ON THE ADAPTIVE SELECTION OF THE PARAMETER IN STABILIZED FINITE ELEMENT APPROXIMATIONS[*]

MARK AINSWORTH[†], ALEJANDRO ALLENDES[‡], GABRIEL R. BARRENECHEA[§], AND RICHARD RANKIN[¶]

Abstract. A systematic approach is developed for the selection of the stabilization parameter for stabilized finite element approximation of the Stokes problem, whereby the parameter is chosen to minimize a computable upper bound for the error in the approximation. The approach is applied in the context of both a single fixed mesh and an adaptive mesh refinement procedure. The optimization is carried out by a derivative-free optimization algorithm and is based on minimizing a new fully computable error estimator. Numerical results are presented illustrating the theory and the performance of the estimator, together with the optimization algorithm.

Key words. stabilized finite element method, stabilization parameter, computable error bounds, derivative-free optimization

AMS subject classifications. 65N15, 65N30

DOI. 10.1137/110837796

1. Introduction. The numerical approximation of the Stokes problem generally follows one of two complementary approaches. The first consists of using discrete velocity-pressure spaces satisfying the discrete inf-sup condition. Many such methods are available in the literature (see [20, 10] for extensive reviews). However, one perceived drawback of this approach is the fact that the discrete spaces cannot be of the same polynomial order in both variables while maintaining stability. The second approach consists of adding so-called stabilizing terms to the discrete formulation using an equal (or the lowest unequal) order velocity-pressure combination. These stabilizing terms can depend on residuals of the equation at the element level or can simply be based on compensating for the inf-sup deficiency of the pressure. For extensive reviews on different alternatives for stabilized finite element methods, see [6, 30].

One characteristic feature of stabilized methods is the presence of a positive constant multiplying the stabilization term. Naturally, the question of the selection of the actual value of the stabilization parameter in practical computation arises, which, although not affecting the rate of convergence, can have a significant impact on the absolute value of the error. Considerable effort has been expended in the quest to avoid having to make an ad hoc decision about the specific choice of the parameter.

[†]Division of Applied Mathematics, Brown University, Providence, RI 02912 (mark_ainsworth@brown.edu).

[‡]Departamento de Matemática, Universidad Técnica Federico Santa María, Casilla 110-V, Valparaíso, Chile (alejandro.allendes@usm.cl). This author was supported by the Faculty of Science of Strathclyde University and Comisión Nacional de Investigación Científica y Tecnológica – CONICYT (Chile) through a research studentship and grant 79090008, Fortalecimiento del grupo de Análisis y Modelamiento Matemático, Valparaíso.

[§]Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, Scotland (gabriel.barrenechea@strath.ac.uk).

[¶]Computational and Applied Mathematics Department, Rice University, Houston, TX 77005 (Richard.A.Rankin@rice.edu).

Variational multiscale methods (including residual-free bubbles and, recently, Petrov Galerkin Enriched methods [22, 11, 4, 5, 3]) may be regarded as a systematic approach to the selection of an explicit, closed form of the value of the stabilization parameter, thereby rendering the methods parameter-free.

In this work our approach is based on the premise that the best parameter is the one for which the error is minimal. Of course, the true value of the error is generally unknown. However, if a computable quantity $\eta(\alpha)$ is available, which depends on the value α of the stabilization parameter, and for which a two-sided bound on the true error in an appropriate norm $\|(e_V, e_P)\|$ holds, i.e., there exists a constant $c > 0$ independent of α, for which

$$(1.1) \qquad\qquad c\,\eta(\alpha) \le \|(e_V, e_P)\| \le \eta(\alpha),$$

then we may, in lieu of minimizing the true error, seek the value of α which minimizes the value of $\eta(\alpha)$. This approach constitutes an objective, rational approach to the selection of the stabilization parameter. Of course, the quality of the results will reflect the quality of the bounds (1.1): If $c = 1$, then we would be minimizing the true error, but this is, of course, unrealistic. In any case, what is really needed is for the value of α at which $\eta(\alpha)$ has a minimum to coincide with the value of α at which the true error has a minimum. The development of a method for defining a computable quantity $\eta(\alpha)$, satisfying (1.1) up to higher-order oscillation terms, is a key component of our approach and occupies section 5. As a fringe benefit, we obtain an expression for $\eta(\alpha)$ which enables us to estimate the error for virtually any existing stabilized method for the Stokes problem.

The search for the optimal value of the stabilization parameter has been considered before. For example, in [8] a residual based a posteriori error estimator was also minimized in order to obtain a value for the stabilization parameter (see also [25] for convection-diffusion problems), while in [31] the value is chosen by minimizing the condition number of the associated Schur complement system for the pressure field.

The development of the measure $\eta(\alpha)$ is only one part of the story and we must also select an algorithm for approximating its minima. The expression for $\eta(\alpha)$ depends on the stabilized finite element approximation obtained using a particular value α for the stabilization parameter. Thus, each evaluation of $\eta(\alpha)$ entails the computation of a finite element approximation. Furthermore, one does not have ready access to derivative information. These considerations suggest the use of a derivative-free optimization (DFO) approach to search for the value α_{opt} for which $\eta(\alpha)$ is minimized.

Numerical examples are presented showing the performance of the approach in the case of a fixed mesh for a variety of stabilized finite element methods. The results indicate that the algorithm does indeed enable us to select a near-optimal value of the stabilization parameter. The approach is then applied to the case of finite element approximation on a sequence of adaptively refined meshes, again yielding satisfactory results.

We summarize the main features of the work:
- the construction of a fully computable a posteriori error bound which is robust with respect to the stabilization parameter and which is applicable to a wide range of stabilized finite element methods;
- the selection of the value of the stabilization parameter based on minimizing the a posteriori error bound; and
- the combination of this approach with an adaptive refinement procedure.

The paper is organized as follows. In the next section we give some preliminaries that will be needed throughout the manuscript along with a description of the different stabilized finite element methods considered. Sections 3 and 4 are devoted to the optimization procedure, the summary of the main results of the a posteriori analysis, and numerical examples. Finally, the technical proofs related to the a posteriori error estimate are given in section 5, and conclusions are drawn in section 6.

2. Preliminaries. For a bounded open domain, $G \subset \mathbb{R}^d$, where $d = 1, 2$; $L^2(G)$ denotes the space of square integrable functions over G, $L_0^2(G)$ represents functions belonging to $L^2(G)$ with zero average in G, $H^1(G)$ is the usual Sobolev space, and $H_0^1(G)$ denotes the subspace of $H^1(G)$ consisting of functions whose trace is zero on the boundary of G. Let $(\cdot, \cdot)_G$ denote the inner product in $L^2(G)$ (or in $L^2(G)^2$, $L^2(G)^{2 \times 2}$ when necessary). The norm of the space $H^m(G)$ is denoted by $\| \cdot \|_{m,G}$, with the convention that when $m = 0$ we have that $\| \cdot \|_{0,G} = \| \cdot \|_{L^2(G)}$. We use bold letters to denote the vector-valued counterparts of the Sobolev and Lebesgue spaces, e.g., $\boldsymbol{H}_0^1(G) = H_0^1(G) \times H_0^1(G)$, and use an extra underaccent to denote their matrix-valued counterparts, e.g., $\underset{\approx}{\boldsymbol{L}}^2(G) = L^2(G)^{2 \times 2}$.

Let $\Omega \subset \mathbb{R}^2$ be an open polygonal domain with boundary Γ. Let $\{\mathcal{P}\}$ be a family of partitions of Ω, where each partition is built up using shape regular triangles K such that $\overline{\Omega} = \bigcup_{K \in \mathcal{P}} K$ and the nonempty intersection of two distinct elements is either a single common edge or vertex of both elements.

2.1. Notation. For convenience, we shall summarize all the notation used throughout the manuscript related to the partition of the domain.

For a fixed partition $\mathcal{P}$ let
- $\mathcal{E}$ denote the set of all edges;
- $\mathcal{E}_I \subset \mathcal{E}$ denote the set of interior edges;
- $\mathcal{E}_\Gamma \subset \mathcal{E}$ denote the set of boundary edges;
- $\mathcal{V}$ denote the set $\{\boldsymbol{x}_n\}$ of all element vertices;
- $\Omega_n = \{K \in \mathcal{P} : \boldsymbol{x}_n \in \overline{K} \text{ for a fixed } \boldsymbol{x}_n \in \mathcal{V}\}$;
- $\mathcal{E}_n = \{\gamma \in \mathcal{E} : \boldsymbol{x}_n \in \overline{\gamma} \text{ for a fixed } \boldsymbol{x}_n \in \mathcal{V}\}$;
- λ_n denote the usual barycentric coordinate associated to the vertex $\boldsymbol{x}_n \in \mathcal{V}$ and let $\lambda_{\mathbf{n}}^{(\mathbf{1})} = \begin{pmatrix} \lambda_n & 0 \end{pmatrix}^T$ and $\lambda_{\mathbf{n}}^{(\mathbf{2})} = \begin{pmatrix} 0 & \lambda_n \end{pmatrix}^T$.

For an element $K \in \mathcal{P}$ let
- $\mathbb{P}_n(K)$ denote the space of polynomials on K of total degree at most n;
- $\mathcal{E}_K$ denote the set containing the individual edges of the element;
- $\tilde{\Omega}_K = \{K' \in \mathcal{P} : \overline{K'} \cap \overline{K} \neq \emptyset\}$;
- $\Omega_K = \{K' \in \mathcal{P} : \mathcal{E}_K \cap \mathcal{E}_{K'} \neq \emptyset\}$;
- $|K|$ denote the area of K;
- $\overline{v}_K = \frac{1}{|K|} \int_K v \, d\boldsymbol{x}$ denote the mean value of v on the element K;
- h_K denote the diameter of K;
- $\hat{\boldsymbol{n}}_\gamma^K$ denote the unit exterior normal vector to edge $\gamma \in \mathcal{E}_K$.

For an edge $\gamma \in \mathcal{E}$ let
- $\mathbb{P}_n(\gamma)$ denote the space of polynomials on γ of total degree at most n;
- $\mathcal{V}_\gamma = \{\boldsymbol{x}_n \in \mathcal{V} : \boldsymbol{x}_n \in \overline{\gamma}\}$ denote the set of endpoints of an edge γ;
- $|\gamma|$ denote the length of γ.

We also define the projection operator $\Pi_K : \boldsymbol{L}^2(K) \to \mathbb{P}_1(K)^2$ by

$$(2.1) \qquad (\boldsymbol{t} - \Pi_K \boldsymbol{t}, \boldsymbol{\theta})_K = 0 \quad \forall \, \boldsymbol{\theta} \in \mathbb{P}_1(K)^2.$$

2.2. Model problem. We are interested in the following Stokes problem: For given data $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, find a velocity $\boldsymbol{u}$ and a pressure field p such that $\boldsymbol{u} = \boldsymbol{0}$ on Γ and

$$(2.2) \qquad -\nu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} \quad \text{and} \quad \nabla \cdot \boldsymbol{u} = 0 \qquad \text{in } \Omega \,,$$

where $\nu > 0$ is the fluid viscosity. The weak formulation of the Stokes problem then reads as follows: Find $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ such that

$$(2.3) \qquad \mathcal{A}(\boldsymbol{u}, p; \boldsymbol{v}, q) = \mathcal{F}(\boldsymbol{v}) \quad \text{for all } (\boldsymbol{v}, q) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega),$$

where

$$(2.4) \quad \mathcal{A}(\boldsymbol{u}, p; \boldsymbol{v}, q) = \nu(\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{\nabla}\boldsymbol{v})_\Omega - (p, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega + (q, \boldsymbol{\nabla} \cdot \boldsymbol{u})_\Omega \text{ and } \mathcal{F}(\boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})_\Omega.$$

The well-posedness of problem (2.3) is a consequence of two facts: the bilinear form $\nu(\boldsymbol{\nabla}\boldsymbol{u}, \boldsymbol{\nabla}\boldsymbol{v})_\Omega$ is coercive on $\boldsymbol{H}_0^1(\Omega)$ owing to Poincaré's inequality and hence is also coercive on the subspace

$$(2.5) \qquad \boldsymbol{X} = \left\{ \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) : \ \boldsymbol{\nabla} \cdot \boldsymbol{v} = 0 \right\},$$

and there exists a constant $\beta > 0$ such that

$$(2.6) \qquad \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \setminus \{\boldsymbol{0}\}} \frac{(q, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega}{\|\boldsymbol{\nabla}\boldsymbol{v}\|_{0,\Omega}} \geq \beta \|q\|_{0,\Omega} \quad \forall \, q \in L_0^2(\Omega).$$

The constant β is known as the inf-sup constant for the domain Ω. For more details concerning the well-posedness of problem (2.3), see [20].

2.3. Stabilized finite element methods. For nonnegative integers m, let

$$X_h^m = \left\{ v \in L^2(\Omega) : \ v_{|K} \in \mathbb{P}_m(K) \ \forall \ K \in \mathcal{P} \right\}$$

and $\boldsymbol{X}_h^m = (X_h^m)^2$. Let $\boldsymbol{V}_h = \boldsymbol{X}_h^1 \cap \boldsymbol{H}_0^1(\Omega)$ and let $Q_h \subset X_h^1 \cap L_0^2(\Omega)$. A stabilized finite element approximation of the Stokes problem then reads as follows: Find a pair $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ such that

$$(2.7) \qquad \mathcal{A}(\boldsymbol{u}_h, p_h; \boldsymbol{v}, q) + \alpha \mathcal{S}(\boldsymbol{u}_h, p_h, \boldsymbol{f}; q) = \mathcal{F}(\boldsymbol{v}, q) \quad \forall \, (\boldsymbol{v}, q) \in \boldsymbol{V}_h \times Q_h,$$

where $\mathcal{S}(\boldsymbol{u}_h, p_h, \boldsymbol{f}; q)$ is the stabilization term and the parameter α is a positive constant, sometimes referred to as the *stabilization parameter*. Often, the developers of a particular stabilized method give a recommendation α_{rec} for the value of the stabilization parameter to be used in practical computations, but in some cases no such value is identified. Many stabilized finite element methods are available, and below we give examples of stabilized finite element methods which can be used to approximate the solution of the Stokes problem. We shall employ various combinations of discrete velocity-pressure spaces, depending on the particular choice of stabilization (see Table 1).

• Galerkin least-squares type methods (GLS) [24, 30, 23, 19, 26]: The stabilizing term is given by

$$\mathcal{S}(\boldsymbol{u}_h, p_h, \boldsymbol{f}; q) = -\sum_{K \in \mathcal{P}} \frac{h_K^2}{\nu} (\boldsymbol{f} + \Delta\boldsymbol{u}_h - \nabla p_h, \nabla q)_K + \sum_{\gamma \in \mathcal{E}_I} \frac{|\gamma|}{\nu} (\llbracket p_h \rrbracket_\gamma, \llbracket q \rrbracket_\gamma)_\gamma,$$

TABLE 1
Discrete velocity-pressure space combinations used in conjunction with the stabilized formulations.

Method	Velocity space $\boldsymbol{V}_h$	Pressure space Q_h
$\mathbb{P}_1^2 - \mathbb{P}_0$	$\boldsymbol{X}_h^1 \cap \boldsymbol{H}_0^1(\Omega)$	$X_h^0 \cap L_0^2(\Omega)$
$\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$	$\boldsymbol{X}_h^1 \cap \boldsymbol{H}_0^1(\Omega)$	$X_h^1 \cap L_0^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$
$\mathbb{P}_1^2 - \mathbb{P}_1^{dis}$	$\boldsymbol{X}_h^1 \cap \boldsymbol{H}_0^1(\Omega)$	$X_h^1 \cap L_0^2(\Omega)$

where $[\![v]\!]_\gamma$ denotes the jump of v across γ, and may be used in conjunction with a $\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$ or a $\mathbb{P}_1^2 - \mathbb{P}_1^{dis}$ or a $\mathbb{P}_1^2 - \mathbb{P}_0$ pair. For this type of method the stabilization parameter is often recommended to be taken as $\alpha_{rec} = 1/24$.

• Brezzi and Pitkäranta (BP) [12]: The stabilizing term reads

$$\mathcal{S}\left(\boldsymbol{u}_h, p_h, \boldsymbol{f}; q\right) = \sum_{K \in \mathcal{P}} \frac{h_K^2}{\nu}\left(\nabla p_h, \nabla q\right)_K$$

for a $\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$ pair. For this method the authors do not recommend a particular choice of α, so in the absence of further information, we select $\alpha_{rec} = 1$.

• Polynomial pressure methods (PPS) [9, 17]: The stabilizing term reads

$$\mathcal{S}\left(\boldsymbol{u}_h, p_h, \boldsymbol{f}; q\right) = \sum_{K \in \mathcal{P}} \frac{1}{\nu}\left((I - \Psi)p_h, (I - \Psi)q\right)_K,$$

and the operator Ψ may be taken as $(\Psi v)_{|K} = \overline{v}_K$ for the $\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$ pair or a Clément-like interpolant for the $\mathbb{P}_1^2 - \mathbb{P}_0$ pair. (See section 6 in [9] for more details about the operator Ψ.) This method is recommended with a stabilization parameter $\alpha_{rec} = 1$. See also [7] for a more general version of this class, called local projection stabilized methods.

All the previous methods constitute stable and convergent schemes. However, alternative methods exist based on discretizing a regularization of the basic Stokes problem. Such methods, while stable, are inconsistent and nonconvergent in general, but can nevertheless deliver useful approximations if the value of the regularization parameter α is judiciously selected. One example of such a method is the following:

• Penalty pressure-type methods (PEPS) [13]: The stabilizing term reads

$$\mathcal{S}\left(\boldsymbol{u}_h, p_h, \boldsymbol{f}; q\right) = \sum_{K \in \mathcal{P}} \left(p_h, q\right)_K$$

and may be used in conjunction with a $\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$ or a $\mathbb{P}_1^2 - \mathbb{P}_1^{dis}$ or a $\mathbb{P}_1^2 - \mathbb{P}_0$ pair. We take $\alpha_{rec} = 1$.

The estimators that we obtain are valid for all the above mentioned methods. Moreover, they remain valid in the case of nonhomogeneous Dirichlet data $\boldsymbol{u} = \boldsymbol{u}_D$ on Γ, for given $\boldsymbol{u}_D \in \boldsymbol{X}_h^1 \cap \boldsymbol{H}^1(\Omega)$. Furthermore, from now on, c and C will denote positive constants which are independent of any mesh size, the viscosity ν, and the stabilization parameter α.

3. An algorithm for selecting the stabilization parameter on a given mesh. Although the a priori rate of convergence of a stabilized method is independent of the value of the stabilization parameter (provided the discrete problem is well-posed), the absolute value of the error varies depending on the choice of the parameter. In order to illustrate this point we consider two example problems.

Problem 1. We consider $\Omega = (0,1)^2$ (the unit square). For this domain, a lower bound of 0.38 for the value of the inf-sup constant β is proved in [32]. We took $\nu = 1$ and let $\boldsymbol{f}$ be such that the exact velocity and pressure are given by $\boldsymbol{u} = [x^2(x-1)^2 y(y-1)(2y-1), -y^2(y-1)^2 x(x-1)(2x-1)]$ and $p = xy(1-x)(1-y)-1/36$.

Problem 2. We consider the T-shaped domain $\Omega = ((-1.5, 1.5) \times (0,1)) \cup ((-0.5, 0.5) \times (-2, 1))$. From [32] we have that 0.1 is a lower bound for the inf-sup constant β for this domain. We took $\nu = 1$ and let $\boldsymbol{f} = \boldsymbol{0}$ and imposed Dirichlet boundary conditions of $\boldsymbol{u} = (y, 0)$ on $x = \pm 1.5$, $\boldsymbol{u} = (1,0)$ on $y = 1$ and $\boldsymbol{u} = (0,0)$ elsewhere on the boundary.

We shall present results for Problems 1 and 2 obtained using meshes S to S_d and T_a to T_d shown in Figures 1 and 2, respectively.

The following norm is used to measure the velocity and the pressure errors:

$$\|(\boldsymbol{e_V}, e_P)\| = \left(\nu^2 \|\boldsymbol{\nabla} \boldsymbol{e_V}\|_{0,\Omega}^2 + \beta^2 \|e_P\|_{0,\Omega}^2\right)^{1/2}.$$

The values of norms of the errors obtained for various stabilized schemes and various values of the stabilization parameter on fixed meshes from Figures 1 and 2 are illustrated in Figures 3 and 4 for Problem 1. For Problem 2, since we do not have at hand the analytical solution, we build a numerical reference solution and compute the errors with respect to it. For that, we solved Problem 2 using a Taylor–Hood discretization ($\mathbb{P}_2^{cts} - \mathbb{P}_1^{cts}$) on a highly refined mesh which we obtained following the same refinements as in Figure 2 (refine about the reentrant corners) until we built a mesh with 261,348 elements, and used that solution as an "exact" solution to compute the errors with respect to. The results are depicted in Figure 5. It is clear that in some cases the choice of the parameter α can significantly affect the errors. In particular, an inappropriate choice can result in a loss of a factor of two, or sometimes much more, in the accuracy compared with a more judicious choice. In terms of practical computation this means that a careful choice of α can sometimes be at least as effective as a global mesh refinement.
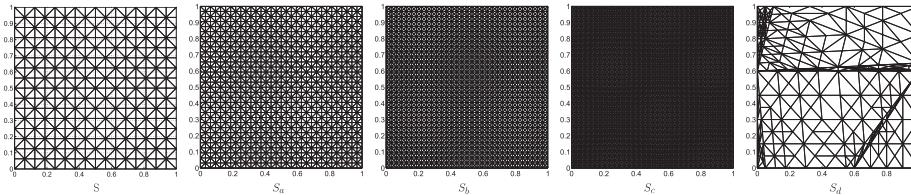
FIG. 1. *Uniform meshes S with 512, S_a with 2048, S_b with 4096, and S_c with 8192 elements and distorted mesh S_d with 412 elements for Problem* 1.
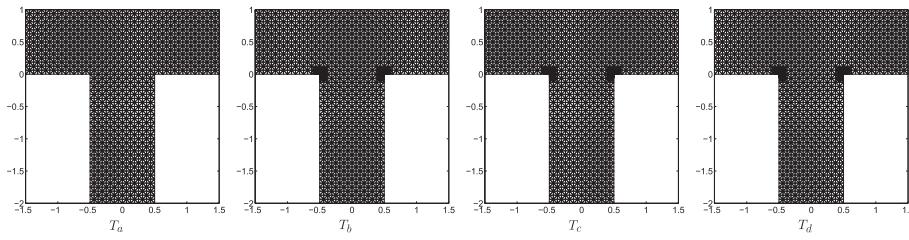
FIG. 2. *Meshes T_a with 2560, T_b with 5076, T_c with 7108 and T_d with 11,006 elements for Problem* 2.
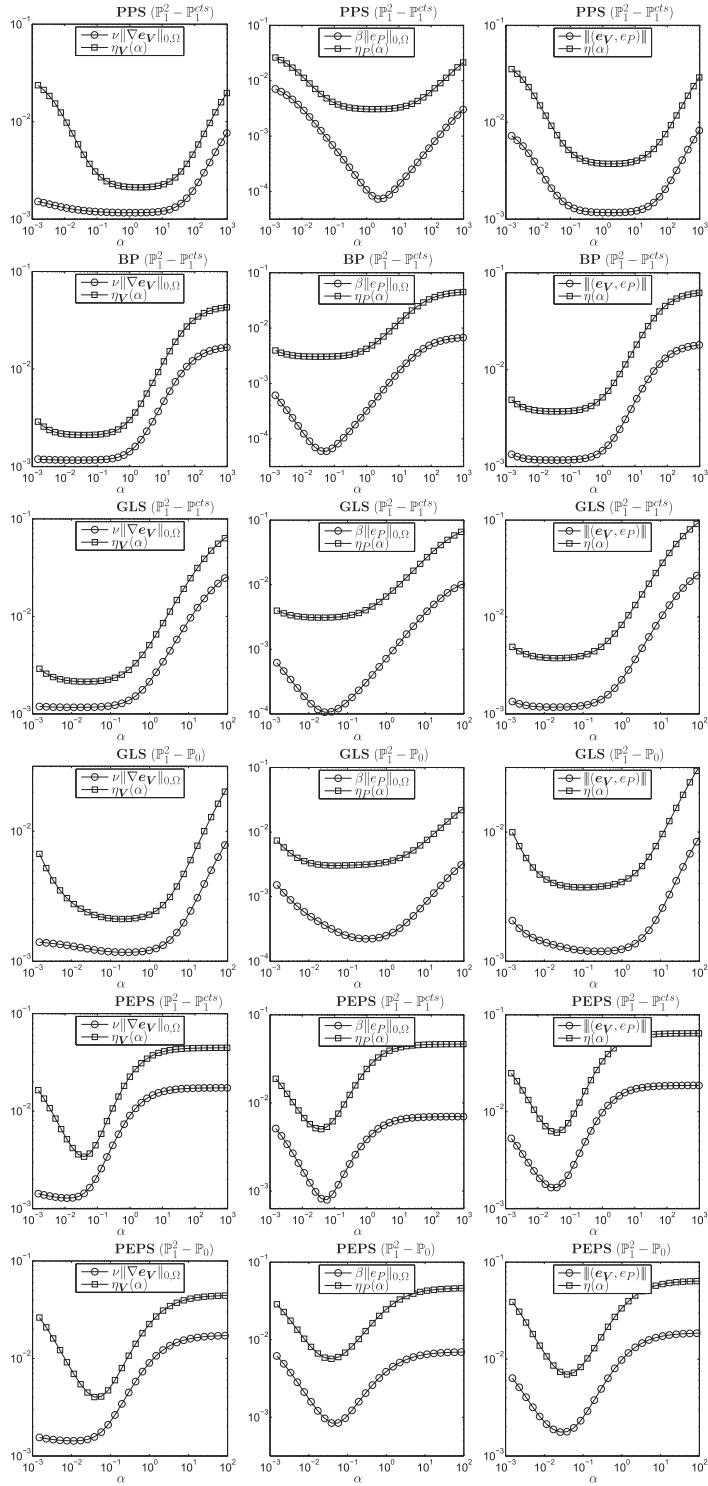
FIG. 3. *Comparison of the estimated and true errors for different values of α using mesh S_c for Problem* 1.
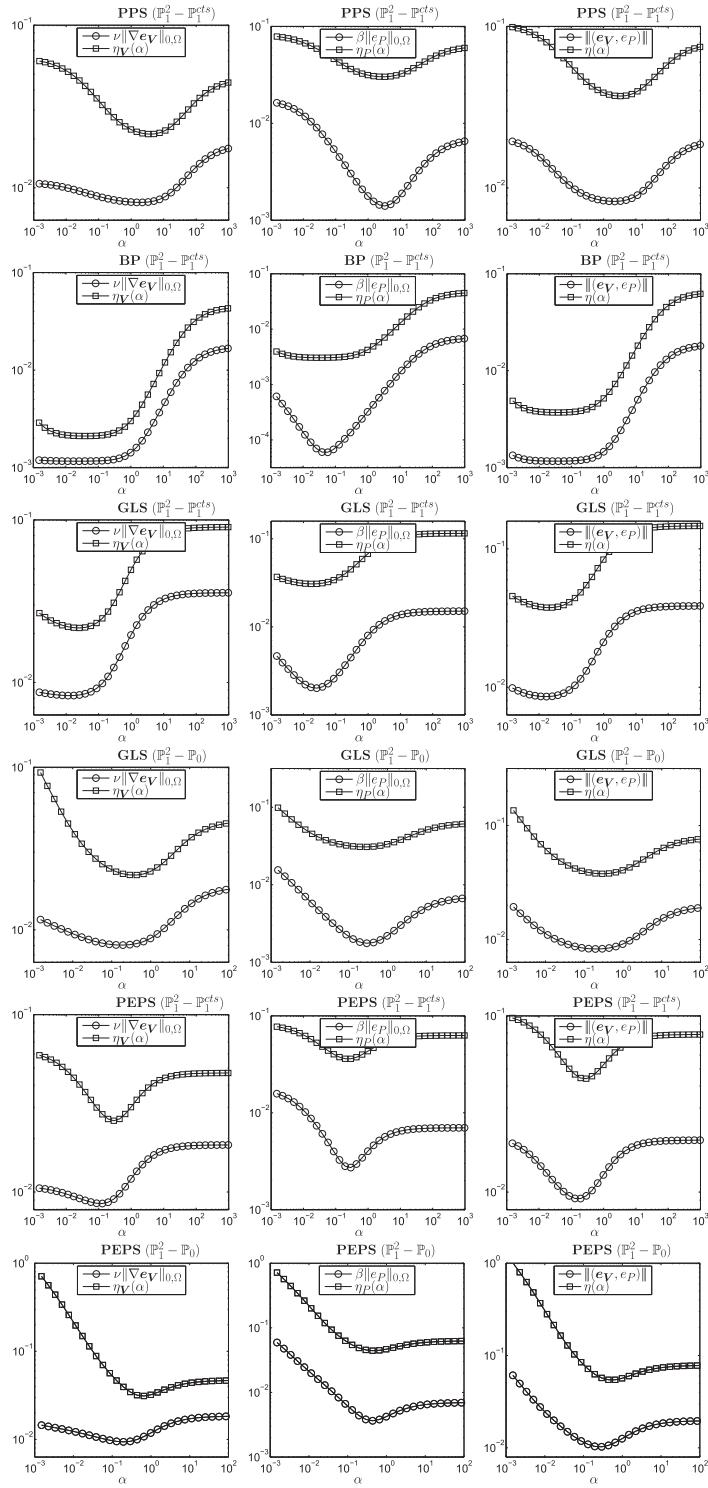
FIG. 4. *Comparison of the estimated and true errors for different values of α using mesh S_d for Problem* 1.
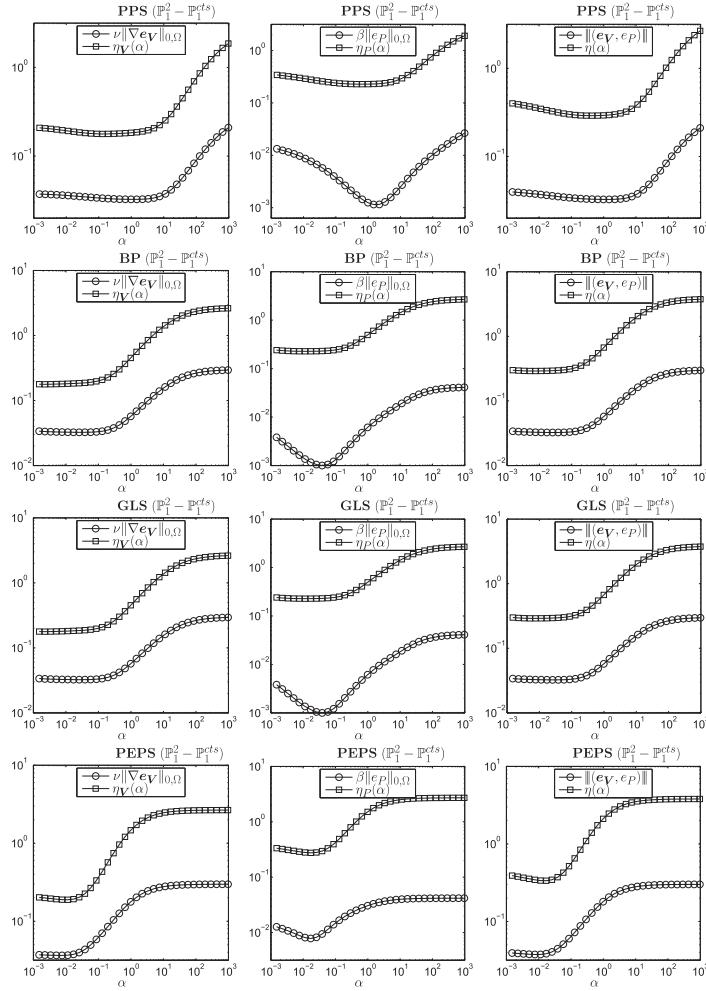
FIG. 5. *Comparison of the estimated and true errors for different values of α using mesh T_c for Problem 2.*

Ideally, we would like to select the best value of the stabilization parameter for each problem, each mesh, and each discretization scheme. The following result, which summarizes the findings of section 5, will be helpful in this respect.

THEOREM 3.1. *Let $\alpha > 0$ and $\boldsymbol{e_V} = \boldsymbol{u} - \boldsymbol{u}_h$ and $e_P = p - p_h$ denote the error in the velocity and pressure approximations obtained using a stabilized finite element formulation. Then,*

$$(3.1) \qquad \nu \left\| \boldsymbol{\nabla} \boldsymbol{e_V} \right\|_{0,\Omega} \leq \eta_{\boldsymbol{V}}(\alpha), \quad \beta \left\| e_P \right\|_{0,\Omega} \leq \eta_P(\alpha), \quad and \quad \left\| (\boldsymbol{e_V}, e_P) \right\| \leq \eta(\alpha),$$

where the velocity and pressure error estimators are given by

$$(3.2) \qquad \eta_{\boldsymbol{V}}(\alpha)^2 = \Phi_c^2 + \nu^2 \Phi_{nc}^2 \quad and \quad \eta_P(\alpha)^2 = \left(\Phi_c^* + \nu \Phi_{nc} \right)^2,$$

respectively, and the total error estimator is given by

$$(3.3) \qquad \eta(\alpha) = \left(\eta_{\boldsymbol{V}}(\alpha)^2 + \eta_P(\alpha)^2 \right)^{1/2},$$

with Φ_c, Φ_{nc} and Φ_c^*, *being given in* (5.21), (5.22), *and* (5.29), *respectively. Moreover, there exists a positive constant c independent of* α, ν *and any mesh size, such that*

$$(3.4) \qquad c\,\eta_K(\alpha)^2 \leq \sum_{K' \in \tilde{\Omega}_K} \left(\nu^2 \|\boldsymbol{\nabla e_V}\|_{0,K'}^2 + \beta^2 \|e_P\|_{0,K'}^2 + h_{K'}^2 \|\boldsymbol{f} - \Pi_{K'}\boldsymbol{f}\|_{0,K'}^2 \right),$$

where $\eta_K(\alpha)$ *is given in* (4.1).

Proof. The upper bounds follow from Theorems 5.2 and 5.3. The proof of (3.4) is given in section 5.5. $\quad\square$

It should be borne in mind that the estimator $\eta(\alpha)$ is computed using the finite element approximation obtained using the value α as the stabilization parameter. The values of the quantities $\eta(\alpha)$, $\eta_{\boldsymbol{V}}(\alpha)$, and $\eta_P(\alpha)$ are shown along with the true errors in Figures 3 to 5. We first observe that both components of the error and the estimator seem sensitive to the choice of α and that both components of the estimator have a similar qualitative behavior to the individual errors. We observe as well that the total error $\|(\boldsymbol{e_V}, e_P)\|$ and the complete estimator $\eta(\alpha)$ are more in agreement than both the components, in terms of sensitivity to the choice of α. Significantly, both exhibit minima at roughly the same locations. This correlation suggests selecting the stabilization parameter α to minimize the upper bound $\eta(\alpha)$ for the true error $\|(\boldsymbol{e_V}, e_P)\|$. While the values of the estimated and true errors may differ, the proximity of the minimizers means that the resulting choice of α will be near optimal.

It remains to select an appropriate method for obtaining the minimizer of $\eta(\alpha)$. We use the trust-region DFO algorithm (see [14] and references therein) to approximate the minimizer of $\eta(\alpha)$. For the reader's convenience, we give a brief description of the method, which is described in full detail in [15, 16].

We begin by choosing constants $\varepsilon_{\mathcal{D}}, \Lambda, \Delta_{max} > 0$, $0 \leq \text{tol}_0 \leq \text{tol}_1 < 1$, $0 < \omega_0 < 1 < \omega_1$ and a trust-region radius $\Delta_0 \in (0, \Delta_{max}]$. Construct a fully quadratic model (in the sense of section 3 in [16]) by evaluating $\eta(\alpha)$ at a set of three sample points $\boldsymbol{\alpha}_0 = \{\alpha_1, \alpha_2, \alpha_3\}$ to obtain a quadratic interpolant, given by

$$m_0(\alpha) = c_0 + \alpha g_0 + \alpha^2 H_0,$$

where $c_0, g_0, H_0 \in \mathbb{R}$. Denote by $\mathcal{D}_0(\alpha) = \max\{|g_0 + 2\alpha H_0|, |2H_0|\}$ and choose any initial point χ_0 from the sample points; in our case we take the one with minimum value of $\eta(\alpha)$. If there are two such choices for χ_0, then choose the one maximizing $\mathcal{D}_0(\chi_0)$, and if there are still two choices, either is used at random. If there are three such choices, then use a model-improvement algorithm (Algorithm 6.2 from [16]), based on moving the sample points in order to obtain a fully quadratic model. Set $k = 0$.

If $\mathcal{D}_k(\chi_k) \leq \varepsilon_{\mathcal{D}}$ call Algorithm 6.2 from [16] to obtain a new quadratic model; otherwise compute the step s_k that sufficiently reduces the model $m_k(\alpha)$ by solving the trust-region problem

$$\min_{s \in (-\Delta_k, \Delta_k)} m(\chi_k + s).$$

Compute $\eta(\chi_k + s_k)$ and define

$$\rho_k = \frac{\eta(\chi_k) - \eta(\chi_k + s_k)}{m(\chi_k) - m(\chi_k + s_k)}.$$

If $\rho_k \geq \text{tol}_1$ or if both $\rho_k \geq \text{tol}_0$ and the model is fully quadratic, then the new iterate $\chi_{k+1} = \chi_k + s_k$ replaces the sample point with the largest value of $\eta(\alpha)$, resulting in a new sample set $\boldsymbol{\alpha}_{k+1}$ from which we obtain a new fully quadratic model $m_{k+1}(\alpha)$; otherwise use Algorithm 6.2 from [16] and define $m_{k+1}(\alpha)$ to be the (possibly improved) model.

Update the trust-region radius as follows. Set

$$\Delta_{k+1} \in \begin{cases} \{\min\{\omega_1\Delta_k, \Delta_{max}\}\} & \text{if } \rho_k \geq \text{tol}_1 \text{ and } \Delta_k < \Lambda\mathcal{D}_k(\chi_k), \\ [\Delta_k, \min\{\omega_1\Delta_k, \Delta_{max}\}] & \text{if } \rho_k \geq \text{tol}_1 \text{ and } \Delta_k \geq \Lambda\mathcal{D}_k(\chi_k), \\ \{\omega_0\Delta_k\} & \text{if } \rho_k < \text{tol}_1 \text{ and } m_k \text{ is fully quadratic}, \\ \{\Delta_k\} & \text{if } \rho_k < \text{tol}_1 \text{ and } m_k \text{ is not fully quadratic}. \end{cases}$$

Take $\alpha_{opt} = \arg\min\{\eta(\alpha) : \alpha \in \boldsymbol{\alpha}_{k+1}\}$, increment k, and repeat the algorithm.

In Figure 6 the DFO search is presented for the **GLS** ($\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$) and **PEPS** ($\mathbb{P}_1^2 - \mathbb{P}_0$) methods using mesh S_c from Figure 1 and mesh T_d from Figure 2. We measure the gain using the approximation α_{opt} of the optimal value for the stabilization parameter compared with the recommended value α_{rec} by calculating the percentage gains, i.e.,

$$\mathcal{G}_{\boldsymbol{V}} = 100\frac{\eta_{\boldsymbol{V}}(\alpha_{rec}) - \eta_{\boldsymbol{V}}(\alpha_{opt,\boldsymbol{V}})}{\eta_{\boldsymbol{V}}(\alpha_{rec})}\%, \quad \text{where} \quad \alpha_{opt,\boldsymbol{V}} \approx \arg\min\{\eta_{\boldsymbol{V}}(\alpha)\},$$

$$\mathcal{G}_P = 100\frac{\eta_P(\alpha_{rec}) - \eta_P(\alpha_{opt,P})}{\eta_P(\alpha_{rec})}\%, \quad \text{where} \quad \alpha_{opt,P} \approx \arg\min\{\eta_P(\alpha)\},$$

$$\mathcal{G} = 100\frac{\eta(\alpha_{rec}) - \eta(\alpha_{opt})}{\eta(\alpha_{rec})}\%, \qquad \text{where} \quad \alpha_{opt} \approx \arg\min\{\eta(\alpha)\}.$$

The findings of performing the DFO search on fixed meshes are shown in Table 2, where we present the percentage gains and the approximations of the optimal value for the stabilization parameters $\alpha_{opt}, \alpha_{opt,\boldsymbol{V}}$, and $\alpha_{opt,P}$. We notice that the optimal values achieved with the separate estimators are different, each one providing a significant gain on the estimator. On the other hand, the value achieved by the total estimator is (at least in most cases) between those two values and can provide a significant reduction on the estimator. The numerical results from Figures 3–5 suggest that this reduction, of up to 80% in some cases with respect to the reference value for α, induces a significant gain on the error as well.

4. Selection of the stabilization parameter on a sequence of adaptively refined meshes. The results in the previous section are concerned with fixed meshes. We now apply the approach in the context of an adaptive mesh refinement procedure, driven using the local error indicator

(4.1) $$\eta_K(\alpha)^2 = \Phi_{c,K}^2 + \nu^2\Phi_{nc,K}^2 + \left(\Phi_{c,K}^* + \nu\Phi_{nc,K}\right)^2,$$

where Φ_c is given by (5.21), Φ_{nc} is given by (5.22), and Φ_c^* is given by (5.29).

Ideally, one would optimize over α on every mesh constructed throughout the adaptive refinement procedure. In practice, the cost of such a procedure would be prohibitive and, fortunately, is unnecessary. Instead we optimize the choice of α *once* on the initial mesh, and then retain this value on all the subsequent adaptively refined meshes.

In Figures 7 and 8 we present the results obtained using both the *idealized algorithm* and the proposed *practical algorithm* to approximate the same problems considered in the previous section. We define the effectivity index $\Theta(\alpha) = \eta(\alpha)/\|(\boldsymbol{e_V}, e_P)\|$
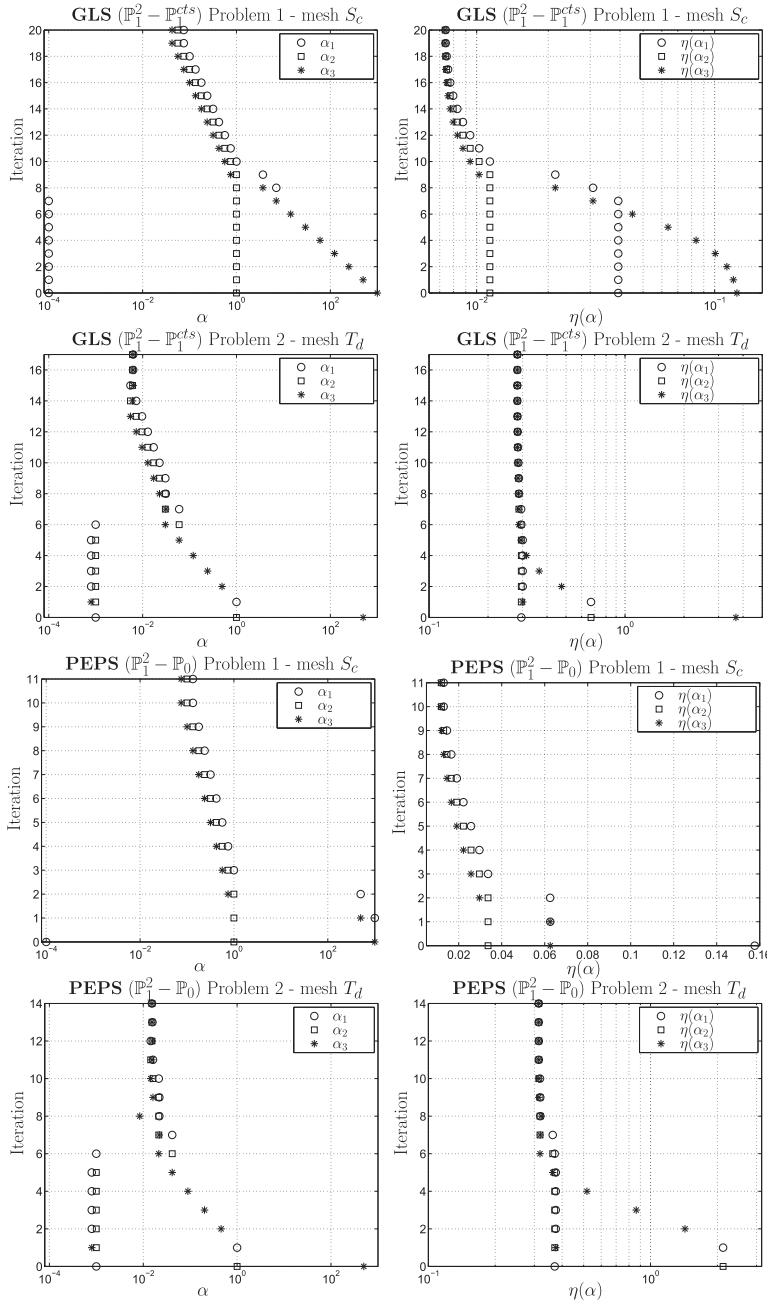
FIG. 6. *DFO search for Problems* 1 *and* 2 *on fixed meshes* S_c *and* T_d, *respectively.*

and show the performance of the algorithms for a variety of stablized methods. The meshes used to start the algorithms are shown in Figure 9. The results show the good behavior of the error estimator and that the performance of both algorithms is virtually identical, indicating that the optimal choice of α changes little from the value obtained based on the initial coarse mesh.

TABLE 2

Percentage gains $\mathcal{G}$, $\mathcal{G}_V$, $\mathcal{G}_P$, and α_{opt}'s for Problem 1 using the fixed meshes S_d, S, S_a, S_b, and S_c and for Problem 2 using the fixed meshes T_a, T_b, T_c, and T_d.

| | PPS ($\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$) $\alpha_{rec}=1$ | | BP ($\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$) $\alpha_{rec}=1$ | | GLS ($\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$) $\alpha_{rec}=1/24$ | | GLS ($\mathbb{P}_1^2 - \mathbb{P}_0$) $\alpha_{rec}=1/24$ | | PEPS ($\mathbb{P}_1^2 - \mathbb{P}_1^{cts}$) $\alpha_{rec}=1$ | | PEPS ($\mathbb{P}_1^2 - \mathbb{P}_0$) $\alpha_{rec}=1$ | |

Problem 1

Mesh	$\mathcal{G}_V$	$\alpha_{opt,V}$	$\mathcal{G}_V$	$\alpha_{opt,V}$	$\mathcal{G}_V$	$\alpha_{opt,V}$	$\mathcal{G}_V$	$\alpha_{opt,V}$	$\mathcal{G}_V$	$\alpha_{opt,V}$	$\mathcal{G}_V$	$\alpha_{opt,V}$
S_d	6.26	3.9034	29.25	0.0956	0.82	0.0312	18.9	0.457624	17.25	0.2826	2.55	0.765648
S	2.83	2.0909	47.78	0.0418	2.17	0.0235	12.27	0.190141	48.46	0.1780	42.40	0.1816
S_a	1.36	2.0044	40.94	0.0402	2.74	0.0060	7.89	0.287962	71.82	0.0728	66.5	0.1181
S_b	0.09	0.6655	36.81	0.0138	2.74	0.0060	3.4382	0.197	86.27	0.0106	84.39	0.0039
S_c	0.69	1.9764	30.22	0.0397	0.76	0.0224	6.2856	0.1902	84.98	0.0376	81.27	0.0676

Mesh	$\mathcal{G}_P$	$\alpha_{opt,P}$	$\mathcal{G}_P$	$\alpha_{opt,P}$	$\mathcal{G}_P$	$\alpha_{opt,P}$	$\mathcal{G}_P$	$\alpha_{opt,P}$	$\mathcal{G}_P$	$\alpha_{opt,P}$	$\mathcal{G}_P$	$\alpha_{opt,P}$
S_d	3.86	2.9589	30.3	0.04714	2.33	0.0188	10.6	0.313205	17.95	0.2319	4.55	0.586223
S	1.7	1.8343	43.51	0.04906	2.99	0.0203	4.68	0.159632	42.56	0.1242	37.50	0.1422
S_a	0.85	1.7745	37.45	0.04751	2.82	0.0086	1.16	0.104272	64.44	0.0652	60.06	0.0904
S_b	0.04	0.7686	34.15	0.01704	2.82	0.0086	6.27	0.000101	79.29	0.0225	85.11	0.0001
S_c	0.45	1.7600	28.17	0.04672	1.00	0.0201	0.04	0.0476	79.1	0.0447	76.30	0.053

Mesh	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}
S_d	4.62	3.2509	30.04	0.0515	1.85	0.0205	13.8	0.284102	17.63	0.2495	3.59	0.669955
S	2.06	1.9203	44.98	0.0498	2.7	0.0214	6.88	0.207045	44.5	0.1591	38.77	0.1577
S_a	1.01	1.8501	38.63	0.0482	1.68	0.0209	3.25	0.158682	67.39	0.0683	62.4	0.1033
S_b	0.05	0.7298	35.04	0.0163	2.78	0.0080	0.03	0.034763	82	0.0201	84.78	0.000101
S_c	2	1.8302	28.89	0.0365	0.92	0.0209	1.35	0.093298	81.77	0.0363	78.34	0.0591

Problem 2

Mesh	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}	$\mathcal{G}$	α_{opt}
T_a	8.3	0.1953	50.75	0.0025	15.7	0.0032	14.2	0.0089	72.9	0.0217	79.9	0.0188
T_b	1.5	0.3189	56.6	0.005	3.9	0.005	8.57	0.0129	84.1	0.0146	84.8	0.0149
T_c	0.78	0.4039	56.89	0.0074	2.32	0.0055	8.2	0.0133	84.2	0.015	85.03	0.0146
T_d	0.94	0.4219	57.9	0.0062	3.25	0.0062	9.5	0.0127	84.6	0.0146	85.1	0.0153

Idealized algorithm: Adaptive mesh refinement and DFO search.

1: Construct mesh $\mathcal{P}_0$. Set $i = 0$.

2: Performing the DFO algorithm on the fixed mesh $\mathcal{P}_i$, compute $\alpha_{opt}^{(i)}$.

3: For each element K in $\mathcal{P}_i$, compute a local error indicator $\eta_K(\alpha_{opt}^{(i)})$.

4: Triangle K is marked for refinement if

$$\eta_K(\alpha_{opt}^{(i)}) \geq \tfrac{1}{2} \max_{K \in \mathcal{P}_i} \eta_K(\alpha_{opt}^{(i)}).$$

5: From step **4** deduce a new mesh using longest edge bisection refinement.

6: Set $i \leftarrow i + 1$ and return to step **2**.

Practical algorithm: Adaptive mesh refinement and DFO search.

1: Construct mesh $\mathcal{P}_0$.

2: Performing the DFO algorithm on the fixed mesh $\mathcal{P}_0$, compute $\alpha_{opt}^{(0)}$ and set $i = 0$.

3: For each element K in $\mathcal{P}_i$, compute a local error indicator $\eta_K(\alpha_{opt}^{(0)})$.

4: Triangle K is marked for refinement if

$$\eta_K(\alpha_{opt}^{(0)}) \geq \tfrac{1}{2} \max_{K \in \mathcal{P}_i} \eta_K(\alpha_{opt}^{(0)}).$$

5: From step **4** deduce a new mesh using longest edge bisection refinement.

**6:** Set $i \leftarrow i + 1$ and return to step **3**.
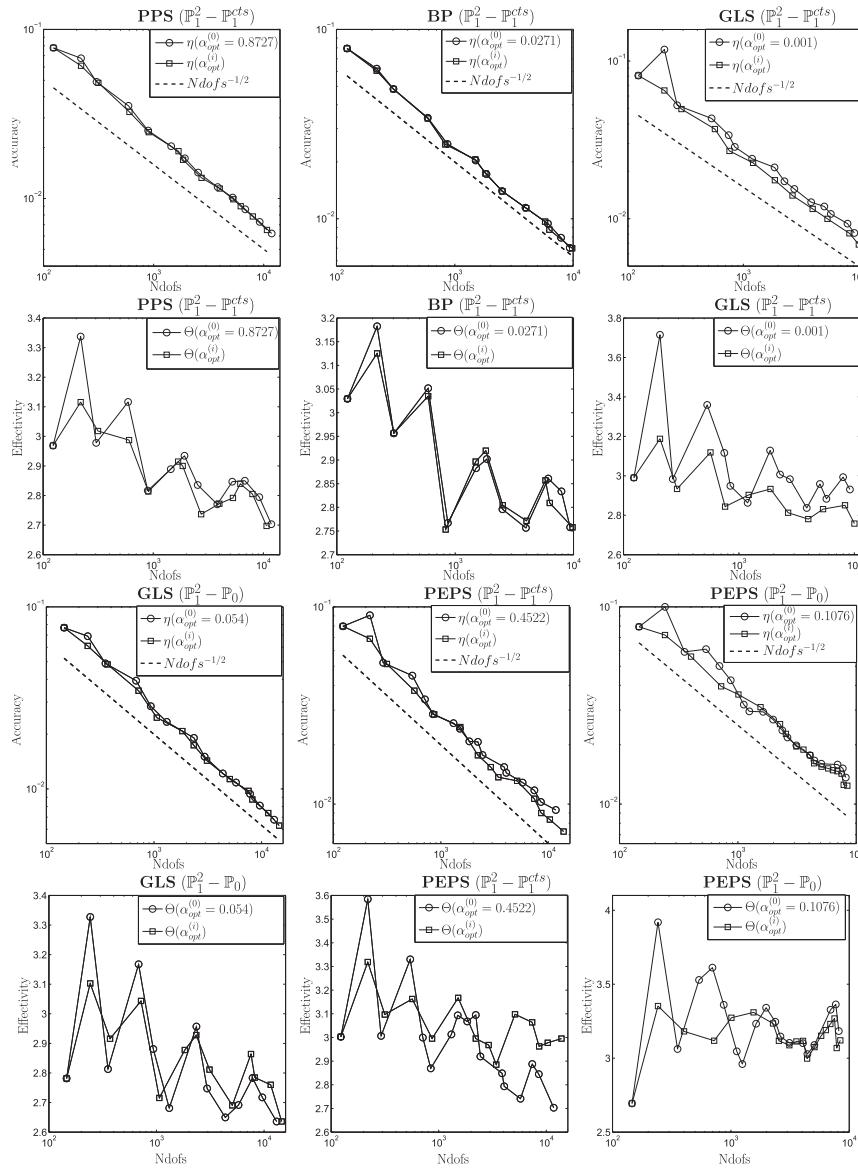
FIG. 7. *Performance of the idealized and practical algorithms applied to Problem* 1.

5. A posteriori error estimator.
We now turn to the derivation of the error estimator $\eta(\alpha)$ used in the selection of the stabilization parameter.

5.1. The error equation. Recall that $\boldsymbol{e_V} = \boldsymbol{u} - \boldsymbol{u}_h \in \boldsymbol{H}_0^1(\Omega)$ and $e_P = p - p_h \in L_0^2(\Omega)$ denote the errors in velocity and pressure, respectively, where $(\boldsymbol{u}, p)$ is the solution of (2.3) and $(\boldsymbol{u}_h, p_h)$ is the solution of (2.7). Thanks to (2.3) and (2.4), the errors satisfy, for all $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$ and $q \in L_0^2(\Omega)$,

$$(5.1) \quad \mathcal{A}(\boldsymbol{e_V}, e_P; \boldsymbol{v}, q) = \sum_{K \in \mathcal{P}} \Big((\boldsymbol{f}, \boldsymbol{v})_K - \nu(\boldsymbol{\nabla} \boldsymbol{u}_h, \boldsymbol{\nabla} \boldsymbol{v})_K + (p_h, \boldsymbol{\nabla} \cdot \boldsymbol{v})_K - (q, \boldsymbol{\nabla} \cdot \boldsymbol{u}_h)_K \Big).$$
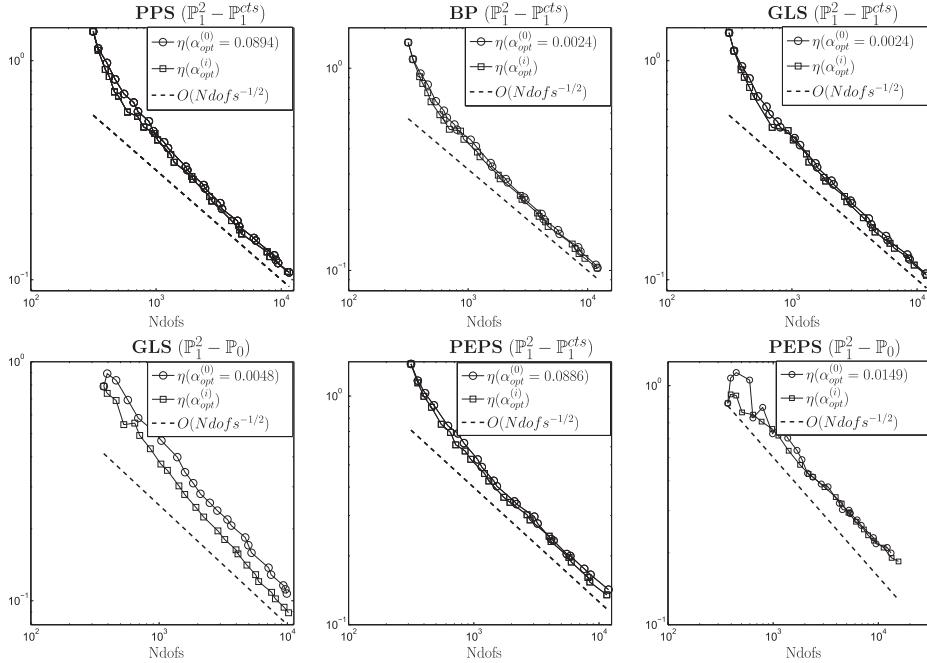
FIG. 8. *Performance of the idealized and practical algorithms applied to Problem* 2.
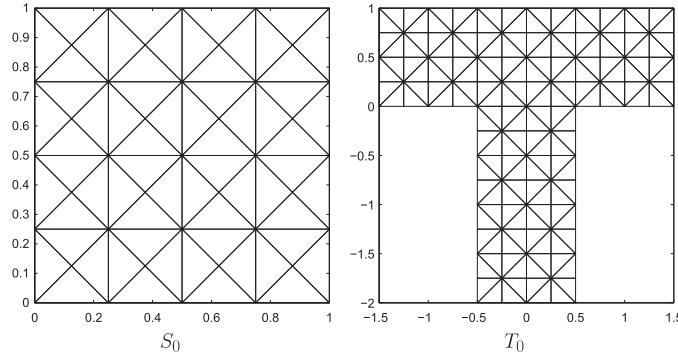

FIG. 9. *Initial meshes S_0 and T_0 for Problems 1 and 2, respectively, for the idealized and practical algorithms.*

Let $\{\boldsymbol{g}_{\gamma,K}\}$ be a set of equilibrated boundary fluxes which are such that $\boldsymbol{g}_{\gamma,K} \in \mathbb{P}_1(\gamma)^2$ for each $\gamma \in \mathcal{E}_K$ for all $K \in \mathcal{P}$,

$$(5.2) \qquad \boldsymbol{g}_{\gamma,K} + \boldsymbol{g}_{\gamma,K'} = \boldsymbol{0} \quad \text{if} \quad \gamma = \mathcal{E}_K \cap \mathcal{E}_{K'} \quad \text{for distinct} \quad K, K' \in \mathcal{P},$$

and, for all $K \in \mathcal{P}$, satisfy the first-order equilibration condition

$$(5.3) \quad (\boldsymbol{f},\boldsymbol{\theta})_K + \sum_{\gamma \in \mathcal{E}_K} \left(\boldsymbol{g}_{\gamma,K},\boldsymbol{\theta}\right)_\gamma - \nu\left(\boldsymbol{\nabla}\boldsymbol{u}_h, \boldsymbol{\nabla}\boldsymbol{\theta}\right)_K + (p_h, \boldsymbol{\nabla}\cdot\boldsymbol{\theta})_K = 0 \;\; \forall \boldsymbol{\theta} \in \mathbb{P}_1(K)^2.$$

A process to obtain such equilibrated boundary fluxes will be described in section 5.3. Now, (5.2) allows us to incorporate the boundary fluxes into (5.1) and integrate

by parts to yield

$$
\mathcal{A}(\boldsymbol{e_V}, e_P; \boldsymbol{v}, q)
$$

(5.4)
$$
= \sum_{K \in \mathcal{P}} \left((\boldsymbol{r}_K, \boldsymbol{v})_K + \sum_{\gamma \in \mathcal{E}_K} (\boldsymbol{R}_{\gamma,K}, \boldsymbol{v})_\gamma + (\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v})_K - (q, \boldsymbol{\nabla} \cdot \boldsymbol{u}_h)_K \right)
$$

for all $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$ and $q \in L_0^2(\Omega)$, where the interior residuals $\boldsymbol{r}_K \in \mathbb{P}_1(K)^2$ are given by

(5.5)
$$
\boldsymbol{r}_K = \Pi_K \boldsymbol{f} - \nabla p_h \qquad \text{on } K,
$$

and the boundary residuals $\boldsymbol{R}_{\gamma,K} \in \mathbb{P}_1(\gamma)^2$ are given by

(5.6)
$$
\boldsymbol{R}_{\gamma,K} = \boldsymbol{g}_{\gamma,K} - \nu \boldsymbol{\nabla} \boldsymbol{u}_{h|K} \hat{\boldsymbol{n}}_\gamma^K + p_{h|K} \hat{\boldsymbol{n}}_\gamma^K \qquad \text{on } \gamma \in \mathcal{E}_K.
$$

Let $\underset{\approx}{\boldsymbol{\sigma}}_K^* \in \mathbb{P}_2(K)^{2\times 2}$ be the unique solution to the following problem:

$$
\begin{aligned}
\text{minimize} \quad & \left\| \underset{\approx}{\boldsymbol{\sigma}}_K^* \right\|_{0,K} \\
\text{subject to} \quad & \underset{\approx}{\boldsymbol{\sigma}}_K^* \hat{\boldsymbol{n}}_\gamma^K = \boldsymbol{R}_{\gamma,K} \text{ on each } \gamma \in \mathcal{E}_K, \\
& -\mathbf{div} \underset{\approx}{\boldsymbol{\sigma}}_K^* = \boldsymbol{r}_K \text{ in } K.
\end{aligned}
$$

Note that since integration by parts and (2.1) allow us to write (5.3) as

(5.7)
$$
(\boldsymbol{r}_K, \boldsymbol{\theta})_K + \sum_{\gamma \in \mathcal{E}_K} (\boldsymbol{R}_{\gamma,K}, \boldsymbol{\theta})_\gamma = 0 \ \forall \ \boldsymbol{\theta} \in \mathbb{P}_1(K)^2,
$$

the existence of such a $\underset{\approx}{\boldsymbol{\sigma}}_K^*$ follows from [29]. Moreover, integration by parts also yields that

(5.8)
$$
\left(\underset{\approx}{\boldsymbol{\sigma}}_K^*, \boldsymbol{\nabla} \boldsymbol{v} \right)_K = (\boldsymbol{r}_K, \boldsymbol{v})_K + \sum_{\gamma \in \mathcal{E}_K} (\boldsymbol{R}_{\gamma,K}, \boldsymbol{v})_\gamma \quad \forall \ \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega).
$$

Also, let

$$
\underset{\approx}{\boldsymbol{\sigma}}_K = \underset{\approx}{\boldsymbol{\sigma}}_K^* - \mathbf{curl}\boldsymbol{\zeta} - \frac{1}{2}\text{tr}\left(\underset{\approx}{\boldsymbol{\sigma}}_K^* - \mathbf{curl}\boldsymbol{\zeta} \right) \boldsymbol{I},
$$

where tr denotes the trace of a matrix, $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{\zeta} \in \boldsymbol{H}_0^1(K) \cap \mathbb{P}_3(K)^2$ is such that $\left\| \underset{\approx}{\boldsymbol{\sigma}}_K \right\|_{0,K}$ is minimized. From [27], we know that

(5.9)
$$
\left(\underset{\approx}{\boldsymbol{\sigma}}_K, \boldsymbol{\nabla} \boldsymbol{v} \right)_K = (\boldsymbol{r}_K, \boldsymbol{v})_K + \sum_{\gamma \in \mathcal{E}_K} (\boldsymbol{R}_{\gamma,K}, \boldsymbol{v})_\gamma \quad \forall \ \boldsymbol{v} \in \boldsymbol{X}.
$$

Formulas for the quantities involving $\underset{\approx}{\boldsymbol{\sigma}}_K^*$ and $\underset{\approx}{\boldsymbol{\sigma}}_K$ which are required to compute our estimator, namely, $\left\| \underset{\approx}{\boldsymbol{\sigma}}_K^* \right\|_{0,K}$ and $\left\| \underset{\approx}{\boldsymbol{\sigma}}_K \right\|_{0,K}$, are given in section 5.4.

5.2. Guaranteed upper bounds for the errors. In order to obtain a guaranteed upper bound for the errors we use the Helmholtz-type decomposition of the gradient of the velocity error from [18], given by

(5.10)
$$
\boldsymbol{\nabla} \boldsymbol{e_V} = \boldsymbol{\nabla} \boldsymbol{e}_c + \underset{\approx}{\boldsymbol{e}}_{nc},
$$

where $\boldsymbol{e}_c \in \boldsymbol{X}$ is uniquely defined by

$$(5.11) \qquad (\boldsymbol{\nabla} \boldsymbol{e}_c, \boldsymbol{\nabla} \boldsymbol{v})_\Omega = (\boldsymbol{\nabla} \boldsymbol{e_V}, \boldsymbol{\nabla} \boldsymbol{v})_\Omega \qquad \forall \, \boldsymbol{v} \in \boldsymbol{X},$$

while the remainder part $\underset{\approx}{\boldsymbol{e}}_{nc}$ belongs to the closed subspace

$$(5.12) \qquad \underset{\approx}{\boldsymbol{Y}} = \left\{ \underset{\approx}{\boldsymbol{w}} \in \underset{\approx}{\boldsymbol{L}}^2(\Omega) : \left(\underset{\approx}{\boldsymbol{w}}, \boldsymbol{\nabla} \boldsymbol{v} \right)_\Omega = 0 \, \forall \, \boldsymbol{v} \in \boldsymbol{X} \right\}$$

of $\underset{\approx}{\boldsymbol{L}}^2(\Omega)$. As a consequence of (5.10) and (5.11) we obtain an orthogonal decomposition of the error and hence it satisfies

$$(5.13) \qquad \|\boldsymbol{\nabla} \boldsymbol{e_V}\|_{0,\Omega}^2 = \|\boldsymbol{\nabla} \boldsymbol{e}_c\|_{0,\Omega}^2 + \left\| \underset{\approx}{\boldsymbol{e}}_{nc} \right\|_{0,\Omega}^2.$$

We will call $\boldsymbol{e}_c$ the *conforming part of the error* and $\underset{\approx}{\boldsymbol{e}}_{nc}$ the *nonconforming part of the error*.

Taking $q = 0$ in (5.4), and using (5.11), (5.9), and (2.1), allows us to say that for all $\boldsymbol{v} \in \boldsymbol{X}$,

$$(5.14) \qquad \nu(\boldsymbol{\nabla} \boldsymbol{e}_c, \boldsymbol{\nabla} \boldsymbol{v})_\Omega = \sum_{K \in \mathcal{P}} \left(\left(\underset{\approx}{\boldsymbol{\sigma}}_K, \boldsymbol{\nabla} \boldsymbol{v} \right)_K + (\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v} - \overline{\boldsymbol{v}}_K)_K \right).$$

Hence,

$$(5.15) \qquad \nu^2 \|\boldsymbol{\nabla} \boldsymbol{e}_c\|_{0,\Omega}^2 \le \sum_{K \in \mathcal{P}} \left(\left\| \underset{\approx}{\boldsymbol{\sigma}}_K \right\|_{0,K} + \frac{h_K}{\pi} \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K} \right)^2,$$

upon applying the Cauchy–Schwarz inequality, applying the optimal Poincaré inequality [28]

$$(5.16) \qquad \|\boldsymbol{v} - \overline{\boldsymbol{v}}_K\|_{0,K} \le \frac{h_K}{\pi} \|\boldsymbol{\nabla} \boldsymbol{v}\|_{0,K},$$

and taking $\boldsymbol{v} = \boldsymbol{e}_c \in \boldsymbol{X}$.

To estimate the nonconforming part of the velocity error $\underset{\approx}{\boldsymbol{e}}_{nc}$, we need the following result from [18, Lemma 2].

LEMMA 5.1. *For each $\underset{\approx}{\boldsymbol{w}} \in \underset{\approx}{\boldsymbol{Y}}$ there exists a function $w \in L_0^2(\Omega)$ such that*

$$(5.17) \qquad \left(\underset{\approx}{\boldsymbol{w}}, \boldsymbol{\nabla} \boldsymbol{v} \right)_\Omega = (w, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega \quad \forall \, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega),$$

i.e., $\boldsymbol{\nabla} \cdot \underset{\approx}{\boldsymbol{w}} = \nabla w$. Moreover, w satisfies

$$(5.18) \qquad \|w\|_{0,\Omega} \le \frac{1}{\beta} \left\| \underset{\approx}{\boldsymbol{w}} \right\|_{0,\Omega},$$

where β is the inf-sup constant in (2.6).

Taking $\underset{\approx}{\boldsymbol{w}} \in \underset{\approx}{\boldsymbol{Y}}$ and using (5.10) and (5.12) yields

$$\left(\underset{\approx}{\boldsymbol{e}}_{nc}, \underset{\approx}{\boldsymbol{w}} \right)_\Omega = \left(\boldsymbol{\nabla}(\boldsymbol{e_V} - \boldsymbol{e}_c), \underset{\approx}{\boldsymbol{w}} \right)_\Omega = \left(\boldsymbol{\nabla} \boldsymbol{e_V}, \underset{\approx}{\boldsymbol{w}} \right)_\Omega.$$

It then follows from the previous lemma that

$$\left(\underset{\approx}{\boldsymbol{e}}_{nc}, \underset{\approx}{\boldsymbol{w}} \right)_\Omega = (\boldsymbol{\nabla} \cdot \boldsymbol{e_V}, w)_\Omega = -(\boldsymbol{\nabla} \cdot \boldsymbol{u}_h, w)_\Omega \le \frac{1}{\beta} \|\boldsymbol{\nabla} \cdot \boldsymbol{u}_h\|_{0,\Omega} \|\underset{\approx}{\boldsymbol{w}}\|_{0,\Omega}$$

upon observing that $\boldsymbol{u} \in \boldsymbol{X}$. By taking $\underset{\approx}{\boldsymbol{w}} = \underset{\approx}{\boldsymbol{e}}_{nc} \in \underset{\approx}{\boldsymbol{Y}}$, we can then arrive at the upper bound for the nonconforming error,

$$(5.19) \qquad \|\underset{\approx}{\boldsymbol{e}}_{nc}\|_{0,\Omega} \le \frac{1}{\beta} \|\boldsymbol{\nabla} \cdot \boldsymbol{u}_h\|_{0,\Omega}.$$

Hence, from (5.13), (5.15), and (5.19) we obtain the following result.

THEOREM 5.2. *The gradient of the velocity error e_V can be bounded above as*

$$(5.20) \qquad \nu^2 \|\nabla e_V\|_{0,\Omega}^2 \le \Phi_c^2 + \nu^2 \Phi_{nc}^2,$$

where Φ_c is the conforming estimator defined by

$$(5.21) \qquad \Phi_c = \left(\sum_{K \in \mathcal{P}} \Phi_{c,K}^2 \right)^{1/2} \quad with \quad \Phi_{c,K} = \left\| \underset{\approx}{\boldsymbol{\sigma}}_K \right\|_{0,K} + \frac{h_K}{\pi} \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K},$$

and Φ_{nc} is the nonconforming estimator defined by

$$(5.22) \qquad \Phi_{nc} = \left(\sum_{K \in \mathcal{P}} \Phi_{nc,K}^2 \right)^{1/2} \quad with \quad \Phi_{nc,K} = \frac{1}{\beta} \|\nabla \cdot \boldsymbol{u}_h\|_{0,K}.$$

It remains to give the upper a posteriori error bound for e_P. Taking $q = 0$ in (5.4), and using (5.8) and (2.1), allows us to say that for all $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$,

$$\nu(\nabla e_V, \nabla \boldsymbol{v})_\Omega - (e_P, \nabla \cdot \boldsymbol{v})_\Omega = \sum_{K \in \mathcal{P}} \left(\left(\underset{\approx}{\boldsymbol{\sigma}}_K^*, \nabla \boldsymbol{v} \right)_K + (\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v} - \overline{\boldsymbol{v}}_K)_K \right).$$

Splitting the gradient of the test function $\nabla \boldsymbol{v} = \nabla \boldsymbol{v}_c + \underset{\approx}{\boldsymbol{v}}_{nc}$ as in (5.10) and noticing that since $\boldsymbol{v}_c \in \boldsymbol{X}$, $\left(\underset{\approx}{\boldsymbol{\sigma}}_K^*, \nabla \boldsymbol{v}_c \right)_K = \left(\underset{\approx}{\boldsymbol{\sigma}}_K, \nabla \boldsymbol{v}_c \right)_K$ and $\nabla \cdot \boldsymbol{v} = \mathrm{tr}(\nabla \boldsymbol{v}_c + \underset{\approx}{\boldsymbol{v}}_{nc}) = \mathrm{tr}(\underset{\approx}{\boldsymbol{v}}_{nc})$, where tr denotes the trace of a matrix, then allows us to obtain

$$
\begin{aligned}
(5.23) \qquad & \nu \left((\nabla e_c, \nabla \boldsymbol{v}_c)_\Omega + \left(\underset{\approx}{\boldsymbol{e}}_{nc}, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_\Omega \right) - \left(e_P, \mathrm{tr} \left(\underset{\approx}{\boldsymbol{v}}_{nc} \right) \right)_\Omega \\
& = \sum_{K \in \mathcal{P}} \left(\left(\underset{\approx}{\boldsymbol{\sigma}}_K, \nabla \boldsymbol{v}_c \right)_K + \left(\underset{\approx}{\boldsymbol{\sigma}}_K^*, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_K + (\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v} - \overline{\boldsymbol{v}}_K)_K \right).
\end{aligned}
$$

Now, let $\boldsymbol{\phi}_K \in \boldsymbol{V}_K$ be a solution of the local problem

$$(5.24) \qquad (\nabla \boldsymbol{\phi}_K, \nabla \boldsymbol{v})_K = (\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v} - \overline{\boldsymbol{v}}_K)_K \quad \forall \, \boldsymbol{v} \in \boldsymbol{V}_K,$$

where $\boldsymbol{V}_K = \left\{ \boldsymbol{v} \in \boldsymbol{H}^1(K) : \boldsymbol{v} = \boldsymbol{0} \text{ on } \mathcal{E}_\Gamma \cap \mathcal{E}_K \right\}$. Notice that from (5.24) and (5.16) it easily follows that

$$(5.25) \qquad \|\nabla \boldsymbol{\phi}_K\|_{0,K} \le \frac{h_K}{\pi} \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K}.$$

The same splitting of $\nabla \boldsymbol{v}$ then allows us to say that for all $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$,

$$(\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v} - \overline{\boldsymbol{v}}_K)_K = (\nabla \boldsymbol{\phi}_K, \nabla \boldsymbol{v})_K = (\nabla \boldsymbol{\phi}_K, \nabla \boldsymbol{v}_c)_K + \left(\nabla \boldsymbol{\phi}_K, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_K.$$

Moreover, taking $\boldsymbol{v} = \boldsymbol{v}_c$ in (5.24) then yields that

$$(\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v} - \overline{\boldsymbol{v}}_K)_K = (\boldsymbol{f} - \Pi_K \boldsymbol{f}, \boldsymbol{v}_c - \overline{(\boldsymbol{v}_c)}_K)_K + \left(\nabla \boldsymbol{\phi}_K, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_K.$$

Inserting this expression into (5.23) and using (5.14) then allows us to conclude that

$$
\begin{aligned}
& - \left(e_P, \mathrm{tr} \left(\underset{\approx}{\boldsymbol{v}}_{nc} \right) \right)_\Omega \\
& = \sum_{K \in \mathcal{P}} \left(\left(\underset{\approx}{\boldsymbol{\sigma}}_K^*, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_K + \left(\nabla \boldsymbol{\phi}_K, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_K \right) - \nu \left(\underset{\approx}{\boldsymbol{e}}_{nc}, \underset{\approx}{\boldsymbol{v}}_{nc} \right)_\Omega \\
(5.26) \qquad & \le \left(\left(\sum_{K \in \mathcal{P}} \left(\left\| \underset{\approx}{\boldsymbol{\sigma}}_K^* \right\|_{0,K} + \frac{h_K}{\pi} \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K} \right)^2 \right)^{1/2} + \nu \Phi_{nc} \right) \left\| \underset{\approx}{\boldsymbol{v}}_{nc} \right\|_{0,\Omega}
\end{aligned}
$$

upon using the Cauchy–Schwarz inequality, (5.25), (5.19), and (5.22).

Finally, thanks to the inf-sup condition, we have

$$(5.27) \qquad \beta \|e_P\|_{0,\Omega} \leq \sup_{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \setminus \{\boldsymbol{0}\}} \frac{-(e_P, \boldsymbol{\nabla} \cdot \boldsymbol{v})_\Omega}{\|\boldsymbol{\nabla} \boldsymbol{v}\|_{0,\Omega}} \leq \sup_{\underaccent{\approx}{\boldsymbol{v}}_{nc} \in \underaccent{\approx}{\boldsymbol{Y}}} \frac{-(e_P, \mathrm{tr}(\underaccent{\approx}{\boldsymbol{v}}_{nc}))_\Omega}{\|\underaccent{\approx}{\boldsymbol{v}}_{nc}\|_{0,\Omega}}.$$

Hence, from (5.27) and (5.26) we obtain the following upper bound for the pressure error.

THEOREM 5.3. *The error for the pressure can be bounded above as*

$$(5.28) \qquad \beta \|e_P\|_{0,\Omega} \leq \Phi_c^* + \nu \Phi_{nc},$$

where Φ_c^* is defined by

$$(5.29) \quad \Phi_c^* = \left(\sum_{K \in \mathcal{P}} \left(\Phi_{c,K}^* \right)^2 \right)^{1/2} \quad with \quad \Phi_{c,K}^* = \left\| \underaccent{\approx}{\boldsymbol{\sigma}}_K^* \right\|_{0,K} + \frac{h_K}{\pi} \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K}.$$

From Theorems 5.2 and 5.3 we obtain the following result.

THEOREM 5.4. *The velocity and the pressure errors can be bounded above as*

$$(5.30) \qquad \|(\boldsymbol{e_V}, e_P)\|^2 \leq \Phi_c^2 + \nu^2 \Phi_{nc}^2 + (\Phi_c^* + \nu \Phi_{nc})^2.$$

This estimator presents a significant improvement over the one from [1], since the bounds are tighter.

Remark 5.1. An alternative approach can be obtained using Lemma 3.1 in [21]. For simplicity we consider the case when $\nu = 1$, where the above cited result yields

$$\|(\boldsymbol{e_V}, e_P)\| \leq \frac{\sqrt{5} + 1}{2} \sup_{(\boldsymbol{v}, q) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)} \frac{\mathcal{A}(\boldsymbol{e_V}, e_P; \boldsymbol{v}, q)}{\|(\boldsymbol{v}, q)\|},$$

and then using (5.4), (5.8), (2.1), (5.16), the Cauchy–Schwarz inequality, and the definitions of Φ_c^* and Φ_{nc} given in (5.29) and (5.22), respectively, we obtain

$$\mathcal{A}(\boldsymbol{e_V}, e_P; \boldsymbol{v}, q) \leq \left((\Phi_c^*)^2 + \Phi_{nc}^2 \right)^{1/2} \|(\boldsymbol{v}, q)\|.$$

As an immediate consequence it follows that

$$(5.31) \qquad \|(\boldsymbol{e_V}, e_P)\|^2 \leq \left(\frac{\sqrt{5} + 1}{2} \right)^2 \left((\Phi_c^*)^2 + \Phi_{nc}^2 \right).$$

Notice that if we were just interested in the estimation of the velocity field, then the estimation using the orthogonal decomposition (5.10) leads to a tighter upper bound, i.e., from Theorem 5.2 we obtain

$$\|\boldsymbol{\nabla} \boldsymbol{e_V}\|_{0,\Omega}^2 \leq \Phi_c^2 + \Phi_{nc}^2 < \left(\frac{\sqrt{5} + 1}{2} \right)^2 \left((\Phi_c^*)^2 + \Phi_{nc}^2 \right).$$

Likewise, if one wishes to only estimate the pressure error, a superior upper bound again follows by using the orthogonal decomposition, i.e., from Theorem 5.3 we obtain

$$\beta^2 \|e_P\|_{0,\Omega}^2 \leq (\Phi_c^* + \Phi_{nc})^2 < \left(\frac{\sqrt{5} + 1}{2} \right)^2 \left((\Phi_c^*)^2 + \Phi_{nc}^2 \right).$$

If one wishes to estimate the combined error, (5.30) and (5.31) yield

$$\|(e_V, e_P)\| \le \eta(\alpha)^2 = \Phi_c^2 + \Phi_{nc}^2 + (\Phi_c^* + \Phi_{nc})^2,$$

$$\|(e_V, e_P)\| \le \tilde{\eta}(\alpha)^2 = \left(\frac{\sqrt{5}+1}{2}\right)^2 \left((\Phi_c^*)^2 + \Phi_{nc}^2\right),$$

which in turn gives

$$0.618 \approx \frac{2}{1+\sqrt{5}} \le \frac{\theta}{\tilde{\theta}} = \frac{\eta(\alpha)}{\tilde{\eta}(\alpha)} \le \frac{2\sqrt{3}}{1+\sqrt{5}} \approx 1.07,$$

where $\theta = \eta(\alpha)/\|(e_V, e_P)\|$ and $\tilde{\theta} = \tilde{\eta}(\alpha)/\|(e_V, e_P)\|$ are the effectivity indices. Hence,

$$1 \le \theta \le 1.07 \, \tilde{\theta} \qquad \text{and} \qquad 1 \le \tilde{\theta} \le 1.618 \, \theta,$$

leading to the conclusion that the estimator $\eta(\alpha)$ is in general a sharper bound.

5.3. Computation of the equilibrated boundary fluxes. We now describe a procedure for obtaining a set of boundary fluxes $\{\boldsymbol{g}_{\gamma,K}\}$ satisfying (5.2) and (5.3). The procedure is a slight extension of the procedure described in section 6.4 of [2] and so we just briefly outline the main steps.

Let

$$\langle \boldsymbol{J} \rangle_{\gamma,K} = \left\{ \begin{array}{ll} \frac{1}{2} (\boldsymbol{J}_{\gamma,K} - \boldsymbol{J}_{\gamma,K'}) & \text{if } \gamma \in \mathcal{E}_K \cap \mathcal{E}_{K'} \text{ for } K' \ne K, \\ \boldsymbol{J}_{\gamma,K} & \text{if } \gamma \in \mathcal{E}_K \cap \mathcal{E}_\Gamma, \end{array} \right.$$

and

$$[\boldsymbol{J}]_\gamma = \left\{ \begin{array}{ll} \frac{1}{2} (\boldsymbol{J}_{\gamma,K} + \boldsymbol{J}_{\gamma,K'}) & \text{if } \gamma \in \mathcal{E}_K \cap \mathcal{E}_{K'} \text{ for distinct } K, K' \in \mathcal{P}, \\ \boldsymbol{0} & \text{if } \gamma \in \mathcal{E}_K \cap \mathcal{E}_\Gamma \text{ for } K \in \mathcal{P}, \end{array} \right.$$

with $\boldsymbol{J}_{\gamma,K} = \nu \boldsymbol{\nabla} \boldsymbol{u}_{h|K} \hat{\boldsymbol{n}}_\gamma^K - p_{h|K} \hat{\boldsymbol{n}}_\gamma^K$ and $\boldsymbol{J}_{\gamma,K'} = \nu \boldsymbol{\nabla} \boldsymbol{u}_{h|K'} \hat{\boldsymbol{n}}_\gamma^{K'} - p_{h|K'} \hat{\boldsymbol{n}}_\gamma^{K'}$. Since the flux $\boldsymbol{g}_{\gamma,K} \in \mathbb{P}_1(\gamma)^2$, it is uniquely determined by the moments

$$(5.32) \qquad \mu_{K,n}^{\gamma,i} = \left(\boldsymbol{g}_{\gamma,K}, \boldsymbol{\lambda}_n^{(i)}\right)_\gamma \text{ for } n : \boldsymbol{x}_n \in \mathcal{V}_\gamma \text{ and } i = 1, 2.$$

Moreover, we can satisfy (5.2) and (5.3) by taking the moments
(5.33)
$$\mu_{K,n}^{\gamma,i} = \left\{ \begin{array}{ll} \frac{1}{2} \left(\xi_{K,n}^{(i)} - \xi_{K',n}^{(i)}\right) + \left(\langle \boldsymbol{J} \rangle_{\gamma,K}, \boldsymbol{\lambda}_n^{(i)}\right)_\gamma & \text{if } \gamma \in \mathcal{E}_K \cap \mathcal{E}_{K'} \text{ for } K' \ne K, \\ \xi_{K,n}^{(i)} + \left(\boldsymbol{J}_{\gamma,K}, \boldsymbol{\lambda}_n^{(i)}\right)_\gamma & \text{if } \gamma \in \mathcal{E}_K \cap \mathcal{E}_\Gamma, \end{array} \right.$$

where the $\xi_{K,n}^{(i)}$ are obtained by solving the system of equations

$$(5.34) \qquad \frac{1}{2} \sum_{K' \in \Omega_n \cap \Omega_K} \left(\xi_{K,n}^{(i)} - \xi_{K',n}^{(i)}\right) + \sum_{\gamma \in \mathcal{E}_K \cap \mathcal{E}_\Gamma \cap \mathcal{E}_n} \xi_{K,n}^{(i)} = \Delta_{K,n}^{(i)} \quad \forall \, K \in \Omega_n,$$

where
(5.35)
$$\Delta_{K,n}^{(i)} = \nu \left(\boldsymbol{\nabla} \boldsymbol{u}_h, \boldsymbol{\nabla} \boldsymbol{\lambda}_n^{(i)}\right)_K - \left(p_h, \boldsymbol{\nabla} \cdot \boldsymbol{\lambda}_n^{(i)}\right)_K - \left(\boldsymbol{f}, \boldsymbol{\lambda}_n^{(i)}\right)_K - \sum_{\gamma \in \mathcal{E}_K} \left(\langle \boldsymbol{J} \rangle_{\gamma,K}, \boldsymbol{\lambda}_n^{(i)}\right)_\gamma.$$

The above system consists of $\sharp\Omega_n$ equations for $\sharp\Omega_n$ unknowns, where $\sharp\Omega_n$ denotes the cardinality of Ω_n. Note that even though the linear system (5.34) will not always have a unique solution, the same arguments used in [2] allow us to conclude that (5.34) will always have at least one solution and that any solution to (5.34) yields unique moments defined by (5.33) which are such that

(5.36)

$$
h_K^{1/2} \left\| \boldsymbol{R}_{\gamma,K} \right\|_{0,\gamma} = h_K^{1/2} \left\| \boldsymbol{g}_{\gamma,K} - \boldsymbol{J}_{\gamma,K} \right\|_{0,\gamma}
$$

$$
\leq C \sum_{n:\boldsymbol{x}_n \in \mathcal{V}_\gamma} \sum_{K' \in \Omega_n} \left(h_{K'} \| \boldsymbol{r}_{K'} \|_{0,K'} + h_{K'}^{1/2} \sum_{\gamma' \in \mathcal{E}_{K'} \cap \mathcal{E}_n} \| [\boldsymbol{J}]_{\gamma'} \|_{0,\gamma'} \right).
$$

We note that the fact that the stabilization parameter α does not appear explicitly in (5.33), (5.34), or (5.35) means that, as well as being independent of ν and any mesh size, the positive constant C in the preceding inequality is independent of α.

5.4. Expressions for $\|\underset{\approx}{\boldsymbol{\sigma}}_K^*\|_{0,K}$ and $\|\underset{\approx}{\boldsymbol{\sigma}}_K\|_{0,K}$. Let the edges and vertices of element K be labeled as in Figure 10 and let the tangent vectors $\boldsymbol{t}_1 = \boldsymbol{x}_3 - \boldsymbol{x}_2$, $\boldsymbol{t}_2 = \boldsymbol{x}_1 - \boldsymbol{x}_3$, and $\boldsymbol{t}_3 = \boldsymbol{x}_2 - \boldsymbol{x}_1$. Let

$$
\hat{\underset{\approx}{\boldsymbol{\sigma}}}_K = \sum_{l=1}^2 \left(\sum_{i=1}^3 \underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_i}^{(l)} + c_l \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(l)} \right),
$$

where

$$
\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_1}^{(l)} = \frac{1}{2\,|K|} \left(\left(\boldsymbol{R}_{\gamma_1,K}, \boldsymbol{\lambda}_2^{(l)} \right)_{\gamma_1} \left((2\lambda_3 - 3\lambda_2 - \lambda_1)\lambda_3 \underset{\approx}{\boldsymbol{\tau}}_2^{(l)} + (4\lambda_2 - \lambda_3 - 7\lambda_1)\lambda_2 \underset{\approx}{\boldsymbol{\tau}}_3^{(l)} \right) \right.
$$

$$
\left. - \left(\boldsymbol{R}_{\gamma_1,K}, \boldsymbol{\lambda}_3^{(l)} \right)_{\gamma_1} \left((4\lambda_3 - \lambda_2 - 7\lambda_1)\lambda_3 \underset{\approx}{\boldsymbol{\tau}}_2^{(l)} + (2\lambda_2 - 3\lambda_3 - \lambda_1)\lambda_2 \underset{\approx}{\boldsymbol{\tau}}_3^{(l)} \right) \right)
$$

with $\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_2}^{(l)}$ and $\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_3}^{(l)}$ being defined by permuting the indices,

$$
\underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(l)} = -\frac{4}{|K|} \left(\lambda_2 \lambda_3 \underset{\approx}{\boldsymbol{\tau}}_1^{(l)} + \lambda_3 \lambda_1 \underset{\approx}{\boldsymbol{\tau}}_2^{(l)} + \lambda_1 \lambda_2 \underset{\approx}{\boldsymbol{\tau}}_3^{(l)} \right),
$$

and $\begin{bmatrix} c_1 & c_2 \end{bmatrix} = \boldsymbol{B}^T \boldsymbol{A}^{-1}$, where

$$
\boldsymbol{A} = \begin{bmatrix} \left(\underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(1)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(1)} \right)_K & \left(\underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(1)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(2)} \right)_K \\ \left(\underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(2)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(1)} \right)_K & \left(\underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(2)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(2)} \right)_K \end{bmatrix} \text{ and } \boldsymbol{B} = \begin{pmatrix} \sum_{l=1}^2 \sum_{i=1}^3 \left(\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_i}^{(l)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(1)} \right)_K \\ \sum_{l=1}^2 \sum_{i=1}^3 \left(\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_i}^{(l)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(2)} \right)_K \end{pmatrix}
$$
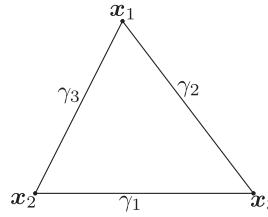
FIG. 10. *The labeling of the edges and vertices of element K.*

with

$$(5.37) \quad \underset{\approx}{\boldsymbol{\tau}}_i^{(1)} = \left[\begin{array}{c} \boldsymbol{t}_i^T \\ \boldsymbol{0} \end{array} \right] - \frac{\varrho}{2} \text{tr} \left(\left[\begin{array}{c} \boldsymbol{t}_i^T \\ \boldsymbol{0} \end{array} \right] \right) \boldsymbol{I} \text{ and } \underset{\approx}{\boldsymbol{\tau}}_i^{(2)} = \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{t}_i^T \end{array} \right] - \frac{\varrho}{2} \text{tr} \left(\left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{t}_i^T \end{array} \right] \right) \boldsymbol{I},$$

where $\boldsymbol{0} = \left[\begin{array}{cc} 0 & 0 \end{array} \right]$ and the values taken by ϱ are specified below. Now, (5.7) and tedious but straightforward calculations can be used to show that if we take $\varrho = 0$ in (5.37), then $\underset{\approx}{\boldsymbol{\sigma}}_K^* = \hat{\underset{\approx}{\boldsymbol{\sigma}}}_K$. It then follows that if we take $\varrho = 1$ in (5.37), then $\underset{\approx}{\boldsymbol{\sigma}}_K = \hat{\underset{\approx}{\boldsymbol{\sigma}}}_K$.

Let $\underset{\approx}{\boldsymbol{D}} : \underset{\approx}{\boldsymbol{F}} = D_{11}F_{11} + D_{12}F_{12} + D_{21}F_{21} + D_{22}F_{22}$. Then, for $l, m = 1, 2$,

$$\left(\underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(l)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(m)} \right)_K = \frac{1}{720 \, |K|} \Big(128 \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_2^{(m)} \right) + 64 \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right)$$
$$+ 64 \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_2^{(m)} \right) + 128 \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right) \Big)$$

and, for $i = 1, 2, 3$ and $l, m = 1, 2$,

$$\left(\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_i}^{(l)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,0}^{(m)} \right)_K = \frac{1}{720 \, |K|} \left(\vec{S}_i^{(l)} \right)^T \boldsymbol{M}_{i0}^{(l,m)},$$

where

$$\vec{S}_1^{(l)} = \left(\begin{array}{cc} \left(\boldsymbol{R}_{\gamma_1,K}, \boldsymbol{\lambda}_2^{(l)} \right)_{\gamma_1} & \left(\boldsymbol{R}_{\gamma_1,K}, \boldsymbol{\lambda}_3^{(l)} \right)_{\gamma_1} \end{array} \right)^T$$

with $\vec{S}_2^{(l)}$ and $\vec{S}_3^{(l)}$ being defined by permuting the indices and

$$\boldsymbol{M}_{10}^{(l,m)} = \left[\begin{array}{c} -16 \\ -48 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_2^{(m)} \right) + \left[\begin{array}{c} 8 \\ -56 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right)$$
$$+ \left[\begin{array}{c} 56 \\ -8 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_2^{(m)} \right) + \left[\begin{array}{c} 48 \\ 16 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right)$$

with $\boldsymbol{M}_{20}^{(l,m)}$ and $\boldsymbol{M}_{30}^{(l,m)}$ being defined by permuting the indices. Moreover, for $i, j = 1, 2, 3$ and $l, m = 1, 2$,

$$\left(\underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_i}^{(l)}, \underset{\approx}{\boldsymbol{\sigma}}_{K,\gamma_j}^{(m)} \right)_K = \frac{1}{720 \, |K|} \left(\vec{S}_i^{(l)} \right)^T \boldsymbol{M}_{ij}^{(l,m)} \vec{S}_j^{(m)},$$

where

$$\boldsymbol{M}_{11}^{(l,m)} = \left[\begin{array}{cc} 26 & -42 \\ -42 & 114 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_2^{(m)} \right) + \left[\begin{array}{cc} -9 & 7 \\ -17 & -9 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right)$$
$$+ \left[\begin{array}{cc} -9 & -17 \\ 7 & -9 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_2^{(m)} \right) + \left[\begin{array}{cc} 114 & -42 \\ -42 & 26 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right)$$

with $\boldsymbol{M}_{22}^{(l,m)}$ and $\boldsymbol{M}_{33}^{(l,m)}$ being defined by permuting the indices and

$$\boldsymbol{M}_{12}^{(l,m)} = \left[\begin{array}{cc} -5 & -5 \\ 15 & -25 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right) + \left[\begin{array}{cc} 54 & -22 \\ -78 & 54 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_2^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_1^{(m)} \right)$$
$$+ \left[\begin{array}{cc} -13 & 31 \\ -1 & -13 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_3^{(m)} \right) + \left[\begin{array}{cc} -25 & -5 \\ 15 & -5 \end{array} \right] \left(\underset{\approx}{\boldsymbol{\tau}}_3^{(l)} : \underset{\approx}{\boldsymbol{\tau}}_1^{(m)} \right)$$

with $\boldsymbol{M}_{23}^{(l,m)}$ and $\boldsymbol{M}_{31}^{(l,m)}$ being defined by permuting the indices. So, upon taking $\varrho = 0$ in (5.37),

$$\left\|\underset{\approx}{\boldsymbol{\sigma}}_K^*\right\|_{0,K}^2 = \sum_{l=1}^2 \sum_{i=1}^3 \sum_{j=1}^3 \left(\boldsymbol{\sigma}_{K,\gamma_i}^{(l)}, \boldsymbol{\sigma}_{K,\gamma_j}^{(l)}\right) - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B},$$

and, upon taking $\varrho = 1$ in (5.37),

$$\left\|\underset{\approx}{\boldsymbol{\sigma}}_K\right\|_{0,K}^2 = \sum_{l=1}^2 \sum_{m=1}^2 \sum_{i=1}^3 \sum_{j=1}^3 \left(\boldsymbol{\sigma}_{K,\gamma_i}^{(l)}, \boldsymbol{\sigma}_{K,\gamma_j}^{(m)}\right) - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B}.$$

Furthermore, standard arguments can be used to show that

$$(5.38) \qquad \left\|\underset{\approx}{\boldsymbol{\sigma}}_K\right\|_{0,K} \leq \left\|\underset{\approx}{\boldsymbol{\sigma}}_K^*\right\|_{0,K} \leq C h_K^{1/2} \sum_{\gamma \in \mathcal{E}_K} \|\boldsymbol{R}_{\gamma,K}\|_{0,\gamma}.$$

Remark 5.2. Notice that there is no need to reconstruct the real boundary flux $\boldsymbol{g}_{\gamma,K}$ in order to compute $\left\|\underset{\approx}{\boldsymbol{\sigma}}_K^*\right\|_{0,K}$ and $\left\|\underset{\approx}{\boldsymbol{\sigma}}_K\right\|_{0,K}$ since the formulas for these quantities involve only the moments of the fluxes weighted against the basis functions on the edge γ, i.e., we only need $\mu_{K,i}^{\gamma,l}$, given by (5.32), since

$$\left(\boldsymbol{R}_{\gamma,K}, \boldsymbol{\lambda}_i^{(l)}\right)_\gamma = \mu_{K,i}^{\gamma,l} + \left(-\nu \boldsymbol{\nabla} \boldsymbol{u}_{h|K} \hat{\boldsymbol{n}}_\gamma^K + p_{h|K} \hat{\boldsymbol{n}}_\gamma^K, \boldsymbol{\lambda}_i^{(l)}\right)_\gamma.$$

5.5. Local lower bounds. First we focus on bounding $\Phi_{c,K}$ and $\Phi_{c,K}^*$. Since, $\|\underset{\approx}{\boldsymbol{\sigma}}_K\|_{0,K} \leq \|\underset{\approx}{\boldsymbol{\sigma}}_K^*\|_{0,K}$, using (5.38) yields that

$$(5.39) \qquad \Phi_{c,K}^2 \leq \left(\Phi_{c,K}^*\right)^2 \leq C \left(h_K \sum_{\gamma \in \mathcal{E}_K} \|\boldsymbol{R}_{\gamma,K}\|_{0,\gamma}^2 + h_K^2 \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K}^2\right).$$

By applying standard bubble function arguments [2, 33] to (5.1), it can be proved that for all $K \in \mathcal{P}$,

$$h_K \|\boldsymbol{r}_K\|_{0,K} \leq C \left(\nu \|\boldsymbol{\nabla} \boldsymbol{e}_V\|_{0,K} + \beta \|e_P\|_{0,K} + h_K \|\boldsymbol{f} - \Pi_K \boldsymbol{f}\|_{0,K}\right)$$

and

$$h_K^{1/2} \sum_{\gamma \in \mathcal{E}_K} \|[\boldsymbol{J}]_\gamma\|_{0,\gamma} \leq C \sum_{K' \in \Omega_K} \left(\nu \|\boldsymbol{\nabla} \boldsymbol{e}_V\|_{0,K'} + \beta \|e_P\|_{0,K'} + h_{K'} \|\boldsymbol{f} - \Pi_{K'} \boldsymbol{f}\|_{0,K'}\right),$$

where, as well as being independent of ν and any mesh size, C is independent of α as the stabilization parameter α does not appear explicitly in (5.1). Combining these bounds with (5.36) then allows us to see that
(5.40)
$$h_K \sum_{\gamma \in \mathcal{E}_K} \|\boldsymbol{R}_{\gamma,K}\|_{0,\gamma}^2 \leq C \sum_{K' \in \tilde{\Omega}_K} \left(\nu^2 \|\boldsymbol{\nabla} \boldsymbol{e}_V\|_{0,K'}^2 + \beta \|e_P\|_{0,K'}^2 + h_{K'}^2 \|\boldsymbol{f} - \Pi_{K'} \boldsymbol{f}\|_{0,K'}^2\right).$$

Using (5.40) to bound (5.39) allows us to conclude that

$$(5.41) \qquad \begin{aligned} \Phi_{c,K}^2 &\leq \left(\Phi_{c,K}^*\right)^2 \\ &\leq C \sum_{K' \in \tilde{\Omega}_K} \left(\nu^2 \|\boldsymbol{\nabla} \boldsymbol{e}_V\|_{0,K'}^2 + \beta \|e_P\|_{0,K'}^2 + h_{K'}^2 \|\boldsymbol{f} - \Pi_{K'} \boldsymbol{f}\|_{0,K'}^2\right). \end{aligned}$$

The bound for $\Phi_{nc,K}$ can be arrived at upon noting that since the solution $\boldsymbol{u} \in \boldsymbol{X}$,

$$(5.42) \qquad \Phi_{nc,K} = \frac{1}{\beta}\|\boldsymbol{\nabla} \cdot \boldsymbol{u}_h\|_{0,K} = \frac{1}{\beta}\|\boldsymbol{\nabla} \cdot (\boldsymbol{u}_h - \boldsymbol{u})\|_{0,K} \leq \frac{\sqrt{2}}{\beta}\,\|\boldsymbol{\nabla}\boldsymbol{e_V}\|_{0,K}\,.$$

The local efficiency from (3.4) follows at once using (5.41) and (5.42).

6. Conclusion. A computable a posteriori error estimator is derived which can be applied to a wide class of stabilized finite element approximations of the Stokes problem. The estimator generally provides tight guaranteed upper bounds on the velocity and pressure errors along with a local lower bound, both of which are independent of the stabilization parameter. The estimator is then used as the basis for the selection of the stabilization parameter in conjunction with a trust-region DFO search. Numerical results revealed that the approach yields a good approximation to the optimal value for α. Moreover, it was seen that the best choice is not always the same as the recommended value.

REFERENCES

[1] M. AINSWORTH, A. ALLENDES, G. R. BARRENECHEA, AND R. RANKIN, *Computable error bounds for nonconforming Fortin–Soulie finite element approximation of the Stokes problem*, IMA J. Numer. Anal., 32 (2012), pp. 417–447.

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

[3] R. ARAYA, G. R. BARRENECHEA, AND F. VALENTIN, *Stabilized finite element methods based on multiscale enrichment for the Stokes problem*, SIAM J. Numer. Anal., 44 (2006), pp. 322–348.

[4] C. BAIOCCHI, F. BREZZI, AND L. P. FRANCA, *Virtual bubbles and Galerkin-least-squares type methods (Ga. L. S.)*, Comput. Methods Appl. Mech. Engrg., 105 (1993), pp. 125–141.

[5] G. R. BARRENECHEA AND F. VALENTIN, *Consistent local projection stabilized finite element methods*, SIAM J. Numer. Anal, 48 (2010), pp. 1801–1825.

[6] T. BARTH, P. B. BOCHEV, M. GUNZBURGER, AND J. SHADID, *A taxonomy of consistently stabilized finite element methods for the Stokes problem*, SIAM J. Sci. Comput., 25 (2004), pp. 1585–1607.

[7] R. BECKER AND M. BRAACK, *A finite element pressure gradient stabilization for the Stokes equations based on local projections*, Calcolo, 38 (2001), pp. 173–199.

[8] C. BERNARDI, V. GIRAULT, AND V. HECHT, *A posteriori analysis of a penalty method and application to the Stokes problem*, Math. Models Methods Appl. Sci., 13 (2003), pp. 1599–1628.

[9] P. B. BOCHEV, C. R. DOHRMANN, AND M. D. GUNZBURGER, *Stabilization of low-order mixed finite elements for the Stokes equations*, SIAM J. Numer. Anal, 44 (2006), pp. 82–101.

[10] D. BOFFI AND L. GASTALDI, *Mixed Finite Elements, Compatibility Conditions, and Applications*, Springer-Verlag, Berlin, 2006.

[11] F. BREZZI, T. J. R. HUGHES, L. D. MARINI, A. RUSSO, AND E. SÜLI, *A priori error analysis of finite element method with residual-free bubbles for advection-dominated equations*, SIAM J. Numer. Anal., 36 (1999), pp. 1933–1948.

[12] F. BREZZI AND J. PITKÄRANTA, *On the stabilization of finite element approximations of the Stokes equations*, in Efficient Solution of Elliptic Systems, W. Hackbush, ed., Vieweg, Braunschweig, 1984, pp. 11–19.

[13] G. F. CAREY AND R. KRISHNAN, *Penalty approximation of Stokes flow*, Comput. Methods Appl. Mech. Engrg., 35 (1982), pp. 169–206.

[14] A. R. CONN, K. SCHEINBERG, AND P. TOINT, *On the convergence of derivative-free methods for unconstrained optimization*, in Approximation Theory and Optimization, Tributes to M. J. D. Powell, M. D. Buhmann and A. Iserles eds., Cambridge University Press, Cambridge, UK, (1997), pp. 83–108.

[15] A. R. CONN, K. SCHEINBERG, AND L. N. VICENTE, *Geometry of interpolation sets in derivative free optimization*, Math. Program. Ser. B 111 (2008), pp. 141–172.

[16] A. R. CONN, K. SCHEINBERG, AND L. N. VICENTE, *Global convergence of general derivative-free trust-region algorithms to first and second order critical points*, SIAM J. Optim., 20 (2009), pp. 387–415.

[17] C. R. DOHRMANN AND P. B. BOCHEV, *A stabilized finite element method for the Stokes problem based on polynomial pressure projections*, Internat. J. Numer. Methods Fluids, 46 (2004), pp. 183–201.

[18] W. DÖRFLER AND M. AINSWORTH, *Reliable a posteriori error control for nonconforming finite element approximation of Stokes flow*, Math. Comput., 74 (2005), pp. 1599–1616.

[19] J. DOUGLAS AND J. WANG, *An absolutely stabilized finite element method for the Stokes problem*, Math. Comput., 52 (1989), pp. 495–508.

[20] V. GIRAULT AND P. A. RAVIART, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Ser. Comput. Math., 5 Springer, New York, 1986.

[21] A. HANNUKAINEN, R. STENBERG, AND M. VOHRALÍK, *A unified framework for a posteriori error estimation for the Stokes problem*, Numer. Math., 122 (2012), pp. 725–769.

[22] T. J. R. HUGHES, G. R. FEIJOO, L. MAZZEI, AND J. QUINCY, *The variational multiscale method: A paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg., 166 (1998), pp. 3–24.

[23] T. J. R. HUGHES, L. P. FRANCA, AND M. BALESTRA, *A new finite element formulation for computational fluids dynamics: V. Circumventing the Babuska–Brezzi condition: A stable Petrov–Galerkin formulation of the Stokes problem accommodating equal-order interpolations*, Comput. Methods Appl. Mech. Engrg., 59 (1986), pp. 85–99.

[24] T. J. R. HUGHES AND L. P. FRANCA, *A new finite element formulation for computational fluid dynamics: VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity/pressure spaces*, Comput. Methods Appl. Mech. Engrg., 65 (1987), pp. 85–96.

[25] V. JOHN, P. KNOBLOCH, AND S. B. SAVESCU, *A posteriori optimization of parameters in stabilized methods for convection-diffusion problems—Part* I, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 2916–2929.

[26] N. KECHKAR AND D. SILVESTER, *Analysis of locally stabilised mixed finite element methods for the Stokes problem*, Math. Comp., 58 (1992), pp. 1–10.

[27] K.-Y. KIM, *Flux reconstruction for the P2 nonconforming finite element method with application to a posteriori error estimation*, Appl. Numer. Math., 62 (2012), pp. 1701–1717.

[28] L. E. PAYNE AND H. F. WEINBERGER, *An optimal Poincaré inequality for convex domains*, Arch. Rational Mech. Anal., 5 (1960), pp. 286–292.

[29] P. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for* 2*nd order elliptic problems*. in Mathematical Aspects of Finite Element Methods, Lecture Notes in Math. 606, Springer Verlag, Berlin, 1977.

[30] H. ROOS, M. STYNES, AND L. TOBISKA, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, Springer Ser. Comput. Math., Springer-Verlag, Berlin, 2008.

[31] D. SILVESTER, *Optimal low order finite element methods for incompressible flow*, Comput. Methods Appl. Mech. Engrg., 111 (1994), pp. 357–368.

[32] G. STOYAN, *Towards discrete Velte decompositions and narrow bounds for the inf-sup constants*, Comput. Math. Appl., 38 (1999), pp. 243–261.

[33] R. VERFÜRTH, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Chichester, UK, 1996.