A POSTERIORI ERROR ESTIMATION FOR AN INTERIOR PENALTY TYPE METHOD EMPLOYING H(DIV) ELEMENTS FOR THE STOKES EQUATIONS

JUNPING WANG*, YANQIU WANG†, AND XIU YE‡

Abstract. This paper establishes a posteriori error analysis for the Stokes equations discretized by an interior penalty type method using H(div) finite elements. The a posteriori error estimator is then employed for designing two grid refinement strategies: one is locally based and the other is globally based. The locally based refinement technique is believed to be able to capture local singularities in the numerical solution. The numerical formulations for the Stokes problem make use of H(div) conforming elements of the Raviart–Thomas type. Therefore, the finite element solution features a full satisfaction of the continuity equation (mass conservation). The result of this paper provides a rigorous analysis for the method's reliability and efficiency. In particular, an H^1-norm a posteriori error estimator is obtained, together with upper and lower bound estimates. Numerical results are presented to verify the new theory of a posteriori error estimators.

Key words. finite element methods, Stokes problem, a posteriori estimates

AMS subject classifications. Primary, 65N15, 65N30, 76D07; Secondary, 35B45, 35J50

1. Introduction. In this paper, the authors are concerned with a posteriori error analysis for numerical solutions of the Stokes equations discretized by an interior penalty type method using H(div) finite elements. The Stokes equations under study seeks a velocity $\mathbf{u}$ and a pressure p satisfying

$$-\Delta \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega, \tag{1.1}$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega, \tag{1.2}$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \partial\Omega, \tag{1.3}$$

where Δ, ∇, and $\nabla\cdot$ denote the Laplacian, gradient, and divergence operators, respectively; $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ is the region occupied by the fluid; $\mathbf{f} \in (L^2(\Omega))^d$ is the unit external volumetric force acting on the fluid. For simplicity, the method will be presented for two-dimensional problems ($d = 2$) on polygonal domains. An extension to three dimensions can be made formally for general polyhedral domains.

In the engineering society, it is required for numerical schemes to retain the original physical properties, such as mass and energy conservation. For the Stokes equations, such a requirement translates to a discretization scheme satisfying the incompressible constraint equation (1.2) exactly on the computational domain. However, constructing such a finite element space in H^1 is quite challenging, and the resulting spaces are often not computationally friendly. Recently, a new approach, which uses H(div) conforming elements, has been developed for the Stokes [22] and Navier-Stokes equations [23]. Numerical solutions of this method satisfy the continuity equation (1.2)

*Division of Mathematical Sciences, National Science Foundation, Arlington, VA 22230 (jwang@nsf.gov). The research of Wang was supported by the NSF IR/D program, while working at the Foundation. However, any opinion, finding, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

†Department of Mathematics, Oklahoma State University, Stillwater, OK 74075 (yqwang@math.okstate.edu).

‡Department of Mathematics, University of Arkansas at Little Rock, Little Rock, AR 72204 (xxye@ualr.edu). This research was supported in part by National Science Foundation Grant DMS-0813571

exactly. The discrete velocity, which lies in an $H(\mathrm{div})$ conforming finite element space, has continuous normal component across internal edges. The tangential continuity is imposed weakly, using the idea similar to the one in interior penalty methods.

The idea of employing $H(div)$ conforming elements to the Stokes equations has been explored by several researchers in the last two decades. All the existing work applies the $H(div)$ conforming elements to either a stress-velocity or a stress-velocity-pressure formulation of Stokes and Navier-Stokes equations, where the stress is in $H(div)$. In [5, 6], a pseudostress-velocity formulation has been proposed and solved by $H(div)$ elements. Its extension to the pseudostress-velocity-pressure formulation has been considered in [11], together with a priori and a posteriori error analysis. In [10], an augmented formulation using the $H(div)$ conforming element method has been developed. Also, the dual-mixed method has been studied in [12, 13]. In all these work, the $H(div)$ element was used to approximate the stress or stress-type dual variables. We would like to point out that our method is different from the existing ones in that we use the velocity-pressure formulation and the $H(div)$ element is used to approximate the velocity directly.

The goal of this paper is to obtain an a posteriori error estimator for the $H(\mathrm{div})$ finite element method developed in [22], and to provide its upper and lower bound estimates. In the analysis, we borrow many ideas from some previous a posteriori error analysis for Stokes equations, including [20], which uses conforming finite elements, and [8, 9, 7, 14, 15, 21], which use nonconforming finite elements. Our contributions in this paper are: (1) successfully established a posteriori error estimator for the $H(\mathrm{div})$ finite element method, (2) proposed and tested a new grid refinement method using local information in order to capture local singularities, (3) conducted a series of numerical experiments for the refinement strategies. We hope that the numerical results presented in this paper will shine some light on a further development of computational techniques in fluid dynamics.

The paper is organized as follows. In Section 2, we introduce some notations for scalar, vector, and matrix Sobolev spaces. In Section 3, the $H(\mathrm{div})$ finite element formulation and its a priori error estimate are stated. Section 4 is dedicated to an establishment and analysis of an a posteriori error estimator. In Section 5, we shall present two grid refinement strategies: one is locally based and the other is globally based. Finally in Section 6, we present some numerical results and offer some of our own observations.

2. Notations. Let D be a polygon. We use standard definitions for the Sobolev spaces $H^s(D)$ and their associated inner products $(\cdot,\cdot)_{s,D}$, norms $\|\cdot\|_{s,D}$, and seminorms $|\cdot|_{s,D}$, for $s \geq 0$. The space $H^0(D)$ coincides with $L^2(D)$, for which the norm and the inner product are also denoted by $\|\cdot\|_D$ and $(\cdot,\cdot)_D$, respectively. For convenience, when $D = \Omega$, we usually suppress D in the subscript. Denote $L^2_0(\Omega)$ to be the subspace of $L^2(\Omega)$ consisting of functions with mean value zero.

Notice that all above definitions can be extended to the case of vector-valued or matrix-valued functions, through product spaces. We use the same notation for their norms and inner products. Also, all these definitions can be transported from a polygon D to an edge e. Similar notation system will be employed, for example, $\|\cdot\|_{s,e}$ and $\|\cdot\|_e$.

Throughout the paper, we follow the convention that a bold face Latin character

denotes a vector. For vector function $\mathbf{v} \in \mathbb{R}^2$, define

$$\nabla \mathbf{v} = \begin{pmatrix} \frac{\partial v_1}{\partial x} & \frac{\partial v_1}{\partial y} \\ \frac{\partial v_2}{\partial x} & \frac{\partial v_2}{\partial y} \end{pmatrix}, \qquad \mathbf{curl}\ \mathbf{v} = \begin{pmatrix} -\frac{\partial v_1}{\partial y} & \frac{\partial v_1}{\partial x} \\ -\frac{\partial v_2}{\partial y} & \frac{\partial v_2}{\partial x} \end{pmatrix},$$
$$\nabla \cdot \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}.$$

Define the space $H(\mathrm{div}; \Omega)$ to be the set of vector-valued functions on Ω which, together with their divergence, are square integrable; i.e.,

$$H(\mathrm{div}; \Omega) = \left\{ \mathbf{v} : \ \mathbf{v} \in (L^2(\Omega))^2,\ \nabla \cdot \mathbf{v} \in L^2(\Omega) \right\}.$$

The norm in $H(\mathrm{div}; \Omega)$ is defined by

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

Let $\mathcal{T}_h$ be a geometrically conformal triangulation of the domain Ω, i.e., the intersection of any two triangles in $\mathcal{T}_h$ is either empty, a common vertex, or a common edge. Denote h_K to be the diameter of triangle $K \in \mathcal{T}_h$, and h to be the maximum of all h_K. We also assume $\mathcal{T}_h$ is shape regular, that is, for each $K \in \mathcal{T}_h$, the ratio between h_K and the diameter of the inscribed circle is bounded above. This ensures that the scaling arguments and the inverse inequalities work on each triangle.

Define the finite element spaces V_h and W_h for the velocity and pressure variables, respectively, by

$$V_h = \{ \mathbf{v} \in H(\mathrm{div}; \Omega) : \ \mathbf{v}|_K \in V_k(K) \quad \forall K \in \mathcal{T}_h;\ \mathbf{v} \cdot \mathbf{n}|_{\partial\Omega} = 0 \}$$

$$W_h = \{ q \in L_0^2(\Omega) : \ q|_K \in W_k(K) \quad \forall K \in \mathcal{T}_h \},$$

where $\mathbf{n}$ is the outward normal direction, $(V_k(K), W_k(K))$ can be any existing $H(\mathrm{div})$ conforming finite element pairs [4] of order $k \geq 1$. For example, the Raviart–Thomas elements (RT_k) [16] or the Brezzi–Douglas–Marini elements (BDM_k) [3]. For BDM_1 element, W_h consists of all piecewise constants on $\mathcal{T}_h$. Also, notice that for all $\mathbf{v} \in V_h$, it has continuous normal component $\mathbf{v} \cdot \mathbf{n}$ across internal edges, while its tangential component is not necessarily continuous.

For vectors $\mathbf{v}$, $\mathbf{n} \in \mathbb{R}^2$, denote $\mathbf{v} \otimes \mathbf{n} = \{v_i n_j\}_{1 \leq i,j \leq 2}$ to be the vector tensor product. For matrices σ, $\tau \in \mathbb{R}^{2\times 2}$, define $\sigma : \tau = \sum_{i,j=1}^{2} \sigma_{ij} \tau_{ij}$. Notice that

$$\mathbf{v} \cdot (\sigma \mathbf{n}) = (\mathbf{v} \otimes \mathbf{n}) : \sigma\,.$$

Later we will use the above equation without explicit mentioning.

Let e be an interior edge shared by two elements K_1 and K_2 in $\mathcal{T}_h$. Denote unit normal vectors $\mathbf{n}_1$, $\mathbf{n}_2$ and tangential directions $\mathbf{t}_1$, $\mathbf{t}_2$, respectively, on e for K_1 and K_2 (as shown in Figure 2.1). Define the average $\{\cdot\}$ and jump $[\cdot]$ on e for scalar function q, vector function $\mathbf{v}$ and matrix function σ, respectively, by

$$\{q\} = \frac{1}{2}(q|_{\partial K_1} + q|_{\partial K_2}), \qquad [q\mathbf{n}] = q|_{\partial K_1}\mathbf{n}_1 + q|_{\partial K_2}\mathbf{n}_2,$$
$$\{\mathbf{v}\} = \frac{1}{2}(\mathbf{v}|_{\partial K_1} + \mathbf{v}|_{\partial K_2}), \qquad \{\sigma\} = \frac{1}{2}(\sigma|_{\partial K_1} + \sigma|_{\partial K_2}),$$
$$[\sigma\mathbf{n}] = \sigma|_{\partial K_1}\mathbf{n}_1 + \sigma|_{\partial K_2}\mathbf{n}_2, \qquad [\sigma\mathbf{t}] = \sigma|_{\partial K_1}\mathbf{t}_1 + \sigma|_{\partial K_2}\mathbf{t}_2.$$

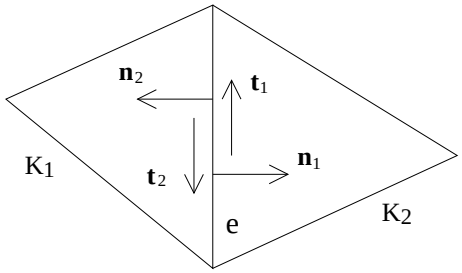


FIG. 2.1. *Normal and tangential vectors for neighboring triangles.*

We also define a scalar average $\{\!\{\varepsilon(\cdot)\}\!\}$ and a matrix valued jump $[\![\cdot]\!]$ for a vector function $\mathbf{v}$ by

$$\{\!\{\varepsilon(\mathbf{v})\}\!\} = \frac{1}{2}\left(\mathbf{n}_1 \cdot \nabla(\mathbf{v}\cdot\mathbf{t}_1)|_{\partial K_1} + \mathbf{n}_2 \cdot \nabla(\mathbf{v}\cdot\mathbf{t}_2)|_{\partial K_2}\right),$$
$$[\![\mathbf{v}]\!] = \mathbf{v}|_{\partial K_1}\otimes\mathbf{n}_1 + \mathbf{v}|_{\partial K_2}\otimes\mathbf{n}_2.$$

If e is a boundary edge, the above definitions should be modified such that both the average and the jump are equal to the one-sided values on e. For example

$$\{q\} = q|_e, \quad [q\mathbf{n}] = q|_e\mathbf{n}.$$

Other terms should be modified in the same fashion.

Denote by $\mathcal{E}_h$ the set of all edges in $\mathcal{T}_h$, and $\mathcal{E}_h^0 := \{e \in \mathcal{E}_h,\, e \nsubseteq \partial\Omega\}$ the set of all interior edges. Let $V(h) = V_h + (H^s(\Omega)\cap H_0^1(\Omega))^2$, with $s > \frac{3}{2}$, where the summation means the mathematical sum of functions from each subspace. For $\mathbf{v} \in V(h)$, define $\nabla_h\mathbf{v}$ to be the function whose restriction to each element $K \in \mathcal{T}_h$ is given by the standard gradient $\nabla\mathbf{v}$.

3. Finite element scheme and a priori error estimate. We use the numerical scheme proposed and analyzed in [22], where details on convergence analysis can be found. For simplicity of presentation, this paper uses a slightly different notation in describing the numerical schemes. To this end, we introduce two bilinear forms on $V(h)\times V(h)$ as follows

$$a_s(\mathbf{w},\mathbf{v}) = (\nabla_h\mathbf{w},\nabla_h\mathbf{v}) + \sum_{e\in\mathcal{E}_h}\int_e \left(\alpha h_e^{-1}[\![\mathbf{w}]\!] : [\![\mathbf{v}]\!] - \{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!] - \{\nabla\mathbf{v}\} : [\![\mathbf{w}]\!]\right) ds,$$

$$a_{ns}(\mathbf{w},\mathbf{v}) = (\nabla_h\mathbf{w},\nabla_h\mathbf{v}) + \sum_{e\in\mathcal{E}_h}\int_e \left(\alpha h_e^{-1}[\![\mathbf{w}]\!] : [\![\mathbf{v}]\!] - \{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!] + \{\nabla\mathbf{v}\} : [\![\mathbf{w}]\!]\right) ds,$$

where $\alpha > 0$ is a parameter to be determined later, and h_e is the length of the edge e. It is not hard to verify that the above two bilinear forms are exactly the same as those stated in [22]. For reader's convenience, a brief explanation is given in Appendix A for such a verification.

As usual, there is a bilinear form on $V(h)\times L_0^2(\Omega)$ given by

$$b(\mathbf{v},q) = (\nabla\cdot\mathbf{v},q).$$

The $H(\mathrm{div})$ finite element scheme for (1.1)–(1.3) seeks $(\mathbf{u}_h;p_h)\in V_h\times W_h$ such that

$$a(\mathbf{u}_h,\mathbf{v}) - b(\mathbf{v},p_h) = (\mathbf{f},\mathbf{v}) \qquad \forall\mathbf{v}\in V_h, \tag{3.1}$$

$$b(\mathbf{u}_h,q) = 0 \qquad \forall q\in W_h, \tag{3.2}$$

where $a(\cdot,\cdot)$ can be taken as either $a_s(\cdot,\cdot)$ or $a_{ns}(\cdot,\cdot)$. It has been proved in [22] that the above system (3.1)–(3.2) is well posed for non-symmetric bilinear form $a_{ns}(\cdot,\cdot)$ with $\alpha > 0$, and for symmetric bilinear form $a_s(\cdot,\cdot)$ with α large enough.

It is not hard to see that the solution $(\mathbf{u};\, p)$ of (1.1)–(1.3) also satisfies

$$a(\mathbf{u},\mathbf{v}) - b(\mathbf{v},p) = (\mathbf{f},\mathbf{v}) \qquad \forall \mathbf{v} \in V_h, \tag{3.3}$$

$$b(\mathbf{u},q) = 0 \qquad \forall q \in W_h. \tag{3.4}$$

Subtracting (3.1)–(3.2) from (3.3)–(3.4) gives the following error equations

$$a(\mathbf{u}-\mathbf{u}_h,\mathbf{v}) - b(\mathbf{v},p-p_h) = 0 \qquad \forall \mathbf{v} \in V_h, \tag{3.5}$$

$$b(\mathbf{u}-\mathbf{u}_h,q) = 0 \qquad \forall q \in W_h. \tag{3.6}$$

To investigate the approximation property of the above numerical scheme, we introduce a norm on $V(h)$ as follows:

$$|||\mathbf{v}|||^2 = \|\nabla_h \mathbf{v}\|^2 + \sum_{e\in\mathcal{E}_h} h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2 + \sum_{e\in\mathcal{E}_h} h_e\|\{\!\{\varepsilon(\mathbf{v})\}\!\}\|_e^2.$$

Let Π_h be the interpolation into V_h associated with the usual degrees of freedom (see [4] for details), and Q_h be the L^2 projection from $L_0^2(\Omega)$ onto W_h. It is well-known that $(\nabla\cdot)\Pi_h = Q_h(\nabla\cdot)$. Furthermore, the following a priori error estimate has been proved in [22]:

THEOREM 3.1. *Let* $(\mathbf{u};p)$ *be the solution of* (1.1)–(1.3) *and* $(\mathbf{u}_h;p_h) \in V_h \times W_h$ *be obtained from* (3.1)–(3.2). *Then, there exists a constant* C *independent of* h *such that*

$$|||\mathbf{u}-\mathbf{u}_h||| + \|p-p_h\| \le C\left(|||\mathbf{u}-\Pi_h\mathbf{u}||| + \|p-Q_hp\|\right). \tag{3.7}$$

For example, Theorem 3.1 implies an error estimate of $|||\mathbf{u}-\mathbf{u}_h||| + \|p-p_h\| = O(h)$ when the BDM_1 element is used in the numerical discretization with an exact solution $(\mathbf{u};\, p) \in (H^2 \cap H_0^1)^2 \times (H^1 \cap L_0^2)$.

4. A posteriori error estimator. The goal of this section is to derive an a posteriori error estimator for the finite element formulation (3.1)–(3.2). A detailed presentation will be given only for the symmetric formulation $a_s(\cdot,\cdot)$; the non-symmetric case of $a_{ns}(\cdot,\cdot)$ can be handled analogously without any difficulty. For simplicity of notation, we use "$\lesssim$" to denote "less than or equal to up to a constant independent of the mesh size, variables, or other parameters appearing in the inequality".

On each edge e, we introduce the following "jumps":

$$\mathbf{J}_1(\nabla\mathbf{u}_h\cdot\mathbf{n} - p_h\mathbf{n}) = \begin{cases} [\nabla\mathbf{u}_h\mathbf{n}] - [p_h\mathbf{n}], & \text{if } e \in \mathcal{E}_h^0, \\ \mathbf{0}, & \text{otherwise,} \end{cases}$$

and

$$\mathbf{J}_2(\mathbf{u}_h) = \begin{cases} [\![\mathbf{u}_h]\!], & \text{if } e \in \mathcal{E}_h^0, \\ 2\mathbf{u}_h \otimes \mathbf{n}, & \text{otherwise.} \end{cases}$$

Define a local error estimator on each element $K \in \mathcal{T}_h$ by

$$\eta_K^2 = h_K^2\|\mathbf{f}_h + \Delta\mathbf{u}_h - \nabla p_h\|_K^2 + \frac{1}{2}\sum_{e\in\partial K}\left(h_e\|\mathbf{J}_1(\nabla\mathbf{u}_h\cdot\mathbf{n} - p_h\mathbf{n})\|_e^2 + h_e^{-1}\|\mathbf{J}_2(\mathbf{u}_h)\|_e^2\right), \tag{4.1}$$

and define a global error estimator $\eta^2 = \sum_{K\in\mathcal{T}_h} \eta_K^2$. Here and in what follows of this paper, $\mathbf{f}_h$ is the L^2 projection of the load function $\mathbf{f}$ into the velocity space defined locally on each element. It will be seen that $\mathbf{f}_h$ can be a projection of f into any polynomial space defined on each individual element K.

For any $K \in \mathcal{T}_K$ and one of its edges e, it can be proved by using the trace theorem and the scaling argument that for every $q \in H^1(K)$, we have the following estimate (see Theorem 3.10 in [1] or Equation (2.4) in [2])

$$\|q\|_e^2 \lesssim h_K^{-1}\|q\|_K^2 + h_K\|\nabla q\|_K^2. \tag{4.2}$$

Another useful estimate can be stated as follows.

LEMMA 4.1. *Let* $\mathbf{v} \in V_h$, *then*

$$h_e\|[\nabla \mathbf{v}\,\mathbf{t}]\|_e^2 \lesssim h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2. \tag{4.3}$$

Proof. First, let e be an interior edge shared by triangles K_1 and K_2 in $\mathcal{T}_h$. On edge e, define $\mathbf{q} = \mathbf{v}|_{K_1} - \mathbf{v}|_{K_2}$. Then, by the inverse inequality, we have

$$h_e\|[\nabla \mathbf{v}\mathbf{t}]\|_e^2 = h_e\|\mathbf{q}'\|_e^2 \lesssim h_e^{-1}\|\mathbf{q}\|_e^2 = h_e^{-1}\|[\![\mathbf{v}]\!]\|_e^2.$$

The proof for boundary edges is similar. ∎

4.1. Reliability of the estimator. Let $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$ and $\epsilon = p - p_h$. It has been proved in [8] that every matrix-valued function in $(L^2(\Omega))^4$, and hence $\nabla_h\mathbf{e}$, admits the following decomposition:

$$\nabla_h\mathbf{e} = \nabla\mathbf{r} - qI + \mathbf{curl}\,\mathbf{s}, \tag{4.4}$$

where $\mathbf{r} \in H_0^1(\Omega)^2$ is divergence-free, $\mathbf{s} \in H^1(\Omega)^2$, $q \in L_0^2(\Omega)$, and I is the 2×2 identity matrix. Furthermore, the following bound holds [8]

$$\|\mathbf{r}\|_1 + \|\mathbf{s}\|_1 \lesssim \|\nabla_h\mathbf{e}\|_\Omega. \tag{4.5}$$

Since $\mathbf{e}$ is exactly divergence-free, it follows that

$$(\nabla_h\mathbf{e},\ qI) = (\nabla\cdot\mathbf{e}, q) = 0. \tag{4.6}$$

Therefore, we have from (4.4) and (4.6) that

$$(\nabla_h\mathbf{e},\ \nabla_h\mathbf{e}) = (\nabla_h\mathbf{e},\ \nabla\mathbf{r}) + (\nabla_h\mathbf{e},\ \mathbf{curl}\,\mathbf{s}). \tag{4.7}$$

For any vector-valued function $\mathbf{v} \in (H_0^1(\Omega))^2$, denote by $\mathbf{v}_I$ the Clément type interpolation onto continuous piecewise linears on $\mathcal{T}_h$, preserving the homogeneous boundary condition; details for such an interpolation can be found from [17]. When RT_k or BDM_k elements, with $k \geq 1$, are used in the numerical discretization scheme, we see that the above mentioned interpolation satisfies $\mathbf{v}_I \in (H_0^1(\Omega))^2 \cap V_h$. Clearly, the jump term $[\![\mathbf{v}_I]\!] = \mathbf{0}$ vanishes on every $e \in \mathcal{E}_h$. Furthermore, we have the following approximation property:

$$\begin{aligned} |\mathbf{v}_I|_{1,K} &\lesssim |\mathbf{v}|_{1,K}, \\ \|\mathbf{v} - \mathbf{v}_I\|_K &\lesssim h_K|\mathbf{v}|_{1,K}, \\ \|\mathbf{v} - \mathbf{v}_I\|_e &\lesssim h_e^{1/2}\|\mathbf{v}\|_{1/2,e}. \end{aligned} \tag{4.8}$$

THEOREM 4.2. *Let $(\mathbf{u};p)$ and $(\mathbf{u}_h;p_h)$ be the solution of (1.1)-(1.3) and (3.1)-(3.2), respectively. Then we have the following upper bound for $\|\nabla_h \mathbf{e}\|$:*

$$\|\nabla_h \mathbf{e}\| \lesssim \eta + (\sum_{K\in\mathcal{T}_h} h_K^2 \|\mathbf{f}-\mathbf{f}_h\|_K^2)^{1/2}. \tag{4.9}$$

Proof. Recall the decomposition (4.4). By setting $\mathbf{v} = \mathbf{r}_I$ in the error equation (3.5) and using $[\![\mathbf{r}_I]\!] = \mathbf{0}$, we have

$$(\nabla_h \mathbf{e},\ \nabla \mathbf{r}_I) - \sum_{e\in\mathcal{E}_h} \int_e \{\nabla \mathbf{r}_I\} : [\![\mathbf{u}_h]\!] - (\nabla\cdot \mathbf{r}_I,\ \epsilon) = 0. \tag{4.10}$$

Using (4.10), integration by parts, and the fact that $\mathbf{r}$ is divergence-free, we have

$$\begin{aligned}
&(\nabla_h \mathbf{e},\ \nabla \mathbf{r}) = (\nabla_h \mathbf{e},\ \nabla \mathbf{r} - \nabla \mathbf{r}_I) + \sum_{e\in\mathcal{E}_h} \int_e \{\nabla \mathbf{r}_I\} : [\![\mathbf{u}_h]\!] ds + (\nabla\cdot \mathbf{r}_I,\ \epsilon) \\
&= \sum_{K\in\mathcal{T}_h} (-\Delta \mathbf{u} + \Delta \mathbf{u}_h, \mathbf{r}-\mathbf{r}_I)_K + \sum_{K\in\mathcal{T}_h} \int_{\partial K} (\nabla \mathbf{e}\,\mathbf{n})\cdot(\mathbf{r}-\mathbf{r}_I) ds \\
&\quad + \sum_{e\in\mathcal{E}_h} \int_e \{\nabla \mathbf{r}_I\} : [\![\mathbf{u}_h]\!] ds - (\nabla\cdot(\mathbf{r}-\mathbf{r}_I),\ \epsilon) \\
&= \sum_{K\in\mathcal{T}_h} (-\Delta \mathbf{u} + \Delta \mathbf{u}_h, \mathbf{r}-\mathbf{r}_I)_K - \sum_{e\in\mathcal{E}_h^0} \int_e [\nabla \mathbf{u}_h \mathbf{n}]\cdot(\mathbf{r}-\mathbf{r}_I) ds + \sum_{e\in\mathcal{E}_h} \int_e \{\nabla \mathbf{r}_I\} : [\![\mathbf{u}_h]\!] ds \\
&\quad + \sum_{K\in\mathcal{T}_h} (\nabla p - \nabla p_h,\ \mathbf{r}-\mathbf{r}_I)_K - \sum_{K\in\mathcal{T}_h} \int_{\partial K} (\mathbf{r}-\mathbf{r}_I)\cdot\mathbf{n}(p-p_h)\ ds \\
&= \sum_{K\in\mathcal{T}_h} (\mathbf{f} + \Delta \mathbf{u}_h - \nabla p_h,\ \mathbf{r}-\mathbf{r}_I)_K - \sum_{e\in\mathcal{E}_h^0} \int_e ([\nabla \mathbf{u}_h \mathbf{n}] - [p_h \mathbf{n}])\cdot(\mathbf{r}-\mathbf{r}_I) ds \\
&\quad + \sum_{e\in\mathcal{E}_h} \int_e \{\nabla \mathbf{r}_I\} : [\![\mathbf{u}_h]\!] ds
\end{aligned}$$

Applying equations (4.2), (4.5), (4.8), the trace theorem, and the standard inverse inequality to the above gives

$$|(\nabla_h \mathbf{e},\ \nabla \mathbf{r})| \lesssim \left(\eta + (\sum_{K\in\mathcal{T}_h} h_K^2 \|\mathbf{f}-\mathbf{f}_h\|_K^2)^{1/2}\right) \|\nabla_h \mathbf{e}\|.$$

Next, it follows from the integration by parts that

$$\begin{aligned}
(\nabla_h \mathbf{e},\ &\mathbf{curl}\ \mathbf{s}) = -(\nabla_h \mathbf{u}_h,\ \mathbf{curl}\ \mathbf{s}) \\
&= -(\nabla_h \mathbf{u}_h,\ \mathbf{curl}\ (\mathbf{s}-\mathbf{s}_I)) - (\nabla_h \mathbf{u}_h,\ \mathbf{curl}\ \mathbf{s}_I) \\
&= \sum_{K\in\mathcal{T}_h} \int_{\partial K} (-(\nabla \mathbf{u}_h\,\mathbf{t})\cdot(\mathbf{s}-\mathbf{s}_I) - \mathbf{u}_h\cdot(\mathbf{curl}\ \mathbf{s}_I\,\mathbf{n}))\, ds \\
&= \sum_{e\in\mathcal{E}_h} \int_e (-[\nabla \mathbf{u}_h \mathbf{t}]\cdot(\mathbf{s}-\mathbf{s}_I) - [\![\mathbf{u}_h]\!] : \{\mathbf{curl}\ \mathbf{s}_I\})\, ds.
\end{aligned}$$

Thus, by using Lemma 4.1, we obtain

$$|(\nabla_h \mathbf{e}, \mathbf{curl}\, \mathbf{s})| \lesssim \left(\left(\sum_{e\in\mathcal{E}_h} h_e \|[\nabla \mathbf{u}_h \mathbf{t}]\|_e^2 \right)^{1/2} + \left(\sum_{e\in\mathcal{E}_h} h_e^{-1} \|[\![\mathbf{u}_h]\!]\|_e^2 \right)^{1/2} \right) \|\nabla_h \mathbf{e}\|$$
$$\lesssim \eta \, \|\nabla_h \mathbf{e}\|.$$

Combining (4.7) with the above estimates completes a proof of the lemma. □

As to the pressure error, we have the following result.

THEOREM 4.3. *Let* $(\mathbf{u}; p)$ *and* $(\mathbf{u}_h; p_h)$ *be the solution of (1.1)-(1.3) and (3.1)-(3.2), respectively. Then we have the following upper bound for* $\|\epsilon\|$*:*

$$\|\epsilon\| \lesssim \eta + (\sum_{K\in\mathcal{T}_h} h_K^2 \|\mathbf{f} - \mathbf{f}_h\|_K^2)^{1/2}. \tag{4.11}$$

Proof. First, we recall the following continuous version of the inf-sup condition:

$$\|p - p_h\| \lesssim \sup_{\mathbf{v}\in[H_0^1(\Omega)]^2} \frac{b(\mathbf{v}, p - p_h)}{\|\mathbf{v}\|_1}. \tag{4.12}$$

For $\mathbf{v} \in H_0^1(\Omega)^2$, using integration by parts, equations (3.1), (4.2), (4.5), (4.8), (4.9), the trace theorem, and the inverse inequality, we have

$$\begin{aligned}
&(\nabla\cdot\mathbf{v}, p - p_h) = -(\mathbf{f}, \mathbf{v}) + (\nabla\mathbf{u}, \nabla\mathbf{v}) - (\nabla\cdot\mathbf{v}, p_h)\\
=& -(\mathbf{f}, \mathbf{v}) + (\nabla_h\mathbf{e}, \nabla\mathbf{v}) + (\nabla_h\mathbf{u}_h, \nabla\mathbf{v}) - (\nabla\cdot\mathbf{v}, p_h)\\
&+ (\mathbf{f}, \mathbf{v}_I) - a(\mathbf{u}_h, \mathbf{v}_I) + b(\mathbf{v}_I, p_h)\\
=& -(\mathbf{f}, \mathbf{v}) + (\nabla_h\mathbf{e}, \nabla\mathbf{v}) + (\nabla_h\mathbf{u}_h, \nabla\mathbf{v}) - (\nabla\cdot\mathbf{v}, p_h)\\
&+ (\mathbf{f}, \mathbf{v}_I) - (\nabla_h\mathbf{u}_h, \nabla\mathbf{v}_I) + \sum_{e\in\mathcal{E}_h}\int_e \{\nabla\mathbf{v}_I\} : [\![\mathbf{u}_h]\!] ds + (\nabla\cdot\mathbf{v}_I, p_h)\\
=& -(\mathbf{f}, \mathbf{v} - \mathbf{v}_I) + (\nabla_h\mathbf{e}, \nabla\mathbf{v}) - \sum_{K\in\mathcal{T}_h} (\Delta\mathbf{u}_h, \mathbf{v} - \mathbf{v}_I)_K + \sum_{K\in\mathcal{T}_h} (\nabla p_h,\ \mathbf{v} - \mathbf{v}_I)_K\\
&+ \sum_{K\in\mathcal{T}_h}\int_{\partial K} (\nabla\mathbf{u}_h\,\mathbf{n} - p_h\mathbf{n})\cdot(\mathbf{v} - \mathbf{v}_I) ds + \sum_{e\in\mathcal{E}_h}\int_e \{\nabla\mathbf{v}_I\} : [\![\mathbf{u}_h]\!] ds\\
=& -\sum_{K\in\mathcal{T}_h} (\mathbf{f} + \Delta\mathbf{u}_h - \nabla p_h, \mathbf{v} - \mathbf{v}_I) + (\nabla_h\mathbf{e}, \nabla\mathbf{v})\\
&+ \sum_{e\in\mathcal{E}_h^0}\int_e ([\nabla\mathbf{u}_h]_{\mathbf{n}} - [p_h])\cdot(\mathbf{v} - \mathbf{v}_I) ds + \sum_{e\in\mathcal{E}_h}\int_e \{\nabla\mathbf{v}_I\} : [\![\mathbf{u}_h]\!] ds.
\end{aligned}$$

Thus, we obtain from the standard Cauchy-Schwarz inequality that

$$|(\nabla\cdot\mathbf{v}, p - p_h)| \lesssim \left(\eta + (\sum_{K\in\mathcal{T}_h} h_K^2 \|\mathbf{f} - \mathbf{f}_h\|_K^2)^{1/2} \right) \|\mathbf{v}\|_1,$$

which, together with the inequality (4.12), completes the proof of Theorem 4.3. □

By the definition of $\mathbf{f}_h$ and the approximation property of L^2 projection, it is not hard to see that $(\sum_{K\in\mathcal{T}_h} h_K^2 \|\mathbf{f} - \mathbf{f}_h\|_K^2)^{1/2}$ has higher order in h_K than $\|\nabla_h\mathbf{e}\| + \|\epsilon\|$, as long as $\mathbf{f}$ is smooth enough. Hence, theorems 4.2 and 4.3 imply that the a posteriori error estimator η is reliable in that the error term $\|\nabla_h\mathbf{e}\| + \|\epsilon\|$ must be small when η is small. Moreover, the former is controlled by the latter in their magnitude.

4.2. Efficiency of the estimator. For each triangular element $K \in \mathcal{T}_h$, denote by ϕ_K the following bubble function

$$\phi_K = \begin{cases} 27\,\lambda_1\lambda_2\lambda_3 & \text{in } K, \\ 0 & \text{in } \Omega\backslash K, \end{cases}$$

where λ_i, $i = 1; 2; 3$, are barycentric coordinates on K. Needless to say, the finite element partition $\mathcal{T}_h$ is assumed to contain triangular elements only. It is clear that $\phi_K \in H_0^1(\Omega)$ and satisfies the following properties (Section I.2.12 in [19]):

- For any polynomial q with degree at most m, there exist positive constants c_m and C_M, depending only on m, such that

$$c_m\|q\|_K^2 \le \int_K q^2\phi_K\,dx \le \|q\|_K^2, \tag{4.13}$$

$$\|\nabla(q\phi_K)\|_K \le C_m h_K^{-1}\|q\|_K. \tag{4.14}$$

For each $e \in \mathcal{E}_h^0$, we can analogously define an edge bubble function ϕ_e. Let K_1 and K_2 be two triangles sharing the edge e. To this end, denote by $\Omega_e = K_1 \cup K_2$ the union of the elements K_1 and K_2. Assume that in K_i, $i = 1, 2$, the barycentric coordinates associated with the two ends of e are $\lambda_1^{K_i}$ and $\lambda_2^{K_i}$, respectively. The edge bubble function can be defined as follows

$$\phi_e = \begin{cases} 4\lambda_1^{K_1}\lambda_2^{K_1} & \text{in } K_1, \\ 4\lambda_1^{K_2}\lambda_2^{K_2} & \text{in } K_2, \\ 0 & \text{in } \Omega\backslash\Omega_e. \end{cases}$$

It is obviously that $\phi_e \in H_0^1(\Omega)$ and satisfies the following properties (Section I.2.12 in [19]):

- For any polynomial q with degree at most m, there exist positive constants d_m, D_M and E_m, depending only on m, such that

$$d_m\|q\|_e^2 \le \int_e q^2\phi_e\,ds \le \|q\|_e^2, \tag{4.15}$$

$$\|\nabla(q\phi_e)\|_{\Omega_e} \le D_m h_e^{-1/2}\|q\|_e, \tag{4.16}$$

$$\|q\phi_e\|_{\Omega_e} \le E_m h_e^{1/2}\|q\|_e. \tag{4.17}$$

Regarding the efficiency of the numerical schemes under consideration, we have the following result.

THEOREM 4.4. *For every $K \in \mathcal{T}_h$ or $e \in \mathcal{E}_h^0$,*

$$h_K\|\mathbf{f}_h + \Delta\mathbf{u}_h - \nabla p_h\|_K \lesssim \|\nabla\mathbf{e}\|_K + \|\epsilon\|_K + h_K\|\mathbf{f} - \mathbf{f}_h\|_K, \tag{4.18}$$

$$h_e^{1/2}\|[\nabla\mathbf{u}_h\mathbf{n}] - [p_h\mathbf{n}]\|_e \lesssim \sum_{K\in\Omega_e}\left(\|\nabla\mathbf{e}\|_K + \|\epsilon\|_K + h_K\|\mathbf{f} - \mathbf{f}_h\|_K\right). \tag{4.19}$$

Proof. Let $\mathbf{w}_K = (\mathbf{f}_h + \Delta\mathbf{u}_h - \nabla p_h)\phi_K$. It is clear that

$$(\mathbf{f},\ \mathbf{w}_K) = (\nabla\mathbf{u},\ \nabla\mathbf{w}_K) - (\nabla\cdot\mathbf{w}_K,\ p),$$

which implies

$$\begin{aligned}(\mathbf{f}-\mathbf{f}_h,\ \mathbf{w}_K)+(\mathbf{f}_h,\ \mathbf{w}_K)-(\nabla\mathbf{u}_h,\ \nabla\mathbf{w}_K)+(\nabla\cdot\mathbf{w}_K,\ p_h)\\ =(\nabla\mathbf{e},\ \nabla\mathbf{w}_K)-(\nabla\cdot\mathbf{w}_K,\ \epsilon).\end{aligned}$$

Using integration by parts, the above equation becomes

$$(\mathbf{f}_h+\Delta\mathbf{u}_h-\nabla p_h,\ \mathbf{w}_K)=(\nabla\mathbf{e},\ \nabla\mathbf{w}_K)-(\nabla\cdot\mathbf{w}_K,\ \epsilon)-(\mathbf{f}-\mathbf{f}_h,\ \mathbf{w}_K).$$

By the definition of $\mathbf{w}_K$, the Cauchy-Schwartz inequality, and inequalities (4.13)-(4.14), we have

$$\begin{aligned}\|\mathbf{f}_h+\Delta\mathbf{u}_h-\nabla p_h\|_K^2 &\lesssim |(\nabla\mathbf{e},\ \nabla\mathbf{w}_K)-(\nabla\cdot\mathbf{w}_K,\ \epsilon)-(\mathbf{f}-\mathbf{f}_h,\ \mathbf{w}_K)|\\ &\lesssim \left(h_K^{-1}\|\nabla\mathbf{e}\|_K+h_K^{-1}\|\epsilon\|_K+\|\mathbf{f}-\mathbf{f}_h\|_K\right)\|\mathbf{f}_h+\Delta\mathbf{u}_h-\nabla p_h\|_K.\end{aligned}$$

This completes the proof of inequality (4.18).

As to an estimate for the jump term $\|[\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}]\|_e$, we consider a similar function $\mathbf{w}_e$ defined by

$$\mathbf{w}_e=([\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}])\phi_e.$$

Here, in the multiplication, the function $[\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}]$ should be extended as a constant function in the normal direction of the edge e; i.e., $\mathbf{w}_e$ takes the polynomial $[\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}]$ on e and all lines parallel with e. It is not hard to see that

$$\begin{aligned}\sum_{K\in\Omega_e}&((\mathbf{f},\ \mathbf{w}_e)_K-(\nabla\mathbf{u}_h,\nabla\mathbf{w}_e)_K+(\nabla\cdot\mathbf{w}_e,p_h))\\ &=\sum_{K\in\Omega_e}\left((\nabla\mathbf{e},\nabla\mathbf{w}_e)_K-(\nabla\cdot\mathbf{w}_e,\epsilon)_K\right).\end{aligned}$$

Using integration by parts, we have

$$\begin{aligned}\sum_{K\in\Omega_e}&((\mathbf{f}+\Delta\mathbf{u}_h-\nabla p_h,\ \mathbf{w}_e)_K-\int_e([\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}])\cdot\mathbf{w}_e\,ds\\ &=\sum_{K\in\Omega_e}\left((\nabla\mathbf{e},\nabla\mathbf{w}_e)_K-(\nabla\cdot\mathbf{w}_e,\epsilon)_K\right).\end{aligned}$$

Therefore, by the definition of $\mathbf{w}_e$, the Cauchy-Schwartz inequality, and inequalities (4.15)-(4.17), we have

$$\begin{aligned}\|[\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}]\|_e^2\lesssim(&h_e^{1/2}\|\mathbf{f}-\mathbf{f}_h\|_{\Omega_e}+h_e^{1/2}\|\mathbf{f}_h+\Delta\mathbf{u}_h-\nabla p_h\|_{\Omega_e}\\ &+h_e^{-1/2}\|\nabla_h\mathbf{e}\|_{\Omega_e}+h_e^{-1/2}\|\epsilon\|_{\Omega_e})\|[\nabla\mathbf{u}_h\mathbf{n}]-[p_h\mathbf{n}]\|_e.\end{aligned}$$

Inequality (4.19) follows from the above, inequality (4.18), and the fact that $\mathcal{T}_h$ is shape regular. This completes the proof of the theorem. □

Summing over all the elements yields the following lower bound for the error term $|||\mathbf{e}|||+\|\epsilon\|$.

THEOREM 4.5. *Let $(\mathbf{u};p)$ and $(\mathbf{u}_h;p_h)$ be the solution of (1.1)-(1.3) and (3.1)-(3.2), respectively. Then we have the following lower bound estimate:*

$$\eta\lesssim|||\mathbf{e}|||+\|\epsilon\|+\left(\sum_{K\in\mathcal{T}_h}h^2\|\mathbf{f}-\mathbf{f}_h\|_K^2\right)^{1/2}.$$

Combining theorems 4.2, 4.3 and 4.5, and noticing that $(\sum_{K\in\mathcal{T}_h} h_K^2\|\mathbf{f}-\mathbf{f}_h\|_K^2)^{1/2}$ is a higher order term, we can conclude that, theoretically, η is a good indicator for $|||\mathbf{e}||| + \|\epsilon\|$.

5. Two strategies in local grid refinement. The local a priori error estimator η_K as defined in (4.1) can be used to provide algorithms for local grid refinement. Two different refinement strategies are considered in this study. The first one is based on a comparison of each error η_K with the maximum value of all the error estimators. The strategy can be described as follow.

Local Refinement by "Maximum Strategy":

1. Given a current triangular mesh, error estimators η_K on each triangle, and a threshold $\theta \in (0,1)$ (e.g., $\theta = 0.5$). One computes the maximum error $\eta_{max} = \max \eta_K$.
2. For each triangle K, if $\eta_K \geq \theta\, \eta_{max}$, refine this triangle uniformly by connecting the center of three edges.
3. The previous step will generate "hanging nodes". Use bisection to get a conformal mesh. More precisely, one needs to check every unrefined triangle K and perform the following modifications:
   - If K has one "hanging node", then bisect it once.
   - If K has two "hanging nodes", then bisect it twice.
   - Take a special care to prevent the occurrence of degenerated triangles; i.e., to guarantee that the new mesh preserves the shape regularity. This can be done by adding extra "hanging nodes" if the current bisection results in degenerated triangles.

The second refinement strategy is based on a comparison of η_K with those for its neighbors. To explain the main idea, let $\rho > 0$ be a prescribed distance parameter and set

$$\mathcal{T}_{\rho,K} = \{\tilde{K} :\ 0 < \|K - \tilde{K}\| \leq \rho\},$$

where $\|K - \tilde{K}\|$ stands for the distance of the centers of K and $\tilde{K}$. With a given threshold $\theta > 1$, we mark a triangle K for uniform refinement if

$$\eta_K \geq \theta\, \eta_{N(K)},$$

where $\eta_{N(K)}$ is the average of the local error indicator on all the neighboring triangles $\tilde{K} \in \mathcal{T}_{\rho,K}$.

The following is such a refinement strategy that was numerically investigated in this study.

Local Refinement by "Local Strategy":

1. Given a current triangular mesh, error estimators η_K on each triangle, and a threshold $\theta > 1.0$ (e.g., $\theta = 1.3$). One computes an error indicator $\eta_{N(K)}$ as the average of the local error indicator on neighboring triangles that share a vertex or an edge with K, not including K itself.
2. For each triangle K, if $\eta_K \geq \theta\, \eta_{N(K)}$, refine this triangle uniformly by connecting the center of three edges.
3. The previous step will generate "hanging nodes". Use bisection to get a conformal mesh. More precisely, one needs to check each unrefined triangle K and perform the following modifications:

- If K has one "hanging node", then bisect it once.
- If K has two "hanging nodes", then bisect it twice.
- Take a special care to prevent the occurrence of degenerated triangles; i.e., to guarantee that the new mesh preserves the shape regularity. This can be done by adding extra "hanging nodes" if the current bisection results in degenerated triangles.

The refinement method by the "local strategy" is based on the observation that local estimators on the exact location of the singularity tend to be much larger than on surrounding regions. Since the singularity usually occurs on a lower dimensional manifold, for example, a point or an edge in two dimension, it is safe to use such a local strategy to locate singularities. In our numerical experiments to be presented in the coming section, we have found that $\theta = 1.3$ is a practical choice.

6. Some numerical results. The goal of this section is to report some numerical results for the refinement strategies as discussed in the previous section. Without loss of generality, we discuss only numerical results arising from the non-symmetric formulation with $a_{ns}(\cdot,\cdot)$. As mentioned before, the non-symmetric formulation is well-posed for any $\alpha > 0$, while the symmetric formulation is so only for α sufficiently large. Although numerical tests in [24] have indicated that the symmetric form works for α greater than a moderate number (usually between 1 and 5), they have also suggested that the non-symmetric formulation is more stable than the symmetric one with respect to the selection of α values (see examples given in [24]).

For the numerical results to be presented in this section, the parameter α is set to be 5, and the Stokes equations are discretized by the BDM_1 element for the velocity. The GMRES iterative method is used for solving the resulting linear algebraic system, and a relative residual of 10^{-8} is set to be the stopping criteria. It should be pointed out that, for the test problems and the value of α chosen in this paper, numerical results show that the a posteriori error indicators for both the symmetric and non-symmetric formulations behave almost exactly the same. Hence the results for the symmetric formulation are omitted. However, we did not explore the cases when α is too small or too large.

The numerical formulation in this paper can be easily extended to problems with non-homogeneous Dirichlet boundary conditions; details have been given in [24]. Accordingly, changes need to be made for the a posteriori error estimator η_K. For the Dirichlet boundary condition $\mathbf{u} = \mathbf{g}$ on $\partial\Omega$, we must modify $\mathbf{J}_2(\mathbf{u}_h)$ as follows:

$$\mathbf{J}_2(\mathbf{u}_h) = \begin{cases} [\![\mathbf{u}_h]\!] & \text{if } e \in \mathcal{E}_h^0, \\ (\mathbf{u}_h - \mathbf{g}) \otimes \mathbf{n} & \text{otherwise.} \end{cases}$$

Another small modification is that, in the definition of η_K, the term $h_K^2 \|\mathbf{f}_h + \Delta\mathbf{u}_h - \nabla p_h\|_K^2$ must be computed as $2|K|\,\|\mathbf{f}_h + \Delta\mathbf{u}_h - \nabla p_h\|_K^2$, where $|K|$ is the area of the triangle K. This modification is purely for implementation purpose since the two formulas are equivalent mathematically. For the BDM_1 element, we also have $\Delta\mathbf{u}_h - \nabla p_h = 0$. In the experiment, $\|\mathbf{f}_h\|_K$ is calculated by computing $\|\mathbf{f}\|_K$ using a high order numerical integration, which is exact for up to 7th order polynomials. The effect is equivalent to considering $\mathbf{f}_h$ to be a piecewise cubic interpolation of $\mathbf{f}$.

6.1. Test problems. We consider three test problems, all are defined on $\Omega = (0,1) \times (0,1)$. Two of them have exact solutions given as:

$$\textbf{Test problem 1:} \quad \mathbf{u} = \begin{pmatrix} -2x^2y(x-1)^2(2y-1)(y-1) \\ xy^2(2x-1)(x-1)(y-1)^2 \end{pmatrix}, \quad p = 0,$$

and

$$\textbf{Test problem 2:} \quad \mathbf{u} = \begin{pmatrix} \frac{3}{2}\sqrt{r}\left(\cos\frac{\theta}{2} - \cos\frac{3\theta}{2}\right) \\ \frac{3}{2}\sqrt{r}\left(3\sin\frac{\theta}{2} - \sin\frac{3\theta}{2}\right) \end{pmatrix}, \qquad p = -6r^{-1/2}\cos\frac{\theta}{2}.$$

The third one is a lid driven cavity problem. Clearly, the solution of test problem 1 is smooth and satisfies the homogeneous Dirichlet boundary condition. The solution to test problem 2 satisfies

$$-\Delta\mathbf{u} + \nabla p = \mathbf{0}.$$

The Dirichlet boundary condition can be set by using the value of $\mathbf{u}$ on the boundary. Test problem 2 has a corner singularity of order 0.5 at the origin $(0, 0)$.

The 2D lid driven cavity problem describes the flow in a rectangular container which is driven by the uniform motion of the top lid [18]. Because of the discontinuous velocity boundary condition at two top corners, it is know that the exact solution $(\mathbf{u}; p)$ does not even belong to $(H^1)^2 \times L^2$, if there is any in a certain sense. Indeed, the weak formulation does not hold for this problem. However, the discrete problem is still well-posed and provides a certain approximation to the actual solution. In two dimensional case, the discontinuous boundary condition also results in corner singularities at the two top corners.

6.2. On uniform meshes. In this experiment, test problems 1 and 2 are solved on uniform meshes. The mesh is generated by dividing Ω into $n \times n$ sub-rectangles, and then dividing each sub-rectangle into two triangles by connecting the diagonal lines with the negative slope. For test problem 1, since the solution is smooth, we expect theoretically that $\eta = O(h)$, $\|\nabla\mathbf{e}\| = O(h)$, $\|\epsilon\| = O(h)$, and $\|\mathbf{e}\| = O(h^2)$. For test problem 2, only η and $\|\mathbf{e}\|$ are computed for each mesh. Due to the corner singularity, we expect them to have asymptotic order between $0 \sim 1$ and $1 \sim 2$, respectively. Numerical results for these two test problems are reported in Tables 6.1 and 6.2. They agree with the theoretical prediction.

TABLE 6.1
Error norms for test problem 1 on $n \times n$ uniform triangular meshes.

n	η	$\|\nabla_h\mathbf{e}\|$	$\|\mathbf{e}\|$	$\|\epsilon\|$
20	4.7471e-02	7.3535e-3	7.2677e-05	6.4306e-03
24	3.9947e-02	6.1326e-3	5.0784e-05	5.4066e-03
28	3.4465e-02	5.2582e-3	3.7477e-05	4.6615e-03
32	3.0298e-02	4.6016e-3	2.8790e-05	4.0957e-03
36	2.7025e-02	4.0904e-3	2.2807e-05	3.6518e-03
40	2.4388e-02	3.6813e-3	1.8512e-05	3.2944e-03
44	2.2219e-02	3.3464e-3	1.5326e-05	3.0005e-03
48	2.0403e-02	3.0674e-3	1.2897e-05	2.7546e-03
52	1.8860e-02	2.8312e-3	1.1002e-05	2.5459e-03
Asym. Order $O(h^k)$, $k =$	0.9671	0.9991	1.9763	0.9707

6.3. Adaptive refinements for test problem 2. We perform adaptive refinements for test problem 2, which has a corner singularity. The two refinement strategies as explained in Section 5 are employed in this investigation. The errors

TABLE 6.2
Error norms for test problem 2 on $n \times n$ uniform triangular meshes.

n	dofs	GMRES iter.	η	$\|\mathbf{e}\|$
16	2112	1208	2.9080e+00	6.1029e-03
20	3280	1793	2.6112e+00	4.3923e-03
24	4704	2398	2.3893e+00	3.3533e-03
28	6384	3181	2.2138e+00	2.6698e-03
32	8320	4032	2.0685e+00	2.1894e-03
36	10512	4856	1.9438e+00	1.8528e-03
40	12960	5871	1.8384e+00	1.5957e-03
44	15664	6980	1.7417e+00	1.4233e-03
48	18624	8182	1.6540e+00	1.3460e-03
Asym. Order $O(h^k)$, $k =$	-1.9819	-1.8270	0.5113	1.4136

from the "Maximum Strategy" are reported in Table 6.3, and corresponding meshes are presented in Figure 6.1. The errors from the "Local Strategy" are reported in Table 6.4, and corresponding meshes are given in Figure 6.2. We also draw the results from tables 6.3-6.4 in Figure 6.3, to provide a direct image of the relation between η and the number of degrees of freedom. After eliminating several starting levels, the computed asymptotic orders of η in terms of the number of degrees of freedom are $O(N^{-0.5097})$ and $O(N^{-0.4518})$, respectively, for the maximum strategy and the local strategy.

By comparing Table 6.2, 6.3 and 6.4, it can be seen that the adaptive refinement is much more efficient in reducing the error for this problem. The mesh refinements shown in Figure 6.1 and 6.2 indicate that the local error indicator η_K has located the singularity accurately. It seems that the "Local Strategy" gives similar refinements to the "Maximum Strategy" for this test problem. The values of θ, when lying in a reasonable range, controls the refinement scale in both strategies.

TABLE 6.3
Test problem 2, adaptive refinement using the "Maximum strategy".

# Triangles	dofs	GMRES iter.	η	$\|\mathbf{e}\|$
128	544	360	4.0264	1.6785e-2
140	594	405	3.3045	9.8765e-3
152	644	454	2.9531	8.5347e-3
174	734	544	2.5339	6.7235e-3
191	805	635	2.3725	6.4171e-3
284	1180	969	1.9054	4.5336e-3
398	1642	1328	1.6162	4.1063e-3
1056	4296	3180	1.0208	2.2767e-3

We also draw the graph of the velocity and pressure on the finest mesh from the "Local Strategy" in Figure 6.4. They agree well with the graph of the exact solution. The graph of numerical solution using the "Maximum Strategy" is similar, and thus omitted.

It has been known that an averaging post-processing of the pressure usually improves the approximation for the pressure. Hence we would also like to check the

FIG. 6.1. *Test problem 2, adaptive refinement using the "Maximum strategy". Initial mesh and meshes after several refinements.*

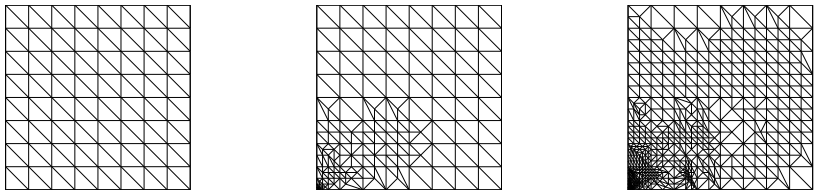


TABLE 6.4
Test problem 2, adaptive refinement using the "Local strategy", $\theta = 1.3$.

# Triangles	dofs	GMRES iter.	η	$\|\mathbf{e}\|$
128	544	360	4.0264	1.6785e-2
140	594	405	3.3045	9.8765e-3
178	748	532	2.6805	6.4625e-3
272	1130	834	2.1786	4.4451e-3
439	1803	1730	1.7596	3.5040e-3
820	3338	2537	1.3610	2.6289e-3

robustness of our adaptive mesh refinements under such a pressure-averaging process, for which the main idea can be described as follows.

- After a solution has been computed on the current mesh, for each triangle K in the mesh, identify a neighborhood of K including triangles sharing one vertex or one edge with K, including K itself.
- Sum up the products of the computed pressure, which is piecewise constant, with the area of triangles over all triangles in the neighborhood of K. Then divide the summation by the total area of the involved triangles. Assign this value to the post-processed pressure on K.

This pressure-averaging process is applied before one computes the error indicators. The rest of the adaptive mesh refinement remains unchanged.

We report the results for test problem 2 with pressure-averaging, using both maximum and local strategies, in Tables 6.5, 6.6 and Figures 6.5, 6.6, 6.7. We point out that after the pressure-averaging process, the a posteriori error η is different from the one without post-processing, while the L^2 norm of the error for the velocity $\|\mathbf{e}\|$ remains unchanged. Our numerical results show that the error indicator works well with the pressure-averaging method.

6.4. Adaptive refinements for the driven cavity problem. Since the exact solution for driven cavity problem is not even in $(H^1)^2 \times L^2$, we anticipate that η does not decrease when the mesh is refined. Our numerical experiments show that the local error indicator η_K is able to locate both corner singularities for this problem. The meshes are plotted in Figure 6.8, 6.9, 6.10, 6.11, and 6.12.

Notice that when using the "Local Strategy", changing the value of θ clearly affects the refinement dramatically. Indeed, small oscillations of local error indicators have been observed near the corner singularities such that the "Local Strategy" with $\theta = 1.3$ often locates an oscillation instead of the singularity. We pointed out that such oscillations might be inherited from the numerical discretization scheme due to

FIG. 6.2. *Test problem 2, adaptive refinement using the "Local strategy", $\theta = 1.3$. Initial mesh and meshes after several refinements.*

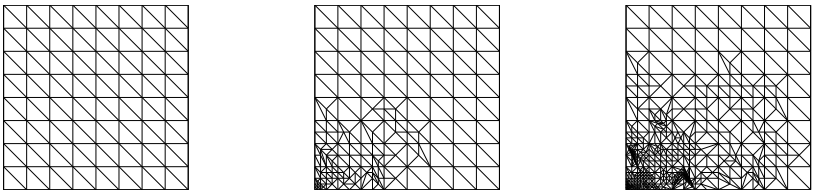
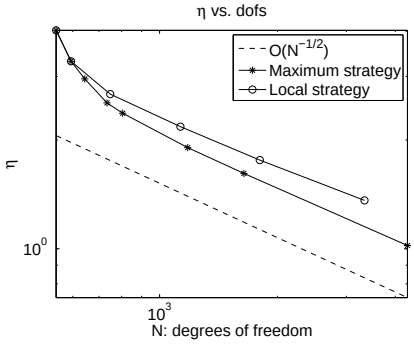


FIG. 6.3. *Test problem 2, adaptive refinement, relation between η and the number of degrees of freedom, using both maximum and local strategy.*



the severe singularity of the boundary conditions. By setting $\theta = 1.5$, the method can focus more on the real singularity, as shown in Figure 6.10.

Next, we draw the velocity and pressure profiles for the driven cavity problem computed by using different refinement strategies. The solutions in Tables 6.13, 6.14, 6.15 and 6.16 are drawn on the finest reported mesh of each adaptive strategy. We would like readers to draw conclusions from reading the results.

Finally, we compare the present solutions with the numerical solution as computed by using the popular Taylor-Hood finite element method. Figure 6.17 gives the velocity profiles, which contain the plot of u_1 on the vertical centerline of the domain and u_2 on the horizontal centerline. The solution of the Taylor-Hood element is computed on a 64×64 triangular mesh, and is denoted by stars in the figure. The velocity profile of the solutions from adaptive mesh refinements using different strategies are drawn in curves. From the plots, one can see that the solutions match very well with each other. This is of course a comparison of the solution at smooth areas. But it is evident that the proposed local refinement methods do provide competitive numerical solutions for the Stokes equation.

Appendix A. Equivalence of the weak formulation to the one stated in [22]. For simplicity, we only consider the case for the symmetric bilinear form $a_s(\cdot,\cdot)$. In [22], $a_s(\cdot,\cdot)$ is defined as following:

$$a_s(\mathbf{w},\mathbf{v}) = (\nabla_h \mathbf{w}, \nabla_h \mathbf{v}) + \sum_{e\in\mathcal{E}_h} \int_e \left(\alpha h_e^{-1} [\mathbf{w}\cdot\mathbf{t}]\,[\mathbf{v}\cdot\mathbf{t}] - \{\!\{\varepsilon(\mathbf{w})\}\!\}[\mathbf{v}\cdot\mathbf{t}] - \{\!\{\varepsilon(\mathbf{v})\}\!\}[\mathbf{w}\cdot\mathbf{t}]\right) ds,$$

FIG. 6.4. *Test problem 2, adaptive refinement using the "Local Strategy", $\theta = 1.3$. Velocity and pressure.*

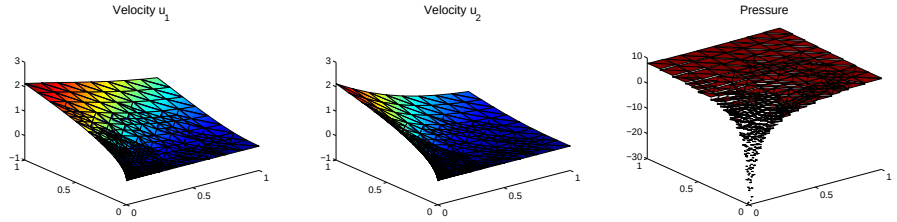


TABLE 6.5
Test problem 2, adaptive refinement using the "Maximum strategy", with pressure-averaging.

# Triangles	dofs	GMRES iter.	η	$\|\mathbf{e}\|$
128	544	360	2.0492	1.6785e-2
156	660	452	1.7587	8.4449e-3
184	776	547	1.5944	6.6796e-3
230	964	709	1.4127	5.4049e-3
306	1272	977	1.2289	4.7228e-3
434	1788	1397	1.0421	3.4815e-3
651	2663	2046	0.8330	2.8027e-3
1484	6026	4441	0.5546	2.2096e-3

where on internal edge e,

$$[\mathbf{v}\cdot\mathbf{t}] = \mathbf{v}|_{\partial K_1}\cdot\mathbf{t}_1 + \mathbf{v}|_{\partial K_2}\cdot\mathbf{t}_2,$$

and on boundary edges it only takes the value on one side. The definitions for V_h, W_h and $b(\cdot,\cdot)$ remain the same.

Therefore, we only need to show that for all $\mathbf{w}$, $\mathbf{v}\in V(h)$ and $e\in\mathcal{E}_h$,

$$[\![\mathbf{w}]\!] : [\![\mathbf{v}]\!] = [\mathbf{w}\cdot\mathbf{t}]\,[\mathbf{v}\cdot\mathbf{t}], \tag{A.1}$$

$$\{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!] = \{\!\{\varepsilon(\mathbf{w})\}\!\}[\mathbf{v}\cdot\mathbf{t}]. \tag{A.2}$$

For $\mathbf{v}\in V(h)$, denote

$$\delta\mathbf{v} = \begin{cases} \mathbf{v}|_{\partial K_1} - \mathbf{v}|_{\partial K_2} & \text{on internal edge } e,\\ \mathbf{v}|_e & \text{on boundary edge } e,\end{cases}$$

$$\bar{\mathbf{v}} = \begin{cases} \frac{1}{2}(\mathbf{v}|_{\partial K_1} + \mathbf{v}|_{\partial K_2}) & \text{on internal edge } e,\\ \mathbf{v}|_e & \text{on boundary edge } e.\end{cases}$$

Since vectors in $V(h)$ have continuous normal components across e, clearly $\delta\mathbf{v}\cdot\mathbf{n} = 0$ on internal edges. By the homogeneous boundary condition, we also know that both $\delta\mathbf{v}\cdot\mathbf{n}$ and $\bar{\mathbf{v}}\cdot\mathbf{n}$ vanish on boundary edges.

Therefore, on every $e\in\mathcal{E}_h$, we have

$$[\![\mathbf{w}]\!] : [\![\mathbf{v}]\!] = \delta\mathbf{w}\cdot\delta\mathbf{v} = (\delta\mathbf{w}\cdot\mathbf{t}_1)(\delta\mathbf{v}\cdot\mathbf{t}_1) = [\mathbf{w}\cdot\mathbf{t}]\,[\mathbf{v}\cdot\mathbf{t}],$$

and

$$\{\!\{\varepsilon(\mathbf{w})\}\!\}[\mathbf{v}\cdot\mathbf{t}] = (\nabla\bar{\mathbf{w}}\,\mathbf{n}_1)\cdot\delta\mathbf{v} = (\nabla\bar{\mathbf{w}}) : (\delta\mathbf{v}\otimes\mathbf{n}_1) = \{\nabla\mathbf{w}\} : [\![\mathbf{v}]\!].$$

FIG. 6.5. *Test problem 2, adaptive refinement using the "Maximum strategy", with pressure-averaging. Initial mesh and meshes after several refinements.*

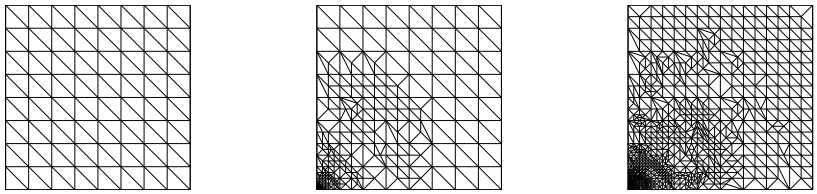


TABLE 6.6
Test problem 2, adaptive refinement using the "Local strategy", $\theta = 1.3$, with pressure-averaging.

# Triangles	dofs	GMRES iter.	η	$\|\mathbf{e}\|$
128	544	360	2.0492	1.6785e-2
138	584	397	1.6994	9.2067e-3
172	722	508	1.5129	6.4453e-3
250	1038	750	1.4051	4.8342e-3
370	1522	1129	1.2399	4.0372e-3
662	2698	2022	1.0375	2.9473e-3

Combining the above, the weak form given in in this paper is equivalent to the one from [22]. Hence Theorem 3.1 follows directly from [22].

REFERENCES

[1] S. AGMON, *Lectures on elliptic boundary problems*, Van Nostrand, Princeton, New Jersey, 1965.

[2] D.N. ARNOLD, AN INTERIOR PENALTY FINITE ELEMENT METHOD WITH DISCONTINUOUS ELEMENTS, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[3] F. BREZZI, J. DOUGLAS, AND L. MARINI, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.

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

[5] Z. CAI, C. TONG, P.S. VASSILEVSKI AND C. WANG, *Mixed finite element methods for incompressible flow: Stationary Stokes equations*, Numer. Meth. Part. Diff. Eq., 26 (2009), pp. 957–978.

[6] Z. CAI, C. WANG AND S. ZHANG, *Mixed Finite Element Methods for Incompressible Flow: Stationary NavierStokes Equations*, SIAM J. Numer. Anal., 48 (2010), pp. 79–94.

[7] C. CARSTENSEN AND S. FUNKEN *A posteriori error control in low-order finite element discretisations of incompressible stationary flow problems*, Math. Comp., 70 (2001), pp. 1353–1381.

[8] E. DARI, R. DURÁN AND C. PADRA *Error estimators for nonconforming finite element approximations of the Stokes problem*, Math. Comp., 64 (1995), pp. 1017–1033.

[9] W. DÖRFLER AND M. AINSWORTH *Reliable a posteriori error control for nonconforming finite element approximation of Stokes flow*, Math. Comp., 74 (2005), pp. 1599–1619.

[10] L. FIGUEROA, G.N. GATICA, A. MARQUEZ, *Augmented mixed finite element methods for the stationary Stokes equations*, SIAM J. Sci. Comput., 31 (2008), pp. 1082–1119.

[11] G.N. GATICA, A. MARQUEZ, AND M.A. SANCHEZ, *Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 1064–1079.

[12] J.S. HOWELL, *Dual-mixed finite element approximation of Stokes and nonlinear Stokes problems using trace free velocity gradients*, J. Comp. Appl. Math., 231 (2009), pp. 780–792.

[13] J.S. HOWELL, *Approximation of generalized Stokes problems using dual-mixed finite elements without enrichment*, Int. J. Numer. Meth. Fluids, (2010), DOI 10.1002/fld.2356.

[14] C. PADRA, *A posteriori error estimators for nonconforming approximation of some quasi-Newtonian flows*, SIAM J. Numer. Anal., 34 (1997), pp. 1600–1615.

FIG. 6.6. *Test problem 2, adaptive refinement using the "Local strategy", $\theta = 1.3$, with pressure-averaging. Initial mesh and meshes after several refinements.*

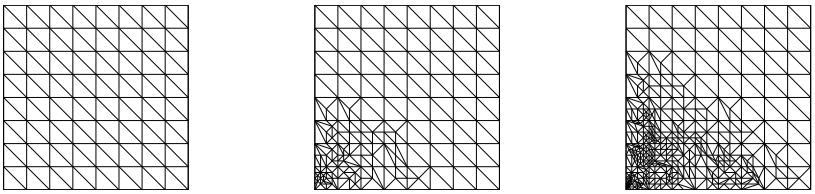
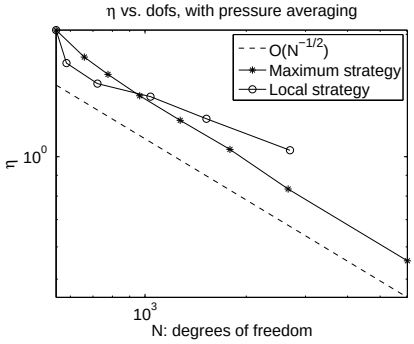


FIG. 6.7. *Test problem 2, adaptive refinement, relation between η and the number of degrees of freedom, with pressure-averaging, using both maximum and local strategy.*



[15] M. PARASCHIVOIU AND A.T. PATERA, *A posteriori bounds for linear functional outputs of Crouzeix-Raviart finite element discretizations of the incompressible Stokes problem*, Int. J. Numer. Methods Fluids, 32 (2000), pp. 823–849.

[16] P. RAVIART AND J. THOMAS, *A mixed finite element method for second order elliptic problems*, Mathematical Aspects of the Finite Element Method, I. Galligani, E. Magenes, eds., Lectures Notes in Math., Vol. 606, Springer-Verlag, New York, 1977.

[17] L.R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth function satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[18] P.N. SHANKAR AND M.D. DESHPANDE, *Fluid Mechanics in the driven cavity*, Annu. Rev. Fluid Mech., 32 (2000), pp. 93–136.

[19] R. VERFÜRTH, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Teubner Skripten zur Numerik. B.G. Willey-Teubner, Stuttgart, 1996.

[20] R. VERFÜRTH, *A posteriori error estimators for the Stokes equations*, Numer. Math., 55 (1989), pp. 309–325.

[21] R. VERFÜRTH, *A posteriori error estimators for the Stokes equations II. non-conforming discetizations*, Numer. Math., 60 (1991), pp. 235–249.

[22] J. WANG AND X. YE, *New finite element methods in computational fluid dynamics by H(div) elements*, SIAM J. Numer. Anal., 45 (2007), pp. 1269–1286.

[23] J. WANG, X. WANG AND X. YE, *Finite element methods for the Navier-Stokes equations by H(div) elements*, J. Comput. Math., 26 (2008), pp. 410–436.

[24] J. WANG, Y. WANG AND X. YE, *A robust numerical method for Stokes equations based on divergence-free H(div) finite element methods*, SIAM J. Sci. Comp., 31 (2009), pp. 2784–2802.

FIG. 6.8. *Driven cavity problem, adaptive refinement using the "Maximum strategy". Initial mesh and meshes after several refinements.*

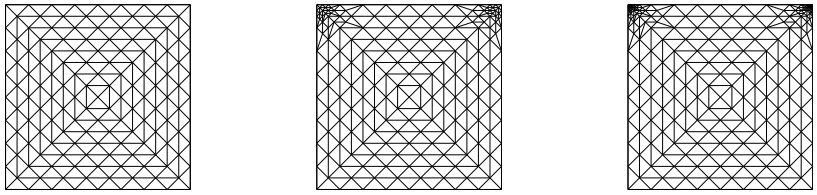


FIG. 6.9. *Driven cavity problem, adaptive refinement using the "Local strategy", $\theta = 1.3$. Initial mesh and meshes after several refinements.*

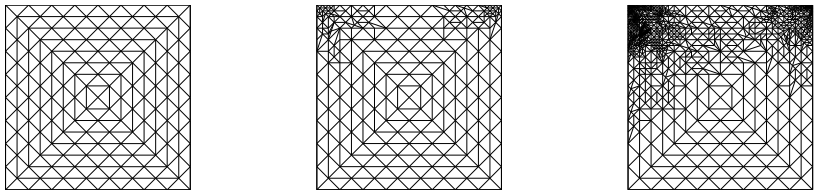


FIG. 6.10. *Driven cavity problem, adaptive refinement using the "Local strategy", $\theta = 1.5$. Initial mesh and meshes after each refinement.*

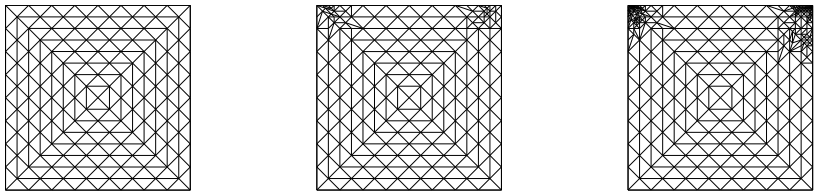


FIG. 6.11. *Driven cavity problem, adaptive refinement using the "Maximum strategy", with pressure-averaging. Initial mesh and meshes after several refinements.*

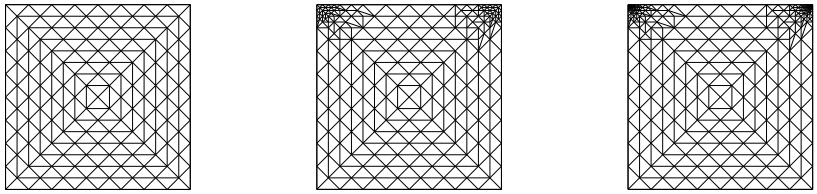


FIG. 6.12. *Driven cavity problem, adaptive refinement using the "Local strategy", $\theta = 1.5$, with pressure-averaging. Initial mesh and meshes after several refinements.*

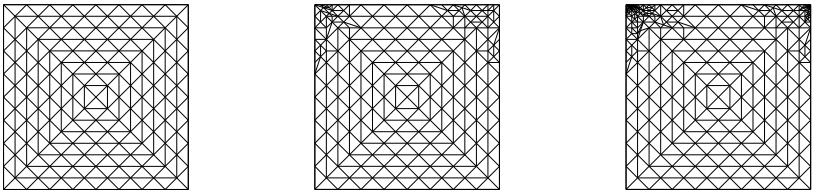


FIG. 6.13. *Driven cavity problem, adaptive refinement using the "Maximum strategy". Velocity and pressure.*

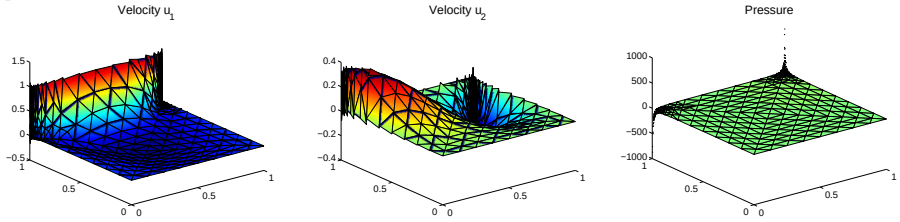


FIG. 6.14. *Driven cavity problem, adaptive refinement using the "Local Strategy", $\theta = 1.5$. Velocity and pressure.*

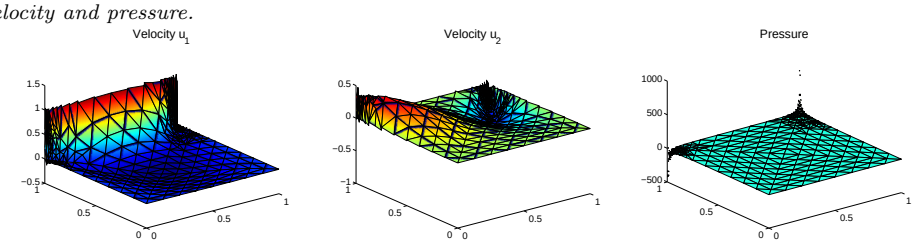


FIG. 6.15. *Driven cavity problem, adaptive refinement using the "Maximum strategy", with pressure-averaging. Velocity and pressure.*

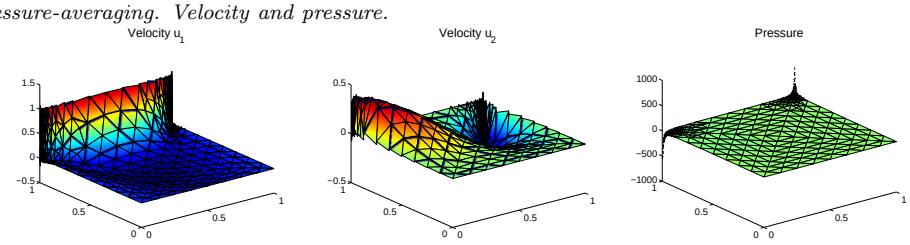


FIG. 6.16. *Driven cavity problem, adaptive refinement using the "Local Strategy", $\theta = 1.5$, with pressure-averaging. Velocity and pressure.*

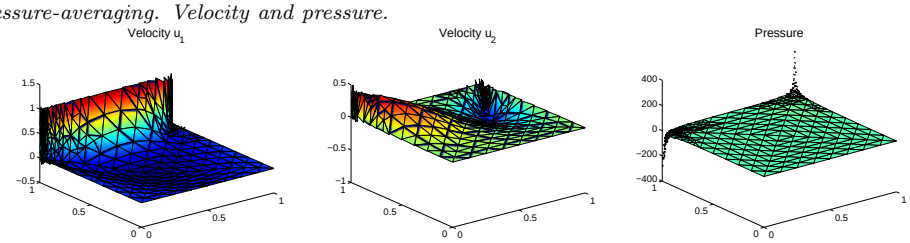


FIG. 6.17. *Velocity profiles of the driven cavity problem. The curves are velocity profiles of present solutions after 5 refinements, using different strategies. The stars denote the velocity profile of the solution using the Taylor-Hood element.*