AN OPTIMAL ADAPTIVE FINITE ELEMENT METHOD FOR THE STOKES PROBLEM

YAROSLAV KONDRATYUK† AND ROB STEVENSON‡

Abstract. A new adaptive finite element method for solving the Stokes equations is developed, which is shown to converge with the best possible rate. The method consists of 3 nested loops. The outermost loop consists of an adaptive finite element method for solving the pressure from the (elliptic) Schur complement system that arises by eliminating the velocity. Each of the arising finite element problems is a Stokes-type problem, with the pressure being sought in the current trial space and the divergence-free constraint being reduced to orthogonality of the divergence to this trial space. Such a problem is still continuous in the velocity field. In the middle loop, its solution is approximated using the Uzawa scheme. In the innermost loop, the solution of the elliptic system for the velocity field that has to be solved in each Uzawa iteration is approximated by an adaptive finite element method.

Key words. adaptive finite element method, convergence rates, computational complexity, a posteriori error estimators, Uzawa iteration, Stokes equations

AMS subject classifications. 65N30, 65N50, 65N15, 65Y20, 41A25

DOI. 10.1137/06066566X

1. Introduction. Often the solution of a boundary value problem exhibits singularities, e.g., due to a nonsmooth boundary. Then, because of the lacking (Sobolev) smoothness of the solution, finite element methods based on quasi-uniform partitions converge with a rate smaller than allowed by the polynomial degrees used in the finite element spaces. This can be repaired when suitable refinements are made near those singularities. The optimal distribution of the size of elements as a function of the distance to a singularity depends on the strength of the singularity, which is generally unknown.

With adaptive finite element methods (AFEMs), a sequence of nested partitions is created, where, when creating the next partition, the decision of where to refine is made on the basis of an a posteriori estimator of the error in the current finite element approximation. Although they had been used successfully for more than 25 years, in more than one space dimension, even for second order elliptic equations, their convergence was not shown before the works of Dörfler [Dör96] and that of Morin, Nochetto, and Siebert [MNS00]. Convergence alone, however, does not show that the use of an adaptive method for a solution that has singularities improves upon, or even competes with, that of a nonadaptive one. Recently, after the derivation of such a result by Binev, Dahmen, and DeVore [BDD04] for an AFEM extended with a so-called coarsening or derefinement routine, in [Ste07] we could prove that standard AFEMs converge with the best possible rate in linear complexity.
In this paper, as a model saddle point problem, we consider the Stokes equations

\[
\begin{cases}
-\Delta u + \nabla p = f & \text{on } \Omega \subset \mathbb{R}^d, \\
\text{div} u = 0 & \text{on } \Omega, \\
\mathbf{u} = 0 & \text{on } \partial \Omega.
\end{cases}
\]

(Although, in this introduction, we write equations in strong form, actually we always mean the corresponding variational formulations.) In [DDU02], Dahlke, Dahmen, and Urban analyzed an adaptive wavelet method for solving these equations. The starting point was the application of the Uzawa iteration on the continuous level, i.e., given some \( p_0 \), compute for \( j = 0, 1, \ldots \)

\[
\begin{cases}
-\Delta u_{j+1} = f - \nabla p_j & \text{on } \Omega \\
\mathbf{u}_{j+1} = 0 & \text{on } \partial \Omega.
\end{cases}
\]

\( p_{j+1} = p_j - \text{div} \mathbf{u}_j. \)

Of course, this iteration cannot be performed exactly, and in each iteration the solution of the elliptic system was approximated using an adaptive wavelet method within decreasing tolerances as the iteration proceeds. Convergence was shown, and by the inclusion of coarsening steps, even optimal rates and linear complexity were demonstrated. Since no Galerkin discretizations were formed from the mixed problem, the so-called LBB stability for the discrete problem was not required.

In [BMN02], Bänsch, Morin, and Nochetto studied the above solution method with the adaptive wavelet method replaced by an AFEM. They proved convergence, and despite the fact that they did not include coarsening, in numerical experiments they observed optimal rates, at least when the elliptic problems were solved not too accurately. When prescribing an a priori tolerance of the form \( \gamma^j \) in the \( j \)th iteration, it was necessary to take \( \gamma \) in the range \([\approx 0.95, 1)\). By the addition of coarsening to this method, in [Kon06] optimal computational complexity was demonstrated.

When starting this work, our aim was to prove optimal computational complexity of basically the method from [BMN02], i.e., without coarsening. For a reason that will be indicated later (in Remark 6.6), we did not succeed in doing this. Instead, we prove this result for a somewhat more complicated algorithm involving an additional outer loop, without coarsening though. Below we briefly describe the loops of our algorithm starting from the outermost one.

The pressure \( p \) can be found as the solution of the Schur complement equation that one obtains by eliminating the velocity \( \mathbf{u} \) from the Stokes equations. This equation is elliptic, with corresponding energy norm equivalent to the \( L^2(\Omega) \)-norm. Given a finite element space \( P_{\sigma_i} \), where \( \sigma_i \) denotes the underlying finite element partition, the best approximation from this space to \( p \) with respect to this energy norm is the Galerkin solution \( p_i \in P_{\sigma_i} \). With \( Q_{\sigma_i} \), denoting the \( L^2(\Omega) \)-orthogonal projection onto \( P_{\sigma_i} \), this \( p_i \) can be shown to be the unique solution of

\[
\begin{cases}
-\Delta \mathbf{u}^{(i)} + \nabla p_i = f & \text{on } \Omega, \\
Q_{\sigma_i} \text{div} \mathbf{u}^{(i)} = 0 & \text{on } \Omega, \\
\mathbf{u}^{(i)} = 0 & \text{on } \partial \Omega,
\end{cases}
\]

i.e., the Stokes equations in which the divergence-free condition has been relaxed. We refer to this system as the reduced Stokes equations.

Concerning the velocity, this is still a problem posed over an infinite-dimensional space. Inside the inner loops, a sufficiently accurate approximate solution will be
AN OPTIMAL ADAPTIVE FINITE ELEMENT METHOD

749
determined, which here again is denoted as \((u^{(i)}, p_i)\). We will have that \(p_i \in P_{\sigma_i}\),
while \(u^{(i)}\) is from a finite element space \(V_{\tau_i}\) with respect to a partition \(\tau_i\) that is a
refinement of \(\sigma_i\). It will be shown that the energy norm of \(p - p_i\) is equivalent to the
\textit{a posteriori error estimator} \(\|\text{div} u^{(i)}\|_{L_2(\Omega)}\), and that for any refinement \(\sigma_{i+1}\) of
\(\sigma_i\), the energy norm of \(p_{i+1} - p_i\) is equivalent to \(\|Q_{\sigma_{i+1}} \text{div} u^{(i)}\|_{L_2(\Omega)}\). Now following
along the lines of [Dör96, MNS00] for Poisson-type problems, by selecting \(\sigma_{i+1}\) such
that for some \(\theta \in (0, 1]\), \(\|Q_{\sigma_{i+1}} \text{div} u^{(i)}\|_{L_2(\Omega)} \geq \theta \|\text{div} u^{(i)}\|_{L_2(\Omega)}\) (the “bulk criterion”),
the so-called saturation property is guaranteed, and a linearly convergent sequence \((p_i)_i\) towards \(p\) is obtained. Moreover, by taking \(\theta\) to be small enough, and \(\sigma_{i+1}\)
with quasi-minimal cardinality, following along the lines of [Ste07] convergence of the \(p_i \in P_{\sigma_i}\)
towards \(p\) with the optimal rate is shown.

Compared to the adaptive methods for Poisson-type problems, a complication is
that to find such a \(\sigma_{i+1}\), it is generally not sufficient to search it within the set of
partitions that can be created by refining each element of \(\sigma_i\) some fixed number of
times. On the other hand, by considering finite element spaces such that \(\text{div} V_{\tau_i} \subseteq P_{\tau_i}\),
we know that we never have to subdivide elements that are in \(\tau_i\). To determine \(\sigma_{i+1}\)
with the required properties we apply an adaptive tree algorithm developed by Binev
and DeVore in [BD04].

For solving the reduced Stokes problem for given \(i\), i.e., fixed \(\sigma_i\), we follow the
approach from [BMN02] for the full Stokes problem. That is, we apply Uzawa, where
the pressure update then reads as \(p_{j+1}^{(i)} = p_j^{(i)} - Q_{\sigma_j} \text{div} u_j^{(i)}\). Each of the inner elliptic
systems \(-\Delta u_j^{(i)} = f - \nabla p_j^{(i)}\) on \(\Omega\), \(u = 0\) on \(\partial \Omega\) is solved with a standard AFEM
producing increasingly more refined partitions, with \(\sigma_j\) being the initial partition for
the first call. Already having optimal control over \#\(\sigma_i\), we are now able to prove
also optimal rates of the velocity approximations towards \(u\). Note that other than in
[BMN02], we have two different partitions underlying pressure and velocity approximations,
with the latter being always a refinement of the first. Throughout the algorithm, both partitions become increasingly more refined, i.e., no derefinements
are made.

This paper is organized as follows: In section 2, we recall some properties of the
Stokes problem. In section 3, we define the finite element spaces that we will use. We
give some properties of a procedure for refining partitions, which is a generalization
to arbitrary space dimensions of the newest vertex bisection method in two space
dimensions. An overview of the solution method will be given in section 4. In section 5,
a posteriori error estimators are derived for the various problems that occur in our
solution method. In section 6, the adaptive refinement routines for pressure and
velocity partitions are given. In section 7, we give the detailed description of the
adaptive method in the simplified situation that the right-hand side \(f\) is piecewise
polynomial with respect to the initial partition. We prove convergence with the
optimal rates. In this section, we assume that the arising finite-dimensional linear
systems are solved exactly, ignoring the question of computational complexity. In
section 8, we give the method for general right-hand sides, and replace the direct
solvers by iterative solution methods, with which we end up with a method of optimal
computational complexity. Finally, in section 9, we present numerical results, and
compare them with those obtained with the method from [BMN02]. As we will see,
in this example both methods give similar results.

In this paper, by \(C \lesssim D\) we will mean that \(C\) can be bounded by a multiple of
\(D\), independently of parameters which \(C\) and \(D\) may depend on. Obviously, \(C \gtrsim D\)
is defined as \(D \lesssim C\), and \(C \approx D\) as \(C \lesssim D\) and \(C \gtrsim D\).
2. **Stokes problem.** Let $\Omega$ be a polygonal domain in $\mathbb{R}^d$. We consider the Stokes problem in variational form: With
\[
\mathbb{V} := H_0^1(\Omega)^d, \quad \mathbb{P} := L_{2,0}(\Omega) := \left\{ q \in L_2(\Omega) : \int_\Omega q = 0 \right\},
\]
and given a $f \in \mathbb{V}'$, throughout this paper $u \in \mathbb{V}$ (the velocity) and $p \in \mathbb{P}$ (the pressure) will denote the solutions of
\[
(a, b, q) = \left\{ q \in L_2(\Omega) : \int_\Omega q = 0 \right\},
\]
where $a : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$, $b : \mathbb{V} \times \mathbb{P} \to \mathbb{R}$ are defined by
\[
a(w, v) := \int_\Omega \nabla w : \nabla v, \quad b(v, q) := -\int_\Omega q \text{div} v.
\]
It is well known that
\[
\|v\|_\mathbb{V} := a(v, v)^{\frac{1}{2}} \approx \|v\|_{H^1(\Omega)} \quad (v \in \mathbb{V}),
\]
b is bounded, and
\[
\beta := \inf_{0 \neq q \in \mathbb{P}} \sup_{0 \neq v \in \mathbb{V}} \frac{b(v, q)}{\|v\|_\mathbb{V} \|q\|_{L_2(\Omega)}} > 0.
\]
As a consequence, the Stokes problem is well-posed, meaning that
\[
\|w\|_\mathbb{V} + \|r\|_{L_2(\Omega)} \approx \sup_{0 \neq (v, q) \in \mathbb{V} \times \mathbb{P}} \frac{a(w, v) + b(v, r) + b(w, q)}{\|w\|_\mathbb{V} + \|r\|_{L_2(\Omega)}} \quad (w \in \mathbb{V}, r \in \mathbb{P}).
\]

**Remark 2.1.** Since $a(\cdot, \cdot)$ is elliptic on the whole of $\mathbb{V}$ instead of only on the space of the divergence-free velocities, clearly the Stokes problem (2.1) with $\mathbb{P}$ replaced by any subspace is also well-posed, uniformly in the choice of such a subspace.

Defining $A : \mathbb{V} \to \mathbb{V}'$, $B : \mathbb{V} \to \mathbb{P}'$, and $B' : \mathbb{P} \to \mathbb{V}'$ by $(Av)(w) = a(v, w)$, $(Bv)(q) = b(v, q) = (B'q)(v)$, problem (2.1) can be equivalently written as
\[
\begin{bmatrix}
A & B' \\
B & 0
\end{bmatrix}
\begin{bmatrix}
u \\
p
\end{bmatrix}
= \begin{bmatrix} f \\
p
\end{bmatrix}
\]
and, with the Schur complement $S := BA^{-1}B'$, $p$ is also uniquely determined by
\[
Sp = BA^{-1}f.
\]

**Lemma 2.2.** $(S_q)(q) = \sup_{0 \neq v \in \mathbb{V}} \frac{b(v, q)^2}{a(v, v)}$ and $\|q\|_\mathbb{V} := (S_q)(q)^{\frac{1}{2}} \approx \|q\|_{L_2(\Omega)}$ ($q \in \mathbb{P}$).

**Proof.** With $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{H^1(\Omega)}$ (or $\langle \cdot, \cdot \rangle = a(\cdot, \cdot)$), let $R : \mathbb{V} \to \mathbb{V}'$ be the mapping such that $g(v) = \langle v, Rg \rangle$ ($v \in \mathbb{V}$, $g \in \mathbb{V}'$). Writing $B' = RB'$, $\tilde{A} = RA$, we have
\[
\sup_{0 \neq v \in \mathbb{V}} \frac{b(v, q)^2}{a(v, v)} = \sup_{0 \neq v \in \mathbb{V}} \frac{(B'q)(v)^2}{(Av)(v)} = \sup_{0 \neq v \in \mathbb{V}} \frac{\langle v, B'q \rangle^2}{\langle v, Av \rangle} = \sup_{0 \neq v \in \mathbb{V}} \frac{\langle w, \tilde{A}^{-\frac{1}{2}}B'q \rangle^2}{\langle w, w \rangle}.
\]

The second statement follows from the ellipticity of $a$, the boundedness of $b$, and (2.2). \qed

For $g \in \mathbb{V}'$, we set $\|g\|_{\mathbb{V}'} := \sup_{0 \neq v \in \mathbb{V}} \frac{|g(v)|}{\|v\|_{\mathbb{V}'} \|v\|_\mathbb{V}}$. Equipped with norms $\|\cdot\|_\mathbb{V}$ and $\|\cdot\|_{\mathbb{V}'}$, respectively, $A : \mathbb{V} \to \mathbb{V}'$ is an isomorphism.

Functions $g \in L_2(\Omega)^d$ will be interpreted as functionals by means of $g(v) := \int_\Omega g \cdot v$. 

---

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
3. **Finite element approximation.** Given some fixed \( m \in \mathbb{N}_{>0} \), and partitions \( \tau \) and \( \sigma \) of \( \bar{\Omega} \) into essentially disjoint (closed) \( d \)-simplices, we will search approximations for \( u \) and \( p \) from the finite element spaces

\[
V_\tau := \mathbb{V} \cap \prod_{T \in \tau} P_m(T)^d \quad \text{and} \quad P_\sigma := \mathbb{P} \cap \prod_{T \in \sigma} P_{m-1}(T),
\]

respectively. For doing so, furthermore we will approximate the right-hand side \( f \) by functions from

\[
V_\tau^* := \prod_{T \in \tau} P_{m-1}(T)^d.
\]

At any moment in our algorithm we will have that \( \tau \supseteq \sigma \), meaning that \( \tau = \sigma \) or that it is a proper refinement of \( \sigma \). Note that \((V_\tau, P_\tau)\) is not an LBB stable pair.

Sometimes we will view \( \mathbb{V} \) and \( \mathbb{P} \) formally as finite element spaces corresponding to the infinitely fine partition \( \infty \), and denote them as \( \mathbb{V}_\infty \) and \( \mathbb{P}_\infty \), respectively.

**Remark 3.1.** The fact that the approximate pressure is a piecewise polynomial of degree not larger than \( m - 1 \) will be used only in the forthcoming Proposition 5.2. It is most likely that also there higher degree polynomials can be permitted at the expense of having a more complicated refinement rule for the velocity partitions (it will be needed to create more interior vertices; cf. Figure 5.1). On the other hand, at least for our analysis, it will be essential that \( P_\tau \supseteq \text{div} V_\tau \) (cf. Remark 6.1).

We will exclusively consider partitions that can be created by a certain recursive bisection procedure starting from some fixed conforming initial partition \( \tau_0 \). The procedure we consider is a generalization to any space dimension of the well-known newest vertex bisection rule in two space dimensions (cf. [Mit89]). Here we recall a few properties of this generalized newest vertex bisection method that are relevant for the current exposition, and we refer to [Ste06] for a complete description of the method.

The way of bisecting any simplex in any partition that can be created by the method is uniquely determined. It depends only on a local numbering of the vertices of the simplices in the initial partition \( \tau_0 \). As a consequence, any partition can be represented by a subtree of the infinite binary master tree \( T_* \) having as roots the simplices of \( \tau_0 \), and in which the parent–child relation corresponds to the unique bisection of the parent. The partitions that can be created are uniformly shape regular, dependent only on \( \tau_0 \).

For applying a posteriori error estimators, we will need the partitions \( \tau \) underlying the velocity approximation spaces \( V_\tau \) to be conforming. So in the following,

\[
\tau, \tau', \tilde{\tau}, \text{ etc., will always denote conforming partitions.}
\]

Bisecting one or more simplices in a conforming partition \( \tau \) generally results in a nonconforming partition \( \varrho \). Conformity has to be restored by (recursively) bisecting any simplex \( T \in \varrho \) that contains a vertex \( v \) of a \( T' \in \varrho \) that does not coincide with any vertex of \( T \) (such a \( v \) is called a hanging vertex). This process, called completion, results into the smallest conforming refinement of \( \varrho \).

Our adaptive method will be of the following form:

\[
\text{for } j := 1 \text{ to } M \\
\text{do create some, possibly nonconforming refinement } \varrho_j \text{ of } \tau_{j-1} \\
\text{complete } \varrho_j \text{ to its smallest conforming refinement } \tau_j \\
\text{endfor}
\]
As we will see, we will be able to bound \( \sum_{j=1}^{M} \# \varrho_j - \# \tau_{j-1} \). Because of the additional bisections made in the completion steps, however, generally \( \# \tau_M - \# \tau_0 \) will be larger. The following crucial result shows that these additional bisections inflate the total number of simplices by at most an absolute constant factor. It is valid with a proper local numbering of the vertices of the simplices in the initial partition \( \tau_0 \), which we will assume in the following. In two dimensions such a numbering exists for any conforming partition. In more than two dimensions, it always exists after some initial refinement that inflates the number of simplices by not more than a constant factor.

**Theorem 3.2** (see [Ste06], generalizes upon [BDD04, Theorem 2.4] for \( d = 2 \)).

\[
\# \tau_M - \# \tau_0 \lesssim \sum_{j=1}^{M} \# \varrho_j - \# \tau_{j-1},
\]

dependent only on \( \tau_0 \), and in particular thus independent of \( M \).

**Remark 3.3.** Note that this result in particular implies that any descendant \( \varrho \) of \( \tau_0 \) has a conforming refinement \( \tau \) with \( \# \tau \lesssim \# \varrho \), dependent only on \( \tau_0 \) and \( d \).

We end this section by introducing some more notation. The smallest partition that is a refinement of partitions \( \varrho_1 \) and \( \varrho_2 \), i.e., their smallest common refinement, will be denoted as \( \varrho_1 \cup \varrho_2 \). For a partition \( \tau \) (thus a conforming one), \( E_\tau \) will denote the set of all internal \((d-1)\)-dimensional hyperfaces in \( \tau \). For \( T \in \tau \), \( F_\tau(T) \) denotes the union of \( T \) and the set of its neighbors in \( \tau \), i.e., those simplices that share a \((d-1)\)-dimensional hyperface with \( T \).

### 4. Overview of the solution method.

For a partition \( \sigma_i \), we consider the Galerkin problem of finding \( p^{(i)} \in P_\sigma \), such that

\[
(Sp^{(i)})(q) = (BA^{-1}f)(q) \quad (q \in P_\sigma).
\]

With \( u^{(i)} := A^{-1}(f - B'p^{(i)}) \), this problem is equivalent to the semidiscrete problem of finding \( (u^{(i)}, p^{(i)}) \in \mathbb{V} \times P_\sigma \), such that

\[
(a(u^{(i)}, v) + b(v, p^{(i)}) + b(u^{(i)}, q) = f(v) \quad (v \in \mathbb{V}, q \in P_\sigma).
\]

Since this is just the Stokes problem in which the divergence-free constraint has been relaxed, we will refer to this problem as the **reduced Stokes problem**. The solution \( p^{(i)} \) is the best approximation to \( p \) from \( P_\sigma \), with respect to \( \| \cdot \|_P \), and by creating a suitable adaptively refined sequence of partitions \( \tau_0 = \sigma_0 \subset \sigma_1 \subset \cdots \), a convergent sequence \( (p^{(i)})_i \) towards \( p \) is obtained.

The reduced Stokes problem, however, cannot be solved exactly. With the Riesz operator \( R_\sigma : P_\sigma \to P_\sigma \) being defined by \( g(q) = \langle q, R_\sigma q \rangle_{L^2(\Omega)} \) (\( q \in P_\sigma \)), it can be written as \( R_\sigma, Sp^{(i)} = R_\sigma, BA^{-1}f \). Equipping \( P_\sigma \), with \( \langle \cdot, \cdot \rangle_{L^2(\Omega)} \), the operator \( R_\sigma, S : P_\sigma \to P_\sigma \) is symmetric, bounded, and positive definite, with a spectrum in \([\beta^2, 1]\) (cf. [NP04]). So to solve the reduced Stokes problem approximately, we may apply Richardson’s iteration:

\[
p_{j+1}^{(i)} = p_j^{(i)} - (R_\sigma, Sp_j^{(i)} - R_\sigma, BA^{-1}f) \quad (j = 0, 1, \ldots).
\]

With \( Q_\sigma : P \to P_\sigma \), denoting the \( L^2(\Omega) \)-orthogonal projector onto \( P_\sigma \), we have \( R_\sigma, B = -Q_\sigma, \text{div} \). Writing \( u_j^{(i)} := A^{-1}(f - B'p_j^{(i)}) \), we arrive at the equivalent formulation

\[
\begin{align*}
\langle a(u_j^{(i)}, v) = f(v) - b(v, p_j^{(i)}) \quad (v \in \mathbb{V}), \\
p_{j+1}^{(i)} = p_j^{(i)} - Q_\sigma, \text{div}u_j^{(i)},
\end{align*}
\]

**Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.**
known as the Uzawa iteration. The properties of $R_\sigma, S$ show that
\begin{equation}
\| p^{(i)} - p^{(i)}_{j+1} \|_{L^2(\Omega)} \leq [1 - \beta^2]\| p^{(i)} - p^{(i)}_{j} \|_{L^2(\Omega)}.
\end{equation}

Also the Uzawa iteration for solving the reduced Stokes problem cannot be performed exactly since it involves solving an elliptic problem posed over $V$. To solve this problem approximately, we will again consider Galerkin approximations: Given a partition $\tau_{j,k}^{(i)}$, let $u_{j,k}^{(i)} \in V_{\tau_{j,k}^{(i)}}$ be the solution of
\begin{equation}
a(u_{j,k}^{(i)}, v) = f(v) - b(v, p_{j}^{(i)}) \quad (v \in V_{\tau_{j,k}^{(i)}}).
\end{equation}

It is the best approximation to $u_{j}^{(i)}$ from $V_{\tau_{j,k}^{(i)}}$ with respect to $\| \cdot \|_V$, and by creating a suitable adaptively refined sequence of partitions $\sigma_i \subseteq \tau_{j,0}^{(i)} \subset \tau_{j,1}^{(i)} \subset \cdots$, a convergent sequence $(u_{j,k}^{(i)})_k$ towards $u_{j}^{(i)}$ is obtained. To guarantee that $u_{j,k+1}^{(i)}$ is indeed a better approximation than $u_{j,k}^{(i)}$, we will require that $f \in V'$ be sufficiently well approximated by a vector field from $V_{\tau_{j,k}^{(i)}}$. To implement the latter requirement, instead of working with $f$, we will replace it by suitable piecewise polynomial vector fields of degree $m - 1$, which should become increasingly more accurate when the iterations proceed.

Finally, instead of solving the finite-dimensional linear systems (4.4) exactly, in order to obtain a method of (quasi-)optimal computational complexity, we will use approximate solutions obtained by employing optimal iterative solvers.

In order to stop each of the nested loops in time, as well as for both Galerkin problems, to create adaptively refined partitions such that the corresponding approximations converge towards the solution, and such that these partitions have quasi-optimal cardinalities, we need a posteriori error estimators, which will be discussed in the next section. On the one hand, stopping a loop in time means that it should not stop too early in order to guarantee convergence of the overall process. On the other hand, the iterations should not proceed too long in order to control both the cardinalities of the partitions, which grow by the refinements, and the computational complexity.

In view of our solution method, we fix some notation. Throughout this paper, given some $r \in \mathbb{P}$, where we have in mind an approximation to $p$, and a partition $\tau$, $u^\tau_r \in V_\tau$ will denote the solution of the (discretized) elliptic problem
\begin{equation}
\begin{bmatrix}
a(u^\tau_r, v_\tau) = f(v_\tau) - b(v_\tau, r) \quad (v_\tau \in V_\tau).
\end{bmatrix}
\end{equation}

As a special case of (4.5), $u^\tau = u^\tau_\infty \in V$ thus denotes the solution of
\begin{equation}
a(u^\tau, v) = f(v) - b(v, r) \quad (v \in V).
\end{equation}

So the subscript $\tau$ in the notation $u^\tau_r$ refers to its membership of the finite element space $V_\tau$, whereas the superscript $r$ in the notation $u^\tau_r$ and $u^\tau$ refers to their dependence on the approximate pressure $r$ on the right-hand side. Note that $u = u^p$. Often we will consider (4.5) and (4.6) for $r = r_\sigma \in \mathbb{P}_\sigma$ for some $\sigma \subseteq \tau$ giving rise to the notation $u^\tau_{\sigma}^r$ and $u^\tau_{\sigma}$.

Given a partition $\sigma$, $(u^{p_{\sigma}}, p_{\sigma}) \in V \times \mathbb{P}_\sigma$ will denote the solution of the reduced Stokes problem
\begin{equation}
a(u^{p_{\sigma}}, v) + b(v, p_{\sigma}) + b(u^{p_{\sigma}}, q_{\sigma}) = f(v) \quad (v \in V, q_{\sigma} \in \mathbb{P}_\sigma).
\end{equation}

Note that $a(u^{p_{\sigma}}, v) = f(v) - b(v, p_{\sigma}) \quad (v \in V)$ so that the notation $u^{p_{\sigma}}$ is consistent with (4.6).
5. A posteriori error estimators.

5.1. A posteriori error estimator for the inner elliptic problem. For a given right-hand side \( f \) and partitions \( \tau \supseteq \sigma \) and \( r_\sigma \in \mathcal{P}_\sigma \), we study an a posteriori error estimator for \( \| u^{r_\sigma} - u^{r_\sigma}_\tau \|_V \). Since even the Galerkin problem (4.5) will be solved only inexactly, whenever possible we even estimate \( \| u^{r_\sigma} - w_\tau \|_V \) for some general \( w_\tau \in \mathcal{V}_\tau \).

For \( T \in \tau \), we set the local error indicator

\[
\eta_T(f, r_\sigma, w_\tau) := \text{diam}(T)^2 \| f - \nabla r_\sigma + \Delta w_\tau \|_{L_2(T)^d}^2 + \text{diam}(T) \| [r_\sigma \mathbf{n} - \nabla w_\tau \cdot \mathbf{n}]_{\partial T} \|_{L_2(\partial T)^d}^2,
\]

where \([ \cdot ]_{\partial T}\) denotes the jump of its argument over \( \partial T \) in the direction of \( \mathbf{n} \), being a unit vector normal to \( \partial T \). This jump is defined to be zero over \( \partial \Omega \). We set the elliptic error estimator

\[
\mathcal{E}^E(\tau, f, r_\sigma, w_\tau) := \left[ \sum_{T \in \tau} \eta_T(f, r_\sigma, w_\tau) \right]^{\frac{1}{2}}.
\]

Note that the definition of the error estimator requires \( f \in L_2(\Omega)^d \), which we therefore assume here.

In Proposition 5.1 below, for \( \tau' \supseteq \tau \) it is shown that the sum of the local error indicators corresponding to the simplices that were refined when creating \( \tau' \) from \( \tau \), or those that have nonempty intersection with such simplices, is an upper bound for \( \| u^{r_\sigma} - u^{r_\sigma}_\tau \|_V \). Substituting \( \tau' = \infty \), it generalizes the known upper bound for \( \| u^{r_\sigma} - u^{r_\sigma}_\tau \|_V \) [BMN02, Lemma 5.1, eq. (5.4)], [Ver96]. Since our generalization is not so difficult to derive, we refer to the extended preprint version [KS07] of this work for a proof.

**PROPOSITION 5.1.** Let \( \tau' \supseteq \tau \supseteq \sigma \) be partitions, and let \( r_\sigma \in \mathcal{P}_\sigma \), \( f \in L_2(\Omega)^d \), and

\[
\mathcal{F} = \mathcal{F}(\tau, \tau') := \{ T \in \tau : T \cap \tilde{T} \neq \emptyset \text{ for some } \tilde{T} \in \tau \text{ that has been refined in } \tau' \}.
\]

Then we have

\[
\| u^{r_\sigma}_\tau - u^{r_\sigma}_\tau \|_V \leq C_1 \left[ \sum_{T \in \mathcal{F}} \eta_T(f, r_\sigma, u^{r_\sigma}_\tau) \right]^{\frac{1}{2}}
\]

for some absolute constant \( C_1 > 0 \). Note that \( \# \mathcal{F} \leq \# \tau' - \# \tau \).

In particular, by taking \( \tau' = \infty \), we have \( \| u^{r_\sigma} - u^{r_\sigma}_\tau \|_V \leq C_1 \mathcal{E}^E(\tau, f, r_\sigma, u^{r_\sigma}_\tau) \).

Next we study whether the error estimator also provides a lower bound for the error \( \| u^{r_\sigma} - u^{r_\sigma}_\tau \|_V \) and, when \( \tau' \) is a sufficient refinement of \( \tau \), for \( \| u^{r_\sigma}_\tau - u^{r_\sigma}_\tau \|_V \). In order to derive such estimates, for the time being we further restrict the type of right-hand sides to piecewise polynomials of degree \( m - 1 \) with respect to \( \tau \). We will call \( \tau' \supseteq \tau \) a full refinement with respect to \( T \in \tau \), when

all \( \tilde{T} \in \mathcal{F}_\tau(T) \), as well as all \((d - 1)\)-dimensional hyperfaces of \( T \), contain a vertex of \( \tau' \) in their interiors.

See Figure 5.1 for an illustration of the case \( d = 2 \). The following proposition was shown in [BMN02, Lemma 5.3]. (Actually, a somewhat stronger condition on the
refinement was imposed there, but not used; a more general $f$ was considered at the expense of an additional “oscillation” term in the expression; and, finally, there our general $w_\tau \in V_\tau$ reads as $u_\tau^r$, whose additional property of being the solution of (4.5) was not used though.)

**Proposition 5.2.** Let $\tau \supseteq \sigma$ be partitions, let $r_\sigma \in P_\sigma$, and let us assume that $f \in V_\tau^\star$. Let $\tau' \supset \tau$ be a full refinement of $\tau$ with respect to $T \in \tau$. Then for any $w_\tau \in V_\tau$, we have

$$\eta_T(f, r_\sigma, w_\tau) \lesssim \sum_{T' \in \mathcal{F}_{\tau}(T)} |u_\tau^r - w_\tau|_{H^1(T')}^2.$$  

As a straightforward consequence we have the following.

**Corollary 5.3.** In the situation of Proposition 5.2, let $\tau' \supseteq \tau$ be a full refinement of $\tau$ with respect to all $T$ from some $F \subset \tau$. Then

$$c_2 \left[ \sum_{T \in F} \eta_T(f, r_\sigma, w_\tau) \right]^{\frac{1}{2}} \leq \| u_\tau^r - w_\tau \|_V$$

for some absolute constant $c_2 > 0$. In particular, we have

$$c_2 \mathcal{E}_E(\tau, f, r_\sigma, w_\tau) \leq \| u_\tau^r - w_\tau \|_V. \quad (5.1)$$

Finally in this subsection, we investigate the stability of the elliptic error estimator.

**Proposition 5.4.** Let $\tau \supseteq \sigma$ be partitions, and let $r_\sigma \in P_\sigma$, $f \in L_2(\Omega)^d$, and $v_\tau, w_\tau \in V_\tau$. Then

$$c_2 [\mathcal{E}_E(\tau, f, r_\sigma, v_\tau) - \mathcal{E}_E(\tau, f, r_\sigma, w_\tau)] \leq \| v_\tau - w_\tau \|_V.$$  

**Proof.** For $g \in L_2(\Omega)^d$, $q_\sigma \in P_\sigma$, by two applications of the triangle inequality in the form $\| \cdot \| - \| \cdot \|^2 \leq \| \cdot - \cdot \|^2$, first for vectors and then for functions, we have

$$|\mathcal{E}_E(\tau, f, r_\sigma, v_\tau) - \mathcal{E}_E(\tau, g, q_\sigma, w_\tau)| \leq \mathcal{E}_E(\tau, f - g, r_\sigma - q_\sigma, v_\tau - w_\tau).$$

By substituting $g = f$ and $q_\sigma = r_\sigma$, and by applying (5.1), the proof is completed. \qed
5.2. A posteriori error estimator for the (reduced) Stokes problem. For given \( f \in L_2(\Omega)^d \) and partitions \( \tau \supseteq \sigma \) and \( \varrho \), where we think of either \( \varrho = \infty \) (full Stokes) or \( \varrho = \sigma \) (reduced Stokes), \( r_\sigma \in P_\sigma \) being an approximation for \( p_\varrho \), we study an a posteriori error estimator for \( \| u^{p_\varrho} - u^{r_\sigma}_\tau \|_V + \| p_\varrho - r_\sigma \|_p \), or whenever possible, for this quantity with \( u^{r_\sigma}_\tau \) replaced by a general \( w_\tau \in V_\tau \). We set the estimator

\[
E^S(\varrho, \tau, f, r_\sigma, w_\tau) := E^E(\tau, f, r_\sigma, w_\tau) + \| Q_{\varrho} \text{div} w_\tau \|_{L_2(\Omega)}.
\]

**Proposition 5.5.** For partitions \( \tau \supseteq \sigma \) and \( \varrho, r_\sigma \in P_\sigma \), and \( f \in L_2(\Omega)^d \), we have

\[
\| u^{p_\varrho} - u^{r_\sigma}_\tau \|_V + \| p_\varrho - r_\sigma \|_p \leq C_3 E^S(\varrho, \tau, f, r_\sigma, u^{r_\sigma}_\tau)
\]

for some absolute constant \( C_3 > 0 \).

The proof given in [Ver96] (cf. [BMN02, Lemma 4.1]) for \( \varrho = \infty \) easily generalizes to \( \varrho \subseteq \infty \). More details can be found in [KS07].

The first statement of the following proposition was shown in [Ver96] (cf. [BMN02, Lemma 4.1]) for \( \varrho = \infty \), but the proof generalizes immediately to general partitions \( \varrho \). The second statement follows easily by using \( Q_{\varrho} \text{div} u^{p_\varrho} = 0 \).

**Proposition 5.6.** Let \( \tau \supseteq \sigma \) and \( \varrho \) be partitions, let \( r_\sigma \in P_\sigma \) and \( w_\tau \in V_\tau \), and let us assume that \( f \in \mathcal{V}_\tau \). Then

\[
\eta_T(f, r_\sigma, w_\tau) \leq \sum_{T \in F_\tau(T)} \left[ \| u^{p_\varrho} - w_\tau \|^2_{H^1(T)^d} + \| p_\varrho - r_\sigma \|^2_{L_2(T)^d} \right] \quad (T \in \tau)
\]

and

\[
c_4 E^S(\varrho, \tau, f, r_\sigma, w_\tau) \leq \| u^{p_\varrho} - w_\tau \|_V + \| p_\varrho - r_\sigma \|_p
\]

for some absolute constant \( c_4 > 0 \).

The last result in this subsection provides an a posteriori error estimator for the outer elliptic problem.

**Proposition 5.7.** For a partition \( \varrho \), and for \( r \in P \), we have

\[
c_6 \| Q_{\varrho} \text{div} u^r \|_{L_2(\Omega)} \leq \| p_\varrho - r \|_p \leq C_5 \| Q_{\varrho} \text{div} u^r \|_{L_2(\Omega)}
\]

for some absolute constants \( C_5, c_6 > 0 \).

**Proof.** Use \( \sup_{0 \neq \psi \in V} \frac{b(\psi, p_\varrho - r)}{\| \psi \|_V} = \sup_{0 \neq \psi \in V} \frac{a(u^{p_\varrho} - u^r, \psi)}{\| \psi \|_V} = \| u^{p_\varrho} - u^r \|_V \), and thus

\[
\beta \leq \frac{\| u^{p_\varrho} - u^r \|_V}{\| p_\varrho - r \|_p} \leq 1,
\]

and

\[
\| p_\varrho - r \|_p + \| u^{p_\varrho} - u^r \|_V \approx \sup_{0 \neq (\psi, q_\varrho) \in V \times P_\varrho} \frac{a(u^{p_\varrho} - u^r, \psi) + b(\psi, p_\varrho - r) + b(u^{p_\varrho} - u^r, q_\varrho)}{\| \psi \|_V + \| q_\varrho \|_p}
\]

\[
= \| Q_{\varrho} \text{div} u^r \|_{L_2(\Omega)}. \quad \Box
\]

Clearly, the evaluation of this estimator \( \| Q_{\varrho} \text{div} u^r \|_{L_2(\Omega)} \) is not feasible, and so later \( u^r \) will be replaced by an approximation.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
6. Adaptive refinements resulting in error reduction. For both elliptic problems $Sp = BA^{-1}f$ and $a(u', v) = f(v) - b(v, r)$ ($v \in V$), the latter for some given $r \in P$, we construct adaptive refinement routines based on the a posteriori error estimators. Given (approximate) Galerkin solutions from $\mathbb{P}_{\sigma}$ or $V_{\tau}$, respectively, they produce refinements $\tilde{\sigma} \supset \sigma$ or $\tilde{\tau} \supset \tau$ such that the Galerkin solutions with respect to these partitions have strictly smaller errors. Moreover, we will give bounds on the number of refined simplices which eventually will lead to the conclusion that our adaptive Stokes solver generates quasi-optimal partitions.

6.1. Adaptive pressure refinements. With $C_5, c_6$ being the constants from Proposition 5.7, for some absolute constants

\[
\frac{d}{1 - \left(c_6^2 \frac{C_5^2}{d} \right)^2}, \quad \theta \in \left(0, 1 - d \left(1 - \frac{1}{\theta^2} \right) \right), \quad D \geq 1,
\]

we assume that we have the following routine available. We think of its argument $w$ as being an approximation to $u_{\tau\sigma}$, where $r_{\sigma}$ is some approximation to $p$ from $\mathbb{P}_{\sigma}$.

\[
\text{REFpres}[\sigma, w] \rightarrow \tilde{\sigma}
\]

$\% \sigma$ is a partition and $w \in V$.

Select a partition $\tilde{\sigma} \supseteq \sigma$ with

\[
\|Q_{\tilde{\sigma}} \nabla w\|_{L^2(\Omega)} \geq \theta \|\nabla w\|_{L^2(\Omega)},
\]

such that

\[
\#\tilde{\sigma} - \#\sigma \leq D(\#\tilde{\sigma} - \#\sigma)
\]

for any $\tilde{\sigma} \supseteq \sigma$ with

\[
\|Q_{\tilde{\sigma}} \nabla w\|_{L^2(\Omega)} \geq \sqrt{1 - d \left(1 - \theta^2\right)} \|\nabla w\|_{L^2(\Omega)}.
\]

Remark 6.1. Eventually, we will make calls $\tilde{\sigma} := \text{REFpres}[\sigma, w]$ only for the argument $w$ from $V_{\tau}$ for some $\tau \supset \sigma$. Since then $\nabla w \in \mathbb{P}_{\tau}$, in those cases we may always assume that $\sigma \subseteq \tilde{\sigma} \subseteq \tau$. The fact that the partition underlying the pressure approximation is always contained in or equal to that underlying the velocity approximation will be essential for the forthcoming adaptive refinement routine for reducing the error in the inner elliptic problem.

The benefit of $\text{REFpres}$ can be seen from the following two lemmas.

Lemma 6.2. Let $\sigma$ be a partition, and let $r_{\sigma} \in \mathbb{P}_{\sigma}$. Then for $\tilde{\sigma} = \text{REFpres}[\sigma, u^{\tau\sigma}]$, we have

\[
\|p - p_{\tilde{\sigma}}\|_{P} \leq \left[1 - \frac{C_5^2}{c_6^2} \right]^\frac{1}{2} \|p - r_{\sigma}\|_{P}.
\]

Moreover,

\[
\#\tilde{\sigma} - \#\sigma \leq D(\#\tilde{\sigma} - \#\tau_0)
\]

for any partition $\tilde{\sigma}$ for which

\[
\inf_{q_{\sigma} \in \mathbb{P}_{\sigma}} \|p - q_{\sigma}\|_{P} \leq \left[1 - \frac{C_5^2}{c_6^2} \left(1 - d \left(1 - \theta^2\right)\right)\right]^\frac{1}{2} \|p - r_{\sigma}\|_{P}.
\]

(Note that (6.1) implies that $\frac{C_5^2}{c_6^2} \left(1 - d \left(1 - \theta^2\right)\right) < 1.$)
Proof. Recall that \( p_\sigma \) denotes the solution of a reduced Stokes problem, i.e., it is the best approximation to \( p \) from \( \mathbb{P}_\sigma \) with respect to \( \| \cdot \|_p \). Therefore, the first statement follows from

\[
\| p - r_\sigma \|_p^2 = \| p - p_\sigma \|_p^2 + \| p_\sigma - r_\sigma \|_p^2
\]

and \( \| p_\sigma - r_\sigma \|_p \geq c_6 \| Q_\sigma \text{div} u^r \|_{L^2(\Omega)} \geq c_6 \theta \| \text{div} u^r \|_{L^2(\Omega)} \geq \frac{c_6 \theta}{c_5} \| p - r_\sigma \|_p \) by Proposition 5.7.

For a \( \bar{\sigma} \) satisfying (6.4), let \( \bar{\sigma} = \sigma \cup \bar{\sigma} \). Then from \( \| p - p_\sigma \|_p \leq \inf_{q_\sigma \in \mathbb{P}_\sigma} \| p - q_\sigma \|_p \) and Proposition 5.7, with \( \lambda := \frac{c_5^2}{c_6} (1 - d(1 - \theta^2)) \) we have

\[
C_5^2 \| Q_\sigma \text{div} u^r \|_{L^2(\Omega)}^2 \geq \| p_\sigma - r_\sigma \|_p^2 = \| p - r_\sigma \|_p^2 - \| p - p_\sigma \|_p^2 \\
\geq \lambda \| p - r_\sigma \|_p^2 \geq \lambda \lambda c_6^2 \| \text{div} u^r \|_{L^2(\Omega)}^2.
\]

Noting that \( \frac{c_5^2}{c_6} = 1 - d(1 - \theta^2) \), by construction of \( \bar{\sigma} \), we conclude that

\[
\#\bar{\sigma} - \#\sigma \leq D[\#\bar{\sigma} - \#\sigma] \leq D[\#\bar{\sigma} - \#\sigma].
\]

Now we generalize Lemma 6.2 to the practical relevant situation in which only an approximation to \( u^r \) is available.

**Lemma 6.3.** Let \( \omega \in (0, \theta) \) be a constant, let \( \sigma \) be a partition, let \( r_\sigma \in \mathbb{P}_\sigma \), and let \( w \in \mathbb{V} \) with

\[
\| \text{div} u^r \|_{L^2(\Omega)} \leq \omega \| \text{div} w \|_{L^2(\Omega)}.
\]

Then for \( \bar{\sigma} = \text{REFpres}[\sigma, w] \), we have

\[
\| p - p_\sigma \|_p \leq \left[ 1 - \frac{c_5^2 (\theta - \omega)^2}{c_5^2 (1 + \omega)^2} \right]^\frac{1}{2} \| p - r_\sigma \|_p.
\]

Moreover, if \( \omega \) is sufficiently small such that \( \frac{\omega + \sqrt{1 - d(1 - \theta^2)}}{1 - \omega} < \frac{c_5}{c_6} \), then

\[
\#\bar{\sigma} - \#\sigma \leq D[\#\bar{\sigma} - \#\sigma]
\]

for any partition \( \bar{\sigma} \) for which, with \( \xi := \left[ 1 - \left( \frac{\omega + \sqrt{1 - d(1 - \theta^2)}}{1 - \omega} \right) \frac{c_5}{c_6} \right]^\frac{1}{2} \),

\[
\inf_{q_\sigma \in \mathbb{P}_\sigma} \| p - q_\sigma \|_p \leq \xi \| p - r_\sigma \|_p.
\] (6.5)

Proof. The proof is similar to that of Lemma 6.2. For the first part use that

\[
\| Q_\sigma \text{div} u^r \|_{L^2(\Omega)} \geq \| Q_\sigma \text{div} w \|_{L^2(\Omega)} - \omega \| \text{div} w \|_{L^2(\Omega)} \geq \frac{\omega - \omega}{1 + \omega} \| \text{div} u^r \|_{L^2(\Omega)},
\]

and, for the second part, with any \( \bar{\sigma} \) satisfying (6.5) and \( \bar{\sigma} = \sigma \cup \bar{\sigma} \), that

\[
C_5 \| Q_\sigma \text{div} w \|_{L^2(\Omega)} + \omega \| \text{div} w \|_{L^2(\Omega)} \geq C_5 \| Q_\sigma \text{div} u^r \|_{L^2(\Omega)} \geq \| p_\sigma - r_\sigma \|_p
\]

\[
= \left[ \| p - r_\sigma \|_p^2 - \| p - p_\sigma \|_p^2 \right]\frac{1}{2} \geq \sqrt{1 - \xi^2} \| p - r_\sigma \|_p \geq \xi \sqrt{1 - \xi^2} \| \text{div} u^r \|_{L^2(\Omega)}
\]

\[
\geq (1 - \omega)c_6 \sqrt{1 - \xi^2} \| \text{div} w \|_{L^2(\Omega)}.
\]

or, equivalently, \( \| Q_\sigma \text{div} w \|_{L^2(\Omega)} \geq \sqrt{1 - d(1 - \theta^2)} \| \text{div} w \|_{L^2(\Omega)} \).
As mentioned previously, we will make calls \( \hat{\sigma} := \text{REFpres}[\sigma, w] \) only for the argument \( w \) from \( V_\tau \) for some \( \tau \supseteq \sigma \), so that \( \text{div} w \in \mathbb{P}_r \). Obviously, if \( \text{REFpres} \) is implemented as the selection for some \( \theta \in (0, c_6/C_5) \) of the smallest partition \( \hat{\sigma} \supseteq \sigma \) such that (6.2) is valid, it satisfies its requirements with \( d = 1 = D \). Yet, at least a naive implementation of this algorithm would require computing \( \| Q_\sigma \text{div} w \|_{L_2(\Omega)} \) for all partitions \( \sigma \subsetneq \hat{\sigma} \subsetneq \tau \), which is prohibitively expensive.

Recalling that any partition corresponds to a subtree of the infinite binary master tree \( T_\bullet \), that is determined by the initial partition \( \tau_0 \), alternatively one may apply an adaptive tree approximation algorithm from [BD04]. Note that (6.2) and (6.3) are equivalent to
\[
\|(I - Q_\sigma)\text{div} w\|_{L_2(\Omega)} \leq \sqrt{1 - \theta^2} \|\text{div} w\|_{L_2(\Omega)} \text{ and } \|(I - Q_\sigma)\text{div} w\|_{L_2(\Omega)} \leq \sqrt{d(1 - \theta^2)} \|\text{div} w\|_{L_2(\Omega)},
\]
respectively. Prescribing a \( \theta \in (0, c_6/C_5) \), a call of the thresholding second algorithm starting with initial partition \( \sigma \), “error functional”
\( e(T) = \inf_{\tilde{\sigma} \in \mathcal{P}_{n-1}(T)} \|(\text{div} w)_{|T} - g\|_{L_2(T)} \) \((T \in T_\bullet)\), and tolerance \( (1 - \theta^2) \|\text{div} w\|_{L_2(\Omega)}^2 \) fulfills the requirements on \( \text{REFpres} \) for any given \( d < 1 \) and \( D := D(d) < \infty \), independent of \( \sigma \), and thus for \( d, \theta, D \) as in (6.1). This was demonstrated to the authors by Binev [Bin07], using the properties of \( T_\bullet \), being a binary tree and \( e(T) \geq e(T_1) + e(T_2) \) whenever \( T_1 \cup T_2 = T \). (The proof of [BD04, Corollary 5.4] shows the properties mentioned in \( \text{REFpres} \) for some specific constants \( d < 1 < D \) that might not fulfill the conditions from (6.1).) After precomputing the values \( e(T) \) for any simplex \( T \) from any partition \( \sigma \subsetneq \hat{\sigma} \subsetneq \tau \), which can be done in \( O(\#\tau) \) operations, this algorithm produces \( \hat{\sigma} \) as in (6.2) in \( O(\#\sigma) \) additional operations. With this implementation of \( \text{REFpres} \) we thus have that for \( \sigma \subsetneq \tau \) and \( w \in V_\tau \), the call \( \text{REFpres}[\sigma, w] \) takes \( O(\#\tau) \) operations.

6.2. Adaptive velocity refinements. For some fixed
\[
\zeta \in \left(0, \frac{c_2}{C_1}\right),
\]
we will make use of the following routine to determine a suitable adaptive refinement for an update of the velocity:

\( \text{REFvel}[\tau, g, r_\sigma, w_\tau] \to \tilde{\tau} \)

\( \% \ \tau \) is a partition, \( g \in L_2(\Omega)^d \), \( r_\sigma \in \mathbb{P}_r \) for some \( \sigma \subseteq \tau \), and \( w_\tau \in V_\tau \).
Select a set \( F \subset \tau \) with, up to some absolute factor, minimal cardinality such that

\[
\sum_{T \in F} \eta_T(g, r_\sigma, w_\tau) \geq \zeta^2 \mathcal{E}^E(\tau, g, r_\sigma, w_\tau)^2.
\]

Construct the smallest \( \tilde{\tau} \supset \tau \) which is a full refinement with respect to all \( T \in F \).

The next lemma will show the benefit of \( \text{REFvel} \). It applies under the (unrealistic) assumptions that \( f \in V_\tau^* \) and that the Galerkin problems are solved exactly. In Lemma 8.2, inexact Galerkin solutions will be allowed, and the given right-hand side \( f \in V^* \) will be replaced by approximations from \( V_\tau^* \) and \( V_\tau^* \), respectively.

Note that when \( f \in V_\tau^* \), the computation of all \( \eta_T(f, r_\tau, w_\tau) \) \((T \in \tau)\) can be done in \( O(\#\tau) \) operations. By doing an approximate sorting of the \( \eta_T(f, r_\tau, w_\tau) \) by their values, \( \text{REFvel}[\tau, f, \ldots] \) can be implemented in \( O(\#\tau) \) operations (cf. [Ste07]).

**Lemma 6.4.** Let \( \tau \supset \sigma \) be partitions, let \( f \in V_\tau^* \), and let \( r_\sigma \in \mathbb{P}_r \). Then for \( \tilde{\tau} = \text{REFvel}[\tau, f, r_\sigma, u_\tau^*] \), we have

\[
\|u^\tau - u_\tau^*\|_V \leq \left[1 - \frac{\zeta^2}{C_1}\right] \|u^\tau - u_\tau^e\|_V.
\]
Moreover, if \( \zeta < \frac{c_2^2}{c_2^2} \), and, for some absolute constant \( \vartheta > 0 \),
\[
\|u^{r_\sigma} - u^{f, r_\sigma}\|_V \geq \vartheta \|u - u^{r_\sigma}\|_V,
\]
then for the set of marked simplices \( F \) inside \( \text{REFvel} \), we have
\[
\# F \lesssim \# \bar{\tau} + \# \bar{\sigma} + \# \sigma
\]
for any partitions \( \bar{\tau} \) and \( \bar{\sigma} \) for which
\[
\inf_{v_{r_\sigma} \in V_{r_\sigma}} \|u - v_{r_\sigma}\|_V \leq \frac{1}{2} \left[ 1 - \frac{C_2^2 c_2^2}{c_2^2} \right]^\frac{1}{2} \|u^{r_\sigma} - u^{f, r_\sigma}\|_V,
\]
\[
\inf_{q_{\sigma} \in V_{\sigma}} \|p - q_{\sigma}\|_V \leq \frac{1}{2} \left[ 1 - \frac{C_2^2 c_2^2}{c_2^2} \right]^\frac{1}{2} \|u^{r_\sigma} - u^{f, r_\sigma}\|_V.
\]

**Remark 6.5.** Note that the bound on \( \# F \) in terms of \( u \) and \( p \) (via \( \bar{\tau} \) and \( \bar{\sigma} \)) and \( \sigma \) can be shown only when (the variational formulation of) \( -\Delta u^{r_\sigma} = f - \nabla r_\sigma \) is solved not too accurately, a restriction imposed by (6.7). This will be enforced by a timely stopping of the AFEM for solving this elliptic problem guided by a posteriori error estimators (cf. (7.20)). By assuming that \( u \) and \( p \) are in certain approximation classes, i.e., that these functions can be approximated with certain rates by finite element functions with respect to the best partitions, later we will derive quasi-optimal bounds for \( \# \bar{\tau} \) and \( \# \bar{\sigma} \), as well as for \( \# \sigma \) via Lemma 6.3, and so in the end on \( \# F \). Without imposing (6.7), we would arrive at a similar bound on \( \# F \) only when we would assume that all approximations \( u^{r_\sigma} \) to \( u \) corresponding to all approximate pressures \( r_\sigma \) that are created in the adaptive method are similarly easy to approximate as \( u \) itself, which is an unverifiable assumption.

**Proof.** The first statement follows from
\[
\|u^{r_\sigma} - u^{f, r_\sigma}\|_V^2 = \|u^{r_\sigma} - u^{f, r_\sigma}\|_V^2 + \|u^{f, r_\sigma} - u^{r_\sigma}\|_V^2,
\]
and \( \|u^{f, r_\sigma} - u^{r_\sigma}\|_V \geq c_2^2 \xi^E(\tau, f, r_\sigma, u^{r_\sigma}) \geq \frac{c_2^2}{c_2^2} \|u^{r_\sigma} - u^{f, r_\sigma}\|_V \) by Corollary 5.3 and Proposition 5.1.

Let \( \bar{\tau} \) be a partition for which
\[
\inf_{v_{r_\sigma} \in V_{r_\sigma}} \|u^{r_\sigma} - v_{r_\sigma}\|_V \leq \left[ 1 - \frac{C_2^2 c_2^2}{c_2^2} \right]^\frac{1}{2} \|u^{r_\sigma} - u^{f, r_\sigma}\|_V,
\]
and let \( \bar{\tau} = \tau \cup \bar{\tau} \). Then with \( \bar{F} = \overline{F}(\tau, \bar{\tau}) \) from Proposition 5.1, we have
\[
C_2^2 \sum_{T \in \bar{F}} \eta_T(f, r_\sigma, u^{r_\sigma}) \geq \|u^{r_\sigma} - u^{f, r_\sigma}\|_V^2 = \|u^{r_\sigma} - u^{f, r_\sigma}\|_V^2 - \|u^{r_\sigma} - u^{f, r_\sigma}\|_V^2
\]
\[
\geq \frac{c_2^2}{c_2^2} \|u^{r_\sigma} - u^{f, r_\sigma}\|_V^2 \geq C_2^2 \xi^E(\tau, f, r_\sigma, u^{r_\sigma})^2.
\]
By construction of \( \bar{F} \), we infer that
\[
\# F \lesssim \bar{F} \lesssim \# \bar{\tau} - \# \tau \leq \# \bar{\tau} - \# \tau_0.
\]

It remains to bound \( \# \bar{\tau} - \# \tau_0 \) for a \( \bar{\tau} \) as in (6.11). With \( \bar{\sigma} \) as in (6.10), we write \( u^{r_\sigma} = (u^{r_\sigma} - u^{f, r_\sigma}) + (u^{f, r_\sigma} - u) + u \) and approximate each of the three terms within
tolerance $\frac{1}{3} \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right]^\frac{1}{2}$ with finite element functions (the second one with zero). From (5.2), we have

$$\| \mathbf{u} - \mathbf{u}^{p*} \|_V \leq \| p - p_\sigma \|_V \leq \frac{1}{3} \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right]^\frac{1}{2} \| \mathbf{u}^{r*} - \mathbf{u}^{r*}_r \|_V.$$

The vector field $\mathbf{w} = \mathbf{u}^{r*} - \mathbf{u}^{p*}$ solves

$$a(\mathbf{w}, \mathbf{v}) = -b(\mathbf{v}, r_\sigma - p_\sigma) \quad (\mathbf{v} \in \mathcal{V}).$$

With $\hat{\sigma} = \sigma \cup \bar{\sigma}$, the error in its best approximation $\mathbf{w}_{\hat{\sigma}}$ from $\mathcal{V}_{\hat{\sigma}}$ can be bounded by

$$\| \mathbf{w} - \mathbf{w}_{\hat{\sigma}} \|_V \leq \| \mathbf{w} \|_V \leq \| \mathbf{u}^{r*} - \mathbf{u} \|_V + \| \mathbf{u}^{p*} - \mathbf{u} \|_V \leq \left( \varrho^{-1} + \frac{1}{3} \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right] \right) \| \mathbf{u}^{r*} - \mathbf{u}^{r*}_r \|_V.$$

With $\hat{\sigma}_0 = \hat{\sigma}$, for $k = 1, 2, \ldots$ let $\hat{\sigma}_k \supset \hat{\sigma}_{k-1}$ be the smallest partition that is a full refinement with respect to all $T \in \hat{\sigma}_{k-1}$. Then using the fact that $r_\sigma - p_\sigma \in \mathbb{P}_{\hat{\sigma}_0}$, as in the first part of this lemma, now with $\zeta = 1$, we have $\mathbf{w} - \mathbf{w}_{\hat{\sigma}_k} \|_V \leq \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right]^\frac{1}{2} \mathbf{w} - \mathbf{w}_{\hat{\sigma}} \|_V$, where $\mathbf{w}_{\hat{\sigma}_k}$ denotes the best approximation to $\mathbf{w}$ from $\mathcal{V}_{\hat{\sigma}_k}$. With $k$ being the smallest integer with $\left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right]^\frac{1}{2} (\varrho^{-1} + \frac{1}{3} \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right] \leq \frac{1}{3} \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right] \leq \frac{1}{3} \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right]$, we conclude that

$$\inf_{\mathbf{v}_r \in \mathcal{V}_r} \| \mathbf{u}^{r*} - (\mathbf{v}_r + \mathbf{w}_{\hat{\sigma}_k}) \|_V \leq \| \mathbf{u}^{r*} - \mathbf{u}^{p*} - \mathbf{w}_{\hat{\sigma}_k} \|_V + \| \mathbf{u}^{p*} - \mathbf{u} \|_V + \inf_{\mathbf{v}_r \in \mathcal{V}_r} \| \mathbf{u} - \mathbf{v}_r \|_V \leq \left[ 1 - \frac{C_1^2 c^2}{c_2^2} \right]^\frac{1}{2} \| \mathbf{u}^{r*} - \mathbf{u}^{r*}_r \|_V.$$

Since $\mathbf{v}_r + \mathbf{w}_{\hat{\sigma}_k} \in \mathcal{V}_r \cup \bar{\mathcal{V}}_{\hat{\sigma}_k}$, and $\#(\hat{\tau} \cup \hat{\sigma}_k) \leq \#\hat{\tau} + \#\bar{\sigma} + \#\sigma$ (dependent on $k$ and thus on $\varrho$), in view of (6.11) and (6.12) the proof is complete. \hfill \Box

Remark 6.6. If in (6.7) $\varrho > \left[ 1 - \frac{C_1 c^2}{c_2^2} \right]^{-\frac{1}{2}}$, then by a simplification of the above proof, instead of (6.8), one obtains that

$$\#E \leq \# \hat{\tau} - \#\tau_0$$

for any $\hat{\tau}$ with

$$\inf_{\mathbf{v}_r \in \mathcal{V}_r} \| \mathbf{u} - \mathbf{v}_r \|_V \leq \left( \left[ 1 - \frac{C_1 c^2}{c_2^2} \right]^{-\frac{1}{2}} - \varrho^{-1} \right) \| \mathbf{u}^{r*} - \mathbf{u}^{r*}_r \|_V,$$

which bound on $\#E$ is in particular independent of the pressure $p$. It is, however, not clear whether under the condition $\| \mathbf{u}^{r*} - \mathbf{u}^{r*}_r \|_V \geq \varrho \| \mathbf{u} - \mathbf{u}^{r*}_r \|_V$ for such $\varrho$ the inner elliptic problem is solved with sufficient accuracy to obtain a convergent inexact Uzawa algorithm for solving the reduced Stokes problem of finding $(\mathbf{u}^{p*}, p_\sigma)$. Knowing that $\# \sigma$ is controlled in our outermost loop that produces Galerkin approximations to $p$, Lemma 6.4 provides a way to avoid the condition that $\varrho > \left[ 1 - \frac{C_1 c^2}{c_2^2} \right]^{-\frac{1}{2}}$. This point is the exact reason why we did not succeed in proving quasi-optimality of the Uzawa iteration for solving the full Stokes problem, i.e., without our outermost loop. Indeed, with that method there is no separate control over the partitions that underlie the pressure approximations.

For \( s > 0 \), we define the approximation class

\[
\mathcal{A}_s^v = \{ v \in \mathcal{V} : |v|_{\mathcal{A}_s^v} := \sup_{\varepsilon > 0} \inf_{\tau : \varepsilon < \varepsilon} \frac{|\#\tau - \#\tau_0|^s}{\varepsilon} < \infty \}
\]

and equip it with norm \( \|v\|_{\mathcal{A}_s^v} := \|v\|_{\mathcal{V}} + |v|_{\mathcal{A}_s^v} \). So \( \mathcal{A}_s^v \) is the class of vector fields that can be approximated within any given tolerance \( \varepsilon > 0 \) by a \( \mathcal{V} \) for some partition \( \tau \) with \( \#\tau - \#\tau_0 \leq \varepsilon^{-1/s} |u|_{\mathcal{A}_s^v}^{1/s} \). Similarly, we define

\[
\mathcal{A}_s^q = \{ q \in \mathbb{P} : |q|_{\mathcal{A}_s^q} := \sup_{\varepsilon > 0} \inf_{\sigma : \varepsilon < \varepsilon} \frac{|\#\sigma - \#\tau_0|^s}{\varepsilon} < \infty \}
\]

and equip it with norm \( \|q\|_{\mathcal{A}_s^q} := \|q\|_p + |q|_{\mathcal{A}_s^q} \).

Because of the polynomial degrees of our approximations, only for \( s \leq m/d \) can membership of \( u \in \mathcal{A}_s^u \) or \( p \in \mathcal{A}_s^p \) be enforced by imposing suitable smoothness conditions on \( u \) or \( p \), respectively. These smoothness conditions, however, are much milder than requiring that \( u \in H^{1+m}(\Omega)^d \) or \( p \in H^m(\Omega) \), which would be needed when only uniformly refined partitions were considered. The approximation classes can be (nearly) characterized as certain Besov spaces (see [BDDP02] for details). In any case for \( d = 2 \), polygonal domains, and sufficiently smooth \( f \), it is known (see [Dah99]) that \( u \) and \( p \) have sufficient Besov smoothness so that they are in \( \mathcal{A}_s^u \) or \( \mathcal{A}_s^p \) for any \( s < m/d \), and thus also in the presence of reentrant corners.

The results derived in section 6.2 concerning adaptive velocity refinements were valid under the assumption that \( f \) was piecewise polynomial of degree \( m - 1 \) with respect to the current partition. In order to make our exposition not too complicated, in this section we will assume that \( f \) is piecewise polynomial of degree \( m - 1 \) with respect to any partition that we encounter, i.e., that is in \( \mathcal{V}_{\tau_0} \). In the next section, we will remove this restriction. Furthermore, in this section we assume that the arising finite-dimensional linear systems are solved exactly, i.e., we do not care about the computational cost. In the next section, by applying iterative solvers, we will show quasi-optimal computational complexity.

The following algorithm is an implementation of the solution method that was announced in section 4 with the simplifications mentioned above. Note that the do-loop in the algorithm actually consists of 3 nested loops over \( i \), \( j \), and \( k \). The loop over \( i \) concerns an adaptive method for solving \( p \) from the Schur complement equation. The loop over \( j \) concerns the Uzawa method for solving \( p^{\sigma} \) from the reduced Stokes problem. Finally, the loop over \( k \) concerns an adaptive finite element method for approximating the solution \( u^{p^{\sigma}} \) in \( \mathcal{V} \) of \( a(u^{p^{\sigma}}, v) = f(v) - b(v, p^{(i)}) \) \( (v \in \mathcal{V}) \). We have formulated these loops as one loop to deal efficiently with the complicated stopping criteria. For example, the innermost one stops when either \( E^{E_s}(\cdots) \leq \alpha E^{E_s}(\sigma_i, \ldots) \) or \( E^{S}(\sigma, \ldots) \leq \kappa E^{S}(\infty, \ldots) \) or \( E^{S}(\infty, \ldots) \leq \varepsilon \), i.e., when either

\[
\|u^{p^{(i)}} - u_{j,k}^{(i)}\|_V \leq C_1 \alpha c_4^{-1} \|p_{j}^{(i)} - p_{j}^{(i)}\| + \|u^{p_{j}^{(i)}} - u_{j,k}^{(i)}\|_V \quad \text{or}
\]

\[
\|p^{\sigma} - p_{j}^{(i)}\|_p + \|u^{p_{j}^{(i)}} - u_{j,k}^{(i)}\|_V \leq C_3 \kappa c_4^{-1} \|p - p_{j}^{(i)}\|_p + \|u - u_{j,k}^{(i)}\|_V \quad \text{or}
\]

\[
\|p - p_{j}^{(i)}\|_p + \|u - u_{j,k}^{(i)}\|_V \leq \varepsilon.
\]
\textbf{STOKESSOLVE}_0[f, \varepsilon] \rightarrow [\sigma_j^{(i)}, p_j^{(i)}, \tau_{j,k}, u_{j,k}]

% For this preliminary version of the adaptive solver it is assumed
% that \( f \in V_{\tau_0} \).
% Let the parameter \( \zeta \) from \text{REFvel} satisfy \( \zeta \in (0, \frac{\omega}{2}) \), and \( \theta \) from \text{REFpres}
% satisfy \( \theta \in (0, [1 - \frac{1-c^2d^2}{d}]^{\frac{1}{2}}) \). For some \( \omega \in (0, \theta) \) small enough such that
% \( \frac{\omega^+\sqrt{1-d(1-\sigma^2)}}{1-\alpha} < \frac{\omega}{C_{\zeta}} \), fix some sufficiently small constants \( \kappa, \alpha > 0 \) such that
% \( \kappa < 1, C_1 \frac{\kappa}{\sqrt{\kappa}} \leq \omega, \kappa C_1 < c_4, [1 - \frac{2\kappa C_1}{c_4-\kappa c_1}]^{-1}[1 - \frac{c^2(\theta-\omega)^2}{(1+\kappa)^2}]^{\frac{1}{2}} < \alpha C_1 < c_4, \)
% and \( 1 - \beta^2 + \frac{2\alpha C_1}{c_4-\alpha C_1} < 1. \)
% \( p_0^{(0)} := 0, \sigma_0 := \tau_0^{(0)} := \tau_0 \)
i := i := j := k := 0

do \ u_{j,k}^{(i)} := u_{j,k}^{(i)}% \ i.e., \ u_{j,k}^{(i)} \in V_{\tau_{j,k}}, a(u_{j,k}^{(i)}, v) = f(v) - b(v, p_j^{(i)}), \ (v \in V_{\tau_{j,k}})
% if \ C_2 \mathcal{E}(\infty, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \varepsilon \ then \ stop
% elseif \ C_2 \mathcal{E}(\sigma_j, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \kappa \mathcal{E}(\infty, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}^{(i)}) \ then
\quad \sigma_i+1 := \text{REFpres}[(\sigma_i, u_{j,k}^{(i)})]
\quad p_0^{(i+1)} := p_j^{(i)}, \tau_{0,0}^{(i+1)} := \tau_{j,k}^{(i)}
i++, j := k := 0
% elseif \ C_2 \mathcal{E}(\tau_{j,k}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \alpha \mathcal{E}(\sigma_j, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}^{(i)}) \ then
\quad p_j^{(i+1)} := p_j^{(i)} - Q_{\sigma_j} \text{div} u_{j,k}^{(i)}
\quad \tau_{j+1,0}^{(i)} := \tau_{j,k}^{(i)}
i++, k := 0
% else \ \tau_{j,k+1}^{(i)} := \text{REFvel}[(\tau_{j,k}, f, p_j^{(i)}, u_{j,k}^{(i)})]
k++
enddo

\textbf{Theorem 7.1.} (I) Let \( f \in V_{\tau_0} \). Then \([\sigma_j^{(i)}, p_j^{(i)}, \tau_{j,k}, u_{j,k}^{(i)}] := \text{STOKESSOLVE}_0 f, \varepsilon \) terminates, and \( \|u - u_{j,k}^{(i)}\|_V + \|p - p_j^{(i)}\|_V \leq \varepsilon \). (II) If, for some \( s > 0, p \in A_\varepsilon^s \), then \#\( \sigma_j^{(i)} - \#\tau_0 \leq \varepsilon^{-1/s} \|p\|_{A_\varepsilon^s}^{1/s} \), dependent only on \( \tau_0 \), and on \( s \) when it tends to 0 or infinity. If, in addition, for some \( \bar{s} > 0, u \in \mathcal{A}_{\bar{s}}^s \), then with \( \bar{s} = \text{min}(s, \bar{s}) \), \#\( \tau_{j,k}^{(i)} - \#\tau_0 \leq \varepsilon^{-1/s}(\|p\|_{A_\bar{s}^s}^{1/s} + \|u\|_{A_\bar{s}^s}^{1/s}) \), dependent only on \( \tau_0 \), and on \( \bar{s} \) when it tends to 0 or infinity.

\textbf{Remark 7.2.} In view of the assumptions, note that generally \#\( \sigma_j^{(i)} - \#\tau_0 \) is at most a constant multiple larger than this expression for the best partition \( \sigma_j^{(i)} \) giving rise to such an error in the pressure. Similarly, generally \#\( \tau_{j,k}^{(i)} - \#\tau_0 \) is at most a constant multiple larger than this expression for the best partition \( \tau_{j,k}^{(i)} \) on which \( p \) and \( u \) can be approximated by piecewise polynomials of degree \( m - 1 \), or continuous piecewise polynomials of degree \( m \) with errors less than or equal to \( \varepsilon \) in \( \|\cdot\|_p \) or \( \|\cdot\|_V \), respectively.

\textbf{Proof.} (I) Given \( i \) and \( j, k = k(i, j) \) will denote the maximum value attained by \( k \) for those \( i \) and \( j \). Given an \( i, j = j(i) \) will denote the maximum value attained by \( j \) for that \( i \), and \( k = k(i) := k(i, j(i)) \) is the the maximum value attained by \( k \) for that \( i \). Finally, \( \xi \) will denote the maximum value attained by \( i \).
At the moment \( C_3 \mathcal{E}^S(\infty, \tau_{j,k}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \varepsilon \) is passed, i.e., the loop over \( i \), and thus the algorithm terminates, then, by definition \((i, j, k) = (i, j(i), k(i))\), and we have \( \| u - u_{j,k}^{(i)} \| + \| p - p_j^{(i)} \| \leq \varepsilon \) by Proposition 5.5. Next, we show that indeed the loop over \( i \) terminates. In doing so, we first assume that each inner loop over \( j \), which will be shown afterwards.

The inequality \( \mathcal{E}^S(\sigma, \tau_{j,k}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \kappa \mathcal{E}^S(\infty, \tau_{j,k}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \) is equivalent to \( \mathcal{E}^E(\tau_{j,k}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq (1 - \kappa)^{-1}(\kappa \| \mathbf{div} u_{j,k}^{(i)} \|_{L^2(\Omega)} - \| Q_{\sigma} \mathbf{div} u_{j,k}^{(i)} \|_{L^2(\Omega)}) \). So at the moment that this test is passed, i.e., a loop over \( j \) terminates, then by definition \((j, k) = (j(i), k(i))\), and by Propositions 5.1, 5.6, and 5.5,

\[
\begin{align*}
(7.1) \quad & C_1^{-1} \| u^{(i)}_j - u_{j,k}^{(i)} \|_L^2 \leq \mathcal{E}^E(\tau_{j,k}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \frac{\kappa}{1 - \kappa} \| \mathbf{div} u_{j,k}^{(i)} \|_{L^2(\Omega)}, \\
(7.2) \quad & C_1^{-1} \| u^{(i)}_j - u_{j,k}^{(i)} \|_L^2 \leq \kappa c_4^{-1}(\| u - u_{j,k}^{(i)} \| + \| p - p_j^{(i)} \|), \\
(7.3) \quad & C_3^{-1} \| p_{\sigma_i} - p_{\sigma_{i+1}} \|_{L^2} \leq \kappa c_4^{-1}(\| u - u_{j,k}^{(i)} \| + \| p - p_j^{(i)} \|).
\end{align*}
\]

Furthermore, by (7.2), \( \kappa \leq c_4 C_1^{-1} \), and (5.2), we have

\[
\| u - u_{j,k}^{(i)} \|_L^2 \leq \| u - u^{(i)}_j \|_L^2 + \frac{\kappa c_4}{c_4 - \kappa c_4}(\| u - u^{(i)}_j \| + \| p - p_j^{(i)} \|)
\]

and so by (7.3),

\[
\| p_{\sigma_i} - p_{\sigma_{i+1}} \|_{L^2} \leq \frac{2\kappa c_4}{c_4 - \kappa c_4}(\| p - p_j^{(i)} \|).
\]

By (7.1), by \( \| \mathbf{div} \cdot \|_{L^2(\Omega)} \leq \| \cdot \|_{V} \) on \( V \), and by \( C_1 \frac{\kappa}{1 - \kappa} \leq \omega \), Lemma 6.3 shows that for the output \( \sigma_{i+1} \) of the routine REFpres called subsequently, it holds that

\[
(7.6) \quad \| p - p_{\sigma_{i+1}} \| \leq \left[ 1 - \frac{c_4^2(\theta - \omega)^2}{C_5^2(1 + \omega)^2} \right]^\frac{1}{2} \| p - p_j^{(i)} \|.
\]

Furthermore, since \( \omega + \sqrt{1 - (1 - \theta^2)} \) \( \frac{c_4}{c_5} \) and in view of the definition of \( \mathcal{A}_\sigma^p \), we have that

\[
(7.7) \quad \# \sigma_{i+1} - \# \sigma_i \leq \| p - p_{\sigma_{i+1}} \|^{-1/s} \| p \|_{\mathcal{A}_\sigma^p}^{1/s},
\]

which will be used in the second part of the proof. Concerning (7.6), it is not \( p_{\sigma_{i+1}} \) that is computed in the next loop over \( j \), but instead its approximation \( p_{(i+1)}^{(j)} \). Combining (7.5), with \( i \) reading as \( i + 1 \), together with (7.6) shows that with \( \rho_1 := \left[ 1 - \frac{2c_4^2}{c_4 - \kappa c_4} \right]^{-1/2} - \left[ 1 - \frac{c_4^2(\theta - \omega)^2}{C_5^2(1 + \omega)^2} \right]^\frac{1}{2} \leq 1,
\]

\[
(7.8) \quad \| p - p_{(i+1)}^{(j)} \| \leq \rho_1 \| p - p_j^{(i)} \|.
\]

Since \( c_4 \mathcal{E}^S(\infty, \tau_{j,k}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \leq \| u - u_{j,k}^{(i)} \| + \| p - p_j^{(i)} \|, \) together (7.8) and (7.4) show that indeed the loop over \( i \), and thus the algorithm, terminates once we have
shown that each loop over \( j \) does, which is shown next. In doing so, we first assume that each inner loop over \( k \) terminates, which will be shown afterwards.

At the moment \( \mathcal{E}^E(\tau_{j,k}, f, p_j(i), u_{j,k}) \leq \alpha \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j(i), u_{j,k}) \) is passed, i.e., a loop over \( k \) terminates, then by definition \( k = k(i, j) \), and

\[
(7.9) \quad \|u^{(i)} - u_{j,k}^{(i)}\|_V \leq \alpha C_1 c_4^{-1}(\|p_{\sigma_i} - p_j^{(i)}\|_P + \|u^{\sigma_i} - u_{j,k}^{(i)}\|_V).
\]

Estimating \( \|u^{\sigma_i} - u_{j,k}^{(i)}\|_V \leq \|u^{\sigma_i} - u^{(i)}\|_V + \|u^{(i)} - u_{j,k}^{(i)}\|_V \), and applying (7.9) and \( \|u^{\sigma_i} - u^{(i)}\|_V \leq \|p_{\sigma_i} - p_j^{(i)}\|_P \) ((5.2)), we obtain that

\[
(7.10) \quad \|u^{\sigma_i} - u_{j,k}^{(i)}\|_V \leq \frac{c_4 + \alpha C_1}{c_4 - \alpha C_1} \|p_{\sigma_i} - p_j^{(i)}\|_P
\]

and, by substituting this in (7.9), that

\[
(7.11) \quad \|u^{(i)} - u_{j,k}^{(i)}\|_V \leq \frac{2\alpha C_1}{c_4 - \alpha C_1} \|p_{\sigma_i} - p_j^{(i)}\|_P.
\]

For \( p_{j+1}^{(i)} := p_{j}^{(i)} - Q_{\sigma_i} \text{div} u_{j,k}^{(i)} \) that is evaluated subsequently, from \( \|\text{div} \|L_2(\Omega) \| \leq \|\cdot\|_V \), (7.11), and the statement (4.3) concerning convergence of the exact Uzawa iteration, with \( \rho_2 := 1 - \beta^2 + \frac{2\alpha C_1}{c_4 - \alpha C_1} < 1 \), we have

\[
(7.12) \quad \|p_{\sigma_i} - p_{j+1}^{(i)}\|_P \leq \rho_2 \|p_{\sigma_i} - p_j^{(i)}\|_P.
\]

Since \( c_4 \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \|u^{\sigma_i} - u_{j,k}^{(i)}\|_V + \|p_{\sigma_i} - p_j^{(i)}\|_P \), together (7.11) and (7.12) show that indeed each loop over \( j \) terminates by either \( \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \kappa \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) or \( C_3 \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \varepsilon \) once we have shown that each loop over \( k \) terminates, which is shown next.

With \( \rho_3 := \frac{\rho_2}{\kappa} < 1 \), Lemma 6.4 shows that

\[
(7.13) \quad \|u_{j,k+1}^{(i)} - u_{j,k}^{(i)}\|_V \leq \rho_3 \|u_j^{(i)} - u_{j,k}^{(i)}\|_V.
\]

Since \( c_2 \mathcal{E}^E(\tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \|u_{j,k}^{(i)} - u_{j,k}^{(i)}\|_V \), from (7.13) we conclude that indeed each loop over \( k \) terminates by either \( \mathcal{E}^E(\tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \alpha \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) or \( \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \kappa \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) or \( C_3 \mathcal{E}^S(\sigma, \tau_{j,k}, f, p_j^{(i)}, u_{j,k}) \) \( \leq \varepsilon \).

With this, part (I) of the theorem is proved.

Before starting with the second part, we collect some estimates for \( \|p - p_{j0}^{(i)}\|_P \), \( \|p_{\sigma_i} - p_j^{(i)}\|_P \), and \( \|u_{j,k}^{(i)} - u_{j,0}^{(i)}\|_V \), i.e., the initial values for the recursions (7.8), (7.12), and (7.13) over \( i, j, \) and \( k \), respectively.

From (7.5) we infer that

\[
(7.14) \quad \|p - p_{j0}^{(i)}\|_P \leq \left[1 - \frac{2\alpha C_1}{c_4 - \alpha C_1}\right]^{-1}\|p - p_0\|_P \leq \left[1 - \frac{2\alpha C_1}{c_4 - \alpha C_1}\right]^{-1}\|p\|_P.
\]

For \( j = 0 \) and \( i = 0 \), we have that

\[
(7.15) \quad \|p_{\sigma_i} - p_{j0}^{(i)}\|_P \leq \|p_{\sigma_i} - p\|_P + \|p\|_P \leq 2\|p\|_P
\]
and, for \( j = 0 \) and \( i > 0 \), that

\[
\|p_{\sigma} - p_0^{(i)}\|_p = \|p_{\sigma} - p_{2(i-1)}^{(i-1)}\|_p \leq \|p_{\sigma} - p\|_p + \|p - p_{2(i-1)}^{(i-1)}\|_p
\]

(7.16)

by (7.6).

For \( k = j = i = 0 \), we have

\[
\|u^{p_0^{(0)}} - u_0^{(0)}\|_V \leq \|u^{p_0^{(0)}}\|_V = \|f\|_V.
\]

(7.17)

For \( k = j = 0 \) and \( i > 0 \), we have \( p_0^{(i)} = p_{2(i-1)}^{(i-1)} \) and \( \tau_{0,0}^{(i)} = \tau_{2(i-1),k(i-1)}^{(i-1)} \), and so

\[
\|u^{p_0^{(i)}} - u_{0,0}^{(i)}\|_V = \|u^{p_{2(i-1)}^{(i-1)}} - u_{2(i-1),k(i-1)}^{(i-1)}\|_V \leq \frac{2C_1\kappa}{4 - \kappa C_1} \|p - p_{2(i-1)}^{(i-1)}\|_p
\]

by (7.2) and (7.4). For \( k = 0 \) and \( j > 0 \), we have

\[
\|u^{p_0^{(i)}} - u_{j,0}^{(i)}\|_V \leq \|u^{p_0^{(i)}} - u_{j-1,k(j-1)}^{(i)}\|_V
\]

\[
\leq \|u^{p_0^{(i)}} - u_{j-1,k(j-1)}^{(i)}\|_V + \|u_{j-1,k(j-1)}^{(i)} - u_{j-1,k(j-1)}^{(i)}\|_V.
\]

Now using that

\[
\|u^{p_0^{(i)}} - u_{j-1,k(j-1)}^{(i)}\|_V \leq \|p_{j}^{(i)} - p_{j-1}^{(i)}\|_p \leq \|u^{\sigma} - u_{j-1,k(j-1)}^{(i)}\|_V
\]

\[
\leq \|u^{\sigma} - u_{j-1,k(j-1)}^{(i)}\|_V + \|u_{j-1,k(j-1)}^{(i)} - u_{j-1,k(j-1)}^{(i)}\|_V
\]

by (5.2) and

\[
\|u_{j-1,k(j-1)}^{(i)} - u_{j-1,k(j-1)}^{(i)}\|_V \leq \frac{2\alpha C_1}{4 - \kappa C_1} \|p_{\sigma} - p_{j-1}^{(i)}\|_p
\]

by (7.11), we find that

\[
\|u^{p_0^{(i)}} - u_{j,0}^{(i)}\|_V \leq (1 + \frac{4\alpha C_1}{4 - \kappa C_1}) \|p_{\sigma} - p_{j-1}^{(i)}\|_p.
\]

(7.19)

(II) At the moment of a call \( \sigma_{i+1} := \text{REFpres}[\sigma_i, u_{j,k}^{(i)}] \), we have

\[
\|p - p_{\sigma}^{(i)}\|_p \geq c_4C^\epsilon(\infty, \tau_{\sigma}^{(i)}, \hat{f}, p_{\sigma}^{(i)}, u_{\sigma}^{(i)}) > c_4C_3^{-1}\epsilon,
\]

so that in view of (7.7) and (7.8), we have

\[
\#\sigma_{\perp} - \#\tau_0 \leq \sum_{i=0}^{j} \#\sigma_{\perp} - \#\sigma_{i-1} \leq \epsilon^{-1/s} |p_{\sigma_{\perp}}^{1/s}.
\]

At the moment of a call \( \tau_{j,k+1}^{(i)} := \text{REFvel}[\tau_{j,k}^{(i)}, f, p_{j}^{(i)}, u_{j,k}^{(i)}] \), we have

\[
\|u_{j,k}^{(i)} - u_{j,k}^{(i)}\|_V > c_2\epsilon^{\min(\tau_{j,k}^{(i)}, f, p_{j}^{(i)}, u_{j,k}^{(i)})} > c_2\alpha C^\epsilon(\sigma_{\perp}, \tau_{\sigma}^{(i)}, \hat{f}, p_{\sigma}^{(i)}, u_{\sigma}^{(i)})
\]

\[
> c_2\alpha C^\epsilon(\infty, \tau_{\sigma}^{(i)}, \hat{f}, p_{\sigma}^{(i)}, u_{\sigma}^{(i)}) \geq c_2\alpha C_3^{-1}\epsilon - \|u - u_{j,k}^{(i)}\|_V,
\]

(7.20)
where for the first time in this proof we use the fact that the innermost iteration is stopped in time. In view of the definitions of $\mathcal{A}_s^\tau$ and $\mathcal{A}_V^\tau$, Lemma 6.4 now shows that the set of marked simplices $E_{i,j,k} = E$ inside $\text{REFel}^{(i)}_{j,k}f, p_j^{(i)}, u_{j,k}^{(i)}$ satisfies

\begin{equation}
#E_{i,j,k} \leq \|u^{(i)}_0 - u_{j,k}^{(i)}\|_V^{-1/\delta}(\|u\|_{\mathcal{A}_s^\tau}^{1/\delta} + \|p\|_{\mathcal{A}_V^\tau}^{1/\delta}) + \#\sigma_i - \#\tau_0 + \#\tau_0.
\end{equation}

From $\|u^{(0)} - u_{j,k}^{(0)}\|_V \leq \|u^{(0)}\|_V \leq \|u\|_V + \|p\|_p$, we have

$$#\tau_0 \leq 1 \leq \|u^{(0)}_0 - u_{j,k}^{(i)}\|_V^{-1/\delta}(\|u\|_{\mathcal{A}_s^\tau}^{1/\delta} + \|p\|_{\mathcal{A}_V^\tau}^{1/\delta}).$$

For $i > 0$, we have $\#\sigma_i - \#\tau_0 \leq \|p - p_j^{(i)}\|_p^{1/\delta}$ by (7.7). By (7.18), (7.16), and (7.19), and the decrease of $\|p_{\sigma_i} - p_j^{(i)}\|_p$ and $\|u^{(j)}_0 - u_{j,k}^{(i)}\|_V$ as a function of $j$ or $k$, respectively, we have

$$\|p - p_j^{(i)}\|_p \geq \|u^{(j)}_0 - u_{j,k}^{(i)}\|_V$$

uniformly in $i, j, k$. We conclude that

$$#E_{i,j,k} \leq \|u^{(i)}_0 - u_{j,k}^{(i)}\|_V^{-1/\delta}(\|u\|_{\mathcal{A}_s^\tau}^{1/\delta} + \|p\|_{\mathcal{A}_V^\tau}^{1/\delta}),$$

and so, by Theorem 3.2 and (7.13),

$$#\tau_j^{(i)} \leq \|u^{(i)}_0 - u_{j,k}^{(i)}\|_V^{-1/\delta}(\|u\|_{\mathcal{A}_s^\tau}^{1/\delta} + \|p\|_{\mathcal{A}_V^\tau}^{1/\delta}) \sum_{i=0}^j \sum_{j=0}^{i} \sum_{k=0}^{j} \|u^{(i)}_j - u_{j,k}^{(i)}\|_V^{-1/\delta}$$

$$\leq \|u\|_{\mathcal{A}_s^\tau}^{1/\delta} + \|p\|_{\mathcal{A}_V^\tau}^{1/\delta}) \sum_{i=0}^j \sum_{j=0}^{i} \|u^{(i)}_j - u_{j,k}^{(i)}\|_V^{-1/\delta}.$$

We will bound this expression by using the fact that each of the three nested iterations is stopped in time.

For any $i$ and $j$, by the definition of $\mathcal{E}_i$, we have

$$\mathcal{E}_i^{E}(\tau_{i,j}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \geq \alpha \mathcal{E}_{\sigma_i}^{S}(\tau_{i,j}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)})$$

$$\geq \alpha \kappa \mathcal{E}_{\sigma_i}^{S}(\tau_{i,j}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) > \alpha \kappa C_3^{-1} \varepsilon,$$

or

$$\|u^{(i)}_j - u_{j}^{(i)}\|_V \geq \|u^{(i)}_j - u_{j}^{(i)}\|_V + \|p_{\sigma_i} - p_j^{(i)}\|_p \geq \|u - u_{j}^{(i)}\|_V + \|p - p_j^{(i)}\|_p \geq \varepsilon.$$

Similarly, for any $i$, by the definition of $\mathcal{J}$, we have

$$\mathcal{E}_{\sigma_i}^{S}(\tau_{i,j}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) \geq \kappa \mathcal{E}_{\sigma_i}^{S}(\tau_{i,j}^{(i)}, f, p_j^{(i)}, u_{j,k}^{(i)}) > \kappa C_3^{-1} \varepsilon,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
or
\[ \|p_{\sigma_i} - p_{L_{1-1}}^{(i)}\|_\varepsilon \approx \|u^{e_i} - u_{L_{1-1,\varepsilon}}^{(i)}\|_\varepsilon + \|p_{\sigma_i} - p_{L_{1-1}}^{(i)}\|_\varepsilon \geq \|u - u_{L_{1-1,\varepsilon}}^{(i)}\|_\varepsilon + \|p - p_{L_{1-1}}^{(i)}\|_\varepsilon \varepsilon, \]
where for \( \approx \) we used (7.10). Finally, by the definition of \( \Delta \), we have
\[ E^S(\infty, \tau_{L_{1-1}}, f, p_{L_{1-1}}, u_{L_{1-1,\varepsilon}^{(i)}}) > C_3^{-1} \varepsilon, \]
or
\[ \|p - p_{L_{1-1}}^{(i)}\|_\varepsilon \approx \|u - u_{L_{1-1,\varepsilon}^{(i)}}\|_\varepsilon + \|p - p_{L_{1-1}}^{(i)}\|_\varepsilon \varepsilon, \]
where for \( \approx \) we used (7.4). By using in addition (7.12) and (7.8), we find that
\[
\sum_{i=0}^{j} \sum_{j=0}^{i} \|u^{e_i} - u_{L_{1-1,\varepsilon}^{(i)}}\|_\varepsilon^{-1/s} \\
\leq \sum_{i=0}^{j} \sum_{j=0}^{i} \|p_{\sigma_i} - p_{L_{1-1}}^{(i)}\|_\varepsilon^{-1/s} + \sum_{i=0}^{j} \|p - p_{L_{1-1}}^{(i)}\|_\varepsilon^{-1/s} + \varepsilon^{-1/s} \\
\leq \sum_{i=0}^{j} \|p_{\sigma_i} - p_{L_{1-1}}^{(i)}\|_\varepsilon^{-1/s} + \|p - p_{L_{1-1}}^{(i-1)}\|_\varepsilon^{-1/s} + \varepsilon^{-1/s} \\
\leq \sum_{i=0}^{j-1} \|p - p_{L_{1-1}}^{(i)}\|_\varepsilon^{-1/s} + \|p_{\sigma_i} - p_{L_{1-1}}^{(i)}\|_\varepsilon^{-1/s} + \varepsilon^{-1/s} \\
\leq \sum_{i=0}^{j-1} \|p - p_{L_{1-1}}^{(i)}\|_\varepsilon^{-1/s} + \varepsilon^{-1/s} \leq \varepsilon^{-1/s},
\]
which completes the proof of the theorem.

8. A practical adaptive method for the Stokes problem. In the previous section, we assumed that with respect to any current partition \( \tau' \), the right-hand side \( f \) was in \( V^*_{\tau'} \). This led to the severe limitation of \( f \) being in \( V^*_{\tau_0} \). In the practical method, given a general right-hand side \( f \) and any current partition \( \tau' \), \( f \) is going to be replaced by a sufficiently accurate approximation from \( V^*_{\tau} \) or, when such an approximation does not exist, by an approximation from \( V^*_{\tau'} \) for some \( \tau' \supset \tau \). We assume the availability of the following routine, \( \text{RHS} \).

\( \text{RHS}[\tau, \eta] \rightarrow [\tau', f_{\eta}] \)

\% \( \eta > 0 \). The output consists of an \( f_{\eta} \in V^*_{\tau'} \), where \( \tau' = \tau \), or, if necessary, \( \% \tau' \supset \tau \), such that \( \|f - f_{\eta}\|_\varepsilon \leq \eta \).

Assuming that \( p \in A_s^\varepsilon \) and \( u \in A_s^\varepsilon \) for some \( s, \varepsilon > 0 \), the cost of approximating the right-hand side \( f \) using \( \text{RHS} \) will generally not dominate the other cost of our practical adaptive method only then when there is some constant \( c_f \) such that, with \( \varepsilon = \max(s, \varepsilon) \), for any \( \eta > 0 \) and any partition \( \tau \), for \( [\tau', \cdot] := \text{RHS}[\tau, \eta] \), it holds that
\[
\#\tau' - \#\tau \leq c_f^{1/\varepsilon} \eta^{-1/\varepsilon},
\]
and the number of arithmetic operations required by the call is \( O(\#\tau') \). We will call such an RHS \( \tilde{s} \)-optimal with constant \( c_\tilde{s} \). Obviously, given \( \tilde{s} \), such a routine can exist only when \( f \in \mathcal{A}_\tilde{V}_\tau \), defined by

\[
\mathcal{A}_\tilde{V}_\tau = \{ g \in \mathcal{V}' : \sup_{\tau' \in \Omega} \inf_{\tau \in \Omega} \inf_{g \in \mathcal{V} : \| g - g \|_{\mathcal{V}} \leq \varepsilon} \| \#\tau - \#\tau_0 \|_\tilde{s} < \infty \}.
\]

Knowing that \( f \in \mathcal{A}_\tilde{V}_\tau \) is a different thing than knowing how to construct suitable approximations. If \( \tilde{s} \in [1/d,(m+1)/d] \) and \( f \in H^{\tilde{s}d-1}(\Omega)^d \), then \( f \in \mathcal{A}_\tilde{V}_\tau \), and \( f, \tau \), constructed as the best approximation from \( \mathcal{V}_\tau \), to \( f \) with respect to \( L_2(\Omega)^d \) using (the smallest common refinement of the input partition and) uniform refined partitions \( \tau' \) are known to converge at the required rate. For general \( f \in \mathcal{A}_\tilde{V}_\tau \), however, a realization of a suitable routine RHS has to depend on \( f \) at hand.

The maximal values of \( s, \tilde{s} \), and \( \tilde{s} \) for which membership \( p \in \mathcal{A}_s \), \( u \in \mathcal{A}_\tilde{V}_\tau \), and \( f \in \mathcal{A}_\tilde{V}_\tau \), can be expected are \( m/d, m/d, \) and \( (m+1)/d \), respectively. Therefore, usually one can expect that asymptotically calls of RHS will not give rise to refinements.

Remark 8.1. Instead of replacing \( f \) by a piecewise polynomial approximation both for setting up the Galerkin system for the inner elliptic problem and for computing the a posteriori error estimators, one may work with the original \( f \). In that case, generally one commits quadrature errors, which one may view as a consequence of an implicit replacement of \( f \) by a piecewise polynomial approximation. This approach, however, is restricted to \( f \in L_2(\Omega)^d \), since otherwise the error estimators are not defined.

Since we are going to consider modified right-hand sides, we introduce the following notation: Given \( r \in \mathbb{P} \), \( g \in \mathcal{V}' \), and a partition \( \tau \), \( u_{\tau}^g \in \mathcal{V}_\tau \) will denote the solution of the Galerkin problem

\[
\begin{align*}
o(u_{\tau}^g,v_{\tau}) &= g(v_{\tau}) - b(v_{\tau},r) & (v_{\tau} \in \mathcal{V}_\tau).
\end{align*}
\]

(8.1)

So in view of the notation introduced in (4.5), we have \( u_{\tau}^{r,f} = u_{\tau}^g \). Note that \( \| u_{\tau}^{r,f} - u_{\tau}^g \|_{\mathcal{V}} \leq \| f - g \|_{\mathcal{V}} \).

In the previous section, we solved the arising Galerkin systems exactly. When aiming at a method of optimal computational complexity we cannot afford this. Therefore, we assume that we have an iterative solver of optimal type available:

\[
\text{GALSOLVE}[\tau,f_r,r_\tau,w_\tau^{(0)},\eta] \rightarrow w_\tau
\]

\% \( \tau \) is a partition, \( f_r \in \mathcal{V}_\tau \), \( r_\tau \in \mathbb{P} \), for some \( \sigma \supseteq \tau \), \( w_\tau^{(0)} \in \mathcal{V}_\tau \), the latter being an \% initial approximation for an iterative solver, and \( \eta > 0 \).

\% The output \( w_\tau \in \mathcal{V}_\tau \) satisfies

\[
\| u_{\tau}^{r,f_r} - w_\tau \|_{\mathcal{V}} \leq \eta.
\]

\% The call requires \( \lesssim \max\{1, \log(\eta^{-1}\| u_{\tau}^{r,f_r} - w_\tau^{(0)} \|_{\mathcal{V}}) \} \#\tau \) arithmetic operations.

Additive or multiplicative multigrid methods with local smoothing are known to be of this type [WC06].

In Lemma 6.4, the adaptive refinement routine \texttt{REFvel} was analyzed. For a call \( \tau = \texttt{REFvel}[\tau,f,r_\sigma,u_{\tau}^g] \), it was shown that \( \| u^{r_\sigma} - u_{\tau}^g \|_{\mathcal{V}} \) was strictly smaller than \( \| u^{r_\sigma} - u_{\tau}^g \|_{\mathcal{V}} \), and for \( u_{\tau}^g \) not too close to \( u^{r_\sigma} \), a bound for the number of marked simplices inside the call was given. In the following lemma, these results are extended to the situation in which we work with approximations to \( f \), and where we have only inexact Galerkin solutions available.
Lemma 8.2. There exist constants $\chi_1 = \chi_1(\zeta, C_1, c_2) > 0$ and $\lambda = \lambda(\chi_1, C_1, c_2) \in (0, \frac{1}{2} \left[1 - \frac{C_2^1 C_2^2}{C_2^3} \right]^\frac{1}{2})$ such that if for $f \in \mathcal{V}'$, partitions $\tau \supseteq \sigma, r_\sigma \in \mathcal{P}_\sigma, f_\tau \in \mathcal{V}_{\tau}^*, w_\tau \in \mathcal{V}_{\tau}$,

\begin{equation}
|f - f_\tau|_{\mathcal{V}'} + \|u_{\tau}^{r_\tau} - f_\tau - w_\tau\|_{\mathcal{V}} \leq \chi_1 \mathcal{E}(\tau, f_\tau, r_\sigma, w_\tau)
\end{equation}

and, for some absolute constant $\theta > 0$,

\[\|u_{\tau}^{r_\tau} - w_\tau\|_{\mathcal{V}} \geq \theta \|u - u_{\tau}^{r_\tau}\|_{\mathcal{V}},\]

then the set of marked simplices $\hat{E}$ inside the call $\hat{\tau} := \text{REFvel}([\tau, f_\tau, r_\sigma, w_\tau])$ satisfies

\[\# \hat{E} \leq \# \bar{\tau} + \# \bar{\sigma} + \# \sigma\]

for any partitions $\bar{\tau}$ and $\bar{\sigma}$ for which

\begin{equation}
\inf_{v_\tau, \sigma \in \mathcal{V}_\tau} \|u - v_\tau\|_{\mathcal{V}} \leq \lambda \|u_{\tau}^{r_\tau} - w_\tau\|_{\mathcal{V}}, \quad \inf_{q_\sigma \in \mathcal{P}_\sigma} \|p - q_\sigma\|_{\mathcal{V}} \leq \lambda \|u_{\tau}^{r_\tau} - w_\tau\|_{\mathcal{V}}.
\end{equation}

Furthermore, given a $\mu \in \left([1 - \frac{C_2^1 C_2^2}{C_2^3}]^\frac{1}{2}, 1\right)$,

there exists an $\chi_2 = \chi_2(\mu, \zeta, C_1, c_2) > 0$, such that if (5.2) is valid with $\chi_1$ reading as $\chi_2$, and for $\tau' \supseteq \bar{\tau}$, $f_\tau' \in \mathcal{V}'$, and $w_{\tau'} \in \mathcal{V}_{\tau'}$,

\[|f - f_\tau'|_{\mathcal{V}'} + \|u_{\tau'}^{r_\tau'} - f_\tau'|_{\mathcal{V}} \leq \chi_2 \mathcal{E}(\tau, f_\tau, r_\sigma, w_\tau),\]

then

\[\|u_{\tau}^{r_\tau} - w_{\tau'}\|_{\mathcal{V}} \leq \mu \|u_{\tau}^{r_\tau} - w_\tau\|_{\mathcal{V}}.
\]

Proof. Following along the lines of [Ste07, proof of Lemma 6.1], for suitable constants $\chi_1$ and $\lambda$, one can show that $\# \hat{E} \leq \# \bar{\tau} - \# r_0$ for any partition $\bar{\tau}$ with $\inf_{v_\tau, \sigma \in \mathcal{V}_\tau} \|u_{\tau}^{r_\tau} - v_\tau\|_{\mathcal{V}} \leq \lambda \|u_{\tau}^{r_\tau} - w_\tau\|_{\mathcal{V}}$. Then, following the proof of Lemma 6.4, we infer that $\# \bar{\tau} - \# r_0 \leq \# \bar{\tau} + \# \bar{\sigma} + \# \sigma$, with $\bar{\tau}$ and $\bar{\sigma}$ from (8.3).

The second statement can be proven as in [Ste07, Lemma 6.2].

Now we are ready to formulate our practical adaptive Stokes solver, STOKESOLVE.

STOKESOLVE($f, \varepsilon$) \rightarrow [$\sigma_j^{(i)}$, $p_j^{(i)}$, $r_{j,k}^{(i)}$, $w_{j,k}^{(i)}$]

% Let the parameter $\zeta$ from $\text{REFvel}$ satisfy $\zeta \in (0, \frac{C_2^1}{C_2^3})$, and let $\theta$ from $\text{REFpres}$
% satisfy $\theta \in (0, [1 - \frac{C_2^1 C_2^2}{C_2^3}]^\frac{1}{2})$. Let $\kappa, \alpha, \chi > 0$ be sufficiently small constants.
$\sigma_0 := \tau_0^{(0)} := \tau_0$, $p_0 := 0$, $w_{0,0} := 0$, $\delta = \|f\|_{\mathcal{V}}$, $i := j := k := 0$

do \{$f_{j,k}^{(i)}, r_{j,k}^{(i)}$\} := GALSSOLVE($r_{j,k}^{(i)}$, $f_{j,k}^{(i)}$, $p_j^{(i)}$, $w_{j,k}^{(i)}$, $\delta / 2$)

if $C_2^5 \mathcal{E}(\sigma_i, r_{j,k}^{(i)}, f_{j,k}^{(i)}, p_j^{(i)}, w_{j,k}^{(i)}) + (1 + C_3^2 (2c_2^{-1} + 1))\delta / 2 \leq \varepsilon$ then stop

else $\mathcal{E}(\sigma_i, r_{j,k}^{(i)}, f_{j,k}^{(i)}, p_j^{(i)}, w_{j,k}^{(i)}) + \delta \leq \kappa \mathcal{E}(\sigma_i, r_{j,k}^{(i)}, f_{j,k}^{(i)}, p_j^{(i)}, w_{j,k}^{(i)})$

\sigma_{i+1} := $\text{REFpres}([\sigma_i, w_{j,k}^{(i)}])$

$p_0^{(i+1)} := p_j^{(i)}$, $\tau_{0,0}^{(i+1)} := r_{j,k}^{(i)}$, $w_{0,0}^{(i+1)} := w_{j,k}^{(i)}$
\[ \delta := \beta(e^S(\infty, \tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) + \delta) \]

\[ i++, \ j := k := 0 \]

\begin{itemize}
  \item \textbf{elseif} \( E^E(\tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) + \delta \leq \alpha E^S(\sigma_i, \tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) \) \text{ then}
    \[ p^{(i)}_{j+1} := p^{(i)}_j - Q\sigma, \ \text{div} w^{(i)}_{j,k} \]
    \[ \tau^{(i)}_{j+1,0} := \tau^{(i)}_{j,k}, \ w^{(i)}_{j+1,0} := w^{(i)}_{j,k} \]
    \[ \delta := \beta(e^E(\sigma_i, \tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) + \delta) \]
    \[ j++, \ k := 0 \]
  \item \textbf{elseif} \( \delta \leq \chi E^E(\tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) \)
    \[ \tau^{(i)}_{j,k+1} := \text{REFvel}(\tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}), \ w^{(i)}_{j,k+1} := w^{(i)}_{j,k} \]
    \[ \delta := \beta(e^E(\tau^{(i)}_{j,k}, f^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) + \delta) \]
    \[ k+++ \]
  \item \textbf{else} \( \delta := \delta/2 \)
\end{itemize}

enddo

Compared to the preliminary version \textsc{stokeSSolve}0, in \textsc{stokeSSolve} the exact solution \( u^{(i)}_{j,k} \) of the inner elliptic problem is replaced by an approximate one, \( w^{(i)}_{j,k} \), which is moreover computed using an approximation \( f^{(i)}_{j,k} \in V_{j,k}^{(i)} \) of the right-hand side \( f \) in \( V' \). As a consequence, the results on the a posteriori error estimators from section 5 cannot be applied directly. Using the stability of the problems defining \( u \) and \( p \) with respect to perturbations in \( f \), and the stability of the error estimator \( E^E \) demonstrated in Proposition 5.4, for sufficiently small \( \alpha \) and \( \kappa \), stopping by either \( E^E(\ldots) + \delta \leq \alpha E^S(\sigma_i, \ldots), E^S(\sigma_i, \ldots) + \delta \leq \kappa E^S(\infty, \ldots) \) or \( E^E(\infty, \ldots) + (1+C_3(2c_2^{-1} + 1))\delta/2 \leq \varepsilon \) means that

\[ ||u^{(i)}_{j,k} - w^{(i)}_{j,k}||_V \leq D_1(\alpha)||p^{\sigma_i} - p^{(i)}_j||_V + ||u^{P^{\sigma_i}} - w^{(i)}_{j,k}||_V, \]

\[ ||p^{\sigma_i} - p^{(i)}_j||_V + ||u^{P^{\sigma_i}} - w^{(i)}_{j,k}||_V \leq D_2(\kappa)||p - p^{(i)}_j||_V + ||u - w^{(i)}_{j,k}||_V, \]

or

\[ ||p - p^{(i)}_j||_V + ||u - w^{(i)}_{j,k}||_V \leq \varepsilon, \]

respectively, with \( D_1(\alpha), D_2(\kappa) > 0 \) being some constants that tend to zero when \( \alpha \) or \( \kappa \) tend to zero.

Since the tolerance \( \delta/2 \) for the approximations for the right-hand side and for the approximate solution of the inner Galerkin problem decreases until \( \delta \leq \chi E^E(\ldots) \), where, for \( \beta \) small enough, the next \( \delta \) respects the same bound, the second part of Lemma 8.2 shows that for each \( i \) and \( j \), \( w^{(i)}_{j,k} \) converges linearly towards \( u^{(i)}_{j,k} \). Here, we silently assumed that by halving \( \delta \), at some point \( \delta \leq \chi E^E(\ldots) \) is valid. This, however, is not necessarily true since \( E^E(\ldots) \) changes as well. For example, think of the (unlikely) situation in which we have reached a partition on which \( u^{(i)}_{j,k} \) can be represented exactly. Yet, in that case we have also linear convergence of \( w^{(i)}_{j,k} \) towards \( u^{(i)}_{j,k} \). Indeed, if during the process of halving \( \delta \) it remains larger than \( \chi E^E(\ldots) \), then \( \chi E^E(\ldots) + \delta \), being up to some constant factor an upper bound for \( ||u^{(i)}_{j,k} - w^{(i)}_{j,k}||_V \), decreases linearly.

Based on these observations, similarly as in the proof of part (I) of Theorem 8.3, for sufficiently small \( \kappa \) and \( \alpha \), one shows convergence of \textsc{stokeSSolve}, with
\[ \|p - p^{(i)}_j\|_p + \| u - w^{(i)}_{j,k}\|_V \leq \varepsilon \text{ at termination.} \]

When \( \text{REFvel} \) is called, from \( \delta \leq \chi \epsilon^E(\cdot, \cdot) \) for sufficiently small \( \chi \), \( \epsilon^E(\cdot, \cdot) + \delta > \alpha \epsilon^S(\sigma_i, \cdot, \cdot) \) and \( \epsilon^S(\sigma_i, \cdot, \cdot) + \delta > \kappa \epsilon^S(\sigma_i, \cdot, \cdot) \), we have

\[ \|u^{(i)}_{j,k} - w^{(i)}_{j,k}\|_V \geq \epsilon^E(\cdot, \cdot) - \delta \geq \epsilon^S(\cdot, \cdot) + \delta \geq \epsilon^S(\sigma_i, \cdot, \cdot) + \delta \geq \epsilon^S(\sigma_i, \cdot, \cdot) + \delta \geq \varepsilon, \]

meaning that we may apply the first part of Lemma 8.2 to bound the cardinality of the set \( E_{i,j,k} \) of marked simplices. Assuming that \( p \in A^e_k \) and \( u \in A^v_k \), as in (7.21) we find

\[ \#E_{i,j,k} \lesssim \|u^{(i)}_{j,k} - w^{(i)}_{j,k}\|_V^{-1/8}(|u|^{1/8}_{A^e_k} + |p|^{1/8}_{A^v_k}) + \#\sigma_i - \#\tau_0 + \#\tau_0. \]

Other than in \( \text{STOKESSOLVE}_0 \), in \( \text{STOKESSOLVE} \) refinements can also be made by calls \( [f^{(i)}_{j,k}, \tau^{(i)}_{j,k}, w^{(i)}_{j,k}] := \text{RHS}[\delta/2, \tilde{\tau}^{(i)}_{j,k}] \). Assuming that \( \text{RHS} \) \( \tilde{s} \)-optimal with constant \( c_f \), then \( \#\tilde{\tau}^{(i)}_{j,k} - \#\tau^{(i)}_{j,k} \leq c_f^{-1/8}(\delta/2)^{-1/8} \). Using the fact that such a call can be made only when

\[ \delta \geq \epsilon^E(\cdot, \cdot) + \delta \geq \epsilon^S(\sigma_i, \cdot, \cdot) + \delta \geq \epsilon^S(\sigma_i, \cdot, \cdot) + \delta \geq \varepsilon, \]

by applying similar techniques to those in the proof of Theorem 7.1 one can prove that \( \text{STOKESSOLVE} \) outputs quasi-optimal partitions.

Finally, one can prove that \( \text{STOKESSOLVE} \) is of quasi-optimal computational complexity. The main point is that with a call \( \text{GALSOLVE}(r^{(i)}_{j,k}, p^{(i)}_{j,k}, f^{(i)}_{j,k}, w^{(i)}_{j,k}, \delta/2) \), it holds that \( \|u^{(i)}_{j,k} - w^{(i)}_{j,k}\|_V \lesssim \delta/2 \), so that the error has to be reduced by a constant factor only.

Along the lines indicated above, we end up with the following theorem. We have chosen not to include a full proof, since this would require another level of technicalities on top of those from the proof of Theorem 7.1.

**Theorem 8.3.** (I) \( [\sigma^{(i)}_j, p^{(i)}_j, \tau^{(i)}_{j,k}, w^{(i)}_{j,k}] := \text{STOKESSOLVE}^0[f, \varepsilon] \) terminates, and \( \|u - w^{(i)}_{j,k}\|_V + \|p - p^{(i)}_j\|_V \leq \varepsilon \). (II) If, for some \( s > 0 \), \( p \in A^e_k \), then \( \#\sigma^{(i)}_j - \#\tau_0 \leq \varepsilon^{-1/8}|p|^{1/8}_{A^e_k} \), dependent only on \( \tau_0 \), and on \( s \) when it tends to 0 or infinity. If, in addition, for some \( \tilde{s} > 0 \), \( u \in A^v_k \), and with \( \tilde{s} = \min(s, \tilde{s}) \), \( \text{RHS} \) is \( \tilde{s} \)-optimal with constant \( c_f \), then \( \#\tilde{\tau}^{(i)}_{j,k} - \#\tau_0 \leq \varepsilon^{-1/8}(|p|^{1/8}_{A^e_k} + \|u|^{1/8}_{A^v_k} + c^{1/8}_{vcf}) \), dependent only on \( \tau_0 \), and on \( s \) when it tends to 0 or infinity. Moreover, when \( \varepsilon \lesssim \|f\|_V, \) the number of arithmetic operations and storage locations required by the call is also bounded by a multiple of \( \varepsilon^{-1/8}(|p|^{1/8}_{A^e_k} + \|u|^{1/8}_{A^v_k} + c^{1/8}_{vcf}) \).

**9. Numerical experiment.** We consider \( d = 2 \), the L-shaped domain \( \Omega = (0,1)^2\setminus(\frac{1}{3},1)^2 \), \( f : x \mapsto 25(4x_2 - 1, 1-4x_1) \), and \( m = 1, \) i.e., continuous piecewise linear approximation for the velocity, piecewise constant approximation for the pressure, and an initial partition into 24 triangles.

The “bulk chasing” parameters \( \theta \) and \( \zeta \) inside \( \text{REFpres} \) or \( \text{REFvel} \) were chosen to be \( 0.7 \) and \( \sqrt{0.3} \), respectively. Since \( f \) is smooth, we followed the approach discussed in Remark 8.1 and skipped the calls of \( \text{RHS} \). Furthermore, instead of solving each arising finite-dimensional linear system within tolerance \( \delta/2 \), for the first value of \( \delta/2 \) obtained by successively halving that is less than \( \chi \epsilon^E(\tau^{(i)}_{j,k}, \tilde{\tau}^{(i)}_{j,k}, p^{(i)}_j, w^{(i)}_{j,k}) \), we always approximately solved it by 3 multigrid iterations with local smoothing (cf. [WC06]).
starting with the previously computed approximate velocity. The parameters \( \kappa \) and \( \alpha \) are chosen to be 0.88 and 0.9, respectively.

For comparison, we also implemented the adaptive Uzawa method for solving the full Stokes problem, i.e., \texttt{STOKESSOLVE} without the outermost loop over \( i \), i.e., without the part starting from the first \texttt{elseif} statement until the second one, and all remaining occurrences of \( \sigma \) replaced by \( \infty \). In particular the pressure update \( p_{j+1}^{(i)} := p_j^{(i)} + Q \sigma \text{div} w^{(i)} \) then reads \( p_{j+1} := p_j + \text{div} w_{j,k} \). This is the algorithm studied in [BMN02], apart from the replacement of a priori prescribed tolerances by a posteriori ones.

In Figure 9.1 (left), we plotted the full Stokes error estimator \( E_S(\infty, \tau_{j,k}^{(i)}, f, P_{j,k}^{(i)}, w_{j,k}^{(i)}) \) vs. \# \( \tau_{j,k}^{(i)} \) (or \( E_S(\infty, \tau_{j,k}^{(i)}, f, p_j, w_{j,k}^{(i)}) \) vs. \# \( \tau_{j,k}^{(i)} \)). Ignoring the fact that generally \( w_{j,k}^{(i)} \neq p_{j,k}^{(i)} \) or \( w_{j,k}^{(i)} \neq u_{j,k}^{(i)} \), modulo a constant factor this estimator is an upper bound for \( \| u - w_{j,k}^{(i)} \|_V + \| p - P_{j,i}^{(i)} \|_P \) (or \( \| u - w_{j,k}^{(i)} \|_V + \| p - p_j \|_P \)). (For \( f \in V_{j,k}^{(i)} \) (or \( f \in \mathcal{V}_{j,k}^{(i)} \)), it would also be a lower bound.) In Figure 9.1 (right), we plotted the pressure error estimator \( \| \text{div} w_{j,k}^{(i)} \|_{L_2(\Omega)} \) vs. \# \( \tau \) (or \( \| \text{div} w_{j,k}^{(i)} \|_{L_2(\Omega)} \) vs. \# \( \tau_{j,k} \)). Ignoring the fact that generally \( w_{j,k}^{(i)} \neq p_{j,k}^{(i)} \) or \( w_{j,k}^{(i)} \neq u_{j,k}^{(i)} \), this estimator is equivalent to \( \| p - P_{j,i}^{(i)} \|_P \) (or \( \| p - p_j \|_P \)). As predicted by the theory, the approximations produced by \texttt{STOKESSOLVE} converge with the best possible rates. In this example, we observe that the same is true for the adaptive Uzawa method.

Our implementation is partly written in C and, for our convenience, partly in MATLAB, making use of the PDE toolbox. This toolbox was modified such that it performs newest vertex bisection. Due to the data structures used in the MATLAB part, our code is not of optimal computational complexity; the time needed for a call of \texttt{REFvel} is not proportional to the cardinality of its output partition. Subtracting the times spent for this routine, we observed computing times that are proportional to the cardinality of the final partition.

In Figure 9.2 we show two partitions \( \tau_j \) that were produced by \texttt{STOKESSOLVE}. Note that these partitions are nonconforming. In Figure 9.3 we give two partitions \( \tau_{j,k}^{(i)} \). Finally, in Figure 9.4 plots of \( p \) and \( u \) are given.
Fig. 9.2. Pressure partitions with 1965 or 4235 triangles.

Fig. 9.3. Velocity partitions with 1842 or 6126 triangles.

Fig. 9.4. Pressure $p$ and the norm $|u|$ of the velocity.

Acknowledgments. The authors are indebted to Dr. Haijun Wu (Nanjing University, China) for making available to us his code implementing multigrid on locally refined partitions. We thank Dr. Peter Binev (University of South Carolina, USA) for proving a new estimate for an adaptive tree algorithm that we needed for the REFPRES routine.

REFERENCES


