J Sci Comput (2013) 55:125–148
DOI 10.1007/s10915-012-9625-4

Convergence and Optimality of the Adaptive Nonconforming Linear Element Method for the Stokes Problem

Jun Hu · Jinchao Xu

Received: 22 October 2011 / Revised: 19 March 2012 / Accepted: 29 June 2012 /
Published online: 18 July 2012

Abstract In this paper, we analyze the convergence and optimality of a standard adaptive nonconforming linear element method for the Stokes problem. After establishing a special quasi-orthogonality property for both the velocity and the pressure in this saddle point problem, we introduce a new prolongation operator to carry through the discrete reliability analysis for the error estimator. We then use a specially defined interpolation operator to prove that, up to oscillation, the error can be bounded by the approximation error within a properly defined nonlinear approximate class. Finally, by introducing a new parameter-dependent error estimator, we prove the convergence and optimality estimates.

Keywords Adaptive finite element method · Convergence · Optimality · The Stokes problem

1 Introduction

The adaptive finite element method plays an important role in the numerical solution for partial differential equations [1, 2, 42]. The convergence and optimality of the adaptive method have been much studied in recent years. For the Poisson equation and its variants, the theory is well-developed [9, 15, 19, 20, 26, 35–38, 40, 41]. However, for many other important problems this is not the case. Among these under studied problems is the Stokes problem, the main subject of this paper.

J. Hu (✉)
LMAM, Peking University, Beijing 100871, P.R. China
e-mail: hujun@math.pku.edu.cn

J. Hu · J. Xu
School of Mathematical Sciences, Peking University, Beijing 100871, P.R. China

J. Xu
Beijing International Center for Mathematical Research, Peking University, Beijing, P.R. China

J. Xu
Department of Mathematics, Pennsylvania State University, University Park, PA 16801, USA
e-mail: xu@math.psu.edu

The convergence analysis of the adaptive finite element method of the Poisson equation is based on the orthogonality property [19, 26, 35, 36], such orthogonality can be weakened to some quasi-orthogonality for the nonconforming and mixed methods [4, 6, 14, 15, 17, 20, 21, 29, 31, 34, 39]. The Stokes problem, as a saddle point problem with two variables (velocity and pressure), lacks the usual orthogonality or quasi-orthogonality that holds for the positive and definite problem. As a result, it is not obvious how the technique for nonconforming and mixed methods for the Poisson equation can be carried over to the Stokes problem. Although the mixed formulation of the Poisson equation is also a saddle point problem, analyses of this formulation's convergence and optimality [4, 17, 20] are not so different from that for the primary formulation of the Poisson equation. The reason is that only the stress variable, which can be decoupled from the primary variable, needs to be involved in the analysis. This is not, however, the case for the Stokes problem under consideration here because the two variables, velocity and pressure, are coupled and cannot be separated in analyses of the convergence and optimality. To circumvent this difficulty, Bänsch, Morin, and Nochetto developed a modified adaptive procedure in which the Uzawa algorithm on the continuous level is used as the outer iteration [3, 32, 33]. See also [24] for adaptive wavelet methods.

The optimality of the adaptive finite element method for the Poisson equation is analyzed based on discrete reliability (see [19, 40, 41] and the references therein). Basically, we need one restriction operator and one prolongation operator in order to analyze the discrete reliability. For the conforming method, a natural candidate for the prolongation operator is the usual inclusion operator, and for the restriction operator a Scott–Zhang-type can be used as it has both the local projection property and the global and uniform boundedness property. For the nonconforming method under consideration here, however, it is a challenge to come up with a prolongation operator that has both the local projection property and the global and uniform boundedness property. For the nonconforming linear element method for the Poisson equation, such a difficulty can be circumvented using the discrete Helmholtz decomposition [6, 39]. However, the Helmholtz decomposition seems not applicable for the problem under consideration because the existence of such a decomposition is unclear for the general case.

The first convergence and optimality analysis of a standard adaptive finite element method for the Stokes problem was presented in a technical report [30] in 2007 by the authors of this paper. The analysis was based on some special relation between the nonconforming P_1 element and the lowest Raviart–Thomas element for the Stokes problem and one prolongation operator between the discrete spaces. But we later found a gap in our discrete reliability analysis caused by the prolongation operator used therein. A convergence and optimality analysis was published in [5] in 2011; however, we also found a gap in their analysis similar to that in our earlier report [30] (see Appendix for more details).

The present paper is an improved version of [30] with simplified and corrected proofs. Its purpose is to provide a rigorous analysis of the convergence and optimality of the adaptive nonconforming linear element method for the Stokes problem. The main idea is to establish the orthogonality or quasi-orthogonality of both the velocity variable and the pressure variable. The nonconformity of the discrete velocity space is the main difficulty in establishing the desired quasi-orthogonality property and the discrete reliability estimate. To overcome this difficulty we take two steps, (1) we establish the quasi-orthogonality for both the velocity and pressure variables by using a special conservative property of the nonconforming linear element, and (2) we introduce a new prolongation operator that has both the projection property and the uniform boundedness property for the discrete reliability analysis. To analyze optimality within the standard nonlinear approximate class [19], we define a new interpolation operator to bound the consistency error and prove that the consistency error can be bounded by the approximation error up to oscillation. This in fact implies that the nonlinear approximate class used in [30] is *equivalent* to the standard nonlinear approximate class [7, 19]. Finally, by introducing a new parameter-dependent error estimator, we prove convergence and optimality estimates for the Stokes problem.

The rest of the paper is organized as follows. In Sect. 2 we present the Stokes problem and its nonconforming linear finite element method, and recall a posteriori error estimate according to [12, 13, 16, 25]. We prove the quasi-orthogonality in Sect. 3 and then show the reduction of some total error in Sect. 4 in terms of a new parameter-dependent estimator. We introduce a new prolongation operator to establish discrete reliability in Sect. 5. And, we show optimality of the adaptive nonconforming linear element method in Sect. 6.

2 The Adaptive Nonconforming Linear Element

Let us first introduce some notations. We use the standard gradient and divergence operators $\nabla r := (\partial r/\partial x, \partial r/\partial y)$ for a scalar function r, and $\operatorname{div} \boldsymbol{\psi} := \partial \psi_1/\partial x + \partial \psi_2/\partial y$ for a vector function $\boldsymbol{\psi} = (\psi_1, \psi_2)$. Given a polygonal domain $\Omega \subset \mathbb{R}^2$ with the boundary $\partial\Omega$, we use the standard notation for Sobolev spaces, such as $H^1(\Omega)$ and $L^2(\Omega)$. We define

$$H_0^1(\Omega) := \left\{ v \in H^1(\Omega), v = 0 \text{ on } \partial\Omega \right\}, \quad \text{and}$$

$$L_0^2(\Omega) := \left\{ q \in L^2(\Omega), \int_\Omega q\, dx = 0 \right\}.$$

In addition, we denote $(\cdot, \cdot)_{L^2(\Omega)}$ as the usual L^2 inner product of functions in the space $L^2(\Omega)$, and $\|\cdot\|_{L^2(\Omega)}$ the L^2 norm.

Suppose that $\overline{\Omega}$ is covered exactly by a sequence of shape-regular triangulations $\mathcal{T}_k$ ($k \geq 0$) consisting of triangles in $2D$ (see [11, 22]), and that this sequence is produced by some adaptive algorithm where $\mathcal{T}_k$ is some nested refinement of $\mathcal{T}_{k-1}$ by the newest vertex bisection [40, 41]. Let $\mathcal{E}_k$ be the set of all edges in $\mathcal{T}_k$; $\mathcal{E}_k(\Omega)$ the set of interior edges; $\mathcal{E}(K)$ the set of edges of any given element K in $\mathcal{T}_k$; and $h_K = |K|^{1/2}$ the size of the element $K \in \mathcal{T}_k$ where $|K|$ is the area of element K. ω_K is the union of elements $K' \in \mathcal{T}_k$ that share an edge with K, and ω_E is the union of elements that share a common edge E. Given any edge $E \in \mathcal{E}_k(\Omega)$ with the length h_E, we assign one fixed unit normal $\nu_E := (\nu_1, \nu_2)$ and tangential vector $\tau_E := (-\nu_2, \nu_1)$. For E on the boundary, we choose $\nu_E := \nu$, the unit outward normal to Ω. Once ν_E and τ_E are fixed on E, in relation to ν_E we define the elements $K_- \in \mathcal{T}_k$ and $K_+ \in \mathcal{T}_k$, with $E = K_+ \cap K_-$. Given $E \in \mathcal{E}_k(\Omega)$ and some $\mathbb{R}^d$-valued function v defined in Ω, with $d = 1, 2$, we denote $[v] := (v|_{K_+})|_E - (v|_{K_-})|_E$ as the jump of v across E, where $v|_K$ is the restriction of v on K and $v|_E$ is the restriction of v on E.

2.1 The Stokes Problem and Its Nonconforming Linear Element

The Stokes problem is defined as follows: Given $g \in L^2(\Omega)^2$, find $(u, p) \in V \times Q := (H_0^1(\Omega))^2 \times L_0^2(\Omega)$ such that

$$a(u, v) + b(v, p) + b(u, q) = (g, v)_{L^2(\Omega)} \quad \text{for any } (v, q) \in V \times Q, \tag{2.1}$$

where u and p are the velocity and pressure of the flow, respectively, and

$$a(u, v) := \mu(\nabla u, \nabla v)_{L^2(\Omega)} \quad \text{and} \quad b(v, q) := (\operatorname{div} v, q)_{L^2(\Omega)}, \tag{2.2}$$

where $\mu > 0$ is the viscosity coefficient of the flow.

Given $\omega \subset \mathbb{R}^2$ and some integer ℓ, denote $P_\ell(\omega)$ as the space of polynomials of degree $\leq \ell$ over ω. We define

$$V_k := \Big\{ v_k \in L^2(\Omega)^2, v_k|_K \in P_1(K)^2 \text{ for any } K \in \mathcal{T}_k, \int_E [v_k]\, ds = 0 \text{ for any } E \in \mathcal{E}_k(\Omega), \text{ and } \int_E v_k\, ds = 0 \text{ for any } E \in \mathcal{E}_k \cap \partial\Omega \Big\},$$

$$Q_k := \big\{ q_k \in Q, q_k|_K \in P_0(K) \text{ for any } K \in \mathcal{T}_k \big\}.$$

Since V_k is not a subspace of $H^1(\Omega)^2$, the gradient and divergence operators are defined element by element with respect to $\mathcal{T}_k$, and denoted by ∇_k and div_k. Define the piecewise smooth space

$$H^1(\mathcal{T}_k) := \big\{ v \in L^2(\Omega), v|_K \in H^1(K) \text{ for any } K \in \mathcal{T}_k \big\}. \tag{2.3}$$

The discrete bilinear forms read

$$a_k(u, v) := \mu(\nabla_k u, \nabla_k v)_{L^2(\Omega)} \quad \text{and} \quad b_k(v, q) := (\operatorname{div}_k v, q)_{L^2(\Omega)} \tag{2.4}$$

for any $u, v \in (H^1(\mathcal{T}_k))^2$, and $q \in Q$.

The nonconforming P_1 element, proposed in [23], for the Stokes problem is as follows: Given $g \in L^2(\Omega)^2$, find $(u_k, p_k) \in V_k \times Q_k$ such that

$$a_k(u_k, v) + b_k(v, p_k) + b_k(u_k, q) = (g, v)_{L^2(\Omega)} \quad \text{for any } (v, q) \in V_k \times Q_k. \tag{2.5}$$

Let $\mathrm{id} \in \mathbb{R}^{2\times 2}$ be the identity matrix. Define

$$\sigma_k := \mu \nabla_k u_k + p_k \,\mathrm{id}.$$

Then, we have

$$(\sigma_k, \nabla_k v_k)_{L^2(\Omega)} = (g, v_k)_{L^2(\Omega)} \quad \text{for any } v_k \in V_k. \tag{2.6}$$

2.2 The a Posteriori Error Estimate

To recall the a posteriori error estimator of the nonconforming P_1 element, we define the residual $\mathrm{R}_{k-1}(\cdot)$ by

$$\mathrm{R}_{k-1}(v) := (g, v)_{L^2(\Omega)} - a_k(u_{k-1}, v) - b_k(v, p_{k-1}) \quad \text{for any } v \in H^1(\mathcal{T}_k)^2, \tag{2.7}$$

with the solution (u_{k-1}, p_{k-1}) of (2.5) on the mesh $\mathcal{T}_{k-1}$, which is a coarser and nested mesh of $\mathcal{T}_k$. It follows from the definition of (u_{k-1}, p_{k-1}) that

$$\mathrm{R}_{k-1}(v_{k-1}) = 0 \quad \text{for any } v_{k-1} \in V_{k-1}.$$

Given $K \in \mathcal{T}_k$, we define the element estimator

$$\eta_K(u_k, p_k) := h_K \|g\|_{L^2(K)} + \Big(\sum_{E \subset \partial K} h_K \big\| [\nabla_k u_k \tau_E] \big\|^2_{L^2(E)} \Big)^{1/2}. \tag{2.8}$$

Given $S_k \subset \mathcal{T}_k$, we define the estimator over it by

$$\eta^2(u_k, p_k, S_k) := \sum_{K \in S_k} \eta_K^2(u_k, p_k). \tag{2.9}$$

Given any $K \in \mathcal{T}_k$, denote g_K as the L^2 projection of g onto $P_0(K)$. We define the oscillation

$$\operatorname{osc}^2(g, \mathcal{T}_k) := \sum_{K \in \mathcal{T}_k} h_K^2 \|g - g_K\|_{L^2(K)}^2. \tag{2.10}$$

The reliability and efficiency of the estimator $\eta(u_k, p_k, \mathcal{T}_k)$ can be found in [12, 13, 16, 25], as stated in the following lemma.

Lemma 2.1 *Let* (u, p) *and* (u_k, p_k) *be the solutions of the Stokes problem* (2.1) *and the discrete problem* (2.5), *respectively. Then,*

$$\left\|\nabla_k(u - u_k)\right\|_{L^2(\Omega)}^2 + \|p - p_k\|_{L^2(\Omega)}^2 \lesssim \eta^2(u_k, p_k, \mathcal{T}_k), \tag{2.11}$$

$$\eta^2(u_k, p_k, \mathcal{T}_k) \lesssim \left\|\nabla_k(u - u_k)\right\|_{L^2(\Omega)}^2 + \|p - p_k\|_{L^2(\Omega)}^2 + \operatorname{osc}^2(g, \mathcal{T}_k). \tag{2.12}$$

Remark 2.2 For the Stokes problem, the estimator usually involves the pressure approximation. For the nonconforming P_1 element, as shown in the above lemma, we can decouple the pressure from the velocity [25].

Here and throughout the paper, we use the notations $\lesssim$ and $\approxeq$. When we write

$$A_1 \lesssim B_1, \quad \text{and} \quad A_2 \approxeq B_2,$$

possible constants C_1, c_2 and C_2 exist such that

$$A_1 \le C_1 B_1, \quad \text{and} \quad c_2 B_2 \le A_2 \le C_2 B_2.$$

2.3 The Adaptive Nonconforming Finite Element Method

The adaptive algorithm is defined as follows: Let $\mathcal{T}_0$ be an initial shape-regular triangulation, a right-side $g \in L^2(\Omega)^2$, a tolerance ϵ, and a parameter $0 < \theta < 1$.

Algorithm 2.1 $[\mathcal{T}_N, u_N, p_N] =$ **ANFEM**$(\mathcal{T}_0, g, \epsilon, \theta)$
$\eta = \epsilon, k = 0$

WHILE $\eta \ge \epsilon$, **DO**

(1) Solve (2.5) on $\mathcal{T}_k$ to get the solution (u_k, p_k).
(2) Compute the error estimator $\eta = \eta(u_k, p_k, \mathcal{T}_k)$.
(3) Mark the minimal element set $\mathcal{M}_k$ such that

$$\eta^2(u_k, p_k, \mathcal{M}_k) \ge \theta\, \eta^2(u_k, p_k, \mathcal{T}_k). \tag{2.13}$$

(4) Refine each triangle $K \in \mathcal{M}_k$ by the newest vertex bisection to get $\mathcal{T}_{k+1}$ and set $k =: k + 1$.

END WHILE

$\mathcal{T}_N = \mathcal{T}_k$.
END ANFEM

3 Quasi-orthogonality

The quasi-orthogonality property is the main ingredient for the convergence analysis of the adaptive nonconforming method under consideration. In this section we establish such a property by exploring the conservative property of the nonconforming linear element and by confirming that the stress is piecewise constant. To this end, we define a canonical interpolation operator Π_k for the nonconforming space V_k and a restriction operator I_{k-1} from V_k to the coarser space V_{k-1}. Given $v \in V$, we define the interpolation $\Pi_k v \in V_k$ by

$$\int_E \Pi_k v ds := \int_E v ds \quad \text{for any } E \in \mathcal{E}_k. \tag{3.1}$$

In this paper, the above property is referred to as the conservative property. This property is crucial for the analysis herein. A similar conservative property was first explored in [29] to analyze the quasi-orthogonality property of the Morley element.

The interpolation admits the following estimate:

$$\|v - \Pi_k v\|_{L^2(K)} \lesssim h_K \|\nabla v\|_{L^2(K)} \quad \text{for any } K \in \mathcal{T}_k \text{ and } v \in V. \tag{3.2}$$

Given $v_k \in V_k$, we define the restriction interpolation $I_{k-1} v_k \in V_{k-1}$ by

$$\int_E I_{k-1} v_k ds := \sum_{l=1}^{\ell} \int_{E_l} v_k ds, \quad E \in \mathcal{E}_{k-1} \text{ with } E = E_1 \cup E_2 \cup \cdots \cup E_\ell \text{ and } E_i \in \mathcal{E}_k. \tag{3.3}$$

The properties of the restriction operator I_{k-1} are summarized in the following lemma.

Lemma 3.1 *Let the restriction operator I_{k-1} be defined in (3.3). Then,*

$$I_{k-1} v_k = v_k \quad \textit{for any } K \in \mathcal{T}_k \cap \mathcal{T}_{k-1}, v_k \in V_k, \tag{3.4}$$

$$\|I_{k-1} v_k - v_k\|_{L^2(K)} \lesssim h_K \|\nabla_k v_k\|_{L^2(K)} \quad \textit{for any } K \in \mathcal{T}_{k-1} \backslash \mathcal{T}_k, v_k \in V_k. \tag{3.5}$$

Proof The property (3.4) directly follows from the definition of the restriction interpolation. Only the estimate (3.5) needs to be proved. In fact, both sides of (3.5) are semi-norms of the restriction $(V_k)_K$ of V_k on K. If the right-hand side vanishes for some $v \in (V_k)_K$, then v_k is a piecewise constant vector over K with respect to $\mathcal{T}_k$. Given the average continuity of v_k across the internal edges of $\mathcal{T}_k$, it follows that v_k is a constant vector on K. Therefore, the left-hand side also vanishes for the same v_k. The desired result then follows a scaling argument. □

Remark 3.2 An alternative proof for the inequality (3.5) follows the discrete Poincare inequality established in [10] for the scalar function, which is further investigated in [39]. Notice that the positive constant of (3.5) is independent of the ratio

$$\gamma := \max_{K \in \mathcal{T}_{k-1} \backslash \mathcal{T}_k} \max_{\mathcal{T}_k \ni T \subset K} \frac{h_K}{h_T}, \tag{3.6}$$

see [39, Lemma 4.1] for more details.

Lemma 3.3 *Let* (u_{k-1}, p_{k-1}) *be the solution of the discrete problem* (2.5) *on the mesh* $\mathcal{T}_{k-1}$. *It, therefore, holds that*

$$\left|\mathrm{R}_{k-1}(v_k)\right| \lesssim \left(\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|^2_{L^2(K)}\right)^{1/2}\|\nabla_k v_k\|_{L^2(\Omega)} \quad \textit{for any } v_k\in V_k. \tag{3.7}$$

Proof For the reader's convenience, we recall the definition of the residual as follows:

$$\mathrm{R}_{k-1}(v_k) = (g, v_k)_{L^2(\Omega)} - (\sigma_{k-1}, \nabla_k v_k)_{L^2(\Omega)}. \tag{3.8}$$

To analyze the right-hand side of the above equation, we set $v_{k-1} = I_{k-1}v_k$. As σ_{k-1} is a piecewise constant tensor with respect to the mesh $\mathcal{T}_{k-1}$, the definition of the interpolation operator I_{k-1} in (3.3) leads to

$$\int_E (v_k - v_{k-1})\cdot\sigma_{k-1}\nu_E ds = 0 \quad \text{for any } E\in\mathcal{E}_{k-1}. \tag{3.9}$$

For any $E\in\mathcal{E}_k$ that lies in the interior of some $K\in\mathcal{T}_{k-1}$, the integral average of v_k over E is continuous and σ_{k-1} is a constant on K. Then,

$$\int_E [v_k - v_{k-1}]\cdot\sigma_{k-1}\nu_E ds = 0. \tag{3.10}$$

By integrating parts on the fine mesh $\mathcal{T}_k$ and using (3.9) and (3.10), we get

$$\left(\nabla_k(v_k - v_{k-1}), \sigma_{k-1}\right)_{L^2(\Omega)} = 0. \tag{3.11}$$

Inserting this identity into (3.8) and adopting the discrete problem (2.5), we employ properties (3.4) and (3.5) of the interpolation operator I_{k-1} to derive

$$\begin{aligned}\left|\mathrm{R}_{k-1}(v_k)\right| &= \left|(g, v_k - v_{k-1})_{L^2(\Omega)}\right| \le \sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k}\|g\|_{L^2(K)}\|v_k - v_{k-1}\|_{L^2(K)}\\ &\lesssim \sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K\|g\|_{L^2(K)}\|\nabla_k v_k\|_{L^2(K)},\end{aligned} \tag{3.12}$$

which completes the proof. □

Lemma 3.4 (Quasi-orthogonality of the velocity) *Let* (u_k, p_k) *and* (u_{k-1}, p_{k-1}) *be the discrete solutions of* (2.5) *on* $\mathcal{T}_k$ *and* $\mathcal{T}_{k-1}$, *respectively. Then,*

$$\left|a_k(u - u_k, u_k - u_{k-1})\right| \lesssim \left\|\nabla_k(u - u_k)\right\|_{L^2(\Omega)}\left(\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|^2_{L^2(K)}\right)^{1/2}.$$

Proof The Stokes problem (2.1) and the discrete problem (2.5) give

$$a_k(u - u_k, u_k - u_{k-1}) = \left(\nabla_k(u - u_k), \sigma_k - \sigma_{k-1}\right)_{L^2(\Omega)}. \tag{3.13}$$

Given that $(\mathrm{div}_k(u - u_k), p_k - p_{k-1})_{L^2(\Omega)} = 0$, let $v_k = \Pi_k(u - u_k)$. And, $\sigma_k - \sigma_{k-1}$ is a piecewise constant tensor with respect to the fine mesh $\mathcal{T}_k$; therefore, by the definition of the interpolation operator Π_k in (3.1), we integrate by parts on $\mathcal{T}_k$ to obtain

$$\big(\nabla_k\big((u-u_k)-v_k\big),\sigma_k-\sigma_{k-1}\big)_{L^2(\Omega)}=0. \tag{3.14}$$

From the discrete problem (2.5), we have

$$a_k(u-u_k,u_k-u_{k-1})=(g,v_k)_{L^2(\Omega)}-(\nabla_k v_k,\sigma_{k-1})_{L^2(\Omega)}=\mathrm{R}_{k-1}(v_k). \tag{3.15}$$

The term on the right-hand side of (3.15) can be estimated by the inequality (3.7) as follows:

$$\begin{aligned}\big|\mathrm{R}_{k-1}(v_k)\big| &\lesssim \sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K\|g\|_{L^2(K)}\|\nabla_k v_k\|_{L^2(K)}\\ &\lesssim \sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K\|g\|_{L^2(K)}\big\|\nabla_k(u-u_k)\big\|_{L^2(K)},\end{aligned}$$

which completes the proof. □

Lemma 3.5 (Quasi-orthogonality of the pressure) *Let* (u_k,p_k) *and* (u_{k-1},p_{k-1}) *be the discrete solutions of* (2.5) *on* $\mathcal{T}_k$ *and* $\mathcal{T}_{k-1}$, *respectively. Then,*

$$\begin{aligned}&\big|(p-p_k,p_k-p_{k-1})_{L^2(\Omega)}\big|\\ &\quad\lesssim\bigg(\bigg(\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|_{L^2(K)}^2\bigg)^{1/2}+\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\Omega)}\bigg)\|p-p_k\|_{L^2(\Omega)}.\end{aligned} \tag{3.16}$$

Remark 3.6 The quasi-orthogonality of the pressure herein is different from those for the nonstandard method of the Poisson equation [14, 15, 20] by the fact that both $\|\nabla_k(u_k-u_{k-1})\|_{L^2(\Omega)}$ and $\|p-p_k\|_{L^2(\Omega)}$ appear on the right-hand side of (3.16).

Proof Let $\Pi_{0,k}$ be the L^2 projection operator from $L_0^2(\Omega)$ onto Q_k. It follows from the discrete inf-sup condition that there exists $v_k\in V_k$ with

$$\operatorname{div}_k v_k=\Pi_{0,k}p-p_k,\quad\text{and}\quad\|\nabla_k v_k\|_{L^2(\Omega)}\lesssim\|\Pi_{0,k}p-p_k\|_{L^2(\Omega)}. \tag{3.17}$$

Since $p_k-p_{k-1}\in Q_k$, it follows from the continuous problem (2.1), the discrete problem (2.5), and the definition of the residual (2.7) that

$$(p-p_k,p_k-p_{k-1})_{L^2(\Omega)}=(\operatorname{div}_k v_k,p_k-p_{k-1})_{L^2(\Omega)}=\mathrm{R}_{k-1}(v_k)+a_k(u_{k-1}-u_k,v_k).$$

We use the estimates in (3.7) and (3.17) to get

$$\begin{aligned}&\big|(p-p_k,p_k-p_{k-1})_{L^2(\Omega)}\big|\\ &\quad\lesssim\bigg(\bigg(\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|_{L^2(K)}^2\bigg)^{1/2}+\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\Omega)}\bigg)\|p-p_k\|_{L^2(\Omega)},\end{aligned}$$

which completes the proof. □

4 The Convergence of the ANFEM

To prove the convergence of the adaptive algorithm, we first prove the reduction of the error between the two nested meshes, $\mathcal{T}_k$ and $\mathcal{T}_{k-1}$, where $\mathcal{T}_k$ is the refinement of the coarser mesh $\mathcal{T}_{k-1}$ with (2.13) by the newest vertex bisection. In order to control the volume part $\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|^2_{L^2(K)}$ appearing in Lemmas 3.4 and 3.5, we introduce the following modified estimator:

$$\tilde{\eta}^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}) := \sum_{K\in\mathcal{T}_{k-1}} \big(\beta_1 h_K^2\|g\|^2_{L^2(K)} + \eta_K^2(u_{k-1}, p_{k-1})\big) \tag{4.1}$$

with the positive constant $\beta_1 > 0$ to be determined later. Note that this modified estimator is introduced only for the convergence analysis and that the final convergence and optimal complexity will be proved for Algorithm 2.1.

Note that the volume residual $\sum_{K\in\mathcal{T}_{k-1}} h_K^2\|g\|^2_{L^2(K)}$ does not contain the unknowns. Hence, we add it to settle down the lacking of the Galerkin-orthogonality or quasi-orthogonality. We stress that the Galerkin-orthogonality or quasi-orthogonality is an essential ingredient for the convergence analysis of the adaptive conforming, nonconforming, and mixed methods for the Poisson-like problems [14, 15, 19, 20, 26, 35, 36]. This is another reason that we need a modified estimator as in (4.1).

We list three standard components for the convergence analysis of the adaptive method, which can be proved by following the arguments, for instance, in [15, 19, 26].

Lemma 4.1 *Let $\mathcal{T}_k$ be some refinement of $\mathcal{T}_{k-1}$ from Algorithm 2.1, then $\rho > 0$ and a positive constant $\beta \in (1-\rho\theta, 1)$ exist, such that*

$$\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_k) \le \beta\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}) + (1-\rho\theta-\beta)\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}). \tag{4.2}$$

Proof The result can be proved by following the idea in [15, 19, 26]. The details are only given for the readers' convenience. In fact, we have

$$\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_k) = \eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}\cap\mathcal{T}_k) + \eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_k\backslash\mathcal{T}_{k-1}). \tag{4.3}$$

For any $K \in \mathcal{T}_{k-1}\backslash\mathcal{T}_k$, we only need to consider the case where K is subdivided into $K_1, K_2 \in \mathcal{T}_k$ with $|K_1| = |K_2| = \frac{1}{2}|K|$. As $[\nabla_{k-1}u_{k-1}\tau_E] = 0$ over the interior edge $E = K_1\cap K_2 \in \mathcal{E}_k$, we have

$$\begin{aligned}
&\sum_{i=1}^{2}\eta_{K_i}^2(u_{k-1}, p_{k-1}) \\
&:= \sum_{i=1}^{2}\Bigg(h_{K_i}\|g\|_{L^2(K_i)} + \Bigg(\sum_{\mathcal{E}_k\ni E\subset\partial K_i} h_{K_i}\big\|[\nabla_{k-1}u_{k-1}\tau_E]\big\|^2_{L^2(E)}\Bigg)^{1/2}\Bigg)^2 \\
&\le \frac{1}{2^{1/2}}\eta_K^2(u_{k-1}, p_{k-1}) \\
&:= \frac{1}{2^{1/2}}\Bigg(h_K\|g\|_{L^2(K)} + \Bigg(\sum_{\mathcal{E}_{k-1}\ni E\subset\partial K} h_K\big\|[\nabla_{k-1}u_{k-1}\tau_E]\big\|^2_{L^2(E)}\Bigg)^{1/2}\Bigg)^2.
\end{aligned} \tag{4.4}$$

Consequently,

$$\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k}\sum_{i=1}^{2}\eta_{K_i}^2(u_{k-1},p_{k-1})\le\frac{1}{2^{1/2}}\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1}\backslash\mathcal{T}_k). \tag{4.5}$$

Let $\rho=1-\frac{1}{2^{1/2}}$, therefore, we obtain

$$\eta^2(u_{k-1},p_{k-1},\mathcal{T}_k)\le\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1})-\rho\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1}\backslash\mathcal{T}_k). \tag{4.6}$$

Choosing the positive parameter β with $1-\rho\theta<\beta<1$, we combine the above inequality and the bulk criterion (2.13) to achieve the desired result. □

Lemma 4.2 *Let $\mathcal{T}_k$ be some refinement of $\mathcal{T}_{k-1}$ produced in Algorithm 2.1, then there exists $\rho>0$ such that*

$$\sum_{K\in\mathcal{T}_k}h_K^2\|g\|_{L^2(K)}^2\le\sum_{K\in\mathcal{T}_{k-1}}h_K^2\|g\|_{L^2(K)}^2-\rho\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k}h_K^2\|g\|_{L^2(K)}^2. \tag{4.7}$$

Proof This can be proved by a similar argument proposed in the previous lemma. □

Lemma 4.3 (Continuity of the estimator) *Let u_k and u_{k-1} be the solutions to the discrete problem (2.5) on the meshes $\mathcal{T}_k$ and $\mathcal{T}_{k-1}$ obtained from Algorithm 2.1. Given any positive constant ϵ, there exists a positive constant $\beta_2(\epsilon)$ dependent on ϵ such that*

$$\eta^2(u_k,p_k,\mathcal{T}_k)\le(1+\epsilon)\eta^2(u_{k-1},p_{k-1},\mathcal{T}_k)+\frac{1}{\beta_2(\epsilon)}\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\Omega)}^2. \tag{4.8}$$

Proof Given any $K\in\mathcal{T}_k$, it follows from the definitions of $\eta_K(u_k,p_k)$ and $\eta_K(u_{k-1},p_{k-1})$ in (4.4) that

$$\begin{aligned}&\big|\eta_K(u_k,p_k)-\eta_K(u_{k-1},p_{k-1})\big|\\&=\left|\Bigg(\sum_{\mathcal{E}_k\ni E\subset\partial K}h_K\big\|[\nabla_k u_k\tau_E]\big\|_{L^2(E)}^2\Bigg)^{1/2}-\Bigg(\sum_{\mathcal{E}_k\ni E\subset\partial K}h_K\big\|[\nabla_{k-1}u_{k-1}\tau_E]\big\|_{L^2(E)}^2\Bigg)^{1/2}\right|\\&\le\Bigg(\sum_{\mathcal{E}_k\ni E\subset\partial K}h_K\big\|\big[\nabla_k(u_k-u_{k-1})\tau_E\big]\big\|_{L^2(E)}^2\Bigg)^{1/2}.\end{aligned}$$

Given $E\in\mathcal{E}_k$, let $K_1,K_2\in\mathcal{T}_k$ be the two elements that take E as one edge. Then, we use the trace theorem and the fact that $\nabla_k(u_k-u_{k-1})$ is a piecewise constant tensor to get

$$\begin{aligned}&\big\|\big[\nabla_k(u_k-u_{k-1})\tau_E\big]\big\|_{L^2(E)}\\&\quad\le\big\|\nabla_k(u_k-u_{k-1})\tau_E|_{K_1}\big\|_{L^2(E)}+\big\|\nabla_k(u_k-u_{k-1})\tau_E|_{K_2}\big\|_{L^2(E)}\\&\quad\lesssim h_K^{-1/2}\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\omega_E)},\end{aligned} \tag{4.9}$$

which gives

$$\eta_K(u_k,p_k)\le\eta_K(u_{k-1},p_{k-1})+C_{Con}\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\omega_K)}, \tag{4.10}$$

for some positive constant C_{Con}. Given any positive constant ϵ, we apply the Young inequality to get

$$\eta_K^2(u_k, p_k) \le (1+\epsilon)\eta_K^2(u_{k-1}, p_{k-1}) + \frac{C_{Con}^2(1+\epsilon)}{\epsilon}\big\|\nabla_k(u_k - u_{k-1})\big\|^2_{L^2(\omega_K)}. \tag{4.11}$$

A summation over all elements in $\mathcal{T}_k$ completes the proof with $\beta_2(\epsilon) = \frac{M\epsilon}{C_{Con}^2(1+\epsilon)}$, where the positive constant M depends on the finite overlapping of the patches ω_K. □

In the following theorem, we prove the convergence of the adaptive nonconforming finite element method for the Stokes problem. The main ingredients are the quasi-orthogonality of both the velocity and the pressure in Lemmas 3.4 and 3.5, and the relations of the estimators between two the meshes $\mathcal{T}_k$ and $\mathcal{T}_{k-1}$ presented in Lemmas 4.1–4.3.

Theorem 4.4 *Let* (u, p) *and* (u_k, p_k) *be the solutions of* (2.1) *and* (2.5). *Then* $\gamma_1, \gamma_2, \beta_1 > 0$ *and* $0 < \alpha < 1$ *exist, such that*

$$\begin{aligned}
&\big\|\nabla_k(u-u_k)\big\|^2_{L^2(\Omega)} + \gamma_1\|p-p_k\|^2_{L^2(\Omega)} + \gamma_2\tilde{\eta}^2(u_k, p_k, \mathcal{T}_k) \\
&\quad\le \alpha\Big(\big\|\nabla_{k-1}(u-u_{k-1})\big\|^2_{L^2(\Omega)} + \gamma_1\big\|p-p_{k-1}\big\|^2_{L^2(\Omega)} \\
&\qquad + \gamma_2\tilde{\eta}^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1})\Big).
\end{aligned} \tag{4.12}$$

Proof First, we adopt the quasi-orthogonality of both the velocity and the pressure. Denote the multiplication constant in Lemma 3.4 by C_{QOV}. As

$$\begin{aligned}
\big\|\nabla_k(u-u_k)\big\|^2_{L^2(\Omega)} &= \big\|\nabla_k(u-u_{k-1})\big\|^2_{L^2(\Omega)} - \big\|\nabla_k(u_k-u_{k-1})\big\|^2_{L^2(\Omega)} \\
&\quad - 2\big(\nabla_k(u-u_k), \nabla_k(u_k-u_{k-1})\big)_{L^2(\Omega)},
\end{aligned} \tag{4.13}$$

it follows from the quasi-orthogonality of the velocity in Lemma 3.4 and the Young inequality that

$$\begin{aligned}
&(1-\delta_1)\big\|\nabla_k(u-u_k)\big\|^2_{L^2(\Omega)} \\
&\quad\le \big\|\nabla_{k-1}(u-u_{k-1})\big\|^2_{L^2(\Omega)} - \big\|\nabla_k(u_k-u_{k-1})\big\|^2_{L^2(\Omega)} \\
&\qquad + C_1(\delta_1)\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|^2_{L^2(K)},
\end{aligned} \tag{4.14}$$

where $C_1(\delta_1) = \frac{C_{QOV}^2}{\delta_1}$ for any positive constant $0 < \delta_1 < 1$. Denote the multiplication constant in Lemma 3.5 by C_{QOP}. From the quasi-orthogonality of the pressure proved in Lemma 3.5 and the Young inequality, we have

$$\begin{aligned}
(1-\delta_2-\delta_3)\|p-p_k\|^2_{L^2(\Omega)} &\le \|p-p_{k-1}\|^2_{L^2(\Omega)} - \|p_k-p_{k-1}\|^2_{L^2(\Omega)} \\
&\quad + \frac{1}{\beta_3(\delta_3)}\big\|\nabla_k(u_k-u_{k-1})\big\|^2_{L^2(\Omega)} \\
&\quad + C_2(\delta_2)\sum_{K\in\mathcal{T}_{k-1}\backslash\mathcal{T}_k} h_K^2\|g\|^2_{L^2(K)};
\end{aligned} \tag{4.15}$$

here $\beta_3(\delta_3) = \frac{\delta_3}{C_{QOP}^2}$ and $C_2(\delta_2) = \frac{C_{QOP}^2}{\delta_2}$ for any constants $0 < \delta_2, \delta_3 < 1$. Then we multiply the inequality (4.14) by $\gamma_1 > 0$ and the inequality (4.15) by $\gamma_2 > 0$ to obtain

$$\begin{aligned}
&\gamma_1(1-\delta_1)\big\|\nabla_{k-1}(u-u_k)\big\|_{L^2(\Omega)}^2 + \gamma_2(1-\delta_2-\delta_2)\|p-p_k\|_{L^2(\Omega)}^2 \\
&\quad \le \gamma_1\big\|\nabla_{k-1}(u-u_{k-1})\big\|_{L^2(\Omega)}^2 + \gamma_2\|p-p_{k-1}\|_{L^2(\Omega)}^2 \\
&\qquad - \left(\gamma_1 - \frac{\gamma_2}{\beta_3(\delta_3)}\right)\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\Omega)}^2 \\
&\qquad - \gamma_2\|p_k-p_{k-1}\|_{L^2(\Omega)}^2 + \big(\gamma_1 C_1(\delta_1) + \gamma_2 C_2(\delta_2)\big) \\
&\qquad \times \sum_{K\in\mathcal{T}_{k-1}\setminus\mathcal{T}_k} h_K^2\|g\|_{L^2(K)}^2.
\end{aligned} \tag{4.16}$$

For the presentation, we introduce some short-hand notations for any positive constants $\gamma_3, \gamma_4 > 0$:

$$\begin{aligned}
\mathfrak{G}_k(u_k,p_k) &:= \gamma_1(1-\delta_1)\big\|\nabla_{k-1}(u-u_k)\big\|_{L^2(\Omega)}^2 + \gamma_2(1-\delta_2-\delta_3)\|p-p_k\|_{L^2(\Omega)}^2 \\
&\quad + \gamma_3\eta^2(u_k,p_k,\mathcal{T}_k) + \gamma_4\sum_{K\in\mathcal{T}_k} h_K^2\|g\|_{L^2(K)}^2, \\
\overline{\mathfrak{G}}_{k-1}(u_{k-1},p_{k-1}) &:= \gamma_1\big\|\nabla_{k-1}(u-u_{k-1})\big\|_{L^2(\Omega)}^2 + \gamma_2\|p-p_{k-1}\|_{L^2(\Omega)}^2 \\
&\quad + \gamma_3\beta\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1}) + \gamma_4\sum_{K\in\mathcal{T}_{k-1}} h_K^2\|g\|_{L^2(K)}^2.
\end{aligned} \tag{4.17}$$

Second, we use the continuity of the estimators from Lemmas 4.1–4.3 to cancel both the term $\|\nabla_k(u_k-u_{k-1})\|_{L^2(\Omega)}$ and the volume estimator. In fact, from (4.2) and (4.8), we have

$$\begin{aligned}
\eta^2(u_k,p_k,\mathcal{T}_k) &\le \beta\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1}) + \frac{1}{\beta_2(\epsilon)}\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\Omega)}^2 \\
&\quad + \big((1-\rho\theta-\beta)(1+\epsilon)+\epsilon\beta\big)\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1}).
\end{aligned} \tag{4.18}$$

Then we combine the above inequality with the inequalities (4.16) and (4.7) to obtain

$$\begin{aligned}
\mathfrak{G}_k(u_k,p_k) &\le \overline{\mathfrak{G}}_{k-1}(u_{k-1},p_{k-1}) - \left(\gamma_1 - \frac{\gamma_2}{\beta_3(\delta_3)} - \frac{\gamma_3}{\beta_2(\epsilon)}\right)\big\|\nabla_k(u_k-u_{k-1})\big\|_{L^2(\Omega)}^2 \\
&\quad - \gamma_2\|p_k-p_{k-1}\|_{L^2(\Omega)}^2 + \gamma_3\big((1-\rho\theta-\beta)(1+\epsilon)+\epsilon\beta\big)\eta^2(u_{k-1},p_{k-1},\mathcal{T}_{k-1}) \\
&\quad + \big(\gamma_1 C_1(\delta_1) + \gamma_2 C_2(\delta_2) - \gamma_4\rho\big)\sum_{K\in\mathcal{T}_{k-1}\setminus\mathcal{T}_k} h_K^2\|g\|_{L^2(K)}^2.
\end{aligned}$$

It remains to prove that the positive constants $\delta_i, i = 1,2,3,$ $\gamma_i, i = 1,2,3,4,$ ϵ, β, and β_1 exist such that the contraction (4.12) holds for some constant $0 < \alpha < 1$. Further it is possible that the constant dependent on the choices of the aforementioned parameters but independent of the meshsize h and the level k. This will be achieved in the following three steps.

Step 1 For the second, fourth, and fifth terms on the right-hand side of the above inequality to vanish, we set

$$\begin{aligned}\gamma_2 &= \left(\gamma_1 - \frac{\gamma_3}{\beta_2(\epsilon)}\right)\beta_3(\delta_3) \quad \text{with } \gamma_1 > \frac{\gamma_3}{\beta_2(\epsilon)},\\ \gamma_4 &= \big(\gamma_1 C_1(\delta_1) + \gamma_2 C_2(\delta_2)\big)/\rho,\\ \beta &= (1-\rho\theta)(1+\epsilon).\end{aligned} \tag{4.19}$$

Note that γ_2, γ_4, and β will be determined after $\delta_i, i = 1, 2, 3$, γ_1, γ_3, and ϵ have been specified. In the following, we assume that ϵ is fixed in such a way that $0 < \beta < 1$. Also, we let γ_1 and γ_3 be fixed such that $\gamma_1 > \frac{\gamma_3}{\beta_2(\epsilon)}$ and $\gamma_2 > 0$. Hence, we have

$$\mathfrak{G}_k(u_k, p_k) \le \overline{\mathfrak{G}}_{k-1}(u_{k-1}, p_{k-1}).$$

Let the positive constant α with $\beta < \alpha < 1$ be determined later. We define

$$\begin{aligned}&\mathfrak{R}_{k-1}(u_{k-1}, p_{k-1})\\ &:= \big(1-\alpha(1-\delta_1)\big)\gamma_1 \big\|\nabla_{k-1}(u-u_{k-1})\big\|^2_{L^2(\Omega)} + \gamma_2\big(1-\alpha(1-\delta_2-\delta_3)\big)\|p-p_{k-1}\|^2_{L^2(\Omega)}\\ &\quad + \gamma_3(\beta-\alpha)\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}) + \gamma_4(1-\alpha)\sum_{K\in\mathcal{T}_{k-1}} h_K^2\|g\|^2_{L^2(K)}.\end{aligned}$$

Then we perform the decomposition $\overline{\mathfrak{G}}_{k-1}(u_{k-1}, p_{k-1}) = \alpha\mathfrak{G}_{k-1}(u_{k-1}, p_{k-1}) + \mathfrak{R}_{k-1}(u_{k-1}, p_{k-1})$ to get

$$\mathfrak{G}_k(u_k, p_k) \le \alpha\mathfrak{G}_{k-1}(u_{k-1}, p_{k-1}) + \mathfrak{R}_{k-1}(u_{k-1}, p_{k-1}).$$

Step 2 Now we only need to show that it is possible to choose $\alpha < 1$ such that $\mathfrak{R}_{k-1}(u_{k-1}, p_{k-1}) \le 0$. This can be achieved by selecting parameters $\delta_i, i = 1, 2, 3$. To this end, we recall the reliability of $\eta(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1})$ in Lemma 2.1 with the multiplication coefficient C_{Rel}:

$$\big\|\nabla_{k-1}(u-u_{k-1})\big\|^2_{L^2(\Omega)} + \|p-p_{k-1}\|^2_{L^2(\Omega)} \le C_{Rel}\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}). \tag{4.20}$$

Further, we take $\delta_1 = \delta_2 + \delta_3$ with $0 < \delta_1 < \min(\frac{\gamma_3(1-\beta)}{C_{Rel}(\gamma_1+\gamma_2)}, 1)$. Then, we take

$$\alpha := \frac{(\gamma_1+\gamma_2)C_{Rel} + \gamma_3\beta + \gamma_4}{(1-\delta_1)(\gamma_1+\gamma_2)C_{Rel} + \gamma_3 + \gamma_4}.$$

It is straightforward to see that $\beta < \alpha < 1$. As

$$\sum_{K\in\mathcal{T}_{k-1}} h_K^2\|g\|^2_{L^2(K)} \le \eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}), \tag{4.21}$$

we obtain

$$\begin{aligned}&\mathfrak{R}_{k-1}(u_{k-1}, p_{k-1})\\ &\le \Big(\big(1-\alpha(1-\delta_1)\big)(\gamma_1+\gamma_2)C_{Rel} + \gamma_3(\beta-\alpha) + \gamma_4(1-\alpha)\Big)\eta^2(u_{k-1}, p_{k-1}, \mathcal{T}_{k-1}) = 0.\end{aligned}$$

This proves that

$$\mathfrak{G}_k(u_k, p_k) \le \alpha \mathfrak{G}_{k-1}(u_{k-1}, p_{k-1}).$$

Step 3 Finally, we take $\beta_1 := \gamma_4/\gamma_3$ and rearrange $\gamma_2 := \gamma_2(1-\delta_2-\delta_3)/(1-\delta_1)\gamma_1$, $\gamma_3 := \gamma_3/(1-\delta_1)\gamma_1$, which completes the proof. □

5 The Discrete Reliability

In this section, we prove the discrete reliability. The analysis needs some prolongation operator from V_k to $V_{k+\ell}$ with some integer $\ell \ge 1$. Some further notations are needed. Given $E \in \mathcal{E}_{k+\ell}$, the edge patch $\omega_{E,k}$ of E with respect to the mesh $\mathcal{T}_k$ is defined as

$$\omega_{E,k} := \{K \in \mathcal{T}_k, E \subset \partial K \text{ or } E \text{ lies in the interior of } K\}. \tag{5.1}$$

Let $\xi_E = \operatorname{card}(\omega_{E,k})$. We define the prolongation interpolation $I'_{k+\ell} v_k \in V_{k+\ell}$ for any $v_k \in V_k$, as

$$\int_E I'_{k+\ell} v_k ds := \frac{1}{\xi_E} \sum_{K \in \omega_{E,k}} \int_E (v_k|_K) ds \quad \text{for any } E \in \mathcal{E}_{k+\ell}. \tag{5.2}$$

For the interpolation operator $I'_{k+\ell}$, we have

$$I'_{k+\ell} v_k = v_k \quad \text{for any } K \in \mathcal{T}_k \cap \mathcal{T}_{k+\ell} \text{ and } v_k \in V_{k+\ell}. \tag{5.3}$$

As we will see in Remark 5.3 below, we cannot directly use the prolongation operator $I'_{k+\ell}$ in the analysis of the discrete reliability. An averaging operator is needed. Denote $\mathcal{N}_k$ as the set of internal vertexes of the mesh $\mathcal{T}_k$, and denote $S_k \subset H_0^1(\Omega)$ as the conforming linear element space over $\mathcal{T}_k$. Given $Z \in \mathcal{N}_k$, the nodal patch $\omega_{Z,k}$ is defined by

$$\omega_{Z,k} := \{K \in \mathcal{T}_k, Z \in K\}. \tag{5.4}$$

Denote $\phi_Z \in S_k$ as the canonical basis function associated to Z, which satisfies $\phi(Z) = 1$ and $\phi(Z') = 0$ for vertex Z' of $\mathcal{T}_k$ other than Z. We define

$$\mathcal{E}_Z := \{E \in \mathcal{E}_k, Z \in \mathcal{N}_k \text{ is one end point of } E\}. \tag{5.5}$$

The idea of [10] leads to the definition of the following averaging operator $\Pi : V_k \to (S_k)^2$:

$$\Pi v_k := \sum_{Z \in \mathcal{N}_k} v_Z \phi_Z \quad \text{for any } v_k \in V_k, \tag{5.6}$$

where

$$v_Z = \frac{1}{\xi_Z} \sum_{K \in \omega_{Z,k}} (v_k|_K)(Z) \quad \text{with } \xi_Z = \operatorname{card}(\omega_{Z,k}). \tag{5.7}$$

Given any $K \in \mathcal{T}_k$, we have

$$\|\Pi v_k - v_k\|_{L^2(K)} + h_K \left\|\nabla(\Pi v_k - v_k)\right\|_{L^2(K)}$$

$$\lesssim h_K^{3/2}\Big(\sum_{T\in\mathcal{T}_k\&T\cap K\neq\emptyset}\sum_{E\subset\partial T}\big\|[\nabla_k v_k \tau_E]\big\|^2_{L^2(E)}\Big)^{1/2}, \tag{5.8}$$

for any $v_k \in V_k$, see [10] for the proof. Define

$$\Omega_{\mathcal{R}} := \text{interior}\Big(\bigcup\big\{K : K \in \mathcal{T}_k\backslash\mathcal{T}_{k+\ell},\ \partial K \cap \partial(\mathcal{T}_k \cap \mathcal{T}_{k+\ell}) = \emptyset\big\}\Big).$$

The main idea herein is to take the mixture of the prolongation operators $I'_{k+\ell}$ and Π. More precisely, we use Π in the region $\Omega_{\mathcal{R}}$ where the elements of $\mathcal{T}_k$ are refined and take $I'_{k+\ell}$ in the region $\mathcal{T}_{k+\ell} \cap \mathcal{T}_k$, and we define some mixture in the layers between them. This leads to the prolongation operator $J_{k+\ell} : V_k \to V_{k+\ell}$ as follows:

$$J_{k+\ell}v_k := \begin{cases} \Pi_{k+\ell}\Pi v_k & \text{on } \Omega_{\mathcal{R}},\\ I'_{k+\ell}v_k & \text{on } \mathcal{T}_k \cap \mathcal{T}_{k+\ell},\\ v_{k+\ell,tr} & \text{on } \Omega\backslash(\Omega_{\mathcal{R}} \cup (\mathcal{T}_k \cap \mathcal{T}_{k+\ell})),\end{cases}$$

where $v_{k+\ell,tr}$ is defined as

$$\int_E v_{k+\ell,tr}ds := \begin{cases} \int_E \Pi v_k ds & \text{if } E \subset \partial\Omega_{\mathcal{R}}\\ \int_E I'_{k+\ell}v_k ds & \text{otherwise}\end{cases} \quad \text{for any } E \in \mathcal{E}_{k+\ell}.$$

Lemma 5.1 *For any $v_k \in V_k$, it holds that*

$$\big\|\nabla_{k+\ell}(J_{k+\ell}v_k - v_k)\big\|^2_{L^2(\Omega)} \lesssim \sum_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}}\ \sum_{E\subset\partial K\&E\not\subset\partial(\mathcal{T}_k\cap\mathcal{T}_{k+\ell})} h_K\big\|[\nabla_k v_k \tau_E]\big\|^2_{L^2(E)}. \tag{5.9}$$

Proof As $J_{k+\ell}v_k = \Pi v_k$ on $\Omega_{\mathcal{R}}$ and $J_{k+\ell}v_k = v_k$ on $\mathcal{T}_{k-1} \cap \mathcal{T}_{k+\ell}$, from (5.3) and (5.8), we only need to estimate $\|\nabla(J_{k+\ell}v_k - v_k)\|_{L^2(K)} = \|\nabla(v_{k+\ell,tr} - v_k)\|_{L^2(K)}$ for $\mathcal{T}_{k+\ell} \ni K \subset \Omega\backslash(\Omega_{\mathcal{R}} \cup (\mathcal{T}_k \cap \mathcal{T}_{k+\ell}))$. Given $E \in \mathcal{E}_{k+\ell}$, let φ_E be the canonical basis function of the nonconforming P_1 element on $\mathcal{T}_{k+\ell}$, which satisfies $\int_E \varphi_E ds = |E|$ and $\int_{E'} \varphi_E ds = 0$ for any $E' \in \mathcal{E}_{k+\ell}$ other than E. A direct calculation yields

$$\|\varphi_E\|_{L^2(\Omega)} + h_E\|\nabla_{k+\ell}\varphi_E\|_{L^2(\Omega)} \lesssim h_E.$$

Let $v'_E := \int_E v_{k+\ell,tr}|_K ds$ and $v_E := \int_E v_k|_K ds$; thus we have

$$\big\|\nabla(v_{k+\ell,tr} - v_k)\big\|_{L^2(K)} \lesssim \sum_{E\subset\partial K}\big|v'_E - v_E\big|/h_E. \tag{5.10}$$

Next we bound the terms $|v'_E - v_E|$ for $E \in \mathcal{E}_{k+\ell}$.

Case 1 $E \subset \partial\Omega_{\mathcal{R}}$. Let $F \in \mathcal{E}_k$ be the mother of edge E in the sense of $E \subset F$. Let $T \in \mathcal{T}_k$ be the mother of K in the sense of $K \subset T$. Denote the vertexes of T as $Z_i, i = 1, 2, 3$. Without losing generality, we assume that Z_1 and Z_2 are two endpoints of F. Then, the trace of $v_k|_T$ on F can be expressed as

$$v_k|_F = (v_k|_T)(Z_1)\phi_{Z_1} + (v_k|_T)(Z_2)\phi_{Z_2}. \tag{5.11}$$

Note that

$$\Pi v_k|_F = v_{Z_1}\phi_{Z_1} + v_{Z_2}\phi_{Z_2}. \tag{5.12}$$

We recall that v_{Z_i} are defined in (5.7) and that ϕ_{Z_i} are the canonical basis functions associated with vertexes Z_i for the conforming linear element. Therefore

$$\begin{aligned} \left|v'_E - v_E\right| &= \left|\int_E (\Pi v_k|_F - v_k|_F)ds\right| \\ &= \left|\int_E \big(\big(v_{Z_1} - (v_k|_T)(Z_1)\big)\phi_{Z_1} + \big(v_{Z_2} - (v_k|_T)(Z_2)\big)\phi_{Z_2}\big)ds\right| \\ &\lesssim h_E\left(\sum_{i=1}^{2}\sum_{E'\in\mathcal{E}_{Z_i}} h'_E\left\|[\nabla_k v_k \tau_{E'}]\right\|^2_{L^2(E')}\right)^{1/2}. \end{aligned} \tag{5.13}$$

Case 2 $E \nsubseteq \partial\Omega_{\mathcal{R}}$. Again, let $F \in \mathcal{E}_k$ be the mother of E in the sense of $E \subset F$. Then, we simply have

$$\left|v'_E - v_E\right| \lesssim h_F^{3/2}\left\|[\nabla_k v_k \tau_F]\right\|_{L^2(F)}. \tag{5.14}$$

By inserting the estimates of $|v'_E - v_E|$ from (5.13) and (5.14) into (5.10), we complete the proof. □

We define the ratio γ as follows:

$$\gamma := \max_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}} \max_{\mathcal{T}_{k+\ell}\ni T\subset K} \frac{h_K}{h_T}. \tag{5.15}$$

One observation herein is that γ is bounded for the element $K \in \mathcal{T}_k$, which lies in the layer $\Omega\backslash(\Omega_{\mathcal{R}} \cup (\mathcal{T}_k \cap \mathcal{T}_{k+\ell}))$.

Lemma 5.2 *The following discrete reliability holds*:

$$\left\|\nabla_{k+\ell}(u_{k+\ell} - u_k)\right\|_{L^2(\Omega)} + \|p_{k+\ell} - p_k\|_{L^2(\Omega)} \lesssim \eta(u_k, p_k, \mathcal{T}_k\backslash\mathcal{T}_{k+\ell}). \tag{5.16}$$

Remark 5.3 If we directly take the prolongation operator $I'_{k+\ell}$ to analyze this discrete reliability, the constant for the *established* discrete reliability will depend on the ratio γ (see Appendix for an example).

Proof For any $v_{k+\ell} \in V_{k+\ell}$, we have the following decomposition:

$$\begin{aligned} &\mu\left\|\nabla_{k+\ell}(u_{k+\ell} - u_k)\right\|^2_{L^2(\Omega)} \\ &\quad = a_{k+\ell}(u_{k+\ell} - u_k, u_{k+\ell} - v_{k+\ell}) + a_{k+\ell}(u_{k+\ell} - u_k, v_{k+\ell} - u_k). \end{aligned} \tag{5.17}$$

We will first estimate the first term on the right-hand side of the above equation. It follows the discrete problem (2.5) that

$$a_{k+\ell}(u_{k+\ell} - u_k, u_{k+\ell} - v_{k+\ell}) = \mathrm{R}_k(u_{k+\ell} - v_{k+\ell}) - b_{k+\ell}(u_{k+\ell} - v_{k+\ell}, p_{k+\ell} - p_k). \tag{5.18}$$

The first term on the right-hand side of (5.18) can be bounded as in (3.7):

$$\left|\mathrm{R}_k(u_{k+\ell} - v_{k+\ell})\right| \lesssim \left(\sum_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}} h_K^2\|g\|^2_{L^2(K)}\right)^{1/2}\left\|\nabla_{k+\ell}(u_{k+\ell} - v_{k+\ell})\right\|_{L^2(\Omega)}. \tag{5.19}$$

Now we turn to the second term on the right hand side of (5.18). Thanks to the discrete inf-sup condition, we use the discrete problem (2.5) to get

$$\begin{aligned}\|p_{k+\ell}-p_k\|_{L^2(\Omega)} &\lesssim \sup_{0\neq v_{k+\ell}\in V_{k+\ell}} \frac{b_{k+\ell}(v_{k+\ell}, p_{k+\ell}-p_k)}{\|\nabla_{k+\ell} v_{k+\ell}\|_{L^2(\Omega)}} \\ &\lesssim \sup_{0\neq v_{k+\ell}\in V_{k+\ell}} \frac{\mathrm{R}_k(v_{k+\ell})}{\|\nabla_{k+\ell} v_{k+\ell}\|_{L^2(\Omega)}} + \big\|\nabla_{k+\ell}(u_{k+\ell}-u_k)\big\|_{L^2(\Omega)}. \quad (5.20)\end{aligned}$$

An application of the Cauchy–Schwarz inequality leads to

$$\big|b_{k+\ell}(u_{k+\ell}-v_{k+\ell}, p_{k+\ell}-p_k)\big| \le \|p_{k+\ell}-p_k\|_{L^2(\Omega)}\big\|\nabla_{k+\ell}(u_{k+\ell}-v_{k+\ell})\big\|_{L^2(\Omega)}. \quad (5.21)$$

After inserting (5.18), (5.19), (5.20), and (5.21) into (5.17), we use the triangle and Young inequalities to derive

$$\begin{aligned}&\big\|\nabla_{k+\ell}(u_{k+\ell}-u_k)\big\|^2_{L^2(\Omega)} + \|p_{k+\ell}-p_k\|^2_{L^2(\Omega)} \\ &\quad\lesssim \sum_{K\in\mathcal{T}_k\setminus\mathcal{T}_{k+\ell}} h_K^2\|g\|^2_{L^2(K)} + \inf_{v_{k+\ell}\in V_{k+\ell}}\big\|\nabla_{k+\ell}(u_k-v_{k+\ell})\big\|^2_{L^2(\Omega)}. \quad (5.22)\end{aligned}$$

An application of (5.9) bounds the second term on the right-hand side of (5.22). This completes the proof. □

With γ_1 from Theorem 4.4, we define the following energy norm:

$$|||v,q|||^2 := \|\nabla v\|^2_{L^2(\Omega)} + \gamma_1\|q\|^2_{L^2(\Omega)}, \quad \text{for any } (v,q)\in V\times Q. \quad (5.23)$$

We denote its piecewise version by $|||\cdot|||_{k+\ell}$.

The following lemma gives links between the error reduction to the bulk criterion.

Lemma 5.4 *Let $\mathcal{T}_{k+\ell}$ be the refinement of $\mathcal{T}_k$ with the following reduction*:

$$\begin{aligned}&|||u-u_{k+\ell}, p-p_{k+\ell}|||^2_{k+\ell} + \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_{k+\ell}) \\ &\quad\le \alpha'\big(|||u-u_k, p-p_k|||^2_k + \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_k)\big), \quad (5.24)\end{aligned}$$

with $0<\alpha'<1$ and the positive constant γ_2 from Theorem 4.4. There exists $0<\theta_*<1$ with

$$\theta_*\eta^2(u_k,p_k,\mathcal{T}_k) \le \eta^2(u_k,p_k,\mathcal{T}_k\setminus\mathcal{T}_{k+\ell}). \quad (5.25)$$

Proof It follows (5.24) and the definitions of the norms $|||\cdot|||_k$ and $|||\cdot|||_{k+\ell}$ that

$$\begin{aligned}&\big(1-\alpha'\big)\big(|||u-u_k, p-p_k|||^2_k + \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_k)\big) \\ &\quad\le |||u-u_k, p-p_k|||^2_k + \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_k) - |||u-u_{k+\ell}, p-p_{k+\ell}|||^2_{k+\ell} - \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_{k+\ell}) \\ &\quad= \big\|\nabla_{k+\ell}(u_k-u_{k+\ell})\big\|^2_{L^2(\Omega)} + \gamma_1\|p_k-p_{k+\ell}\|^2_{L^2(\Omega)} + \frac{2}{\mu}a_{k+\ell}(u-u_{k+\ell}, u_{k+\ell}-u_k) \\ &\qquad+ 2\gamma_1(p-p_{k+\ell}, p_{k+\ell}-p_k)_{L^2(\Omega)} + \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_k) - \gamma_2\,\mathrm{osc}^2(g,\mathcal{T}_{k+\ell}) \\ &\quad= I_1+I_2+I_3+I_4+I_5.\end{aligned}$$

The first two terms, I_1 and I_2, are estimated by the discrete reliability in Lemma 5.2,

$$|||u_{k+\ell} - u_k|||_{k+\ell}^2 + \gamma_1 \| p_k - p_{k+\ell}\|_{L^2(\Omega)}^2 \le C_{Drel}\eta^2(u_k, p_k, \mathcal{T}_k \backslash \mathcal{T}_{k+\ell}), \tag{5.26}$$

where the coefficient C_{Drel} is from Lemma 5.2. The third term I_3 can be estimated by the quasi-orthogonality of the velocity in Lemma 3.4. In fact, let the multiplication constant therein be the coefficient C_{QOV}, so that we have

$$\begin{aligned} &\left|\frac{2}{\mu} a_{k+\ell}(u - u_{k+\ell}, u_{k+\ell} - u_k)\right| \\ &\quad \le 2C_{QOV} \left\|\nabla_{k+\ell}(u - u_{k+\ell})\right\|_{L^2(\Omega)} \left(\sum_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}} h_K^2 \|g\|_{L^2(K)}^2\right)^{1/2} \\ &\quad \le \frac{1-\alpha'}{2}\left\|\nabla_{k+\ell}(u - u_{k+\ell})\right\|_{L^2(\Omega)}^2 + \frac{2(C_{QOV})^2}{1-\alpha'} \sum_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}} h_K^2 \|g\|_{L^2(K)}^2. \end{aligned} \tag{5.27}$$

Next, we use the quasi-orthogonality of the pressure in Lemma 3.5 to analyze the fourth term, I_4. Denote the constant of Lemma 3.5 by C_{QOP}, and we obtain

$$\begin{aligned} &\left|2\gamma_1(p - p_{k+\ell}, p_{k+\ell} - p_k)_{L^2(\Omega)}\right| \\ &\le 2\gamma_1 C_{QOP}\left(\left(\sum_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}} h_K^2 \|g\|_{L^2(K)}^2\right)^{1/2} + \left\|\nabla_{k+\ell}(u_{k+\ell} - u_k)\right\|_{L^2(\Omega)}\right)\|p - p_{k+\ell}\|_{L^2(\Omega)} \\ &\le \frac{2\gamma_1(C_{QOP})^2}{1-\alpha'}\left(\left(\sum_{K\in\mathcal{T}_k\backslash\mathcal{T}_{k+\ell}} h_K^2 \|g\|_{L^2(K)}^2\right)^{1/2} + \left\|\nabla_{k+\ell}(u_{k+\ell} - u_k)\right\|_{L^2(\Omega)}\right)^2 \\ &\quad + \frac{1-\alpha'}{2}\gamma_1\|p - p_{k+\ell}\|_{L^2(\Omega)}^2. \end{aligned}$$

Hence it follows from (5.26) that

$$\begin{aligned} &\left|2\gamma_1(p - p_{k+\ell}, p_{k+\ell} - p_k)_{L^2(\Omega)}\right| \\ &\quad \le \frac{1-\alpha'}{2}\gamma_1\|p - p_{k+\ell}\|_{L^2(\Omega)}^2 \\ &\qquad + \frac{2\gamma_1(C_{QOP})^2(1 + C_{Drel}^{1/2})^2}{1-\alpha'}\eta^2(u_k, p_k, \mathcal{T}_k\backslash\mathcal{T}_{k+\ell}). \end{aligned} \tag{5.28}$$

A direct calculation leads to

$$\gamma_2\left|\operatorname{osc}^2(f, \mathcal{T}_k) - \operatorname{osc}^2(f, \mathcal{T}_{k+\ell})\right| \le \gamma_2\eta^2(u_k, p_k, \mathcal{T}_k\backslash\mathcal{T}_{k+\ell}), \tag{5.29}$$

we combine (5.26)–(5.29), and (5.24) with the efficiency of the estimator, which proves the desired result by the parameter

$$\theta_* = \frac{(1-\alpha')^2 C_{Eff}}{2(2(C_{QOV})^2 + 2\gamma_1(C_{QOP})^2(1 + C_{Drel}^{1/2})^2 + (1-\alpha')(C_{Drel} + \gamma_2))},$$

with the efficiency constant C_{Eff} of the estimator $\eta(u_k, p_k, \mathcal{T}_k)$ from Lemma 2.1. □

6 The Optimality of the ANFEM

In this section, we address the optimality of the adaptive nonconforming linear element method under consideration. We need to control the consistency error $\kappa(\sigma, \mathcal{T})$ defined by

$$\kappa(\sigma, \mathcal{T}) = \sup_{v_{\mathcal{T}} \in V_{\mathcal{T}}} \frac{(g, v_{\mathcal{T}})_{L^2(\Omega)} - (\sigma, \nabla_{\mathcal{T}} v_{\mathcal{T}})_{L^2(\Omega)}}{\|\nabla_{\mathcal{T}} v_{\mathcal{T}}\|_{L^2(\Omega)}} \quad \text{with } \sigma = \mu \nabla u + p \, \mathrm{id}, \tag{6.1}$$

where $\mathcal{T}$ is some refinement of the initial mesh $\mathcal{T}_0$ by the newest vertex bisection. The following conforming finite element space is needed:

$$P_3(\mathcal{T}) := \left\{ v \in \left(H_0^1(\Omega)\right)^2, v|_K \in \left(P_3(K)\right)^2, \text{ for any } K \in \mathcal{T} \right\}. \tag{6.2}$$

Then, there exists an interpolation operator $\Pi_{\mathcal{T}} : V_{\mathcal{T}} \to P_3(\mathcal{T})$ with the following properties [28, Lemma A.3]:

$$\begin{aligned} &\int_E (v_{\mathcal{T}} - \Pi_{\mathcal{T}} v_{\mathcal{T}}) \cdot c_E \, ds = 0 \quad \text{for any } c_E \in \left(P_1(E)\right)^2, \\ &\int_K (v_{\mathcal{T}} - \Pi_{\mathcal{T}} v_{\mathcal{T}}) dx = 0, \end{aligned} \tag{6.3}$$

for any edge E and element K of $\mathcal{T}$. In addition, we have

$$\|v_{\mathcal{T}} - \Pi_{\mathcal{T}} v_{\mathcal{T}}\|_{L^2(K)} + h_K \|\nabla \Pi_{\mathcal{T}} v_{\mathcal{T}}\|_{L^2(K)} \lesssim h_K \|\nabla_{\mathcal{T}} v_{\mathcal{T}}\|_{L^2(\omega_K)}. \tag{6.4}$$

For any $s_{\mathcal{T}} \in V_{\mathcal{T}}$ and $q_{\mathcal{T}} \in Q_{\mathcal{T}}$, we define $\sigma_{\mathcal{T}} = \mu s_{\mathcal{T}} + q_{\mathcal{T}}$. The idea of [27, Lemma 2.1] leads to the following decomposition:

$$\begin{aligned} &(g, v_{\mathcal{T}})_{L^2(\Omega)} - (\sigma, \nabla_{\mathcal{T}} v_{\mathcal{T}})_{L^2(\Omega)} \\ &\quad = (g, v_{\mathcal{T}} - \Pi_{\mathcal{T}} v_{\mathcal{T}})_{L^2(\Omega)} - \left(\sigma - \sigma_{\mathcal{T}}, \nabla_{\mathcal{T}}(v_{\mathcal{T}} - \Pi_{\mathcal{T}} v_{\mathcal{T}})\right)_{L^2(\Omega)} \\ &\qquad + \left(\sigma_{\mathcal{T}}, \nabla_{\mathcal{T}}(v_{\mathcal{T}} - \Pi_{\mathcal{T}} v_{\mathcal{T}})\right)_{L^2(\Omega)} \end{aligned} \tag{6.5}$$

for any $v_{\mathcal{T}} \in V_{\mathcal{T}}$. By the properties (6.3) and (6.4), we obtain

$$\kappa(\sigma, \mathcal{T}) \lesssim \inf_{(v_{\mathcal{T}}, q_{\mathcal{T}}) \in V_{\mathcal{T}} \times Q_{\mathcal{T}}} |||u - v_{\mathcal{T}}, p - q_{\mathcal{T}}|||_{\mathcal{T}} + \operatorname{osc}(g, \mathcal{T}). \tag{6.6}$$

This implies that the nonlinear approximate class used in [30] is *equivalent* to the standard nonlinear approximate class [7, 19]. Hence, we can introduce the following semi-norm:

$$\mathfrak{E}^2(N; u, p, g) := \inf_{\mathcal{T} \in \mathbb{T}_N} \left(\inf_{(v_{\mathcal{T}}, q_{\mathcal{T}}) \in V_{\mathcal{T}} \times Q_{\mathcal{T}}} |||u - v_{\mathcal{T}}, p - q_{\mathcal{T}}|||_{\mathcal{T}}^2 + \gamma_2 \operatorname{osc}^2(g, \mathcal{T}) \right). \tag{6.7}$$

Then the nonlinear approximate class $\mathbb{A}_s$ can be defined by

$$\mathbb{A}_s := \left\{ (u, p, g), |u, p, g|_s := \sup_{N > N_0} N^s \mathfrak{E}(N; u, p, g) < +\infty \right\}. \tag{6.8}$$

We must stress that this is the first time the standard nonlinear approximate class [19] has been used to analyze the adaptive nonconforming finite element method. In the relevant literature, the discrete solution of the discrete problem has been used to define the nonlinear approximate class [5, 6, 34, 39]. Let $(u_{\mathcal{T}}, p_{\mathcal{T}})$ be the approximation solution of (2.5) on the mesh $\mathcal{T}$. It follows from the Strang Lemma [22]

$$|||u - u_{\mathcal{T}}, p - p_{\mathcal{T}}|||_{\mathcal{T}} \lesssim \inf_{(v_{\mathcal{T}}, q_{\mathcal{T}}) \in V_{\mathcal{T}} \times Q_{\mathcal{T}}} |||u - v_{\mathcal{T}}, p - q_{\mathcal{T}}|||_{\mathcal{T}} + \kappa(\sigma, \mathcal{T}),$$

and the following fact

$$\inf_{(v_{\mathcal{T}}, q_{\mathcal{T}}) \in V_{\mathcal{T}} \times Q_{\mathcal{T}}} |||u - v_{\mathcal{T}}, p - q_{\mathcal{T}}|||_{\mathcal{T}} + \kappa(\sigma, \mathcal{T}) \lesssim |||u - u_{\mathcal{T}}, p - p_{\mathcal{T}}|||_{\mathcal{T}},$$

that the nonlinear approximate class of [5] is equivalent to $\mathbb{A}_s$ of (6.8). A similar method herein proves that the nonlinear approximate class of [6, 34, 39] is equivalent to the standard nonlinear approximate class [19].

Remark 6.1 After we submitted the revised version to the journal, we learnt about that a different argument of [18] shows that the nonlinear approximate class of [6, 34, 39] is equivalent to the standard nonlinear approximate class [19].

Thanks to (6.6), we have

$$\begin{aligned} &|||u - u_{k-1}, p - p_{k-1}|||_{k-1}^2 \\ &\quad \lesssim \inf_{(v_{k-1}, q_{k-1}) \in V_{k-1} \times Q_{k-1}} |||u - v_{k-1}, p - q_{k-1}|||_{k-1}^2 + \operatorname{osc}^2(g, \mathcal{T}_{k-1}). \end{aligned} \tag{6.9}$$

A straightforward investigation shows that if $\mathcal{T}_k$ is any refinement of $\mathcal{T}_{k-1}$, then it holds that

$$\begin{aligned} &\inf_{(v_k, q_k) \in V_k \times Q_k} |||u - v_k, p - q_k|||_k^2 + \gamma_2 \operatorname{osc}^2(g, \mathcal{T}_k) \\ &\quad \le C_3 \Big(\inf_{(v_{k-1}, q_{k-1}) \in V_{k-1} \times Q_{k-1}} |||u - v_{k-1}, p - q_{k-1}|||_{k-1}^2 + \gamma_2 \operatorname{osc}^2(g, \mathcal{T}_{k-1}) \Big). \end{aligned} \tag{6.10}$$

With these preparations, following [29], we have the following optimality:

Theorem 6.2 *Let (u, p) be the solution of Problem* (2.1), *and let $(\mathcal{T}_k, V_k \times Q_k, (u_k, p_k))$ be the sequence of meshes, finite element spaces, and discrete solutions produced by the adaptive finite element methods. If $(u, p, g) \in \mathbb{A}_s$ with*

$$\theta \le \frac{C_{Eff}}{2(2(C_{QOV})^2 + 2\gamma_1 (C_{QOP})^2 (1 + C_{Drel}^{1/2})^2 + C_{Drel} + \gamma_2)}.$$

Then, it holds that

$$|||u - u_N, p - p_N|||_N^2 + \gamma_2 \operatorname{osc}^2(g, \mathcal{T}_N) \lesssim |u, p, g|_s^2 (\#\mathcal{T}_N - \#\mathcal{T}_0)^{-2s}. \tag{6.11}$$

Acknowledgements The first author was supported by NSFC 10971005, and in part by NSFC 11031006. The second author was supported in part, by NSFC-10528102, NSF DMS 0915153, and DMS 0749202, and by the PSU-PKU Joint Center for Computational Mathematics and Applications.

Appendix: A Counter Example

We present an example in this appendix to show that if the prolongation operator I_h' defined by (5.2) is directly used to analyze the discrete reliability of the estimator, the constant for the *established* discrete reliability could depend on some key mesh refinement ratio

$$\gamma := \max_{K\in\mathcal{T}_H\backslash\mathcal{T}_h} \max_{\mathcal{T}_h\ni T\subset K} \frac{h_K}{h_T},$$

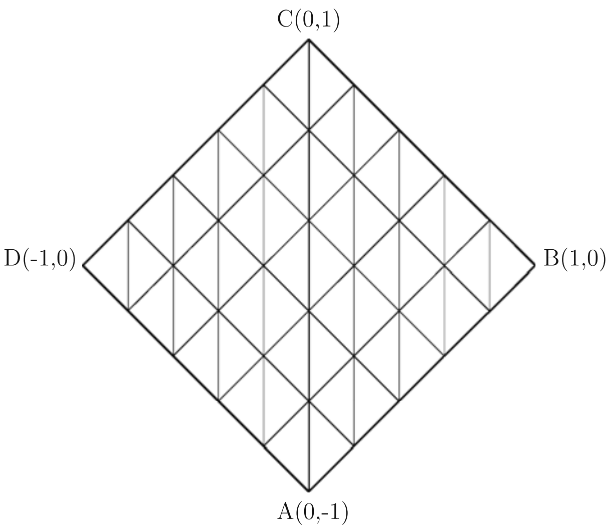
where $\mathcal{T}_H$ is some regular triangulation of Ω into triangles and $\mathcal{T}_h$ is some refinement of $\mathcal{T}_H$. To this end, we first give an example to demonstrate that there are generally no positive constants C independent of γ such that the following estimate holds true:

$$\sum_{E\in\mathcal{E}_h\backslash\mathcal{E}_H} \int_E [u_H]\left\{\frac{\partial v_h}{\partial \nu_E}\right\} ds \le C\left(\sum_{E\in\mathcal{E}_H\backslash\mathcal{E}_h} h_E^{-1}\left\|[u_H]\right\|_{L^2(E)}^2\right)^{1/2} \|\nabla_h v_h\|_{L^2(\Omega)}, \tag{A.1}$$

where $u_H \in V_H$ is the finite element solution of the velocity on the mesh $\mathcal{T}_H$ and v_h is some element of V_h over the nested fine mesh $\mathcal{T}_h$. As usual, $\mathcal{E}_H$ (resp. $\mathcal{E}_h$) is the set of the edges of $\mathcal{T}_H$ (resp. $\mathcal{T}_h$). Denote V_H (resp. V_h) as the nonconforming linear element space with respect to $\mathcal{T}_H$ (resp. $\mathcal{T}_h$). Denote $[\cdot]$ as the jump of some function across the edge E and $\{\cdot\}$ as the average of some function across the edge E. In addition, denote ν_E as the unit normal vector to E with the length h_E.

In the following, an example is given to show that $u_H \in V_H$ and $v_h \in V_h$ exist such that the above constant C depends on the ratio γ. For simplicity, let $\mathcal{T}_H$ consist of two triangles $\triangle ABC$ and $\triangle ACD$ as in Fig. 1. Let $\mathcal{T}_h$ be a uniform triangulation of Ω into $2 \times N^2$ triangles, cf. Fig. 1 for the case $N = 5$. We stress that the idea and result can be easily extended to the mesh with the newest vertex bisection. For the sake of simplicity, let $N = 2k + 1$ with some nonnegative integer k. Let Z_i, $i = -k, \dots, k$, be the nodes of $\mathcal{T}_h$ whose coordinates are $(\frac{1}{N}, \frac{2i}{N})$. Let ϕ_{Z_i} be the nodal basis function of the conforming linear element space defined over $\mathcal{T}_h$ such that $\phi_{Z_i}(Z_i) = 1$ and $\phi_{Z_i}(Z) = 0$ for any node Z other

Fig. 1 The meshes $\mathcal{T}_H$ and $\mathcal{T}_h$



than Z_i. We choose $u_H \in V_H$ such that the jump is $[u_H] = y$ over the edge AC. We choose v_h as follows:

$$v_h := \sum_{i=-k}^{k} \text{sign}(i)\phi_{Z_i} \quad \text{with sign}(i) := \begin{cases} 1 & \text{if } i > 0, \\ 0 & \text{if } i = 0, \\ -1 & \text{if } i < 0. \end{cases} \tag{A.2}$$

Note that $\{\frac{\partial \phi_{Z_i}}{\partial \nu_E}\} = N/2$ over the edge AC for $i = -k, \ldots, k$. A direct calculation gives

$$\int_{AC} [u_H] \left\{ \frac{\partial v_h}{\partial \nu_E} \right\} ds = N/2 - \frac{1}{2N}. \tag{A.3}$$

On the other hand, a direct calculation leads to

$$\|\nabla_h v_h\|^2_{L^2(\Omega)} \leq 4N. \tag{A.4}$$

This indicates that the constant C in (A.1) should be $\mathcal{O}(\sqrt{N})$, which depends on the ratio $\gamma = \mathcal{O}(N)$ for this example.

For the analysis of the discrete reliability, a direct application of the prolongation operator I'_h as defined in (5.2) will lead to a similar estimate like (A.1), and, as a result, the constant for the *established* discrete reliability based on such an estimate will depend on the ratio γ. Note that in the analysis of optimality of the adaptive method it is possible to know that $\mathcal{T}_h$ is some refinement of $\mathcal{T}_H$ only by the newest vertex bisection [8, 19, 40]. Note, too, that there is no guarantee that γ is bounded. Therefore, the proof of the discrete reliability based on the prolongation operator I'_h as presented in [5, 30] may not lead to a uniform estimate as claimed.

References

1. Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. Wiley, New York (2000)
2. Babuskă, I., Vogelius, M.: Feedback and adaptive finite element solution of one-dimensional boundary value problems. Numer. Math. **44**, 75–102 (1984)
3. Bänsch, E., Morin, P., Nochetto, R.: An adaptive Uzawa FEM for the Stokes problem: convergence without the inf-sup condition. SIAM J. Numer. Anal. **40**, 1207–1229 (2002)
4. Becker, R., Mao, S.P.: An optimally convergent adaptive mixed finite element method. Numer. Math. **111**, 35–54 (2008)
5. Becker, R., Mao, S.P.: Quasi-optimality of adaptive nonconforming finite element methods for the Stokes equations. SIAM J. Numer. Anal. **49**, 970–991 (2011)
6. Becker, R., Mao, S.P., Shi, Z.C.: A convergent nonconforming adaptive finite element method with quasi-optimal complexity. SIAM J. Numer. Anal. **47**, 4639–4659 (2010)
7. Binev, P., Dahmen, W., DeVore, R., Petrushev, P.: Approximation classes for adaptive methods. Serdica Math. J. **28**, 391–416 (2002)
8. Binev, P., Dahmen, W., DeVore, R.: Adaptive finite element methods with convergence rate. Numer. Math. **97**, 219–268 (2004)
9. Bonito, A., Nochetto, R.H.: Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. SIAM J. Numer. Anal. **48**, 734–771 (2010)
10. Brenner, S.: Poincaré-Friedrichs inequality for piecewise H^1 functions. SIAM J. Numer. Anal. **41**, 306–324 (2003)
11. Brenner, S., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Springer, New York (1994)
12. Carstensen, C.: A unifying theory of a posteriori finite element error control. Numer. Math. **100**, 617–637 (2005)

13. Carstensen, C., Funken, S.A.: A posteriori error control in low-order finite element discretisations of incompressible stationary flow problems. Math. Comput. **70**, 1353–1381 (2001)
14. Carstensen, C., Hoppe, R.H.W.: Error reduction and convergence for an adaptive mixed finite element method. Math. Comput. **75**, 1033–1042 (2006)
15. Carstensen, C., Hoppe, R.H.W.: Convergence analysis of an adaptive nonconforming finite element method. Numer. Math. **103**, 251–266 (2006)
16. Carstensen, C., Hu, J.: A unifying theory of a posteriori error control for nonconforming finite element methods. Numer. Math. **107**, 473–502 (2007)
17. Carstensen, C., Rabus, H.: An optimal adaptive mixed finite element method. Math. Comput. **80**, 649–667 (2011)
18. Carstensen, C., Peterseim, D., Schedensack, M.: Comparison results of finite element methods for the Poisson model problem (2011). Available at http://www.math.hu-berlin.de/~Peterseim/files/Comparison.pdf
19. Cascon, J.M., Kreuzer, C., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal. **46**, 2524–2550 (2008)
20. Chen, L., Holst, M., Xu, J.: Convergence and optimality of adaptive mixed finite element methods. Math. Comput. **78**, 35–53 (2009)
21. Chen, H.X., Xu, X.J., Hoppe, R.H.W.: Convergence and quasi-optimality of adaptive nonconforming finite element methods for some nonsymmetric and indefinite problems. Numer. Math. **116**, 383–419 (2010)
22. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978). Reprinted as SIAM Classics in Applied Mathematics, 2002
23. Crouzeix, M., Raviart, P.-A.: Conforming and nonconforming finite element methods for solving the stationary Stokes equations. RAIRO. Anal. Numér. **7**, 33–76 (1973)
24. Dahlke, S., Dahmen, W., Urban, K.: Adaptive wavelet methods for saddle point problems—Optimal convergence rates. SIAM J. Numer. Anal. **40**, 1230–1262 (2002)
25. Dari, E., Duran, R., Padra, C.: Error estimators for nonconforming finite element approximations of the Stokes problem. Math. Comput. **64**, 1017–1033 (1995)
26. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**, 1106–1124 (1996)
27. Gudi, T.: A new error analysis for discontinuous finite element methods for linear problems. Math. Comput. **79**, 2169–2189 (2010)
28. Girault, V., Raviart, P.A.: Finite Element Methods for Navier–Stokes Equations, Theorem and Algorithms. Springer, Berlin (1986)
29. Hu, J., Shi, Z.C., Xu, J.C.: Convergence and optimality of the adaptive Morley element method. Numer. Math. (2012). doi:10.1007/s00211-012-0445-0. See also, Hu, J. and Shi, Z.C. and Xu, J.C., Convergence and optimality of the adaptive Morley element method. Research Report **19** (2009). School of Mathematical Sciences and Institute of Mathematics, Peking University. Available online from May 2009. http://www.math.pku.edu.cn:8000/var/preprint/7280.pdf
30. Hu, J., Xu, J.C.: Convergence of adaptive conforming and nonconforming finite element methods for the perturbed Stokes equation. Research report, School of Mathematical Sciences and Institute of Mathematics, Peking University (2007). Also available online from December 2007. http://www.math.pku.edu.cn:8000/var/preprint/7297.pdf
31. Huang, J.G., Huang, X.H., Xu, Y.F.: Convergence of an adaptive mixed finite element method for Kirchhof plate bending problems. SIAM J. Numer. Anal. **49**, 574–607 (2010)
32. Kondratyuk, Y.: Adaptive finite element algorithms for the Stokes problem: Convergence rates and optimal computational complexity. Preprint 1346, Department of Mathematics, Utrecht University (2006)
33. Kondratyuk, Y., Stevenson, R.: An Optimal Adaptive Finite Element Method for the Stokes Problem. Preprint (2007)
34. Mao, S.P., Zhao, X.Y., Shi, Z.C.: Convergence of a standard adaptive nonconforming finite element method with optimal complexity. Appl. Numer. Math. **60**, 673–688 (2010)
35. Mekchay, K., Nochetto, R.H.: Convergence of adaptive finite element methods for general second order linear elliptic PDEs. SIAM J. Numer. Anal. **43**, 1043–1068 (2005)
36. Morin, P., Nochetto, R., Siebert, K.: Data oscillation and convergence of adaptive FEM. SIAM J. Numer. Anal. **38**, 466–488 (2000)
37. Morin, P., Nochetto, R.H., Siebert, K.G.: Convergence of adaptive finite element methods. SIAM Rev. **44**, 631–658 (2002)
38. Morin, P., Nochetto, R.H., Siebert, K.G.: Local problems on stars: A posteriori error estimators, convergence, and performance. Math. Comput. **72**, 1067–1097 (2003)
39. Rabus, H.: A natural adaptive nonconforming FEM is of quasi–optimal complexity. Comput. Methods Appl. Math. **10**, 316–326 (2010)

40. Stevenson, R.: Optimality of a standard adaptive finite element method. Found. Comput. Math. **7**, 245–269 (2007)
41. Stevenson, R.: The completion of locally refined simplicial partitions created by bisection. Math. Comput. **77**, 227–241 (2008)
42. Verfürth, R.: A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Technique. Wiley-Teubner, New York (1996)