k^*,u_k^*,p_k^*,T))^{1/2}\le\sqrt{|\mathcal{T}_l^+|}\max_{T\in\mathcal{T}_l^+}\eta_{k,3}(\sigma_k^*,u_k^*,p_k^*,T)\le\sqrt{|\mathcal{T}_l^+|}\max_{T\in\mathcal{M}_k}\eta_k(\sigma_k^*,u_k^*,U_k^*,p_k^*,P_k^*,T).$$

Appealing again to Lemma 4.3, we can choose $k_2>l_2$ for some large fixed l_2 such that when $k>k_2$ $c_3(\sum_{T\in\mathcal{T}_l^+}\eta_{k,3}^2(\sigma_k^*,u_k^*,p_k^*,T))^{1/2}\|\mu\|_{H^2(\Omega)}$ is smaller than any given positive number. Hence

$$(\alpha\nabla\sigma_k^*,\nabla(\mu-I_k\mu))-(\nabla u_k^*,\nabla p_k^*(\mu-I_k\mu))\to 0\quad\forall\mu\in\widetilde{\mathcal{A}}.\tag{4.11}$$

Using the $H^1(\Omega)$-convergence of $\{\sigma_k^*\}$ from Theorem 4.2, we have

$$(\alpha\nabla\sigma_k^*,\nabla(\mu-\sigma_k^*))\to(\alpha\nabla\sigma_\infty^*,\nabla(\mu-\sigma_\infty^*))\quad\forall\mu\in\widetilde{\mathcal{A}}.\tag{4.12}$$

The convergence of $\{p_k^*\}$ to p_∞^* in $H^1(\Omega)$ in Theorem 4.3, the stability estimate (3.4) and the box constraint in $\widetilde{\mathcal{A}}$ yield

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

and this together with Theorem 4.2 implies

$$(\mu\nabla u_k^*,\nabla p_k^*)=(\mu\nabla u_k^*,\nabla(p_k^*-p_\infty^*))+(\mu\nabla u_k^*,\nabla p_\infty^*)\to(\mu\nabla u_\infty^*,\nabla p_\infty^*)\quad\forall\mu\in\widetilde{\mathcal{A}}.\tag{4.13}$$

By elementary calculations, we derive

$$(\sigma_k^*\nabla u_k^*,\nabla p_k^*)-(\sigma_\infty^*\nabla u_\infty^*,\nabla p_\infty^*)=(\sigma_k^*\nabla u_k^*,\nabla(p_k^*-p_\infty^*))+((\sigma_k^*-\sigma_\infty^*)\nabla u_k^*,\nabla p_\infty^*)$$
$$+(\sigma_\infty^*\nabla(u_k^*-u_\infty^*),\nabla p_\infty^*).$$

Repeating the arguments for (4.13) yields that for the first and third terms there hold

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

The stability estimate (3.4), the pointwise convergence of $\{\sigma_k^*\}$ of Theorem 4.2 and Lebesgue's dominated convergence theorem [15] show

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

Hence

$$(\sigma_k^*\nabla u_k^*,\nabla p_k^*)\to(\sigma_\infty^*\nabla u_\infty^*,\nabla p_\infty^*).\tag{4.14}$$

Now by passing both sides of (4.10) to the limit and combining (4.11)-(4.14), we obtain

$$\alpha(\nabla\sigma_\infty^*,\nabla(\mu-\sigma_\infty^*))_{L^2(\Omega)}-(\nabla u_\infty^*,\nabla p_\infty^*(\mu-\sigma_\infty^*))_{L^2(\Omega)}\ge 0\quad\forall\mu\in\widetilde{\mathcal{A}}.$$

By means of the density of $C^\infty(\overline{\Omega})$ in $H^1(\Omega)$ and the construction via a standard mollifier [15], for any $\mu\in\mathcal{A}$ there exists a sequence $\{\mu^n\}\subset\widetilde{\mathcal{A}}$ such that $\|\mu^n-\mu\|_{H^1(\Omega)}\to 0$ as $n\to\infty$. Then by Lebesgue's dominated convergence theorem [15], we deduce

$$(\alpha\nabla\sigma_\infty^*,\nabla\mu^n)\to(\alpha\nabla\sigma_\infty^*,\nabla\mu)\quad\text{and}\quad(\mu^n\nabla u_\infty^*,\nabla p_\infty^*)\to(\mu\nabla u_\infty^*,\nabla p_\infty^*)$$

after possibly passing to a subsequence. The desired result follows from the preceding two estimates. □

Finally, by combining Theorems 4.2 and 4.3, with Lemmas 4.5 and 4.6, we arrive at the main theoretical result of this part: the sequence of solutions generated by the AFEM contains a subsequence convergent to a solution of the continuous necessary optimality system (2.5).

Theorem 4.4. *The sequence of discrete solutions* $\{(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*)\}$ *generated by Algorithm 1 has a subsequence* $\{(\sigma_{k_m}^*, u_{k_m}^*, U_{k_m}^*, p_{k_m}^*, P_{k_m}^*)\}$ *converging to a solution* $(\sigma^*, u^*, U^*, p^*, P^*)$ *to the continuous optimality system* (2.5) *in the following sense:*

$$\|\sigma_{k_m}^* - \sigma^*\|_{H^1(\Omega)},\ \|(u_{k_m}^* - u^*, U_{k_m}^* - U^*)\|_{\mathbb{H},*},\ \|(p_{k_m}^* - p^*, P_{k_m}^* - P^*)\|_{\mathbb{H},*} \to 0 \quad \text{as } m \to \infty.$$

5 Numerical experiments and discussions

Now we present numerical results to illustrate the convergence and efficiency of the adaptive algorithm. The setup of the numerical experiments is as follows. The domain Ω is taken to be a square $\Omega = (-1, 1)^2$. There are sixteen electrodes $\{e_l\}_{l=1}^L$ (with $L = 16$) evenly distributed along the boundary Γ, each of the length 1/4, thus occupying one half of the boundary Γ. The contact impedances $\{z_l\}_{l=1}^L$ on the electrodes $\{e_l\}_{l=1}^L$ are all set to unit, and the background conductivity σ_0 is taken to be $\sigma_0 \equiv 1$. For each example, we measure the electrode voltages U for the first ten sinusoidal input currents, in order to gain enough information about the true conductivity $\sigma^\dagger$. Then the noisy data U^δ is generated by adding componentwise Gaussian noise to the exact data $U(\sigma^\dagger)$ as follows

$$U_l^\delta = U_l(\sigma^\dagger) + \epsilon \max_l |U_l(\sigma^\dagger)|\xi_l, \quad l = 1, \ldots, L,$$

where ϵ is the noise level, and $\{\xi_l\}$ follow the standard normal distribution. In all the experiments, the marking strategy (3.6) in the module `MARK` is represented by a specific maximum strategy, cf. Remark 3.2, i.e., mark a subset $\mathcal{M}_k \subseteq \mathcal{T}_k$, i.e., the refinement set, such that

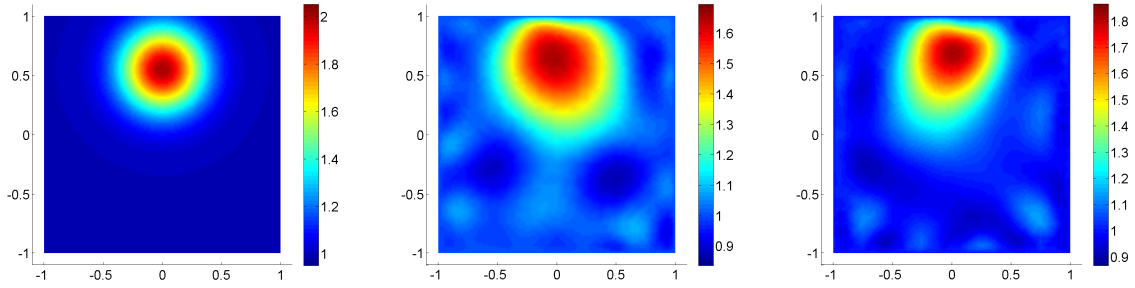
$$\eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*, \mathcal{M}_k) \geq \theta \eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*, \mathcal{T}_k),$$

with a threshold $\theta \in [0, 1]$, and it contains all elements T with $\eta_k(\sigma_k^*, u_k^*, U_k^*, p_k^*, P_k^*, T)$ in a descending order. In the computation, we fix the threshold θ at $\theta = 0.7$. The discrete nonlinear optimization problems (3.1)-(3.2) are solved by a nonlinear conjugated gradient method and the initial guess of the conductivity at the coarsest mesh $\mathcal{T}_1$ is initialized to the background $\sigma_0 = 1$, and then the reconstruction on the mesh $\mathcal{T}_{k-1}$ is interpolated to the mesh $\mathcal{T}_k$ to warm start the conjugate gradient iteration for the discrete formulation on the mesh $\mathcal{T}_k$. Throughout the adaptive loop, the regularization parameter α in the Tikhonov model (2.3) is fixed and it is determined in a trial-and-error manner.

Example 1. *The true conductivity* $\sigma^\dagger$ *is given by* $\sigma^\dagger(x) = \sigma_0(x) + e^{-8(x_1^2 + (x_2 - 0.55)^2)}$*, with the background conductivity* $\sigma_0(x) = 1$*.*

In this example, the true conductivity field $\sigma^\dagger$ consists of a very smooth blob in a constant background, and the profile is shown in Fig. 1(a). The final recovered conductivity fields from the voltage measurements with $\epsilon = 0.1\%$ noise in the data are shown in Fig. 1. For both uniform and adaptive refinements, the reconstructions capture well the location and height of the blob: it is very smooth, due to the use of a smoothness prior. Consequently, it does not induce any grave solution singularity. The reconstructions by the uniform refinement and that by the adaptive one are similar to each other in terms of location and magnitude, both suffering from slight loss of the contrast.

Next we examine the adaptive refinement more closely. On a very coarse initial mesh $\mathcal{T}_0$, which is a uniform triangulation of the domain Ω, cf. the first panel of Fig. 2(a), the recovered conductivity tends to have pronounced oscillations around the boundary, since the forward solution is not accurately resolved. In particular, the discretization error can compromise the reconstruction accuracy, and it induces large errors in the location and height of the recovered blob. This motivates the use of the adaptive strategy. The meshes during the adaptive iteration and the corresponding recoveries are shown

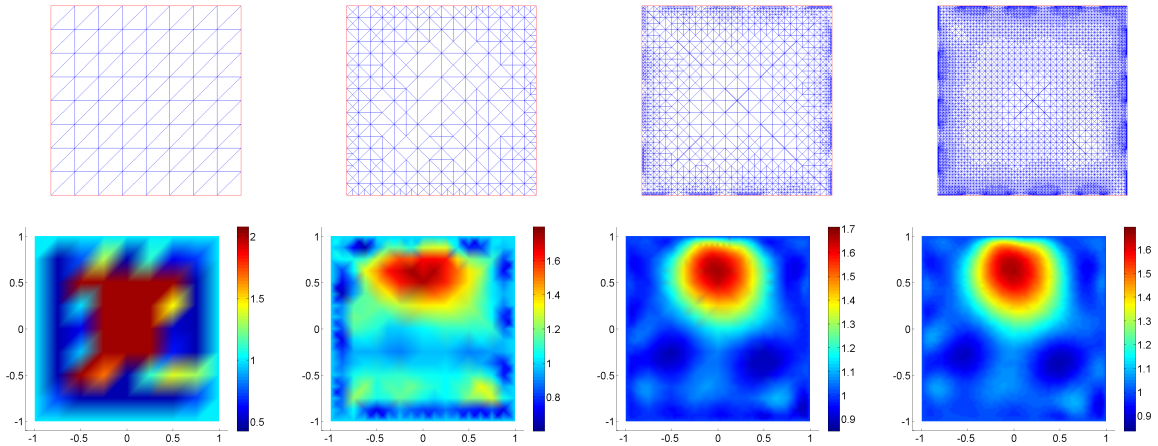


(a) true conductivity (b) adaptive refinement (c) uniform refinement

Figure 1: The final reconstructions by the uniform and adaptive refinements for example 1 with $\epsilon = 0.1\%$ noise in the data. The degree of freedom is 9818 and 16641 for the adaptive and uniform refinement, respectively. The regularization parameter α is fixed at $\alpha = 2.5 \times 10^{-4}$.

in Fig. 2. The refinement step first concentrates only on the region around the electrode surface. This is expected from the abrupt change of the boundary condition over there, and thus the solution to the direct problem has weak singularity [19]. Afterwards, the AFEM starts to refine also the interior of the domain, simultaneously with the boundary region. Accordingly, the spurious oscillations in the reconstruction are suppressed as the adaptive iteration proceeds. Interestingly, the central part of the domain is refined only slightly during the whole refinement procedure. This concurs with the observation that the center is much harder to resolve from the boundary data. Hence, the refinement step tends to adapt automatically to the resolving power of the conductivity field in different regions.

In Fig. 3, we plot the $L^2(\Omega)$ error of the reconstructions versus the degree of freedom N of the mesh $\mathcal{T}_k$ for the adaptive and uniform refinement, where the reference solution is computed on a much finer mesh. It is observed that with the same degree of freedom, the adaptive algorithm can give much more accurate reconstructions than the uniform one. This shows clearly the efficiency of the adaptive algorithm.

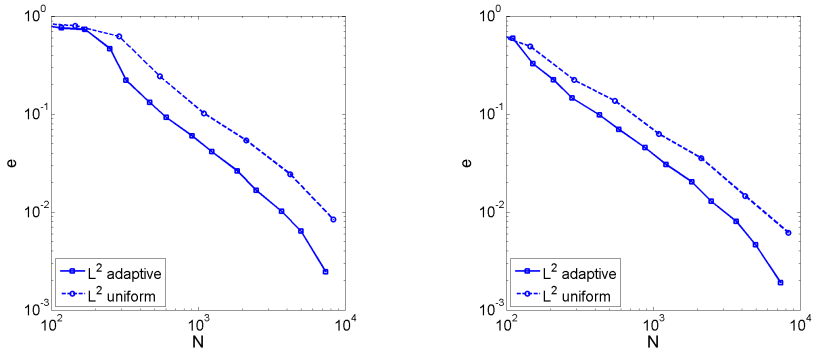


(a) 1st step (b) 5th step (c) 10th step (d) 15th step

Figure 2: The recovered conductivity distributions σ_k^* during the adaptive refinement, for example 1 with $\epsilon = 0.1\%$ noise in the data. The regularization parameter α is fixed at $\alpha = 2.5 \times 10^{-4}$.

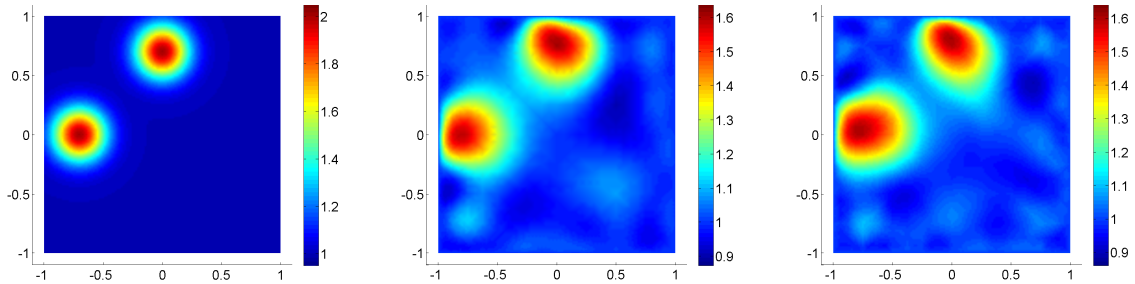
A second example contains two neighboring smooth blobs.

Example 2. *The true conductivity* $\sigma^\dagger$ *is given by* $\sigma^\dagger(x) = \sigma_0(x) + e^{-20((x_1+0.7)^2+x_2^2)} + e^{-20(x_1^2+(x_2-0.7)^2)}$, *and the background conductivity* $\sigma_0(x) = 1$.



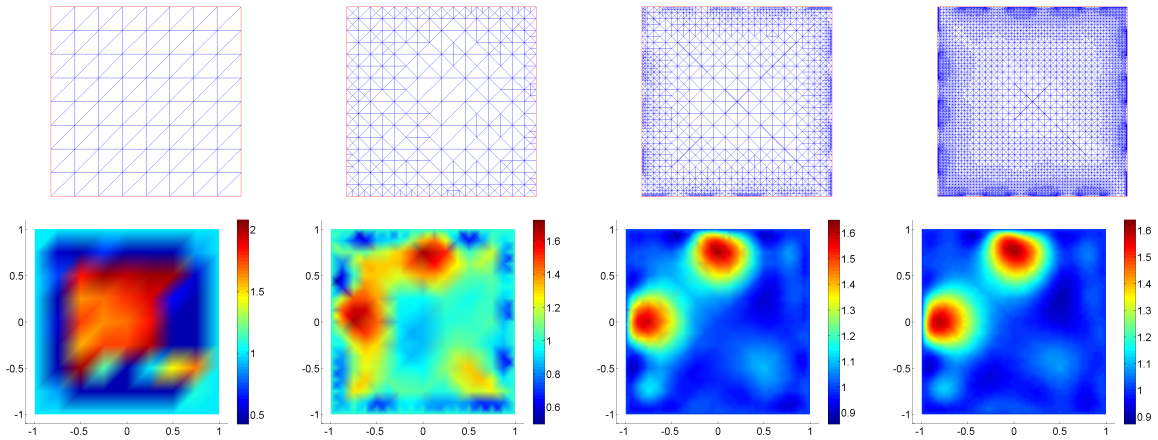
(a) $\epsilon = 1 \times 10^{-3}$ (b) $\epsilon = 1 \times 10^{-2}$

Figure 3: The $L^2(\Omega)$ error e versus the degree of freedom N of the mesh, for example 1 at two different noise levels, using the adaptive refinement (solid line) and uniform refinement (dashed line).



(a) true conductivity (b) adaptive refinement (c) uniform refinement

Figure 4: The final reconstructions by the adaptive and uniform refinements for example 2 with $\epsilon = 0.1\%$ noise in the data. The degree of freedom is 9803 and 16641 for the adaptive and uniform refinement, respectively. The regularization parameter α is fixed at $\alpha = 2.5 \times 10^{-4}$.

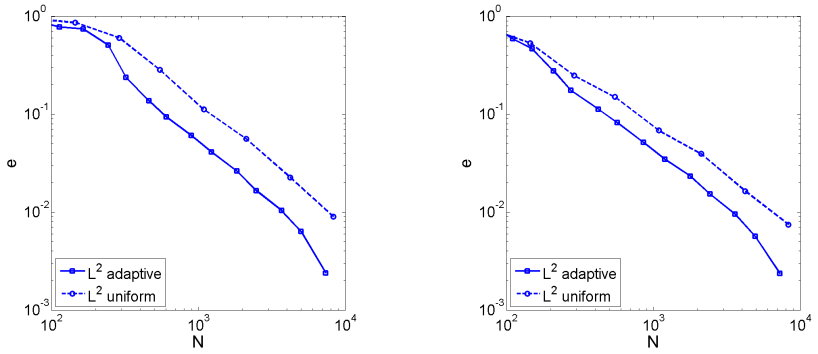


(a) 1st step (b) 5th step (c) 10th step (d) 15th step

Figure 5: The recovered conductivity during the adaptive refinement, for example 2 with $\epsilon = 0.1\%$ noise in the data. The regularization parameter is fixed at $\alpha = 2.5 \times 10^{-4}$.

Like before, the true conductivity $\sigma^\dagger$ is very smooth (cf. Fig. 4(a) for the profile), and thus the smoothness penalty is suitable. Overall, the observations from example 1 remain valid: the reconstructed coefficient captures very well the supports of the inclusions, and the magnitude is also reasonable. The reconstruction by the adaptive algorithm is comparable with that based on uniform refinement, but requiring far less degrees of freedom.

At the initial stage, the mesh refinement mainly occurs in the region around electrode surfaces, where the weak solution singularity appears. However, as the iterative refinement proceeds, the region away from the boundary is also refined simultaneously, but to a lesser degree, which is especially pronounced for the central part of the domain. In case of a very coarse initial mesh, the recovery even fails to identify the number of inclusions, but as the AFEM proceeds, the spurious oscillations disappear, and then it can identify reasonably the locations and magnitudes of the blobs from the reconstructions, cf. Fig. 5. In Fig. 6, we show the $L^2(\Omega)$ error of the reconstructions versus the degree of freedom N of the mesh $\mathcal{T}_k$ for the adaptive and uniform refinement. These plots fully show the efficiency of the adaptive algorithm, for both $\epsilon = 0.1\%$ and $\epsilon = 1\%$ noise in the data.



(a) $\epsilon = 1 \times 10^{-3}$ (b) $\epsilon = 1 \times 10^{-2}$

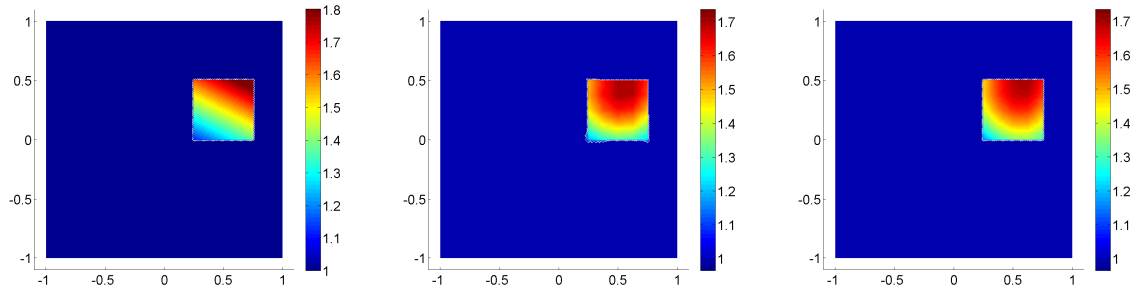
Figure 6: The $L^2(\Omega)$ error versus the degree of freedom N of the mesh, for example 2 at two different noise levels, using the adaptive refinement (solid line) and uniform refinement (dashed line).

Now we consider a discontinuous conductivity field as the last example.

Example 3. *The true conductivity $\sigma^\dagger$ is given by $\sigma^\dagger(x) = \sigma_0(x) + (x_1/2 + x_2)\chi_{\Omega'}$, where $\chi_{\Omega'}$ is the characteristic function of the set $\Omega' = (1/4, 3/4) \times (0, 1/2)$, and the back ground conductivity $\sigma_0(x) = 1$.*

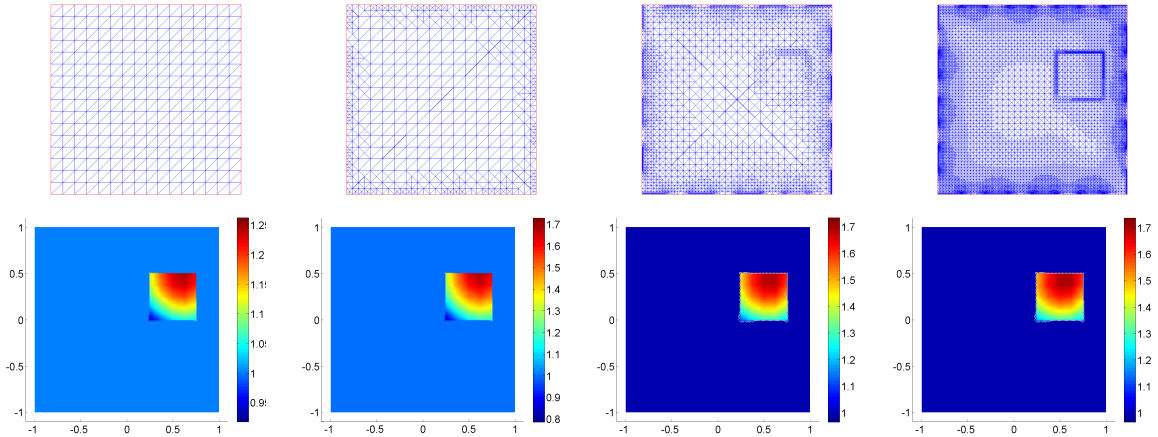
Since the $H^1(\Omega)$ penalty imposes a global smoothness condition, it is unsuitable for recovering discontinuous conductivity fields. Hence, in this example we assume that the support Ω' of the true conductivity field $\sigma^\dagger$ is known, and aim at determining the variation within the support using the $H^1(\Omega')$ semi-norm penalty. With minor modifications, the adaptive algorithm and the convergence proof can be extended (with the only change lying in the variational inequality now defined only on the subdomain Ω', and the estimator $\eta^2_{\mathcal{T},3}(\sigma^*_{\mathcal{T}}, u^*_{\mathcal{T}}, p^*_{\mathcal{T}}, T)$ is only for elements in Ω').

The numerical results for the example are presented in Figs. 7, 8 and 9. The observations from the preceding two examples remain largely valid. The magnitude of the conductivity is slightly reduced, but otherwise the profile is reasonable, and visually the reconstructions by the adaptive and the uniform refinements are close to each other, cf. Figs. 7(b) and 7(c). Even though the conductivity field σ is discontinuous, the adaptive algorithm first mainly resolves the singularity due to the abrupt change of boundary conditions, i.e., the weak singularity of the solution around the boundary is dominant, cf. Fig. 8(b). As the adaptive iteration proceeds, the algorithm then starts to refine the region near the boundary of Ω': first the part close to the boundary $\partial\Omega$, and then the part away from $\partial\Omega$, cf. Figs. 8(c) and 8(d). This is consistent with the observation that the further away from the boundary, the more challenging is it to be resolved. The gain of computational efficiency is shown in Fig. 9: the $L^2(\Omega)$-error decreases faster with the increase of degree of freedom for the adaptive algorithm than the uniform refinement.



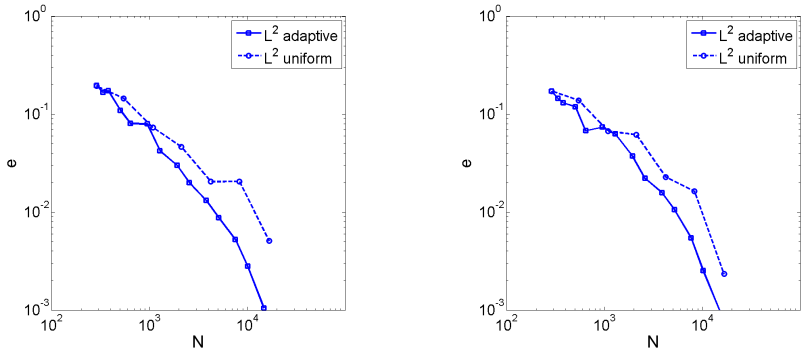
(a) true conductivity (b) adaptive refinement (c) uniform refinement

Figure 7: The final reconstructions by the adaptive and uniform refinements for example 3 with $\epsilon = 0.1\%$ noise in the data. The degree of freedom is 19608 and 33025 for the adaptive and uniform refinement, respectively. The regularization parameter α is fixed at $\alpha = 3.2 \times 10^{-3}$.



(a) 1st step (b) 5th step (c) 10th step (d) 15th step

Figure 8: The recovered conductivity during the adaptive refinement, for example 3 with $\epsilon = 0.1\%$ noise in the data. The regularization parameter is fixed at $\alpha = 3.2 \times 10^{-3}$.



(a) $\epsilon = 1 \times 10^{-3}$ (b) $\epsilon = 1 \times 10^{-2}$

Figure 9: The $L^2(\Omega)$ error versus the degree of freedom N of the mesh, for example 3 at two different noise levels, using the adaptive refinement (solid line) and uniform refinement (dashed line).

6 Concluding remarks

In this work, we have developed a novel adaptive finite element method for the electrical impedance tomography of recovering the conductivity distribution from the boundary voltage measurements, using the complete electrode model. The inverse problem is formulated as an output least-squares problem with a Sobolev smoothness penalty. The weak solution singularity around the electrode surfaces and low-regularity conductivity motivate the use of the adaptive refinement techniques. We have derived a residual-type a posteriori error estimator, which involves the state solution, adjoint solution and the identifying conductivity, and established the convergence of the sequence of solutions generated by the adaptive technique that the accumulation point solves the continuous optimality system. The efficiency and convergence of the proposed adaptive algorithm is fully confirmed by the numerical experiments.

This work represents only a first step towards the rigorous adaptive finite element method for nonlinear inverse problems associated with partial differential equations. There are several research problems deserving further study. First, the proposed algorithm is only for the smoothness penalty, which is essential in the development and convergence analysis of the algorithm. It is of much interest to derive and to analyze adaptive algorithms for nonsmooth penalties, e.g., total variation and sparsity. Second, numerically one observes that the algorithm can approximate a (local/global) minimizer of the continuous optimization well, instead of only a solution to the necessary optimality condition, cf. Theorem 4.4. This is still theoretically to be justified. Third, the reliability and optimality of the adaptive algorithm for nonlinear inverse problems are completely open, which seems not fully understood even for linear ones, e.g., flux reconstruction. The optimality issue is especially important in the context of inverse problems, where the data has only finite precision.

Acknowledgements

The work of B. Jin was partially supported by UK Engineering and Physical Sciences Research Council grant EP/M025160/1. The work of Y. Xu was partially supported by National Natural Science Foundation of China (11201307), MOE of China through Specialized Research Fund for the Doctoral Program of Higher Education (20123127120001), E-Institute of Shanghai Universities (E03004) and Innovation Program of Shanghai Municipal Education Commission (13YZ059). The work of J. Zou was substantially supported by Hong Kong RGC grants (projects 14306814 and 405513). The authors are very grateful to Mr. Chun-Man Yuen for his great help in carrying out the numerical experiments in this work.

References

[1] A. Adler, R. Gaburro, and W. Lionheart. Electrical impedance tomography. In *Handbook of Mathematical Methods in Imaging*, (editor: O. Scherzer), Springer-Verlag, 2011.

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

[3] W. Bangerth and A. Joshi. Adaptive finite element methods for the solution of inverse problems in optical tomography. *Inverse Problems*, 24(3):034011, 22 pp., 2008.

[4] R. Becker and B. Vexler. A posteriori error estimation for finite element discretization of parameter identification problems. *Numer. Math.*, 96:435–459, 2004.

[5] 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 (electronic), 2006.

[6] L. Beilina and C. Johnson. A posteriori error estimation in computational inverse scattering. *Math. Models Methods Appl. Sci.*, 15(1):23–35, 2005.

[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 and M. V. Klibanov. *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*. Springer-Verlag, New York, 2012.

[10] 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.

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

[12] 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.

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

[14] M. M. Dunlop and A. M. Stuart. The Bayesian formulation of EIT: analysis and algorithms. preprint, arXiv:1508.04106, 2015.

[15] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, Boca Raton, 1992.

[16] 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.

[17] 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.

[18] A. Griesbaum, B. Kaltenbacher, and B. Vexler. Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization. *Inverse Problems*, 24(2):025025, 20, 2008.

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

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

[21] M. Hintermüller and R. H. W. Hoppe. Goal-oriented adaptivity in pointwise state constrained optimal control of partial differential equations. *SIAM J. Control Optim.*, 48(8):5468–5487, 2010.

[22] M. Hintermüller, R. H. W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM, Control Optim. Calc. Var.*, 14(3):540–560, 2008.

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

[24] 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.

[25] 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.

[26] B. Kaltenbacher, A. Kirchner, and S. Veljović. Goal oriented adaptivity in the IRGNM for parameter identification in PDEs: I. reduced formulation. *Inverse Problems*, 30(4):0450011, 26, 2014.

[27] 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.

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

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

[30] R. Li, W. Liu, H. Ma, and T. Tang. Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM J. Control Optim.*, 41(5):1321–1349, 2002.

[31] W. Liu and N. Yan. A posteriori error estimates for distributed convex optimal control problems. *Adv. Comput. Math.*, 15(1-4):285–309, 2001.

[32] 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.

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

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

[35] 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.

[36] R. Verfürth. *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester, New York, Stuttgart, 1996.

[37] R. Winkler and A. Rieder. Resolution-controlled conductivity discretization in electrical impedance tomography. *SIAM J. Imaging Sci.*, 7(4):2048–2077, 2014.

[38] Y. Xu and J. Zou. Analysis of an adaptive finite element method for recovering the Robin coefficient. *SIAM J. Control Optim.*, 53(2):622–644, 2015.

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