INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2005; **49**:877–903
Published online 3 August 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1029

A mixed finite element method for the generalized Stokes problem

Rommel Bustinza[1,‡], Gabriel N. Gatica[1,*,†] and María González[2,§]

[1]*Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile*
[2]*Departamento de Matemáticas, Universidade da Coruña, 15071 A Coruña, España, Spain*

SUMMARY

We present and analyse a new mixed finite element method for the generalized Stokes problem. The approach, which is a natural extension of a previous procedure applied to quasi-Newtonian Stokes flows, is based on the introduction of the flux and the tensor gradient of the velocity as further unknowns. This yields a two-fold saddle point operator equation as the resulting variational formulation. Then, applying a slight generalization of the well known Babuška–Brezzi theory, we prove that the continuous and discrete formulations are well posed, and derive the associated *a priori* error analysis. In particular, the finite element subspaces providing stability coincide with those employed for the usual Stokes flows except for one of them that needs to be suitably enriched. We also develop an *a posteriori* error estimate (based on local problems) and propose the associated adaptive algorithm to compute the finite element solutions. Several numerical results illustrate the performance of the method and its capability to localize boundary layers, inner layers, and singularities. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: mixed finite elements; Stokes problem; *a priori* error estimate; *a posteriori* error estimate

1. INTRODUCTION

The generalized Stokes problem, which is a Stokes-like linear system with a dominating zeroth order term, arises naturally in the time discretization of the corresponding non-steady equations, and hence it plays a fundamental role in the numerical simulation of viscous incompressible flows (laminar and turbulent). Indeed, the most expensive part of the solution procedure for the time-dependent Navier–Stokes equations reduces to solving the generalized Stokes problem at each nonlinear iteration. In order to define it explicitly, we first let Ω be a

*Correspondence to: Gabriel N. Gatica, GI[2]MA, Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile.
†E-mail: ggatica@ing-mat.udec.cl
‡E-mail: rbustinz@ing-mat.udec.cl
§E-mail: mgtaboad@udc.es

Contract/grant sponsor: CONICYT-Chile

Received 1 February 2005
Revised 25 May 2005
Accepted 26 May 2005

bounded open subset of $\mathbb{R}^2$ with Lipschitz continuous boundary Γ. Then, given $\mathbf{f} \in [L^2(\Omega)]^2$ and $\mathbf{g} \in [H^{1/2}(\Gamma)]^2$, we look for the velocity $\mathbf{u} := (u_1, u_2)^t$ and the pressure p of a fluid occupying the region Ω, such that

$$\begin{aligned} \alpha\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \operatorname{div}(\mathbf{u}) &= 0 \quad \text{in } \Omega \\ \mathbf{u} &= \mathbf{g} \quad \text{on } \Gamma \end{aligned} \tag{1}$$

where ν is a positive constant called kinematic viscosity of the fluid and α is a positive parameter proportional to the inverse of the time-step. Throughout the rest of the paper we assume that $\alpha \geqslant \nu$. Also, we remark that, as a consequence of the incompressibility of the fluid, the Dirichlet datum $\mathbf{g}$ must satisfy the compatibility condition $\int_\Gamma \mathbf{g} \cdot \mathbf{n}\, ds = 0$, where $\mathbf{n}$ is the unit outward normal to Γ.

In recent years considerable effort has gone into the design and study of efficient numerical methods to solve (1). The new proposed algorithms apply and combine different techniques, which include Uzawa's schemes, splittings of boundary conditions, fictitious domains, domain decomposition, stabilization, and preconditioning (see, e.g. References [1–7], and the references therein). A common feature to these papers is that they all deal with the usual pressure–velocity variational formulation of the problem, in which the unknowns live in $L^2(\Omega)$ and $H^1(\Omega)$, respectively. In particular, this means that the finite element subspace for the velocity needs to be a subset of the continuous functions. In addition, the Dirichlet boundary condition, being essential and non-homogeneous, cannot be incorporated either in the continuous and discrete formulations or in the definitions of the spaces involved, and therefore one is necessarily led to a non-conforming Galerkin scheme. Certainly, the latter concern refers to the theoretical analysis of the method since the interpolation of essential boundary conditions causes no problems in the practical implementation of the corresponding Galerkin scheme.

On the other hand, within a dual-mixed setting the velocity becomes an unknown in $L^2(\Omega)$, which gives more flexibility to choose the associated finite element subspace (for instance, piecewise constant functions become a feasible choice). Furthermore, the Dirichlet boundary condition, being now natural, is incorporated directly into the right-hand sides (linear functionals) of the continuous and discrete formulations, and hence the error analysis arising from a non-conforming scheme is avoided. Another important advantage of using dual-mixed methods lies on the possibility of introducing further unknowns with a physical interest (for instance, the flux). These unknowns are then approximated directly, which avoids any numerical postprocessing yielding additional sources of error.

As a recent example of the above, we recall here that in References [8, 9] we introduce and analyse a dual-mixed formulation for a class of quasi-Newtonian Stokes flows whose kinematic viscosities are nonlinear monotone functions of the gradient of the velocity. The mixed finite element method proposed there simply relies on the introduction of the flux and the tensor gradient of the velocity as auxiliary unknowns, which yields a two-fold saddle point operator equation as the resulting variational formulation. Therefore, the abstract theory developed in Reference [10], which is a slight generalization of the well known Babuška–Brezzi theory, is applied to prove that the continuous and discrete schemes are well posed. In particular, it is shown that the stability of the Galerkin scheme only requires low-order finite

element subspaces: it suffices to use Raviart–Thomas spaces of order zero to approximate the flux and piecewise constant functions to approximate the other unknowns. In addition, since the monotonicity certainly includes the linear case, we also obtain as a by-product a new mixed finite element method for the linear Stokes equation (problem (1) with $\alpha = 0$).

Up to the authors' knowledge, there are no dual-mixed methods available in the current literature for the generalized Stokes problem. This gap is somehow filled by the present work, and hence we believe this is one of the main motivations for it. More precisely, the purpose of this paper is to extend the analysis from References [8, 9] to our problem (1) with moderately large values of the parameter α, which includes, similarly as we did in Reference [9], the derivation of *a posteriori* error estimates based on local problems. The rest of the paper is organized as follows. In Section 2, we derive the continuous dual-mixed variational formulation of problem (1) and prove that it is well posed. Then, in Section 3, we present and analyse the corresponding mixed finite element scheme. Again, we use Raviart–Thomas spaces of order zero to approximate the flux and piecewise constant functions to approximate the velocity and the pressure. However, in order to guarantee the stability of the Galerkin scheme, we need to include in the approximation space of the tensor gradient of the velocity the deviator of the vector Raviart–Thomas space of order zero. In this way, we prove that the discrete scheme has a unique solution and derive quasi-optimal error estimates and the corresponding rates of convergence. Next, in Section 4, we develop an implicit reliable and quasi-efficient *a posteriori* error estimate, and a fully explicit reliable one, and propose the adaptive algorithm associated to the latter to compute the finite element solutions. Finally, several numerical results are reported in Section 5.

In what follows, given any Hilbert space H, we denote by H^2 and $H^{2\times 2}$ the spaces of vectors and tensors of order two, respectively, with entries in H, provided with the product norms induced by the norm of H. In addition, for any $\boldsymbol{\tau} := (\tau_{ij}), \boldsymbol{\zeta} := (\zeta_{ij}) \in \mathbb{R}^{2\times 2}$, we denote $\mathrm{tr}(\boldsymbol{\tau}) := \tau_{11} + \tau_{22}$ and $\boldsymbol{\tau} : \boldsymbol{\zeta} := \sum_{i,j=1}^{2} \tau_{ij}\zeta_{ij}$. The deviator of tensor $\boldsymbol{\tau}$ is denoted by $\mathrm{dev}(\boldsymbol{\tau}) := \boldsymbol{\tau} - \frac{1}{2}\mathrm{tr}(\boldsymbol{\tau})\mathbf{I}$. We remark that $\mathrm{tr}(\mathrm{dev}(\boldsymbol{\tau})) = 0$.

2. THE CONTINUOUS VARIATIONAL FORMULATION

We first proceed as in Reference [8] and introduce two additional unknowns in Ω, namely, the tensor gradient of the velocity $\mathbf{t} := \nabla\mathbf{u}$ and the flux $\boldsymbol{\sigma} := \nu\nabla\mathbf{u} - p\mathbf{I}$, where $\mathbf{I}$ is the identity matrix in $\mathbf{R}^{2\times 2}$. It follows that the equilibrium equation becomes

$$\alpha\mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) = \mathbf{f} \quad \text{in } \Omega \tag{2}$$

where $\boldsymbol{\sigma} := \nu\mathbf{t} - p\,\mathbf{I}$ and **div** denotes the vector divergence operator. In addition, since $\mathrm{div}(\mathbf{u}) = \mathrm{tr}(\mathbf{t})$ in Ω, we can rewrite the incompressibility condition as

$$\mathrm{tr}(\mathbf{t}) = 0 \quad \text{in } \Omega \tag{3}$$

Now, multiplying the relation $\mathbf{t} = \nabla\mathbf{u}$ by a tensor $\boldsymbol{\tau}$, integrating by parts and using that $\mathbf{u} = \mathbf{g}$ on Γ, we get:

$$\int_\Omega \boldsymbol{\tau} : \mathbf{t} + \int_\Omega \mathbf{div}(\boldsymbol{\tau}) \cdot \mathbf{u} = \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{g} \rangle \quad \forall \boldsymbol{\tau} \in H(\mathbf{div}; \Omega) \tag{4}$$

Hereafter, $\langle\cdot,\cdot\rangle$ denotes the duality pairing of $[H^{-1/2}(\Gamma)]^2$ and $[H^{1/2}(\Gamma)]^2$ with $[L^2(\Gamma)]^2$ as pivot space, and $H(\mathbf{div}\,;\Omega)$ is the space of tensors $\boldsymbol{\tau}\in[L^2(\Omega)]^{2\times 2}$ satisfying $\mathbf{div}\,(\boldsymbol{\tau})\in[L^2(\Omega)]^2$. We also recall that $H(\mathbf{div}\,;\Omega)$, endowed with the inner product $\langle\boldsymbol{\zeta},\boldsymbol{\tau}\rangle_{H(\mathbf{div}\,;\Omega)}:=\langle\boldsymbol{\zeta},\boldsymbol{\tau}\rangle_{[L^2(\Omega)]^{2\times 2}}+\langle\mathbf{div}\,\boldsymbol{\zeta},\mathbf{div}\,\boldsymbol{\tau}\rangle_{[L^2(\Omega)]^2}$, is a Hilbert space, where $\langle\cdot,\cdot\rangle_{[L^2(\Omega)]^{2\times 2}}$ and $\langle\cdot,\cdot\rangle_{[L^2(\Omega)]^2}$ stand for the usual inner products of $[L^2(\Omega)]^{2\times 2}$ and $[L^2(\Omega)]^2$, respectively.

Then, testing the relation $\boldsymbol{\sigma}=\nu\mathbf{t}-p\,\mathbf{I}$, Equations (2) and (3), and reordering appropriately the resulting equations and (4), we obtain the following mixed variational formulation: Find $(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p)\in[L^2(\Omega)]^{2\times 2}\times[L^2(\Omega)]^2\times H(\mathbf{div}\,;\Omega)\times L^2(\Omega)$ such that

$$\begin{aligned}
\nu\int_\Omega \mathbf{t}:\mathbf{s} \qquad\qquad & -\int_\Omega \boldsymbol{\sigma}:\mathbf{s} \quad -\int_\Omega p\,\mathrm{tr}\,(\mathbf{s}) &&= 0\\
\alpha\int_\Omega \mathbf{u}\cdot\mathbf{v} \quad & -\int_\Omega \mathbf{div}\,(\boldsymbol{\sigma})\cdot\mathbf{v} &&= \int_\Omega \mathbf{f}\cdot\mathbf{v}\\
-\int_\Omega \boldsymbol{\tau}:\mathbf{t} \quad & -\int_\Omega \mathbf{div}\,(\boldsymbol{\tau})\cdot\mathbf{u} &&= -\langle\boldsymbol{\tau}\mathbf{n},\mathbf{g}\rangle\\
-\int_\Omega q\,\mathrm{tr}\,(\mathbf{t}) \qquad\qquad & &&= 0
\end{aligned} \tag{5}$$

for all $(\mathbf{s},\mathbf{v},\boldsymbol{\tau},q)\in[L^2(\Omega)]^{2\times 2}\times[L^2(\Omega)]^2\times H(\mathbf{div}\,;\Omega)\times L^2(\Omega)$.

Next, we observe that the variational formulation (5) is not uniquely solvable since given any solution $(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p)$ of this problem and any $c\in\mathbb{R}$, $(\mathbf{t},\mathbf{u},\boldsymbol{\sigma}+c\mathbf{I},p-c)$ also becomes a solution. Therefore, in order to guarantee uniqueness we proceed as in Reference [11] and require that $\int_\Omega \mathrm{tr}\,(\boldsymbol{\sigma})=0$, which leads to the introduction of a new unknown, a Lagrange multiplier $\xi\in\mathbb{R}$. Thus, from now on we consider the following mixed variational formulation of (1): Find $(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p,\xi)\in[L^2(\Omega)]^{2\times 2}\times[L^2(\Omega)]^2\times H(\mathbf{div}\,;\Omega)\times L^2(\Omega)\times\mathbb{R}$ such that

$$\begin{aligned}
\nu\int_\Omega \mathbf{t}:\mathbf{s} \qquad\qquad & -\int_\Omega \boldsymbol{\sigma}:\mathbf{s} \quad -\int_\Omega p\,\mathrm{tr}\,(\mathbf{s}) &&= 0\\
\alpha\int_\Omega \mathbf{u}\cdot\mathbf{v} \quad & -\int_\Omega \mathbf{div}\,(\boldsymbol{\sigma})\cdot\mathbf{v} &&= \int_\Omega \mathbf{f}\cdot\mathbf{v}\\
-\int_\Omega \boldsymbol{\tau}:\mathbf{t} \quad & -\int_\Omega \mathbf{div}\,(\boldsymbol{\tau})\cdot\mathbf{u} \quad +\xi\int_\Omega \mathrm{tr}\,(\boldsymbol{\tau}) &&= -\langle\boldsymbol{\tau}\mathbf{n},\mathbf{g}\rangle\\
-\int_\Omega q\,\mathrm{tr}\,(\mathbf{t}) \qquad\qquad & &&= 0\\
\eta\int_\Omega \mathrm{tr}\,(\boldsymbol{\sigma}) \qquad\qquad & &&= 0
\end{aligned} \tag{6}$$

for all $(\mathbf{s},\mathbf{v},\boldsymbol{\tau},q,\eta)\in[L^2(\Omega)]^{2\times 2}\times[L^2(\Omega)]^2\times H(\mathbf{div}\,;\Omega)\times L^2(\Omega)\times\mathbb{R}$. We remark here that taking $\boldsymbol{\tau}=\mathbf{I}$ in the third equation and applying (3) and the compatibility condition for the Dirichlet datum, we find *a priori* that $\xi=0$. However, we keep this artificial unknown to ensure the symmetry and the well-posedness of the whole formulation.

In order to prove the unique solvability of the variational formulation (6), we write it now as a system of operator equations with a two-fold saddle point structure. To this end, we

first define the spaces $X_1 := [L^2(\Omega)]^{2\times 2} \times [L^2(\Omega)]^2$, $M_1 := H(\mathbf{div}\,;\Omega)$, and $M := L^2(\Omega) \times \mathbb{R}$. Then, we introduce the operators and functionals $\mathbf{A}_1 : X_1 \to X_1'$, $\mathbf{B}_1 : X_1 \to M_1'$, $\mathbf{B}^p : X_1 \to L^2(\Omega)$, $\mathbf{B}^\xi : M_1 \to \mathbb{R}$, $\mathbf{F}_1 \in X_1'$, and $\mathbf{F}_2 \in M_1'$, as suggested by the structure of (6), so that this problem can be stated as: Find $((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi)) \in X_1 \times M_1 \times M$ such that

$$\begin{aligned} [\mathbf{A}_1(\mathbf{t},\mathbf{u}),(\mathbf{s},\mathbf{v})] + [\mathbf{B}_1(\mathbf{s},\mathbf{v}),\boldsymbol{\sigma}] + [\mathbf{B}^p(\mathbf{s},\mathbf{v}),p] &= [\mathbf{F}_1,(\mathbf{s},\mathbf{v})] \\ [\mathbf{B}_1(\mathbf{t},\mathbf{u}),\boldsymbol{\tau}] + [\mathbf{B}^\xi(\boldsymbol{\tau}),\xi] &= [\mathbf{F}_2,\boldsymbol{\tau}] \\ [\mathbf{B}^p(\mathbf{t},\mathbf{u}),q] + [\mathbf{B}^\xi(\boldsymbol{\sigma}),\eta] &= 0 \end{aligned} \tag{7}$$

for all $((\mathbf{s},\mathbf{v}),\boldsymbol{\tau},(q,\eta)) \in X_1 \times M_1 \times M$, where $[\cdot,\cdot]$ denotes the duality pairing induced by the operators and functionals used in each case.

We now let $X := X_1 \times M_1$, identify X' with $X_1' \times M_1'$, and define $\mathbf{A} : X \to X'$ as the matrix operator

$$\mathbf{A} := \begin{bmatrix} \mathbf{A}_1 & \mathbf{B}_1' \\ \mathbf{B}_1 & \mathbf{O} \end{bmatrix} \tag{8}$$

where $\mathbf{B}_1' : M_1 \to X_1'$ is the adjoint of $\mathbf{B}_1$, and $\mathbf{O}$ denotes, from now on, a generic null operator/functional. Hence, (7) can be set equivalently as: Find $((\mathbf{t},\mathbf{u},\boldsymbol{\sigma}),(p,\xi)) \in X \times M$ such that

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}' \\ \mathbf{B} & \mathbf{O} \end{bmatrix} \begin{bmatrix} (\mathbf{t},\mathbf{u},\boldsymbol{\sigma}) \\ (p,\xi) \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{O} \end{bmatrix} \tag{9}$$

where $\mathbf{B} : X \to M$ is defined by $[\mathbf{B}(\mathbf{r},\mathbf{w},\boldsymbol{\zeta}),(q,\eta)] := [\mathbf{B}^p(\mathbf{r},\mathbf{w}),q] + [\mathbf{B}^\xi(\boldsymbol{\zeta}),\eta]$, $\mathbf{B}' : M \to X'$ is the adjoint of $\mathbf{B}$, and $\mathbf{F} \in X'$ is defined by $[\mathbf{F},(\mathbf{s},\mathbf{v},\boldsymbol{\tau})] := [\mathbf{F}_1,(\mathbf{s},\mathbf{v})] + [\mathbf{F}_2,\boldsymbol{\tau}]$ for all $(\mathbf{s},\mathbf{v},\boldsymbol{\tau})$, $(\mathbf{r},\mathbf{w},\boldsymbol{\zeta}) \in X$ and for all $(q,\eta) \in M$. In this way, the two-fold saddle point structure of (7) becomes clear from (8) and (9) since $\mathbf{A}$ itself has the saddle point structure.

We now apply the abstract theory from Reference [10] (see also the related results given in References [12, 13]) to establish the solvability and continuous dependence of (7).

Theorem 1

Problem (7) has a unique solution $((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi)) \in X_1 \times M_1 \times M$. Moreover, there exists a positive constant $C(\alpha,\nu) = O(\alpha^3/\nu)$, independent of the solution, such that

$$\|((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi))\|_{X_1\times M_1\times M} \leqslant C(\alpha,\nu)\{\|\mathbf{F}_1\| + \|\mathbf{F}_2\|\} \tag{10}$$

Proof

We observe first that the operators $\mathbf{A}_1$, $\mathbf{B}_1$ and $\mathbf{B}$ are all linear and bounded. In particular, it is easy to see that $\|\mathbf{A}_1\| = O(\alpha)$ and that both $\|\mathbf{B}_1\|$ and $\|\mathbf{B}\|$ are of $O(1)$. In addition, since $\alpha \geqslant \nu$, we deduce that $\mathbf{A}_1$ is X_1-elliptic with ellipticity constant ν. Thus, according to the linear version of Theorem 2.4 in Reference [10] (see also Theorem 2 in Reference [12]), it only remains to show that $\mathbf{B}$ and $\mathbf{B}_1$ satisfy the corresponding inf–sup conditions on $X \times M$ and on the kernel of $\mathbf{B}$, respectively.

Indeed, given $(q,\eta)\in M$ we get lower bounds for

$$\sup_{(\mathbf{s},\mathbf{v},\boldsymbol{\tau})\in X\setminus\{0\}}\frac{[\mathbf{B}(\mathbf{s},\mathbf{v},\boldsymbol{\tau}),(q,\eta)]}{\|(\mathbf{s},\mathbf{v},\boldsymbol{\tau})\|_X}$$

by taking $(\mathbf{s},\mathbf{v},\boldsymbol{\tau})=(0,0,\eta\mathbf{I})$ and $(\mathbf{s},\mathbf{v},\boldsymbol{\tau})=(-q\,\mathbf{I},0,0)$, which yields the inf–sup condition for $\mathbf{B}$.

Next, we realize that the null space of the operator $\mathbf{B}$ is $\tilde{X}=\tilde{X}_1\times\tilde{M}_1$, where $\tilde{X}_1:=\{(\mathbf{s},\mathbf{v})\in X_1:\mathrm{tr}\,(\mathbf{s})=0 \text{ in } \Omega\}$ and $\tilde{M}_1:=\{\boldsymbol{\tau}\in M_1:\int_\Omega \mathrm{tr}\,(\boldsymbol{\tau})=0\}$. Thus, given $\boldsymbol{\tau}\in\tilde{M}_1$ we get now lower bounds for

$$\sup_{(\mathbf{s},\mathbf{v})\in\tilde{X}_1\setminus\{0\}}\frac{[\mathbf{B}_1(\mathbf{s},\mathbf{v}),\boldsymbol{\tau}]}{\|(\mathbf{s},\mathbf{v})\|_{X_1}}$$

by taking $(\mathbf{s},\mathbf{v})=(\mathbf{0},-\mathbf{div}\,(\boldsymbol{\tau}))$ and $(\mathbf{s},\mathbf{v})=(-\mathrm{dev}\,(\boldsymbol{\tau}),\mathbf{0})$, which, using Lemma 3.1 in Reference [14], yields the inf–sup condition for $\mathbf{B}_1$. This lower bound is also obtained when $\mathbf{div}\,(\boldsymbol{\tau})=0$ or $\mathrm{dev}\,(\boldsymbol{\tau})=\mathbf{0}$. We omit details.

Finally, we remark that the order of the continuous dependence constant $C(\alpha,\nu)$ follows from the analysis provided in Section 2 of Reference [10] and from a particular case of Proposition 2.3 in Reference [15]. □

3. THE MIXED FINITE ELEMENT SCHEME

In what follows we assume, for simplicity, that Γ is a polygonal curve. Then, we let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of $\bar{\Omega}$ by triangles T of diameter h_T such that $h:=\max\{h_T:T\in\mathcal{T}_h\}$ and $\bar{\Omega}=\cup\{T:T\in\mathcal{T}_h\}$. Also, we let $X_{1,h}^{\mathbf{t}}$, $X_{1,h}^{\mathbf{u}}$, $M_{1,h}$, and M_h^p be finite element subspaces for the unknowns $\mathbf{t}$, $\mathbf{u}$, $\boldsymbol{\sigma}$, and p, respectively, and define $X_{1,h}:=X_{1,h}^{\mathbf{t}}\times X_{1,h}^{\mathbf{u}}$ and $M_h:=M_h^p\times\mathbb{R}$. Then, the Galerkin scheme associated with problem (7) reads: Find $((\mathbf{t}_h,\mathbf{u}_h),\boldsymbol{\sigma}_h,(p_h,\xi_h))\in X_{1,h}\times M_{1,h}\times M_h$ such that

$$\begin{aligned}
&[\mathbf{A}_1(\mathbf{t}_h,\mathbf{u}_h),(\mathbf{s}_h,\mathbf{v}_h)]+[\mathbf{B}_1(\mathbf{s}_h,\mathbf{v}_h),\boldsymbol{\sigma}_h]+[\mathbf{B}^p(\mathbf{s}_h,\mathbf{v}_h),p_h]&&=[\mathbf{F}_1,(\mathbf{s}_h,\mathbf{v}_h)]\\
&[\mathbf{B}_1(\mathbf{t}_h,\mathbf{u}_h),\boldsymbol{\tau}_h]+[\mathbf{B}^\xi(\boldsymbol{\tau}_h),\xi_h]&&=[\mathbf{F}_2,\boldsymbol{\tau}_h]\\
&[\mathbf{B}^p(\mathbf{t}_h,\mathbf{u}_h),q_h]+[\mathbf{B}^\xi(\boldsymbol{\sigma}_h),\eta_h]&&=0
\end{aligned}\tag{11}$$

for all $((\mathbf{s}_h,\mathbf{v}_h),\boldsymbol{\tau}_h,(q_h,\eta_h))\in X_{1,h}\times M_{1,h}\times M_h$.

Our purpose is to define these finite element subspaces so that (11) becomes well posed. As suggested by the linear version of Theorem 3.2 in Reference [10], it suffices to prove the ellipticity of $\mathbf{A}_1$ on $X_{1,h}$ and the discrete inf–sup conditions for $\mathbf{B}$ and $\mathbf{B}_1$. In order to establish these properties we show below that the same arguments of the continuous case can be applied again, which will yield to determine the appropriate finite element spaces for each unknown. To begin with, we realize that there is nothing else to prove for $\mathbf{A}_1$ since the ellipticity of this bilinear form is certainly valid on any subspace of X_1. Next, in order to extend the proof of the continuous inf–sup condition for $\mathbf{B}$ to the discrete case, we require that $(0,0,\eta\,\mathbf{I})$ and $(q_h\,\mathbf{I},0,0)$ belong to $X_{1,h}^{\mathbf{t}}\times X_{1,h}^{\mathbf{u}}\times M_{1,h}$ for any $(q_h,\eta)\in M_h^p\times\mathbb{R}$, that is

$$\eta\mathbf{I}\in M_{1,h}\quad\forall\eta\in\mathbb{R}\quad\text{and}\quad q_h\mathbf{I}\in X_{1,h}^{\mathbf{t}}\quad\forall q_h\in M_h^p\tag{12}$$

On the other hand, it is easy to see that the discrete kernel of the bilinear form $\mathbf{B}$ is given by $\tilde{X}^{\mathbf{t}}_{1,h} \times X^{\mathbf{u}}_{1,h} \times \tilde{M}_{1,h}$, where $\tilde{X}^{\mathbf{t}}_{1,h} := \{s_h \in X^{\mathbf{t}}_{1,h} : \int_\Omega q_h \operatorname{tr}(s_h) = 0 \ \forall q_h \in M^p_h\}$ and $\tilde{M}_{1,h} := \{\boldsymbol{\tau}_h \in M_{1,h} : \int_\Omega \operatorname{tr}(\boldsymbol{\tau}_h) = 0\}$.

We note here that $\tilde{M}_{1,h}$ is clearly a subspace of $\tilde{M}_1$ and hence the equivalence of $\|\boldsymbol{\tau}_h\|_{[L^2(\Omega)]^{2\times 2}}$ and $\|\operatorname{dev}(\boldsymbol{\tau}_h)\|_{[L^2(\Omega)]^{2\times 2}}$ also holds for each $\boldsymbol{\tau}_h \in \tilde{M}_{1,h}$. Thus, in order to extend now the proof of the continuous inf–sup condition for $\mathbf{B}_1$ to the discrete case, we need that

$$\mathbf{div}(\boldsymbol{\tau}_h) \in X^{\mathbf{u}}_{1,h} \quad \text{and} \quad \operatorname{dev}(\boldsymbol{\tau}_h) \in \tilde{X}^{\mathbf{t}}_{1,h} \quad \forall \boldsymbol{\tau}_h \in \tilde{M}_{1,h} \tag{13}$$

Since (12) and (13) do not impose any explicit condition on the elements of M^p_h, we choose this subspace of $L^2(\Omega)$ as the simplest possible one, that is, as the piecewise constant functions on the triangulation $\mathscr{T}_h$. Similarly, since the first restriction of (12) is satisfied if the piecewise constant tensors are included in $M_{1,h}$, we just choose this subspace of $H(\mathbf{div}; \Omega)$ as the Raviart–Thomas space of order zero (see References [11, 16]). Because of this choice of $M_{1,h}$, and in order to satisfy the first requirement of (13), we realize that it suffices to take $X^{\mathbf{u}}_{1,h}$ as the space of piecewise constant vectors on $\mathscr{T}_h$.

Finally, taking into account the choices already made for M^p_h and $M_{1,h}$, and observing that the trace of any deviator is zero, we find that the remaining conditions in (12) and (13) are accomplished if $X^{\mathbf{t}}_{1,h}$ is chosen so that its restriction on each triangle $T \in \mathscr{T}_h$ becomes the local space $\mathscr{A}_0(T) := \langle\{\mathbf{I}\}\rangle \oplus \operatorname{dev}([\mathscr{RT}_0(T)\mathscr{RT}_0(T)]^{\mathrm{t}})$, where $\langle\,\rangle$ is used hereafter to denote spanning, and

$$\mathscr{RT}_0(T) := \left\langle \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\} \right\rangle$$

is the local Raviart–Thomas space of order zero. Moreover, it is not difficult to see that

$$\mathscr{A}_0(T) := [\mathbf{P}_0(T)]^{2\times 2} \oplus \left\langle \left\{ \begin{pmatrix} x_1 & 2x_2 \\ 0 & -x_1 \end{pmatrix}, \begin{pmatrix} -x_2 & 0 \\ 2x_1 & x_2 \end{pmatrix} \right\} \right\rangle \tag{14}$$

where $\mathbf{P}_0(T)$ denotes the space of constant functions defined on T.

According to the above analysis, our finite element subspaces are given by

$$X^{\mathbf{t}}_{1,h} := \{\mathbf{s} \in [L^2(\Omega)]^{2\times 2} : \mathbf{s}|_T \in \mathscr{A}_0(T) \ \forall T \in \mathscr{T}_h\}$$

$$X^{\mathbf{u}}_{1,h} := \{\mathbf{v} \in [L^2(\Omega)]^2 : \mathbf{v}|_T \in [\mathbf{P}_0(T)]^2 \ \forall T \in \mathscr{T}_h\}$$

$$M_{1,h} := \{\boldsymbol{\tau} \in H(\mathbf{div}; \Omega) : (\tau_{i1} \ \ \tau_{i2})^{\mathrm{t}}|_T \in \mathscr{RT}_0(T) \ \forall i \in \{1,2\} \quad \forall T \in \mathscr{T}_h\}$$

and

$$M^p_h := \{q \in L^2(\Omega) : \ q|_T \in \mathbf{P}_0(T) \ \forall T \in \mathscr{T}_h\}$$

The well posedness of the discrete problem (11) and the corresponding quasi-optimal error estimate can then be established.

Theorem 2
Problem (11) has a unique solution $((\mathbf{t}_h,\mathbf{u}_h),\boldsymbol{\sigma}_h,(p_h,\xi_h))\in X_{1,h}\times M_{1,h}\times M_h$. Moreover, there exists a positive constant $\hat{C}(\alpha,\nu)=O(\alpha^3/\nu)$, independent of h, such that

$$\|((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi))-((\mathbf{t}_h,\mathbf{u}_h),\boldsymbol{\sigma}_h,(p_h,\xi_h))\|$$
$$\leqslant \hat{C}(\alpha,\nu)\inf_{((\mathbf{s}_h,\mathbf{v}_h),\boldsymbol{\tau}_h,q_h)\in X_{1,h}\times M_{1,h}\times M_h^p}\|((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},p)-((\mathbf{s}_h,\mathbf{v}_h),\boldsymbol{\tau}_h,q_h)\| \tag{15}$$

where $((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi))$ is the unique solution of the continuous problem (7).

Proof
We first remark, since we are dealing with a linear problem, that the Céa estimate (15) is equivalent to stability of the Galerkin scheme (11). Now, as shown by our previous analysis, the present finite element subspaces guarantee the ellipticity of $\mathbf{A}_1$ on $X_{1,h}$, with the same constant ν, as well as the discrete inf–sup conditions for $\mathbf{B}$ and $\mathbf{B}_1$, with constants depending only on Ω. Hence, using again that $\|\mathbf{A}_1\|=O(\alpha)$ and that both $\|\mathbf{B}_1\|$ and $\|\mathbf{B}\|$ are of $O(1)$, we can apply the linear version of Theorem 3.2 in Reference [10] to deduce the unique solvability of (11) and the corresponding stability with a constant behaving like $O(\alpha^3/\nu)$. □

Next, we have the following result on the rate of convergence of the solution of (11).

Theorem 3
Let $((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi))$ and $((\mathbf{t}_h,\mathbf{u}_h),\boldsymbol{\sigma}_h,(p_h,\xi_h))$ be the unique solutions of the continuous and discrete formulations, respectively. Assume that $\mathbf{t}\in[H^1(\Omega)]^{2\times 2}$, $\mathbf{u}\in[H^1(\Omega)]^2$, $\boldsymbol{\sigma}\in[H^1(\Omega)]^{2\times 2}$, $\mathbf{div}(\boldsymbol{\sigma})\in[H^1(\Omega)]^2$, and $p\in H^1(\Omega)$. Then there exists a positive constant $\bar{C}(\alpha,\nu)=O(\alpha^3/\nu)$, independent of h, such that

$$\|((\mathbf{t},\mathbf{u}),\boldsymbol{\sigma},(p,\xi))-((\mathbf{t}_h,\mathbf{u}_h),\boldsymbol{\sigma}_h,(p_h,\xi_h))\|_{X_1\times M_1\times M}$$
$$\leqslant \bar{C}(\alpha,\nu)h(\|\mathbf{t}\|_{[H^1(\Omega)]^{2\times 2}}+\|\boldsymbol{\sigma}\|_{[H^1(\Omega)]^{2\times 2}}+\|\mathbf{u}\|_{[H^1(\Omega)]^2}+\|\mathbf{div}(\boldsymbol{\sigma})\|_{[H^1(\Omega)]^2}+\|p\|_{H^1(\Omega)})$$

Proof
It is a consequence of the Céa estimate (15) and the well-known approximation properties of the subspaces $X_{1,h}^{\mathbf{t}}$, $X_{1,h}^{\mathbf{u}}$, $M_{1,h}$ and M_h^p, which follow from classical error estimates for projection and equilibrium interpolation operators (see, e.g. Reference [16]).

4. *A POSTERIORI* ERROR ANALYSIS

We now develop an *a posteriori* error analysis (based on suitable local problems) and derive reliable estimates for the mixed finite element solution introduced in the previous section. Similarly as in Reference [9], our approach follows the technique from Reference [17], which is a modification of the original Bank–Weiser method proposed in Reference [18].

Let us first introduce some notations. We denote by E_h the set of all the edges of the triangulation $\mathcal{T}_h$, define $E_h(\Gamma):=\{e\in E_h: e\subseteq\Gamma\}$, and given $T\in\mathcal{T}_h$, we let $E(T):=\{e\in E_h: e\subseteq\partial T\}$.

In addition, the inner product of $H(\mathbf{div}\,;T)$ is denoted by $\langle\cdot,\cdot\rangle_{H(\mathbf{div}\,;T)}$, and $\mathbf{n}_T$ stands for the unit outward normal to ∂T.

On the other hand, given a polygonal domain $\mathcal{S}\subseteq\mathbb{R}^2$ and $m\in(1,\infty)$, the Sobolev space $W^{1,m}(\mathcal{S})$ is the Banach space of functions $v\in L^m(\mathcal{S})$ such that the first-order distributional derivatives of v are functions of $L^m(\mathcal{S})$. A Sobolev imbedding theorem establishes that $W^{1,m}(\mathcal{S})\subseteq C(\bar{\mathcal{S}})$ if $m>2$ (see Reference [19] for details). Also, it is well known that the trace theorem ensures the existence of a linear continuous map $\gamma: W^{1,m}(\mathcal{S})\to L^m(\partial\mathcal{S})$ such that $\gamma v=v|_{\partial\mathcal{S}}$ for each $v\in W^{1,m}(\mathcal{S})\cap C(\bar{\mathcal{S}})$. The range of γ, which is a strict subspace of $L^m(\partial\mathcal{S})$, is denoted by $W^{1-1/m,m}(\partial\mathcal{S})$. In particular, when $\mathcal{S}:=T\in\mathcal{T}_h$ and $m=2$, we use the standard notation and write $H^{1/2}(\partial T)$ instead of $W^{1/2,2}(\partial T)$.

Now, given an edge $e\in E(T)$, $H_0^1(e)$ stands for the closure in $H^1(e)$ of the space of infinitely differentiable functions with compact support in e. We recall here that the interpolation space with index $\frac{1}{2}$ between $H_0^1(e)$ and $L^2(e)$ is $H_{00}^{1/2}(e)$ (cf. Reference [19]). The space $H_{00}^{1/2}(e)$ may be alternatively defined as the subspace of functions in $H^{1/2}(e)$ whose extensions by zero to the rest of ∂T belong to $H^{1/2}(\partial T)$. We will also need the dual space of $H^{1/2}(\partial T)$, which is denoted by $H^{-1/2}(\partial T)$. To this respect, we remark that the restriction of an element in $H^{-1/2}(\partial T)$ over e does not belong in general to $H^{-1/2}(e)$, but to $H_{00}^{-1/2}(e)$, dual space of $H_{00}^{1/2}(e)$ pivotal to $L^2(e)$, and which is, therefore, larger than $H^{-1/2}(e)$. According to this, in what follows we set $\langle\cdot,\cdot\rangle_e$ for the duality pairing between $[H_{00}^{-1/2}(e)]^2$ and $[H_{00}^{1/2}(e)]^2$ with $[L^2(e)]^2$ as pivot space, and we let $\langle\cdot,\cdot\rangle_{\partial T}$ be the duality pairing between $[H^{-1/2}(\partial T)]^2$ and $[H^{1/2}(\partial T)]^2$ with $[L^2(\partial T)]^2$ as pivot space.

We now define the Riesz projection of the error with respect to the inner product of $X:=X_1\times M_1$ as the unique element $(\bar{\mathbf{t}},\bar{\mathbf{u}},\bar{\boldsymbol{\sigma}})\in X$ such that

$$\langle(\bar{\mathbf{t}},\bar{\mathbf{u}},\bar{\boldsymbol{\sigma}}),(\mathbf{s},\mathbf{v},\boldsymbol{\tau})\rangle_X=[\mathbf{A}(\mathbf{t}-\mathbf{t}_h,\mathbf{u}-\mathbf{u}_h,\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),(\mathbf{s},\mathbf{v},\boldsymbol{\tau})]+[\mathbf{B}(\mathbf{s},\mathbf{v},\boldsymbol{\tau}),(p-p_h,\xi-\xi_h)]$$

for all $(\mathbf{s},\mathbf{v},\boldsymbol{\tau})\in X$, where $\mathbf{A}$ and $\mathbf{B}$ are the bilinear forms defined in Section 2, and

$$\langle(\bar{\mathbf{t}},\bar{\mathbf{u}},\bar{\boldsymbol{\sigma}}),(\mathbf{s},\mathbf{v},\boldsymbol{\tau})\rangle_X:=\langle\bar{\mathbf{t}},\mathbf{s}\rangle_{[L^2(\Omega)]^{2\times 2}}+\langle\bar{\mathbf{u}},\mathbf{v}\rangle_{[L^2(\Omega)]^2}+\langle\bar{\boldsymbol{\sigma}},\boldsymbol{\tau}\rangle_{H(\mathbf{div}\,;\Omega)}$$

In what follows, we assume that there exists $m>2$ such that the Dirichlet datum $\mathbf{g}\in[H^{1/2}(\Gamma)\cap W^{1-1/m,m}(\Gamma)]^2$ and let φ_h be a given auxiliary function in $[H^1(\Omega)\cap W^{1,m}(\Omega)]^2$ such that $\varphi_h(\bar{\mathbf{x}})=\mathbf{g}(\bar{\mathbf{x}})$ for each vertex $\bar{\mathbf{x}}$ of $\mathcal{T}_h$ lying on Γ. In addition, for each $T\in\mathcal{T}_h$ we let $\hat{\boldsymbol{\sigma}}_T\in H(\mathbf{div}\,;T)$ be the unique solution of the local problem

$$\langle\hat{\boldsymbol{\sigma}}_T,\boldsymbol{\tau}\rangle_{H(\mathbf{div}\,;T)}=\mathcal{F}_{h,T}(\boldsymbol{\tau})\quad\forall\boldsymbol{\tau}\in H(\mathbf{div}\,;T)\tag{16}$$

where $\mathcal{F}_{h,T}\in H(\mathbf{div}\,;T)'$ is defined by

$$\mathcal{F}_{h,T}(\boldsymbol{\tau}):=\int_T\boldsymbol{\tau}:\mathbf{t}_h+\int_T\mathbf{u}_h\cdot\mathbf{div}\,(\boldsymbol{\tau})-\xi_h\int_T\mathrm{tr}\,(\boldsymbol{\tau})-\langle\boldsymbol{\tau}\mathbf{n}_T,\varphi_h\rangle_{\partial T}+\sum_{e\in E(T)\cap E_h(\Gamma)}\langle\boldsymbol{\tau}\mathbf{n}_T,\varphi_h-\mathbf{g}\rangle_e$$

The following lemma provides an upper bound for $\|(\bar{\mathbf{t}},\bar{\mathbf{u}},\bar{\boldsymbol{\sigma}})\|_X$.

Lemma 4
There holds

$$\|(\bar{\mathbf{t}},\bar{\mathbf{u}},\bar{\boldsymbol{\sigma}})\|_X^2\leqslant\sum_{T\in\mathcal{T}_h}\hat{\theta}_T^2\tag{17}$$

where

$$\hat{\theta}_T^2 := \|\hat{\boldsymbol{\sigma}}_T\|^2_{H(\mathbf{div};T)} + \|\boldsymbol{\sigma}_h - \nu\mathbf{t}_h + p_h\mathbf{I}\|^2_{[L^2(T)]^{2\times 2}} + \|\mathbf{f} + \mathbf{div}(\boldsymbol{\sigma}_h) - \alpha\mathbf{u}_h\|^2_{[L^2(T)]^2}$$

Proof
The proof utilizes the same arguments of Lemma 3.1 in Reference [9]. We only remark here that the regularity hypotheses on $\mathbf{g}$ and φ_h guarantee, by virtue of a Sobolev imbedding theorem, that $\mathbf{g}$ and φ_h are both continuous, and hence the associated interpolation condition for the vertices of $\mathscr{T}_h$ lying on Γ makes sense. We omit further details. □

A priori estimates for the solutions of the local problems (16) are given in the following lemma.

Lemma 5
There exists $C>0$, independent of h, α, ν, and T, such that

$$\|\hat{\boldsymbol{\sigma}}_T\|^2_{H(\mathbf{div};T)} \leqslant C\left(\|\mathbf{t}_h - \nabla\varphi_h\|^2_{[L^2(T)]^{2\times 2}} + \|\mathbf{u}_h - \varphi_h\|^2_{[L^2(T)]^2} + h_T^2\,|\xi_h|^2 + \sum_{e\in E(T)\cap E_h(\Gamma)} \|\varphi_h - \mathbf{g}\|^2_{[H^{1/2}_{00}(e)]^2}\right)$$

Furthermore, for any $\mathbf{z}\in[H^1(\Omega)\cap W^{1,m}(\Omega)]^2$, with $m>2$, such that $\mathbf{z}=\mathbf{g}$ on Γ, we get

$$\|\hat{\boldsymbol{\sigma}}_T\|^2_{H(\mathbf{div};T)} \leqslant C\left(\|\mathbf{t}_h - \nabla\mathbf{z}\|^2_{[L^2(T)]^{2\times 2}} + \|\mathbf{u}_h - \mathbf{z}\|^2_{[L^2(T)]^2} + h_T^2|\xi_h|^2 + \|\mathbf{J}_{h,T}(\mathbf{z})\|^2_{[H^{1/2}(\partial T)]^2}\right)$$

where

$$\mathbf{J}_{h,T}(\mathbf{z}) := \begin{cases} \mathbf{0} & \text{on } \partial T\cap\Gamma \\ \mathbf{z}-\varphi_h & \text{otherwise} \end{cases}$$

Proof
It is clear from (16) that $\|\hat{\boldsymbol{\sigma}}_T\|_{H(\mathbf{div};T)} = \|\mathscr{F}_{h,T}\|_{H(\mathbf{div};T)'}$. Therefore, the first estimate for $\|\hat{\boldsymbol{\sigma}}_T\|_{H(\mathbf{div};T)}$ follows after applying Gauss' formula to the expression $\langle\boldsymbol{\tau}\mathbf{n}_T, \varphi_h\rangle_{\partial T}$ appearing in the definition of the functional $\mathscr{F}_{h,T}$, and reordering the resulting terms so that the Cauchy–Schwarz inequality can be applied conveniently. The proof of the second estimate is similar. One just needs to observe, after simple computations, that

$$-\langle\boldsymbol{\tau}\mathbf{n}_T, \varphi_h\rangle_{\partial T} + \sum_{e\in E(T)\cap E_h(\Gamma)} \langle\boldsymbol{\tau}\mathbf{n}_T, \varphi_h - \mathbf{g}\rangle_e = -\langle\boldsymbol{\tau}\mathbf{n}_T, \mathbf{z}\rangle_{\partial T} + \langle\boldsymbol{\tau}\mathbf{n}_T, \mathbf{J}_{h,T}(\mathbf{z})\rangle$$

The rest proceeds as in the first case, applying now Gauss' formula to $\langle\boldsymbol{\tau}\mathbf{n}_T, \mathbf{z}\rangle_{\partial T}$. □

The above lemmata allow us to establish next the main *a posteriori* error estimates.

Theorem 6
There exists a positive constant $C_0(\alpha,\nu) = O(\alpha^3/\nu)$, independent of h, such that

$$\|(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p,\xi) - (\mathbf{t}_h,\mathbf{u}_h,\boldsymbol{\sigma}_h,p_h,\xi_h)\|_{X_1\times M_1\times M} \leqslant C_0(\alpha,\nu)\left\{\sum_{T\in\mathscr{T}_h}\tilde{\theta}_T^2\right\}^{1/2} \tag{18}$$

and

$$\|(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p,\xi)-(\mathbf{t}_h,\mathbf{u}_h,\boldsymbol{\sigma}_h,p_h,\xi_h)\|_{X_1\times M_1\times M}\leqslant C_0(\alpha,\nu)\left\{\sum_{T\in\mathscr{T}_h}\theta_T^2\right\}^{1/2} \tag{19}$$

where for each triangle $T\in\mathscr{T}_h$, we define

$$\begin{aligned}\tilde{\theta}_T^2 :=& \|\hat{\boldsymbol{\sigma}}_T\|^2_{H(\mathbf{div};T)}+\|\boldsymbol{\sigma}_h-\nu\mathbf{t}_h+p_h\mathbf{I}\|^2_{[L^2(T)]^{2\times 2}}\\ &+\|\mathbf{f}+\mathbf{div}\,(\boldsymbol{\sigma}_h)-\alpha\mathbf{u}_h\|^2_{[L^2(T)]^2}+\|\mathrm{tr}\,(\mathbf{t}_h)\|^2_{L^2(T)}\end{aligned} \tag{20}$$

and

$$\begin{aligned}\theta_T^2 :=& \|\mathbf{t}_h-\nabla\varphi_h\|^2_{[L^2(T)]^{2\times 2}}+\|\mathbf{u}_h-\varphi_h\|^2_{[L^2(T)]^2}+h_T^2|\xi_h|^2\\ &+\sum_{e\in E(T)\cap E_h(\Gamma)}\|\varphi_h-\mathbf{g}\|^2_{[H^{1/2}_{00}(e)]^2}+\|\boldsymbol{\sigma}_h-\nu\mathbf{t}_h+p_h\mathbf{I}\|^2_{[L^2(T)]^{2\times 2}}\\ &+\|\mathbf{f}+\mathbf{div}\,(\boldsymbol{\sigma}_h)-\alpha\mathbf{u}_h\|^2_{[L^2(T)]^2}+\|\mathrm{tr}\,(\mathbf{t}_h)\|^2_{L^2(T)}\end{aligned} \tag{21}$$

Proof
It is similar to the proof of Theorem 2.1 in Reference [9]. Indeed, the continuous dependence result given in Theorem 1 (cf. (10)) is equivalent to the global inf–sup condition for the linear operator obtained by adding the three equations of the left-hand side of (7). Hence, by applying this condition to the error $(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p,\xi)-(\mathbf{t}_h,\mathbf{u}_h,\boldsymbol{\sigma}_h,p_h,\xi_h)$, and using the definition of the Riesz projection $(\bar{\mathbf{t}},\bar{\mathbf{u}},\bar{\boldsymbol{\sigma}})\in X$, the definition of the operator $\mathbf{B}$, the third equation of the continuous problem (7), and Lemmas 4 and 5, we obtain estimates (18) and (19). The details are omitted. □

According to the previous theorem, we now introduce the *reliable a posteriori* error estimates

$$\tilde{\boldsymbol{\theta}}:=\left\{\sum_{T\in\mathscr{T}_h}\tilde{\theta}_T^2\right\}^{1/2}\quad\text{and}\quad\boldsymbol{\theta}:=\left\{\sum_{T\in\mathscr{T}_h}\theta_T^2\right\}^{1/2} \tag{22}$$

In addition, we establish next that $\tilde{\boldsymbol{\theta}}$ is *quasi-efficient*, which means that it is efficient up to a term depending on the traces $(\mathbf{u}-\varphi_h)$ on the edges of $\mathscr{T}_h$. We also remark that, besides the regularity and interpolation conditions, Lemmas 4, 5 and Theorem 6 do not require any further assumptions on φ_h. However, as we show below, the above-mentioned *quasi-efficiency* will restrict the possible choices of this auxiliary function.

Lemma 7
Assume that $\mathbf{u}\in[W^{1,m}(\Omega)]^2$, with $m>2$. Then there exists a positive constant $C_1(\alpha)=O(\alpha^2)$, independent of h, such that for all $T\in\mathscr{T}_h$

$$\begin{aligned}\tilde{\theta}_T^2 \leqslant& C_1(\alpha)\{\|\mathbf{t}-\mathbf{t}_h\|^2_{[L^2(T)]^{2\times 2}}+\|\mathbf{u}-\mathbf{u}_h\|^2_{[L^2(T)]^2}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{H(\mathbf{div};T)}\\ &+\|p-p_h\|^2_{L^2(T)}+h_T^2|\xi-\xi_h|^2+\|\mathbf{J}_{h,T}(\mathbf{u})\|^2_{[H^{1/2}(\partial T)]^2}\}\end{aligned} \tag{23}$$

and hence

$$\tilde{\boldsymbol{\theta}}^2 \leqslant C_1(\alpha)\left\{\|(\mathbf{t},\mathbf{u},\boldsymbol{\sigma},p,\xi)-(\mathbf{t}_h,\mathbf{u}_h,\boldsymbol{\sigma}_h,p_h,\xi_h)\|^2_{X_1\times M_1\times M}+\sum_{T\in\mathscr{T}_h}\|\mathbf{J}_{h,T}(\mathbf{u})\|^2_{[H^{1/2}(\partial T)]^2}\right\}$$

where

$$\mathbf{J}_{h,T}(\mathbf{u}):=\begin{cases}\mathbf{0} & \text{on } \partial T\cap\Gamma\\ \mathbf{u}-\varphi_h & \text{otherwise}\end{cases}$$

Proof
It is similar to the proof of Lemma 4.1 in Reference [9]. We omit the details here. □

At this point we remark that, although the reliable *a posteriori* error estimate $\tilde{\boldsymbol{\theta}}$ is also *quasi-efficient*, its eventual applicability is limited by the fact that it requires the knowledge of the exact solutions $\hat{\boldsymbol{\sigma}}_T$ of the local problems (16), which live all in the infinite dimensional space $H(\mathbf{div};T)$. This difficulty can be partially overcome by using h, p or $h-p$ versions of the finite element method to solve (16) approximately, which naturally yields approximations of the local indicators $\tilde{\theta}_T$, and hence of $\tilde{\boldsymbol{\theta}}$. We will go back to this point in Section 5.

On the other hand, the advantage of $\boldsymbol{\theta}$, which is not necessarily *quasi-efficient*, is that it does not need the exact or any approximate solution of (16), but a suitable φ_h. In order to choose this auxiliary function, we adopt the criterion of enforcing the *quasi-efficiency* of $\tilde{\boldsymbol{\theta}}$ so that it becomes closer to full efficiency. As established in Lemma 7, this criterion requires that the traces $\varphi_h|_{\partial T}$ be as close as possible to the exact traces $\mathbf{u}|_{\partial T}$, for each $T\in\mathscr{T}_h$. According to this, and taking into account the finite element subspaces defining the discrete scheme, we propose next a heuristic procedure to choose φ_h. Hereafter, given a nonnegative integer l, $\mathbf{P}_l(T)$ denotes the space of polynomials of degree $\leqslant l$ defined on T.

We first compute local functions $\varphi_{h,T}$, for each $T\in\mathscr{T}_h$, satisfying:

1. $\varphi_{h,T}\in[\mathbf{P}_2(T)]^2$.
2. $\nabla\varphi_{h,T}$ is the $[L^2(T)]^{2\times 2}$-projection of $\mathbf{t}_h|_T$ onto the space $\nabla[\mathbf{P}_2(T)]^2$.
3. $\varphi_{h,T}(\bar{\mathbf{x}}_T)=\mathbf{u}_h|_T$, where $\bar{\mathbf{x}}_T$ is the barycentre of T.

It is easy to see that each $\varphi_{h,T}$ is uniquely determined by the above conditions. Then, similarly as in Reference [20], we now define φ_h as the *continuous average* of the functions $\varphi_{h,T}$. In other words, φ_h is the unique function in $[C(\bar{\Omega})]^2$ satisfying:

1. $\varphi_h|_T\in[\mathbf{P}_2(T)]^2$ for each $T\in\mathscr{T}_h$.
2. For each vertex $\bar{\mathbf{x}}$ of $\mathscr{T}_h$ lying on Γ and for each middle point $\bar{\mathbf{x}}$ of the edges $e\in E_h(\Gamma)$, $\varphi_h(\bar{\mathbf{x}})=\mathbf{g}(\bar{\mathbf{x}})$.
3. For each vertex $\bar{\mathbf{x}}$ of $\mathscr{T}_h$ lying in Ω and for each middle point $\bar{\mathbf{x}}$ of the edges $e\in E_h-E_h(\Gamma)$, $\varphi_h(\bar{\mathbf{x}})$ is the average of the values $\varphi_{h,T}(\bar{\mathbf{x}})$ on all the triangles $T\in\mathscr{T}_h$ to which $\bar{\mathbf{x}}$ belongs.

We observe here that the computations of the local functions $\varphi_{h,T}$ (through a projection argument) and the global P_2-interpolant φ_h are standard procedures in the finite element method. Moreover, they are comparable (in complexity) to the computations of approximate solutions of the local problems (16) via higher order $H(\mathbf{div};T)$ subspaces.

5. NUMERICAL RESULTS

In this section, we provide several numerical examples illustrating the performance of the mixed finite element scheme (11) and the fully explicit *a posteriori* error estimate $\boldsymbol{\theta}$ (cf. (21) and (22)) with the above described choice of the auxiliary function φ_h.

Hereafter, N is the number of degrees of freedom defining the subspaces $X_{1,h}$, $M_{1,h}$, and M_h, that is, $N := 9\,(\text{number of triangles of } \mathscr{T}_h) + 2\,(\text{number of edges of } \mathscr{T}_h) + 1$, which leads asymptotically to 12 unknowns per triangle. In order to compare this amount with the number of unknowns employed by more traditional approaches, we consider the $\mathbf{P}_1$-iso$\mathbf{P}_1$ finite element pair for the usual primal-mixed variational formulation of the generalized Stokes problem. In this case, the pressure is approximated by continuous piecewise linear elements on a triangular mesh, and the velocity is approximated by continuous piecewise linear elements on the finer mesh obtained by refining each element in the pressure mesh into four elements using the midpoints of each side. It is not difficult to see that these finite element subspaces lead asymptotically to 4.5 degrees of freedom per triangle. We emphasize then that the additional cost of 7.5 unknowns per triangle in our present mixed approach is certainly justified by the fact that two other quantities of physical interest ($\mathbf{t}$ and $\boldsymbol{\sigma}$) are approximated directly, without any need of numerical postprocessing. In addition, on the contrary to the $\mathbf{P}_1$-iso$\mathbf{P}_1$ approximation where one has to impose continuity requirements at each node for the pressure p_h and at each node and midpoint for the components of the velocity $\mathbf{u}_h$, we only need here to consider the continuity of the normal components of the flux $\boldsymbol{\sigma}_h$ at the edges of the triangulation. This yields therefore a much simpler computation of the global stiffness matrix of the present mixed finite element method since all the sub-matrices connecting only the other three main unknowns ($\mathbf{t}_h$, $\mathbf{u}_h$, and p_h) become *block-diagonal* (in a general sense including rectangular matrices). Consequently, computing in particular the inverse of the sub-matrix generated by $\mathbf{t}_h$ and $\mathbf{u}_h$ is straightforward, which allows to eliminate these unknowns, thus simplifying significantly the solution procedure of the whole discrete system (11).

We now provide further notations. The individual and global errors are defined as follows:

$$\mathbf{e}(\mathbf{t}) := \|\mathbf{t} - \mathbf{t}_h\|_{[L^2(\Omega)]^{2\times 2}}, \quad \mathbf{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{[L^2(\Omega)]^2}$$

$$\mathbf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\mathbf{div};\Omega)}, \quad \mathbf{e}(p) := \|p - p_h\|_{L^2(\Omega)}, \quad \mathbf{e}(\xi) := |\xi - \xi_h|$$

and

$$\mathbf{e} := \{[\mathbf{e}(\mathbf{t})]^2 + [\mathbf{e}(\boldsymbol{\sigma})]^2 + [\mathbf{e}(p)]^2 + [\mathbf{e}(\mathbf{u})]^2 + [\mathbf{e}(\xi)]^2\}^{1/2}$$

where $(\mathbf{t}, \mathbf{u}, \boldsymbol{\sigma}, p, \xi)$ and $(\mathbf{t}_h, \mathbf{u}_h, \boldsymbol{\sigma}_h, p_h, \xi_h)$ are the unique solutions of the continuous and discrete mixed formulations, respectively. In addition, the effectivity index associated to the *a posteriori* error estimate $\boldsymbol{\theta}$ is given by $\mathbf{e}/\boldsymbol{\theta}$. Also, given two consecutive triangulations with degrees of freedom N and N', and corresponding total errors given by $\mathbf{e}$ and $\mathbf{e}'$, the experimental rate of convergence is defined by $\boldsymbol{\gamma} := -2[(\log(\mathbf{e}/\mathbf{e}'))/(\log(N/N'))]$.

On the other hand, the adaptive algorithm used in the mesh refinement process is the following (see Reference [21]):

1. Start with a coarse mesh $\mathscr{T}_h$.
2. Solve the discrete problem (11) for the current mesh $\mathscr{T}_h$.

3. Compute the auxiliary function φ_h.
4. Compute θ_T for each triangle $T \in \mathscr{T}_h$.
5. Evaluate stopping criterion and decide to finish or go to next step.
6. Use *blue–green* procedure to refine each $T' \in \mathscr{T}_h$ whose indicator $\theta_{T'}$ satisfies

$$\theta_{T'} \geqslant \tfrac{1}{2} \max\{\theta_T : T \in \mathscr{T}_h\}$$

7. Define resulting mesh as current mesh $\mathscr{T}_h$ and go to step 2.

We remark here that the $H^{1/2}$-norm appearing in the computation of θ_T (cf. (21)) is approximated by means of an interpolation estimate, that is, given $e \in E(T) \cap E_h(\Gamma)$, we consider the bound

$$\|\varphi_h - \mathbf{g}\|^2_{[H^{1/2}_{00}(e)]^2} \leqslant \|\varphi_h - \mathbf{g}\|_{[L^2(e)]^2} \|\varphi_h - \mathbf{g}\|_{[H^1_0(e)]^2}$$

The numerical results presented below were obtained in a *Compaq Alpha ES40 Parallel Computer* using a MATLAB code. We consider four examples of the generalized Stokes problem (1), with different choices of the parameters α and ν for three of them. The data $\mathbf{f}$ and $\mathbf{g}$ are chosen so that the velocity $\mathbf{u} := (u_1, u_2)^{\mathrm{t}}$ and the pressure p are the ones shown below in Table I. Example 1 considers smooth solutions for both unknowns. Then, Examples 2 and 4 illustrate the case of velocities with boundary and inner layers, respectively. More precisely, $\mathbf{u}$ has a boundary layer around the origin in Example 2, while $\mathbf{u}$ has an inner layer around the line $x_2 = 0.5 - x_1$ in Example 4. Finally, Example 3 considers the case of $\mathbf{u}$ with a singularity around the exterior neighbourhood of the boundary point (1,1). Certainly, the velocities of the four examples are divergence free.

Before commenting on the numerical results obtained, we observe in advance that the information on the individual error $\mathbf{e}(\xi)$ is not displayed below since it converges very rapidly to zero in all the examples considered.

In Tables II and III we give the individual and global errors, the effectivity index $\mathbf{e}/\boldsymbol{\theta}$, and the experimental rate of convergence $\boldsymbol{\gamma}$ for the uniform refinement as applied to Example 1 with pairs of parameters $(\alpha, \nu) = (10, 1)$ and $(\alpha, \nu) = (100, 1)$. In addition, Figures 1 and 3 show the corresponding individual errors vs the degrees of freedom N. We observe here that the rates of convergence behave as predicted by the theory, that is, of $O(h)$, and that, due to the order of the constant $\bar{C}(\alpha, \nu)$ in Theorem 3, some of these rates begin to deteriorate as α increases. We also notice that the dominant components of the global error are given by $\mathbf{e}(\mathbf{t})$ and $\mathbf{e}(\boldsymbol{\sigma})$. Further, the effectivity indexes remain bounded above and below as N increases (with smaller lower bounds for bigger α), which confirms the reliability of $\boldsymbol{\theta}$ and provides numerical evidences for it being efficient. However, as shown in Figures 2 and 4, and due to

Table I. Domain and exact solution for each example.

Example	Ω	$u_1(x_1, x_2)$	$u_2(x_1, x_2)$	$p(x_1, x_2)$
1	$(-1, 1)^2$	$-e^{x_1}(x_2 \cos(x_2) + \sin(x_2))$	$e^{x_1} x_2 \sin(x_2)$	$2e^{x_1} \sin(x_2)$
2	$(0, 1)^2$	$-\sqrt{\alpha} e^{-\sqrt{\alpha}(x_1 + x_2)}$	$-u_1(x_1, x_2)$	$2e^{2x_1 - 1} \sin(2x_2 - 1)$
3	$(-1, 1)^2$	$-(2.1 - x_1 - x_2)^{-1/3}$	$-u_1(x_1, x_2)$	$2e^{x_1} \sin(x_2)$
4	$(0, 1)^2$	$-\sqrt{\alpha} e^{-\sqrt{\alpha}(0.5 - x_1 - x_2)^2}$	$-u_1(x_1, x_2)$	$2e^{2x_1 - 1} \sin(2x_2 - 1)$

Table II. Individual errors, total error $\mathbf{e}$, effectivity index, and global rate of convergence for the uniform refinement (Example 1, $\alpha = 10, \nu = 1$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{e}(p)$	$\mathbf{e}$	$\mathbf{e}/\boldsymbol{\theta}$	γ
29	3.5762	2.4040	7.7467	3.6114	9.5720	0.3968	—
105	2.2827	1.2595	4.3739	2.0922	5.5051	0.4343	0.8598
401	1.3824	0.6473	2.5351	1.2701	3.2203	0.4912	0.8002
1569	0.7853	0.3244	1.2887	0.6462	1.6735	0.5029	0.9596
6209	0.4206	0.1616	0.5951	0.2791	0.7969	0.4755	1.0787
24705	0.2163	0.0806	0.2776	0.1196	0.3803	0.4525	1.0710
98561	0.1092	0.0402	0.1343	0.0548	0.1860	0.4423	1.0338

Table III. Individual errors, total error $\mathbf{e}$, effectivity index, and global rate of convergence for the uniform refinement (Example 1, $\alpha = 100, \nu = 1$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{e}(p)$	$\mathbf{e}$	$\mathbf{e}/\boldsymbol{\theta}$	γ
29	3.3494	2.3697	10.3296	4.1105	11.8504	0.0500	—
105	2.0578	1.2424	7.7385	3.1434	8.6916	0.0700	0.4818
401	1.1765	0.6395	5.8418	2.3861	6.4509	0.1011	0.4449
1569	0.6821	0.3224	3.3689	1.4000	3.7254	0.1158	0.8048
6209	0.3909	0.1613	1.4540	0.6076	1.6317	0.1013	1.2003
24705	0.2109	0.0806	0.5048	0.2096	0.5914	0.0734	1.4697
98561	0.1084	0.0402	0.1765	0.0720	0.2230	0.0553	1.4095

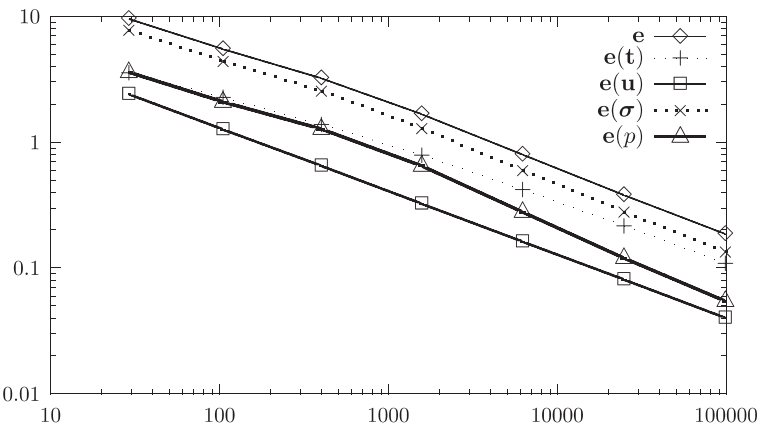


Figure 1. Errors vs N for the uniform refinement (Example 1, $\alpha = 10, \nu = 1$).

the fact that the solution of Example 1 is smooth, there is no relevant difference between the uniform and adaptive procedures for the global error $\mathbf{e}$ vs N.

The numerical results concerning Example 2 are presented in Tables IV and V where we display the individual and global errors, the effectivity index $\mathbf{e}/\boldsymbol{\theta}$, and the experimental rate of convergence γ for both refinements with $(\alpha, \nu) = (100, 0.5)$ and $(\alpha, \nu) = (1000, 0.5)$. We note

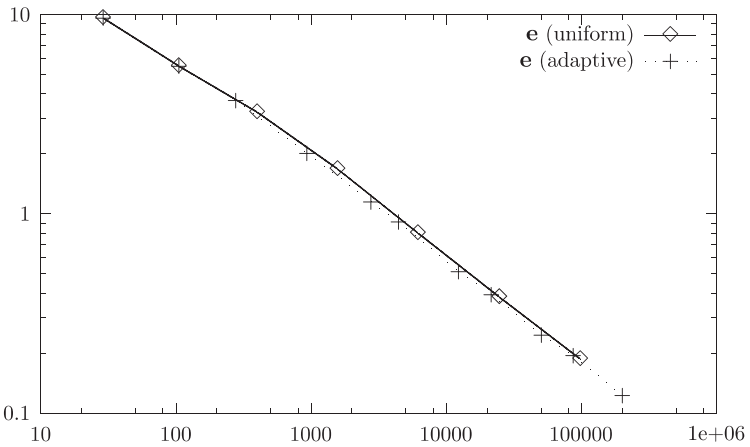


Figure 2. $\mathbf{e}$ vs N for both refinements (Example 1, $\alpha = 10, v = 1$).

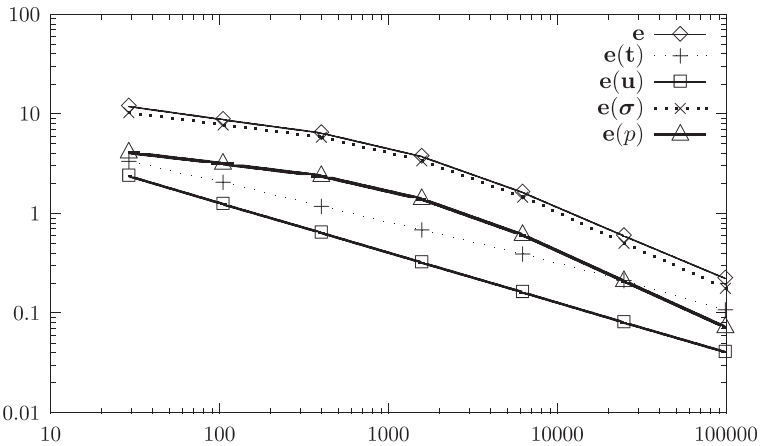


Figure 3. Errors vs N for the uniform refinement (Example 1, $\alpha = 100$, $v = 1$).

from these tables that only $\mathbf{e}(\boldsymbol{\sigma})$ constitutes now the dominant part of $\mathbf{e}$. In addition, we observe a very clear difference between the uniform and adaptive refinements. The global error $\mathbf{e}$ of the latter decreases much faster than that of the former, thus recovering the rate of convergence $O(h)$. As shown in Figures 5 and 6, this fact is even more pronounced for $\alpha = 1000$ where the convergence of the uniform refinement is very slow. However, although the effectivity indexes remain bounded above and below as N increases, for the two pairs of parameters and for both refinements, they also increase as α gets larger. Next, Meshes 1 and 2 display some intermediate meshes obtained with the adaptive refinement algorithm. It is interesting to confirm, as expected, that the procedure is able to recognize the boundary layer around $(0, 0)$. Also, we remark that this refinement is even more localized near

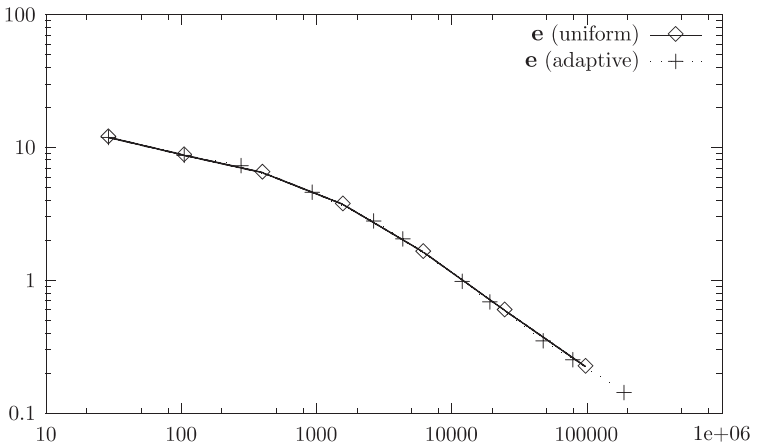


Figure 4. **e** vs N for both refinements (Example 1, $\alpha = 100, \nu = 1$).

Table IV. Individual errors, total error **e**, effectivity index, and global rate of convergence for both refinements (Example 2, $\alpha = 100, \nu = 0.5$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{e}(p)$	**e**	$\mathbf{e}/\boldsymbol{\theta}$	γ
105	5.4391	0.5032	50.7765	3.2059	51.1700	4.5545	—
401	4.7510	0.5047	50.7097	2.4017	50.9908	8.0530	0.0052
1569	2.7950	0.3270	32.8015	1.1490	32.9420	10.4012	0.6405
6209	1.5002	0.1760	17.6482	0.4435	17.7182	11.1071	0.9016
24705	0.7768	0.0896	8.9857	0.1716	9.0213	11.1273	0.9775
98561	0.3937	0.0450	4.5126	0.0740	4.5305	11.0640	0.9955
105	5.4391	0.5032	50.7765	3.2059	51.1700	4.5545	—
251	4.7575	0.5050	50.7475	2.3963	51.0289	7.9226	0.0063
397	2.8246	0.3287	33.0313	1.5115	33.1880	9.5209	1.8766
543	1.7132	0.1857	18.7508	1.2146	18.8690	8.0700	3.6061
1105	1.2094	0.1080	10.9389	0.9160	11.0442	7.3584	1.5077
2642	0.7755	0.0624	6.3504	0.6207	6.4279	6.5904	1.2418
4890	0.6248	0.0518	5.2416	0.3865	5.2931	7.2317	0.6310
11810	0.4113	0.0303	3.0674	0.2027	3.1016	6.3483	1.2122
26486	0.2902	0.0217	2.2001	0.1144	2.2222	6.4654	0.8256
47876	0.2133	0.0150	1.5206	0.0723	1.5373	6.0896	1.2448
99636	0.1528	0.0112	1.1353	0.0447	1.1464	6.3580	0.8004
190594	0.1084	0.0075	0.7664	0.0289	0.7746	6.0338	1.2089

the origin for $\alpha = 1000$, which is due to the fact that this layer becomes thinner as α gets larger.

Next, Tables VI and VII provide the numerical results obtained for Example 3 with $(\alpha, \nu) = (10, 1)$ and $(\alpha, \nu) = (100, 1)$. As for the previous examples, $\mathbf{e}(\boldsymbol{\sigma})$ is again the dominant part of the global error **e**. Also, we observe that the effectivity indexes remain bounded above and below as the number of degrees of freedom N increases, with bounds close to 1.0, which confirms the reliability of $\boldsymbol{\theta}$ and constitutes numerical evidences of its eventual

Table V. Individual errors, total error $\mathbf{e}$, effectivity index, and global rate of convergence for both refinements (Example 2, $\alpha = 1000, \nu = 0.5$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{e}(p)$	$\mathbf{e}$	$\mathbf{e}/\boldsymbol{\theta}$	γ
105	7.0383	0.0643	65.0020	5.0629	65.5777	1.0652	—
401	12.9106	0.3337	334.0185	7.4337	334.3507	7.3412	—
1569	17.7197	0.5382	538.4277	8.1264	538.7808	18.2978	—
6209	12.7694	0.4502	450.3525	5.3283	450.5652	28.8851	0.2599
24705	7.1160	0.2680	268.0732	2.3559	268.1781	34.7400	0.7514
98561	3.7577	0.1404	140.4863	0.9185	140.5396	36.0464	0.9339
105	7.0383	0.0643	65.0020	5.0629	65.5777	1.0652	—
251	12.9142	0.3338	334.0546	6.6592	334.3706	7.3370	—
397	17.7427	0.5382	538.4458	7.0993	538.7851	18.2613	—
543	12.8121	0.4501	450.2576	4.9853	450.4676	28.6842	1.1433
689	7.2888	0.2693	269.4196	2.6824	269.5316	33.3358	4.3135
835	4.7125	0.1543	154.4175	1.8963	154.5011	29.2990	5.7909
1722	2.7302	0.0798	79.8489	1.4523	79.9088	25.8793	1.8217
3839	1.6684	0.0463	46.3600	1.3075	46.4084	24.4862	1.3555
7031	1.2912	0.0363	36.4156	1.1266	36.4560	25.8745	0.7977
14781	0.8664	0.0239	24.0379	1.0019	24.0744	24.4694	1.1169
27310	0.6568	0.0182	18.2701	0.7172	18.2960	25.0696	0.8941
56312	0.4520	0.0123	12.3484	0.5608	12.3694	24.2670	1.0818
110822	0.3375	0.0090	9.0979	0.2983	9.1090	24.2919	0.9038

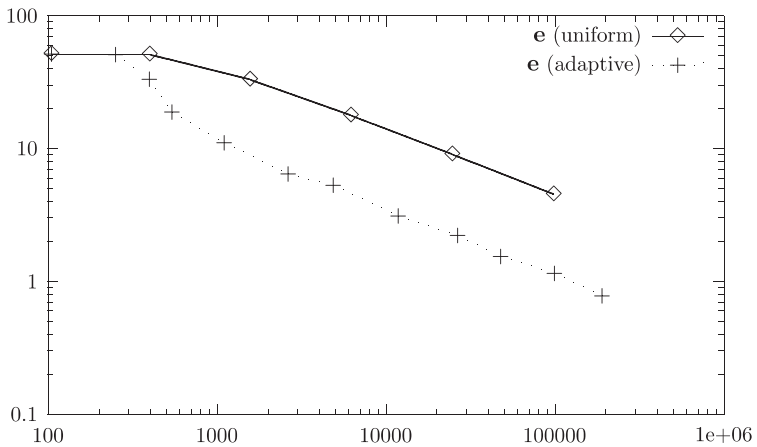


Figure 5. $\mathbf{e}$ vs N for both refinements (Example 2, $\alpha = 100, \nu = 0.5$).

efficiency. In addition, according to the experimental rates of convergence, which are also illustrated in Figures 7 and 8, the adaptive procedure yields again the quasi-optimal rate of convergence $O(h)$ for the global error $\mathbf{e}$. Moreover, as expected, the adaptive refinement algorithm is able to identify the singularities of the problem. In fact, as shown by Meshes 3 and 4, the adapted meshes are highly refined around the boundary point $(1, 1)$, in whose outer

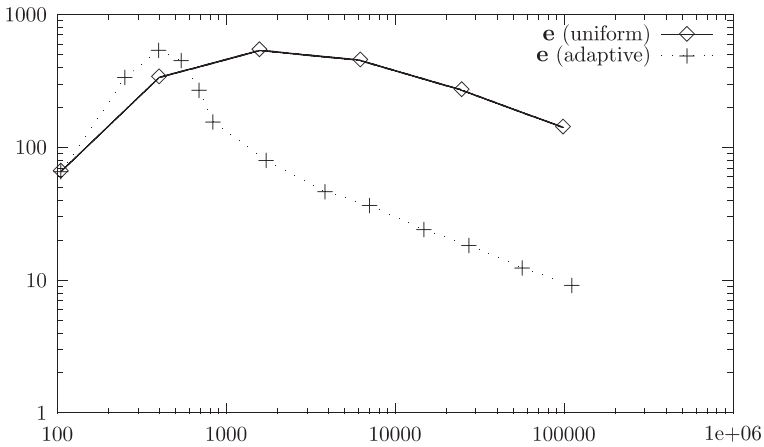
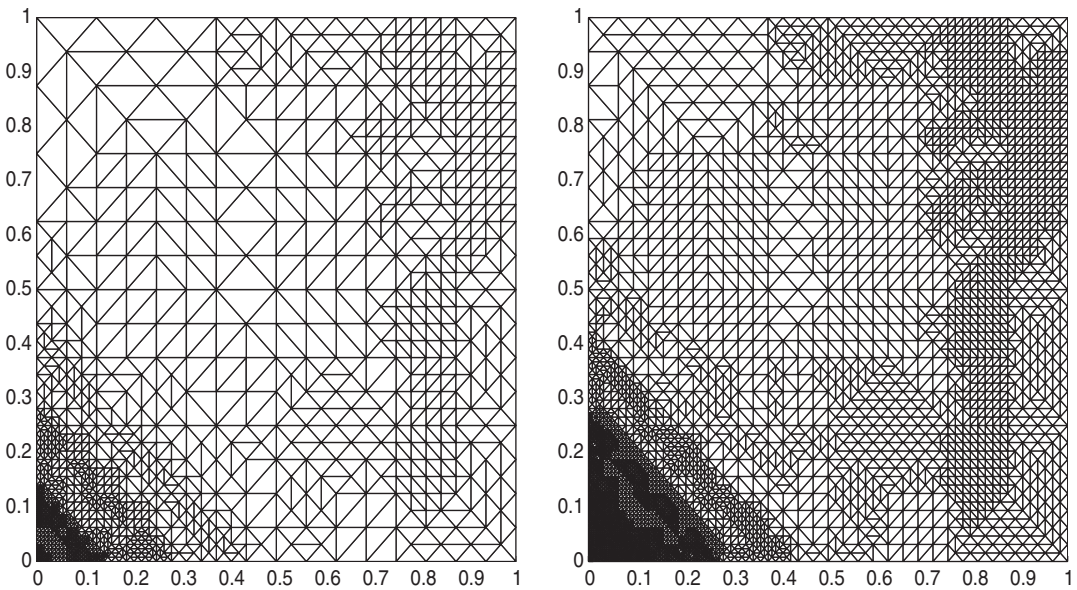


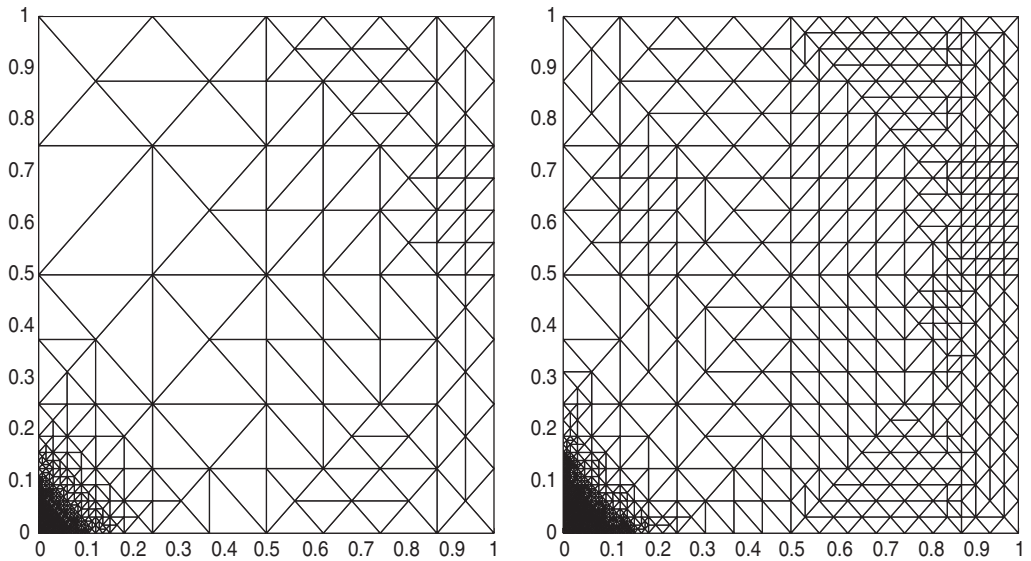
Figure 6. **e** vs N for both refinements (Example 2, $\alpha = 1000, \nu = 0.5$).



Meshes 1. Adapted intermediate meshes with 26 486 and 99 636 degrees of freedom, respectively, for Example 2, $\alpha = 100, \nu = 0.5$.

neighbourhood the singularity lives. Further, similarly as for Example 2, the refinement is even more localized as α gets larger.

Finally, the numerical results concerning Example 4 with $(\alpha, \nu) = (1000, 0.5)$ are collected in Table VIII. The remarks and conclusions here are similar to those for Examples 2 and 3. Again, the effectivity indexes remain bounded, with bounds around 0.11, and, although the



Meshes 2. Adapted intermediate meshes with 14 781 and 56 312 degrees of freedom, respectively, for Example 2, $\alpha = 1000, \nu = 0.5$.

Table VI. Individual errors, total error **e**, effectivity index, and global rate of convergence for both refinements (Example 3, $\alpha = 10, \nu = 1$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\sigma)$	$\mathbf{e}(p)$	$\mathbf{e}$	$\mathbf{e}/\theta$	γ
29	0.7664	0.4607	6.0365	2.8186	6.7219	1.5756	—
105	0.8281	0.2984	5.1990	2.0257	5.6487	1.8869	0.2703
401	0.6839	0.1733	5.3278	1.1694	5.5001	1.3734	0.0398
1569	0.4615	0.0898	5.3663	0.5403	5.4139	1.1420	0.0231
6209	0.2660	0.0447	4.3239	0.2269	4.3382	1.0762	0.3220
24705	0.1407	0.0223	2.7594	0.0984	2.7649	1.0551	0.6523
98561	0.0718	0.0111	1.5160	0.0454	1.5185	1.0483	0.8662
29	0.7664	0.4607	6.0365	2.8186	6.7219	1.5756	—
105	0.8281	0.2984	5.1990	2.0257	5.6487	1.8869	—
325	0.6916	0.1767	5.3524	1.1959	5.5307	1.3761	—
471	0.5949	0.1254	5.5517	0.9228	5.6606	1.1705	—
617	0.5444	0.1105	4.6251	0.8722	4.7393	1.1340	1.3158
763	0.5281	0.1074	3.2728	0.8655	3.4280	1.1900	3.0502
909	0.5244	0.1069	2.4193	0.8646	2.6243	1.3368	3.0515
1055	0.5238	0.1069	2.1416	0.8644	2.3706	1.4524	1.3654
3126	0.3502	0.0570	1.1486	0.4108	1.2704	1.3863	1.1485
8157	0.2154	0.0305	0.6736	0.2004	0.7357	1.3126	1.1390
12794	0.1926	0.0258	0.5409	0.1717	0.5998	1.3094	0.9071
32524	0.1147	0.0148	0.3269	0.0905	0.3583	1.2666	1.1042
54229	0.0949	0.0117	0.2551	0.0726	0.2819	1.2478	0.9378
130544	0.0578	0.0071	0.1621	0.0423	0.1774	1.2398	1.0541

Table VII. Individual errors, total error **e**, effectivity index, and global rate of convergence for both refinements (Example 3, $\alpha = 100$, $v = 1$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{e}(p)$	$\mathbf{e}$	$\mathbf{e}/\boldsymbol{\theta}$	γ
29	0.3670	0.3899	7.2973	2.9958	7.9065	0.2145	—
105	0.4457	0.2537	7.0634	2.8054	7.6175	0.3444	0.0578
401	0.4376	0.1525	6.7047	2.1753	7.0639	0.6181	0.1125
1569	0.3580	0.0849	5.9249	1.2798	6.0727	1.0757	0.2216
6209	0.2400	0.0441	4.4515	0.5381	4.4906	1.3042	0.4388
24705	0.1363	0.0222	2.7790	0.1799	2.7883	1.2948	0.6901
98561	0.0712	0.0111	1.5186	0.0607	1.5215	1.2738	0.8754
29	0.3670	0.3899	7.2973	2.9958	7.9065	0.2145	—
105	0.4457	0.2537	7.0634	2.8054	7.6175	0.3444	0.0578
251	0.4305	0.1595	7.1218	2.4020	7.5300	0.6149	0.0264
789	0.3678	0.0912	6.4211	1.7284	6.6605	1.0505	0.2142
2483	0.2812	0.0501	4.7966	0.9914	4.9063	1.2275	0.5332
2629	0.2520	0.0431	3.4981	0.9790	3.6415	0.9561	10.4354
2947	0.2383	0.0403	2.6186	0.9350	2.7910	0.8150	4.6587
4633	0.2129	0.0327	1.9695	0.7235	2.1093	0.7499	1.2380
11357	0.1568	0.0214	1.3188	0.3226	1.3669	0.7916	0.9675
22547	0.1249	0.0147	0.7832	0.2293	0.8257	0.6380	1.4699
46839	0.0819	0.0101	0.5382	0.1011	0.5538	0.6511	1.0928
89728	0.0673	0.0074	0.3716	0.0744	0.3850	0.5881	1.1184
187253	0.0423	0.0051	0.2704	0.0362	0.2761	0.6391	0.9030

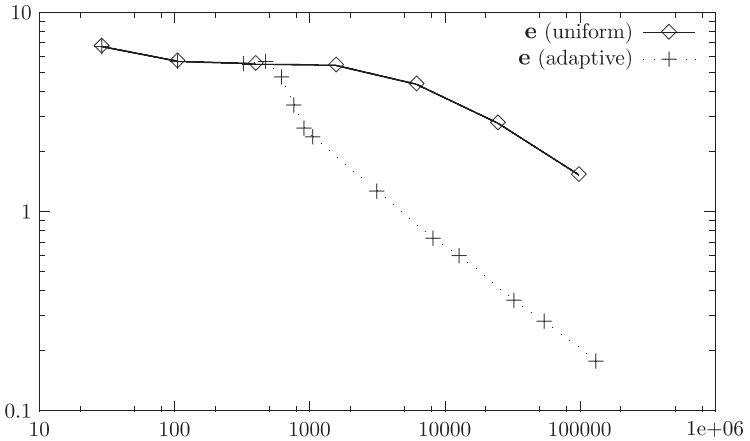


Figure 7. **e** vs N for both refinements (Example 3, $\alpha = 10, v = 1$).

experimental rates of convergence of both refinements approach 1 as N increases, the global error of the adaptive one begins to decrease before than the uniform one. This fact is clearly observed in Figure 9 where the curve **e** vs N is shown. In addition, as expected, the corresponding adaptive refinement algorithm is able to recognize the inner layer of the problem.

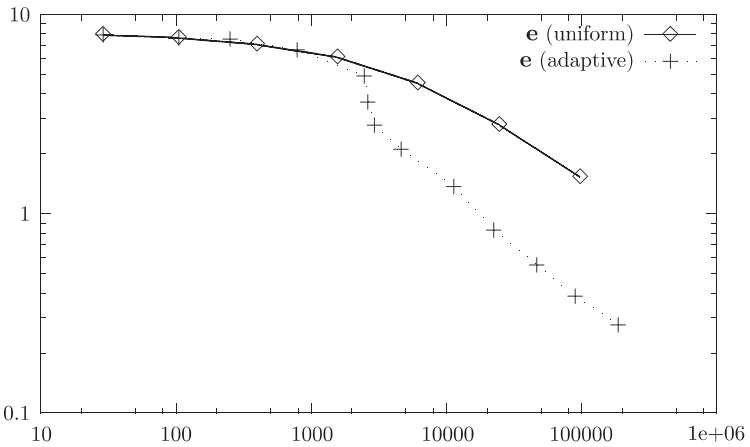
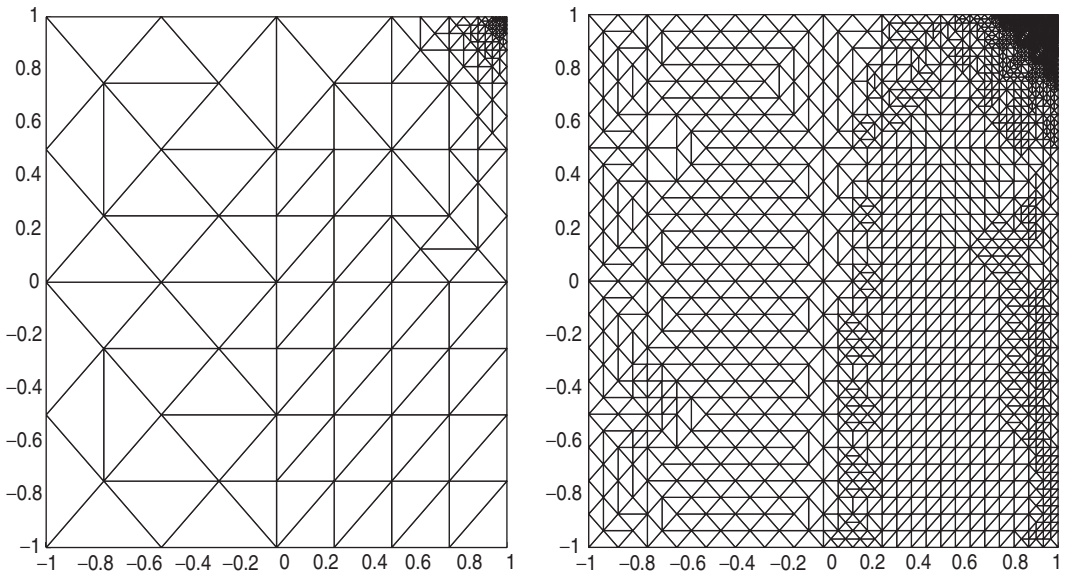


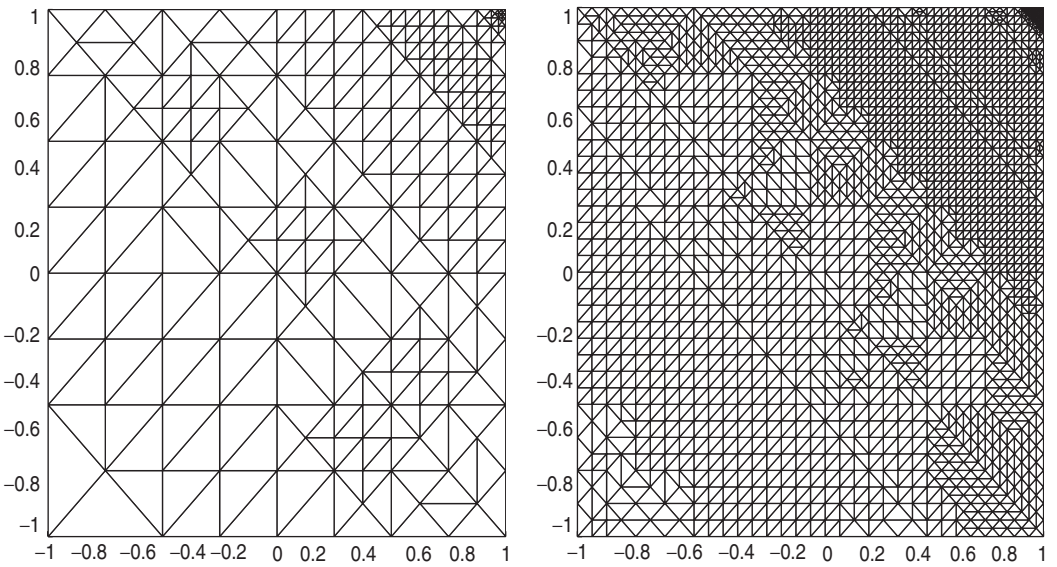
Figure 8. **e** vs N for both refinements (Example 3, $\alpha = 100$, $\nu = 1$).



Meshes 3. Adapted intermediate meshes with 3126 and 54 229 degrees of freedom, respectively, for Example 3, $\alpha = 10$, $\nu = 1$.

Indeed, as can be seen in Meshes 5, the adapted meshes are highly refined around the line $x_2 = 0.5 - x_1$. We also notice here that the refinements identify a thin band exactly on this line, which corresponds to the flat behaviour of the solution caused by the power 2 in the exponent of the exponential function.

On the other hand, although we mentioned before that the solutions of the local problems (16) could be approximated via higher order $H(\mathbf{div}; T)$ subspaces, we show next that this



Meshes 4. Adapted intermediate meshes with 4633 and 46 839 degrees of freedom, respectively, for Example 3, $\alpha = 100$, $\nu = 1$.

Table VIII. Individual errors, total error **e**, effectivity index, and global rate of convergence for both refinements (Example 4, $\alpha = 1000$, $\nu = 0.5$).

N	$\mathbf{e}(\mathbf{t})$	$\mathbf{e}(\mathbf{u})$	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{e}(p)$	$\mathbf{e}$	$\mathbf{e}/\boldsymbol{\theta}$	γ
105	74.3233	10.9499	975.4179	26.3889	978.6626	0.0838	—
401	43.9757	7.2191	790.4335	25.1620	792.0885	0.1017	0.3156
1569	32.5849	4.0695	513.2866	11.7287	514.4697	0.1160	0.6326
6209	17.3618	2.1141	264.9580	5.2669	265.5869	0.1147	0.9613
24705	8.9221	1.0675	132.0790	2.1941	132.4025	0.1130	1.0080
98561	4.5016	0.5350	65.7869	0.8889	65.9489	0.1123	1.0074
105	74.3233	10.9499	975.4179	26.3889	978.6626	0.0838	—
251	47.5896	7.5069	785.4328	25.8358	787.3331	0.0978	0.4992
907	34.5913	4.2248	536.2225	13.2819	537.5178	0.1164	0.5942
2111	20.9361	2.1538	292.7591	8.1437	293.6276	0.1247	1.4315
7141	11.2381	1.1037	148.3617	4.1104	148.8476	0.1240	1.1149
17599	9.4008	0.6518	91.1181	3.1416	91.6579	0.1334	1.0750
30105	5.9042	0.5120	68.4031	2.1752	68.6938	0.1244	1.0744
87607	4.1128	0.2886	41.6759	1.6638	41.9124	0.1367	0.9250
120853	3.5184	0.2520	33.8690	1.3457	34.0788	0.1243	1.2862

additional computational effort would not necessarily improve the efficiency of the *a posteriori* error estimate. To this end, we now identify the main components of $\boldsymbol{\theta}$, which are given by those terms providing, respectively, the *a priori* bounds for the local solutions $\hat{\boldsymbol{\sigma}}_T$, and the residuals of the constitutive, equilibrium, and compressibility equations. More precisely,

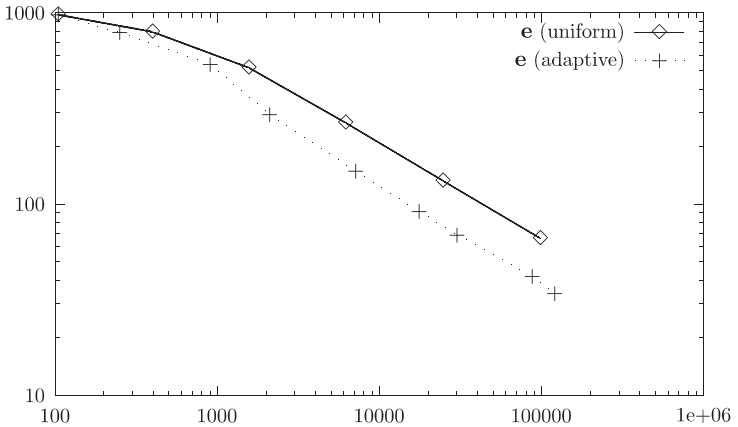


Figure 9. **e** vs N for both refinements (Example 4, $\alpha = 1000, \nu = 0.5$).

according to (21) and (22), we can write $\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\varphi}^2 + \boldsymbol{\theta}_{\mathrm{res}}^2\}^{1/2}$, where

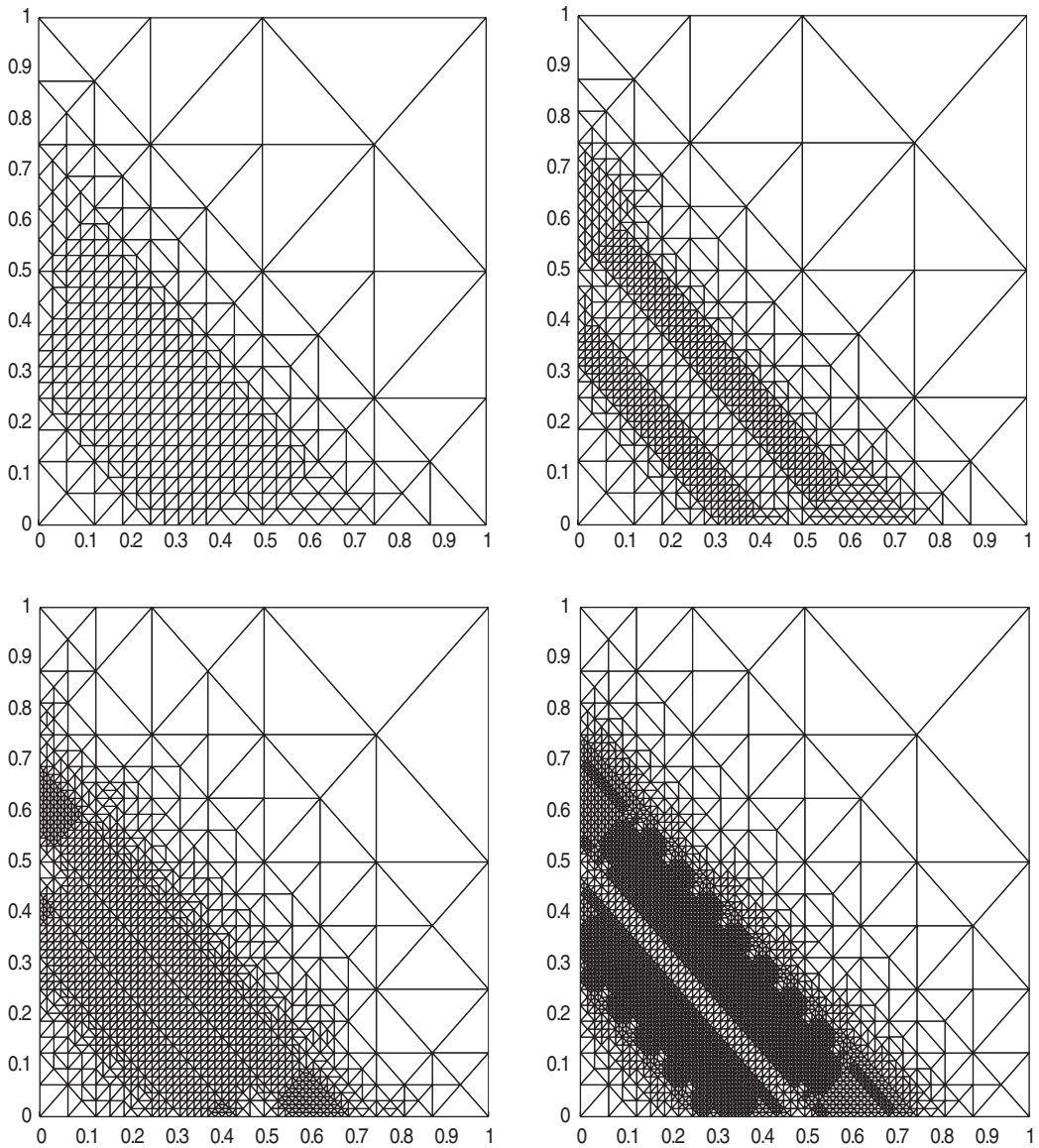
$$\boldsymbol{\theta}_{\varphi}^2 := \sum_{T \in \mathscr{T}_h} \{\|\mathbf{t}_h - \nabla \varphi_h\|_{[L^2(T)]^{2\times 2}}^2 + \|\mathbf{u}_h - \varphi_h\|_{[L^2(T)]^2}^2 + h_T^2 |\xi_h|^2\} + \sum_{e \in E_h(\Gamma)} \|\varphi_h - \mathbf{g}\|_{[H_{00}^{1/2}(e)]^2}^2$$

and

$$\boldsymbol{\theta}_{\mathrm{res}}^2 := \sum_{T \in \mathscr{T}_h} \{\|\boldsymbol{\sigma}_h - \nu \mathbf{t}_h + p_h \mathbf{I}\|_{[L^2(T)]^{2\times 2}}^2 + \|\mathbf{f} + \mathbf{div}(\boldsymbol{\sigma}_h) - \alpha \mathbf{u}_h\|_{[L^2(T)]^2}^2 + \|\mathrm{tr}(\mathbf{t}_h)\|_{L^2(T)}^2\}$$

Then, in Table IX we display the components $\boldsymbol{\theta}_{\varphi}$ and $\boldsymbol{\theta}_{\mathrm{res}}$ for the examples considered in this section. We observe there that in most cases the component $\boldsymbol{\theta}_{\mathrm{res}}$ is the very dominant one. The only exception is the situation concerning Example 2 in which $\boldsymbol{\theta}_{\varphi}$ and $\boldsymbol{\theta}_{\mathrm{res}}$ are of the same order. Consequently, any improvement in the *a priori* bound for $\hat{\boldsymbol{\sigma}}_T$ (even if it could be computed exactly) will not modify the efficiency of the *a posteriori* error estimator. In other words, these examples support the belief that $\tilde{\boldsymbol{\theta}}$ will not necessarily yield better results than those given by $\boldsymbol{\theta}$.

As general remarks, we would like to observe first that the numerical examples presented here behave much better than what the previous theoretical results insinuated. In particular, the order of the constants obtained in Theorems 2 and 3 indicate that the rates of convergence are affected by large values of α, which, nevertheless, was not too severe in the examples. Further, since α is proportional to the inverse of the time-step Δt, the estimates provided in these theorems also indicate that the convergence of time-dependent solutions should deteriorate as Δt decreases. Whether this holds exactly as predicted by the theory or behaves better than that is something to be seen from corresponding numerical results. Now, according to the constants in Theorem 6 and Lemma 7, one would have expected effectivity indexes between $O(\alpha^{-1})$ and $O(\alpha^3/\nu)$. However, as we could see, they all lie on ranges much tighter than that, they do not deteriorate as N increases, and they additionally improve when passing from uniform to adaptive refinements. The above observations yield the conjecture that these



Meshes 5. Adapted intermediate meshes with 7141, 17 599, 30 105 and 87 607 degrees of freedom, respectively, for Example 4, $\alpha = 1000, \nu = 0.5$.

constants are overestimated and that they could be improved. In addition, our conclusion is that the proposed mixed method is perhaps not so competitive for extremely large values of α, but it does constitute a good alternative for moderately large values of this parameter. Finally, we emphasize that the examples provide enough support for the adaptive

Table IX. Main components of the *a posteriori* error estimate $\boldsymbol{\theta}$.

N	$\boldsymbol{\theta}_{\varphi}$	$\boldsymbol{\theta}_{\text{res}}$	N	$\boldsymbol{\theta}_{\varphi}$	$\boldsymbol{\theta}_{\text{res}}$
Example 1 ($\alpha = 10, v = 1$)			*Example 1* ($\alpha = 100, v = 1$)		
29	4.5135	23.6942	29	4.3964	236.8632
105	2.5688	12.4118	105	2.4076	124.1070
401	1.5104	6.3792	401	1.3361	63.7909
1569	0.8615	3.2136	1569	0.7612	32.1361
6209	0.4659	1.6099	6209	0.4345	16.0988
24705	0.2404	0.8053	24705	0.2346	8.0533
98561	0.1213	0.4027	98561	0.1205	4.0271
Example 2 ($\alpha = 1000, v = 0.5$)			*Example 3* ($\alpha = 10, v = 1$)		
105	61.5259	2.1109	29	1.9781	3.7799
251	45.5072	2.4457	105	1.3309	2.6815
397	29.3608	2.9050	325	0.8823	3.9209
543	15.4992	2.5303	471	0.6377	4.7937
689	7.8532	1.9235	617	0.5360	4.1447
835	4.9966	1.6857	763	0.5107	2.8349
1722	2.8452	1.1997	909	0.5060	1.8968
3839	1.7123	0.8124	1055	0.5053	1.5520
7031	1.3031	0.5357	3126	0.3455	0.8487
14781	0.8945	0.4096	8157	0.2140	0.5180
27310	0.6765	0.2737	12794	0.1890	0.4173
56312	0.4681	0.2018	32524	0.1141	0.2589
110822	0.3512	0.1313	54229	0.0942	0.2054
			130544	0.0576	0.1310
Example 3 ($\alpha = 100, v = 1$)			*Example 4* ($\alpha = 1000, v = 0.5$)		
29	1.7393	36.8189	105	75.8505	11672.9488
105	1.1333	22.0862	251	55.7055	8048.8130
251	0.7085	12.2245	907	34.1669	4614.5160
789	0.4471	6.3244	2111	20.4605	2353.0475
2483	0.2940	3.9861	7141	11.3901	1199.6719
2629	0.2491	3.8005	17599	9.3005	687.0050
2947	0.2356	3.4162	30105	5.9360	552.0031
4633	0.2104	2.8048	87607	4.1046	306.5056
11357	0.1580	1.7195	120853	3.5123	274.0432
22547	0.1249	1.2881			
46839	0.0824	0.8465			
89728	0.0680	0.6510			
187253	0.0427	0.4300			

algorithm being much more efficient than a uniform refinement when solving the discrete scheme.

ACKNOWLEDGEMENTS

This research was partially supported by CONICYT-Chile through the FONDAP Program in Applied Mathematics, and by the Dirección de Investigación of the Universidad de Concepción through the Advanced Research Groups Program.

REFERENCES

1. Bakhvalov NS. Solution of the Stokes nonstationary problems by the fictitious domain method. *Russian Journal of Numerical Analysis and Mathematical Modelling* 1995; **10**:163–172.
2. Barrenechea GR, Valentin F. An unusual stabilized finite element method for a generalized Stokes problem. *Numerische Mathematik* 2002; **92**(4):653–677.
3. Cahouet J, Chabard JP. Some fast 3D finite element solvers for the generalized Stokes problem. *International Journal for Numerical Methods in Fluids* 1988; **8**:869–895.
4. Calgaro C, Laminie J. On the domain decomposition method for the generalized Stokes problem with continuous pressure. *Numerical Methods for Partial Differential Equations* 2000; **16**(1):84–106.
5. Kobelkov GM, Olshanskii MA. Effective preconditioning of Uzawa type schemes for a generalized Stokes problem. *Numerische Mathematik* 2000; **86**:443–470.
6. Pal'tsev BV. On rapidly converging iterative methods with incomplete splitting of boundary conditions for a multidimensional singularly perturbed system of Stokes type. *Russian Academy of Sciences*, *Sbornik Mathematics* 1995; **81**:487–531.
7. Sarin V, Sameh A. An efficient iterative method for the generalized Stokes problem. *SIAM Journal on Scientific Computing* 1998; **19**(1):206–226.
8. Gatica GN, González M, Meddahi S. A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. Part I: *a-priori* error analysis. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**(9–11):881–892.
9. Gatica GN, González M, Meddahi S. A low-order mixed finite element method for a class of quasi-Newtonian Stokes flows. Part II: *a-posteriori* error analysis. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**(9–11):893–911.
10. Gatica GN. Solvability and Galerkin approximations of a class of nonlinear operator equations. *Zeitschrift für Analysis und ihre Anwendungen* 2002; **21**(3):761–781.
11. Brezzi F, Fortin M. *Mixed and Hybrid Finite Element Methods*. Springer: Berlin, 1991.
12. Gatica GN. An application of Babuška–Brezzi's theory to a class of variational problems. *Applicable Analysis* 2000; **75**(3–4):297–303.
13. Gatica GN, Heuer N, Meddahi S. On the numerical analysis of nonlinear two-fold saddle point problems. *IMA Journal of Numerical Analysis* 2003; **23**(2):301–330.
14. Arnold DN, Douglas Jr J, Gupta CP. A family of higher order mixed finite element methods for plane elasticity. *Numerische Mathematik* 1984; **45**:1–22.
15. Scheurer B. Existence et approximation de point selles pou certain problemes non-linéaires. *RAIRO Analyse Numérique* 1977; **11**(4):369–400.
16. Roberts JE, Thomas J-M. Mixed and hybrid methods. In *Handbook of Numerical Analysis*, vol. II, Ciarlet y PG, Lions JL (eds). North-Holland: Amsterdam, 1991.
17. Brink U, Stein E. *A posteriori* error estimation in large-strain elasticity using equilibrated local Neumann problems. *Computer Methods in Applied Mechanics and Engineering* 1998; **161**:77–101.
18. Bank RE, Weiser A. Some *a posteriori* error estimators for elliptic partial differential equations. *Mathematics of Computation* 1985; **44**:283–301.
19. Lions JL, Magenes E. Problèmes aux Limites non Homogènes et Applications. I. Dunod: Paris, 1968.
20. Gatica GN, Stephan EP. A mixed-FEM formulation for nonlinear incompressible elasticity in the plane. *Numerical Methods for Partial Differential Equations* 2002; **18**(1):105–128.
21. Verfürth R. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner: Chichester, 1996.
22. Carstensen C, Funken S. *A-posteriori* error control in low-order finite element discretisations of incompressible stationary flows. *Mathematics of Computation* 2001; **70**(236):1353–1381.
23. Ladyzhenskaya O. New equations for the description of the viscous incompressible fluids and solvability in the large for the boundary value problems of them. *Boundary Value Problems of Mathematical Physics*, vol. V. American Mathematical Society: Providence, RI, 1970.

