SIAM J. NUMER. ANAL.
Vol. 47, No. 3, pp. 2132–2156

RECOVERY-BASED ERROR ESTIMATOR FOR INTERFACE PROBLEMS: CONFORMING LINEAR ELEMENTS*

ZHIQIANG CAI† AND SHUN ZHANG†

Abstract. This paper studies a new recovery-based a posteriori error estimator for the conforming linear finite element approximation to elliptic interface problems. Instead of recovering the gradient in the continuous finite element space, the flux is recovered through a weighted L^2 projection onto $H(\mathrm{div})$ conforming finite element spaces. The resulting error estimator is analyzed by establishing the reliability and efficiency bounds and is supported by numerical results. This paper also proposes an adaptive finite element method based on either the recovery-based estimators or the edge estimator through local mesh refinement and establishes its convergence. In particular, it is shown that the reliability and efficiency constants as well as the convergence rate of the adaptive method are independent of the size of jumps.

Key words. a posteriori error estimator, adaptive method, interface problems, finite element

AMS subject classifications. 65N30, 65N15

DOI. 10.1137/080717407

1. Introduction.

The a posteriori error estimators of the recovery type have been extensively studied by many researchers, e.g., [1, 2, 3, 5, 10, 28, 29, 34, 36], due to their many appealing properties: simplicity, universality, and asymptotic exactness. (The universality is in the sense that there is no need for the underlying residual or boundary value problem.) Let $u_\mathcal{U}$ be the current approximation, then the recovery-based estimator is defined as the L^2 norm of the difference between the direct and postprocessed approximations of the gradient ($\nabla u_\mathcal{U}$ and $G(\nabla u_\mathcal{U})$):

$$\eta_{G,K} = \|G(\nabla u_\mathcal{U}) - \nabla u_\mathcal{U}\|_K \quad \text{and} \quad \eta_G = \left(\sum_{K\in\mathcal{T}} \eta_{G,K}^2\right)^{\frac{1}{2}}. \tag{1.1}$$

There are many postprocessing, recovery techniques (see survey article [33] by Zhang and references therein). A simple one is the projection of the direct approximation onto vector-valued continuous finite element space with respect to either a discrete or the L^2 inner product. The popular Zienkiewicz–Zhu (ZZ) estimator [35, 36] can be viewed as one based on the discrete L^2 projection (see [28]). For estimators based on the L^2 projection, see, e.g., [10] and references therein. By employing several multigrid smoothing after this L^2 projection, Bank and Xu were able to prove that the resulting estimator is asymptotically exact in [3] on irregular meshes by establishing a superconvergence result of the recovered gradient. See also [17, 27, 32] for asymptotically exact estimators based on different recovery techniques on irregular meshes. For higher-order elements, Bank, Xu, and Zheng [4] and Naga and Zhang [21] studied recovery-based estimators and established their asymptotic exactness assuming that the solution of the underlying problem is sufficient smooth. A summary

*Received by the editors March 3, 2008; accepted for publication (in revised form) March 4, 2009; published electronically June 10, 2009. This work was supported in part by the National Science Foundation under grants DMS-0511430 and DMS-0810855.
http://www.siam.org/journals/sinum/47-3/71740.html
†Department of Mathematics, Purdue University, West Lafayette, IN 47907-2067 (zcai@math.purdue.edu, zhang@math.purdue.edu).

and bibliographical remarks on the recovery-based error estimators may be found in [2, 10].

Estimators of the recovery type possess a number of attractive features that have led to their popularity. In particular, their ease of implementation, generality, and ability to produce quite accurate estimators have led to their widespread adoption, especially in the engineering community. However, for applications with nonsmooth solutions like elliptic interface problems, it is well known [3] that they overrefine regions where there are no error, and hence, they fail to reduce the global error. This is shown by Ovall in [23, 24] through several interesting and realistic examples. To overcome this difficulty, one often applies the method on each subdomain separately. For reasons why this local approach is not favorable, see detailed discussions in [24]. More importantly, the local approach fails when triangulations do not align with interfaces, which occurs when interfaces are curves/surfaces or have unknown locations.

One of the purposes of this paper is to develop a new recovery-based error estimator that completely resolve this difficulty for the interface problem using a global approach. To do so, we notice that the normal component of the gradient and the tangential components of the flux are discontinuous across interfaces. Hence, recovering either the gradient or the flux in a continuous finite element space means using a continuous function to approximate a discontinuous function. The error of such an approximation could be arbitrarily large at where the approximated function is discontinuous. This is the sole reason for the failure of the existing recovery-based estimators. In order to overcome this difficulty, we recover the flux in the $H(\mathrm{div})$ conforming finite element spaces such as those of Raviart–Thomas (RT) or Brezzi–Douglas–Marini (BDM) [8]. More specifically, we introduce two flux recovery procedures based on the global (weighted) L^2 projection onto and on the local averaging in the RT or BDM spaces. The resulting recovery-based (implicit and explicit) estimators are then analyzed by establishing the reliability and efficiency bounds, where the efficiency bound for the implicit estimator is global. Both the reliability and efficiency constants are proved to be independent of the size of jumps, and hence, the estimators introduced in this paper are robust with respect to the diffusion coefficients. Numerically, for a benchmark test problem, we show that our estimators do not overrefine regions along interfaces and that they are not subject to the constraint on the monotonicity of the diffusion coefficient in our theory. Moreover, we show numerically that the implicit estimator introduced in this paper is accurate and robust with respect to triangulations that do not align with interfaces. Hence, it has a potential to treat problems with interfaces being curves/surfaces or having unknown locations.

The other purpose of the paper is to study adaptive finite element method through local mesh refinement based on either the recovery-based estimators or the edge estimator. Recently, there has been intensive study on the convergence of the adaptive method for conforming finite elements (see, e.g., [16, 20, 22]) and also for nonconforming and mixed elements. Dörfler in [16] proved the first convergence result for the Possion equation under the assumption that the initial mesh is fine enough. His adaptive method is based on the edge estimator. For the adaptive method using the residual-based estimator, this fine initial mesh requirement was removed by Morin, Nochetto, and Siebert in [20] by introducing a marking strategy based on the data oscillation. However, to guarantee that the data oscillation monotonically decreases, the refined element has to contain a node of the finer mesh in its interior. Very recently, the convergence of the standard adaptive method based on the explicit residual-based estimator is established by Cascon et al. [13] without the marking strategy on the data oscillation and hence, without the interior node requirement. In this paper, we

introduce an additional marking strategy based on the weighted element residual and establish the convergence of the resulting adaptive method based on either the explicit recovery-based estimator introduced in this paper or the edge estimator in [12]. Assumptions on the initial mesh and on the interior node are not required. Furthermore, the rate of convergence is independent of the size of jumps.

This paper is organized as follows. The interface problem and its conforming linear finite element approximation are described in section 2. The recovery procedure and the resulting recovery-based a posteriori error estimator are introduced in section 3. Theoretical analysis and numerical experiments are presented in sections 4 and 5, respectively. Finally, the adaptive finite element method is described and analyzed in section 6.

2. Problem and finite element approximation. Consider the following interface problem

$$-\nabla \cdot (\alpha(x)\nabla u) = f \quad \text{in } \Omega, \tag{2.1}$$

with boundary conditions

$$u = g \quad \text{on } \Gamma_D \quad \text{and} \quad \mathbf{n} \cdot (\alpha \nabla u) = 0 \quad \text{on } \Gamma_N, \tag{2.2}$$

where the symbols $\nabla\cdot$ and ∇ stand for the divergence and gradient operators, respectively, f and g are given scalar-valued functions, Ω is a bounded polygonal domain in $\Re^d$ ($d = 2$ or 3), with boundary $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$, $\mathbf{n} = (n_1, \ldots, n_d)$ is the outward unit vector normal to the boundary, and $\alpha(x)$ is positive and piecewise constant on polygonal subdomains of Ω with possible large jumps across subdomain boundaries (interfaces):

$$\alpha(x) = \alpha_i > 0 \quad \text{in } \Omega_i$$

for $i = 1, \ldots, n$. Here, $\{\Omega_i\}_{i=1}^n$ is a partition of the domain Ω with Ω_i being an open polygonal domain. For simplicity, we consider only homogeneous Neumann boundary conditions and piecewise linear data g. Also, we assume that Γ_D is not empty (i.e., $\text{mes}\,(\Gamma_D) \neq 0$).

We use the standard notations and definitions for the Sobolev spaces $H^s(\Omega)^d$ and $H^s(\partial\Omega)^d$ for $s \geq 0$. The standard associated inner products are denoted by $(\cdot, \cdot)_{s,\Omega}$ and $(\cdot, \cdot)_{s,\partial\Omega}$, and their respective norms are denoted by $\|\cdot\|_{s,\Omega}$ and $\|\cdot\|_{s,\partial\Omega}$. (We suppress the superscript d because the dependence on dimension will be clear by context. We also omit the subscript Ω from the inner product and norm designation when there is no risk of confusion.) For $s = 0$, $H^s(\Omega)^d$ coincides with $L^2(\Omega)^d$. In this case, the inner product and norm will be denoted by $\|\cdot\|$ and $(\cdot, \cdot)$, respectively. We will also use the energy norm denoted by

$$|\!|\!| v |\!|\!| = |\!|\!| v |\!|\!|_\Omega = \left\| \alpha^{1/2} \nabla v \right\|_{0,\Omega}.$$

Let

$$H^1_{g,D}(\Omega) := \left\{ v \in H^1(\Omega) \,:\, v = g \text{ on } \Gamma_D \right\}.$$

The corresponding variational form of system (2.1) is to find $u \in H^1_{g,D}(\Omega)$ such that

$$a(u, v) = f(v) \quad \forall\, v \in H^1_{0,D}(\Omega), \tag{2.3}$$

where the bilinear and linear forms are defined by

$$a(u, v) = (\alpha(x)\nabla u, \nabla v) \quad \text{and} \quad f(v) = (f, v),$$

respectively.

For simplicity of presentation, consider only triangular and tetrahedra elements in the respective two and three dimensions. Let $\mathcal{T} = \{K\}$ be a finite element partition of the domain Ω. Assume that the triangulation $\mathcal{T}$ is regular (see [15]); i.e., for all $K \in \mathcal{T}$, there exists a positive constant κ such that

$$h_K \leq \kappa\, \rho_K,$$

where h_K denotes the diameter of the element K and ρ_K the diameter of the largest circle that may be inscribed in K. Note that the assumption of the regularity does not exclude highly, locally refined meshes. Furthermore, assume that interfaces

$$F = \{\partial\Omega_i \cap \partial\Omega_j \,|\, i, j = 1, \ldots, n\}$$

do not cut through any element $K \in \mathcal{T}$. (This assumption is needed for analysis and for explicit estimators but not for implicit estimators introduced in this paper.)

Let $P_k(K)$ be the space of polynomials of degree k on element K. Denote the continuous piecewise linear finite element space associated with the triangulation $\mathcal{T}$ by

$$\mathcal{U} = \left\{v \in H^1(\Omega) :\, v|_K \in P_1(K) \quad \forall\, K \in \mathcal{T}\right\}.$$

Let $\mathcal{U}_g = \{v \in \mathcal{U} : v = g \text{ on } \Gamma_D\}$, then the finite element approximation of (2.3) is to find $u_{\mathcal{U}} \in \mathcal{U}_g$ such that

$$a(u_{\mathcal{U}}, v) = f(v) \quad \forall\, v \in \mathcal{U}_0. \tag{2.4}$$

Define

$$\alpha_{\min} = \min_{1\leq i\leq n} \alpha_i \quad \text{and} \quad \alpha_{\max} = \max_{1\leq i\leq n} \alpha_i.$$

It is well known (see, e.g., [6]) that if the solution u is in $H^s(\Omega)$, $1 \leq s \leq 2$, then the following a priori error estimate holds:

$$\|\!|u - u_{\mathcal{U}}|\!\|_\Omega \leq C \left(\sum_{i=1}^{n} \left\|h^{s-1}\alpha^{1/2}\nabla u\right\|^2_{s-1,\,\Omega_i}\right)^{1/2}, \tag{2.5}$$

with

$$\left\|h^{s-1}\alpha\nabla u\right\|_{s-1,\,\Omega_i} = \left(\sum_{K\in\mathcal{T}\cap\Omega_i} h_K^{2(s-1)}\alpha_K \|\nabla u\|^2_{s-1,K}\right)^{1/2}.$$

Here and thereafter, we use C with or without subscripts in this paper to denote a generic positive constant, possibly different at different occurrences, that is independent of the mesh parameter h_K and the ratio $\alpha_{\max}/\alpha_{\min}$ but may depend on the domain Ω.

3. Flux recovery and error estimator. The flux defined by

$$\boldsymbol{\sigma} = -\alpha(x)\nabla u \quad \text{in } \Omega \tag{3.1}$$

is an important physical quantity which is often the primary concern in practice. For the interface problem in (2.1) with $f \in L^2(\Omega)$, it is easy to see that the normal component of the flux is continuous, but its tangential component is discontinuous across the interfaces. This type of vector-valued functions may be precisely characterized by the following space:

$$H(\mathrm{div};\Omega) = \left\{\boldsymbol{\tau} \in L^2(\Omega)^d \,:\, \nabla\cdot\boldsymbol{\tau} \in L^2(\Omega)\right\} \subset L^2(\Omega)^d,$$

which is a Hilbert space under the norm

$$\|\boldsymbol{\tau}\|_{H(\mathrm{div};\,\Omega)} = \left(\|\boldsymbol{\tau}\|_{0,\Omega}^2 + \|\nabla\cdot\boldsymbol{\tau}\|_{0,\Omega}^2\right)^{\frac{1}{2}}.$$

Denote its subspace by

$$\Sigma = H_N(\mathrm{div};\Omega) = \{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) \,:\, \mathbf{n}\cdot\boldsymbol{\tau} = 0 \text{ on } \Gamma_N\}.$$

In (3.1), dividing by $\alpha(x)$, multiplying a test function $\boldsymbol{\tau}$, and integrating over the domain Ω give the following variational problem: find $\boldsymbol{\sigma} \in \Sigma$ such that

$$b(\boldsymbol{\sigma},\,\boldsymbol{\tau}) = u(\boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau} \in \Sigma, \tag{3.2}$$

where bilinear form $b(\cdot\,,\,\cdot)$ and linear form $u(\cdot)$ are defined by

$$b(\boldsymbol{\sigma},\,\boldsymbol{\tau}) = \left(\alpha^{-1}\boldsymbol{\sigma},\,\boldsymbol{\tau}\right) \quad \forall\, (\boldsymbol{\sigma},\,\boldsymbol{\tau}) \in \Sigma\times\Sigma \quad \text{and} \quad u(\boldsymbol{\tau}) = -(\nabla u,\,\boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau}\in\Sigma,$$

respectively.

3.1. Flux recovery. The recovery procedure introduced in this paper is based on the conforming finite element approximation to the variational problem in (3.2). There are several families of the $H(\mathrm{div};\Omega)$ confirming finite element spaces (see, e.g., [8, 26]). We consider only RT and BDM elements for simplicity.

Denote the local lowest-order RT and BDM spaces on element $K \in \mathcal{T}$ by

$$RT_0(K) = P_0(K)^d + \mathbf{x}\,P_0(K) \quad \text{and} \quad BDM_1(K) = P_1(K)^d,$$

respectively, where $\mathbf{x} = (x_1,\,\ldots,\,x_d)$. Then the standard $H(\mathrm{div};\,\Omega)$ conforming RT and BDM spaces are defined by

$$RT_0 = \{\boldsymbol{\tau} \in \Sigma \,:\, \boldsymbol{\tau}|_K \in RT_0(K) \;\; \forall\, K \in \mathcal{T}\}$$

and

$$BDM_1 = \{\boldsymbol{\tau} \in \Sigma \,:\, \boldsymbol{\tau}|_K \in BDM_1(K) \;\; \forall\, K \in \mathcal{T}\},$$

respectively. For convenience, denote RT_0 and BDM_1 by $\mathcal{V}$. It is well known (see [8]) that $\mathcal{V}$ has the following approximation property:

$$\inf_{\boldsymbol{\tau}\in\, RT_0} \|\boldsymbol{\sigma}-\boldsymbol{\tau}\|_{H(\mathrm{div};\,\Omega)} \le C\left(\|h\,\boldsymbol{\sigma}\|_{1,\Omega}^2 + \|h\,\nabla\cdot\boldsymbol{\sigma}\|_{1,\Omega}^2\right)^{1/2} \tag{3.3}$$

for $\boldsymbol{\sigma} \in H^1(\Omega)^{m\times m} \cap \Sigma$, with $\nabla\cdot\boldsymbol{\sigma} \in H^1(\Omega)^m$ and

$$\inf_{\boldsymbol{\tau}\in\, BDM_1} \|\boldsymbol{\sigma}-\boldsymbol{\tau}\|_{0,\Omega} \le C\,\left\|h^l\,\boldsymbol{\sigma}\right\|_{l,\Omega} \tag{3.4}$$

for $\boldsymbol{\sigma} \in H^l(\Omega)^{m\times m} \cap \Sigma$ and $l \in [1,\,2]$.

3.1.1. Implicit approximation. Let $\bar{u}_{\mathcal{U}} \in \mathcal{U}$ be an approximation of the exact solution $u \in H^1_{g,D}(\Omega)$ of (2.3), then we recover the flux by solving the following problem: find $\boldsymbol{\sigma}_{\mathcal{V}} \in \mathcal{V}$ such that

$$b(\boldsymbol{\sigma}_{\mathcal{V}}, \boldsymbol{\tau}) = \bar{u}_{\mathcal{U}}(\boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau} \in \mathcal{V}, \tag{3.5}$$

with $\bar{u}_{\mathcal{U}}(\boldsymbol{\tau}) = -(\nabla\, \bar{u}_{\mathcal{U}}, \boldsymbol{\tau})$.

THEOREM 3.1. *Let u and $\boldsymbol{\sigma}_{\mathcal{V}}$ be the solutions of* (2.3) *and* (3.5)*, respectively. Then there exists a positive constant C independent of the ratio $\alpha_{\max}/\alpha_{\min}$ such that the following a priori error bound*

$$\left\|\alpha^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathcal{V}})\right\|_{0,\Omega} \leq C \left(\inf_{\boldsymbol{\tau} \in \mathcal{V}} \left\|\alpha^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\tau})\right\|_{0,\Omega} + |\!|\!| u - \bar{u}_{\mathcal{U}} |\!|\!|_{\Omega} \right) \tag{3.6}$$

holds.

Proof. Denote the true errors of the solution and the flux by

$$e = u - \bar{u}_{\mathcal{U}} \quad \text{and} \quad \mathbf{E} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathcal{V}},$$

respectively. Difference between (3.2) and (3.5) gives the following error equation:

$$b(\mathbf{E}, \boldsymbol{\tau}) = u(\boldsymbol{\tau}) - \bar{u}_{\mathcal{U}}(\boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau} \in \mathcal{V},$$

where

$$u(\boldsymbol{\tau}) - \bar{u}_{\mathcal{U}}(\boldsymbol{\tau}) = -(\nabla\, e, \boldsymbol{\tau}). \tag{3.7}$$

Using the above error equation and the Cauchy–Schwarz inequality yields

$$\begin{aligned} \left\|\alpha^{-1/2}\mathbf{E}\right\|^2_{0,\Omega} &= b(\mathbf{E}, \mathbf{E}) = b(\mathbf{E}, \boldsymbol{\sigma} - \boldsymbol{\tau}) + b(\mathbf{E}, \boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}}) \\ &\leq \left\|\alpha^{-1/2}\mathbf{E}\right\|_{0,\Omega} \left\|\alpha^{-1/2}(\boldsymbol{\sigma} - \boldsymbol{\tau})\right\|_{0,\Omega} + u(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}}) - \bar{u}_{\mathcal{U}}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}}) \end{aligned}$$

$\forall\, \boldsymbol{\tau} \in \mathcal{V}$. By the Cauchy–Schwarz and triangle inequalities, we have

$$\begin{aligned} |u(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}}) - \bar{u}_{\mathcal{U}}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}})| &= |(\nabla\, e, \boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}})| \\ &\leq \left\|\alpha^{1/2}\nabla\, e\right\|_{0,\Omega} \left\|\alpha^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{V}})\right\|_{0,\Omega} \\ &\leq |\!|\!| e |\!|\!|_{\Omega} \left(\left\|\alpha^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma})\right\|_{0,\Omega} + \left\|\alpha^{-1/2}\mathbf{E}\right\|_{0,\Omega} \right). \end{aligned}$$

Now, the error bound in (3.6) follows from the above inequalities and the ϵ inequality ($ab \leq \frac{1}{2\epsilon}a^2 + \frac{\epsilon}{2}b^2$). ∎

3.1.2. Explicit approximation. In this subsection, we describe an explicit approximation of the flux based on the RT_0. Denote the set of all edges/faces of the triangulation by

$$\mathcal{E} := \mathcal{E}_{\Omega} \cup \mathcal{E}_{D} \cup \mathcal{E}_{N},$$

where $\mathcal{E}_{\Omega}$ is the set of all interior element edges/faces and $\mathcal{E}_{D}$ and $\mathcal{E}_{N}$ are the set of boundary edges/faces belonging to the respective Γ_D and Γ_N. For each $e \in \mathcal{E}$, denote a unit vector normal to e by $\mathbf{n}_e$. When $e \in \mathcal{E}_{D} \cup \mathcal{E}_{N}$, assume that $\mathbf{n}_e$ is the

unit outward normal vector. The nodal basis function $\boldsymbol{\phi}_e$ of RT_0 corresponding to $e \in \mathcal{E}_\Omega \cup \mathcal{E}_D$ is characterized by

$$\boldsymbol{\phi}_e \cdot \mathbf{n}_{e'}|_{e'} = \delta_{e\,e'} \quad \forall\ e,\, e' \in \mathcal{E}, \tag{3.8}$$

where $\delta_{e\,e'}$ is the Kronecker delta. For each interior edge/face $e \in \mathcal{E}_\Omega$, let K_e^+ and K_e^- be the two elements sharing the common edge/face e such that the unit outward normal vector of K_e^+ coincides with $\mathbf{n}_e$. Let $\mathbf{a}_e^\pm$ be the vertices of $K_e^\pm$ opposite to e. Then the nodal basis function of RT_0 corresponding to $e \in \mathcal{E}_\Omega$ has of the form

$$\boldsymbol{\phi}_e(\mathbf{x}) := \begin{cases} \dfrac{|e|}{d\,|K_e^+|}\left(\mathbf{x} - \mathbf{a}_e^+\right) & \text{for } \ x \in K_e^+, \\ -\dfrac{|e|}{d\,|K_e^-|}\left(\mathbf{x} - \mathbf{a}_e^-\right) & \text{for } \ x \in K_e^-, \\ 0 & \text{elsewhere,} \end{cases} \tag{3.9}$$

where $|e|$ and $|K_e^\pm|$ are the $d-1$ and d measure of e and $K_e^\pm$, respectively. For boundary edge/face $e \in \mathcal{E}_D$, the corresponding nodal basis function is

$$\boldsymbol{\phi}_e(\mathbf{x}) := \begin{cases} \dfrac{|e|}{d\,|K_e^+|}\left(\mathbf{x} - \mathbf{a}_e^+\right) & \text{for } \ x \in K_e^+, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.10}$$

Let $\boldsymbol{\tau} = -\alpha(x)\nabla\,\bar{u}_{\mathcal{U}}$, define its approximation $\hat{\boldsymbol{\sigma}}_{RT_0}(\bar{u}_{\mathcal{U}})$ in RT_0 by

$$\hat{\boldsymbol{\sigma}}_{RT_0}(\bar{u}_{\mathcal{U}}) = \sum_{e \in \mathcal{E}_\Omega \cup \mathcal{E}_D} \hat{\sigma}_e\, \boldsymbol{\phi}_e(\mathbf{x}), \tag{3.11}$$

where $\hat{\sigma}_e$ is the normal component of $\hat{\boldsymbol{\sigma}}_{RT_0}$ on $e \in \mathcal{E}_\Omega \cup \mathcal{E}_D$ defined by

$$\hat{\sigma}_e := \begin{cases} \gamma_e \left(\boldsymbol{\tau}|_{K_e^+} \cdot \mathbf{n}_e\right)\big|_e + (1-\gamma_e)\left(\boldsymbol{\tau}|_{K_e^-} \cdot \mathbf{n}_e\right)\big|_e & \text{for } e \in \mathcal{E}_\Omega, \\ \boldsymbol{\tau}|_e \cdot \mathbf{n}_e & \text{for } e \in \mathcal{E}_D, \end{cases} \tag{3.12}$$

for some constant $\gamma_e \in [0,\, 1]$. We choose

$$\gamma_e = \frac{\sqrt{\alpha_{K_e^-}}}{\sqrt{\alpha_{K_e^+}} + \sqrt{\alpha_{K_e^-}}} \tag{3.13}$$

to ensure that the efficiency constant is independent of the ratio $\alpha_{\max}/\alpha_{min}$ (see Theorem 4.3).

Remark 3.1. If the normal component of $\boldsymbol{\tau}$ is continuously across edge/face e, then the normal component of its approximation defined in (3.12) equals to the normal component of $\boldsymbol{\tau}$; i.e.,

$$\hat{\sigma}_e = \boldsymbol{\tau}|_e \cdot \mathbf{n}_e.$$

3.2. Error estimators. Let $\boldsymbol{\sigma}_{\mathcal{V}}$ be the solution of problem (3.5), for any element $K \in \mathcal{T}$, define the following local a posteriori error indicator by

$$\eta_{\mathcal{V},K} = \left\| \alpha^{-1/2}\boldsymbol{\sigma}_{\mathcal{V}} + \alpha^{1/2}\nabla\bar{u}_{\mathcal{U}} \right\|_{0,K}. \tag{3.14}$$

Then the corresponding global a posteriori error estimator is

$$\eta_{\mathcal{V}} = \left(\sum_{K \in \mathcal{T}} (\eta_{\mathcal{V},K})^2 \right)^{1/2} = \left\| \alpha^{-1/2} \boldsymbol{\sigma}_{\mathcal{V}} + \alpha^{1/2} \nabla \bar{u}_{\mathcal{U}} \right\|_{0,\Omega}. \tag{3.15}$$

It is easy to see that

$$\eta_{\mathcal{V}} = \min_{\boldsymbol{\tau} \in \mathcal{V}} \left\| \alpha^{-1/2} \boldsymbol{\tau} + \alpha^{1/2} \nabla \bar{u}_{\mathcal{U}} \right\|_{0,\Omega}. \tag{3.16}$$

This estimator requires numerical solution of a system of linear equations with a mass matrix. Such a system can be solved very efficiently by several sweeps of the Jacobi iteration or better by preconditioned conjugate gradient method with the Jacobi preconditioner. Note that this estimator does not require the alignment between triangulations and the interfaces of the underlying problem. Hence, it can be applied to problems with interfaces being curves/surfaces or having unknown locations.

Next, based on the explicit approximation in (3.11), we define explicit local a posteriori error indicator by

$$\hat{\eta}_{RT_0,K} = \left\| \alpha^{-1/2} \hat{\boldsymbol{\sigma}}_{RT_0} + \alpha^{1/2} \nabla \bar{u}_{\mathcal{U}} \right\|_{0,K} \tag{3.17}$$

for any $K \in \mathcal{T}$ and explicit global a posteriori error estimator by

$$\hat{\eta}_{RT_0} = \left(\sum_{K \in \mathcal{T}} (\hat{\eta}_{RT_0,K})^2 \right)^{1/2} = \left\| \alpha^{-1/2} \hat{\boldsymbol{\sigma}}_{RT_0} + \alpha^{1/2} \nabla \bar{u}_{\mathcal{U}} \right\|_{0,\Omega}. \tag{3.18}$$

This explicit estimator is similar to that introduced by Luce and Wolhmuth [19], but they differ in the recovery procedure. The latter is more complicated, expensive, and probably accurate than the former. Nevertheless, both the estimators are subject to the alignment assumption between triangulations and interfaces. We study this explicit estimator because it is used in our analysis and it is probably appealing to the engineering community.

4. Reliability and efficiency bounds. In this section, we establish reliability and efficiency bounds for both implicit and explicit estimators under the assumption that triangulations align with interfaces.

4.1. Clément-type interpolation. Clément-type interpolation operators have been intensively studied in the literature (see, e.g., [6, 25]), and they are often used for establishing the reliability bound of a posteriori error estimators. In this section, we follow [6] to define a weighted Clément-type interpolation operator and to state its approximation and stability properties.

To this end, denote by $\mathcal{N}$ and $\mathcal{N}_K$ the sets of all vertices of the triangulation $\mathcal{T}$ and of element $K \in \mathcal{T}$, respectively. For any $z \in \mathcal{N}$, denote by ϕ_z the nodal basis function, let $\omega_z = \operatorname{suppt}(\phi_z)$, and denote by $\hat{\omega}_z$ the union of elements in ω_z, where the coefficient α_K achieves the maximum for $K \subset \omega_z$. For a given function v, define its weighted average over $\hat{\omega}_z$ by

$$⨍_{\hat{\omega}_z} v \, dx = \frac{\int_{\hat{\omega}_z} v \, \phi_z \, dx}{\int_{\hat{\omega}_z} \phi_z \, dx}. \tag{4.1}$$

Now, following [6], define the interpolation operator $I: L^2(\Omega) \to \mathcal{U}_0$ by

$$Iv = \sum_{z \in \mathcal{N}} (\pi_z v)\phi_z(x), \tag{4.2}$$

where the nodal value at z is defined by

$$(Iv)(z) = \pi_z v = \begin{cases} \fint_{\hat{\omega}_z} v\, dx & z \in \mathcal{N} \setminus \Gamma_D, \\ 0 & z \in \mathcal{N} \cap \Gamma_D. \end{cases}$$

Note that the weighted average in (4.1) implies

$$(1,\, \phi_z(v - \pi_z v))_{\hat{\omega}_z} = \int_{\hat{\omega}_z} \phi_z(v - \pi_z v)\, dx = 0 \quad \forall\, z \in \mathcal{N} \setminus \Gamma_D, \tag{4.3}$$

which will be used in the subsequent section in order to handle a term involving the right-hand side function f.

In this and next subsections, assume that the Hypothesis 2.7 in [6] holds. That is, assume that for any two different subdomains $\bar{\Omega}_i$ and $\bar{\Omega}_j$, which share at least one point, there is a connected path passing from $\bar{\Omega}_i$ to $\bar{\Omega}_j$ through adjacent subdomains such that the diffusion coefficient α is monotone along this path. This assumption is weakened to the quasi-monotonicity in [25].

LEMMA 4.1. *For any $K \in \mathcal{T}$, $z \in \mathcal{N}_K$, and $v \in H^1_{0,D}(\Omega)$, there exists a positive constant C independent of the ratio $\alpha_{\max}/\alpha_{\min}$ such that*

$$\|(v - \pi_z v)\phi_z\|_{0,K} \le C\, h_K \alpha_K^{-1/2} |\!|\!| v |\!|\!|_{\Delta_K} \tag{4.4}$$

and that

$$\|(v - \pi_z v)\nabla\phi_z\|_{0,K} \le C\, \alpha_K^{-1/2} |\!|\!| v |\!|\!|_{\Delta_K}, \tag{4.5}$$

where Δ_K is the union of all elements that share at least one vertex with K.

Proof. Using the following inequalities

$$\|\nabla\phi_z\|_{\infty,K} \le C\, h_K^{-1} \quad \text{and} \quad \|\nabla\phi_z\|_{0,K} \le C\, h_K^{(d-2)/2},$$

(4.4) and (4.5) can be proved in a similar fashion as that in [6]. ∎

LEMMA 4.2. *For any $K \in \mathcal{T}$ and $v \in H^1_{0,D}(\Omega)$, the following estimates hold:*

$$\|v - Iv\|_{0,K} \le C\, h_K \alpha_K^{-1/2} |\!|\!| v |\!|\!|_{\Delta_K} \tag{4.6}$$

and

$$\|\nabla(v - Iv)\|_{0,K} \le C\, \alpha_K^{-1/2} |\!|\!| v |\!|\!|_{\Delta_K}. \tag{4.7}$$

Proof. Using the identity $\sum_{z \in \mathcal{N}_K} \phi_z(x) = 1$ in K, we have

$$\|v - Iv\|_{0,K} = \|\sum_{z \in \mathcal{N}_K} (v - \pi_z v)\phi_z\|_{0,K} \le \sum_{z \in \mathcal{N}_K} \|(v - \pi_z v)\phi_z\|_{0,K} \tag{4.8}$$

and

$$\nabla(v - Iv)\big|_K = \nabla \sum_{z \in \mathcal{N}_K} (v - \pi_z v)\phi_z = \sum_{z \in \mathcal{N}_K} \phi_z \nabla v + \sum_{z \in \mathcal{N}_K} (v - \pi_z v)\nabla\phi_z. \tag{4.9}$$

Equation (4.6) is a direct consequence of (4.8), the triangle inequality, and (4.4). Equation (4.7) follows from the triangle inequality, the fact that $\|\phi_z\|_{\infty,K} \le 1$, and (4.5) that

$$\begin{aligned}\|\nabla(v - Iv)\|_{0,K} &\le \sum_{z\in\mathcal{N}_K} \|\phi_z \nabla v\|_{0,K} + \sum_{z\in\mathcal{N}_K} \|(v - \pi_z v)\nabla\phi_z\|_{0,K} \\ &\le d\,\|\nabla v\|_{0,K} + C \sum_{z\in\mathcal{N}_K} \alpha_K^{-1/2} |\!|\!| v |\!|\!|_{\Delta_K} \le C\,\alpha_K^{-1/2} |\!|\!| v |\!|\!|_{\Delta_K}.\end{aligned}$$

This completes the proof of the lemma. ∎

4.2. Reliability. Let

$$H_f = \left(\sum_{z\in\mathcal{N}\cap(F\cup\Gamma_D)} \sum_{K\subset\omega_z} \alpha_K^{-1} h_K^2 \,\|f\|_{0,K}^2 + \sum_{z\in\mathcal{N}\setminus(F\cup\Gamma_D)} \sum_{K\subset\omega_z} \alpha_K^{-1} h_K^2 \left\| f - \fint_{\omega_z} f\,dx \right\|_{0,K}^2 \right)^{1/2}$$

and

$$\hat{H}_f = \left\| \alpha^{-1/2} h f \right\| = \left(\sum_{K\in\mathcal{T}} \alpha_K^{-1} h_K^2 \,\|f\|_{0,K}^2 \right)^{1/2}.$$

Remark 4.1. The second term in H_f is a higher-order term for $f \in L^2(\Omega)$ and so is the first term for $f \in L^p(\Omega)$, with $p > 2$ (see [12]).

LEMMA 4.3. *For any $v \in H^1_{0,D}(\Omega)$, there exists a positive constant C independent of the ratio $\alpha_{\max}/\alpha_{\min}$ such that*

$$|(f,\, v - Iv)| \le C\, H_f\, |\!|\!| v |\!|\!| \tag{4.10}$$

and that

$$|(f,\, v - Iv)| \le C\, \hat{H}_f\, |\!|\!| v |\!|\!|. \tag{4.11}$$

Proof. Equation (4.3) and the fact that $\hat{\omega}_z = \omega_z$ for $z \in \mathcal{N} \setminus (F \cup \Gamma_D)$ give

$$(1,\, \phi_z(v - \pi_z v))_{\omega_z} = 0 \quad \forall\, z \in \mathcal{N} \setminus (F \cup \Gamma_D),$$

which, together with $\sum_{z\in\mathcal{N}} \phi_z(x) = 1$ in Ω, gives

$$\begin{aligned}(f,\, v - Iv) &= \sum_{z\in\mathcal{N}} (f,\, (v - \pi_z v)\phi_z)_{\omega_z} \\ &= \sum_{z\in\mathcal{N}\cap(F\cup\Gamma_D)} (f,\, (v - \pi_z v)\phi_z)_{\omega_z} + \sum_{z\in\mathcal{N}\setminus(F\cup\Gamma_D)} (f,\, (v - \pi_z v)\phi_z)_{\omega_z} \\ &= \sum_{z\in\mathcal{N}\cap(F\cup\Gamma_D)} (f,\, (v - \pi_z v)\phi_z)_{\omega_z} \\ &\quad + \sum_{z\in\mathcal{N}\setminus(F\cup\Gamma_D)} \left(f - \fint_{\omega_z} f\,dx,\, (v - \pi_z v)\phi_z \right)_{\omega_z} \\ &= \sum_{z\in\mathcal{N}\cap(F\cup\Gamma_D)} \sum_{K\subset\omega_z} (f,\, (v - \pi_z v)\phi_z)_K \\ &\quad + \sum_{z\in\mathcal{N}\setminus(F\cup\Gamma_D)} \sum_{K\subset\omega_z} \left(f - \fint_{\omega_z} f\,dx,\, (v - \pi_z v)\phi_z \right)_K.\end{aligned}$$

Now, (4.10) is a direct consequence of the Cauchy–Schwarz inequality and (4.4).

Equation (4.11) follows from the identity

$$(f,\, v - Iv) = \sum_{z\in\mathcal{N}} \sum_{K\subset\omega_z} (f,\, (v-\pi_z v)\phi_z)_K,$$

the Cauchy–Schwarz inequality, and (4.4). ∎

THEOREM 4.1. *Assume that $\bar{u}_{\mathcal{U}} = u_{\mathcal{U}}$ is the solution of* (2.4)*. Then the estimator $\eta_{\mathcal{V}}$ defined in* (3.15) *satisfies the following global reliability bound:*

$$\|\!|e|\!\| \le C\ (\eta_{\mathcal{V}} + H_f) \tag{4.12}$$

and the following bound:

$$\|\!|e|\!\|^2 \le C_r\ \left(\eta_{\mathcal{V}}^2 + \hat{H}_f^2\right), \tag{4.13}$$

where C and C_r are constants independent of the ratio $\alpha_{\max}/\alpha_{\min}$.

Proof. It follows from the orthogonality property of the finite element solution, integration by parts, (2.1), and the Cauchy–Schwarz inequality that

$$\begin{aligned}
\|\!|e|\!\|^2 &= a(e,\, e-Ie) = (\alpha\nabla\,(u-u_{\mathcal{U}}), \nabla(e-Ie)) \\
&= (\alpha\nabla u + \boldsymbol{\sigma}_{\mathcal{V}}, \nabla\,(e-Ie)) - (\boldsymbol{\sigma}_{\mathcal{V}} + \alpha\nabla u_{\mathcal{U}},\, \nabla\,(e-Ie)) \\
&\le (f - \nabla\cdot\boldsymbol{\sigma}_{\mathcal{V}},\, e-Ie) + \eta_{\mathcal{V}}\,\|\!|e-Ie|\!\|,
\end{aligned}$$

which, combining with the fact that $\nabla\cdot(\alpha(x)\nabla u_{\mathcal{U}})|_K = 0\ \forall\ K\in\mathcal{T}$, (4.7), (4.10), and the Cauchy–Schwarz inequality, implies

$$\begin{aligned}
\|\!|e|\!\|^2 &\le (f,\, e - I\,e) - \sum_{K\in\mathcal{T}} (\nabla\cdot(\boldsymbol{\sigma}_{\mathcal{V}} + \alpha\nabla u_{\mathcal{U}}),\, e - I\,e)_K + C\,\eta_{\mathcal{V}}\,\|\!|e|\!\| \\
&\le C\,(H_f + \eta_{\mathcal{V}})\,\|\!|e|\!\| + \left(\sum_{K\in\mathcal{T}} h_K^2\,\left\|\nabla\cdot\left(\alpha^{-1/2}\boldsymbol{\sigma}_{\mathcal{V}} + \alpha^{1/2}\nabla u_{\mathcal{U}}\right)\right\|_{0,K}^2\right)^{1/2} \\
&\quad \left(\sum_{K\in\mathcal{T}} h_K^{-2}\alpha_K \|e-Ie\|_{0,K}^2\right)^{1/2}.
\end{aligned}$$

Using the inverse inequality and (4.6), we then have

$$\|\!|e|\!\|^2 \le C\ (H_f + \eta_{\mathcal{V}})\,\|\!|e|\!\| + C\,\eta_{\mathcal{V}}\,\|\!|e|\!\| = C\,(H_f + \eta_{\mathcal{V}})\,\|\!|e|\!\|,$$

which leads to (4.12).

Using (4.11) instead of (4.10), one can prove the validity of (4.13) in the same way. ∎

THEOREM 4.2. *Under the same assumption of Theorem* 4.1*, the explicit estimator $\hat{\eta}_{RT_0}$ defined in* (3.18) *satisfies the following global reliability bound:*

$$\|\!|e|\!\| \le C\ \left(\hat{\eta}_{RT_0} + H_f\right) \tag{4.14}$$

and the following bound:

$$\|\!|e|\!\|^2 \le C_r\ \left(\hat{\eta}_{RT_0}^2 + \hat{H}_f^2\right), \tag{4.15}$$

where C and C_r are constants independent of the ratio $\alpha_{\max}/\alpha_{\min}$.

Proof. The reliability bounds in (4.14) and (4.15) are an immediate consequence of the respective (4.12) and (4.13) and the fact that

$$\hat{\eta}_{RT_0} \geq \min_{\boldsymbol{\tau} \in RT_0} \left\| \alpha(x)^{-1/2} \boldsymbol{\tau} + \alpha(x)^{1/2} \nabla u_{\mathcal{U}} \right\|_{0,\Omega} = \eta_{RT_0}.$$

This completes the proof of the theorem. ∎

4.3. Efficiency. For any $e \in \mathcal{E}_\Omega$ and any vector-valued function $\boldsymbol{\rho}$ that is piecewise constant with respect to the triangulation $\mathcal{T}$, denote the jump of the normal component of $\boldsymbol{\rho}$ across $e = K_e^+ \cap K_e^-$ by

$$J_e(\boldsymbol{\rho}) = [\boldsymbol{\rho} \cdot \mathbf{n}_e] = (\boldsymbol{\rho}|_{K_e^+} - \boldsymbol{\rho}|_{K_e^-}) \cdot \mathbf{n}_e \,|_e.$$

For any $e \in \mathcal{E} \backslash \mathcal{E}_\Omega$, set

$$J_e(\boldsymbol{\rho}) = 0.$$

For any $e \in \mathcal{E}$, denote by ω_e the union of all elements that share edge/face e. Define a modification of the edge error estimator as follows:

$$\eta_E := \left(\sum_{e \in \mathcal{E}} \eta_e^2 \right)^{1/2}, \qquad \text{with} \quad \eta_e = \left(\frac{2h_e}{\alpha_{K_e^+} + \alpha_{K_e^-}} \int_e |J_e(\alpha \nabla u_{\mathcal{U}})|^2 \, ds \right)^{1/2}, \tag{4.16}$$

where h_e is the diameter of edge/face e. Without assumptions on the distribution of the coefficient α, it was proved by Petzoldt (see equation (5.7) in [25]) that there exists a constant $C > 0$ independent of $\alpha_{\max}/\alpha_{\min}$, h_K, and h_e, such that

$$\eta_e^2 \leq C \left(|\!|\!| e |\!|\!|_{\omega_e}^2 + \sum_{K \in \mathcal{T} \cap \omega_e} \frac{h_K^2}{\alpha_{K_e^+} + \alpha_{K_e^-}} \left\| f - \bar{f}_K \right\|_{0,K}^2 \right), \tag{4.17}$$

where $\bar{f}_K$ is the average of f over K:

$$\bar{f}_K = \frac{1}{|K|} \int_K f \, dx.$$

LEMMA 4.4. *For any element $K \in \mathcal{T}$, the constant vector $\boldsymbol{\tau}$ on K has the following representation in RT_0:*

$$\boldsymbol{\tau} = \sum_{e \in \partial K} \tau_{e,K} \, \boldsymbol{\phi}_e(\mathbf{x}), \tag{4.18}$$

where $\tau_{e,K} = (\boldsymbol{\tau}|_K \cdot \mathbf{n}_e)|_e$ is the normal component of $\boldsymbol{\tau}$ on edge e.

Proof. Since constant vector on K belongs to $RT_0(K)$, $\boldsymbol{\tau}$ can be written as

$$\boldsymbol{\tau} = \sum_{e \in \partial K} \tau_{e,K} \, \boldsymbol{\phi}_e(\mathbf{x}).$$

Now, using (3.8) yields $\tau_e = (\boldsymbol{\tau}|_K \cdot \mathbf{n}_e)|_e$ and hence, the lemma. ∎

THEOREM 4.3. *There exists a constant $C > 0$ independent of $\alpha_{\max}/\alpha_{\min}$ such that*

$$\hat{\eta}_{RT_0,K}^2 \leq C \left(|\!|\!| e |\!|\!|_{\omega_K}^2 + \sum_{T \in \mathcal{T} \cap \omega_K} \frac{h_T^2}{\alpha_T} \left\| f - \bar{f}_T \right\|_{0,T}^2 \right), \tag{4.19}$$

where ω_K is the union of elements sharing a common edge/face with K and that

$$\eta_{BDM_1} \le \eta_{RT_0} \le \hat{\eta}_{RT_0} \le C\,|\!|\!|e|\!|\!|_\Omega + C\left(\sum_{K\in\mathcal{T}} \frac{h_K^2}{\alpha_K}\left\|f-\bar{f}_K\right\|_{0,K}^2\right)^{1/2}. \tag{4.20}$$

Proof. For any element $K \in \mathcal{T}$ and for any edge/face $e \in \partial K$, without loss of generality assume that $\mathbf{n}_e$ is the outward unit vector normal to ∂K. Denote by K_e the adjacent element with common edge/face e. Let $\boldsymbol{\tau} = -\alpha\nabla u_{\mathcal{U}}$, then, for any $\mathbf{x} \in K$, (3.11), (3.12), (3.13), and (4.18) give

$$\begin{aligned}
\hat{\boldsymbol{\sigma}}_{RT_0} - \boldsymbol{\tau} &= \sum_{e\in\partial K} (\hat{\sigma}_e - \tau_{e,K})\ \boldsymbol{\phi}_e(\mathbf{x}) \\
&= \sum_{e\in\partial K\setminus\partial\Omega} (1-\gamma_e)\,(\tau_{e,K_e} - \tau_{e,K})\ \boldsymbol{\phi}_e(\mathbf{x}) \\
&= -\sum_{e\in\partial K\setminus\partial\Omega} (1-\gamma_e) J_e(\boldsymbol{\tau})\,\boldsymbol{\phi}_e(\mathbf{x}) = -\sum_{e\in\partial K\setminus\partial\Omega} \frac{\sqrt{\alpha_K}}{\sqrt{\alpha_K}+\sqrt{\alpha_{K_e}}} J_e(\boldsymbol{\tau})\,\boldsymbol{\phi}_e(\mathbf{x}).
\end{aligned}$$

Since $J_e(\boldsymbol{\tau})$ is constant in K and $\|\boldsymbol{\phi}_e(\mathbf{x})\|_{0,K}^2 \le C|K|$, it then follows from the triangle inequality that

$$\begin{aligned}
\hat{\eta}_{RT_0,K}^2 &= \left\|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{RT_0} - \boldsymbol{\tau}\right)\right\|_{0,K}^2 \le C \sum_{e\in\partial K\setminus\partial\Omega} \frac{1}{\left(\sqrt{\alpha_K}+\sqrt{\alpha_{K_e}}\right)^2} \|J_e(\boldsymbol{\tau})\boldsymbol{\phi}_e(\mathbf{x})\|_{0,K}^2 \\
&\le C \sum_{e\in\partial K\setminus\partial\Omega} \frac{1}{\alpha_K+\alpha_{K_e}} |J_e(\boldsymbol{\tau})|^2\,\|\boldsymbol{\phi}_e(\mathbf{x})\|_{0,K}^2 \\
&\le C \sum_{e\in\partial K\setminus\partial\Omega} \eta_e^2, \qquad (4.21)
\end{aligned}$$

which, together with (4.17), implies (4.19).

The first two inequalities in the global efficiency bound (4.20) follow from the fact that $RT_0 \subset BDM_1$ and (3.16), respectively. To prove the third one in (4.20), summing up (4.19) over all $K \in \mathcal{T}$ gives

$$\begin{aligned}
\hat{\eta}_{RT_0}^2 &= \sum_{K\in\mathcal{T}} \left\|\alpha^{-1/2}\hat{\boldsymbol{\sigma}}_{RT_0} + \alpha^{1/2}\nabla u_{\mathcal{U}}\right\|_{0,K}^2 \le C \sum_{K\in\mathcal{T}} \left(|\!|\!|e|\!|\!|_{\omega_K}^2 + \sum_{T\in\mathcal{T}\cap\omega_K} \frac{h_T^2}{\alpha_T}\left\|f-\bar{f}_T\right\|_{0,T}^2\right) \\
&\le C\,|\!|\!|e|\!|\!|^2 + C\left(\sum_{K\in\mathcal{T}} \frac{h_K^2}{\alpha_K}\left\|f-\bar{f}_K\right\|_{0,K}^2\right),
\end{aligned}$$

which, in turn, implies (4.20). This completes the proof the theorem. ∎

5. Numerical experiments. In this section, we report some numerical results for an interface problem with intersecting interfaces used by many authors, e.g., [18, 20, 14], which is considered as a benchmark problem. Let $\Omega = (-1,1)^2$ and

$$u(r,\theta) = r^\beta \mu(\theta)$$

in the polar coordinates at the origin with

$$\mu(\theta) = \begin{cases} \cos((\pi/2-\sigma)\beta)\cdot\cos((\theta-\pi/2+\rho)\beta) & \text{if} \quad 0 \le \theta \le \pi/2, \\ \cos(\rho\beta)\cdot\cos((\theta-\pi+\sigma)\beta) & \text{if} \quad \pi/2 \le \theta \le \pi, \\ \cos(\sigma\beta)\cdot\cos((\theta-\pi-\rho)\beta) & \text{if} \quad \pi \le \theta \le 3\pi/2, \\ \cos((\pi/2-\rho)\beta)\cdot\cos((\theta-3\pi/2-\sigma)\beta) & \text{if} \quad 3\pi/2 \le \theta \le 2\pi, \end{cases}$$

where σ and ρ are numbers. The function $u(r,\theta)$ satisfies the interface equation in (2.1) with $\Gamma_N = \emptyset$, $f = 0$, and

$$\alpha(x) = \begin{cases} R & \text{in } (0,1)^2 \cup (-1,0)^2, \\ 1 & \text{in } \Omega \setminus ([0,1]^2 \cup [-1,0]^2). \end{cases}$$

The numbers β, R, σ, and ρ satisfy some nonlinear relations (e.g., [20, 14]). For example, when $\beta = 0.1$, then

$$R \approx 161.4476387975881, \quad \rho = \pi/4, \quad \text{and} \quad \sigma \approx -14.92256510455152.$$

Note that when $\beta = 0.1$, this is a difficult problem for computation.

Remark 5.1. This problem does not satisfy Hypothesis 2.7 in [6], and the distribution of its coefficients is not quasi monotone.

Starting with a coarse triangulation $\mathcal{T}_0$, a sequence of meshes is generated by using standard adaptive meshing algorithm that adopts the Dörfler's bulk marking strategy, i.e., Marking Strategy E described in section 7.1 of [16] with $\theta_E = 0.2$. The choice of $\theta_E = 0.2$ is not critical but recommended in [14] for better performance. Marked triangles are refined regularly by dividing each into four congruent triangles. Additionally, irregularly refined triangles are needed in order to make the triangulation admissible. For more details on adaptive mesh refinement algorithms, see, e.g., [11, 7].

Note that the solution $u(r,\theta)$ is only in $H^{1+\beta-\epsilon}(\Omega)$ for any $\epsilon > 0$, and hence, it is very singular for small β at the origin. This suggests that refinement is centered around the origin. The true error can be computed by

$$\text{err} := \left\|\alpha^{1/2}\nabla(u - u_{\mathcal{U}})\right\|_{0,\Omega}^2 = \int_{\partial\Omega} \alpha(\mathbf{n}\cdot\nabla u)(u - 2u_{\mathcal{U}})ds + \sum_{K\in\mathcal{T}} \left\|\alpha^{1/2}\nabla u_{\mathcal{U}}\right\|_{0,K}^2.$$

Since the true solution u is very smooth near the boundary, the integrations on the boundary can be computed very accurately. The relative error estimator will be calculated as the ratio of the estimator and $\|\alpha^{1/2}\nabla u\|_{0,\Omega}$:

$$\text{eff-index} := \frac{\eta}{\|\alpha^{1/2}\nabla(u - u_{\mathcal{U}})\|_0},$$

which is the so-called effectivity index. We will use the following stopping criteria:

$$\text{rel-err} := \left\|\alpha^{1/2}\nabla(u - u_{\mathcal{U}})\right\|_{0,\Omega} \Big/ \left\|\alpha^{1/2}\nabla u\right\|_{0,\Omega} \le \text{tol}.$$

Denote by k the number of levels of refinement and by n the number of vertices of triangulation.

Numerical experiments here will also involve the following error estimators:

(1) ZZ gradient recovery-based estimator [35]:

$$\eta_{ZZ,f,K} = \|G(\nabla u_{\mathcal{U}}) - \nabla u_{\mathcal{U}}\|_{0,K}, \tag{5.1}$$

TABLE 5.1
Comparison of estimators for relative error less than 0.5.

	k	n	Err	Rel-err	η	Eff-index
$\eta_{ZZ,g}$	57	1386	0.2697	0.4773	0.0612	0.2271
$\eta_{ZZ,f}$	21	150	0.2824	0.4998	1.7886	6.3333
η_C	24	153	0.2786	0.4932	1.8485	6.6337
η_{BV}	23	242	0.2702	0.4782	0.5976	2.2121
η_{RT}	27	221	0.2726	0.4824	0.2212	0.8118
η_{BDM}	18	196	0.2736	0.4845	0.1761	0.6434

where $G(\nabla u_\mathcal{U}) \in \mathcal{U}^d$ and its nodal value at vertex $z \in \mathcal{N}$ is defined by

$$(G(\nabla u_\mathcal{U}))_z = \frac{1}{|\omega_z|}\int_{\omega_z} \nabla u_\mathcal{U}\, dx.$$

(2) ZZ flux recovery-based error estimator [35]:

$$\eta_{ZZ,f,K} = \left\|\alpha^{-1/2}(G(-\alpha\nabla u_\mathcal{U}) + \alpha\nabla u_\mathcal{U})\right\|_{0,K}, \tag{5.2}$$

where $G(-\alpha\nabla u_\mathcal{U}) \in \mathcal{U}^d$ and its nodal value at vertex $z \in \mathcal{N}$ is defined by

$$(G(-\alpha\nabla u_\mathcal{U}))_z = \frac{1}{|\omega_z|}\int_{\omega_z} -\alpha\nabla u_\mathcal{U}\, dx.$$

(3) Carstensen flux recovery-based error estimator [9]:

$$\eta_C = \min_{\boldsymbol{\tau}\in\mathcal{U}^d}\left\|\alpha^{-1/2}(\alpha\nabla u_\mathcal{U} + \boldsymbol{\tau})\right\|_{0,\Omega}. \tag{5.3}$$

(4) Bernardi-Verfürth (BV) error estimator [6] (an improved explicit residual based estimator for diffusion problems):

$$\eta_{BV,K} := \left(\frac{1}{2}\sum_{e\in\partial K} h_e\alpha_e^{-1/2}\|[\alpha\nabla u_\mathcal{U}]_e\cdot\mathbf{n}\|_{0,e}^2 + h_K^2\alpha_K^{-1}\|f + \nabla\cdot(\alpha\nabla u_\mathcal{U})\|_{0,K}^2\right)^{1/2}, \tag{5.4}$$

where $\alpha_e = \max_{K\in\omega_e}(\alpha_K)$.

Since $f = 0$ in this example, the BV estimator may also be viewed as an edge estimator.

In the first set of numerical experiments, triangulations align with interfaces of the problem. In particular, we start with the coarsest triangulation $\mathcal{T}_0$ obtained from halving 16 congruent squares by connecting the bottom left and upper right corners. We report numerical results with the stopping criteria tol = 0.5, tol = 0.15, and tol = 0.1. For tol = 0.5, Table 5.1 shows that the $\eta_{ZZ,g}$ estimator needs about six times as many grid points as the rest estimators. Comparing Figures 5.1 and 5.2, it is clear that the $\eta_{ZZ,g}$ estimator introduces unnecessary refinements along the interfaces.

For tol = 0.15, the $\eta_{ZZ,g}$ estimator will generate too many grid points for computers to handle. Numerical results for the rest estimators are reported in Table 5.2. Figures 5.3 and 5.4 show that both the estimators $\eta_{ZZ,f}$ and η_C overrefine regions along the interfaces. This is because the tangential component of the flux is discontinuous, but the recovered flux has continuous tangential component.

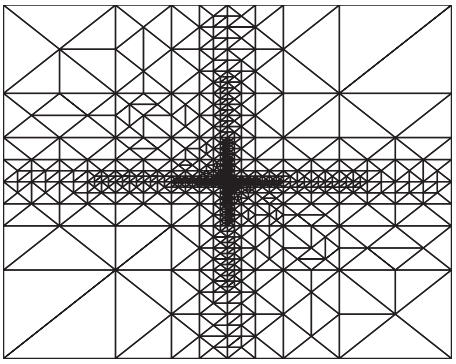

FIG. 5.1. *Mesh generated by* $\eta_{ZZ,g}$.

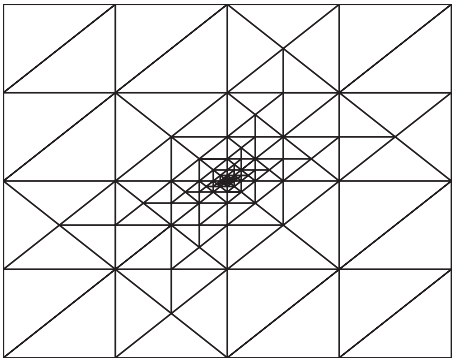

FIG. 5.2. *Mesh generated by* η_{RT}.

TABLE 5.2
Comparison of estimators for relative error less than 0.15.

	k	n	Err	Rel-err	η	Eff-index
$\eta_{ZZ,f}$	137	7593	0.0847	0.1500	0.3818	4.5055
η_C	135	6905	0.0847	0.1500	0.3629	4.2821
η_{BV}	54	1641	0.0826	0.1462	0.2628	3.1819
η_{RT}	60	1676	0.0843	0.1491	0.0959	1.1385
η_{BDM}	54	1652	0.0846	0.1498	0.0695	0.8212

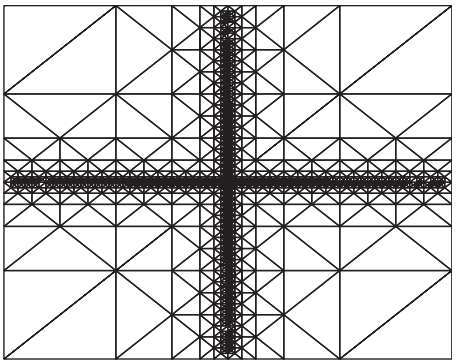

FIG. 5.3. *Mesh generated by* $\eta_{ZZ,f}$.

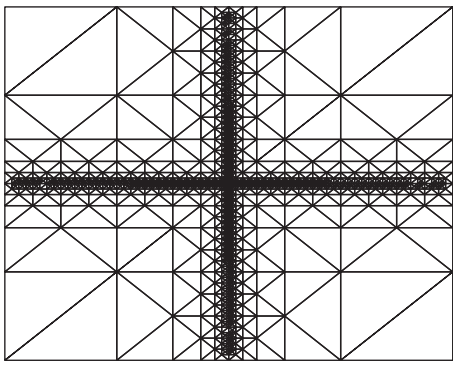

FIG. 5.4. *Mesh generated by* η_C.

TABLE 5.3
Comparison of estimators for relative error less than 0.1.

	k	n	Err	Rel-err	η	Eff-index
η_{BV}	65	3475	0.0542	0.0959	0.1827	3.3723
η_{RT}	69	3467	0.0563	0.0996	0.0671	1.1930
η_{BDM}	64	3454	0.0559	0.0989	0.0473	0.8466

For tol = 0.1, both the estimators $\eta_{ZZ,f}$ and η_C fail. Numerical results for the rest estimators are reported in Table 5.3 and Figures 5.5–5.10. Meshes generated by η_{BV}, η_{RT}, and η_{BDM} are similar. By inspecting the effectivity index, both η_{RT} and η_{BDM} are more accurate than η_{BV}, and they are possibly asymptotically exact.

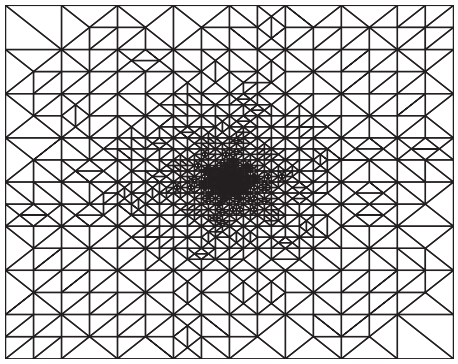

FIG. 5.5. *Mesh generated by* η_{BV}.

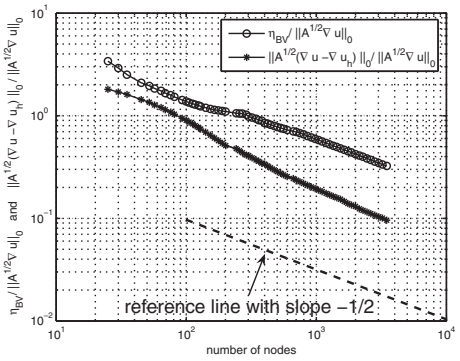

FIG. 5.6. *Error versus estimator* η_{BV}.

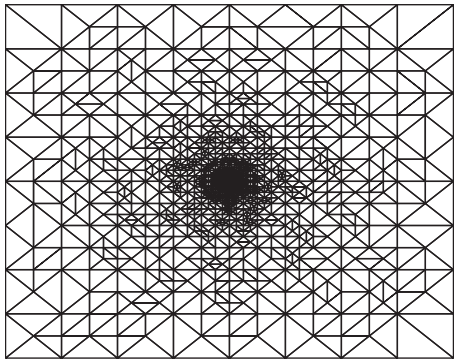

FIG. 5.7. *Mesh generated by* η_{RT}.

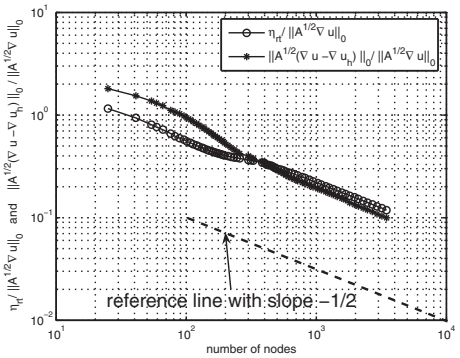

FIG. 5.8. *Error versus estimator* η_{RT}.

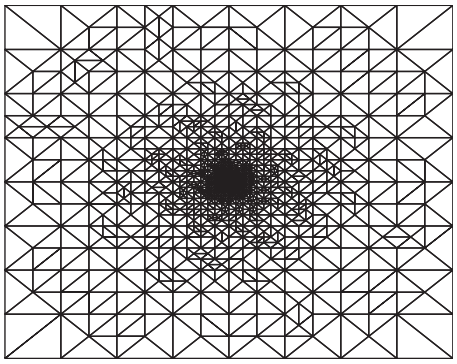

FIG. 5.9. *Mesh generated by* η_{BDM}.

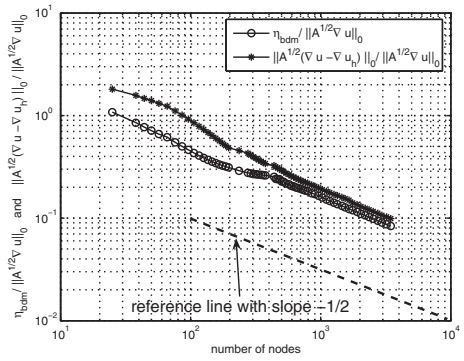

FIG. 5.10. *Error versus estimator* η_{BDM}.

The BV estimator η_{BV} is subject to the assumption that the interfaces do not cut through any element of triangulations and so is the analysis presented in section 4 for the estimators introduced in this paper. However, it is easy to see that the estimator η_{ν} defined in (3.15) is free of this assumption. In practice, it is important to consider the case that triangulations do not align with interfaces because this happens when interfaces are curves/surfaces or their locations are unknown. Hence, the purpose of

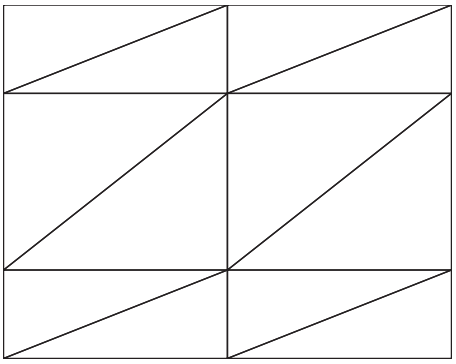

FIG. 5.11. *An initial mesh: Two horizontal lines are* $y = 0.5$ *and* $y = -0.5$.

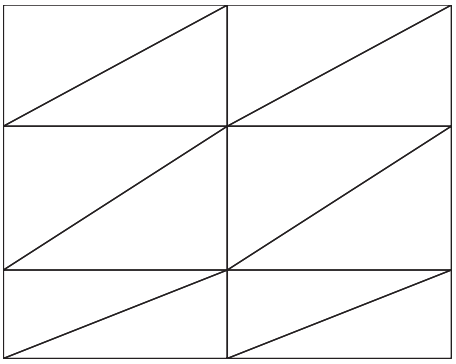

FIG. 5.12. *An initial mesh: Two horizontal lines are* $y = 0.31415926$ *and* $y = -0.5$.

TABLE 5.4
Estimator η_{RT} *with the initial mesh in Figure* 5.11.

	k	n	Err	Rel-err	η	Eff-index
η_{RT}	59	3831	0.0546	0.0966	0.0644	1.1804

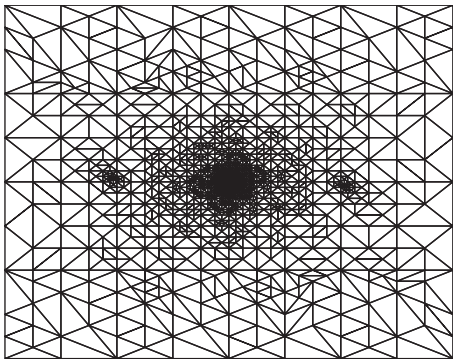

FIG. 5.13. *Mesh generated by* η_{RT} *with the initial mesh Figure* 5.11.

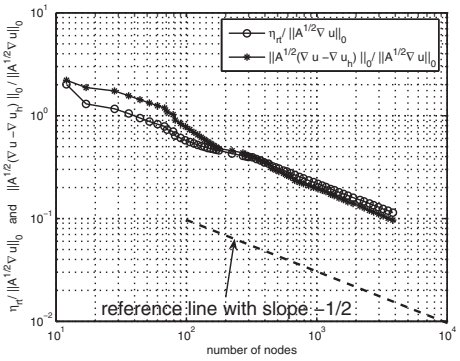

FIG. 5.14. *Error versus estimator* η_{RT} *with the initial mesh Figure* 5.11.

the second set of numerical experiments is to test our estimator η_{RT} for initial meshes not aligning with the interfaces. Consider two initial meshes depicted in Figures 5.11 and 5.12, where two horizontal lines, $y = 0.5$ and $y = -0.5$ in Figure 5.11 and $y = 0.31415926$ and $y = -0.5$ in Figure 5.12, do not coincide with the interface $y = 0$.

For the initial mesh in Figure 5.11, several steps of refinements generate a triangulation that aligns with the interface $y = 0$. Hence, numerical results depicted in Table 5.4 and Figures 5.13 and 5.14 are in a good agreement with those reported previously. For the initial mesh in Figure 5.12, we choose the horizontal line $y = 0.31415926$ so that refinements never generate a triangulation that aligns with the interface $y = 0$. Numerical results for this test are reported in Table 5.5 and Figures 5.15 and 5.16. As expected (see Figure 5.15), the mesh is refined along the interface $y = 0$ due to the nonalignment of the meshes and the interface.

TABLE 5.5
Estimator η_{RT} with the initial mesh in Figure 5.12.

	k	n	Err	Rel-err	η	Eff-index
η_{RT}	78	11249	0.0558	0.0988	0.0641	1.1480

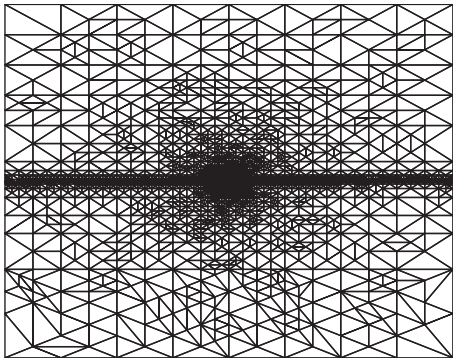

FIG. 5.15. *Mesh generated by η_{RT} with initial mesh Figure* 5.12.

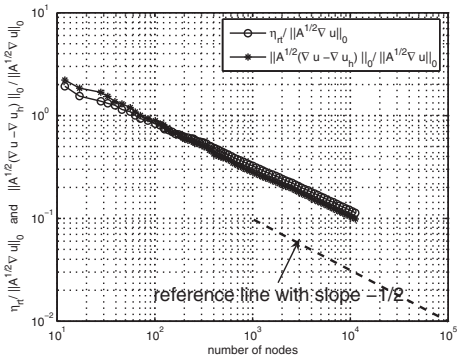

FIG. 5.16. *Error versus estimator η_{RT} with initial mesh Figure* 5.12.

6. Adaptive method. This section proposes an adaptive finite element method and analyzes its convergence. Quantities we used for marking elements for refinement are different from those in [16, 20], and hence, convergence of our adaptive algorithm is not subject to constraints on either the sufficiently small initial mesh [16] or the interior nodes in the refined elements [20].

6.1. Adaptive algorithm. Given an initial triangulation $\mathcal{T}_0$, a sequence of nested conforming triangulations $\mathcal{T}_k$ is generated through the following loop:

$$\textbf{Solve} \longrightarrow \textbf{Estimate} \longrightarrow \textbf{Mark} \longrightarrow \textbf{Refine}.$$

The **Solve** step solves (2.4) in the finite element space corresponding to the triangulation $\mathcal{T}_k$ for the discrete solution $u_k \in \mathcal{U}_g(k)$, where $\mathcal{U}_g(k)$ is the finite element space defined on $\mathcal{T}_k$ accordingly. Here and thereafter, we shall explicitly express the dependence of a quantity on k by either the subscript like u_k or the variable like $\mathcal{U}_g(k)$.

The **Estimate** step computes some *quantities*, and the **Mark** step is to mark elements, where those quantities are large, for refinement. The choice of the quantities used for marking elements is crucial for convergence analysis of the corresponding adaptive algorithms. For example, Dörfler in [16] uses only local indicators (edge-based), and convergence of his algorithm is subject to the sufficiently small initial mesh; Morin, Nochetto, and Siebert in [20] use both local indicators (residual-based) and oscillations, and hence, their algorithm is no longer subject to the initial mesh constraint but is under the interior node assumption. In this paper, we propose to use both local indicators $\eta_K(k)$ (edge- and recovery-based) and weighted element residuals $\|\alpha^{-1/2}h(k)f\|_{0,K} = \alpha_K^{-1/2}h_K(k)\|f\|_{0,K}\ \forall\ K \in \mathcal{T}_k$. Then the corresponding marking strategies are as follows:

- *Marking Strategy E*: Giving a parameter $0 < \theta_E < 1$, construct a *minimal* subset $\hat{\mathcal{T}}_k$ of $\mathcal{T}_k$ such that

$$\sum_{K\in\hat{\mathcal{T}}_k} \eta_K^2(k) \geq \theta_E^2 \eta^2(k), \tag{6.1}$$

  where $\eta^2(k) = \sum_{K\in\mathcal{T}_k} \eta_K^2(k)$;
- *Marking Strategy R*: Giving a parameter $0 < \theta_0 < 1$ and the subset $\hat{\mathcal{T}}_k \subset \mathcal{T}_k$ produced by the Marking Strategy E, enlarge $\hat{\mathcal{T}}_k$ to a *minimal* set (denoted again by $\hat{\mathcal{T}}_k$) such that

$$\sum_{K\in\hat{\mathcal{T}}_k} \left\|\alpha^{-1/2}h(k)f\right\|_{0,K}^2 \geq \theta_0^2 \left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2. \tag{6.2}$$

Finally, the **Refine** step is to refine the elements in $\hat{\mathcal{T}}_k$ obtained in Marking Strategy R to generate a new triangulation $\mathcal{T}_{k+1}$ such that $\mathcal{U}_0(k) \subset \mathcal{U}_0(k+1)$ and that each of its edges/faces contains a node of the finer mesh $\mathcal{T}_{k+1}$ in their interior. Note that some elements in $\mathcal{T}_k \setminus \hat{\mathcal{T}}_k$ adjacent to elements in $\hat{\mathcal{T}}_k$ are also refined to avoid hanging nodes. Note also that new interior nodes in the elements in $\hat{\mathcal{T}}_k$ are not required.

In summary, the adaptive finite element algorithm may be defined as follows.

ADAPTIVE ALGORITHM. For a given initial mesh $\mathcal{T}_0$, choose parameters $\theta_E,\ \theta_0 \in (0,\ 1)$. For $k = 0,\ 1,\ 2, \ldots$, perform

(1) $u_k = \mathbf{Solve}\,(\mathcal{T}_k,\ f,\ g,\ \alpha(x))$.
(2) $\{\eta_K(k),\ \|\alpha^{-1/2}h(k)f\|_{0,K}\}_{K\in\mathcal{T}_k} = \mathbf{Estimate}\,(\mathcal{T}_k,\ u_k,\ f,\ g,\ \alpha(x))$.
(3) $\hat{\mathcal{T}}_k = \mathbf{Mark}\,(\theta_E,\ \theta_0;\ \mathcal{T}_k,\ \{\eta_K(k),\ \|\alpha^{-1/2}h(k)f\|_K\}_{K\in\mathcal{T}_k})$.
(4) $\mathcal{T}_{k+1} = \mathbf{Refine}\,(\mathcal{T}_k,\ \hat{\mathcal{T}}_k)$.

6.2. Convergence analysis. The analysis presented here is similar to that of [20]. To establish the convergence of the adaptive method, we start with the following assumptions on a posteriori error estimators.

Assumption R. Assume that there exists a positive constant C_r such that

$$|||u - u_k|||_\Omega^2 \leq C_r \left(\eta^2(k) + \left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2\right) \tag{6.3}$$

for $k = 1,\ 2,\ \ldots$.

Assumption E. Assume that there exists a positive constant C_l such that

$$C_l\theta_E^2\eta^2(k) \leq |||u_{k+1} - u_k|||_\Omega^2 + \left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2 \tag{6.4}$$

for $k = 1,\ 2,\ \ldots$.

The Assumption R is similar to but weaker than the global reliability bound. This bound is established for the estimators introduced in this paper in section 4.2 and for the edge estimator defined in (4.16) (see the proof of Theorem 5.3 in [25]) with C_r independent of the size of jumps. It also holds for the ZZ estimator, but the constant C_r depends on the size of jumps (see, e.g., [28]). The Assumption E will be verified in the next section for various estimators.

LEMMA 6.1. *Under the Assumptions R and E, we have*

$$|||u_{k+1} - u_k|||_\Omega^2 \geq \delta_1|||u - u_k|||_\Omega^2 - (1+\delta_0)\left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2, \tag{6.5}$$

where $\delta_0 = C_l\theta_E^2$ *and* $\delta_1 = \delta_0/C_r \in (0,\ 1)$.

Proof. Equation (6.5) is a direct consequence of the Assumptions R and E. ∎

Next, we show that the weighted element residual $\|\alpha^{-1/2}h(k)f\|_{0,\Omega}$ as a function of the refinement level k is monotonically decreasing.

LEMMA 6.2. *Let $0 < \gamma_0 < 1$ be the reduction factor of element size associated with one refinement step. Let $\hat{\mathcal{T}}_k$ be a subset of $\mathcal{T}_k$ satisfying Marking Strategy R. If $\mathcal{T}_{k+1}$ is generated by the* Refine step *from $\mathcal{T}_k$, then the following element residual reduction occurs:*

$$\left\|\alpha^{-1/2}h(k+1)f\right\|_{0,\Omega} \leq \zeta \left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}, \tag{6.6}$$

with $\zeta = \sqrt{1-(1-\gamma_0^2)\,\theta_0^2}$.

Proof. Denote the collection of elements in $\mathcal{T}_{k+1}$ contained in elements of $\hat{\mathcal{T}}_k$ by

$$\tilde{\mathcal{T}}_{k+1} = \left\{K \in \mathcal{T}_{k+1} \,|\, K \subset \hat{K} \in \hat{\mathcal{T}}_k\right\}.$$

By the assumption of the lemma, we have

$$h_K(k+1) \leq \gamma_0\, h_{\hat{K}}(k) \quad \forall\ K \in \tilde{\mathcal{T}}_{k+1},$$

which, combining with $\alpha_K = \alpha_{\hat{K}}$ for $K \in \tilde{\mathcal{T}}_{k+1}$, implies

$$\begin{aligned}\sum_{K\in\tilde{\mathcal{T}}_{k+1}} \alpha_K^{-1}h_K^2(k+1)\|f\|_{0,K}^2 &\leq \gamma_0^2 \sum_{\hat{K}\in\hat{\mathcal{T}}_k} \alpha_{\hat{K}}^{-1}h_{\hat{K}}^2(k)\|f\|_{0,\hat{K}}^2 \\ &= \gamma_0^2\left(\left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2 - \sum_{K\in\mathcal{T}_k\setminus\hat{\mathcal{T}}_k} \alpha_K^{-1}h_K^2(k)\|f\|_{0,K}^2\right). \end{aligned} \tag{6.7}$$

Note that

$$\mathcal{T}_{k+1}\setminus\tilde{\mathcal{T}}_{k+1} = \left\{K \in \mathcal{T}_{k+1} \,|\, K \in \mathcal{T}_k\setminus\hat{\mathcal{T}}_k \text{ or } K \subset \hat{K} \in \mathcal{T}_k\setminus\hat{\mathcal{T}}_k\right\}$$

and that for any $K \in \mathcal{T}_{k+1}\setminus\tilde{\mathcal{T}}_{k+1}$,

$$\alpha_K^{-1/2}h_K(k+1) = \alpha_K^{-1/2}h_K(k) \text{ if } K \in \mathcal{T}_k$$

or

$$\alpha_K^{-1/2}h_K(k+1) \leq \alpha_{\hat{K}}^{-1/2}h_{\hat{K}}(k) \text{ if } K \subset \hat{K} \in \mathcal{T}_k\setminus\hat{\mathcal{T}}_k.$$

Hence,

$$\sum_{K\in\mathcal{T}_{k+1}\setminus\tilde{\mathcal{T}}_{k+1}} \alpha_K^{-1}h_K^2(k+1)\|f\|_{0,K}^2 \leq \sum_{K\in\mathcal{T}_k\setminus\hat{\mathcal{T}}_k} \alpha_K^{-1}h_K^2(k)\|f\|_{0,K}^2,$$

which, together with (6.7), yields

$$\begin{aligned}\left\|\alpha^{-1/2}h(k+1)f\right\|_{0,\Omega}^2 &= \sum_{K\in\tilde{\mathcal{T}}_{k+1}} \alpha_K^{-1}h_K^2(k+1)\|f\|_{0,K}^2 + \sum_{K\in\mathcal{T}_{k+1}\setminus\tilde{\mathcal{T}}_{k+1}} \alpha_K^{-1}h_K^2(k+1)\|f\|_{0,K}^2 \\ &\leq \gamma_0^2\left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2 + (1-\gamma_0^2)\sum_{K\in\mathcal{T}_k\setminus\hat{\mathcal{T}}_k} \alpha_K^{-1}h_K^2(k)\|f\|_{0,K}^2.\end{aligned}$$

Combining with the following consequence of (6.2),

$$\sum_{K\in\mathcal{T}_k\setminus\hat{\mathcal{T}}_k}\alpha_K^{-1}h_K^2(k)\|f\|_{0,K}^2\le(1-\theta_0^2)\left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2$$

gives the validity of (6.6). This completes the proof of the lemma. □

Now, we are ready to establish the error reduction property of the adaptive finite element method. For convenience, we introduce some matrix notations. For a vector $\mathbf{y}_k=(y_1,\,y_2)^t$ and a matrix $B=(b_{ij})_{2\times2}$, denote by

$$\|\mathbf{y}_k\|_{l_2}=\sqrt{y_1^2+y_2^2}\quad\text{and}\quad\rho(B)=\max|\lambda(B)|$$

the respective l_2 norm and spectral radius, where $\lambda(B)$ denotes the eigenvalue of B.

THEOREM 6.1. *Let $\delta=\max\{\sqrt{1-\delta_1},\,\zeta\}$ with δ_1 and ζ given in the respective Lemmas* 6.1 *and* 6.2. *Under the Assumption R and Assumption E, the sequence $\{u_k\}$ generated by the adaptive finite element method satisfies the following error reduction property:*

$$|||u-u_k|||_\Omega\le C_0\delta^k, \tag{6.8}$$

with $C_0=\sqrt{|||u-u_0|||_\Omega^2+\|\alpha^{-1/2}h(0)f\|_{0,\Omega}^2}$.

Proof. Since $\mathcal{U}_0(k)\subset\mathcal{U}_0(k+1)$, it follows from the orthogonal property of the finite element approximations and Lemma 6.1 that

$$\begin{aligned}|||u-u_k|||_\Omega^2&=|||u-u_{k+1}|||_\Omega^2+|||u_{k+1}-u_k|||_\Omega^2\\&\ge|||u-u_{k+1}|||_\Omega^2+\delta_1|||u-u_k|||_\Omega^2-(1+\delta_0)\left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2,\end{aligned}$$

which implies

$$|||u-u_{k+1}|||_\Omega^2\le(1-\delta_1)|||u-u_k|||_\Omega^2+(1+\delta_0)\left\|\alpha^{-1/2}h(k)f\right\|_{0,\Omega}^2. \tag{6.9}$$

Let $\mathbf{y}_k=(|||u-u_k|||_\Omega,\,\|\alpha^{-1/2}h(k)f\|_{0,\Omega})^t$ and

$$B=\begin{pmatrix}1-\delta_1 & 1+\delta_0\\ 0 & \zeta\end{pmatrix},$$

then (6.9) and Lemma 6.2 lead to

$$\mathbf{y}_k\le B\,\mathbf{y}_{k-1}=B^k\,\mathbf{y}_0.$$

Hence,

$$|||u-u_k|||_\Omega\le\|\mathbf{y}_k\|_{l_2}\le\rho\left(B^k\right)\,\|\mathbf{y}_0\|_{l_2}.$$

Now, (6.8) is an immediate consequence of the facts that

$$\rho(B^k)=\max\left\{(1-\delta_1)^{k/2},\,\zeta^k\right\}=\delta^k\quad\text{and}\quad C_0=\|\mathbf{y}_0\|_{l_2}.$$

This completes the proof of the theorem. □

6.3. Assumption E. In this section, we verify the Assumption E for several estimators. For simplicity, we analyze only problem (2.1) with pure homogeneous Dirichlet boundary conditions, i.e., $g = 0$ and $\Gamma_D = \partial\Omega$ in (2.2). The extension to mixed boundary conditions for the estimators analyzed in section 4 is straightforward.

Let $\eta_e(k)$ be the edge estimator defined in (4.16); the following lemma establishes a local upper bound. The proof here is similar to those in [16, 20].

LEMMA 6.3. *For every $e \in \mathcal{E}_k$ containing a vertex of $\mathcal{T}_{k+1}$ as its interior point, there exists a positive constant C such that*

$$\eta_e^2(k) \le C \left(|\!|\!| u_{k+1} - u_k |\!|\!|_{\omega_e}^2 + \left\| \alpha^{-1/2} h(k) f \right\|_{0,\omega_e}^2 \right). \tag{6.10}$$

Proof. For any $v \in \mathcal{U}_0(k+1)$, the orthogonality property of the finite element approximation u_{k+1} implies

$$a(u_{k+1} - u_k,\, v) = a(u_{k+1} - u,\, v) + a(u - u_k,\, v) = a(u - u_k,\, v),$$

which, together with integration by parts, (2.1), the fact that $\nabla \cdot (\alpha(x)\nabla u_k)|_K = 0 \;\forall\; K \in \mathcal{T}_k$, and the continuity of the normal component of the exact flux, gives

$$a(u_{k+1} - u_k,\, v) = a(u - u_k,\, v) = \sum_{K \in \mathcal{T}_k} \int_K f v \, dx + \sum_{e \in \mathcal{E}_k} \int_e J_e(\alpha(x)\nabla u_k)\, v \, ds. \tag{6.11}$$

Since $\alpha(x)$ is a piecewise constant, then $J_e(\alpha(x)\nabla u_k)$ on each $e \in \mathcal{E}_k$ is a constant denoted by j_e. Let $\psi_e \in \mathcal{U}_0(k+1)$ be the nodal basis function associated with one of interior nodes on e, then it is easy to see that

$$\begin{aligned} &\operatorname{supp}(\psi_e) \subset \omega_e, \quad \psi_e|_{e'} = 0 \quad \forall\; e \ne e' \in \mathcal{E}_k, \quad \|\psi_e\|_{0,\omega_e} \le C h_e(k), \\ &\|\nabla \psi_e\|_{0,\omega_e} \le C, \quad \text{and} \quad \textstyle\int_e \psi_e \, ds \ge C\, h_e(k), \end{aligned}$$

where $\omega_e = K_e^+ \cup K_e^-$ and C is a positive constant independent of $h_e(k)$. Choosing $v = j_e \psi_e$ in (6.11), we then have

$$\begin{aligned} C\, h_e(k)\, |j_e|^2 &\le \int_e j_e^2 \psi_e \, ds = a(u_{k+1} - u_k,\, j_e\psi_e) - \int_{\omega_e} f j_e \psi_e \, dx \\ &\le C \left(\alpha_{K_e^+} + \alpha_{K_e^-} \right)^{1/2} |j_e| \left(|\!|\!| u_{k+1} - u_k |\!|\!|_{\omega_e} + \left\| \alpha^{-1/2} h(k) f \right\|_{0,\omega_e} \right). \end{aligned}$$

Hence,

$$\eta_e(k) = \frac{\sqrt{2}|j_e|\, h_e(k)}{\left(\alpha_{K_e^+} + \alpha_{K_e^-} \right)^{-1/2}} \le C \left(|\!|\!| u_{k+1} - u_k |\!|\!|_{\omega_e} + \left\| \alpha^{-1/2} h(k) f \right\|_{\omega_e} \right),$$

which leads to (6.10) and hence, the proof of the lemma. ∎

LEMMA 6.4. *For all $K \in \hat{\mathcal{T}}_k$, there exists a positive constant C such that*

$$\hat{\eta}_{RT_0,K}^2(k) \le C \left(|\!|\!| u_{k+1} - u_k |\!|\!|_{\omega_K}^2 + \left\| \alpha^{-1/2} h(k) f \right\|_{0,\omega_K}^2 \right). \tag{6.12}$$

Proof. Equation (6.12) is a direct consequence of Lemma 6.3 and the respective (4.21). ∎

With the local discrete bounds in Lemmas 6.3 and 6.4, it is then straightforward to show the validity of the Assumption E.

LEMMA 6.5. *The Assumption E is valid for the estimators* $\eta_e(k)$ *and* $\hat{\eta}_{RT_0}(k)$.

Proof. Since ω_K contains at most $d+2$ elements, it then follows from (6.1) and (6.12) that

$$\theta_E^2 \hat{\eta}_{RT_0}^2(k) \leq \sum_{K \in \hat{\mathcal{T}}_k} \hat{\eta}_{RT_0,K}^2(k) \leq C \sum_{K \in \hat{\mathcal{T}}_k} \left(\|\!| u_{k+1} - u_k |\!\|_{\omega_K}^2 + \left\| \alpha^{-1/2} h(k) f \right\|_{0,\omega_K}^2 \right)$$

$$\leq (d+2) C \left(\|\!| u_{k+1} - u_k |\!\|_{\Omega}^2 + \left\| \alpha^{-1/2} h(k) f \right\|_{0,\Omega}^2 \right),$$

which proves the validity of the Assumption E with $C_l = \left(C\,(d+2)\right)^{-1}$ for the estimators $\hat{\eta}_{RT_0}(k)$. It may be proved in the same fashion for the estimator $\eta_e(k)$. ∎

REFERENCES

[1] M. AINSWORTH AND A. W. CRAIG, *A posteriori error estimation in finite element method*, Numer. Math., 36 (1991), pp. 429–463.

[2] M. AINSWORTH AND J. T. ODEN, *A Posteriori Error Estimation in Finite Element Analysis*, Pure Appl. Math., Wiley-Interscience, John Wiley & Sons, New York, 2000.

[3] R. BANK AND J. XU, *Asymptotically exact a posteriori error estimators*, *Part* I: *Grids with supperconvergence*, SIAM J. Numer. Anal., 41 (2003), pp. 2294–2312; *Part* II: *General unstructured grids*, SIAM J. Numer. Anal., 41 (2003), pp. 2313–2332.

[4] R. BANK, J. XU, AND B. ZHENG, *Superconvergent derivative recovery for Lagrange triangular elements of degree p on unstructured grids*, SIAM J. Numer. Anal., 45 (2007), pp. 2032–2046.

[5] I. BABUSKA AND T. STROUBOULIS, *The Finite Element Method and Its Reliability*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 2001.

[6] C. BERNARDI AND R. VERFÜRTH, *Adaptive finite element methods for elliptic equations with non-smooth coefficients*, Numer. Math., 85 (2000), pp. 579–608.

[7] S. C. BRENNER AND C. CARSTENSEN, *Finite element methods*, in Encyclopedia of Computational Mechanics, John Wiley and Sons, Vol. 1: Fundamentals, E. Stein, R. de Borst, T. Hughes, eds., 2004, pp. 73–118.

[8] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[9] C. CARSTENSEN, *All first-order averaging technique for a posteriori finite element error control on unstructure grids are efficient and reliable*, Math. Comp., 73 (2003), pp. 1153–1165.

[10] C. CARSTENSEN AND S. BARTELS, *Each averaging technique yields reliable a posteriori error control in FEM on unstructure grids. Part* I: *Low order conforming, nonconforming, and mixed FEM*, Math. Comp., 71 (2002), pp. 945–969.

[11] C. CARSTENSEN, S. BARTELS, AND R. KLOSE, *An experimental survey of a posteriori Courant finite element error control for the Poisson equation*, Adv. Comput. Math., 15 (2001), pp. 79–106.

[12] C. CARSTENSEN AND R. VERFÜRTH, *Edge residuals dominate a posteriori error estimates for low order finite element methods*, SIAM J. Numer. Anal., 36 (1999), pp. 1571–1587.

[13] J. M. CASCON, C. KREUZER, R. H. NOCHETTO, AND K. G. SIEBERT, *Quasi-optimal Convergence Rate for an Adaptive Finite Element Method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.

[14] Z. CHEN AND S. DAI, *On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients*, SIAM J. Sci. Comput., 24 (2002), pp. 443–462.

[15] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[16] W. DÖRFLER, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.

[17] W. HOFFMANN, A. H. SCHATZ, L. B. WAHLBIN, AND G. WITTUM, *Asymptotically exact a posteriori estimators for the pointwise gradient error on each element in irregular meshes,*

Part I: *A smooth problem and globally quasi-uniform meshes*, Math. Comp., 70 (2001), pp. 897–909.

[18] R. B. Kellogg, *On the Poisson equation with intersecting interfaces*, Appl. Anal., 4 (1975), pp. 101–129.

[19] R. Luce and B. I. Wohlmuth, *A local a posteriori error estimator based on equilibrated fluxes*, SIAM J. Numer. Anal., 42 (2004), pp. 1394–1414.

[20] P. Morin, R. H. Nochetto, and K. G. Siebert, *Convergence of adaptive finite element methods*, SIAM Rev., 44 (2002), pp. 631–658.

[21] A. Naga and Z. Zhang, *The polynomial-preserving recovery for higher order finite element methods in* $2D$ *and* $3D$, Discrete Contin. Dyn. Syst. Ser. B, 5-3 (2005), pp. 769–798.

[22] R. H. Nochetto, *Adaptive finite element methods for elliptic PDE*, Lecture notes, Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA, 2006.

[23] J. S. Ovall, *Two Dangers to Avoid When Using Gradient Recovery Methods for Finite Element Error Estimation and Adaptivity*, Technical report 6, Max-Planck-Institute fur Mathematick in den Naturwissenschaften, Bonn, Germany, 2006.

[24] J. S. Ovall, *Fixing a "Bug" in Recovery-type A Posteriori Error Estimators*, Technical report 25, Max-Planck-Institute fur Mathematick in den Naturwissenschaften, Bonn, Germany, 2006.

[25] M. Petzoldt, *A posteriori error estimators for elliptic equations with discontinuous coefficients*, Adv. Comput. Math., 16 (2002), pp. 47–75.

[26] P. A. Raviart and I. M. Thomas, *A Mixed Finite Element Method for Second Order Elliptic Problems*, Lecture Notes in Math. 606, Springer-Verlag, Berlin and New York (1977), pp. 292–315.

[27] A. H. Schatz and L. B.Wahlbin, *Asymptotically exact a posteriori estimators for the pointwise gradient error on each element in irregularmeshes, Part* II: *The piecewise linear case*, Math. Comp., 73 (2004), pp. 517–523.

[28] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Stuttgart, Germany, 1996.

[29] N. Yan and A. Zhou, *Gradient recovery type a posteriori error estimates for finite element approximations on irregular meshes*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 4289–4299.

[30] D. Yu, *Asymptotically exact a posteriori error estimators for elements of bi-even degree*, Chinese J. Numer. Math. Appl., 13, (1991), pp. 64–78.

[31] D. Yu, *Asymptotically exact a posteriori error estimators for elements of bi-odd degree*, Chinese J. Numer. Math. Appl., 13 (1991), pp. 82–90.

[32] Z. Zhang, *A posteriori error estimates on irregular grids based on gradient recovery*, Adv. Comput. Math., 15 (2001), pp. 363–374.

[33] Z. Zhang, *Recovery techniques in finite element methods*, in Adaptive Computations: Theory and Algorithms, Mathematics Monogr. Ser. 6, T. Tang and J. Xu, eds., Science Publisher, New York, 2007, pp. 333–412.

[34] Z. Zhang and J. Z. Zhu, *Analysis of the superconvergent patch recovery technique and a posteriori error estimator in the finite element method. Part* 1, Comput. Methods Appl. Mech. Engrg., 123 (1995), pp. 173–187; *Part* 2, Comput. Methods Appl. Mech. Engrg., 163 (1998), pp. 159–170.

[35] O. C. Zienkiewicz and J. Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis*, Internat. J. Numer. Methods Engrg., 24 (1987), pp. 337–357.

[36] O. C. Zienkiewicz and J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates*, Internat. J. Numer. Methods Engrg., 33 (1992), *Part* 1: *The recovery technique*, pp. 1331–1364; *Part* 2: *Error estimates and adaptivity*, pp. 1365–1382.