A SIMPLE VIRTUAL ELEMENT-BASED FLUX
RECOVERY ON QUADTREE

Shuhao Cao*

Department of Mathematics and Statistics, Washington University in St. Louis
St. Louis, MO 63105, USA

Abstract. In this paper, we introduce a simple local flux recovery for $\mathcal{Q}_k$ finite element of a scalar coefficient diffusion equation on quadtree meshes, with no restriction on the irregularities of hanging nodes. The construction requires no specific ad hoc tweaking for hanging nodes on l-irregular ($l \geq 2$) meshes thanks to the adoption of virtual element families. The rectangular elements with hanging nodes are treated as polygons as in the flux recovery context. An efficient *a posteriori* error estimator is then constructed based on the recovered flux, and its reliability is proved under common assumptions, both of which are further verified in numerics.

1. **Introduction.** In this paper, we consider the following diffusion equation on $\Omega \subset \mathbb{R}^2$,

$$\begin{cases} -\nabla \cdot (\alpha \nabla u) = f, & \text{in } \Omega, \\ \quad\quad\quad u = 0, & \text{on } \partial\Omega. \end{cases} \tag{1}$$

To approximate (1), taking advantage of the adaptive mesh refinement (AMR) to save valuable computational resources, the adaptive finite element method on quadtree mesh is among the most popular ones in the engineering and scientific computing community [20]. Compared with simplicial meshes, quadtree meshes provide preferable performance in the aspects of the accuracy and robustness. There are lots of mature software packages (e.g., [1, 2]) on quadtree meshes. To guide the AMR, one possible way is through the *a posteriori* error estimation to construct computable quantities to indicate the location that the mesh needs to be refined/coarsened, thus to balance the spacial distribution of the error which improves the accuracy per computing power. Residual-based and recovery-based error estimators are among the most popular ones used. In terms of accuracy, the recovery-based error estimator shows more appealing attributes [28, 3].

More recently, newer developments on flux recovery have been studied by many researchers on constructing a post-processed flux in a structure-preserving approximation space. Using (1) as an example, given that the data $f \in L^2(\Omega)$, the flux $-\alpha\nabla u$ is in $\boldsymbol{H}(\text{div}) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \nabla \cdot \boldsymbol{v} \in L^2(\Omega)\}$, which has less continuity constraint than the ones in [28, 3] which are vertex-patch based with the recovered flux being $H^1(\Omega)$-conforming. The $\boldsymbol{H}(\text{div})$-flux recovery shows more robustness than vertex-patch based ones (e.g., [11, 10]).

However, these $\boldsymbol{H}(\mathrm{div})$-flux recovery techniques work mainly on conforming meshes. For nonconforming discretizations on nonmatching grids, some simple treatment of hanging nodes exists by recovering the flux on a conforming mother mesh [22]. To our best knowledge, there is no literature about the local $\boldsymbol{H}(\mathrm{div})$-flux recovery on a multilevel irregular quadtree meshes. One major difficulty is that it is impossible to recover a robust computable polynomial flux to satisfy the $\boldsymbol{H}(\mathrm{div})$-continuity constraint, that is, the flux is continuous in the normal direction on edges with hanging nodes.

More recently, a new class of methods called the virtual element methods (VEM) were introduced in [4, 8], which can be viewed as a polytopal generalization of the tensorial/simplicial finite element. Since then, lots of applications of VEM have been studied by many researchers. A usual VEM workflow splits the consistency (approximation) and the stability of the method as well as the finite dimensional approximation space into two parts. It allows flexible constructions of spaces to preserve the structure of the continuous problems such as higher order continuities, exact divergence-free spaces, and many others. The VEM functions are represented by merely the degrees of freedom (DoF) functionals, not the pointwise values. In computation, if an optimal order discontinuous approximation can be computed elementwisely, then adding an appropriate parameter-free stabilization suffices to guarantee the convergence under common assumptions on the geometry of the mesh.

The adoption of the polytopal element brings many distinctive advantages, for example, treating rectangular element with hanging nodes as polygons allows a simple construction of $\boldsymbol{H}(\mathrm{div})$-conforming finite dimensional approximation space on meshes with multilevel irregularities. We shall follow this approach to perform the flux recovery for a conforming $\mathcal{Q}_k$ discretization of problem (1). Recently, arbitrary level of irregular quadtree meshes have been studied in [21, 26, 15]. Analyses of the residual-based error estimator on 1-irregular (balanced) quadtree mesh can be found, e.g., in [14]. In the virtual element context, Zienkiewicz-Zhu (ZZ)-type recovery techniques are studied for linear elasticity in [18], and for diffusion problems in [24]. In [18, 24], the recovered flux is in $\boldsymbol{H}^1$ and associated with nodal DoFs, thus cannot yield a robust estimate when the diffusion coefficient has a sharp contrast [11, 10]. The first equilibrated flux recovery in $\boldsymbol{H}(\mathrm{div})$ for virtual element methods is studied in [19]. While [19] recovers a flux by solving a mixed problem globally, we opt for a cheap and simple weighted averaging locally.

The major ingredient in our study is an $\boldsymbol{H}(\mathrm{div})$-conforming virtual element space modified from the ones used in [8, 5] (Section 2.2). Afterwards, an $\boldsymbol{H}(\mathrm{div})$-conforming flux is recovered by a robust weighted averaging of the numerical flux, in which some unique properties of the tensor-product type element $\mathcal{Q}_k$ are exploited (Section 3). The *a posteriori* error estimator is constructed based on the projected flux elementwisely. The efficiency of the local error indicator is then proved by bounding it above by the residual-based error indicator (Section 4.1). The reliability of the recovery-based error estimator is then shown under certain assumptions (Section 4.2). These estimates are verified numerically by some common AMR benchmark problems implemented in a publicly available finite element software library iFEM [16] (Section 5).

2. **Preliminaries.**

2.1. **Discretization and notations.** If Ω is not a rectangle, u is extended by 0 to an $\widetilde{\Omega}$ that is rectangular, therefore without loss of generality, we assume Ω is

partitioned into a shape-regular $\mathcal{T} = \{K\}$ with rectangular elements, and $\alpha := \alpha_K$ is assumed to be a piecewise, positive constant with respect to $\mathcal{T}$. The weak form of problem (1) is then discretized in a tensor-product finite element space as follows,

$$(\alpha \nabla u_{\mathcal{T}}, \nabla v_{\mathcal{T}}) = (f, v_{\mathcal{T}}), \quad \forall v_{\mathcal{T}} \in \mathcal{Q}_k(\mathcal{T}) \cap H_0^1(\Omega), \tag{2}$$

in which the standard notation is opted. $(\cdot, \cdot)_D$ denotes the inner product on $L^2(D)$, and $\|\cdot\|_D := \sqrt{(\cdot, \cdot)_D}$, with the subscript omitted when $D = \Omega$. The discretization space is

$$\mathcal{Q}_k(\mathcal{T}) := \{v \in H^1(\Omega) : v|_K \in \mathbb{Q}_k(K), \ \forall K \in \mathcal{T}\}.$$

and on $K = [a, b] \times [c, d]$

$$\mathbb{Q}_k(K) := \mathbb{P}_{k,k}(K) = \{p(x)q(y), \ p \in \mathbb{P}_k([a, b]), q \in \mathbb{P}_k([c, d])\},$$

where $\mathbb{P}_k(D)$ stands for the degree no more than k polynomial defined on D. Henceforth, we shall simply denote $\mathcal{Q}_k(\mathcal{T}) =: \mathcal{Q}_k$ when no ambiguity arises.

On K, the sets of 4 vertices, as well as 4 edges of the same generation with K, are denoted by $\mathcal{N}_K$ and $\mathcal{E}_K$, respectively. The sets of nodes and edges in $\mathcal{T}$ are denoted by $\mathcal{N} := \bigcup_{K \in \mathcal{T}} \mathcal{N}_K$ and $\mathcal{E} := \bigcup_{K \in \mathcal{T}} \mathcal{E}_K$. A node $\boldsymbol{z} \in \mathcal{N}$ is called a hanging node if it is on ∂K but is not counted as a vertex of $K \in \mathcal{T}$, and we denote the set of hanging nodes as $\mathcal{N}_H$

$$\mathcal{N}_H := \{\boldsymbol{z} \in \mathcal{N} : \exists K \in \mathcal{T}, \boldsymbol{z} \in \partial K \backslash \mathcal{N}_K\} \tag{3}$$

Otherwise the node $\boldsymbol{z} \in \mathcal{N}$ is a regular node. If an edge $e \in \mathcal{E}$ contains at most l hanging nodes, the partition $\mathcal{T}$, as well as the element these hanging nodes lie on, is called l-irregular.

For each edge $e \in \mathcal{E}$, a unit normal vector $\boldsymbol{n}_e$ is fixed by specifying its direction pointing rightward for vertical edges, and upward for horizontal edges. If an exterior normal of an element on this edge shares the same orientation with $\boldsymbol{n}_e$, then this element is denoted by K_-, otherwise it is denoted by K_+, i.e., $\boldsymbol{n}_e$ is pointing from K_- to K_+. The intersection of the closures of K_+, K_- is always an edge $e \in \mathcal{E}$. However, we note that by the definition in (3) it is possible that $e \in \mathcal{E}_{K_+}$ but not in $\mathcal{E}_{K_-}$ or vice versa, if there exists a hanging node on e (see e.g., Figure 1). For any function or distribution v well-defined on the two elements, define $[\![v]\!]_e = v^- - v^+$ on an edge $e \not\subset \partial \Omega$, in which v^- and v^+ are defined in the limiting sense $v^\pm = \lim_{\epsilon \to 0^\pm} v(\boldsymbol{x} + \epsilon \boldsymbol{n}_e)$ for $\boldsymbol{x} \in e$. If e is a boundary edge, the function v is extended by zero outside the domain to compute $[\![v]\!]_e$. Furthermore, the following notation denotes a weighted average of v on edge e for a weight $\gamma \in [0, 1]$,

$$\{v\}_e^\gamma := \gamma v^- + (1 - \gamma) v^+.$$

2.2. **Virtual element spaces.** In this subsection, the quadtree mesh $\mathcal{T}$ of interest is embedded into a polygonal mesh $\mathcal{T} \hookrightarrow \mathcal{T}_{\text{poly}} = \{K_{\text{poly}}\}$. On any given quadrilateral element K, for example we consider a $v_{\mathcal{T}} \in \mathbb{Q}_1(K)$, it has 4 degrees of freedom associated with 4 nodes $\{z\}$. Its numerical flux $-\alpha \nabla v_{\mathcal{T}} \cdot \boldsymbol{n}$ is well-defined on the 4 edges $\{e\}$ locally on K, such that on each edge it is a polynomial defined on the whole edge, regardless of the number of hanging nodes on that edge. Using Figure 1 as an example, on the upper right element K, $\nabla v_{\mathcal{T}}|_K \cdot \boldsymbol{n}|_{\overrightarrow{z_2 z_6}} \in \mathbb{P}_1(\overrightarrow{z_2 z_6})$ is a linear function in y-variable.

For the embedded element $K_{\text{poly}} \in \mathcal{T}_{\text{poly}}$, which geometrically coincides with K, it includes all the hanging nodes, while the set of edges are formed accordingly as the edges of the cyclic graph of the vertices. We shall denote the set of all edges on $\mathcal{T}_{\text{poly}}$ as $\mathcal{E}_{\text{poly}}$. Using Figure 1 as example, it is possible to define a flux on K
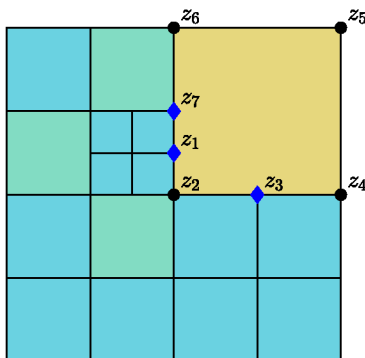
FIGURE 1. For the upper right element $K \in \mathcal{T}$, $\mathcal{N}_K = \{z_2, z_4, z_5, z_6\}$. For $K \in \mathcal{T}_{\text{poly}}$, $\mathcal{N}_K = \{z_i\}_{i=1}^7$

with piecewise linear normal component on $\vec{z_2 z_6}$ which now consists of three edges on ∂K_{poly}.

Subsequently, $K_{\text{poly}} \in \mathcal{T}_{\text{poly}}$ shall be denoted by simply $K \in \mathcal{T}_{\text{poly}}$ in the context of flux recovery, and the notion $e \subset \partial K$ denotes an edge on the boundary of K, which takes into account of the edges formed with one end point or both end points as the hanging nodes.

On $\mathcal{T}_{\text{poly}}$, we consider the following Brezzi-Douglas-Marini-type virtual element modification inspired by the ones used in [8, 5]. The local space on a $K \in \mathcal{T}_{\text{poly}}$ is defined as for $k \geq 1$

$$
\mathcal{V}_k(K) := \Big\{ \boldsymbol{\tau} \in \boldsymbol{H}(\text{div}; K) \cap \boldsymbol{H}(\mathbf{rot}; K) :
$$
$$
\nabla \cdot \boldsymbol{\tau} \in \mathbb{P}_{k-1}(K), \quad \nabla \times \boldsymbol{\tau} = 0, \tag{4}
$$
$$
\boldsymbol{\tau} \cdot \boldsymbol{n}_e \in \mathbb{P}_k(e), \ \forall e \subset \partial K \Big\}.
$$

An $\boldsymbol{H}(\text{div})$-conforming global space for recovering the flux is then

$$
\mathcal{V}_k := \big\{ \boldsymbol{\tau} \in \boldsymbol{H}(\text{div}) : \boldsymbol{\tau}|_K \in \mathcal{V}_k(K), \ \text{on } K \in \mathcal{T}_{\text{poly}} \big\}. \tag{5}
$$

Next we turn to define the degrees of freedom (DoFs) of this space. To this end, we define the set of scaled monomials $\mathbb{P}_k(e)$ on an edge e. e is parametrized by $[0, h_e] \ni s \mapsto \boldsymbol{a} + s \boldsymbol{t}_e$, where $\boldsymbol{a}$ is the starting point of e, and $\boldsymbol{t}_e$ is the unit tangential vector of e. The basis set for $\mathbb{P}_k(e)$ is chosen as:

$$
\mathbb{P}_k(e) := \text{span} \left\{ 1, \frac{s - m_e}{h_e}, \left(\frac{s - m_e}{h_e} \right)^2, \ldots, \left(\frac{s - m_e}{h_e} \right)^k \right\}, \tag{6}
$$

where $m_e = h_e/2$ representing the midpoint when using this parametrization. Similar to the edge case, $\mathbb{P}_k(K)$'s basis set is chosen as follows (see e.g., [4]):

$$
\mathbb{P}_k(K) := \text{span} \left\{ m_\alpha(\boldsymbol{x}) := \left(\frac{\boldsymbol{x} - \boldsymbol{x}_K}{h_K} \right)^{\boldsymbol{\alpha}}, \quad |\boldsymbol{\alpha}| \leq k \right\}. \tag{7}
$$

The degrees of freedom (DoFs) are then set as follows for a $\boldsymbol{\tau} \in \mathcal{V}_k$:

$$(\mathfrak{e})\ k \geq 1 \quad \int_e (\boldsymbol{\tau} \cdot \boldsymbol{n}_e) m \, \mathrm{d}s, \quad \forall m \in \mathbb{P}_k(e), \qquad \text{on}\ e \subset \mathcal{E}_{\text{poly}}.$$

$$(\mathfrak{i})\ k \geq 2 \quad \int_K \boldsymbol{\tau} \cdot \nabla m \, \mathrm{d}\boldsymbol{x}, \quad \forall m \in \mathbb{P}_{k-1}(K)/\mathbb{R} \quad \text{on}\ K \in \mathcal{T}_{\text{poly}}. \tag{8}$$

Remark 1. We note that in our construction, the degrees of freedom to determine the curl of a VEM function originally in [8] are replaced by a curl-free constraint thanks to the flexibility to virtual element. The reason why we opt for this subspace is that the true flux $-\alpha \nabla u$ is locally curl-free since we have assumed that α is a piecewise constant. The unisolvency of the set of DoFs (8) including the curl-part can be found in [8]. While for the modified space (4), a simplified argument is in the proof of Lemma 7.3.

3. **Flux recovery.** As the data $f \in L^2(\Omega)$, the true flux $\boldsymbol{\sigma} = -\alpha \nabla u \in \boldsymbol{H}(\text{div})$. Consequently, we shall seek a postprocessed flux $\boldsymbol{\sigma}_{\mathcal{T}}$ in $\mathcal{V}_k \subset \boldsymbol{H}(\text{div})$ by specifying the DoFs in (8). Throughout this section, whenever considering an element $K \in \mathcal{T}$, we treat it a polygon as $K \in \mathcal{T}_{\text{poly}}$.

3.1. **Virtual element-based flux recovery.** Consider $-\alpha_K \nabla u_{\mathcal{T}}$ which is the numerical flux on K. We note that $-\alpha_K \nabla u_{\mathcal{T}}|_K \in \mathbb{P}_{k-1,k}(K) \times \mathbb{P}_{k,k-1}(K)$. The normal flux on each edge $e \in \mathcal{E}_{\text{poly}}$ is in $\mathbb{P}_k(e)$ as $n_e = (\pm 1, 0)$ and $x = \text{const}$ on vertical edges, $n_e = (0, \pm 1)$ and $y = \text{const}$ on horizontal edges. Therefore, the edge-based DoFs can be computed by a simple averaging thanks to the matching polynomial degrees of the numerical flux to the functions in $\mathcal{V}_k$.

On each $e = \partial K_+ \cap \partial K_-$, define

$$\{-\alpha \nabla u_{\mathcal{T}}\}_e^{\gamma_e} \cdot \boldsymbol{n}_e := \Big(\gamma_e \left(-\alpha_{K_-} \nabla u_{\mathcal{T}}|_{K_-} \right) + (1 - \gamma_e) \left(-\alpha_{K_+} \nabla u_{\mathcal{T}}|_{K_+} \right) \Big) \cdot \boldsymbol{n}_e, \tag{9}$$

where

$$\gamma_e := \frac{\alpha_{K_+}^{1/2}}{\alpha_{K_+}^{1/2} + \alpha_{K_-}^{1/2}}. \tag{10}$$

First for both $k = 1$ and $k \geq 2$ cases, we set the normal component of the recovered flux is set as

$$\boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_e = \{-\alpha \nabla u_{\mathcal{T}}\}_e^{\gamma_e} \cdot \boldsymbol{n}_e. \tag{11}$$

In the lowest order case $k = 1$, $\nabla \cdot \boldsymbol{\sigma}_{\mathcal{T}}$ is a constant on K by (4), thus the construction (11) alone, which consists the edge DoFs $(\mathfrak{e})$ in (8), can determine the divergence $\nabla \cdot \boldsymbol{\sigma}_{\mathcal{T}}$ in K as follows

$$|K| \nabla \cdot \boldsymbol{\sigma}_{\mathcal{T}} = \int_K \nabla \cdot \boldsymbol{\sigma}_{\mathcal{T}} \mathrm{d}\boldsymbol{x} = \int_{\partial K} \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_{\partial K} \mathrm{d}s = \sum_{e \subset \partial K} \int_e \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_{\partial K}|_e \mathrm{d}s. \tag{12}$$

If $k \geq 2$, after the normal component (11) is set, furthermore on each K, denote Π_{k-1} stands for the L^2-projection to $\mathbb{P}_{k-1}(K)$, and we let

$$\nabla \cdot \boldsymbol{\sigma}_{\mathcal{T}} = \Pi_{k-1} f + c_K. \tag{13}$$

The reason to add c_K is that we have set the normal components of the recovered flux first without relying on the divergence information. While in general $\nabla \cdot \boldsymbol{\sigma}_{\mathcal{T}} \neq \Pi_{k-1} f$ as otherwise the divergence theorem will be rendered invalid in (12). As a result, an element-wise constant c_K is added to ensure the compatibility of $\boldsymbol{\sigma}_{\mathcal{T}}$ locally on each K. It is straightforward to verify that c_K has the following form,

and later we shall show that c_K does not affect the efficiency as well as the reliability of the error estimates.

$$c_K = \frac{1}{|K|}\left(-\int_K \Pi_{k-1}f\,\mathrm{d}\boldsymbol{x} + \sum_{e\subset\partial K}\int_e \{-\alpha\nabla u_\mathcal{T}\}_e^{\gamma_e}\cdot\boldsymbol{n}_{\partial K}|_e\mathrm{d}s\right), \qquad (14)$$

Consequently for $k \geq 2$, the set (i) of DoFs can be set as: $\forall q \in \mathbb{P}_{k-1}(K)$

$$\left(\boldsymbol{\sigma}_\mathcal{T},\nabla q\right)_K = -\left(\Pi_{k-1}f + c_K, q\right)_K + \sum_{e\subset\partial K}\left(\{-\alpha\nabla u_\mathcal{T}\}_e^{\gamma_e}\cdot\boldsymbol{n}_{\partial K}|_e, q\right)_e. \qquad (15)$$

3.2. **Locally projected flux.** To the end of constructing a computable local error indicator, inspired by the VEM formulation [8], the recovered flux is projected to a space with a much simpler structure. A local oblique projection $\Pi : \boldsymbol{L}^2(K) \to \nabla\mathbb{P}_k(K)$, $\boldsymbol{\tau} \mapsto \Pi\boldsymbol{\tau}$ is defined as follows:

$$\left(\Pi\boldsymbol{\tau},\nabla p\right)_K = \left(\boldsymbol{\tau},\nabla p\right)_K, \quad \forall p \in \mathbb{P}_k(K)/\mathbb{R}. \qquad (16)$$

Next we are gonna show that this projection operator can be straightforward computed for vector fields in $\mathcal{V}_k(K)$.

3.2.1. $k = 1$. When $k = 1$, we can compute the right hand side of (16) as follows:

$$\left(\boldsymbol{\tau},\nabla p\right)_K = -\left(\nabla\cdot\boldsymbol{\tau}, p\right)_K + \left(\boldsymbol{\tau}\cdot\boldsymbol{n}, p\right)_{\partial K}. \qquad (17)$$

By definition of the space (4) when $k = 1$, $\nabla\cdot\boldsymbol{\tau}$ is a constant on K and can be determined by edge DoFs (e) in (8) similar to (12). Moreover, $p|_e \in \mathbb{P}_1(e)$, thus the boundary term can be evaluated using DoFs (e) in (8).

3.2.2. $k \geq 2$. When $k \geq 2$, the right hand side of (16) can be evaluated following a similar procedure as (17), if we exploit the fact that $\nabla\cdot\boldsymbol{\tau} \in \mathbb{P}_{k-1}(K)$, we have

$$\begin{aligned}\left(\boldsymbol{\tau},\nabla p\right)_K &= -\left(\nabla\cdot\boldsymbol{\tau}, \Pi_{k-1}p\right)_K + \left(\boldsymbol{\tau}\cdot\boldsymbol{n}, p\right)_{\partial K} \\ &= \left(\boldsymbol{\tau},\nabla\Pi_{k-1}p\right)_K + \left(\boldsymbol{\tau}\cdot\boldsymbol{n}, p - \Pi_{k-1}p\right)_{\partial K},\end{aligned} \qquad (18)$$

which can be evaluated using both DoF sets (e) and (i).

4. **A posteriori error estimation.** Given the recovered flux $\boldsymbol{\sigma}_\mathcal{T}$ in Section 3, the recovery-based local error indicator $\eta_{\mathrm{flux},K}$ and the element residual $\eta_{\mathrm{res},K}$ as follows:

$$\begin{aligned}\eta_{\mathrm{flux},K} &:= \left\|\alpha^{-1/2}(\boldsymbol{\sigma}_\mathcal{T} + \alpha\nabla u_\mathcal{T})\right\|_K, \\ \text{and } \eta_{\mathrm{res},K} &:= \left\|\alpha^{-1/2}(f - \nabla\cdot\boldsymbol{\sigma}_\mathcal{T})\right\|_K,\end{aligned} \qquad (19)$$

then

$$\eta_K = \begin{cases} \eta_{\mathrm{flux},K} & \text{when } k = 1, \\ \left(\eta_{\mathrm{flux},K}^2 + \eta_{\mathrm{res},K}^2\right)^{1/2} & \text{when } k \geq 2. \end{cases} \qquad (20)$$

A computable $\widehat{\eta}_{\mathrm{flux},K}$ is defined as:

$$\widehat{\eta}_{\mathrm{flux},K} := \left\|\alpha_K^{-1/2}\Pi(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K\nabla u_\mathcal{T})\right\|_K, \qquad (21)$$

with the oblique projection Π defined in (16). The stabilization part $\widehat{\eta}_{\mathrm{stab},K}$ is

$$\widehat{\eta}_{\mathrm{stab},K} := \left|\alpha_K^{-1/2}(\mathrm{I}-\Pi)(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K\nabla u_\mathcal{T})\right|_{S,K}. \qquad (22)$$

Here $|\cdot|_{S,K} := \left(S_K(\cdot,\cdot)\right)^{1/2}$ is seminorm induced by the following stabilization

$$S_K(\boldsymbol{v},\boldsymbol{w}) := \sum_{e\subset\partial K} h_e\left(\boldsymbol{v}\cdot\boldsymbol{n}_e, \boldsymbol{w}\cdot\boldsymbol{n}_e\right)_e + \sum_{\alpha\in\Lambda}(\boldsymbol{v},\nabla m_\alpha)_K(\boldsymbol{w},\nabla m_\alpha)_K, \qquad (23)$$

where Λ is the index set for the monomial basis of $\mathbb{P}_{k-1}(K)/\mathbb{R}$ with cardinality $k(k+1)/2 - 1$, i.e., the second term in (23) is dropped in the $k = 1$ case. We note that this is a slightly modified version of the standard stabilization for an $\boldsymbol{H}(\mathrm{div})$-function in [8] as we have replaced the edge DoFs by an integral. In Section 7.1 it is shown that the integral-based stabilization still yields the crucial norm equivalence result.

The computable error estimator $\widehat{\eta}$ is then

$$\widehat{\eta}^2 = \begin{cases} \sum_{K \in \mathcal{T}} \left(\widehat{\eta}_{\mathrm{flux},K}^2 + \widehat{\eta}_{\mathrm{stab},K}^2 \right) =: \sum_{K \in \mathcal{T}} \widehat{\eta}_K^2 & \text{when } k = 1, \\ \sum_{K \in \mathcal{T}} \left(\widehat{\eta}_{\mathrm{flux},K}^2 + \widehat{\eta}_{\mathrm{stab},K}^2 + \eta_{\mathrm{res},K}^2 \right) =: \sum_{K \in \mathcal{T}} \widehat{\eta}_K^2 & \text{when } k \geq 2. \end{cases} \tag{24}$$

4.1. **Efficiency.** In this section, we shall prove the proposed recovery-based estimator $\widehat{\eta}_K$ is efficient by bounding it above by the residual-based error estimator. In the process of adaptive mesh refinement, only the computable $\widehat{\eta}_K$ is used as the local error indicator to guide a marking strategy of choice.

Theorem 4.1. *Let $u_\mathcal{T}$ be the solution to problem (2), and $\widehat{\eta}_{\mathrm{flux},K}$ be the error indicator in (24). On $K \in \mathcal{T}_{\mathrm{poly}}$, $\widehat{\eta}_{\mathrm{flux},K}$ can be locally bounded by the residual-based ones:*

$$\widehat{\eta}_{\mathrm{flux},K}^2 \lesssim \mathrm{osc}(f;K)^2 + \eta_{\mathrm{elem},K}^2 + \eta_{\mathrm{edge},K}^2, \tag{25}$$

where

$$\mathrm{osc}(f;K) = \alpha_K^{-1/2} h_K \big\| f - \Pi_{k-1} f \big\|_K,$$

$$\eta_{\mathrm{elem},K} := \alpha_K^{-1/2} h_K \big\| f + \nabla \cdot (\alpha \nabla u_\mathcal{T}) \big\|_K,$$

$$\text{and } \eta_{\mathrm{edge},K} := \left(\sum_{e \subset \partial K} \frac{h_e}{\alpha_K + \alpha_{K_e}} \big\| [\![\alpha \nabla u_\mathcal{T} \cdot \boldsymbol{n}_e]\!] \big\|_e^2 \right)^{1/2}.$$

In the edge jump term, K_e is the element on the opposite side of K with respect to an edge $e \subset \partial K$. The constant depends on k and the number of edges on ∂K.

Proof. Let $\alpha_K^{-1} \Pi(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}) =: \nabla p$ on K, then $p \in \mathbb{P}_k(K)/\mathbb{R}$ and we have

$$\widehat{\eta}_{\mathrm{flux},K}^2 = \big(\Pi(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}), \nabla p \big)_K = \big(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}, \nabla p \big)_K$$

$$= -\big(\nabla \cdot (\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}), p \big)_K + \sum_{e \subset \partial K} \int_e \big(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T} \big) \cdot \boldsymbol{n}_{\partial K} \big|_e \, p \, \mathrm{d}s. \tag{26}$$

By (11), without loss of generality we assume $K = K_-$ (the local orientation of e agrees with the global one, i.e., $\boldsymbol{n}_{\partial K}\big|_e = \boldsymbol{n}_e$), and $K_e = K_+$ which is the element opposite to K with respect to e, and $\gamma_e := \alpha_{K_e}^{1/2}/(\alpha_{K_e}^{1/2} + \alpha_K^{1/2})$, we have on edge $e \subset \partial K$

$$\big(\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T} \big) \cdot \boldsymbol{n}_e = \left((1 - \gamma_e) \alpha_K \nabla u_\mathcal{T}|_K - (1 - \gamma_e) \alpha_{K_e} \nabla u_\mathcal{T}|_{K_e} \right) \cdot \boldsymbol{n}_e$$

$$= \frac{\alpha_K^{1/2}}{\alpha_K^{1/2} + \alpha_{K_e}^{1/2}} [\![\alpha \nabla u_\mathcal{T} \cdot \boldsymbol{n}_e]\!]_e. \tag{27}$$

The boundary term in (26) can be then rewritten as

$$
\int_e \big(\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}\big) \cdot \boldsymbol{n}_e\, p \, \mathrm{d}s
$$
$$
= \int_e \frac{1}{\alpha_K^{1/2} + \alpha_{K_e}^{1/2}} [\![\alpha \nabla u_{\mathcal{T}} \cdot \boldsymbol{n}_e]\!]_e\, \alpha_K^{1/2} p \, \mathrm{d}s \tag{28}
$$
$$
\lesssim \frac{1}{(\alpha_K + \alpha_{K_e})^{1/2}} h_e^{1/2} \big\| [\![\alpha \nabla u_{\mathcal{T}} \cdot \boldsymbol{n}_e]\!] \big\|_e\, \alpha_K^{1/2} h_e^{-1/2} \left\| p \right\|_e.
$$

By a trace inequality on an edge of a polygon (Lemma 7.1), and the Poincaré inequality for $p \in \mathbb{P}_k(K)/\mathbb{R}$, we have,

$$
h_e^{-1/2} \|p\|_e \lesssim h_K^{-1} \|p\|_K + \|\nabla p\|_K \lesssim \|\nabla p\|_K.
$$

As a result,

$$
\sum_{e \subset \partial K} \int_e \big(\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}\big) \cdot \boldsymbol{n}_e\, p \, \mathrm{d}s \lesssim \eta_{\mathrm{edge},K}\, \alpha_K^{1/2} \left\| \nabla p \right\|_e = \eta_{\mathrm{edge},K}\, \widehat{\eta}_{\mathrm{flux},K}.
$$

For the bulk term on K's in (26), when $k = 1$, by (12), the representation in (28), and the Poincaré inequality for $p \in \mathbb{P}_k(K)/\mathbb{R}$ again with $h_K \simeq |K|^{1/2}$, we have

$$
-\big(\nabla \cdot (\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}), p\big)_K \leq |\nabla \cdot (\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}})| \, |K|^{1/2} \, \|p\|_K
$$
$$
\leq \frac{1}{|K|^{1/2}} \left| \int_K \nabla \cdot (\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}) \, \mathrm{d}\boldsymbol{x} \right| \|p\|_K
$$
$$
= \frac{1}{|K|^{1/2}} \left| \sum_{e \subset \partial K} \int_e (\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}) \cdot \boldsymbol{n}_e \, \mathrm{d}s \right| \|p\|_K
$$
$$
\leq \left(\sum_{e \subset \partial K} \frac{1}{\alpha_K^{1/2} + \alpha_{K_e}^{1/2}} \big\| [\![\alpha \nabla u_{\mathcal{T}} \cdot \boldsymbol{n}_e]\!] \big\|_e\, \alpha_K^{1/2} h_e \right) \|\nabla p\|
$$
$$
\lesssim \eta_{\mathrm{edge},K}\, \widehat{\eta}_{\mathrm{flux},K}.
$$

When $k \geq 2$, by (13),

$$
-\big(\nabla \cdot (\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}), p\big)_K = -\big(\Pi_{k-1} f + c_K + \nabla \cdot (\alpha_K \nabla u_{\mathcal{T}}), p\big)_K
$$
$$
\leq \left(\big\| f - \Pi_{k-1} f \big\|_K + \big\| f + \nabla \cdot (\alpha \nabla u_{\mathcal{T}}) \big\|_K + |c_K| |K|^{1/2} \right) \|p\|_K. \tag{29}
$$

The first two terms can be handled by combining the weights $\alpha^{-1/2}$ and h_K from $\|p\|_K \leq h_K \|\nabla p\|_K$. For c_K, it can be estimated straightforwardly as follows

$$
c_K |K|^{1/2} = \frac{1}{|K|^{1/2}} \Big(-\int_K (\Pi_{k-1} f - f) \mathrm{d}\boldsymbol{x} - \int_K \big(f + \nabla \cdot (\alpha \nabla u_{\mathcal{T}})\big) \mathrm{d}\boldsymbol{x}
$$
$$
+ \int_K \nabla \cdot (\alpha \nabla u_{\mathcal{T}}) \mathrm{d}\boldsymbol{x} + \sum_{e \subset \partial K} \int_e \{ -\alpha \nabla u_{\mathcal{T}} \}_e^{\gamma_e} \cdot \boldsymbol{n}_e \mathrm{d}s \Big)
$$
$$
\leq \big\| f - \Pi_{k-1} f \big\|_K + \big\| f + \nabla \cdot (\alpha \nabla u_{\mathcal{T}}) \big\|_K
$$
$$
+ \frac{1}{|K|^{1/2}} \sum_{e \subset \partial K} \int_e (\alpha_K \nabla u_{\mathcal{T}} - \{ \alpha \nabla u_{\mathcal{T}} \}_e^{\gamma_e}) \cdot \boldsymbol{n}_e \mathrm{d}s \tag{30}
$$
$$
\leq \big\| f - \Pi_{k-1} f \big\|_K + \big\| f + \nabla \cdot (\alpha \nabla u_{\mathcal{T}}) \big\|_K
$$
$$
+ \sum_{e \subset \partial K} \frac{\alpha_K^{1/2}}{\alpha_K^{1/2} + \alpha_{K_e}^{1/2}} \big\| [\![\alpha \nabla u_{\mathcal{T}} \cdot \boldsymbol{n}_e]\!] \big\|_e.
$$

The two terms on K can be treated the same way with the first two terms in (29) while the edge terms are handled similarly as in the $k = 1$ case. As a result, we have shown

$$-\big(\nabla \cdot (\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}}), p\big)_K \lesssim \Big(\mathrm{osc}(f; K) + \eta_{\mathrm{elem}, K} + \eta_{\mathrm{edge}, K}\Big) \alpha_K^{1/2} \|\nabla p\|$$

and the theorem follows.   □

Theorem 4.2. *Under the same setting with Theorem 4.1, let $\widehat{\eta}_{\mathrm{stab}, K}$ as the estimator in (22), we have*

$$\widehat{\eta}_{\mathrm{stab}, K}^2 \lesssim \mathrm{osc}(f; K)^2 + \eta_{\mathrm{elem}, K}^2 + \eta_{\mathrm{edge}, K}^2, \tag{31}$$

The constant depends on k and the number of edges on ∂K.

Proof. This theorem follows directly from the norm equivalence Lemma 7.3:

$$\big|\alpha_K^{-1/2}(\mathrm{I} - \Pi)(\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}})\big|_{S, K} \lesssim \big|\alpha_K^{-1/2}(\boldsymbol{\sigma}_{\mathcal{T}} + \alpha_K \nabla u_{\mathcal{T}})\big|_{S, K},$$

while evaluating the DoFs ($\mathfrak{e}$) and ($\mathfrak{i}$) using (11) and (15) reverts us back to the proof of Theorem 4.1.   □

Theorem 4.3. *Under the same setting with Theorem 4.1, on any $K \in \mathcal{T}_{\mathrm{poly}}$ with ω_K defined as the collection of elements in $\mathcal{T}$ which share at least 1 vertex with K*

$$\widehat{\eta}_K \lesssim \mathrm{osc}(f; K) + \big\|\alpha^{1/2} \nabla(u - u_{\mathcal{T}})\big\|_{\omega_K}, \tag{32}$$

with a constant independent of α, but dependent on k and the maximum number of edges in $K \in \mathcal{T}_{\mathrm{poly}}$.

Proof. This is a direct consequence of Theorem 4.1 and 4.2 and the fact that the residual-based error indicator is efficient by a common bubble function argument.
                                                        □

4.2. **Reliability.** In this section, we shall prove that the computable error estimator $\widehat{\eta}$ is reliable under two common assumptions in the *a posteriori* error estimation literature. For the convenience of the reader, we rephrase them here using a "layman" description, for more detailed and technical definition please refer to the literature cited.

Assumption 1 ($\mathcal{T}$ is l-irregular [14]). *Any given $\mathcal{T}$ is always refined from a mesh with no hanging nodes by a quadsecting red-refinement. For any two neighboring elements in $\mathcal{T}$, the difference in their refinement levels is $\leq l$ for a uniformly bounded constant l, i.e., for any edge $e \in \mathcal{E}$, it has at most l hanging nodes.*

By Assumption 1, we denote the father 1-irregular mesh of $\mathcal{T}$ as $\mathcal{T}_1$. On $\mathcal{T}_1$, a subset of all nodes is denoted by $\mathcal{N}_1$, which includes the regular nodes $\mathcal{N}_R$ on $\mathcal{T}_1$, as well as $\mathcal{N}_E$ as the set of end points of edges with a hanging node as the midpoint. By [14, Theorem 2.1], there exists a set of bilinear nodal bases $\{\phi_z\}$ associated with $z \in \mathcal{N}_1$, such that $\{\phi_z\}$ form a partition of unity and can be used to construct a Clément-type quasi-interpolation. Furthermore, the following assumption assures that the Clément-type quasi-interpolant is robust with respect to the coefficient distribution on a vertex patch, when taking nodal DoFs as a weighted average.

Assumption 2 (Quasi-monotonicity of α [6]). *On $\mathcal{T}$, let ϕ_z be the bilinear nodal basis associated with $z \in \mathcal{N}_1$, with $\omega_z := \mathrm{supp}\,\phi_z$. For every element $K \subset \omega_z, K \in \mathcal{T}$, there exists a simply connected element path leading to $\omega_{m(z)}$, which is a Lipschitz*

domain containing the elements where the piecewise constant coefficient α achieves the maximum (or minimum) on ω_z.

Denote

$$\pi_z v = \begin{cases} \dfrac{\int_{\omega_z \cap \omega_{m(z)}} v\phi_z}{\int_{\omega_z \cap \omega_{m(z)}} \phi_z} & \text{if } z \in \Omega, \\ 0 & \text{if } z \in \partial\Omega. \end{cases} \tag{33}$$

We note that if α is a constant on ω_z, $(1, (v - \pi_z v)\phi_z)_{\omega_z} = 0$. A quasi-interpolation $\mathcal{I} : L^2(\Omega) \to \mathcal{Q}_1(\mathcal{T}_1)$ can be defined as

$$\mathcal{I}v := \sum_{z \in \mathcal{N}_1} (\pi_z v)\phi_z. \tag{34}$$

Lemma 4.4 (Estimates for π_z and $\mathcal{I}$). *Under Assumption 1 and 2, the following estimates hold for any $v \in H^1(\omega_K)$*

$$\alpha_K^{1/2} h_K^{-1} \|v - \mathcal{I}v\|_K + \alpha_K^{1/2} \|\nabla \mathcal{I}v\|_K \lesssim \left\|\alpha^{1/2}\nabla v\right\|_{\omega_K}, \tag{35}$$

and for $z \in \mathcal{N}_1$

$$\sum_{K \subset \omega_z} h_z^{-2} \|\alpha^{1/2}(v - \pi_z v)\phi_z\|_K^2 \lesssim \left\|\alpha^{1/2}\nabla v\right\|_{\omega_z}^2, \tag{36}$$

in which $h_z := \max_{K \subset \omega_z} h_K$, and here ω_K denotes the union of elements in $\mathcal{T}_1$ sharing at least a node (hanging or regular) with K.

Proof. The estimate for π_z follows from [6, Lemma 2.8]. For $\mathcal{I}$, its error estimates and stability only rely on the partition of unity property of the nodal basis set $\{\phi_z\}$ (see e.g., [27]), therefore the proof follows the same argument with the ones used on triangulations in [6, Lemma 2.8]. □

Denotes the subset of nodes $\{z\} \subset \mathcal{N}_1$ (i) on the boundary as $\mathcal{N}_{\partial\Omega}$ and (ii) with the coefficient α on patch ω_z as $\mathcal{N}_I$. For the lowest order case, we need the following oscillation term for f

$$\begin{aligned} \operatorname{osc}(f; \mathcal{T})^2 := &\sum_{z \in \mathcal{N}_1 \cap (\mathcal{N}_{\partial\Omega} \cup \mathcal{N}_I)} h_z^2 \|\alpha^{-1/2} f\|_{\omega_z}^2 \\ &+ \sum_{z \in \mathcal{N}_1 \setminus (\mathcal{N}_{\partial\Omega} \cup \mathcal{N}_I)} h_z^2 \|\alpha^{-1/2}(f - f_z)\|_{\omega_z}^2, \end{aligned} \tag{37}$$

with $f_z := \int_{\omega_z} v\phi_z / \int_{\omega_z} \phi_z$.

Theorem 4.5. *Let $u_\mathcal{T}$ be the solution to problem (2), and $\widehat{\eta}$ be the computable error estimator in (24), under Assumption 2 and 1, we have for $k = 1$*

$$\left\|\alpha^{1/2}\nabla(u - u_\mathcal{T})\right\| \lesssim \left(\widehat{\eta}^2 + \operatorname{osc}(f; \mathcal{T})^2\right)^{1/2}. \tag{38}$$

For $k \geq 2$,

$$\left\|\alpha^{1/2}\nabla(u - u_\mathcal{T})\right\| \lesssim \widehat{\eta}, \tag{39}$$

where the constant depends on l and k.

Proof. Let $\varepsilon := u - u_\mathcal{T} \in H_0^1(\Omega)$, and $\mathcal{I}\varepsilon \in \mathcal{Q}_1(\mathcal{T}_1) \subset \mathcal{Q}_1(\mathcal{T})$ be the quasi-interpolant in (34) of ε, then by the Galerkin orthogonality, $\alpha \nabla u + \boldsymbol{\sigma}_\mathcal{T} \in \boldsymbol{H}(\mathrm{div})$, the Cauchy-Schwarz inequality, and the interpolation estimates (35), we have for $k \geq 2$,

$$\left\| \alpha^{1/2} \nabla \varepsilon \right\|^2 = \left(\alpha \nabla(u - u_\mathcal{T}), \nabla(\varepsilon - \mathcal{I}\varepsilon) \right)$$
$$= \left(\alpha \nabla u + \boldsymbol{\sigma}_\mathcal{T}, \nabla(\varepsilon - \mathcal{I}\varepsilon) \right) - \left(\alpha \nabla u_\mathcal{T} + \boldsymbol{\sigma}_\mathcal{T}, \nabla(\varepsilon - \mathcal{I}\varepsilon) \right)$$
$$= \left(f - \nabla \cdot \boldsymbol{\sigma}_\mathcal{T}, \varepsilon - \mathcal{I}\varepsilon \right) - \left(\alpha \nabla u_\mathcal{T} + \boldsymbol{\sigma}_\mathcal{T}, \nabla(\varepsilon - \mathcal{I}\varepsilon) \right)$$
$$\leq \left(\sum_{K \in \mathcal{T}} \alpha_K^{-1} h_K^2 \| f - \nabla \cdot \boldsymbol{\sigma}_\mathcal{T} \|_K^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}} \alpha_K h_K^{-2} \| \varepsilon - \mathcal{I}\varepsilon \|_K^2 \right)^{1/2}$$
$$\left(\sum_{K \in \mathcal{T}} \alpha_K^{-1} \| \alpha \nabla u_\mathcal{T} + \boldsymbol{\sigma}_\mathcal{T} \|_K^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}} \alpha_K \| \nabla(\varepsilon - \mathcal{I}\varepsilon) \|_K^2 \right)^{1/2}.$$
$$\lesssim \left(\sum_{K \in \mathcal{T}} (\eta_{\mathrm{res},K}^2 + \eta_{\mathrm{flux},K}^2) \right)^{1/2} \left(\sum_{K \in \mathcal{T}} \left\| \alpha^{1/2} \nabla \varepsilon \right\|_{\omega_K} \right)^{1/2}.$$

Applying the norm equivalence of η to $\widehat{\eta}$ by Lemma 7.3, as well as the fact that the number of elements in ω_K is uniformly bounded by Assumption 1, yields the desired estimate.

When $k = 1$, the residual term on K can be further split thanks to $\Delta \mathbb{Q}_1(K) = \{0\}$. First we notice that by the fact that $\{\phi_z\}$ form a partition of unity,

$$(f, \varepsilon - \mathcal{I}\varepsilon) = \sum_{z \in \mathcal{N}_1} \sum_{K \subset \omega_z} \left(f, (\varepsilon - \pi_z \varepsilon) \phi_z \right)_K, \tag{40}$$

in which a patch-wise constant f_z (weighted average of f) can be further inserted by the definition of π_z (33) if α is a constant on ω_z. Therefore, by the assumption of α_K being a piecewise constant, splitting (40), we have

$$\left(f - \nabla \cdot \boldsymbol{\sigma}_\mathcal{T}, \varepsilon - \mathcal{I}\varepsilon \right) = \left(f, \varepsilon - \mathcal{I}\varepsilon \right) - \left(\nabla \cdot (\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}), \varepsilon - \mathcal{I}\varepsilon \right)$$
$$= \sum_{z \in \mathcal{N}} \sum_{K \subset \omega_z} \left(f, (\varepsilon - \pi_z \varepsilon) \phi_z \right)_K - \left(\nabla \cdot (\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}), \varepsilon - \mathcal{I}\varepsilon \right)$$
$$\leq \left(\mathrm{osc}(f; \mathcal{T})^2 \right)^{1/2} \left(\sum_{z \in \mathcal{N}_1} \sum_{K \subset \omega_z} h_z^{-2} \| \alpha^{1/2} (\varepsilon - \pi_z \varepsilon) \phi_z \|_K^2 \right)^{1/2}$$
$$+ \left(\sum_{K \in \mathcal{T}} \alpha_K^{-1} h_K^2 \| \nabla \cdot (\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}) \|_K^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}} \alpha_K h_K^{-2} \| \varepsilon - \mathcal{I}\varepsilon \|_K^2 \right)^{1/2}.$$

Applied an inverse inequality in Lemma 7.2 on $\left\| \nabla \cdot (\boldsymbol{\sigma}_\mathcal{T} + \alpha_K \nabla u_\mathcal{T}) \right\|_K$ and the projection estimate for π_z (36), the rest follows the same argument with the one used in the $k \geq 2$ case. $\square$

5. **Numerical examples.** The numerics is prepared using the bilinear element for common AMR benchmark problems. The codes for this paper are publicly available on https://github.com/lyc102/ifem implemented using *i*FEM [16]. The linear algebraic system on an *l*-irregular quadtree is implemented following the conforming prolongation approach [15] by $\mathbf{P}^\top \mathbf{A} \mathbf{P} \mathbf{u} = \mathbf{P}^\top \mathbf{f}$, where $\mathbf{A}$ is the locally assembled stiffness matrix for all nodes in $\mathcal{N}$, $\mathbf{u}$ and $\mathbf{f}$ are the solution vector associated with $\mathcal{N}_R$ and load vector associated with $\mathcal{N}$, respectively. $\mathbf{P} = (\mathbf{I}, \mathbf{W})^\top : \mathbb{R}^{\dim \mathcal{N}_R} \to$

$\mathbb{R}^{\dim \mathcal{N}}$ is a prolongation operator mapping conforming H^1-bilinear finite element function defined on regular nodes to all nodes, the weight matrix $\mathbf{W}$ is assembled locally by a recursive kNN query in $\mathcal{N}_H$, while the polygonal mesh data structure embedding is automatically built during constructing $\mathbf{P}$. For details we refer the readers to https://github.com/lyc102/ifem/tree/master/research/polyFEM.

The adaptive finite element (AFEM) iterative procedure is following the standard

$$\texttt{SOLVE} \longrightarrow \texttt{ESTIMATE} \longrightarrow \texttt{MARK} \longrightarrow \texttt{REFINE}.$$

The linear system is solved by MATLAB `mldivide`. In `MARK`, the Dörfler L^2-marking is used with the local error indicator $\widehat{\eta}_K$ in that the minimum subset $\mathcal{M} \subset \mathcal{T}$ is chosen such that

$$\sum_{K \in \mathcal{M}} \widehat{\eta}_K^2 \geq \theta \sum_{K \in \mathcal{T}} \widehat{\eta}_K^2, \quad \text{for } \theta \in (0, 1).$$

Throughout all examples, we fix $\theta = 0.3$. $\mathcal{T}$ is refined by a red-refinement by quadsecting the marked element afterwards. For comparison, we compute the standard residual-based local indicator for $K \in \mathcal{T}_{\text{poly}}$

$$\eta_{\text{Residual},K}^2 := \alpha_K^{-1} h_K^2 \big\| f + \nabla \cdot (\alpha \nabla u_{\mathcal{T}}) \big\|_K^2 + \frac{1}{2} \sum_{e \subset \partial K} \frac{h_e}{\alpha_K + \alpha_{K_e}} \big\| [\![\alpha \nabla u_{\mathcal{T}} \cdot \boldsymbol{n}_e]\!] \big\|_e^2,$$

Let $\eta_{\text{Residual}}^2 = \sum_{K \in \mathcal{T}} \eta_{\text{Residual},K}^2$. The residual-based estimator η_{Residual} is merely computed for comparison purpose and not used in marking. The AFEM procedure stops when the relative error reaches a threshold. The effectivity indices for different estimators are compared

$$\text{effectivity index} := \eta / \big\| \alpha^{1/2} \nabla \varepsilon \big\|, \quad \text{where } \varepsilon := u - u_{\mathcal{T}}, \ \eta = \eta_{\text{Residual}} \text{ or } \widehat{\eta},$$

i.e., the closer to 1 the effectivity index is, the more accurate this estimator is to measure the error of interest. We use an order 5 Gaussian quadrature to compute $\|\alpha^{1/2} \nabla (u - u_{\mathcal{T}})\|$ elementwisely. The orders of convergence for various η's and $\|\alpha^{1/2} \nabla (u - u_{\mathcal{T}})\|$ are computed, for which r_η and r_{err} are defined as the slope for the linear fitting of $\ln \eta_n$ and $\ln \|\alpha^{1/2} \nabla (u - u_{\mathcal{T},n})\|$ in the asymptotic regime,

$$\ln \eta_n \sim -r_\eta \ln N_n + c_1, \quad \text{and} \quad \ln \|\alpha^{1/2} \nabla (u - u_{\mathcal{T}})\| \sim -r_{\text{err}} \ln N_n + c_2,$$

where the subscript n stands for the number of iteration in the AFEM cycles, $N_n := \#(\mathcal{N}_R \backslash \mathcal{N}_{\partial\Omega})$. r_η and r_{err} are considered optimal when being close to $1/2$.

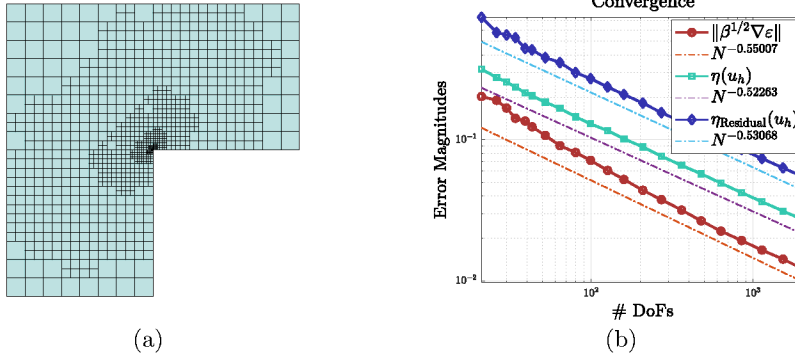(a)   (b)

FIGURE 2. The result of the L-shape example. (a) The adaptively refined mesh with 1014 DoFs. (b) Convergence in Example 1.

5.1. **L-shaped domain.** In this example, a standard AMR benchmark on the L-shaped domain is tested. The true solution $u = r^{2/3}\sin(2\theta/3)$ in polar coordinates on $\Omega = (-1, 1) \times (-1, 1) \backslash [0, 1) \times (-1, 0]$. The AFEM procedure stops if the relative error has reached 0.01. The adaptively refined mesh can be found in Figure 2a. While both estimators show optimal rate of convergence in Figure 2b, the effectivity index for η_{Residual} is 4.52, and is 2.24 for $\widehat{\eta}$.

5.2. **A circular wave front.** The solution $u = \tan^{-1}(\alpha(r - r_0))$ is defined on $\Omega = (0, 1)^2$ with $r := \sqrt{(x + 0.05)^2 + (y + 0.05)^2}$, $\alpha = 100$, and $r_0 = 0.7$. The true solution shows a sharp transition layer (Figure 3a). The result of the convergence can be found in Figure 3b. In this example, the AFEM procedure stops if the relative error has reached 0.05. Additionally, we note that by allowing l-irregular ($l \geq 2$), the AMR procedure shows to be more efficient toward capturing the singularity of the solution. A simple comparison can be found in Figure 4. The effectivity indices for η_{Residual} and $\widehat{\eta}$ are 5.49 and 2.08, respectively.
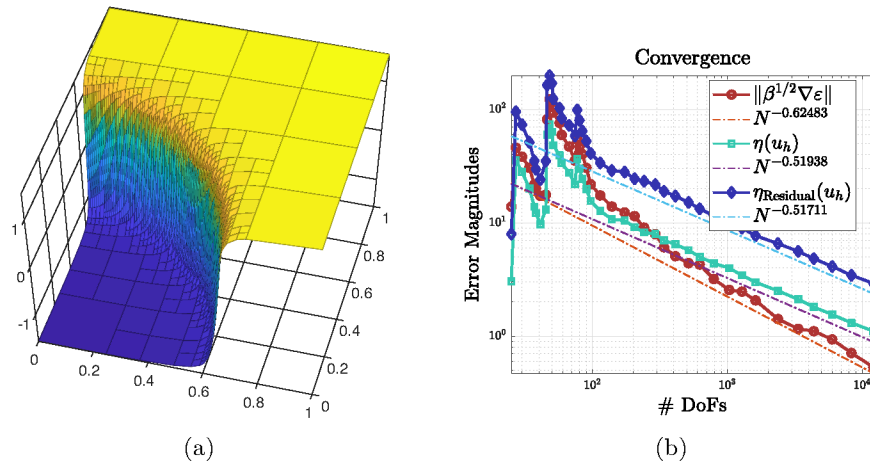

(a)   (b)

FIGURE 3. The result of the circular wave front example. (a) $u_{\mathcal{T}}$ on a 3-irregular mesh with #DoFs = 1996, the relative error is 14.3%. (b) Convergence in Example 2.

5.3. **Kellogg benchmark.** This example is a common benchmark test problem introduced in [9], see also [17, 12]) for elliptic interface problems. The true solution $u = r^\gamma \mu(\theta)$ is harmonic in four quadrants, and $\mu(\theta)$ takes different values within four quadrants:

$$\mu(\theta) = \begin{cases} \cos((\pi/2 - \delta)\gamma) \cdot \cos((\theta - \pi/2 + \rho)\gamma) & \text{if } 0 \leq \theta \leq \pi/2 \\ \cos(\rho\gamma) \cdot \cos((\theta - \pi + \delta)\gamma) & \text{if } \pi/2 \leq \theta \leq \pi \\ \cos(\delta\gamma) \cdot \cos((\theta - \pi - \rho)\gamma) & \text{if } \pi \leq \theta < 3\pi/2 \\ \cos((\pi/2 - \rho)\gamma) \cdot \cos((\theta - 3\pi/2 - \delta)\gamma) & \text{if } 3\pi/2 \leq \theta \leq 2\pi \end{cases}$$

While $\alpha = R$ in the first and third quadrants, and $\alpha = 1$ in the second and fourth quadrants, and the true flux $\alpha\nabla u$ is glued together using $\boldsymbol{H}(\text{div})$-continuity conditions. We choose the following set of coefficients for u

$$\gamma = 0.1, \quad R \approx 161.4476387975881, \quad \rho = \pi/4, \quad \delta \approx -14.92256510455152,$$
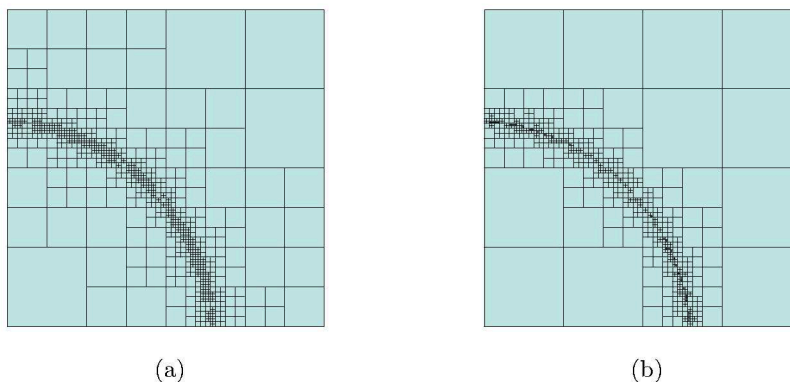
(a)   (b)

FIGURE 4. Comparison of the adaptively refined meshes. (a) 1-irregular mesh, #DoFs = 1083, the relative error is 21.8%. (b) 4-irregular mesh, and #DoFs = 1000, the relative error is 17.8%.

By this choice, this function is very singular near the origin as the maximum regularity it has is $H^{1+\gamma}_{loc}(\Omega\backslash\{\mathbf{0}\})$. Through an integration by parts, it can be computed accurately that $\|\alpha^{1/2}\nabla u\| \approx 0.56501154$. For detailed formula and more possible choices of the parameters above, we refer the reader to [17].

The AFEM procedure for this problem stops when the relative error reaches 0.05, and the resulting mesh and finite element approximation during the refinement can be found in Figure 5, and the AFEM procedure shows optimal rate of convergence in Figure 6. The effectivity index for η_{Residual} is 2.95, and 1.33 for $\widehat{\eta}$.

(a)   (b)
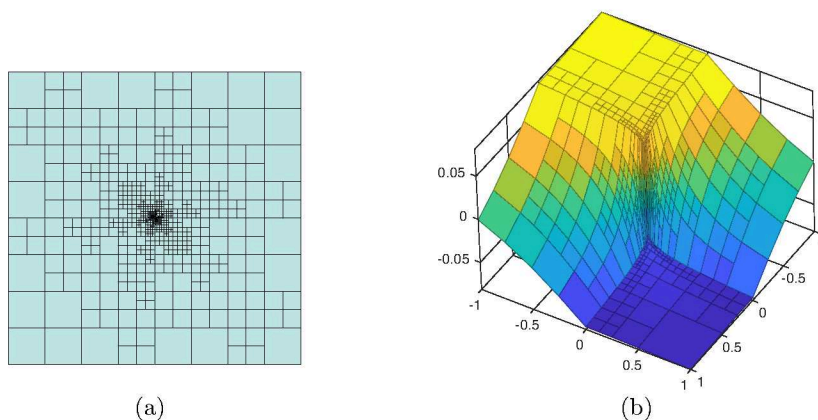
FIGURE 5. The result of the Kellogg example. (a) The adaptively refined mesh with #DoFs = 2001 on which the energy error is 0.0753, this number is roughly 75% of the number of DoFs needed to achieve the same accuracy if using conforming linear finite element on triangular grid (see [17, Section 4]). (b) The finite element approximation with #DoFs = 1736.
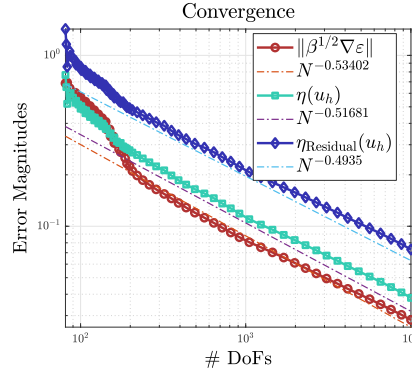
FIGURE 6. The convergence result of the Kellogg example.

6. **Conclusion.** A postprocessed flux with the minimum $\boldsymbol{H}(\mathrm{div})$ continuity requirement is constructed for tensor-product type finite element. The implementation can be easily ported to finite element on quadtree to make use the vast existing finite element libraries in the engineering community. Theoretically, the local error indicator is efficient, and the global estimator is shown to be reliable under the assumptions that (i) the mesh has bounded irregularities, and (ii) the diffusion coefficient is a quasi-monotone piecewise constant. Numerically, we have observed that both the local error indicator and the global estimator are efficient and reliable (in the asymptotic regime), respectively. Moreover, the recovery-based estimator is more accurate than the residual-based one.

However, we do acknowledge that the technical tool involving interpolation is essentially limited to 1-irregular meshes in reliability. A simple weighted averaging has restrictions and is hard to generalize to hp-finite elements, or discretization on curved edges/isoparametric elements. Nevertheless, we have shown that the flexibility of the virtual element framework allows further modification of the space in which we perform the flux recovery to cater the needs.

7. **Appendix.**

7.1. **Inverse estimates and the norm equivalence of a virtual element function.** Unlike the identity matrix stabilization commonly used in most of the VEM literature, for $\boldsymbol{\tau} \in \mathcal{V}_k(K)$, we opt for a mass matrix/DoF hybrid stabilizer approach. Let $\left\| \alpha^{-1/2} \boldsymbol{\tau} \right\|_{h,K}^2 := (\!(\boldsymbol{\tau}, \boldsymbol{\tau})\!)_K$ and

$$(\!(\boldsymbol{\sigma}, \boldsymbol{\tau})\!)_K := \left(\Pi\boldsymbol{\sigma}, \Pi\boldsymbol{\tau} \right)_K + S_K\big((\mathrm{I} - \Pi)\boldsymbol{\sigma}, (\mathrm{I} - \Pi)\boldsymbol{\tau} \big), \tag{41}$$

where $S_K(\cdot, \cdot)$ is defined in (23).

To show the inverse inequality and the norm equivalence used in the reliability bound, on each element, we need to introduce some geometric measures. Consider a polygonal element K and an edge $e \subset \partial K$, let the height l_e which measures how

far from this edge e one can advance to an interior subset of K, and denote $T_e \subset K$ as a right triangle with height l_e and base as edge e.

Proposition 1. *Under Assumption 1, $\mathcal{T}_{poly}$ satisfies (1) The number of edges in every $K \in \mathcal{T}_{poly}$ is uniformly bounded above. (2) For any edge e on every K, l_e/h_e is uniformly bounded below.*

Lemma 7.1 (Trace inequality on small edges [13])**.** *If Proposition 1 holds, for $v \in H^1(K)$ and $K \in \mathcal{T}_{poly}$ we have*

$$h_e^{-1/2} \|v\|_e \lesssim h_K^{-1} \|v\|_K + \|\nabla v\|_K, \quad on \ e \subset K. \tag{42}$$

Proof. The proof follows essentially equation (3.9) in [13, Lemma 3.3] as a standard scaled trace inequality on e toward T_e reads

$$h_e^{-1/2} \|v\|_e \lesssim h_e^{-1} \|v\|_{T_e} + \|\nabla v\|_{T_e} \lesssim h_K^{-1} \|v\|_K + \|\nabla v\|_K.$$

$\square$

Lemma 7.2 (Inverse inequalities)**.** *Under Assumption 1, we have the following inverse estimates for $\boldsymbol{\tau} \in \mathcal{V}_k(K)$ (4) on any $K \in \mathcal{T}_{poly}$ with constants depending on k and the number of edges in K:*

$$\|\nabla \cdot \boldsymbol{\tau}\|_K \lesssim h_K^{-1} \|\boldsymbol{\tau}\|_K, \quad and \quad \|\nabla \cdot \boldsymbol{\tau}\|_K \lesssim h_K^{-1} S_K(\boldsymbol{\tau}, \boldsymbol{\tau})^{1/2}. \tag{43}$$

Proof. The first inequality in (43) can be shown using a bubble function trick. Choose b_K be a bubble function of $T_{e'}$ where e' is the longest edge on ∂K. Denote $p := \nabla \cdot \boldsymbol{\tau} \in \mathbb{P}_{k-1}(K)$, we have

$$\|\nabla \cdot \boldsymbol{\tau}\|_K^2 \lesssim (\nabla \cdot \boldsymbol{\tau}, pb_K) = -(\boldsymbol{\tau}, \nabla(pb_K)) \leq \|\boldsymbol{\tau}\|_K \|\nabla(pb_K)\|_K,$$

and then $\|\nabla(pb_K)\|$ can be estimated as follows

$$\|\nabla(pb_K)\| \leq \|b_K \nabla p\|_K + \|p \nabla b_K\|_K \leq \|b_K\|_{\infty,\Omega} \|\nabla p\|_K + \|p\|_K \|\nabla b_K\|_{\infty,K}.$$

Consequently, the first inequality in (43) follows above by the standard inverse estimate for polynomials $\|\nabla p\|_K \lesssim h_K^{-1} \|p\|_K$, and the properties of the bubble function $\|b_K\|_{\infty,K} = O(1)$, and $\|\nabla b_K\|_{\infty,K} = O(h_K^{-1})$.

To prove the second inequality in (43), by integration by parts we have

$$\|\nabla \cdot \boldsymbol{\tau}\|^2 = (\nabla \cdot \boldsymbol{\tau}, p) = -(\boldsymbol{\tau}, \nabla p) + \sum_{e \subset \partial K} (\boldsymbol{\tau} \cdot \boldsymbol{n}_e, p). \tag{44}$$

Expand $\nabla \cdot \boldsymbol{\tau} = p$ in the monomial basis $p(\boldsymbol{x}) = \sum_{\alpha \in \Lambda} p_\alpha m_\alpha(\boldsymbol{x})$, and denote the mass matrix $\mathbf{M} := \big((m_\alpha, m_\gamma)_K\big)_{\alpha\gamma}$, $\mathbf{p} := (p_\alpha)_{\alpha \in \Lambda}$, it is straightforward to see that

$$\|p\|_K^2 = \mathbf{p}^\top \mathbf{M} \mathbf{p} \geq \mathbf{p}^\top \operatorname{diag}(\mathbf{M}) \mathbf{p} \geq \min_j \mathbf{M}_{jj} \|\mathbf{p}\|_{\ell^2}^2 \simeq h_K^2 \|\mathbf{p}\|_{\ell^2}^2, \tag{45}$$

since $\int_K (x - x_K)^l (y - y_K)^m \, \mathrm{d}x\mathrm{d}y \geq 0$ for the off-diagonal entries of $\mathbf{M}$ due to K being geometrically a rectangle (with additional vertices). As a result, applying the trace inequality in Lemma 7.1 on (44) yields

$$\|\nabla \cdot \boldsymbol{\tau}\|^2 \leq \left(\sum_{\alpha \in \Lambda} (\boldsymbol{\tau}, m_\alpha)_K^2\right)^{1/2} \left(\sum_{\alpha \in \Lambda} p_\alpha^2\right)^{1/2}$$

$$+ \left(\sum_{e \subset \partial K} h_e \|\boldsymbol{\tau} \cdot \boldsymbol{n}_e\|_e^2\right)^{1/2} \left(\sum_{e \subset \partial K} h_e^{-1} \|p\|_e^2\right)^{1/2}$$

$$\lesssim S_K(\boldsymbol{\tau}, \boldsymbol{\tau})^{1/2} \left(\|\mathbf{p}\|_{\ell^2} + h_K^{-1} \|p\|_K + \|\nabla p\|_K\right).$$

As a result, the second inequality in (43) is proved when apply an inverse inequality for $\|\nabla p\|_K$ and estimate (45).   □

Remark 2. While the proof in Lemma 7.2 relies on K being a rectangle, the result holds for a much broader class of polygons by changing the basis of $\mathbb{P}_{k-1}(K)$ from the simple scaled monomials to quasi-orthogonal ones in [25, 7] and apply the isotropic polygon scaling result in [13].

Lemma 7.3 (Norm equivalence). *Under Assumption 1, let Π be the oblique projection defined in (16), then the following relations holds for $\boldsymbol{\tau} \in \mathcal{V}_k(K)$ (4) on any $K \in \mathcal{T}_{poly}$:*

$$\gamma_* \|\boldsymbol{\tau}\|_K \leq \|\boldsymbol{\tau}\|_{h,K} \leq \gamma^* \|\boldsymbol{\tau}\|_K, \tag{46}$$

where both γ_* and γ^* depends on k and the number of edges in K.

Proof. First we consider the lower bound, by triangle inequality,

$$\|\boldsymbol{\tau}\|_K \leq \|\Pi\boldsymbol{\tau}\|_K + \|(\boldsymbol{\tau} - \Pi\boldsymbol{\tau})\|_K.$$

Since $\Pi\boldsymbol{\tau} \in \mathcal{V}^k(K)$, it suffices to establish the following to prove the lower bound in (46)

$$\|\boldsymbol{\tau}\|_K^2 \leq S_K(\boldsymbol{\tau}, \boldsymbol{\tau}), \quad \text{for } \boldsymbol{\tau} \in \mathcal{V}_k(K). \tag{47}$$

To this end, we consider the weak solution to the following auxiliary boundary value problem on K:

$$\begin{cases} \Delta\psi = \nabla \cdot \boldsymbol{\tau} & \text{in } K, \\ \dfrac{\partial\psi}{\partial n} = \boldsymbol{\tau} \cdot \boldsymbol{n}_{\partial K} & \text{on } \partial K. \end{cases} \tag{48}$$

By a standard Helmholtz decomposition result (e.g. Proposition 3.1, Chapter 1[23]), we have $\boldsymbol{\tau} - \nabla\psi = \nabla^\perp\phi$. Moreover, since on ∂K, $0 = \nabla^\perp\phi \cdot \boldsymbol{n} = \nabla\phi \cdot \boldsymbol{t} = \partial\phi/\partial s$, we can further choose $\phi \in H_0^1(K)$. As a result, by the assumption that $\nabla \times \boldsymbol{\tau} = 0$ for $\boldsymbol{\tau}$ in the modified virtual element space (4), we can verify that

$$\|\boldsymbol{\tau} - \nabla\psi\|_K^2 = (\boldsymbol{\tau} - \nabla\psi, \nabla^\perp\phi) = 0.$$

Consequently, we proved essentially the unisolvency of the modified VEM space (4) and $\boldsymbol{\tau} = \nabla\psi$. We further note that ψ in (48) can be chosen in $H^1(K)/\mathbb{R}$ and thus

$$\begin{aligned} \|\boldsymbol{\tau}\|_K^2 &= (\boldsymbol{\tau}, \nabla\psi)_K = (\boldsymbol{\tau}, \nabla\psi)_K \\ &= -(\nabla \cdot \boldsymbol{\tau}, \psi)_K + (\boldsymbol{\tau} \cdot \boldsymbol{n}_{\partial K}, \psi)_{\partial K} \\ &\leq \|\nabla \cdot \boldsymbol{\tau}\|_K \|\psi\|_K + \sum_{e \subset \partial K} \|\boldsymbol{\tau} \cdot \boldsymbol{n}_e\|_e \|\psi\|_e \\ &\leq \|\nabla \cdot \boldsymbol{\tau}\|_K \|\psi\|_K + \left(\sum_{e \subset \partial K} h_e \|\boldsymbol{\tau} \cdot \boldsymbol{n}_e\|_e^2\right)^{1/2} \left(\sum_{e \subset \partial K} h_e^{-1} \|\psi\|_e^2\right)^{1/2} \end{aligned} \tag{49}$$

Proposition 1 allows us to apply an isotropic trace inequality on an edge of a polygon (Lemma 7.1), combining with the Poincaré inequality for $H^1(K)/\mathbb{R}$, we have, on every $e \subset \partial K$,

$$h_e^{-1/2} \|\psi\|_e \lesssim h_K^{-1} \|\psi\|_K + \|\nabla\psi\|_K \lesssim \|\nabla\psi\|_K.$$

Furthermore applying the inverse estimate in Lemma 7.2 on the bulk term above, we have

$$\|\boldsymbol{\tau}\|_K^2 \lesssim S_K(\boldsymbol{\tau}, \boldsymbol{\tau})^{1/2} \|\nabla\psi\|_K,$$

which proves the validity of (47), thus yield the lower bound.

To prove the upper bound, by $\left\|\Pi\boldsymbol{\tau}\right\|_K \leq \|\boldsymbol{\tau}\|_K$, it suffices to establish the reversed direction of (47) on a single edge e and for a single monomial basis $m_\alpha \in \mathbb{P}_{k-1}(K)$:

$$h_e\|\boldsymbol{\tau}\cdot\boldsymbol{n}_e\|_e^2 \lesssim \|\boldsymbol{\tau}\|_K, \quad \text{and} \quad |(\boldsymbol{\tau}, \nabla m_\alpha)_K| \leq \|\boldsymbol{\tau}\|_K. \tag{50}$$

To prove the first inequality, by Proposition 1 again, consider the edge bubble function b_e such that $\operatorname{supp} b_e = T_e$. We can let $b_e = 0$ on $e' \subset \partial K$ for $e' \neq e$. It is easy to verify that:

$$\|\nabla b_e\|_{\infty,K} = O(1/h_e), \text{ and } \|b_e\|_{\infty,K} = O(1). \tag{51}$$

Denote $q_e := \boldsymbol{\tau}\cdot\boldsymbol{n}_e$, and extend it to $\mathring{K}$ by a constant extension in the normal direction rectangular strip $R_e \subset K$ with respect to e (notice $\operatorname{supp} b_e \subset R_e$), we have

$$
\begin{aligned}
\|\boldsymbol{\tau}\cdot\boldsymbol{n}_e\|_e^2 &\lesssim \big(\boldsymbol{\tau}\cdot\boldsymbol{n}_e, b_e q_e\big)_e = x\big(\boldsymbol{\tau}\cdot\boldsymbol{n}_e, b_e q_e\big)_{\partial K} \\
&= \big(\boldsymbol{\tau}, q_e\nabla b_e\big)_K + \big(\nabla\cdot\boldsymbol{\tau}, b_e q_e\big)_K \\
&\leq \|\boldsymbol{\tau}\|_K \|q_e\nabla b_e\|_{T_e} + \|\nabla\cdot\boldsymbol{\tau}\|_K \|q_e b_e\|_{T_e}, \\
&\leq \|\boldsymbol{\tau}\|_K \|q_e\|_{T_e} \|\nabla b_e\|_{\infty,K} + \|\nabla\cdot\boldsymbol{\tau}\|_K \|q_e\|_{T_e} \|b_e\|_{\infty,K}.
\end{aligned}
$$

Now by the fact that $\|q_e\|_{T_e} \lesssim h_e^{1/2} \|q_e\|_e$, the scaling of the edge bubble function in (51), and the first inverse estimate of $\|\nabla\cdot\boldsymbol{\tau}\|_K \lesssim h_K^{-1} \|\boldsymbol{\tau}\|_K$ in Lemma 7.2 yields the first part of (50).

The second inequality in (50) can be estimated straightforward by the scaling of the monomials (7)

$$|(\boldsymbol{\tau}, \nabla m_\alpha)_K| \leq \|\boldsymbol{\tau}\|_K \|\nabla m_\alpha\|_K \leq \|\boldsymbol{\tau}\|_K. \tag{52}$$

Hence, (46) is proved. $\qquad\qquad\square$

REFERENCES

[1] R. Anderson, J. Andrej, A. Barker, J. Bramwell, J.-S. Camier, J. C. V. Dobrev, Y. Dudouit, A. Fisher, T. Kolev, W. Pazner, M. Stowell, V. Tomov, I. Akkerman, J. Dahm, D. Medina and S. Zampini, MFEM: A modular finite element library, *Computers & Mathematics with Applications*, **81** (2021), 42–74.

[2] W. Bangerth, R. Hartmann and G. Kanschat, deal.II – a general purpose object oriented finite element library, *ACM Trans. Math. Software*, **33** (2007), Art. 24, 27 pp.

[3] R. E. Bank and J. Xu, Asymptotically exact a posteriori error estimators, part ii: General unstructured grids, *SIAM J. Numer. Anal.*, **41** (2003), 2313–2332.

[4] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini and A. Russo, Basic principles of virtual element methods, *Mathematical Models and Methods in Applied Sciences*, **23** (2013), 199–214.

[5] L. Beirão da Veiga, F. Brezzi, L. D. Marini and A. Russo, Serendipity face and edge VEM spaces, *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, **28** (2017), 143–180.

[6] C. Bernardi and R. Verfürth, Adaptive finite element methods for elliptic equations with non-smooth coefficient, *Numer. Math.*, **85** (2000), 579–608.

[7] S. Berrone and A. Borio, Orthogonal polynomials in badly shaped polygonal elements for the virtual element method, *Finite Elem. Anal. Des.*, **129** (2017), 14–31.

[8] F. Brezzi, R. S. Falk and L. D. Marini, Basic principles of mixed virtual element methods, *ESAIM Math. Model. Numer. Anal.*, **48** (2014), 1227–1240.

[9] R. Bruce Kellogg, On the Poisson equation with intersecting interfaces, *Applicable Anal.*, **4** (1974), 101–129.

[10] Z. Cai and S. Cao, A recovery-based a posteriori error estimator for **H(curl)** interface problems, *Comput. Methods in Appl. Mech. Eng.*, **296** (2015), 169–195.

[11] Z. Cai and S. Zhang, Recovery-based error estimators for interface problems: conforming linear elements, *SIAM J. Numer. Anal.*, **47** (2009), 2132–2156.

[12] A. Cangiani, E. H. Georgoulis, T. Pryer and O. J. Sutton, A posteriori error estimates for the virtual element method, *Numer. Math.*, **137** (2017), 857–893.

[13] S. Cao and L. Chen, Anisotropic error estimates of the linear nonconforming virtual element methods, *SIAM J. Numer. Anal.*, **57** (2019), 1058–1081.

[14] C. Carstensen and J. Hu, Hanging nodes in the unifying theory of a posteriori finite element error control, *J. Comput. Math.*, **27** (2009), 215–236.

[15] J. Červený, V. Dobrev and T. Kolev, Nonconforming mesh refinement for high-order finite elements, *SIAM J. Sci. Comput.*, **41** (2019), C367–C392.

[16] L. Chen, *iFEM: An Innovative Finite Element Methods Package in MATLAB*, Technical report, 2008, URL https://github.com/lyc102/ifem.

[17] 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), 443–462.

[18] H. Chi, L. Beirão da Veiga and G. H. Paulino, A simple and effective gradient recovery scheme and a posteriori error estimator for the virtual element method (VEM), *Comput. Methods Appl. Mech. Engrg.*, **347** (2019), 21–58.

[19] F. Dassi, J. Gedicke and L. Mascotto, Adaptive virtual element methods with equilibrated fluxes, arXiv preprint, arXiv:2004.11220.

[20] L. Demkowicz, J. T. Oden, W. Rachowicz and O. Hardy, Toward a universal hp adaptive finite element strategy, part 1. constrained approximation and data structure, *Comput. Methods Appl. Mech. Engrg.*, **77** (1989), 79–112.

[21] P. Di Stolfo, A. Schröder, N. Zander and S. Kollmannsberger, An easy treatment of hanging nodes in *hp*-finite elements, *Finite Elem. Anal. Des.*, **121** (2016), 101–117.

[22] A. Ern and M. Vohralík, Flux reconstruction and a posteriori error estimation for discontinuous Galerkin methods on general nonmatching grids, *C. R. Math. Acad. Sci. Paris*, **347** (2009), 441–444.

[23] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, 1986.

[24] H. Guo, C. Xie and R. Zhao, Superconvergent gradient recovery for virtual element methods, *Math. Models Methods Appl. Sci.*, **29** (2019), 2007–2031.

[25] L. Mascotto, Ill-conditioning in the virtual element method: Stabilizations and bases, *Numer. Methods Partial Differential Equations*, **34** (2018), 1258–1281.

[26] P. Šolín, J. Červenỳ and I. Doležel, Arbitrary-level hanging nodes and automatic adaptivity in the *hp*-FEM, *Math. Comput. Simulation*, **77** (2008), 117–132.

[27] R. Verfürth, Error estimates for some quasi-interpolation operators, *M2AN Math. Model. Numer. Anal.*, **33** (1999), 695–713.

[28] O. C. Zienkiewicz and J. Z. Zhu, The superconvergent patch recovery and a posteriori error estimates. part 1: The recovery technique, *Internat. J. Numer. Methods Engrg.*, **33** (1992), 1331–1364.

E-mail address: s.cao@wustl.edu