

Research Article

Dongho Kim, Eun-Jae Park* and Boyoon Seo

Adaptive Multi-level Algorithm for a Class of Nonlinear Problems

<https://doi.org/10.1515/cmam-2023-0088>

Received March 31, 2023; revised December 4, 2023; accepted February 7, 2024

Abstract: In this article, we propose an adaptive mesh-refining based on the multi-level algorithm and derive a unified a posteriori error estimate for a class of nonlinear problems. We have shown that the multi-level algorithm on adaptive meshes retains quadratic convergence of Newton's method across different mesh levels, which is numerically validated. Our framework facilitates to use the general theory established for a linear problem associated with given nonlinear equations. In particular, existing a posteriori error estimates for the linear problem can be utilized to find reliable error estimators for the given nonlinear problem. As applications of our theory, we consider the pseudostress-velocity formulation of Navier–Stokes equations and the standard Galerkin formulation of semilinear elliptic equations. Reliable and efficient a posteriori error estimators for both approximations are derived. Finally, several numerical examples are presented to test the performance of the algorithm and validity of the theory developed.

Keywords: Nonlinear Problem, Navier–Stokes Equations, Pseudostress, A Posteriori Error Estimates, Multi-level Algorithm, Newton's Method

MSC 2020: 65N15, 65N30, 65N50, 76D05

1 Introduction

Adaptive mesh-refining plays an important practical role in accurate and efficient calculation of the numerical solutions of partial differential equations. A posteriori error estimates are an essential ingredient of adaptivity. They are composed of the known values such as the given data of the problem and the computed numerical solutions. Various a posteriori error estimates for finite element methods for second-order elliptic problems have already been studied in [1, 2, 8, 17, 19], for the Stokes problem in [9, 26], and for the Navier–Stokes problem in [12, 14]. Particularly, the nonlinear elliptic equations were considered in [19]. There, a priori and a posteriori error estimates measured in the $L^m(\Omega)$ -norm in the framework of Brezzi–Rappaz–Raviart (BRR) (cf. [4]) were derived. The framework was designed for approximation of branches of nonsingular solutions for a class of nonlinear problems. The theories developed by BRR were applied to study mixed approximation of scalar elliptic problems with gradient nonlinearities in [16, 19]. The theories applied to the velocity-pressure formulation of the Navier–Stokes equations can be found in [13]. The authors of [12, 14] constructed a posteriori error estimators for the two-level method for the Navier–Stokes equations.

In this paper, we derive a unified a posteriori error estimate for the approximate solutions of the multi-level algorithm developed in [22] for a class of nonlinear problems in the framework of BRR. The multi-level algorithm using a two-grid idea is an efficient numerical method designed to resolve nonlinearities. The two-grid algorithm (see [20, 27, 28]) is first applied, and in the subsequent process, uniform mesh refinement is exploited to deliver a desired accuracy and convergence; as the mesh is being refined, the solution on a given mesh is

*Corresponding author: Eun-Jae Park, Department of Computational Science and Engineering, Yonsei University, Seoul 03722, Korea, e-mail: ejpark@yonsei.ac.kr <https://orcid.org/0000-0002-3836-1595>

Dongho Kim, University College, Yonsei University, Seoul 03722, Korea, e-mail: dhkimm@yonsei.ac.kr

Boyoon Seo, Department of Mathematics, Yonsei University, Seoul 03722, Korea, e-mail: mathied@yonsei.ac.kr

exploited as an accurate starting iterate for solutions on the next mesh level. The multi-level algorithm has quadratic convergence in the sense of equation (2.5) in Theorem 2.5. We apply this strategy to efficiently compute numerical solutions by adaptive mesh refinement. We emphasize that the multi-level algorithm on adaptive meshes retains quadratic convergence of Newton's method across different mesh levels, which is validated from the numerical results presented in the last section. The BRR framework facilitates to use the general theory established for a linear problem associated with given nonlinear equations. In particular, existing a posteriori error estimates for the linear problem can be utilized to find reliable error estimators for the nonlinear problem. We mention that this approach is quite general and can be applied to any numerical scheme for nonlinear problems satisfying the BRR framework.

As applications, we consider the pseudostress-velocity formulation of the stationary Navier–Stokes equations (NSE) and the standard Galerkin formulation of semilinear elliptic equations (SEE). Finite element methods for the velocity-pressure and the stress-velocity-pressure formulation of the NSE have been analyzed for decades (cf. [3, 13]). The stress-velocity-pressure formulation is the original physical equations for incompressible Newtonian flows induced by the conservation of momentum and the constitutive law. The mixed finite element methods for the pseudostress-velocity formulation have been recently established in for the Stokes [6, 10] and the Navier–Stokes equations [5, 21]. The pseudostress-velocity formulation has several advantages. The pseudostress is nonsymmetric unlike stress tensor and approximated by the RT_k elements, Raviart–Thomas elements of index $k \geq 0$. Moreover, the approximation of the pressure, the velocity gradient, or even the stress can be algebraically obtained from the approximate value of the pseudostress.

For adaptivity, we derive reliable and efficient residual-based a posteriori error estimators for our multi-level algorithm for the NSE. We mention existing works on the adaptive two-level algorithm. The authors in [12, 14] derived reliable a posteriori error estimators for the two-level methods for the velocity-pressure formulation of the NSE. In comparison to usual adaptive finite element methods [10, 18, 19, 23, 25], the estimators for two-level algorithm contain an additional term which measures the difference between coarse and fine grid solutions. If the additional term measures the difference between consecutive levels and is of higher order (and negligible), then the link between levels does not effect the computation, which is computationally more advantageous. Indeed, this additional term is negligible in our multi-level setting in contrast to [14]. See Remark 3.19 for further details. As the second application of the unified framework, we consider the standard Galerkin method for the SEE and derive reliable a posteriori error estimators. It turns out that behavior of the additional term arising from SEE is similar to that of NSE. Note that our algorithm can be applied to other problems such as the elliptic problem with gradient nonlinearities (cf. [16, 19, 20, 24]).

The remainder of this paper is organized as follows. In the next section, we introduce approximation of branches of nonsingular solutions based on the BRR theory and multi-level algorithm analyzed in [22]. We derive a unified a posteriori error estimate for the approximate solutions of the multi-level algorithm for a class of nonlinear problems in the framework of BRR. Section 3 and Section 4 are devoted to applying the algorithms to the NSE and the SEE, respectively, and constructing a posteriori error estimators of residual type. Section 5 presents our adaptive multi-level algorithm. In Section 6, numerical results are presented to show the performance of the algorithms and to validate the theory developed in this paper. Concluding remarks are given in the last section.

2 BRR Theory and Multi-level Algorithm

In this section, we start with the abstract approximation theory of Brezzi, Rappaz and Raviart for nonlinear problems developed in [4, 7, 13]. Then we introduce the multi-level algorithm analyzed in [22], which is a method to solve resulting nonlinear algebraic equations.

From here to the Remark 2.6, we will use the same parts of what we described in [22]. Let X , \mathcal{X} and Y be real Banach spaces with the norms $\|\cdot\|_X$, $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_Y$, respectively, and assume that $X \hookrightarrow \mathcal{X}$, continuously imbedding. Let $\Lambda \subset \mathbb{R}$ be a compact interval and denote by $\mathcal{L}(X; Y)$ the set of all linear and continuous operators from X into Y . For a given \mathcal{C}^2 map $G: \Lambda \times X \rightarrow Y$, we consider the following nonlinear problem: find $(v, \phi) \in \Lambda \times X$ such that

$$F(v, \phi) := \phi + SG(v, \phi) = 0, \quad (2.1)$$

where $S \in \mathcal{L}(Y; \mathcal{X})$ is a linear operator independent of v . The set $\{(v, \phi(v)) : v \in \Lambda\}$ is called a branch of solutions of (2.1) if $F(v, \phi(v)) = 0$ and the map $v \rightarrow \phi(v)$ is continuous from Λ into X . If the Fréchet derivative $D_\phi F(v, \phi(v))$ of F with respect to ϕ is an isomorphism from X onto \mathcal{X} for all $v \in \Lambda$, then the branch of solutions $\{(v, \phi(v)) : v \in \Lambda\}$ is called nonsingular. To approximate the solution of problem (2.1), we introduce an operator $S_h \in \mathcal{L}(Y; \mathcal{X})$ intended to approximate the linear operator S , where $h > 0$ is a real parameter which will tend to zero. The approximation of nonlinear problem (2.1) is to find $\phi^h \in X$ such that

$$F_h(v, \phi^h) := \phi^h + S_h G(v, \phi^h) = 0. \quad (2.2)$$

We assume that the following hypotheses.

- (H1) G is a \mathcal{C}^2 -mapping from $\Lambda \times X$ into Y , and the second-order Fréchet derivatives of G are bounded on all bounded subsets of $\Lambda \times X$.
- (H2) There exists another Banach space $Z \hookrightarrow Y$, continuously imbedding, such that

$$D_\phi G(v, \phi) \in \mathcal{L}(X; Z) \quad \text{for all } v \in \Lambda \text{ and all } \phi \in X.$$

(H3) For all $y \in Y$, $\lim_{h \rightarrow 0} \|(S - S_h)y\|_{\mathcal{X}} = 0$.

(H4) $\lim_{h \rightarrow 0} \|S - S_h\|_{Z; \mathcal{X}} = 0$.

Under the hypotheses, existence of a unique branch of nonsingular solutions of problem (2.2), its a priori error estimate and its convergence property are stated below. We skip the proof since it is very similar to that of [13, Theorem IV.3.3]. We introduce the following quantities:

$$\begin{aligned} \mu(v) &:= \|D_\phi F(v, \phi(v))^{-1}\|_{\mathcal{X}; X}, \quad \bar{\mu} := \sup_{v \in \Lambda} \mu(v), \\ L(\psi_*(v), \xi) &:= \sup_{v \in \Lambda, \psi \in B(\psi_*(v), \xi)} \|D_\phi^2 G(v, \psi)\|_{X^2; Y}, \end{aligned} \quad (2.3)$$

where $B(\psi_*(v), \xi) := \{\psi \in X; \|\psi_*(v) - \psi\|_X \leq \xi\}$.

Theorem 2.1. *Assume that (H1)–(H4) hold and that $\{(v, \phi(v)) : v \in \Lambda\}$ is a branch of nonsingular solutions of (2.1). Then there exist $\bar{\xi} > 0$ and $\bar{h}(\bar{\xi}) > 0$ such that, for all $h \leq \bar{h}(\bar{\xi})$, there is a \mathcal{C}^2 -function $\phi^h : \Lambda \rightarrow X$ satisfying that $\{(v, \phi^h(v)) : v \in \Lambda\}$ is a branch of nonsingular solutions of (2.2), $\phi^h(v)$ is the only solution of (2.2) in the ball $B(\phi(v), \bar{\xi})$ for all $v \in \Lambda$, and for $K_1 := 4\bar{\mu}(v)$ and a constant $C_S > 0$, the following estimates hold:*

$$\begin{aligned} \|\phi(v) - \phi^h(v)\|_X &\leq K_1 \|F_h(v, \phi(v))\|_{\mathcal{X}} = K_1 \|(S - S_h)G(v, \phi(v))\|_{\mathcal{X}} \quad \text{for all } v \in \Lambda, \\ \|D_\phi F_h(v, \phi^h(v))^{-1}\|_{\mathcal{X}; X} &\leq 2 \|D_\phi F_h(v, \phi(v))^{-1}\|_{\mathcal{X}; X} \leq 4 \|D_\phi F(v, \phi(v))^{-1}\|_{\mathcal{X}; X} \leq K_1, \\ \|S_h\|_{Z; \mathcal{X}} &\leq C_S. \end{aligned}$$

Note that, since Λ is compact, the function $\mu(v)$ is bounded above on Λ . Thus, if needed, we can take K_1 as the constant independent of not only h but also v , i.e., $K_1 := 4\bar{\mu}$. You may refer to [22] for proofs of Lemma 2.2, Theorem 2.5 and Theorem 2.7 that follow in this section.

Lemma 2.2. *Assume that (H1)–(H4) hold and that $D_\phi F(v, \tilde{\phi})$ for $\tilde{\phi} \in X$ is an isomorphism from X onto \mathcal{X} . Then there exists $\xi_* = \xi_*(\tilde{\phi}) > 0$ such that if $\psi \in B(\tilde{\phi}, \xi) \cap X$ for $0 < \xi \leq \xi_*$, then $D_\phi F(v, \psi)$ is an isomorphism from X onto \mathcal{X} and there exists $h_* = h_*(\tilde{\phi}) > 0$ such that if $0 < h \leq h_*$, then $D_\phi F_h(v, \tilde{\phi})$ is an isomorphism from X onto \mathcal{X} with the bound, respectively,*

$$\begin{aligned} \|D_\phi F(v, \psi)^{-1}\|_{\mathcal{X}; X} &\leq 2 \|D_\phi F(v, \tilde{\phi})^{-1}\|_{\mathcal{X}; X}, \\ \|D_\phi F_h(v, \tilde{\phi})^{-1}\|_{\mathcal{X}; X} &\leq 2 \|D_\phi F(v, \tilde{\phi})^{-1}\|_{\mathcal{X}; X}. \end{aligned} \quad (2.4)$$

Remark 2.3. When F is replaced by F_h for h small enough in Lemma 2.2, estimate (2.4) also holds. That is, if $D_\phi F_h(v, \tilde{\phi})$ for $\tilde{\phi} \in X$ is an isomorphism from X onto \mathcal{X} , then there exists $\xi_* = \xi_*(\tilde{\phi}) > 0$ such that, for $\psi \in B(\tilde{\phi}, \xi) \cap X$ with $0 < \xi \leq \xi_*$, $D_\phi F_h(v, \psi)$ is an isomorphism from X onto \mathcal{X} with the bound

$$\|D_\phi F_h(v, \psi)^{-1}\|_{\mathcal{X}; X} \leq 2 \|D_\phi F_h(v, \tilde{\phi})^{-1}\|_{\mathcal{X}; X}.$$

We introduce the multi-level version of the two-grid algorithm. For $j \geq 0$, let $h_j \leq \bar{h}$ given in Theorem 2.1 be mesh sizes, $X_j := X_{h_j}$ finite element spaces such that $X_j \subset X$, and $F_j(v, \phi) := \phi + S_j G(v, \phi)$, where an operator $S_j := S_{h_j}$ is intended to approximate the linear operator S .

Step 1 (Nonlinear solver) Solve nonlinear system on initial coarse mesh.

Find $\phi_0(v) \in X_0$ such that $F_0(v, \phi_0(v)) = 0$.

Step 2 (Linear solver) Update on each finer mesh level j with one Newton iteration.

For $j = 1, 2, \dots, J$, find $\phi_j(v) \in X_j$ such that $D_\phi F_j(v, \phi_{j-1}(v))(\phi_j(v) - \phi_{j-1}(v)) = -F_j(v, \phi_{j-1}(v))$.

Algorithm 1: Multi-level algorithm

Remark 2.4. The two-grid scheme in [27, 28] is based on passing information between finite element equations defined on two grids of different mesh sizes. In the first step, a nonlinear problem itself is solved in a coarse space, i.e., a finite-dimensional space with coarse grids. In the second step, the nonlinear problem is linearized locally at the solution obtained in the coarse space. Then the linearized problem is solved in a fine space. For better performance, this process can be iterated on a sequence of linearized problems with increasing dimensions. In the second step of the multi-level algorithm in [14], the nonlinear problem is linearized locally at the solution ϕ_0 obtained in the fixed coarse space of Step 1. By contrast, in our multi-level algorithm, the nonlinear problem is linearized locally at the solution obtained on the very previous mesh. The advantages of our algorithm are discussed in detail in Remark 3.19.

Theorem 2.5 (A Priori Error Bound). *Assume that (H1)–(H4) hold. Let $\phi^{h_j}(v)$ be nonsingular solutions of (2.2) on X_j . And let $\phi_j(v)$ be solutions obtained from Step 2 of Algorithm 1. Then there exist $\xi > 0$ with $\xi \leq \bar{\xi}$ given in Theorem 2.1 and $h_0^* > 0$ such that, for $h_j \leq h_0^*$, $\phi^{h_j}(v) \in B(\phi(v), \xi/2)$ and $\phi_j(v)$ belongs to the ball $B(\phi^{h_j}(v), \xi/2)$ for $j \geq 1$. Moreover, for the positive constant $K_2 := 4\bar{\mu}C_S L(\phi, \bar{\xi})$ independent of mesh sizes h_j and v , we have the following quadratic relation:*

$$\|\phi^{h_j}(v) - \phi_j(v)\|_X \leq K_2 \|\phi^{h_j}(v) - \phi_{j-1}(v)\|_X^2, \quad (2.5)$$

and for $K_3 := \max\{K_1, 2K_1^2 K_2, 2K_2\}$, we have an a priori estimate

$$\|\phi(v) - \phi_j(v)\|_X \leq K_3 (\|(S - S_j)G(v, \phi(v))\|_{\mathcal{X}} + \|(S - S_j)G(v, \phi(v))\|_{\mathcal{X}}^2 + \|\phi(v) - \phi_{j-1}(v)\|_X^2).$$

Remark 2.6. It should be noted that the above algorithm can be applied even if \mathcal{T}_j is not a uniform refinement of \mathcal{T}_{j-1} and adaptive mesh refining can be exploited along with this algorithm. The multi-level algorithm has the quadratic convergence in the sense of equation (2.5) even on adaptive meshes. This is confirmed by numerical examples in Section 6.

Now, we want to find an a posteriori error bound for solutions obtained from the multi-level algorithm. First, we introduce the following theorem presented and proved in [7, Theorem 2.1]. Let $F: X \rightarrow \mathcal{X}$ be a C^1 mapping and let $v \in X$ be such that $DF(v) \in \mathcal{L}(X; \mathcal{X})$ is an isomorphism. We introduce the notations

$$\epsilon := \|F(v)\|_{\mathcal{X}}, \quad \gamma := \|DF(v)^{-1}\|_{\mathcal{X}; X}, \quad L(\alpha) := \sup_{w \in B(v, \alpha)} \|DF(v) - DF(w)\|_{\mathcal{X}; \mathcal{X}}.$$

Theorem 2.7. *We assume that $2\gamma L(2\gamma\epsilon) \leq 1$. Then $F(x) = 0$ has a unique solution u in the ball $B(v, 2\gamma\epsilon)$ and $DF(u) \in \mathcal{L}(X; \mathcal{X})$ is invertible with*

$$\|DF(u)^{-1}\|_{\mathcal{X}; X} \leq 2\gamma.$$

Moreover, for all $v' \in B(v, 2\gamma\epsilon)$,

$$\|v' - u\|_X \leq 2\gamma \|F(v')\|_{\mathcal{X}}.$$

Theorem 2.8 (A Posteriori Error Bound). *Assume that (H1)–(H4) hold. Let $\phi(v)$ be a nonsingular solution of (2.1) and let $\phi_j(v)$ be the approximate solution produced by Algorithm 1. Then there exists $0 < h_1^* \leq h_0^*$ given in Theorem 2.5 such that, for all $h_j \leq h_1^*$, we have $\phi^{h_j}(v) \in B(\phi(v), \xi/2)$ and $\phi_j(v) \in B(\phi^{h_j}(v), \xi/2)$ for ξ given in Theo-*

rem 2.5, and we have a residual type a posteriori error bound with $K_1 := 4\mu(\nu)$,

$$\|\phi(\nu) - \phi_j(\nu)\|_X \leq 2\|D_\phi F(\nu, \phi_j(\nu))^{-1}\|_{\mathcal{X};X} \|F(\nu, \phi_j(\nu))\|_{\mathcal{X}} \leq K_1 \|F(\nu, \phi_j(\nu))\|_{\mathcal{X}}. \quad (2.6)$$

Proof. The proof follows by applying Theorem 2.7. For this, we first notice that $\phi_j(\nu) \in B(\phi(\nu), \xi)$ for h_j with $h_j \leq h_0^*$ given in Theorem 2.5 and $\phi_j(\nu) \rightarrow \phi(\nu)$ strongly in X as $h_j \rightarrow 0$ thanks to Theorem 2.5 and (H3). By application of Lemma 2.2, $D_\phi F(\nu, \phi_j(\nu))$ is an isomorphism from X onto \mathcal{X} with

$$\|D_\phi F(\nu, \phi_j(\nu))^{-1}\|_{\mathcal{X};X} \leq 2\|D_\phi F(\nu, \phi(\nu))^{-1}\|_{\mathcal{X};X} =: \gamma. \quad (2.7)$$

To apply Theorem 2.7, we set

$$\begin{aligned} \epsilon_j &:= \|F(\nu, \phi_j(\nu))\|_{\mathcal{X}}, \quad \gamma_j := \|D_\phi F(\nu, \phi_j(\nu))^{-1}\|_{\mathcal{X};X}, \\ L_j(\alpha) &:= \sup_{w \in B(\phi_j(\nu), \alpha)} \|D_\phi F(\nu, \phi_j(\nu)) - D_\phi F(\nu, w)\|_{X;X}. \end{aligned}$$

From the mean value theorem and (H1), we have, for $C_1 = \|S\|_{Y;X} L(\phi_j(\nu), \alpha)$ with $L(\cdot, \cdot)$ defined in (2.3),

$$\|D_\phi F(\nu, \phi_j(\nu)) - D_\phi F(\nu, w)\|_{X;X} \leq \|S\|_{Y;X} \|D_\phi G(\nu, \phi_j(\nu)) - D_\phi G(\nu, w)\|_{X;Y} \leq C_1 \|\phi_j(\nu) - w\|_X,$$

which yields that there exist $0 < h'_0 \leq h_0^*$ and $0 < \alpha < \xi$ such that, for $h_j \leq h'_0$,

$$L_j(\alpha) \leq \frac{1}{2\gamma} \quad \text{and} \quad \|\phi(\nu) - \phi_j(\nu)\|_X < \alpha. \quad (2.8)$$

Applying Theorem 2.5, the mean value theorem and (H1) to the identity

$$F(\nu, \phi_j(\nu)) = F(\nu, \phi_j(\nu)) - F(\nu, \phi(\nu)) = \phi_j(\nu) - \phi(\nu) + S(G(\nu, \phi_j(\nu)) - G(\nu, \phi(\nu))),$$

we can find $h_1^* \leq h'_0$ such that, for all $h_j \leq h_1^*$,

$$2\gamma\epsilon_j \leq \alpha. \quad (2.9)$$

From (2.7)–(2.9), we have

$$2\gamma_j L_j(2\gamma_j \epsilon_j) \leq 2\gamma L_j(2\gamma \epsilon_j) \leq 2\gamma L_j(\alpha) \leq 1.$$

Applying Theorem 2.7 with $v = \phi_j(\nu)$, $h_j \leq h_1^*$, we conclude that there exists a unique $\psi \in B(\phi_j, 2\gamma_j \epsilon_j)$ such that

$$F(\psi) = 0 \quad \text{and} \quad \|\psi - \phi_j(\nu)\|_X \leq 2\|D_\phi F(\nu, \phi_j(\nu))^{-1}\|_{\mathcal{X};X} \|F(\nu, \phi_j(\nu))\|_{\mathcal{X}}.$$

From the uniqueness of ψ in $B(\phi_j, \alpha) \subset B(\phi_j, \xi)$ and the fact that $F(\nu, \phi(\nu)) = 0$ and $\phi \in B(\phi_j, \xi)$, we have $\psi = \phi(\nu)$ and (2.6) is proved. \square

3 Application 1: Navier–Stokes Equations

In this section, we introduce the a priori error estimate and analyze an a posteriori error for multi-level algorithms for the pseudostress-velocity formulation of the NSE. The a priori error was analyzed in [21]. The following details are only given for the reader's convenience. First, we introduce some notation and function spaces. Let div and ∇ denote the divergence and gradient operators, respectively. For $\mathbf{v} = (v_1, \dots, v_d) \in \mathbb{R}^d$ and $\boldsymbol{\tau} = (\tau_{ij})_{d \times d} \in \mathbb{R}^{d \times d}$, let $\boldsymbol{\tau}_i = (\tau_{i1}, \dots, \tau_{id})$ denote its i th row for $i = 1, \dots, d$ and define

$$\mathbf{div} \boldsymbol{\tau} := (\text{div} \boldsymbol{\tau}_1, \dots, \text{div} \boldsymbol{\tau}_d), \quad \mathbf{v} \cdot \boldsymbol{\tau} := (\mathbf{v} \cdot \boldsymbol{\tau}_1, \dots, \mathbf{v} \cdot \boldsymbol{\tau}_d),$$

$$\nabla \mathbf{v} := (\nabla v_1, \dots, \nabla v_d)^t = \left(\frac{\partial v_i}{\partial x_j} \right)_{d \times d}, \quad \boldsymbol{\delta} := d \times d \text{ unit matrix.}$$

Given data \mathbf{f} and \mathbf{g} , the stationary and incompressible NSE for the unknown velocity \mathbf{u} and the pressure p reads

$$-\nu \Delta \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega, \quad (3.1)$$

$$\text{div} \mathbf{u} = 0 \quad \text{in } \Omega, \quad (3.2)$$

$$\mathbf{u} = \mathbf{g} \quad \text{on } \partial\Omega. \quad (3.3)$$

The compatibility conditions read

$$\int_{\partial\Omega} \mathbf{g} \cdot \mathbf{n} \, ds = 0 \quad \text{and} \quad \int_{\Omega} p \, dx = 0,$$

where \mathbf{n} is the outward unit normal vector to the boundary.

Let $\mathcal{A}: \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ be the deviatoric operator

$$\mathcal{A}\boldsymbol{\tau} := \boldsymbol{\tau} - \frac{1}{d}(\text{tr } \boldsymbol{\tau})\boldsymbol{\delta}.$$

The pseudostress tensor $\boldsymbol{\sigma} := \nu \nabla \mathbf{u} - p \boldsymbol{\delta}$ allows the pseudostress-velocity formulation of NSE (3.1)–(3.2),

$$\mathcal{A}\left(\frac{\boldsymbol{\sigma}}{\nu}\right) - \nabla \mathbf{u} = \mathbf{0} \quad \text{in } \Omega, \quad (3.4)$$

$$\text{div } \boldsymbol{\sigma} - \mathbf{u} \cdot \mathcal{A}\left(\frac{\boldsymbol{\sigma}}{\nu}\right) = -\mathbf{f} \quad \text{in } \Omega. \quad (3.5)$$

Equation (3.4) is obtained from

$$\text{tr}(\nabla \mathbf{u}) = \text{div } \mathbf{u} = 0 \quad \text{and} \quad \text{tr } \boldsymbol{\sigma} = -dp.$$

The compatibility condition $\int_{\Omega} p \, dx = 0$ implies

$$\int_{\Omega} \text{tr } \boldsymbol{\sigma} \, dx = 0.$$

Scaling pseudostress $\boldsymbol{\sigma}$ and data \mathbf{f} by $\boldsymbol{\sigma}/\nu \rightarrow \boldsymbol{\sigma}$ and $\mathbf{f}/\nu \rightarrow \mathbf{f}$, (3.4)–(3.5) may be rewritten as

$$\mathcal{A}\boldsymbol{\sigma} - \nabla \mathbf{u} = \mathbf{0} \quad \text{in } \Omega, \quad (3.6)$$

$$\text{div } \boldsymbol{\sigma} - \frac{1}{\nu} \mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma} = -\mathbf{f} \quad \text{in } \Omega. \quad (3.7)$$

Assume that Ω is a bounded, open, connected subset of \mathbb{R}^d ($d = 2$ or 3) with the boundary $\partial\Omega$ of convex polygon/polyhedra. We use the standard notation and definitions for the Sobolev spaces $W^{r,p}(\Omega)$ and $W^{r,p}(\partial\Omega)$ for an integer $r \geq 0$ and $p \in [1, \infty]$. The standard associated norms are denoted by $\|\cdot\|_{r,p} := \|\cdot\|_{r,p,\Omega}$ and $\|\cdot\|_{r,p,\partial\Omega}$. For $r = 0$, $W^{r,p}(\Omega)$ coincides with $L^p(\Omega)$. Moreover, the space $W^{r,2}(\Omega)$ will be written in the form $H^r(\Omega)$. Let

$$\begin{aligned} H(\text{div}; \Omega)^d &= \{\boldsymbol{\tau} \in L^2(\Omega)^{d \times d} : \text{div } \boldsymbol{\tau} \in L^2(\Omega)^d\}, \\ W^{0,3}(\text{div}; \Omega)^d &= \{\boldsymbol{\tau} \in W^{0,3}(\Omega)^{d \times d} : \text{div } \boldsymbol{\tau} \in W^{0,3}(\Omega)^d\}, \\ L_0^2(\Omega) &= \left\{ q \in L^2(\Omega) \mid \int_{\Omega} q \, dx = 0 \right\}. \end{aligned}$$

We define spaces

$$\begin{aligned} \hat{H}(\text{div}; \Omega)^d &= \left\{ \boldsymbol{\tau} \in H(\text{div}; \Omega)^d \mid \int_{\Omega} \text{tr } \boldsymbol{\tau} \, dx = 0 \right\}, \\ \hat{W}^{0,3}(\text{div}; \Omega)^d &= \left\{ \boldsymbol{\tau} \in W^{0,3}(\text{div}; \Omega)^d \mid \int_{\Omega} \text{tr } \boldsymbol{\tau} \, dx = 0 \right\}, \\ \hat{H}^r(\Omega)^{d \times d} &= \left\{ \boldsymbol{\tau} \in H^r(\Omega)^{d \times d} \mid \int_{\Omega} \text{tr } \boldsymbol{\tau} \, dx = 0 \right\}. \end{aligned}$$

Then the mixed variational problem of the pseudostress-velocity formulation (3.6)–(3.7) is to find

$$(\boldsymbol{\sigma}, \mathbf{u}) \in \hat{W}^{0,3}(\text{div}; \Omega)^d \times L^2(\Omega)^d$$

such that

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\text{div } \boldsymbol{\tau}, \mathbf{u}) = g_*(\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \hat{H}(\text{div}; \Omega)^d, \quad (3.8)$$

$$(\text{div } \boldsymbol{\sigma}, \mathbf{v}) - \frac{1}{\nu} (\mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma}, \mathbf{v}) = f_*(\mathbf{v}) \quad \text{for all } \mathbf{v} \in L^2(\Omega)^d, \quad (3.9)$$

where (\cdot, \cdot) denotes the $L^2(\Omega)^{d \times d}$ or $L^2(\Omega)^d$ inner product, $g_*(\boldsymbol{\tau}) = \int_{\partial\Omega} (\mathbf{n} \cdot \boldsymbol{\tau}) \cdot \mathbf{g} \, ds$ and $f_*(\mathbf{v}) = - \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, dx$. Note that if we use the trial space $\hat{W}^{0,3}(\text{div}; \Omega)^d \times L^2(\Omega)^d$, then we can guarantee the nonlinear term, $\mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma}$, to be in $L^2(\Omega)^d$ from the imbedding theorem (see [5, Remark 2.1]). The well-posedness and uniqueness of system (3.8)–(3.9) is established by the following well-known theorem (see [5]).

Theorem 3.1. *Let \mathbf{f} be in $L^2(\Omega)^d$. For $\mathbf{g} \in H^{3/2}(\partial\Omega)^d$, system (3.8)–(3.9) has solutions $(\boldsymbol{\sigma}, \mathbf{u})$ which belong to $H^1(\Omega)^{d \times d} \times H^2(\Omega)^d$. Moreover, if $\nu > \nu_0(\Omega; \mathbf{f}, \mathbf{g})$ (see [5] or [13] for the specific value of ν_0), then the solution $(\boldsymbol{\sigma}, \mathbf{u})$ is unique.*

3.1 BRR Framework for NSE

We let

$$G(\nu, \phi) := \left(-\mathbf{g}, -\frac{1}{\nu} \mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma} + \mathbf{f} \right)$$

for $\phi := (\boldsymbol{\sigma}, \mathbf{u})$ and $Y := H^{3/2}(\partial\Omega)^d \times L^2(\Omega)^d$. Let $\Lambda \subset (0, \infty)$ be a compact interval. We consider $Z = Y$. Then condition (H4) implies condition (H3). Note the Fréchet derivative of G at $(\nu, \phi) = (\nu, (\boldsymbol{\sigma}, \mathbf{u}))$ is

$$D_\phi G(\nu, \phi)(\psi) = \left(\mathbf{0}, -\frac{1}{\nu} (\mathbf{u} \cdot \mathcal{A}\boldsymbol{\tau} + \mathbf{v} \cdot \mathcal{A}\boldsymbol{\sigma}) \right)$$

for $\psi = (\boldsymbol{\tau}, \mathbf{v})$. To guarantee that $D_\phi G(\nu, \phi)(\psi) \in Z$ for all ψ , we let $X := L^3(\Omega)^{d \times d} \times L^6(\Omega)^d$. In fact, note that, for any $\psi = (\boldsymbol{\tau}, \mathbf{v}) \in X$,

$$\begin{aligned} \|\mathbf{u} \cdot \mathcal{A}\boldsymbol{\tau} + \mathbf{v} \cdot \mathcal{A}\boldsymbol{\sigma}\|_{0,2} &\leq \|\mathbf{u} \cdot \mathcal{A}\boldsymbol{\tau}\|_{0,2} + \|\mathbf{v} \cdot \mathcal{A}\boldsymbol{\sigma}\|_{0,2} \\ &\leq \|\mathbf{u}\|_{0,6} \|\boldsymbol{\tau}\|_{0,3} + \|\mathbf{v}\|_{0,6} \|\boldsymbol{\sigma}\|_{0,3} \\ &\leq (\|\boldsymbol{\sigma}\|_{0,3}^2 + \|\mathbf{u}\|_{0,6}^2)^{1/2} (\|\boldsymbol{\tau}\|_{0,3}^2 + \|\mathbf{v}\|_{0,6}^2)^{1/2}. \end{aligned}$$

Here we used the triangle inequality and the Hölder inequality. So hypothesis (H2) holds. From the definition, G belongs to the \mathcal{C}^2 -class clearly. So hypothesis (H1) holds.

We let $\mathcal{X} := L^2(\Omega)^{d \times d} \times L^2(\Omega)^d$ and introduce the linear solution operator

$$\begin{aligned} S: Y &\rightarrow \mathcal{X}, \\ (\mathbf{g}^*, \mathbf{f}^*) &\rightarrow S(\mathbf{g}^*, \mathbf{f}^*) = (\boldsymbol{\sigma}, \mathbf{u}) \end{aligned}$$

defined by the solution of

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\text{div } \boldsymbol{\tau}, \mathbf{u}) = \int_{\partial\Omega} (\mathbf{n} \cdot \boldsymbol{\tau}) \cdot \mathbf{g}^* \, ds \quad \text{for all } \boldsymbol{\tau} \in \hat{H}(\text{div}; \Omega)^d, \quad (3.10)$$

$$(\text{div } \boldsymbol{\sigma}, \mathbf{v}) = - \int_{\Omega} \mathbf{f}^* \cdot \mathbf{v} \, dx \quad \text{for all } \mathbf{v} \in L^2(\Omega)^d. \quad (3.11)$$

The following result is well known (see [5, Lemma 5.1]).

Lemma 3.2. *For any $(\mathbf{g}^*, \mathbf{f}^*) \in Y$, Stokes equations (3.10)–(3.11) have a unique solution $(\boldsymbol{\sigma}, \mathbf{u}) = S(\mathbf{g}^*, \mathbf{f}^*)$ which is in $H^1(\Omega)^{d \times d} \times H^2(\Omega)^d \hookrightarrow \mathcal{X}$.*

Clearly, the pair $(\nu, \phi) \in \Lambda \times X$ is a solution of

$$F(\nu, \phi) := \phi + SG(\nu, \phi) = 0 \quad (3.12)$$

if and only if the function $\phi = (\boldsymbol{\sigma}, \mathbf{u})$ is a solution of (3.8)–(3.9).

Theorem 3.3. *For any $(\mathbf{g}, \mathbf{f}) \in Y$, there is a $\nu_0(\Omega; \mathbf{f}, \mathbf{g})$ such that if $\nu > \nu_0(\Omega; \mathbf{f}, \mathbf{g})$, where ν_0 is the fixed value stated in Theorem 3.1, then nonlinear problem (3.12) has a branch of nonsingular solutions.*

Proof. See [21]. □

We now consider mixed finite element approximations. Let $\{\mathcal{T}_h\}$ be a family of shape-regular triangulations of $\bar{\Omega}$ by triangles ($d = 2$) and tetrahedra ($d = 3$), respectively, of diameter h_T , and let $h := \max_{T \in \mathcal{T}_h} h_T$ denote the mesh size. Associated with \mathcal{T}_h , we define a finite subspace of X by

$$X_h := \widehat{\text{RT}}_k^d(\mathcal{T}_h) \times P_k^d(\mathcal{T}_h),$$

where $P_k^d(\mathcal{T}_h) := \prod_{i=1}^d P_k(\mathcal{T}_h)$ (i.e., a product space of polynomials of degree k) and

$$\widehat{\text{RT}}_k^d(\mathcal{T}_h) := \left\{ \tau \in \text{RT}_k^d(\mathcal{T}_h) \mid \int_{\Omega} \text{tr } \tau \, dx = 0 \right\}$$

for $\text{RT}_k^d(\mathcal{T}_h) := \prod_{i=1}^d \text{RT}_k(\mathcal{T}_h)$ (i.e., a product space of Raviart–Thomas space of order k).

For short notation on generic constants C independent of the local or global mesh sizes, for any expressions A and B ,

$$A \lesssim B \quad \text{abbreviates} \quad A \leq CB.$$

We introduce a projection operator over the space W , where $W := \hat{H}^1(\Omega)^{d \times d}$ or $W := \hat{H}(\text{div}; \Omega)^d \cap L^s(\Omega)^{d \times d}$ for some $s > 2$. Let $\tilde{\Pi}_h$ denote the Raviart–Thomas projection operator (see [3]) associated with the degrees of freedom onto $\widehat{\text{RT}}_k^d(\mathcal{T}_h) + \text{span}\{\boldsymbol{\delta}\}$. We define $\Pi_h: W \rightarrow \widehat{\text{RT}}_k^d(\mathcal{T}_h)$ by

$$\Pi_h \tau = \tilde{\Pi}_h \tau - \frac{\int_{\Omega} (\text{tr } \tilde{\Pi}_h \tau) \, dx}{d|\Omega|} \boldsymbol{\delta} \quad \text{for all } \tau \in W,$$

where $|\Omega| = \int_{\Omega} dx$. Then we have $\int_{\Omega} (\text{tr } \Pi_h \tau) \, dx = 0$. We notice that the operator $\tilde{\Pi}_h$ approximates the normal components on element edges such that, for $E \in \mathcal{E}_h$, the set of all edges of \mathcal{T}_h , for $\mathbf{v}_h \in P_k^d(\mathcal{T}_h)$ and for $\tau \in W$,

$$\langle \mathbf{v}_h, \mathbf{n}_E \cdot (\tau - \tilde{\Pi}_h \tau) \rangle_E = 0. \quad (3.13)$$

But, in general, the above relation does not hold for the operator Π_h .

Let \mathbf{P}_h be the L^2 -projection onto $P_k^d(\mathcal{T}_h)$ with the well-known approximation property

$$\|\mathbf{P}_h \mathbf{v} - \mathbf{v}\|_{0,p,T} \leq h_T^r |\mathbf{v}|_{r,p,T}, \quad 0 \leq r \leq k+1, \quad \text{for all } \mathbf{v} \in W^{r,p}(T)^d. \quad (3.14)$$

Then the following lemma holds (see [6] for details).

Lemma 3.4. *The commutative property $\text{div } \Pi_h = \mathbf{P}_h \text{div}$ holds. Furthermore, we have*

$$\|\tau - \Pi_h \tau\|_{0,2} \leq h^r |\tau|_{r,2} \quad \text{for } 1 \leq r \leq k+1, \quad (3.15)$$

$$\|\text{div } \tau - \text{div}(\Pi_h \tau)\|_{0,2} \leq h^r |\text{div } \tau|_{r,2} \quad \text{for } 0 \leq r \leq k+1.$$

The mixed finite element method for approximation of the solution of system (3.8)–(3.9) is defined by finding $(\boldsymbol{\sigma}^h, \mathbf{u}^h) \in X_h$ such that

$$(\mathcal{A} \boldsymbol{\sigma}^h, \tau) + (\mathbf{u}^h, \text{div } \tau) = g_*(\tau) \quad \text{for all } \tau \in \widehat{\text{RT}}_k^d(\mathcal{T}_h), \quad (3.16)$$

$$(\text{div } \boldsymbol{\sigma}^h, \mathbf{v}) - \left(\frac{1}{\nu} \mathbf{u}^h \cdot \mathcal{A} \boldsymbol{\sigma}^h, \mathbf{v} \right) = f_*(\mathbf{v}) \quad \text{for all } \mathbf{v} \in P_k^d(\mathcal{T}_h). \quad (3.17)$$

Next, we define a discrete solution operator $S_h: Y \rightarrow X_h$ by

$$S_h(\mathbf{g}^*, \mathbf{f}^*) = (\boldsymbol{\sigma}_h, \mathbf{u}_h) \in X_h,$$

where $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in X_h$ is the solution of the discrete counterpart of (3.10)–(3.11),

$$(\mathcal{A} \boldsymbol{\sigma}_h, \tau) + (\mathbf{u}_h, \text{div } \tau) = \int_{\partial \Omega} (\mathbf{n} \cdot \tau) \cdot \mathbf{g}^* \, ds \quad \text{for all } \tau \in \widehat{\text{RT}}_k^d(\mathcal{T}_h), \quad (3.18)$$

$$(\text{div } \boldsymbol{\sigma}_h, \mathbf{v}) = - \int_{\Omega} \mathbf{f}^* \cdot \mathbf{v} \, dx \quad \text{for all } \mathbf{v} \in P_k^d(\mathcal{T}_h). \quad (3.19)$$

Now, the discrete nonlinear problem of (3.12) is to find $\phi^h = (\boldsymbol{\sigma}^h, \mathbf{u}^h)$ in X_h such that

$$F_h(\mathbf{v}, \phi^h) := \phi^h + S_h G(\mathbf{v}, \phi^h) = 0. \quad (3.20)$$

Note that $\phi^h = (\boldsymbol{\sigma}^h, \mathbf{u}^h)$ is the solution of (3.20) if and only if $\phi^h \in X_h$ satisfies (3.16)–(3.17).

Theorem 3.5. For any $(\mathbf{g}^*, \mathbf{f}^*) \in Y$, let $(\boldsymbol{\sigma}, \mathbf{u}) = S(\mathbf{g}^*, \mathbf{f}^*)$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h) = S_h(\mathbf{g}^*, \mathbf{f}^*)$ be the solution of (3.10)–(3.11) and (3.18)–(3.19), respectively. Then

$$\lim_{h \rightarrow 0} \|S - S_h\|_{Z; \mathcal{X}} = 0.$$

Moreover, assume that $(\boldsymbol{\sigma}, \mathbf{u}) \in H^r(\Omega)^{d \times d} \times H^r(\Omega)^d$ for $1 \leq r \leq k+1$; then we have

$$\|S(\mathbf{g}^*, \mathbf{f}^*) - S_h(\mathbf{g}^*, \mathbf{f}^*)\|_{\mathcal{X}} \lesssim h^r (\|\boldsymbol{\sigma}\|_{r,2} + \|\mathbf{u}\|_{r,2}) \lesssim h^r (\|\mathbf{f}^*\|_{r-2,2} + \|\mathbf{g}^*\|_{r-\frac{1}{2},2,\partial\Omega}).$$

Proof. See [21]. □

Theorem 3.5 implies that hypotheses (H3) and (H4) hold. We get the following theorem.

Theorem 3.6. Let $(\mathbf{g}, \mathbf{f}) \in Y$. Assume that $v > v_0(\Omega; \mathbf{f}, \mathbf{g})$, where v_0 is the fixed value stated in Theorem 3.3, and that $(v, \phi(v)) = (v, (\boldsymbol{\sigma}(v), \mathbf{u}(v)))$ is a branch of nonsingular solutions of (3.12). Then, for h sufficiently small, there exist a neighborhood \mathcal{O} of the origin in X and a unique C^2 -function $v \rightarrow \phi^h(v) \in X_h$ such that $\{(v, \phi^h(v)) : v \in \Lambda\}$ is a branch of nonsingular solutions of (3.20) and that $\phi(v) - \phi^h(v) \in \mathcal{O}$.

Moreover, assume that $(\boldsymbol{\sigma}(v), \mathbf{u}(v)) \in H^r(\Omega)^{d \times d} \times H^r(\Omega)^d$ for $r = 1, 2, \dots, k+1$. Then we have

$$\|\phi(v) - \phi^h(v)\|_X := \|\boldsymbol{\sigma}(v) - \boldsymbol{\sigma}^h(v)\|_{0,3} + \|\mathbf{u}(v) - \mathbf{u}^h(v)\|_{0,6} \leq h^r (\|\boldsymbol{\sigma}(v)\|_{r,2} + \|\mathbf{u}(v)\|_{r,2}),$$

where $\phi^h(v) = (\boldsymbol{\sigma}^h(v), \mathbf{u}^h(v))$.

Proof. See [21]. □

3.2 Error Estimates for Multi-level Algorithm for NSE

Now, we apply the multi-level algorithm to NSE based on pseudostress-velocity formulation. Note that the Fréchet derivative of F at $(v, \phi) = (v, (\boldsymbol{\sigma}, \mathbf{u}))$ is

$$D_\phi F(v, \phi)(\psi) = (I + SD_\phi G(v, \phi))(\psi) = \psi + S\left(\mathbf{0}, -\frac{1}{v}(\mathbf{u} \cdot \mathcal{A}\boldsymbol{\tau} + \mathbf{v} \cdot \mathcal{A}\boldsymbol{\sigma})\right)$$

for any $\psi = (\boldsymbol{\tau}, \mathbf{v}) \in X$.

We generate a shape-regular triangulation \mathcal{T}_{j+1} from \mathcal{T}_j ($j \geq 1$) with newest vertex bisection or longest edge bisection. For convenience, we omit v from the functions $\phi_j(v)$.

Step 1 (Nonlinear solver) Find $\phi_0 = (\boldsymbol{\sigma}_0, \mathbf{u}_0) \in X_0$, where $X_0 := \widehat{\text{RT}}_k^d(\mathcal{T}_0) \times P_k^d(\mathcal{T}_0)$, such that

$$(\mathcal{A}\boldsymbol{\sigma}_0, \boldsymbol{\tau}) + (\mathbf{div} \boldsymbol{\tau}, \mathbf{u}_0) = g_*(\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \widehat{\text{RT}}_k^d(\mathcal{T}_0), \quad (3.21)$$

$$(\mathbf{div} \boldsymbol{\sigma}_0, \mathbf{v}) - \frac{1}{v}(\mathbf{u}_0 \cdot \mathcal{A}\boldsymbol{\sigma}_0, \mathbf{v}) = f_*(\mathbf{v}) \quad \text{for all } \mathbf{v} \in P_k^d(\mathcal{T}_0), \quad (3.22)$$

Step 2 (Linear solver) For $1 \leq j \leq J$, find $\phi_j = (\boldsymbol{\sigma}_j, \mathbf{u}_j) \in X_j := \widehat{\text{RT}}_k^d(\mathcal{T}_j) \times P_k^d(\mathcal{T}_j)$ such that

$$(\mathcal{A}\boldsymbol{\sigma}_j, \boldsymbol{\tau}) + (\mathbf{div} \boldsymbol{\tau}, \mathbf{u}_j) = g_*(\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \widehat{\text{RT}}_k^d(\mathcal{T}_j), \quad (3.23)$$

$$(\mathbf{div} \boldsymbol{\sigma}_j, \mathbf{v}) - \frac{1}{v}(\mathbf{u}_{j-1} \cdot \mathcal{A}\boldsymbol{\sigma}_j + \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_{j-1}, \mathbf{v}) = -\frac{1}{v}(\mathbf{u}_{j-1} \cdot \mathcal{A}\boldsymbol{\sigma}_{j-1}, \mathbf{v}) + f_*(\mathbf{v}) \quad \text{for all } \mathbf{v} \in P_k^d(\mathcal{T}_j). \quad (3.24)$$

Algorithm 2: Multi-level algorithm for NSE

If \mathcal{T}_j with diameter h_j is uniformly refined triangulations from the previous triangulations \mathcal{T}_{j-1} with diameter h_{j-1} , from the results of Theorem 2.5 and Theorem 3.6, the following theorem for Algorithm 1 can be easily proved.

Theorem 3.7 (A Priori Error Estimates). *Let $(\mathbf{g}, \mathbf{f}) \in Y$. Assume that $\nu > \nu_0(\Omega; \mathbf{f}, \mathbf{g})$, where ν_0 is the fixed value stated in Theorem 3.3, and that $(\nu, \phi(\nu)) = (\nu, (\boldsymbol{\sigma}(\nu), \mathbf{u}(\nu)))$ is a branch of nonsingular solutions of (3.12). Let $\phi_0(\nu) = (\boldsymbol{\sigma}_0(\nu), \mathbf{u}_0(\nu))$ be the solution of (3.21)–(3.22) and $\phi_j(\nu) = (\boldsymbol{\sigma}_j(\nu), \mathbf{u}_j(\nu))$, $1 \leq j \leq J$, solutions of (3.23)–(3.24). If $(\boldsymbol{\sigma}(\nu), \mathbf{u}(\nu)) \in H^r(\Omega)^{d \times d} \times H^r(\Omega)^d$ for $r = 1, 2, \dots, k+1$, then we have*

$$\|\phi(\nu) - \phi_j(\nu)\|_X = \|\boldsymbol{\sigma}(\nu) - \boldsymbol{\sigma}_j(\nu)\|_{0,3} + \|\mathbf{u}(\nu) - \mathbf{u}_j(\nu)\|_{0,6} \lesssim (h_{j-1}^{2r} + h_j^r)(\|\boldsymbol{\sigma}(\nu)\|_{r,2} + \|\mathbf{u}(\nu)\|_{r,2}).$$

Proof. See [21]. □

Theorem 2.8 indeed provides the reliability bound. For brevity, ν is dropped from the functions $\phi_j(\nu)$ and $\bar{\phi}_j(\nu)$. First, observe that

$$\|F(\nu, \phi_j)\|_{\mathcal{X}'} = \|\phi_j + SG(\nu, \phi_j)\|_{\mathcal{X}'} = \|\phi_j - \bar{\phi}_j\|_{L^2(\Omega)^{d \times d} \times L^2(\Omega)^d}, \quad (3.25)$$

where $\bar{\phi}_j := (\bar{\boldsymbol{\sigma}}, \bar{\mathbf{u}}) \in H^1(\Omega)^{d \times d} \times H^2(\Omega)^d$ is the unique solution for the pseudostress-velocity formulation of the Stokes problem

$$(\mathcal{A}\bar{\boldsymbol{\sigma}}, \boldsymbol{\tau}) + (\bar{\mathbf{u}}, \mathbf{div} \boldsymbol{\tau}) = g_*(\boldsymbol{\tau}) \quad \text{for all } \boldsymbol{\tau} \in \hat{H}(\mathbf{div}; \Omega)^d, \quad (3.26)$$

$$(\mathbf{div} \bar{\boldsymbol{\sigma}}, \mathbf{v}) = \frac{1}{\nu}(\mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j, \mathbf{v}) + f_*(\mathbf{v}) \quad \text{for all } \mathbf{v} \in L^2(\Omega)^d. \quad (3.27)$$

The following error equations immediately follow from (3.23)–(3.24) and (3.26)–(3.27):

$$\begin{aligned} (\mathcal{A}\bar{\boldsymbol{\sigma}} - \mathcal{A}\boldsymbol{\sigma}_j, \boldsymbol{\tau}) + (\bar{\mathbf{u}} - \mathbf{u}_j, \mathbf{div} \boldsymbol{\tau}) &= 0 & \text{for all } \boldsymbol{\tau} \in \widehat{\mathbf{RT}}_k^d(\mathcal{T}_j), \\ (\mathbf{div} \bar{\boldsymbol{\sigma}} - \mathbf{div} \boldsymbol{\sigma}_j, \mathbf{v}) &= \frac{1}{\nu}((\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot (\mathcal{A}\boldsymbol{\sigma}_j - \mathcal{A}\boldsymbol{\sigma}_{j-1}), \mathbf{v}) & \text{for all } \mathbf{v} \in P_k^d(\mathcal{T}_j). \end{aligned} \quad (3.28)$$

Remark 3.8. Equation (3.25) says that an a posteriori error estimate for the NSE is controlled by that for the Stokes equations with the additional term appearing in the right-hand side of (3.28). Thus, we can apply existing a posteriori error estimates for the pseudostress-velocity formulation of the Stokes equations to the NSE. Indeed, we apply the residual-based a posteriori error estimates for the two-dimensional Stokes equations analyzed by Carstensen, Kim and Park in [10].

Let us describe the well-known results used throughout this paper. Assume that Ω is a bounded Lipschitz domain with a polygonal boundary. Then there exists a constant $C = C(\Omega)$ such that Poincaré's inequalities read

$$\begin{aligned} \|\mathbf{v} - \mathbf{v}_\Omega\|_{0,2} &\leq C\|\nabla \mathbf{v}\|_{0,2} & \text{for all } \mathbf{v} \in H^1(\Omega)^d, \\ \|\mathbf{v}\|_{0,2} &\leq C\|\nabla \mathbf{v}\|_{0,2} & \text{for all } \mathbf{v} \in H_0^1(\Omega)^d, \end{aligned} \quad (3.29)$$

where \mathbf{v}_Ω is an average of \mathbf{v} over Ω , i.e., $\mathbf{v}_\Omega = \frac{1}{|\Omega|} \int_\Omega \mathbf{v}(\mathbf{y}) \, d\mathbf{y}$.

The general interpolation operator cannot be applied to H^1 functions. But Clément's interpolation which defines a continuous interpolating polynomials using averages of \mathbf{v} instead of point values can be applied to H^1 functions. The nodal values of Clément's interpolating polynomials depend on the value of \mathbf{v} on the adjacent elements through an averaging process. We now consider the error estimate of Clément's interpolation operator. We define $S^0(\mathcal{T}_j)^d \subset L^2(\Omega)^d$ as the piecewise constant and $S^1(\mathcal{T}_j)^d \subset H^1(\Omega)^d$ or $S_0^1(\mathcal{T}_j)^d \subset H_0^1(\Omega)^d$ as continuous and piecewise affine vector functions. For each \mathcal{T}_j , let \mathcal{E}_j denote the set of all edges (faces) of \mathcal{T}_j . In what follows, h_T and h_E stand for the diameters of the triangle $T \in \mathcal{T}_j$ and the edge $E \in \mathcal{E}_j$, respectively. For any $T \in \mathcal{T}_j$ and $E \in \mathcal{E}_j$, we set

$$\omega_T := \bigcup \{T' \in \mathcal{T}_j : \bar{T} \cap \bar{T}' \neq \emptyset\} \quad \text{and} \quad \omega_E := \bigcup \{T \in \mathcal{T}_j : E \subset \bar{T}\}$$

Note that, since the triangulation \mathcal{T}_j is shape-regular, the maximal number of elements in ω_T is h -independently bounded by C_M . Clément's interpolation operator (see [11]) $I_j: H^1(\Omega)^d \rightarrow S^1(\mathcal{T}_j)^d$ satisfies the following estimates:

$$\|\mathbf{v} - I_j \mathbf{v}\|_{m,2,T} \lesssim h_T^{1-m} \|\mathbf{v}\|_{1,2,\omega_T} \quad \text{for all } \mathbf{v} \in H^1(\Omega)^d, \, m = 0, 1 \text{ and all } T \in \mathcal{T}_j, \quad (3.30)$$

$$\|\mathbf{v} - I_j \mathbf{v}\|_{0,2,E} \lesssim h_E^{1/2} \|\mathbf{v}\|_{1,2,\omega_E} \quad \text{for all } \mathbf{v} \in H^1(\Omega)^d, \text{ all } E \in \partial T \text{ and all } T \in \mathcal{T}_j. \quad (3.31)$$

Finally, we state the usual inverse inequalities introduced in [25] which we will use in order to derive local lower bounds for the error. We need some modified results whose proof easily follows from the idea of the paper [25], and we omit the proof.

Theorem 3.9. *There exist constants C_1, \dots, C_5 , which only depend on the shape of the elements, such that the following inequalities hold for all polynomials $\mathbf{v}, \mathbf{w} \in P_k(T)^{d^*}$ and $\mathbf{d} \in P_k(E)^{d^*}$, $d^* = d$ or $d \times d$, and for two real numbers p, q with $1 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$:*

$$C_1 \|\mathbf{v}\|_{0,p,T} \leq \sup_{\mathbf{w} \in P_k(T)^{d^*}} \frac{\int_T \mathbf{v} \mathbf{b}_T \mathbf{w} d\mathbf{x}}{\|\mathbf{w}\|_{0,q,T}} \leq \|\mathbf{v}\|_{0,p,T}, \quad (3.32)$$

$$C_2 \|\mathbf{d}\|_{0,2,E} \leq \|\mathbf{b}_E^{1/2} \mathbf{d}\|_{0,2,E} \leq \|\mathbf{d}\|_{0,2,E}, \quad (3.33)$$

$$\|\mathbf{b}_E P_{\text{ext}} \mathbf{d}\|_{0,q,\omega_E} \leq C_3 h_T^{1/q} \|\mathbf{d}\|_{0,q,E}, \quad (3.34)$$

$$\|\nabla(\mathbf{b}_T \mathbf{v})\|_{0,q,T} \leq C_4 h_T^{-1} \|\mathbf{b}_T \mathbf{v}\|_{0,q,T}, \quad (3.35)$$

$$\|\nabla(\mathbf{b}_E P_{\text{ext}} \mathbf{d})\|_{0,q,\omega_E} \leq C_5 h_T^{-1} \|\mathbf{b}_E P_{\text{ext}} \mathbf{d}\|_{0,q,T}. \quad (3.36)$$

Here, for the barycenters \mathbf{x}_T of T and \mathbf{x}_E of E , two functions $\mathbf{b}_T, \mathbf{b}_E \in C^\infty(T, \mathbb{R})$ satisfy $\mathbf{b}_T \in P_3(T)$, $\mathbf{b}_T(\mathbf{x}_T) = 1$ and $\mathbf{b}_T = 0$ on ∂T , and $\mathbf{b}_E \in P_2(T)$, $\mathbf{b}_E(\mathbf{x}_E) = 1$ and $\mathbf{b}_E = 0$ on $\partial T \setminus E$, respectively, and P_{ext} is an extension operator $P_{\text{ext}}: C(E, \mathbb{R}^{d^*}) \rightarrow C(\omega_E, \mathbb{R}^{d^*})$ with $P_{\text{ext}} \mathbf{w}|_E = \mathbf{w}$.

Let $\mathbf{R}_j(\mathbf{f}) := \mathbf{f} + \mathbf{div} \boldsymbol{\sigma}_j - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A} \boldsymbol{\sigma}_j$, $\bar{\mathbf{e}}_j = \bar{\mathbf{u}} - \mathbf{u}_j$ and $\bar{\boldsymbol{\sigma}}_j = \bar{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_j$, and $\boldsymbol{\Sigma} := \hat{H}(\mathbf{div}; \Omega)^d$ and $\mathbf{V} := L^2(\Omega)^d$. For simplicity, we let $\mathbf{P}_j := \mathbf{P}_{h_j}$ be the L^2 -projection onto $P_k(\mathcal{T}_{h_j})$.

Lemma 3.10. *For the error $\bar{\mathbf{e}}_j$ ($1 \leq j \leq J$), we have*

$$\begin{aligned} \|\bar{\mathbf{e}}_j\|_{0,2}^2 &\leq \sum_{T \in \mathcal{T}_j} h_T^2 \|\mathbf{R}_j(\mathbf{f})\|_{0,2,T}^2 + \sum_{T \in \mathcal{T}_j} h_T^2 \|\mathcal{A} \boldsymbol{\sigma}_j - \nabla \mathbf{u}_j\|_{0,2,T}^2 + \sum_{E \in \mathcal{E}_j \cap \partial \Omega} h_E \|\mathbf{g} - \mathbf{g}_j\|_{0,2,E}^2 \\ &\quad + \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega}^2, \end{aligned}$$

where \mathbf{g}_j is an approximation of \mathbf{g} so that $\langle \mathbf{g}_j, \mathbf{n} \cdot (\boldsymbol{\eta} - \tilde{\Pi}_j \boldsymbol{\eta}) \rangle = 0$.

Proof. Since $\bar{\mathbf{e}} \in L^2(\Omega)^d$, we have unique $\boldsymbol{\eta} \in H^1(\Omega)^{d \times d}$ and $\mathbf{z} \in H_0^1(\Omega)^d \cap H^2(\Omega)^d$ such that

$$\begin{aligned} (\mathcal{A} \boldsymbol{\eta}, \boldsymbol{\tau}) + (\mathbf{div} \boldsymbol{\tau}, \mathbf{z}) &= 0 \quad \text{for all } \boldsymbol{\tau} \in \boldsymbol{\Sigma}, \\ (\mathbf{div} \boldsymbol{\eta}, \mathbf{v}) &= (\bar{\mathbf{e}}, \mathbf{v}) \quad \text{for all } \mathbf{v} \in \mathbf{V}, \end{aligned} \quad (3.37)$$

and

$$\|\mathbf{z}\|_{2,2} \leq \|\bar{\mathbf{e}}\|_{0,2}, \quad \|\boldsymbol{\eta}\|_{1,2} \leq \|\bar{\mathbf{e}}\|_{0,2}. \quad (3.38)$$

Then, by integration by parts, (3.26) and the relation $\mathbf{div} \Pi_j = \mathbf{P}_j \mathbf{div}$, we get

$$\|\bar{\mathbf{e}}\|_{0,2}^2 = (\bar{\mathbf{u}} - \mathbf{u}_j, \mathbf{div} \boldsymbol{\eta}) = (-\mathcal{A} \bar{\boldsymbol{\sigma}}, \boldsymbol{\eta}) + \langle \mathbf{g}, \mathbf{n} \cdot \boldsymbol{\eta} \rangle - (\mathbf{u}_j, \mathbf{div}(\Pi_j \boldsymbol{\eta})).$$

Equations (3.23), (3.27), (3.27) and (3.24) imply

$$\begin{aligned} \|\bar{\mathbf{e}}\|_{0,2}^2 &= -(\mathcal{A} \bar{\boldsymbol{\sigma}} - \mathcal{A} \boldsymbol{\sigma}_j, \boldsymbol{\eta}) - (\mathcal{A} \boldsymbol{\sigma}_j, \boldsymbol{\eta}) + (\mathcal{A} \boldsymbol{\sigma}_j, \Pi_j \boldsymbol{\eta}) + \langle \mathbf{g}, \mathbf{n} \cdot (\boldsymbol{\eta} - \Pi_j \boldsymbol{\eta}) \rangle \\ &= (\mathbf{div}(\bar{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_j), \mathbf{z}) + (\mathcal{A} \boldsymbol{\sigma}_j, \Pi_j \boldsymbol{\eta} - \boldsymbol{\eta}) + \langle \mathbf{g}, \mathbf{n} \cdot (\boldsymbol{\eta} - \Pi_j \boldsymbol{\eta}) \rangle \\ &= -(\mathbf{R}_j(\mathbf{f}), \mathbf{z} - \mathbf{P}_j \mathbf{z}) - \frac{1}{\nu} ((\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1}), \mathbf{P}_j \mathbf{z}) + (\mathcal{A} \boldsymbol{\sigma}_j, \Pi_j \boldsymbol{\eta} - \boldsymbol{\eta}) + \langle \mathbf{g} - \mathbf{g}_j, \mathbf{n} \cdot (\boldsymbol{\eta} - \tilde{\Pi}_j \boldsymbol{\eta}) \rangle. \end{aligned} \quad (3.39)$$

The last term of (3.39) is derived by applying the compatibility condition $\int_{\partial \Omega} \mathbf{g} \cdot \mathbf{n} ds = 0$ and relation (3.13) to the term $\langle \mathbf{g}, \mathbf{n} \cdot (\boldsymbol{\eta} - \Pi_j \boldsymbol{\eta}) \rangle$. To estimate the trilinear term $((\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1}), \mathbf{P}_j \mathbf{z})$, using the Hölder inequality and equation (3.14), then we have

$$\begin{aligned} |((\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1}), \mathbf{P}_j \mathbf{z})| &\leq \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega} \|\mathbf{P}_j \mathbf{z}\|_{0,2,\Omega} \\ &\leq \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega} \|\mathbf{z}\|_{1,2,\Omega}. \end{aligned}$$

Applying estimates (3.14), (3.15) and (3.38) to the above equation (3.39), we conclude the proof. \square

To derive the estimate of the error bound for $\bar{\boldsymbol{\varepsilon}} = \bar{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_j$, we formulate the Helmholtz decomposition suitable for our problems. For this, we first define the curl operator for various kinds of functions. When $d = 2$, we define the curl operator for a scalar function q , a vector function $\mathbf{v} = (v_1, v_2)$ and a tensor function $\boldsymbol{\psi} = (\psi_{ij})_{2 \times 2}$, $\boldsymbol{\psi}_i = (\psi_{i1}, \psi_{i2})$ denote its i th row for $i = 1, 2$, by

$$\mathbf{curl} q := \left(\frac{\partial q}{\partial x_2}, -\frac{\partial q}{\partial x_1} \right), \quad \mathbf{curl} \mathbf{v} := -\frac{\partial v_1}{\partial x_2} + \frac{\partial v_2}{\partial x_1}, \quad \mathbf{Curl} \mathbf{v} := \begin{pmatrix} \mathbf{curl} v_1 \\ \mathbf{curl} v_2 \end{pmatrix}, \quad \mathbf{Curl} \boldsymbol{\psi} := (\mathbf{curl} \boldsymbol{\psi}_1, \mathbf{curl} \boldsymbol{\psi}_2).$$

When $d = 3$, we define the curl operator for a vector function $\mathbf{v} = (v_1, v_2, v_3)$ and a tensor function $\boldsymbol{\psi} = (\psi_{ij})_{3 \times 3}$, $\boldsymbol{\psi}_i = (\psi_{i1}, \psi_{i2}, \psi_{i3})$ by

$$\mathbf{curl} \mathbf{v} := \left(-\frac{\partial v_2}{\partial x_3} + \frac{\partial v_3}{\partial x_2}, -\frac{\partial v_3}{\partial x_1} + \frac{\partial v_1}{\partial x_3}, -\frac{\partial v_1}{\partial x_2} + \frac{\partial v_2}{\partial x_1} \right), \quad \mathbf{Curl} \boldsymbol{\psi} := \begin{pmatrix} \mathbf{curl} \boldsymbol{\psi}_1 \\ \mathbf{curl} \boldsymbol{\psi}_2 \\ \mathbf{curl} \boldsymbol{\psi}_3 \end{pmatrix}.$$

We notice that, for $\boldsymbol{\psi} \in H^1(\omega)^{d \times d}$ and $\mathbf{r} \in H^1(\omega)^{d_*}$, $d_* = 2$ for $d = 2$ and $d_* = 3 \times 3$ for $d = 3$, with the Lipschitz boundary $\partial\omega$,

$$(\mathbf{Curl} \boldsymbol{\psi}, \mathbf{r})_\omega - (\boldsymbol{\psi}, \mathbf{Curl} \mathbf{r})_\omega = \langle \boldsymbol{\gamma}_t \boldsymbol{\psi}, \mathbf{r} \rangle_{\partial\omega}, \quad (3.40)$$

where $\boldsymbol{\gamma}_t \boldsymbol{\psi}$ is denoted by $\mathbf{t} \cdot \boldsymbol{\psi}|_{\partial\omega}$ or $(\boldsymbol{\gamma}_t \boldsymbol{\psi})_i = \mathbf{n} \times \boldsymbol{\psi}_i|_{\partial\omega}$ according to whether $d = 2$ or $d = 3$. Here, when $d = 2$, $\mathbf{t} = (t_1, t_2)$ is tangential on $\partial\omega$ and the direction of \mathbf{t} follows that $t_1 = -n_2, t_2 = n_1$ with the unit outer normal vector $\mathbf{n} = (n_1, n_2)$ to $\partial\omega$ and when $d = 3$, $(\boldsymbol{\gamma}_t \boldsymbol{\psi})_i$ is the i th row of the tensor $\boldsymbol{\gamma}_t \boldsymbol{\psi}$.

Lemma 3.11. *For the error $\bar{\boldsymbol{\varepsilon}}$, let $s := \frac{1}{|\Omega|} \int_\Omega \text{tr} \bar{\boldsymbol{\varepsilon}} \, dx$. Then there exist $\mathbf{z} \in H_0^1(\Omega)^d \cap H^2(\Omega)^d$ with $\text{div} \mathbf{z} = \text{tr} \bar{\boldsymbol{\varepsilon}} - s$, $\tilde{p} \in L_0^2(\Omega)$ and $\mathbf{r} \in H^1(\Omega)^{d_*} / \mathbb{R}^{d_*}$, $d_* = 2$ for $d = 2$ and $d_* = 3 \times 3$ for $d = 3$ such that*

$$\bar{\boldsymbol{\varepsilon}} = \nabla \mathbf{z} - \tilde{p} \boldsymbol{\delta} + \mathbf{Curl} \mathbf{r}, \quad (3.41)$$

with the estimate

$$\|\nabla \mathbf{z}\|_{0,2} + |\mathbf{r}|_{1,2} + \|\tilde{p}\|_{0,2} \leq C \|\bar{\boldsymbol{\varepsilon}}\|_{0,2}, \quad (3.42)$$

where C is a positive constant independent of $\boldsymbol{\varepsilon}$.

Proof. We consider the following Stokes problem: find $\mathbf{z} \in H_0^1(\Omega)^d \cap H^2(\Omega)^d$ and $\tilde{p} \in L_0^2(\Omega)$ such that

$$\begin{aligned} (\nabla \mathbf{z}, \nabla \mathbf{w}) - (\tilde{p}, \text{div} \mathbf{w}) &= (\bar{\boldsymbol{\varepsilon}}, \nabla \mathbf{w}) & \text{for all } \mathbf{w} \in H_0^1(\Omega)^d, \\ (q, \text{div} \mathbf{z}) &= (q, \text{tr} \bar{\boldsymbol{\varepsilon}} - s) & \text{for all } q \in L_0^2(\Omega)^d. \end{aligned}$$

The solutions \mathbf{z} and \tilde{p} satisfy

$$-\Delta \mathbf{z} + \nabla \tilde{p} = -\text{div} \bar{\boldsymbol{\varepsilon}} \quad \text{and} \quad \text{div} \mathbf{z} = \text{tr} \bar{\boldsymbol{\varepsilon}} - s.$$

Hence $\bar{\boldsymbol{\varepsilon}} - \nabla \mathbf{z} + \tilde{p} \boldsymbol{\delta}$ is a divergence-free tensor. Moreover, $\bar{\boldsymbol{\varepsilon}} - \nabla \mathbf{z} + \tilde{p} \boldsymbol{\delta}$ satisfies

$$\langle (\bar{\boldsymbol{\varepsilon}} - \nabla \mathbf{z} + \tilde{p} \boldsymbol{\delta})_i \cdot \mathbf{n}, \mathbf{1} \rangle_{\partial\Omega} = 0, \quad i = 1, \dots, d,$$

where $(\bar{\boldsymbol{\varepsilon}} - \nabla \mathbf{z} + \tilde{p} \boldsymbol{\delta})_i$ is the i th row of the tensor $\bar{\boldsymbol{\varepsilon}} - \nabla \mathbf{z} + \tilde{p} \boldsymbol{\delta}$. These conditions imply that there exist stream functions $r_i \in H^1(\Omega)$ for $d = 2$ or $r_i \in H^1(\Omega)^3$ satisfying $\text{div} r_i = 0$ and $\mathbf{curl} r_i \in H(\text{div}; \Omega)$ for $d = 3$ such that $(\bar{\boldsymbol{\varepsilon}} - \nabla \mathbf{z} + \tilde{p} \boldsymbol{\delta})_i = \mathbf{curl} r_i$ (see [13, Theorems 3.1 and 3.4]). Thus, we conclude that, for the two-dimensional vector $\mathbf{r} = (r_1, r_2)$ or for the 3×3 -tensor $\mathbf{r} = (r_1^t, r_2^t, r_3^t)^t$, we have $\bar{\boldsymbol{\varepsilon}} = \nabla \mathbf{z} - \tilde{p} \boldsymbol{\delta} + \mathbf{Curl} \mathbf{r}$. The regularity theorem of the Stokes problem (cf. [13, Theorem 5.4]) yields $\|\nabla \mathbf{z}\|_{0,2} + \|\tilde{p}\|_{0,2} \leq C \|\bar{\boldsymbol{\varepsilon}}\|_{0,2}$, where the positive constant C is independent of $\bar{\boldsymbol{\varepsilon}}$ (see [13]). From this inequality and the fact that $|\mathbf{r}|_{1,2} = \|\mathbf{Curl} \mathbf{r}\|_{0,2}$ for $d = 2$ and $|\mathbf{r}|_{1,2} \leq C(\|\text{div} \mathbf{r}\|_{0,2} + \|\mathbf{Curl} \mathbf{r}\|_{0,2}) = C \|\mathbf{Curl} \mathbf{r}\|_{0,2}$ for $d = 3$, we easily get $|\mathbf{r}|_{1,2} \leq C \|\bar{\boldsymbol{\varepsilon}}\|_{0,2}$. \square

We define the jump $[\cdot]$ for $E \in \mathcal{E}_j$ and for $\mathbf{v} \in H^1(\cup \mathcal{J}_j)^d$ by

$$[\mathbf{v}]_E := (\mathbf{v}|_{T_+})|_E - (\mathbf{v}|_{T_-})|_E \quad \text{if } E \subset \bar{T}_+ \cap \bar{T}_- \quad (E \in \mathcal{E}_j, T_+, T_- \in \mathcal{J}_j)$$

and \mathbf{n}_E points from T_+ into its neighbor element T_- .

Lemma 3.12. For the error $\bar{\mathbf{e}}_j$ ($1 \leq j \leq J$), we have

$$\begin{aligned} \|\bar{\mathbf{e}}\|_{0,2}^2 &\leq \sum_{T \in \mathcal{T}_j} \|h_T \mathbf{R}_j(\mathbf{f})\|_{0,2,T}^2 + \sum_{T \in \mathcal{T}_j} \|h_T \mathbf{Curl}(\mathcal{A}\boldsymbol{\sigma}_j)\|_{0,2,T}^2 + \sum_{E \in \mathcal{E}_j \setminus \partial\Omega} \|h_E^{1/2} [\gamma_t \mathcal{A}\boldsymbol{\sigma}_j]\|_{0,2,E}^2 \\ &\quad + \sum_{E \in \mathcal{E}_j \cap \partial\Omega} \|h_E^{1/2} \gamma_t(\mathcal{A}\boldsymbol{\sigma}_j - \nabla \mathbf{g})\|_{0,2,E}^2 + \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega}^2. \end{aligned}$$

Proof. In the Helmholtz decomposition (3.41) of $\bar{\mathbf{e}}$,

$$\bar{\mathbf{e}} = \nabla \mathbf{z} - \tilde{\rho} \boldsymbol{\delta} + \mathbf{Curl} \mathbf{r},$$

we choose \mathbf{r} satisfying $\int_{\Omega} \mathbf{r} d\mathbf{x} = \mathbf{0}$ in order to apply Poincaré's inequality (3.29) later.

From the above decomposition and the identity $\operatorname{div} \mathbf{z} = \operatorname{tr} \bar{\mathbf{e}} - s$, we get

$$\operatorname{tr} \bar{\mathbf{e}} = \operatorname{div} \mathbf{z} - \operatorname{tr}(\tilde{\rho} \boldsymbol{\delta} - \mathbf{Curl} \mathbf{r}) = \operatorname{tr} \bar{\mathbf{e}} - s - \operatorname{tr}(\tilde{\rho} \boldsymbol{\delta} - \mathbf{Curl} \mathbf{r}).$$

Thus $\operatorname{tr}(\tilde{\rho} \boldsymbol{\delta} - \mathbf{Curl} \mathbf{r}) = -s$. Since $\int_{\Omega} \operatorname{tr} \tilde{\rho} d\mathbf{x} = 0$ and $\int_{\Omega} \operatorname{tr} \boldsymbol{\sigma}_j d\mathbf{x} = 0$, we have

$$(\operatorname{tr} \bar{\mathbf{e}}, \operatorname{tr}(\tilde{\rho} \boldsymbol{\delta} - \mathbf{Curl} \mathbf{r})) = -s(\operatorname{tr} \tilde{\rho} - \operatorname{tr} \boldsymbol{\sigma}_j, 1) = 0. \quad (3.43)$$

From (3.43), we have the error decomposition

$$\begin{aligned} (\bar{\mathbf{e}}, \bar{\mathbf{e}}) &= (\bar{\mathbf{e}}, \nabla \mathbf{z} - \tilde{\rho} \boldsymbol{\delta} + \mathbf{Curl} \mathbf{r}) \\ &= (\bar{\mathbf{e}}, \nabla \mathbf{z}) - (\operatorname{tr} \bar{\mathbf{e}}, \tilde{\rho}) + (\mathcal{A} \bar{\mathbf{e}}, \mathbf{Curl} \mathbf{r}) + \frac{1}{d} ((\operatorname{tr} \bar{\mathbf{e}}) \boldsymbol{\delta}, \mathbf{Curl} \mathbf{r}) \\ &= (\bar{\mathbf{e}}, \nabla \mathbf{z}) + (\mathcal{A} \bar{\mathbf{e}}, \mathbf{Curl} \mathbf{r}) - \frac{1}{d} (\operatorname{tr} \bar{\mathbf{e}}, \operatorname{tr} \tilde{\rho} \boldsymbol{\delta}) + \frac{1}{d} (\operatorname{tr} \bar{\mathbf{e}}, \operatorname{tr} \mathbf{Curl} \mathbf{r}) \\ &= (\bar{\mathbf{e}}, \nabla \mathbf{z}) + (\mathcal{A} \bar{\mathbf{e}}, \mathbf{Curl} \mathbf{r}) = \text{(I)} + \text{(II)}. \end{aligned} \quad (3.44)$$

Integration by parts, (3.24) and (3.27) imply

$$\text{(I)} = (\bar{\mathbf{e}}, \nabla \mathbf{z}) = (\mathbf{R}_j(\mathbf{f}), \mathbf{z} - \mathbf{P}_j \mathbf{z}) + \frac{1}{v} ((\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1}), \mathbf{P}_j \mathbf{z}).$$

Using inequality (3.14), Cauchy–Schwarz inequality and Sobolev imbedding theorem, we have

$$\text{(I)} \leq \|h_j \mathbf{R}_j(\mathbf{f})\|_{0,2,\Omega} \|\nabla \mathbf{z}\|_{0,2,\Omega} + \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega} \|\mathbf{z}\|_{2,2,\Omega}. \quad (3.45)$$

We now consider the second term (II) of (3.44). For the Clement's operator I_j , we have $\operatorname{div} \mathbf{Curl}(I_j \mathbf{r}) = \mathbf{0}$ in the weak sense and $\mathbf{Curl}(I_j \mathbf{r}) \in H(\operatorname{div}; \Omega)^{d \times d}$. For an arbitrary $\mathbf{r} \in H^1(\Omega)^{d \times d}$, $\mathbf{Curl}(I_j \mathbf{r})$ is not necessarily in $\widehat{\mathbf{RT}}_k^d(\mathcal{T}_j)$ since it may not satisfy the constraint $\int_{\Omega} (\operatorname{tr} \mathbf{Curl}(I_j \mathbf{r})) d\mathbf{x} = 0$. But $\mathbf{Curl}(I_j \mathbf{r}) - \tilde{s} \boldsymbol{\delta}$ belongs to the space $\widehat{\mathbf{RT}}_k^d(\mathcal{T}_j)$, where

$$\tilde{s} = \frac{\int_{\Omega} (\operatorname{tr} \mathbf{Curl}(I_j \mathbf{r})) d\mathbf{x}}{d|\Omega|}.$$

We find, from (3.23) and (3.26),

$$(\mathcal{A} \bar{\mathbf{e}}, \mathbf{Curl}(I_j \mathbf{r}) - \tilde{s} \boldsymbol{\delta}) = -(\bar{\mathbf{e}}, \operatorname{div}(\mathbf{Curl}(I_j \mathbf{r}) - \tilde{s} \boldsymbol{\delta})) = 0.$$

Moreover, from the fact that $(\mathcal{A} \bar{\mathbf{e}}, \tilde{s} \boldsymbol{\delta}) = (\operatorname{tr} \mathcal{A} \bar{\mathbf{e}}, \tilde{s}) = 0$, we get $(\mathcal{A} \bar{\mathbf{e}}, \mathbf{Curl}(I_j \mathbf{r})) = 0$. By the identity $\mathcal{A} \tilde{\rho} = \nabla \tilde{\mathbf{u}}$ and integration by parts (3.40), the second term (II) of (3.44) becomes

$$\text{(II)} = (\mathcal{A} \bar{\mathbf{e}}, \mathbf{Curl}(\mathbf{r} - I_j \mathbf{r})) = (-\mathcal{A} \boldsymbol{\sigma}_j, \mathbf{Curl}(\mathbf{r} - I_j \mathbf{r})) - \langle \gamma_t \nabla \tilde{\mathbf{u}}, \mathbf{r} - I_j \mathbf{r} \rangle$$

Considering integration by parts (3.40) and estimate (3.30)–(3.31), recalling that the number of elements in ω_T is bounded by the constant C_M and using Poincaré's inequality (3.29), we obtain

$$\begin{aligned} \text{(II)} &= - \sum_{T \in \mathcal{T}_j} (\mathbf{Curl}(\mathcal{A} \boldsymbol{\sigma}_j), \mathbf{r} - I_j \mathbf{r})_T + \sum_{T \in \mathcal{T}_j} \langle \gamma_t \mathcal{A} \boldsymbol{\sigma}_j, \mathbf{r} - I_j \mathbf{r} \rangle_{\partial T} - \langle \gamma_t \nabla \mathbf{g}, \mathbf{r} - I_j \mathbf{r} \rangle \\ &\leq \left(\left(\sum_{T \in \mathcal{T}_j} \|h_T \mathbf{Curl}(\mathcal{A} \boldsymbol{\sigma}_j)\|_{0,2,T}^2 \right)^{1/2} + \left(\sum_{E \in \mathcal{E}_j \setminus \partial\Omega} \|h_E^{1/2} [\gamma_t \mathcal{A} \boldsymbol{\sigma}_j]\|_{0,2,E}^2 \right)^{1/2} \right. \\ &\quad \left. + \left(\sum_{E \in \mathcal{E}_j \cap \partial\Omega} \|h_E^{1/2} \gamma_t(\mathcal{A} \boldsymbol{\sigma}_j - \nabla \mathbf{g})\|_{0,2,E}^2 \right)^{1/2} \right) |\mathbf{r}|_{1,2}. \end{aligned} \quad (3.46)$$

Putting (3.45) and (3.46) into (3.44) and using estimate (3.42), we conclude the proof. \square

Combining Corollary 2.8 and Lemmas 3.11–3.12, we obtain the following reliability estimate.

Theorem 3.13 (Reliable A Posteriori Error Estimates). *For the solution $\phi = (\boldsymbol{\sigma}, \mathbf{u})$ of problem (3.8)–(3.9) and the approximate solution $\phi_j = (\boldsymbol{\sigma}_j, \mathbf{u}_j)$ of problem (3.23)–(3.24), we have, for $1 \leq j \leq J$,*

$$\|\phi - \phi_j\|_X^2 \leq \eta_j^2 + \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega}^2,$$

where the error estimator η_j is defined by $\eta_j^2 := \sum_{T \in \mathcal{T}_j} \eta_j(T)^2$ with the local error estimator $\eta_j(T)$ given by

$$\begin{aligned} \eta_j(T)^2 := & h_T^2 \|\mathbf{R}_j(\mathbf{f})\|_{0,2,T}^2 + h_T^2 \|\mathcal{A}\boldsymbol{\sigma}_j - \nabla \mathbf{u}_j\|_{0,2,T}^2 + h_T^2 \|\mathbf{Curl}(\mathcal{A}\boldsymbol{\sigma}_j)\|_{0,2,T}^2 \\ & + \sum_{E \in \partial T \setminus \partial \Omega} \frac{h_E}{2} \|\gamma_{t,E} \mathcal{A}\boldsymbol{\sigma}_j\|_{0,2,E}^2 + \sum_{E \in \partial T \cap \partial \Omega} h_E (\|\mathbf{g} - \mathbf{g}_j\|_{0,2,E}^2 + \|\gamma_{t,E}(\mathcal{A}\boldsymbol{\sigma}_j - \nabla \mathbf{g})\|_{0,2,E}^2). \end{aligned}$$

Remark 3.14. We note that physical quantities such as pressure, gradient velocity and stress can be expressed as follows:

$$p = -\frac{\nu}{d} \operatorname{tr} \boldsymbol{\sigma}, \quad \nabla \mathbf{u} = \mathcal{A}\boldsymbol{\sigma}, \quad \tilde{\boldsymbol{\sigma}} = \nu(\boldsymbol{\sigma} + (\nabla \mathbf{u})^t) = \nu(\boldsymbol{\sigma} + (\mathcal{A}\boldsymbol{\sigma})^t) = \nu(\mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\sigma}^t).$$

From these identities, the approximation of pressure, gradient velocity and stress can be defined by

$$p_j := -\frac{\nu}{d} \operatorname{tr} \boldsymbol{\sigma}_j, \quad (\nabla \mathbf{u})_j := \mathcal{A}\boldsymbol{\sigma}_j, \quad \tilde{\boldsymbol{\sigma}}_j := \nu(\mathcal{A}\boldsymbol{\sigma}_j + (\boldsymbol{\sigma}_j)^t).$$

Then the following relations hold:

$$\begin{aligned} \|p - p_j\|_{0,3} &= \frac{\nu}{d} \|\operatorname{tr} \boldsymbol{\sigma} - \operatorname{tr} \boldsymbol{\sigma}_j\|_{0,3} \leq \frac{\nu}{d} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3}, \\ \|\nabla \mathbf{u} - (\nabla \mathbf{u})_j\|_{0,3} &= \|\mathcal{A}\boldsymbol{\sigma} - \mathcal{A}\boldsymbol{\sigma}_j\|_{0,3} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3}, \\ \|\tilde{\boldsymbol{\sigma}} - \tilde{\boldsymbol{\sigma}}_j\|_{0,3} &\leq \nu \|\mathcal{A}\boldsymbol{\sigma} - \mathcal{A}\boldsymbol{\sigma}_j\|_{0,3} + \nu \|\boldsymbol{\sigma}^t - (\boldsymbol{\sigma}_j)^t\|_{0,3} \leq 2\nu \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3}. \end{aligned}$$

Now, we derive an efficient a posteriori error estimate. Let $\mathbf{f}_j \in P_k(T)^d$ be a polynomial approximation of \mathbf{f} .

Lemma 3.15. *For each $T \in \mathcal{T}_j$ ($1 \leq j \leq J$), we have*

$$h_T \|\mathbf{R}_j(\mathbf{f})\|_{0,2,T} \leq h_T \|\mathbf{u} - \mathbf{u}_j\|_{0,6,T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3,T} + h_T \|\mathbf{f} - \mathbf{f}_j\|_{0,2,T}.$$

Proof. Using inverse inequality (3.32), the identity $\operatorname{div} \boldsymbol{\sigma} = \frac{1}{\nu} \mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma} - \mathbf{f}$ and integration by parts, we have

$$\begin{aligned} C_1 \|\mathbf{f}_j + \operatorname{div} \boldsymbol{\sigma}_j - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j\|_{0,2,T}^2 &\leq \left(\mathbf{f}_j + \operatorname{div} \boldsymbol{\sigma}_j - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j, \mathbf{b}_T \mathbf{R}_j(\mathbf{f}_j) \right)_T \\ &= \left(\mathbf{f}_j - \mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_j - \operatorname{div} \boldsymbol{\sigma} - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j + \frac{1}{\nu} \mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma}, \mathbf{b}_T \mathbf{R}_j(\mathbf{f}_j) \right)_T \\ &= \left(\mathbf{f}_j - \mathbf{f} - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j + \frac{1}{\nu} \mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma}, \mathbf{b}_T \mathbf{R}_j(\mathbf{f}_j) \right)_T + (\boldsymbol{\sigma}_j - \boldsymbol{\sigma}, \nabla \mathbf{b}_T \mathbf{R}_j(\mathbf{f}_j))_T. \end{aligned}$$

By the Hölder inequality, inverse inequality (3.32) and (3.35), and the mean value theorem, we get

$$\begin{aligned} C_1 \|\mathbf{f}_j + \operatorname{div} \boldsymbol{\sigma}_j - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j\|_{0,2,T}^2 &\leq \left(\left\| \frac{1}{\nu} \mathbf{u} \cdot \mathcal{A}\boldsymbol{\sigma} - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j \right\|_{0,2,T} + \|\mathbf{f} - \mathbf{f}_j\|_{0,2,T} \right) \|\mathbf{R}_j(\mathbf{f})\|_{0,2,T} + h_T^{-1} C_4 \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3,T} \|\mathbf{b}_T \mathbf{R}_j(\mathbf{f})\|_{0,3/2,T} \\ &\leq (\|\mathbf{u} - \mathbf{u}_j\|_{0,6,T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3,T}) \|\mathbf{R}_j(\mathbf{f})\|_{0,2,T} + (h_T^{-1} C_4 \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_j\|_{0,3,T} + \|\mathbf{f} - \mathbf{f}_j\|_{0,2,T}) \|\mathbf{R}_j(\mathbf{f})\|_{0,2,T} \end{aligned} \quad (3.47)$$

With an application of the triangle inequality to the identity

$$\mathbf{R}_j(\mathbf{f}) = \mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_j - \frac{1}{\nu} \mathbf{u}_j \cdot \mathcal{A}\boldsymbol{\sigma}_j = (\mathbf{f} - \mathbf{f}_j) + \mathbf{R}_j(\mathbf{f}_j),$$

we complete the proof by (3.47). \square

Since $(\mathcal{A}\boldsymbol{\sigma}_j)|_T \in P_{k+1}(T)^{d \times d}$, we can apply inverse inequalities (3.32)–(3.36) to the polynomial function $(\mathcal{A}\boldsymbol{\sigma}_j)|_T$.

Lemma 3.16. For each $T \in \mathcal{T}_j$ ($1 \leq j \leq J$), we have

$$h_T \|\mathcal{A}\sigma_j - \nabla \mathbf{u}_j\|_{0,2,T} \leq C_1^{-2} C_4 (h_T \|\sigma - \sigma_j\|_{0,3,T} + \|\mathbf{u} - \mathbf{u}_j\|_{0,6,T}).$$

Proof. By inequality (3.32) and integration by parts, we have

$$\begin{aligned} C_1^2 \|\mathcal{A}\sigma_j - \nabla \mathbf{u}_j\|_{0,2,T}^2 &\leq (\mathcal{A}\sigma_j - \nabla \mathbf{u}_j, \mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j))_T \\ &= -(\mathcal{A}\sigma - \mathcal{A}\sigma_j, \mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j))_T + (\nabla \mathbf{u} - \nabla \mathbf{u}_j, \mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j))_T \\ &= -(\mathcal{A}\sigma - \mathcal{A}\sigma_j, \mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j))_T + (\mathbf{u} - \mathbf{u}_j, \operatorname{div}(\mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j)))_T \\ &\leq \|\mathcal{A}\sigma - \mathcal{A}\sigma_j\|_{0,3,T} \|\mathcal{A}\sigma_j - \nabla \mathbf{u}_j\|_{0,3/2,T} + \|\mathbf{u} - \mathbf{u}_j\|_{0,6,T} |\mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j)|_{1,6/5,T}. \end{aligned} \quad (3.48)$$

Applying inverse inequality (3.35),

$$|\mathbf{b}_T(\mathcal{A}\sigma_j - \nabla \mathbf{u}_j)|_{1,6/5,T} \leq C_4 h_T^{-1} \|\mathcal{A}\sigma_j - \nabla \mathbf{u}_j\|_{0,6/5,T}. \quad (3.49)$$

Putting (3.49) into (3.48) and using the relation $\|\mathcal{A}\sigma - \mathcal{A}\sigma_j\|_{0,3,T} \leq \|\sigma - \sigma_j\|_{0,3,T}$, we get

$$h_T \|\mathcal{A}\sigma_j - \nabla \mathbf{u}_j\|_{0,2,T} \leq C_1^{-2} C_4 (h_T \|\sigma - \sigma_j\|_{0,3,T} + \|\mathbf{u} - \mathbf{u}_j\|_{0,6,T}). \quad \square$$

Lemma 3.17. For each $T \in \mathcal{T}_j$ ($1 \leq j \leq J$), we have

$$h_T \|\operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,2,T} \leq C_1^{-2} C_4 \|\sigma - \sigma_j\|_{0,3,T}.$$

Proof. Inequality (3.32), the identity $\operatorname{Curl}(\mathcal{A}\sigma) = \operatorname{Curl}(\nabla \mathbf{u}) = \mathbf{0}$ and integration by parts imply

$$\begin{aligned} C_1^2 \|\operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,2,T}^2 &\leq \|\mathbf{b}_T^{1/2} \operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,2,T}^2 \\ &= -(\operatorname{Curl}(\mathcal{A}\sigma - \mathcal{A}\sigma_j), \mathbf{b}_T \operatorname{Curl}(\mathcal{A}\sigma_j))_T \\ &= -(\mathcal{A}\sigma - \mathcal{A}\sigma_j, \operatorname{Curl}(\mathbf{b}_T \operatorname{Curl}(\mathcal{A}\sigma_j)))_T \\ &\leq \|\sigma - \sigma_j\|_{0,3,T} |\mathbf{b}_T \operatorname{Curl}(\mathcal{A}\sigma_j)|_{1,3/2,T}. \end{aligned} \quad (3.50)$$

We obtain, from inverse inequality (3.35),

$$|\mathbf{b}_T \operatorname{Curl}(\mathcal{A}\sigma_j)|_{1,3/2,T} \leq C_4 h_T^{-1} \|\operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,3/2,T} \leq C_4 h_T^{-1} \|\operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,2,T}. \quad (3.51)$$

Hence, combining (3.50) and (3.51), we conclude the proof. \square

Lemma 3.18. For each $E \in \mathcal{E}_j$ ($1 \leq j \leq J$),

$$h_E^{1/2} \|\gamma_t \mathcal{A}\sigma_j\|_{0,2,E} \leq C \|\sigma - \sigma_j\|_{0,3,\omega_E},$$

where $C = C_2^{-2} C_3 \max\{C_5, C_1^{-2} C_4\}$.

Proof. Set $\mathbf{d} := [\gamma_t \mathcal{A}\sigma_j]$, which is a polynomial along E . Note that $[\gamma_t \mathcal{A}\sigma_j] = -[\gamma_t \mathcal{A}(\sigma - \sigma_j)]$ on each edge E in the weak sense and $P_{\text{ext}} \mathbf{d}|_E = \mathbf{d}$. Using inequality (3.33), integration by parts (3.40) and the identity $\operatorname{Curl}(\mathcal{A}\sigma) = \mathbf{0}$, we have, for any edge $E \notin \Gamma$ and for $T_i \in \omega_E$, $i = 1, 2$,

$$\begin{aligned} C_2^2 \|\mathbf{d}\|_{0,2,E}^2 &\leq \|\mathbf{b}_E^{1/2} \mathbf{d}\|_{0,2,E}^2 = -(\mathbf{b}_E P_{\text{ext}} \mathbf{d}, [\gamma_t \mathcal{A}(\sigma - \sigma_j)])_E \\ &= (\mathcal{A}(\sigma - \sigma_j), \operatorname{Curl}(\mathbf{b}_E P_{\text{ext}} \mathbf{d}))_{\omega_E} - \sum_{i=1,2} (\mathbf{b}_E P_{\text{ext}} \mathbf{d}, \operatorname{Curl}(\mathcal{A}\sigma_j))_{T_i} \\ &\leq \|\sigma - \sigma_j\|_{0,3,\omega_E} |\mathbf{b}_E P_{\text{ext}} \mathbf{d}|_{1,3/2,\omega_E} + \|\mathbf{b}_E P_{\text{ext}} \mathbf{d}\|_{0,2,\omega_E} \left(\sum_{i=1,2} \|\operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,2,T_i} \right). \end{aligned} \quad (3.52)$$

From Lemma 3.17, we have

$$\sum_{i=1,2} \|\operatorname{Curl}(\mathcal{A}\sigma_j)\|_{0,2,T_i} \leq C_1^{-2} C_4 \sum_{i=1,2} h_{T_i}^{-1} \|\sigma - \sigma_j\|_{0,3,T_i}. \quad (3.53)$$

Application of (3.34), (3.36) and (3.53) to (3.52) yield

$$h_E^{1/2} \|\mathbf{d}\|_{0,2,E} \leq C \|\sigma - \sigma_j\|_{0,3,\omega_E},$$

where $C = C_2^{-2} C_3 \max\{C_5, C_1^{-2} C_4\}$. \square

Remark 3.19. By the Hölder inequality and the triangle inequality, we have

$$\begin{aligned} \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega} &\leq \|\mathbf{u}_j - \mathbf{u}_{j-1}\|_{0,6,\Omega} \|\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1}\|_{0,3,\Omega} \\ &\leq (\|\mathbf{u}_j - \mathbf{u}\|_{0,6,\Omega} + \|\mathbf{u} - \mathbf{u}_{j-1}\|_{0,6,\Omega})(\|\boldsymbol{\sigma}_j - \boldsymbol{\sigma}\|_{0,3,\Omega} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{j-1}\|_{0,3,\Omega}). \end{aligned}$$

Assume that $h_{j-1}^2 \leq h_j$ for $j > 0$. Then, by the result of Theorem 3.7,

$$\|\mathbf{u}_j - \mathbf{u}\|_{0,6,\Omega} + \|\mathbf{u} - \mathbf{u}_{j-1}\|_{0,6,\Omega} \leq (h_{j-1}^{2r} + h_j^r) + (h_{j-2}^{2r} + h_{j-1}^r) \leq h_j^r + h_{j-1}^r \leq h_{j-1}^r.$$

Similarly,

$$\|\boldsymbol{\sigma}_j - \boldsymbol{\sigma}\|_{0,3,\Omega} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_{j-1}\|_{0,3,\Omega} \leq h_{j-1}^r.$$

Hence

$$\|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega} = \mathcal{O}(h_{j-1}^{2r}),$$

which shows the asymptotic behavior of the additional term. Note that the additional term

$$\|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot \mathcal{A}(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,2,\Omega} = \mathcal{O}(h_{j-1}^{2r})$$

is of higher order compared to the other error estimates and the error $\|\mathbf{u}_j - \mathbf{u}\|_{0,6,\Omega} + \|\boldsymbol{\sigma}_j - \boldsymbol{\sigma}\|_{0,3,\Omega} = \mathcal{O}(h_j^r)$ if we take $h_j = \mathcal{O}(h_{j-1}^{2-\epsilon})$ for some positive ϵ . In practice, adaptively refining mesh via newest vertex bisection or red refinement satisfies the relation $h_{j-1}^2 \ll \frac{1}{4}h_{j-1} \leq h_j$, which implies that the additional term can be negligible in the adaptive algorithm in contrast to the result in [14], where this additional term is not negligible as refinements proceed. The reason is that, in the second step of his two-grid algorithm, the nonlinear problem is linearized locally at the solution obtained in the fixed coarse space. When the linearized problem in the second step is solved in a fine mesh, as the mesh refinement proceeds, the additional term becomes relatively large. This issue can be resolved by using our multi-level algorithm, where the nonlinear problem is linearized locally at the solution obtained on the very previous mesh.

Combining Lemmas 3.15–3.18, we arrive at the following efficiency estimates.

Theorem 3.20 (Efficient A Posteriori Error Estimate). *For the solution $\phi = (\boldsymbol{\sigma}, \mathbf{u})$ of problem (3.8)–(3.9) and the approximate solution $\phi_j = (\boldsymbol{\sigma}_j, \mathbf{u}_j)$ to problem (3.23)–(3.24), we have*

$$\eta_j \leq \|\phi - \phi_j\|_X + \text{osc}_k(\mathbf{f}, \mathcal{T}_j),$$

where

$$\text{osc}_k(\mathbf{f}, \mathcal{T}_j)^2 := \sum_{T \in \mathcal{T}_j} \min_{\mathbf{f}_j|_T \in P_k(T)^d} h_T^2 \|\mathbf{f} - \mathbf{f}_j\|_{0,T}^2.$$

Note that the derived a posteriori error estimator is efficient and reliable, i.e., the efficiency index

$$\text{EI} := \eta_j / \|\phi - \phi_j\|_X$$

and its reciprocal are bounded for all triangulations.

4 Application 2: Semilinear Elliptic Equations

We consider the semilinear boundary value problems of the form

$$\begin{aligned} -\Delta u + G(\lambda, u) &= 0 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{4.1}$$

where $G: \mathcal{D}(G) \subset \mathbb{R} \times H_0^1(\Omega) \rightarrow L^2(\Omega)$ is a given mapping. This problem has been treated in [7] for $d = 2$. Multiplying by test functions and integrating by parts, we obtain the following weak formulation of problem (4.1): find $u \in H_0^1(\Omega)$ such that

$$\int_{\Omega} \nabla u \nabla v \, dx + \int_{\Omega} G(\lambda, u) v \, dx = 0 \quad \text{for all } v \in H_0^1(\Omega). \tag{4.2}$$

4.1 BRR Framework for SEE

We define the continuous linear operator $S: f \in L^2(\Omega) \rightarrow Sf = w \in H_0^1(\Omega) \cap H^2(\Omega)$ which maps f onto the solution w of the following problem: given a function $f \in L^2(\Omega)$, find the unique solution $w \in H_0^1(\Omega) \cap H^2(\Omega)$ such that

$$\int_{\Omega} \nabla w \nabla v \, dx = \int_{\Omega} f v \, dx \quad \text{for all } v \in H_0^1(\Omega).$$

Introducing the operator S , problem (4.2) can be written in the equivalent way

$$F(\lambda, u) = u + SG(\lambda, u) = 0. \quad (4.3)$$

In order to approximate the solution of problem (4.3), given a finite-dimensional space $X_h \subset H_0^1(\Omega)$, we define the linear operator $S_h: f \in L^2(\Omega) \rightarrow S_h f = w_h \in X_h$ such that

$$\int_{\Omega} \nabla w_h \nabla v_h \, dx = \int_{\Omega} f v_h \, dx \quad \text{for all } v_h \in X_h.$$

Then the approximation of nonlinear problem (4.3) is to find $(\lambda, u^h) \in (\mathbb{R} \times X_h) \cap \mathcal{D}(G)$ such that

$$F_h(\lambda, u^h) = u^h + S_h G(\lambda, u^h) = 0. \quad (4.4)$$

We analyze problem (4.1) with $G(\lambda, u) := G(u) = u^\alpha - f$ for $\alpha = 2$ or 3 and a given function $f \in L^2(\Omega)$, which was analyzed in [20] for $\alpha = 3$. The following details are only given for the reader's convenience. In order to apply Theorem 2.1, we take $X = \mathcal{X} := H_0^1(\Omega)$ and $Y = \mathcal{Z} := L^2(\Omega)$. Since the mapping G is C^∞ and D^2G is bounded on all bounded subsets of $X := H_0^1(\Omega)$, hypothesis (H1) holds. And since $H^1(\Omega)$ is continuously imbedded in $L^p(\Omega)$ with $1 \leq p \leq 6$ for $d \leq 3$ by Sobolev imbedding theorem, the function u^α belongs to $L^2(\Omega)$, i.e., $G(u) \in Y$ and $DG(u) \in \mathcal{L}(X; Z)$ for all $u \in X$. Hence, hypothesis (H2) holds. Let $X_h \subset H_0^1(\Omega)$ be a conforming finite element space of piecewise polynomials of degree $k \geq 1$ defined on the triangulation \mathcal{T}_h . It is well known that, for $f^* \in H^r(\Omega)$ with $2 \leq r \leq k + 1$,

$$\|(S - S_h)f^*\|_{s,2} \leq Ch^{r-s} \|Sf^*\|_{r,2} \leq Ch^{r-s} \|f^*\|_{r-2,2}, \quad s = 0, 1, \quad (4.5)$$

where $\|\cdot\|_{r,2} := \|\cdot\|_{H^r(\Omega)}$ and $\|\cdot\|_{0,2} := \|\cdot\|_{L^2(\Omega)}$, which yields that hypotheses (H3)–(H4) hold. From [7, Theorem 3.1], we have a unique nonsingular solution $u \in H_0^1(\Omega) \cap H^2(\Omega)$ of problem (4.1). From Theorem 2.1 and (4.5), there exists a unique nonsingular solution u^h of (4.4) such that $\|u - u^h\|_{1,2} \leq C\|(S - S_h)G(u)\|_{1,2} \leq Ch\|u\|_{2,2}$. We notice that the error bound $\|u - u^h\|_{0,2} \leq Ch^2\|u\|_{2,2}$ was also derived in [20].

4.2 Error Estimates for Multi-level Algorithm for SEE

We generate a shape-regular triangulation \mathcal{T}_{j+1} from \mathcal{T}_j ($j \geq 1$) with newest vertex bisection or longest edge bisection. For simplicity, we let $X_j := X_{h_j}$. Then, for problem (4.1) with $G(\lambda, u) = u^\alpha - f$, the multi-level algorithm reads as follows.

Step 1 (Nonlinear solver) Find $u_0 \in X_0$ such that

$$\int_{\Omega} \nabla u_0 \nabla v \, dx + \int_{\Omega} u_0^\alpha v \, dx = \int_{\Omega} f v \, dx \quad \text{for all } v \in X_0. \quad (4.6)$$

Step 2 (Linear solver) For $1 \leq j \leq J$, find $u_j \in X_j$ such that

$$\int_{\Omega} \nabla u_j \nabla v \, dx + \alpha \int_{\Omega} u_{j-1}^{\alpha-1} u_j v \, dx = \int_{\Omega} ((\alpha - 1)u_{j-1}^\alpha v + f v) \, dx \quad \text{for all } v \in X_j. \quad (4.7)$$

Theorem 4.1 (A Priori Error Estimates). *Let $f \in Y$. Let u be a solution of (4.1), u_0 the solution of (4.6), and u_j the solution of (4.7). Assume that $u \in H^r(\Omega)$ for $2 \leq r \leq k+1$. Then we have, for a constant $C > 0$ that depends on $\|u\|_{2,2}$ and independent of mesh sizes,*

$$\|u - u_j\|_{s,2} \leq C(h_{j-1}^{2(r-s)} + h_j^{r-s}), \quad s = 0, 1.$$

Proof. See [20, 22]. □

Remark 4.2. By referring to [20], we can prove that the result of (2.5) holds for $X = L^2(\Omega)$ as well as $X = H_0^1(\Omega)$. Moreover, we see that if the exact solution u is in $W_p^2(\Omega)$ with $p > 2$ for $d = 2$ and $p = \frac{12}{5}$ for $d = 3$, then the following result holds:

$$\|u^{h_j} - u_j\|_{1,2} \leq C\|u^{h_j} - u_{j-1}\|_{0,p}^2 (\leq Ch_{j-1}^4), \quad (4.8)$$

which shows improvement $\mathcal{O}(h_{j-1}^4)$ compared with the result $\mathcal{O}(h_{j-1}^2)$ via (2.5).

From Theorem 2.8, we observe that

$$\|F(u_j)\|_X = \|u_j + SG(u_j)\|_X = \|u_j - \bar{u}\|_X,$$

where $\bar{u} \in H_0^1(\Omega) \cap H^2(\Omega)$ is the unique solution for the Galerkin formulation of the linear elliptic equations

$$(\nabla \bar{u}, \nabla v) = (f - u_j^\alpha, v) \quad \text{for all } v \in X. \quad (4.9)$$

The following error equations immediately follow from (4.7)–(4.9). For $w_j := u_j^\alpha - au_j u_{j-1}^{\alpha-1} + (\alpha - 1)u_{j-1}^\alpha$,

$$(\nabla \bar{u} - \nabla u_j, \nabla v) = -(w_j, v) \quad \text{for all } v \in X_j. \quad (4.10)$$

Let $R_j(f) := f + \Delta u_j - u_j^\alpha$ and $\bar{e} = \bar{u} - u_j$.

Theorem 4.3 (Reliable A Posteriori Error Estimates). *For the error \bar{e} , we have $|\bar{e}|_{1,2}^2 \leq \eta_j^2 + A_j^2$, where the error estimator η_j is defined by $\eta_j^2 := \sum_{T \in \mathcal{T}_j} \eta_j(T)^2$ with the local error estimator $\eta_j(T)$ given by*

$$\eta_j(T)^2 := h_T^2 \|R_j(f)\|_{0,2,T}^2 + \sum_{E \in \partial T \cap \partial \Omega} \frac{h_E}{2} \left\| \left[\frac{\partial u_j}{\partial n} \right] \right\|_{0,2,E}^2$$

and $A_j = \|u_j - u_{j-1}\|_{0,3,\Omega}^2$ for $\alpha = 2$ and $A_j = \|u_j - u_{j-1}\|_{0,4,\Omega}^2 \|u_j + 2u_{j-1}\|_{0,3,\Omega}$ for $\alpha = 3$.

Proof. By integration by parts and (4.10), we get

$$\begin{aligned} |\bar{e}|_{1,2}^2 &= (\nabla \bar{u} - \nabla u_j, \nabla \bar{e} - \nabla P_j \bar{e}) + (\nabla \bar{u} - \nabla u_j, \nabla P_j \bar{e}) \\ &= - \sum_{T \in \mathcal{T}_j} \left((\Delta \bar{u} - \Delta u_j, \bar{e} - P_j \bar{e})_T - \frac{1}{2} \sum_{E \in \partial T \cap \partial \Omega} \left\langle \left[\frac{\partial u_j}{\partial n} \right], \bar{e} - P_j \bar{e} \right\rangle_E \right) - (w_j, P_j \bar{e}), \end{aligned}$$

where P_j is L^2 -projection onto X_j . Using the error estimates of L^2 -projection, we conclude the proof. In fact, to estimate the term $(w_j, P_j \bar{e})$, we use Hölder inequality and Sobolev imbedding theorem. Then we have, for $\alpha = 2$,

$$|(w_j, P_j \bar{e})| \leq \|(u_j - u_{j-1})^2\|_{0,3/2} \|P_j \bar{e}\|_{0,3} \leq \|u_j - u_{j-1}\|_{0,3}^2 |\bar{e}|_{1,2},$$

and for $\alpha = 3$,

$$|(w_j, P_j \bar{e})| \leq \|(u_j - u_{j-1})^2\|_{0,2} \|u_j + 2u_{j-1}\|_{0,3} \|P_j \bar{e}\|_{0,6} \leq \|u_j - u_{j-1}\|_{0,4}^2 \|u_j + 2u_{j-1}\|_{0,3} |\bar{e}|_{1,2}. \quad \square$$

Remark 4.4. Assume that $h_{j-1}^2 \leq h_j$ for $j > 0$. Then, by the triangle inequality, the Sobolev imbedding theorem, which is that $H^1(\Omega)$ is continuously imbedded in $L^p(\Omega)$ for $p \in [1, 6]$, and Theorem 4.1, we have, for $q = 3$ or 4 ,

$$\begin{aligned} \|u_j - u_{j-1}\|_{0,q}^2 &\leq \|u_j - u\|_{0,q}^2 + \|u - u_{j-1}\|_{0,q}^2 \leq \|u_j - u\|_{1,2}^2 + \|u - u_{j-1}\|_{1,2}^2 \\ &\leq (h_{j-1}^{2(r-1)} + h_j^{r-1})^2 + (h_{j-2}^{2(r-1)} + h_{j-1}^{r-1})^2 = \mathcal{O}(h_{j-1}^{2(r-1)}), \end{aligned}$$

and

$$\|u_j + 2u_{j-1}\|_{0,3} \leq \|u_j + 2u_{j-1} - 3u\|_{1,2} + 3\|u\|_{1,2} \leq C,$$

where a constant C depends on $\|u\|_{2,2}$ and is independent of mesh sizes. As in the case of the Navier–Stokes equations, the additional term w_j is of higher order $\mathcal{O}(h_{j-1}^{2(r-1)})$ compared to the other error estimates and the error $\|u_j - u\|_{1,2} = \mathcal{O}(h_j^{r-1})$ if we take $h_j = \mathcal{O}(h_{j-1}^{2-\epsilon})$ for some positive ϵ . Efficient a posteriori error estimates can be derived in a way similar to those addressed in the NSE using the usual inverse inequalities introduced in [25].

5 Adaptive Multi-level Algorithm

We introduce an adaptive multi-level algorithm for nonlinear problems based on the local error estimator $\eta_j(T)$ and the global error estimator η_j driven in Section 3 and 4. The adaptive procedure for the nonlinear problem goes as follows: find ϕ_0 from Step 1 of Algorithm 1 on a given initial (uniform) mesh \mathcal{T}_0 , form a mesh \mathcal{T}_1 uniformly refined from \mathcal{T}_0 , and generate an adaptively refined mesh \mathcal{T}_{j+1} from \mathcal{T}_j ($j \geq 1$) by repeating the following process:

SOLVE \rightarrow ESTIMATE \rightarrow MARK \rightarrow REFINES

The step SOLVE means to perform Step 2 of Algorithm 1 on the mesh \mathcal{T}_j ($j \geq 1$). This process is iterated until the global error estimator η_j becomes smaller than a given tolerance. We organize the adaptive multi-level algorithm for the nonlinear problems.

Input: Initial (uniform) mesh \mathcal{T}_0 , parameter θ , and tolerance tol.

Output: Final mesh \mathcal{T}_J and approximate solution ϕ_J for $J \geq 1$.

Find ϕ_0 from Step 1 of Algorithm 1 on \mathcal{T}_0 .

Form a mesh \mathcal{T}_1 uniformly refined from \mathcal{T}_0 .

Obtain ϕ_1 from Step 2 of Algorithm 1 on \mathcal{T}_1 .

Compute the local error estimator $\eta_1(T)$ and global error estimator η_1 .

$j := 1$.

while $\eta_j \geq \text{tol}$ **do**

Mark a set $\mathcal{M}_{\mathcal{T}_j} \subset \mathcal{T}_j$ such that $\sum_{T \in \mathcal{M}_{\mathcal{T}_j}} \eta_j(T)^2 \geq \theta \eta_j^2$ (bulk strategy).

$j := j + 1$.

Generate a conforming triangulations \mathcal{T}_j from \mathcal{T}_{j-1} by refining $T \in \mathcal{M}_{\mathcal{T}_{j-1}}$.

Obtain ϕ_j from Step 2 of Algorithm 1 on \mathcal{T}_j .

Compute the error estimators $\eta_j(T)$ and η_j .

Algorithm 4: Adaptive multi-level algorithm

For the reader’s convenience, we present Figure 1 in order to explain the outline of the (uniform/adaptive) multi-level algorithm. The details of the uniform multi-level algorithm can be found in [22]. Figure 1 depicts the multi-level algorithm compared with the usual nonlinear solver (e.g., Newton–Raphson algorithm). The nonlinear solver finds discrete nonlinear solutions by iteration on a much higher-dimensional fixed space X_j than X_0 . The multi-level algorithm finds the solution ϕ_0 from the nonlinear solver on the low-dimensional space X_0 and then repeatedly finds ϕ_j by only one iteration on X_j with starting value ϕ_{j-1} for $j = 1, \dots, J$. The notations

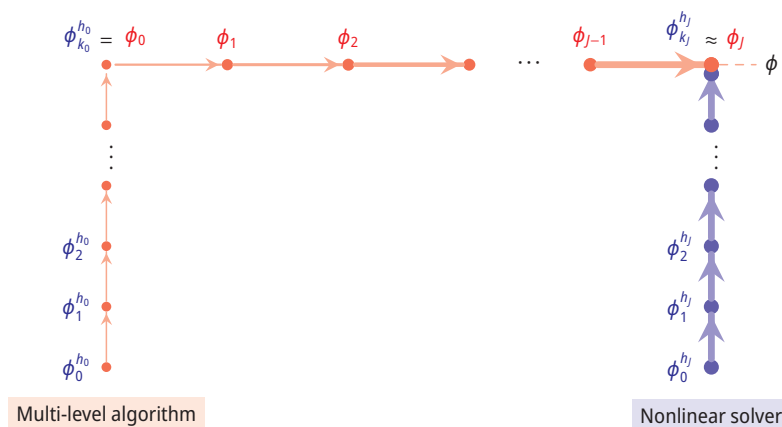


Figure 1: Multi-level process computing $\phi_j \in X_j \subset X, j = 0, 1, \dots, J$, with J finest level

$\{\phi_k^{h_0}\}_{k=0}^{k_0}$ and $\{\phi_k^{h_j}\}_{k=0}^{k_j}$ denote the iteration sequence obtained from the nonlinear solver on the space X_0 and X_j , respectively. The thickness of the arrows indicates the computational costs at the given level; thicker means higher costs.

6 Numerical Experiments

In this section, we perform various numerical experiments to test our adaptive multi-level algorithm and to validate the theory analyzed in this paper. We present the convergence of various errors and indicators on both uniform and adaptive meshes. In particular, we will numerically verify that the adaptive multi-level algorithm has the quadratic convergence in the sense of (2.5).

6.1 Semilinear Elliptic Equations

We approximate the semilinear elliptic problem (4.1) with $G(\lambda, u) = u^\alpha - f$, $\alpha = 2, 3$, by using a conforming finite element space of piecewise linear polynomials. We introduce the following quantities: for $1 \leq j \leq J$,

$$\begin{aligned} E_{j,1}(T) &:= h_T \|\mathbb{R}_j(f)\|_{0,2,T}, \\ E_{j,2}(T) &:= \left(\frac{1}{2} \sum_{E \in \partial T \setminus \partial \Omega} h_E \|\nabla u_j \cdot \mathbf{n}\|_{0,2,E}^2 \right)^{1/2}, \\ E_{j,3}(T) &:= \begin{cases} \|u_j - u_{j-1}\|_{0,3,T}^2, & \alpha = 2, \\ \|u_j - u_{j-1}\|_{0,4,T}^2 \|u_j + 2u_{j-1}\|_{0,3,T}, & \alpha = 3. \end{cases} \end{aligned}$$

We define the local error estimator $\eta_j(T)$ by

$$\eta_j(T) := (E_{j,1}(T)^2 + E_{j,2}(T)^2 + E_{j,3}(T)^2)^{1/2},$$

the global error estimator η_j by

$$\eta_j := \left(\sum_{T \in \mathcal{T}_j} \eta_j(T)^2 \right)^{1/2},$$

and efficiency indices EI_j by

$$\text{EI}_j := \frac{\eta_j}{|u - u_j|_{1,2}}.$$

To show the performance, we consider two test examples: one with a boundary layer, the other with a corner singularity.

Example 6.1 (A Solution with a Boundary Layer). We consider SEE (4.1) with

$$G(\lambda, u) = u^\alpha - f, \quad \alpha = 2, 3, \quad \text{and} \quad \Omega = (0, 1) \times (0, 1).$$

The exact solution u ,

$$u(x, y) = \frac{1}{2}xy(1-x)(1-y)e^{10(x+y-1)},$$

specifies data f .

In Table 1, we present convergence history of multi-level algorithm and nonlinear solver on the uniform meshes. Here, N_{el} means the number of elements at \mathcal{T}_j . The solution obtained from the multi-level algorithm is as accurate as the nonlinear solver at each level j . But we observe that various errors up to low levels ($j \leq 2$) exhibit non-optimal convergence behavior since meshes in low levels are in the pre-asymptotic range. We see that the errors have optimal order of convergence by carrying out more uniform refinement ($j \geq 3$). However, in this uniform refinement, for high levels $j \geq 3$, the number of elements at \mathcal{T}_j increases rapidly and the finite element space X_j requires too many degrees of freedom, which necessitates the adaptive multi-level algorithm.

α	j	$N_{el}(1/h_j)$	Multi-level algorithm			Nonlinear solver				
			$\ u - u_j\ _{0,2}$	CO	$ u - u_j _{1,2}$	CO	$\ u - u^{h_j}\ _{0,2}$	CO	$ u - u^{h_j} _{1,2}$	CO
2	0	256(8)	6.6901e-1		1.9446e+1		6.6901e-1		1.9446e+1	
	1	1024(16)	2.0875e-1	1.68	1.1322e+1	0.78	2.0263e-1	1.72	1.1314e+1	0.78
	2	4096(32)	5.3730e-2	1.96	5.8497e+0	0.95	5.3255e-2	1.93	5.8494e+0	0.95
	3	16384(64)	1.3509e-2	1.99	2.9470e+0	0.99	1.3482e-2	1.98	2.9470e+0	0.99
	4	65536(128)	3.3828e-3	2.00	1.4762e+0	1.00	3.3812e-3	2.00	1.4762e+0	1.00
5	262144 (256)	8.4606e-4	2.00	7.3843e-1	1.00	8.4596e-4	2.00	7.3843e-1	1.00	
3	0	256(8)	6.1077e-1		1.9426e+1		6.1077e-1		1.9426e+1	
	1	1024(16)	2.1849e-1	1.48	1.1368e+1	0.77	1.8097e-1	1.75	1.1307e+1	0.78
	2	4096(32)	5.1247e-2	2.09	5.8541e+0	0.96	4.7263e-2	1.94	5.8500e+0	0.95
	3	16384(64)	1.2075e-2	2.09	2.9472e+0	0.99	1.1936e-2	1.99	2.9471e+0	0.99
	4	65536(128)	2.9973e-3	2.01	1.4762e+0	1.00	2.9917e-3	2.00	1.4762e+0	1.00
5	262144 (256)	7.4873e-4	2.00	7.3844e-1	1.00	7.4840e-4	2.00	7.3844e-1	1.00	

Table 1: Convergence orders (CO) of various errors on uniform meshes for Example 6.1

$\alpha = 2$					$\alpha = 3$				
j	N_{el}	$ u - u_j _{1,2}$	η_j	EI_j	j	N_{el}	$ u - u_j _{1,2}$	η_j	EI_j
0	256	1.9446e+1			0	256	1.9426e+1		
1	1024	1.1320e+1	5.7659e+1	5.09	1	1024	1.1368e+1	5.8250e+1	5.12
2	1076	7.5222e+0	3.8374e+1	5.10	2	1076	7.5233e+0	3.8552e+1	5.12
3	1260	5.1540e+0	2.5581e+1	4.96	3	1257	5.2785e+0	2.6042e+1	4.93
		\vdots	\vdots	\vdots			\vdots	\vdots	\vdots
9	20447	7.4173e-1	3.6179e+0	4.88	9	19876	7.4923e-1	3.6707e+0	4.90
10	36448	5.3738e-1	2.6609e+0	4.95	10	35236	5.4581e-1	2.7078e+0	4.96

Table 2: Efficiency indices on adaptive meshes for Example 6.1

In Table 2 and Figures 2 and 3, we present numerical results from the adaptive multi-level algorithm. Efficiency indices $EI_j = \eta_j/|u - u_j|_{1,2}$ are presented in Table 2. The history of errors $|u - u_j|_{1,2}$ and η_j , and the efficiency indices EI_j are plotted with a logarithmic scale in Figure 2, where the x -axis represents the number of elements. Numerical data for EI_j show that the a posteriori error estimate is numerically reliable and efficient. The superiority of the adaptive multi-level algorithm can be clearly seen in Tables 1 and 2. Indeed, e.g., when $\alpha = 2$, the accuracy on uniform mesh with $N_{el} = 262,144$ in Table 1 is comparable to the accuracy on adaptive mesh with $N_{el} = 36,448$ in Table 2 (see boldface in Tables 1 and 2). Figure 3 shows convergence orders of the indicators. The symbols in the legend of Figure 3 represent the indicators such as

$$E_{j,k} := \left(\sum_{T \in \mathcal{T}_j} E_{j,k}(T)^2 \right)^{1/2}, \quad k = 1, 2, 3.$$

In Figure 3, we observe that the additional term $E_{j,3}$ is of higher order than the other indicators for both uniform and adaptive meshes as mentioned in Remark 4.4.

To examine quadratic convergence behavior in the sense of (2.5) for $X = L^2(\Omega)$ and in the sense of (4.8), we investigate the difference between u_{j-1} and u^{h_j} , and the difference between u_j and u^{h_j} in the L^p or H^1 -norm, and their ratios. In Figure 4, we present these quantities and the ratios RL_j and RH_j , where

$$dL_{j-1} := \|u^{h_j} - u_{j-1}\|_{0,2}, \quad dL_j := \|u^{h_j} - u_j\|_{0,2}, \quad RL_j := \frac{\|u^{h_j} - u_j\|_{0,2}}{\|u^{h_j} - u_{j-1}\|_{0,2}^2},$$

$$dH_{j-1} := \|u^{h_j} - u_{j-1}\|_{0,4}, \quad dH_j := \|u^{h_j} - u_j\|_{1,2}, \quad RH_j := \frac{|u^{h_j} - u_j|_{1,2}}{\|u^{h_j} - u_{j-1}\|_{0,4}^2}.$$

The ratios RL_j and RH_j are almost constant, which means that our adaptive multi-level algorithm retains quadratic convergence of Newton’s method across different mesh levels.

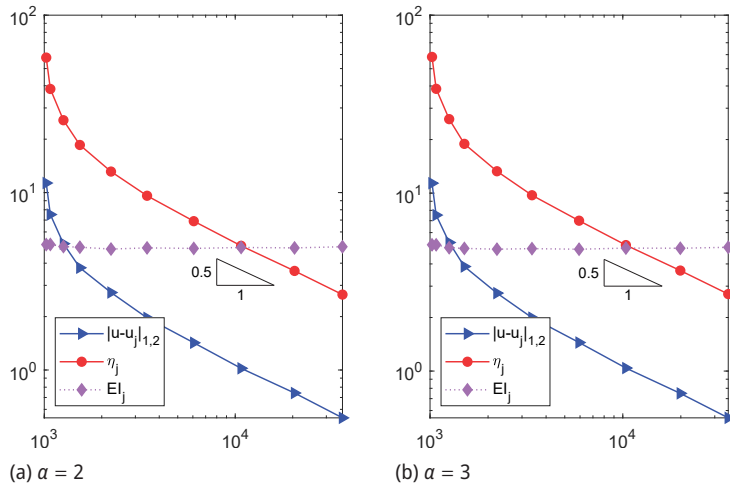


Figure 2: The history of errors $|u - u_j|_{1,2}$ and η_j , and EI_j on adaptive meshes for Example 6.1

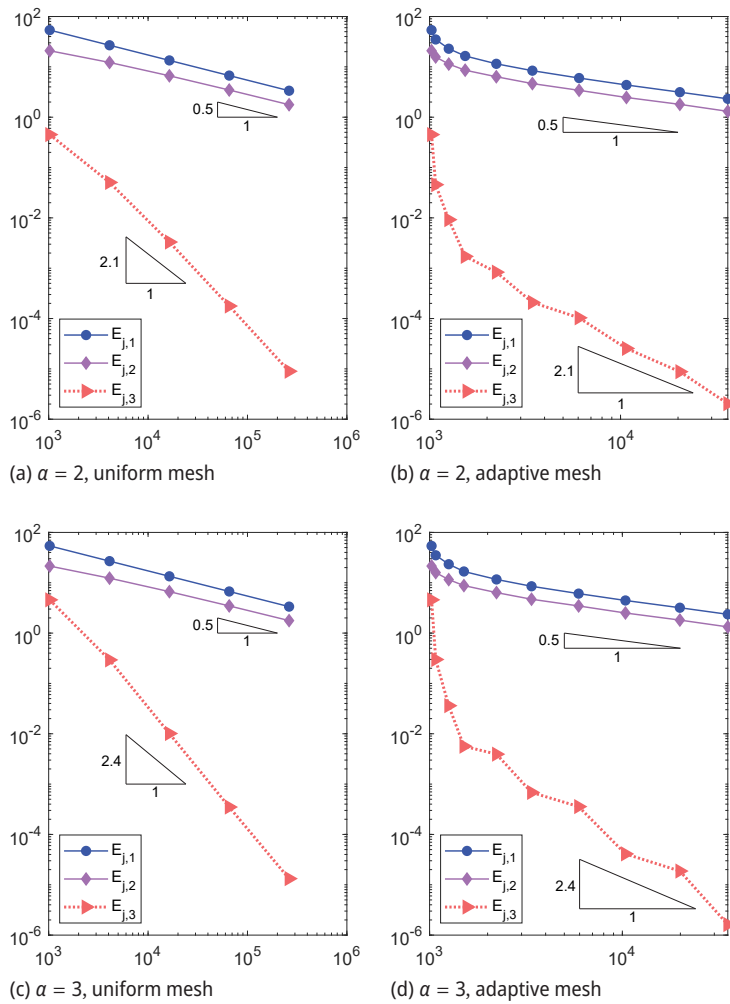


Figure 3: Convergence plots of indicators $E_{j,k}$ for $k = 1, 2, 3$ for Example 6.1

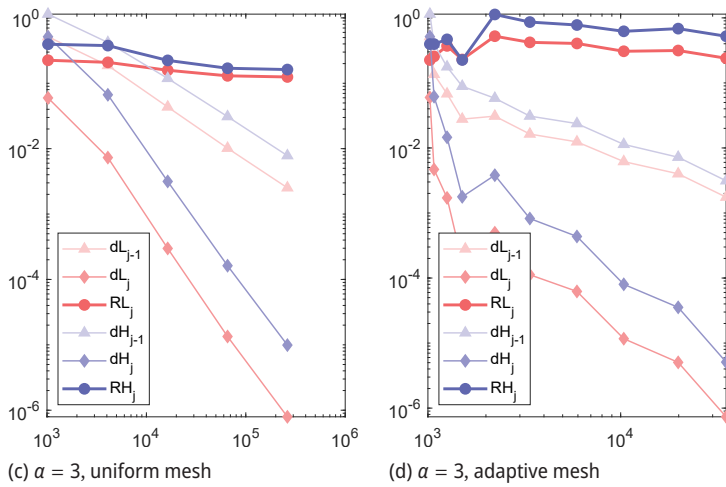
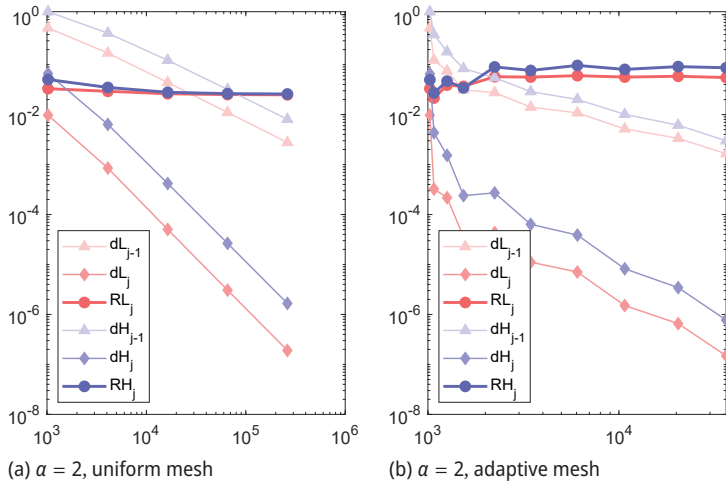


Figure 4: The history of quadratic convergence and various differences for Example 6.1

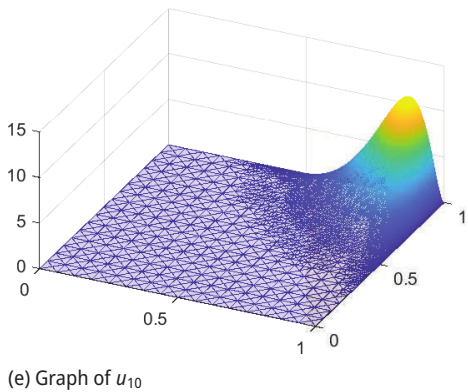
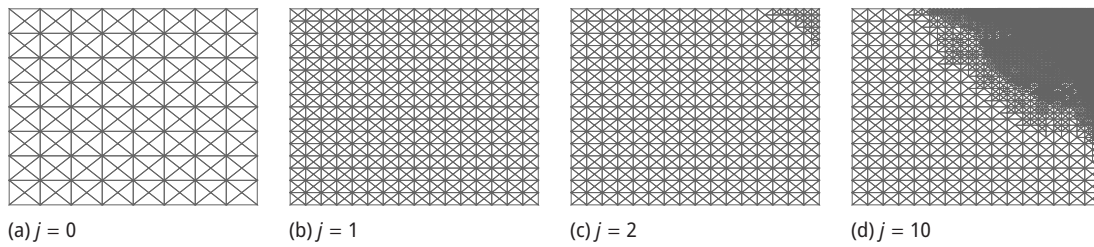


Figure 5: Meshes and u_{10} generated by the adaptive multi-level algorithm for Example 6.1

We present the adaptive meshes \mathcal{T}_j for $j = 0, 1, 2, 10$ and the approximate solution u_{10} on the final mesh \mathcal{T}_{10} for $\alpha = 2$ in Figure 5. We see that the adaptively designed meshes properly express the characteristics of the solution with a boundary layer in the upper right corner.

Example 6.2 (A Solution with a Corner Singularity). We consider SEE (4.1) with $G(\lambda, u) = u^\alpha - f$, $\alpha = 2, 3$, and L -shaped domain $\Omega = (-1, 1) \times (-1, 1) \setminus [0, 1) \times (-1, 0]$. The exact solution u in polar coordinates (r, θ) ,

$$u(r, \theta) = r^{2/3} \sin(2\theta/3) + r^\beta \sin(2\theta/3),$$

specifies data f . On the Dirichlet boundary $\Gamma_D := \{0\} \times [-1, 0] \cup [0, 1] \times \{0\}$, $u = 0$, and on the Neumann boundary $\Gamma_N := \partial\Omega \setminus \Gamma_D$, $\nabla u \cdot \mathbf{n} = g_N$ for the unit outward normal vector \mathbf{n} .

We note that $\Delta v = 0$ for $v := r^{2/3} \sin(2\theta/3) \in H^s(\Omega)$ with $s < 5/3$, and for $\beta > 1$, $w := r^\beta \sin(2\theta/3) \in H^2(\Omega)$ and $G(u) = u^\alpha - f = \Delta w \in L^2(\Omega)$. We take $\beta = 1.01$ for the numerical test. Note also that a priori error estimates for the Poisson model case on the L -shaped domain are well known (see [15, Remark 4.1, Example 4.5 and Problems 4.7]); for any $\epsilon > 0$,

$$\|u - u_h\|_{1,2} \leq h^{2/3-\epsilon} \|u\|_{H^{5/3-\epsilon}(\Omega)}, \quad \|u - u_h\|_{0,2} \leq h^{4/3-\epsilon} \|u\|_{H^{5/3-\epsilon}(\Omega)}. \tag{6.1}$$

To take the Neumann part of the boundary condition into account, we modify indicator $E_{j,2}$ as follows:

$$E_{j,2}(T) := \left(\frac{1}{2} \sum_{E \in \partial T \setminus \partial\Omega} h_E \|\nabla u_j \cdot \mathbf{n}\|_{0,2,E}^2 + \sum_{E \in \partial T \cap \Gamma_N} h_E \|\nabla u_j \cdot \mathbf{n} - g_N\|_{0,2,E}^2 \right)^{1/2}.$$

We only provide results for $\alpha = 2$ as the case $\alpha = 3$ shows similar behavior. We observe from Table 3 that convergence orders on the uniform meshes are in par with the Poisson problem (cf. (6.1)). In Table 4 and Figure 6, we display history of errors on the adaptive meshes for each level j , which shows the full rate of convergence.

j	$N_{el}(1/h_j)$	Multi-level algorithm				Nonlinear solver			
		$\ u - u_j\ _{0,2}$	CO	$ u - u_j _{1,2}$	CO	$\ u - u^h\ _{0,2}$	CO	$ u - u^h _{1,2}$	CO
0	96	2.3413e-2		2.3451e-1		2.3413e-2		2.3451e-1	
1	384	9.7023e-3	1.27	1.4761e-1	0.67	9.7329e-3	1.27	1.4760e-1	0.67
2	1536	3.9676e-3	1.29	9.1644e-2	0.69	3.9724e-3	1.29	9.1644e-2	0.69
3	6144	1.5829e-3	1.33	5.6561e-2	0.70	1.5836e-3	1.33	5.6561e-2	0.70
4	24576	6.2119e-4	1.35	3.4866e-2	0.70	6.2129e-4	1.35	3.4866e-2	0.70
5	98304	2.4160e-4	1.36	2.1517e-2	0.70	2.4161e-4	1.36	2.1517e-2	0.70

Table 3: Convergence orders (CO) of various errors on uniform meshes for Example 6.2

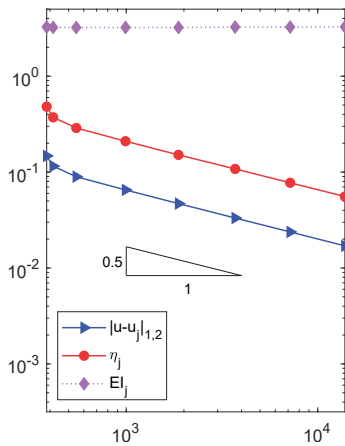


Figure 6: Convergence orders of errors and etas on adaptive meshes for Example 6.2

j	N_{el}	$ u - u_j _{1,2}$	η_j	EI_j
0	96	2.3451e-1		
1	384	1.4761e-1	4.8126e-1	3.26
2	415	1.1532e-1	3.7244e-1	3.23
3	548	8.9599e-2	2.8853e-1	3.22
4	991	6.5174e-2	2.0990e-1	3.22
5	1875	4.6944e-2	1.5114e-1	3.22
6	3705	3.3205e-2	1.0794e-1	3.25
7	7177	2.3774e-2	7.7373e-2	3.25
8	13858	1.7006e-2	5.5668e-2	3.27

Table 4: Errors and etas, and efficiency indices on adaptive meshes for Example 6.2

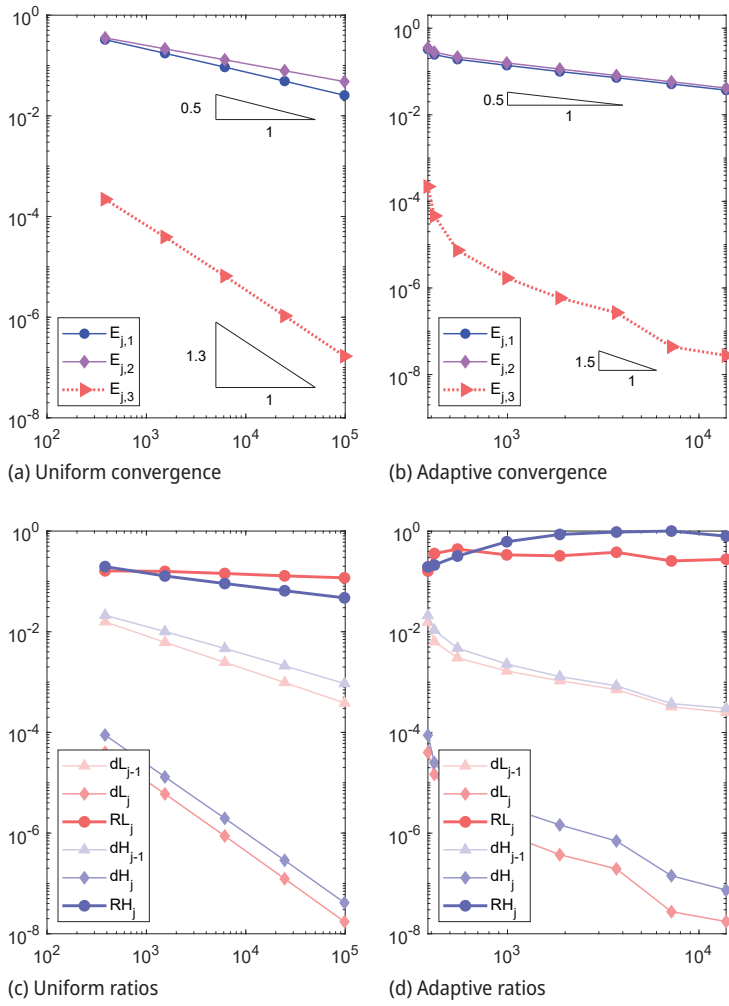


Figure 7: The convergence of indicators E_j and ratios RL_j and RH_j for Example 6.2

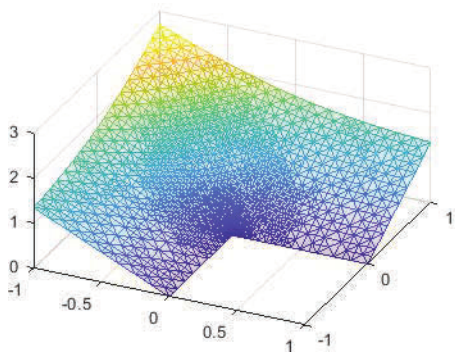
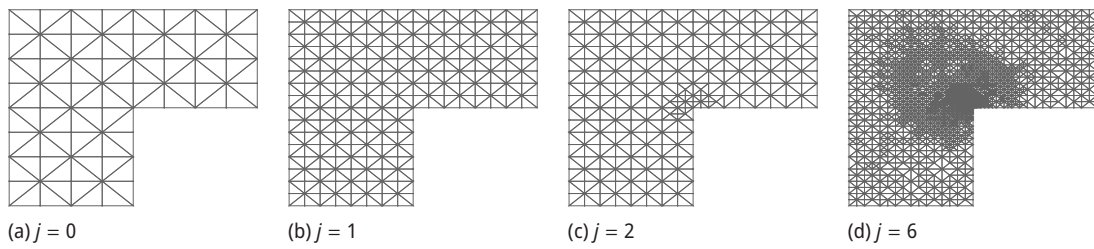


Figure 8: Meshes and u_6 generated by the adaptive multi-level algorithm for Example 6.2

The superiority of the adaptive multi-level algorithm can be clearly seen in Tables 3 and 4. Indeed, the accuracy on uniform mesh with $N_{el} = 24,576$ in Table 3 is comparable to the accuracy on adaptive mesh with $N_{el} = 3,705$ in Table 4 (see boldface in Tables 3 and 4).

In Figure 7 (a) and (b), we present indicators and their convergence orders on uniform and adaptive meshes, respectively. The convergence order of additional term $E_{j,3}$ is higher than that of the other indicators $E_{j,1}$ and $E_{j,2}$. The ratios RL_j and RH_j are almost constant in Figure 7(c) and (d), which means that our adaptive multi-level algorithm retains quadratic convergence of Newton’s method across different mesh levels.

We present adaptive meshes \mathcal{T}_j for $j = 0, 1, 2, 6$ and the approximate solution u_6 on the mesh \mathcal{T}_6 in Figure 8. We see that the adaptively designed meshes well reflect the corner singularity of the solution.

6.2 Navier–Stokes Equations

Now, we consider Navier–Stokes equations (3.1)–(3.3). We use the lowest-order Raviart–Thomas finite element space $X_j = \overline{RT}_0^2(\mathcal{T}_j) \times P_0^2(\mathcal{T}_j)$. We introduce the following error indicators: for $1 \leq j \leq J$,

$$\begin{aligned} E_{j,1}(T) &:= h_T \|R_j(\mathbf{f})\|_{0,2,T}, & E_{j,2}(T) &:= h_T \|A\boldsymbol{\sigma}_j - \nabla \mathbf{u}_j\|_{0,2,T}^2, \\ E_{j,3}(T) &:= h_T \|\mathbf{Curl}(A\boldsymbol{\sigma}_j)\|_{0,2,T}, & E_{j,4}(T) &:= \left(\frac{1}{2} \sum_{E \in T \setminus \partial\Omega} h_E \|[\gamma_t A\boldsymbol{\sigma}_j]\|_{0,2,E}^2\right)^{1/2}, \\ E_{j,5}(T) &:= \left(\sum_{E \in T \cap \partial\Omega} h_E \|\gamma_t(A\boldsymbol{\sigma}_j - \nabla \mathbf{g})\|_{0,2,E}^2\right)^{1/2}, & E_{j,6}(T) &:= \left(\sum_{E \in T \cap \partial\Omega} h_E \|\mathbf{g} - \mathbf{g}_j\|_{0,2,E}^2\right)^{1/2}, \\ E_{j,7}(T) &:= \|(\mathbf{u}_j - \mathbf{u}_{j-1}) \cdot A(\boldsymbol{\sigma}_j - \boldsymbol{\sigma}_{j-1})\|_{0,6/5,T}. \end{aligned}$$

We define the local error estimator $\eta_j(T)$ by

$$\eta_j(T) := \left(\sum_{k=1}^7 E_{j,k}(T)^2\right)^{1/2},$$

the global error estimator η_j by

$$\eta_j := \left(\sum_{T \in \mathcal{T}_j} \eta_j(T)^2\right)^{1/2},$$

and efficiency indices El_j by

$$El_j := \frac{\eta_j}{\|\phi - \phi_j\|_X}.$$

Example 6.3 (A Solution with a Boundary Layer). We consider NSE (3.1)–(3.3) with $\Omega = (0, 1) \times (0, 1)$ and $\nu = 1$. The exact solutions $\mathbf{u} = (u_1, u_2)$ and p ,

$$\begin{aligned} u_1(x, y) &= \frac{1}{2} x^2 y (x - 1)^2 (y - 1) (4y - 10y + 10y^2 - 2) e^{10(x+y-1)}, \\ u_2(x, y) &= -\frac{1}{2} x y^2 (x - 1) (y - 1)^2 (4x - 10x + 10x^2 - 2) e^{10(x+y-1)}, \\ p(x, y) &= \cos(\pi x) \cos(\pi y), \end{aligned}$$

specify data \mathbf{f} and \mathbf{g} .

Table 5 shows convergence history of multi-level algorithm and nonlinear solver on uniform meshes. We observe that the errors $\|\phi - \phi_j\|_X$ up to low levels ($j \leq 2$) exhibit non-optimal convergence behavior since meshes in low levels are in the pre-asymptotic range. We see that the errors have optimal order of convergence by carrying out more uniform refinement ($j \geq 3$). However, in this uniform refinement, for high levels $j \geq 3$, the number of elements at \mathcal{T}_j increases rapidly and the finite element space X_j requires too many degrees of freedom as in Example 6.1; the adaptive multi-level algorithm is beneficial.

j	$N_{el}(1/h_j)$	Multi-level algorithm		Nonlinear solver	
		$\ \phi - \phi_j\ _X$	CO	$\ \phi - \phi^{h_j}\ _X$	CO
0	256(8)	7.903738e+0		7.903738e+0	
1	1024(16)	4.556065e+0	0.795	4.556080e+0	0.795
2	4096(32)	2.436524e+0	0.903	2.436521e+0	0.903
3	16384(64)	1.245096e+0	0.969	1.245096e+0	0.969
4	65536(128)	6.262103e-1	0.992	6.262102e-1	0.992
5	262144(256)	3.135742e-1	0.998	3.135742e-1	0.998

Table 5: Convergence orders (CO) of errors on uniform meshes for Example 6.3

j	N_{el}	$\ \phi - \phi_j\ _X$	η_j	EI_j
0	256	7.903738e+0		
1	1024	4.556065e+0	1.164e+1	2.554
2	1206	2.472931e+0	4.728e+0	1.912
3	1486	1.525288e+0	2.924e+0	1.917
4	2260	1.137552e+0	2.025e+0	1.780
5	3492	7.797460e-1	1.455e+0	1.866
6	6130	5.987671e-1	1.066e+0	1.780
7	10614	4.118791e-1	7.717e-1	1.874
8	20130	3.183675e-1	5.678e-1	1.783
9	35736	2.221120e-1	4.165e-1	1.875

Table 6: Efficiency indices on adaptive meshes for Example 6.3

We present results for the adaptive multi-level algorithm applied to Example 6.3 with a boundary layer in Table 6. The superiority of the adaptive multi-level algorithm can be clearly seen in Table 5 and Table 6. Indeed, the accuracy on uniform mesh with $N_{el} = 65,536$ in Table 5 is comparable to the accuracy on adaptive mesh with $N_{el} = 6,130$ in Table 6 (see boldface in Tables 5 and 6). We present the history of errors $\|\phi - \phi_j\|_X$ and η_j , and the efficiency indices EI_j in Figure 9 (a)

Figure 9 (b) and (c) illustrate the history of indicators on uniform and adaptive meshes, respectively. Here, the indicators $E_{j,k}$ in Figure 9 (b) and (c) are defined by

$$E_{j,k} := \left(\sum_{T \in \mathcal{T}_j} E_{j,k}(T)^2 \right)^{1/2}, \quad k = 1, 2, \dots, 7.$$

We observe that the additional term $E_{j,7}$ has smaller value and is of higher order than the other indicators for both uniform and adaptive meshes as mentioned in Remark 3.19. Similar behavior has been observed in the case of SEE.

To examine the behavior of quadratic convergence in the sense of (2.5), we investigate the difference between ϕ_{j-1} and ϕ^{h_j} , and the difference between ϕ_j and ϕ^{h_j} , and their ratios. These quantities are displayed and particularly ratios RX_j are plotted by bold line in Figure 9 (d). Notations of the legend in Figure 9 (d) are defined by

$$dX_{j-1} := \|\phi^{h_j} - \phi_{j-1}\|_X, \quad dX_j := \|\phi^{h_j} - \phi_j\|_X, \quad RX_j := \frac{\|\phi^{h_j} - \phi_j\|_X}{\|\phi^{h_j} - \phi_{j-1}\|_X^2}.$$

We find that the ratios RX_j are almost constant, which means that our adaptive multi-level algorithm maintains quadratic convergence of Newton’s method across different mesh levels.

In Figure 10, we exhibit some meshes \mathcal{T}_j for $j = 0, 1, 2, 6$ generated by the adaptive multi-level algorithm and a graph of the final approximate solution \mathbf{u}_6 measured in Euclidean norm on \mathcal{T}_6 . We see that the adaptively designed meshes properly express the characteristics of the solution with boundary layers in the upper right corner.

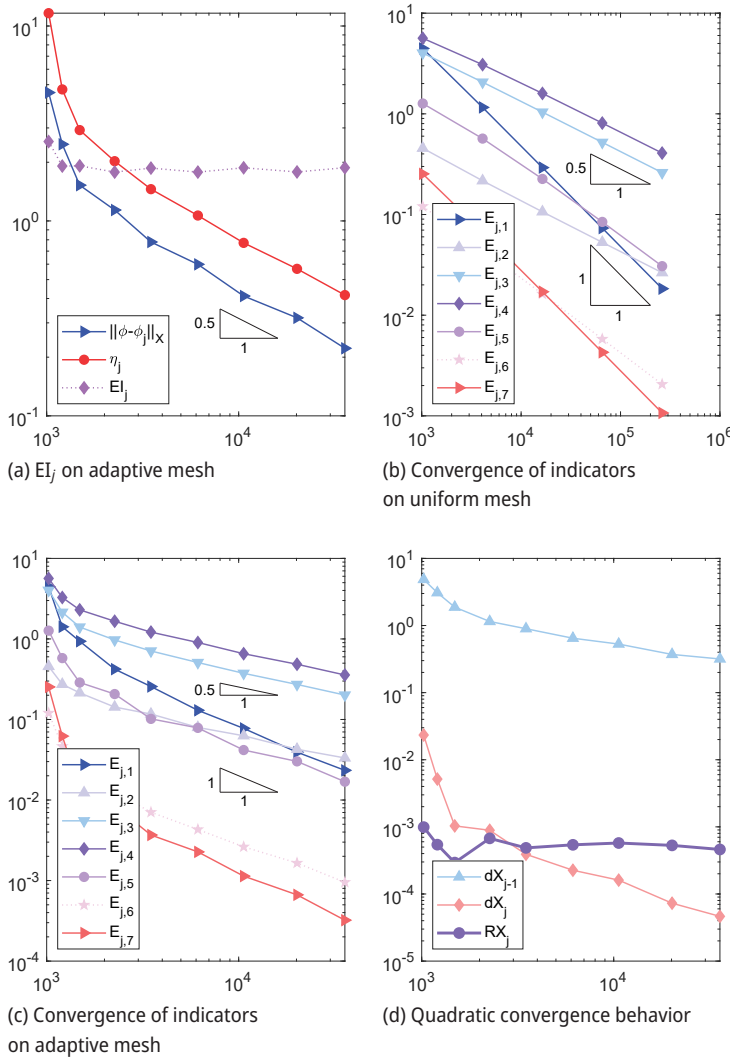


Figure 9: Various behaviors for Example 6.3

7 Conclusion

The adaptive multi-level algorithm analyzed in this paper can be used to efficiently solve nonlinear problems with singularity or layer. We start with a nonlinear system on a very coarse mesh and then solve linearized system on the successively refined meshes with adaptivity. The BRR framework facilitates to use the general theory established for a linear problem associated with given nonlinear equations. In particular, a posteriori error estimates for the linear problem can be utilized to find reliable error estimators for the nonlinear problem.

As applications, the Navier–Stokes equations and the semilinear elliptic equations are considered. Residual type reliable and efficient a posteriori error estimators are constructed for the adaptive multi-level algorithm for the considered problems. With few modifications we apply existing a posteriori error estimators for the linear Stokes equations and elliptic equations corresponding to the NSE and SEE, respectively. Note that the multi-level algorithm with uniform meshes has quadratic convergence in the sense of (2.5). We emphasize that the multi-level algorithm with adaptive meshes retains quadratic convergence of Newton’s method across different mesh levels, which is validated from the numerical results presented in this paper.

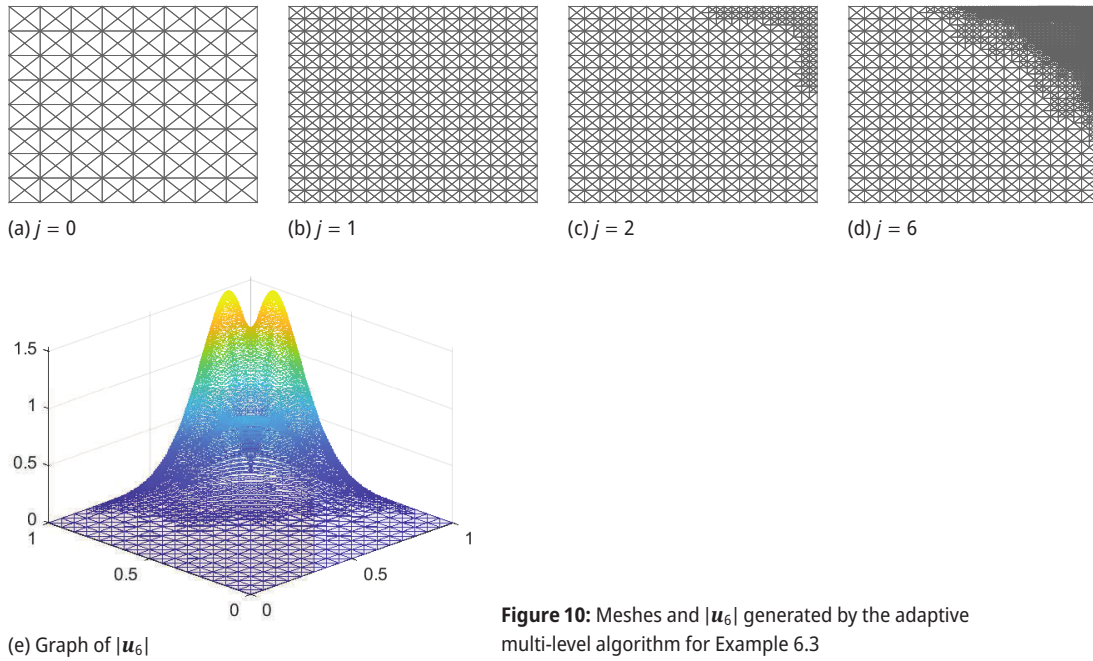


Figure 10: Meshes and $|u_6|$ generated by the adaptive multi-level algorithm for Example 6.3

Data Availability: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding: Dongho Kim was supported by NRF-2021R1F1A106243413. The research of Eun-Jae Park was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (NRF-2022R1A2B5B02002481). Boyoon Seo was supported by NRF-2020R1I1A1A01070361.

References

- [1] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Pure Appl. Math. (New York), Wiley-Interscience, New York, 2000.
- [2] I. Babuška and T. Strouboulis, *The Finite Element Method and its Reliability*, Numer. Math. Sci. Comput., Oxford University, New York, 2001.
- [3] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer, New York, 1991.
- [4] F. Brezzi, J. Rappaz and P.-A. Raviart, Finite-dimensional approximation of nonlinear problems. I. Branches of nonsingular solutions, *Numer. Math.* **36** (1980/81), no. 1, 1–25.
- [5] Z. Cai, C. Wang and S. Zhang, Mixed finite element methods for incompressible flow: Stationary Navier–Stokes equations, *SIAM J. Numer. Anal.* **48** (2010), no. 1, 79–94.
- [6] Z. Cai and Y. Wang, A multigrid method for the pseudostress formulation of Stokes problems, *SIAM J. Sci. Comput.* **29** (2007), no. 5, 2078–2095.
- [7] G. Caloz and J. Rappaz, Numerical analysis for nonlinear and bifurcation problems, in: *Handbook of Numerical Analysis. Vol. V*, North-Holland, Amsterdam (1997), 487–637.
- [8] C. Carstensen, A posteriori error estimate for the mixed finite element method, *Math. Comp.* **66** (1997), no. 218, 465–476.
- [9] C. Carstensen and S. A. Funken, A posteriori error control in low-order finite element discretisations of incompressible stationary flow problems, *Math. Comp.* **70** (2001), no. 236, 1353–1381.
- [10] C. Carstensen, D. Kim and E.-J. Park, A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem, *SIAM J. Numer. Anal.* **49** (2011), no. 6, 2501–2523.
- [11] P. Clément, Approximation by finite element functions using local regularization, *RAIRO Oper. Res.* **9** (1975), no. R-2, 77–84.
- [12] V. Ervin, W. Layton and J. Maubach, A posteriori error estimators for a two-level finite element method for the Navier–Stokes equations, *Numer. Methods Partial Differential Equations* **12** (1996), no. 3, 333–346.
- [13] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes equations*, Springer Ser. Comput. Math. 5, Springer, Berlin, 1986.

- [14] V. John, Residual a posteriori error estimates for two-level finite element methods for the Navier-Stokes equations, *Appl. Numer. Math.* **37** (2001), no. 4, 503–518.
- [15] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University, Cambridge, 1987.
- [16] D. Kim and E.-J. Park, Primal mixed finite-element approximation of elliptic equations with gradient nonlinearities, *Comput. Math. Appl.* **51** (2006), no. 5, 793–804.
- [17] D. Kim and E.-J. Park, A posteriori error estimator for expanded mixed hybrid methods, *Numer. Methods Partial Differential Equations* **23** (2007), no. 2, 330–349.
- [18] D. Kim and E.-J. Park, A posteriori error estimators for the upstream weighting mixed methods for convection diffusion problems, *Comput. Methods Appl. Mech. Engrg.* **197** (2008), no. 6–8, 806–820.
- [19] D. Kim and E.-J. Park, A priori and a posteriori analysis of mixed finite element methods for nonlinear elliptic equations, *SIAM J. Numer. Anal.* **48** (2010), no. 3, 1186–1207.
- [20] D. Kim, E.-J. Park and B. Seo, A unified framework for two-grid methods for a class of nonlinear problems, *Calcolo* **55** (2018), no. 4, Paper No. 45.
- [21] D. Kim, E.-J. Park and B. Seo, Optimal error estimates for the pseudostress formulation of the Navier–Stokes equations, *Appl. Math. Lett.* **78** (2018), 24–30.
- [22] D. Kim, E.-J. Park and B. Seo, Convergence of multi-level algorithms for a class of nonlinear problems, *J. Sci. Comput.* **84** (2020), no. 2, Paper No. 34.
- [23] K. Y. Kim, A posteriori error analysis for locally conservative mixed methods, *Math. Comp.* **76** (2007), no. 257, 43–66.
- [24] F. A. Milner and E.-J. Park, Mixed finite-element methods for Hamilton–Jacobi–Bellman-type equations, *IMA J. Numer. Anal.* **16** (1996), no. 3, 399–412.
- [25] R. Verfürth, A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations, *Math. Comp.* **62** (1994), no. 206, 445–475.
- [26] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Teubner-Wiley, Stuttgart, 1996.
- [27] J. Xu, A novel two-grid method for semilinear elliptic equations, *SIAM J. Sci. Comput.* **15** (1994), no. 1, 231–237.
- [28] J. Xu, Two-grid discretization techniques for linear and nonlinear PDEs, *SIAM J. Numer. Anal.* **33** (1996), no. 5, 1759–1777.