A robust multigrid method for the time-dependent Stokes problem

Stefan Takacs

Abstract In the present paper we propose a coupled multigrid method for generalized Stokes flow problems. Such problems occur as subproblems in implicit time-stepping approaches for time-dependent Stokes problems. The discretized Stokes system is a large-scale linear system whose condition number depends on the grid size of the spatial discretization and of the length of the time step. Recently, for this problem a coupled multigrid method has been proposed, where in each smoothing step a Poisson problem has to be solved (approximately) for the pressure field. In the present paper, we propose a coupled multigrid method where the solution of such sub-problems is not needed. We prove that the proposed method shows robust convergence behavior in the grid size of the spatial discretization and of the length of the time-step.

Keywords Generalized Stokes problem · coupled multigrid methods · robustness

1 Introduction

In the present paper, we consider the following model problem (generalized Stokes flow problem). Let $\Omega \subseteq \mathbb{R}^2$ be a bounded polygonal domain. Assume $f \in [L^2(\Omega)]^2$ and $g \in L^2(\Omega)$ to be given. Find a velocity field $u$ and a pressure distribution $p$ such that

$$
\begin{align*}
-\Delta u + \beta u + \nabla p &= f \text{ in } \Omega, \\
\nabla \cdot u &= g \text{ in } \Omega, \\
\n\end{align*}
$$

$$
\begin{align*}
u &= 0 \text{ on } \partial \Omega
\end{align*}
$$

(1)

The research was funded by the Austrian Science Fund (FWF): J3362-N25.

Stefan Takacs
Visiting Postdoc, Mathematical Institute, University of Oxford, United Kingdom
E-mail: stefan.takacs@numa.uni-linz.ac.at
is satisfied. $\beta > 0$ is assumed to be a given parameter. The problem (1) appears as an auxiliary problem for implicit time-stepping approaches to solve an incompressible, time-dependent Stokes flow problem. In this case, the parameter $\beta$ is proportional to the inverse of the length of the time-step, scaled by a viscosity parameter.

To obtain existence and uniqueness of the solution, we further require $\int_{\Omega} p \, dx = \int_{\Omega} g \, dx = 0$. Note that the analysis presented in this paper is (due to the use of regularity results) restricted to convex domains. However, the numerical method presented in the paper can be applied also to non-convex domains.

The discretization of the problem leads to an indefinite linear system with saddle-point structure. The main goal of this work is to construct and to analyze numerical methods that produce an approximate solution to the problem, where the computational complexity can be bounded by the number of unknowns times a constant which is independent of the grid level (of the spatial discretization) and the choice of $\beta$, in particular for large values of $\beta$.

For the solution of such a saddle-point problem, there are several possibilities. In [5, 11, 13, 14, 17, 26], various kinds of preconditioners have been proposed for this problem which can be combined with a Krylov-subspace method as outer iteration scheme to yield an iterative solver for the problem.

An alternative is to apply a multigrid algorithm directly to the coupled system. Such methods are on the one hand typically quite fast and on the other hand using those methods one does not need an outer iteration scheme. For $\beta = 0$, problem (1) is the standard Stokes problem. For this case several multigrid solvers are available, see, e.g., [13, 22, 6, 25] and the papers cited in [23, 15]. The construction of a multigrid method for $\beta > 0$, particularly if the method is desired to show robust convergence behavior in $\beta$, is more involved, see [12] for an overview and numerical results. Recently, an important step forward has been archived by Olshanskii, who has proposed such a robust multigrid method, see [15]. In the named paper however, a Poisson problem for the pressure has to be solved approximately (for example by applying one V-cycle, cf. Section 8 in [15]) for each step of the smoothing iteration. (We will comment on this in Remarks 2 and 3).

The goal of the present paper is to drop this requirement and, therefore, reduce the computational costs. We propose a multigrid solver where the smoother is a simple linear iteration scheme and we prove that the proposed method is robust in the grid size of the spatial discretization and in the choice of $\beta$. We will present a convergence proof for our multigrid method based on the classical splitting of the analysis into smoothing property and approximation property, see [9].

The results of this paper form a basis of the convergence analysis of an all-at-once multigrid method for the Stokes optimal control problem, cf. [20].

This paper is organized as follows. In Section 2 we will introduce the variational formulation and discuss its discretization. In Section 3 we will introduce the new multigrid approach. In Section 4 we will discuss the choice of the smoother. The proof of the approximation property will be given in Section 5.
Numerical results which illustrate the convergence result will be presented in Section 6. In Section 7 we will close with conclusions.

2 Variational formulation and discretization

Here and in what follows, \( L^2(\Omega) \) and \( H^1(\Omega) \) denote the standard Lebesgue and Sobolev spaces with associated standard norms \( \| \cdot \|_{L^2(\Omega)} \) and \( \| \cdot \|_{H^1(\Omega)} \), respectively. The space \( L^2_0(\Omega) \) is the space of all \( L^2 \)-functions with mean value 0, i.e.,

\[
L^2_0(\Omega) := \left\{ v \in L^2(\Omega) : \int_{\Omega} v \, d\xi = 0 \right\}.
\]

The space \( H^1_0(\Omega) \) is the space of all functions in \( H^1(\Omega) \), vanishing on the boundary. Both spaces are equipped with standard norms, i.e., \( \| \cdot \|_{L^2_0(\Omega)} \) and \( \| \cdot \|_{H^1_0(\Omega)} \).

Using these spaces, we can set up the variational formulation of (1), which reads as follows. Find \( u \in U := [H^1_0(\Omega)]^2 \) and \( p \in P := L^2_0(\Omega) \) such that

\[
(\nabla u, \nabla \tilde{u})_{L^2(\Omega)} + \beta(u, \tilde{u})_{L^2(\Omega)} + (p, \nabla \cdot \tilde{u})_{L^2(\Omega)} = (f, \tilde{u})_{L^2(\Omega)}
\]

holds for all \( \tilde{u} \in U \) and \( \tilde{p} \in P \). Certainly, the variational problem can be rewritten as one variational equation as follows. Find \( x \in X \) such that

\[
B(x, \tilde{x}) = F(\tilde{x}) \quad \text{for all } \tilde{x} \in X,
\]

where \( X := U \times P \) and

\[
B((u,p), (\tilde{u}, \tilde{p})) := (\nabla u, \nabla \tilde{u})_{L^2(\Omega)} + \beta(u, \tilde{u})_{L^2(\Omega)} + (p, \nabla \cdot \tilde{u})_{L^2(\Omega)}
+ (\nabla \cdot u, \tilde{p})_{L^2(\Omega)},
\]

\[
F(\tilde{u}, \tilde{p}) := (f, \tilde{u})_{L^2(\Omega)} + (g, \tilde{p})_{L^2(\Omega)}.
\]

We are interested in finding an approximative solution for equation (2). The convergence analysis follows standard approaches, i.e., we show that the problem in question is well posed in some norm \( \| \cdot \|_X \) (which is for the Poisson problem the \( H^1 \)-norm and for the standard Stokes problem the pair of \( H^1 \) for the velocity and \( L^1 \) for the pressure). In our case this norm will be parameter-dependent. In a second step, we will introduce a properly scaled \( L^2 \)-like norm \( \| \cdot \|_k \), where an inverse inequality shows that it is an upper bound of the norm \( \| \cdot \|_X \). In this norm the convergence of the multigrid method (smoothing property and approximation property) will be shown.

The main focus of the paper is the proof of the approximation property. Throughout the paper we will comment on the differences to the approach presented in [15], which allows us to drop the requirement of solving (approximately) Poisson problems. The key idea of the proof of the approximation property follows classical approaches. However, the way of constructing the norms is non-standard. For simplicity, we follow the abstract framework introduced in [21].

First we introduce the following convenient notation.
Throughout this paper, \( C > 0 \) is a generic constant, independent of the grid level \( k \) and the choice of the parameter \( \beta \). For any scalars \( a \) and \( b \), we write \( a \lesssim b \) (or \( b \gtrsim a \)) if there is a constant \( C > 0 \) such that \( a < Cb \). We write \( a \approx b \) if \( a \lesssim b \lesssim a \).

Let the Hilbert spaces \( X, U \) and \( P \) (introduced above) be equipped with the following norms:

\[
\|x\|_X^2 := \|(u, p)\|_X^2 := \|u\|_L^2 + \|p\|_P^2, \\
\|u\|_U^2 := \|u\|_{H^1(\Omega)}^2 + \beta \|u\|_{L^2(\Omega)}^2 \\
\|p\|_P^2 := \sup_{0 \neq w \in \{H_0^1(\Omega)\}^2} \frac{(p, \nabla \cdot w)_{L^2(\Omega)}}{\|w\|_{H^1(\Omega)}^2 + \beta \|w\|_{L^2(\Omega)}^2}. \tag{3}
\]

Lemma 2.1 in [15] states the following result.

**Lemma 1** The relation

\[
\|x\|_X \lesssim \sup_{0 \neq \tilde{x} \in X} \frac{B(x, \tilde{x})}{\|\tilde{x}\|_X} \lesssim \|x\|_X \quad \text{(A1)}
\]

holds for all \( x \in X \).

Using the following notation, we can express the norms in a nicer way.

**Notation 2** For any Hilbert space \( A \), the symbol \( A^* \) denotes its dual space equipped with the dual norm

\[
\|u\|_{A^*} := \sup_{0 \neq w \in A} \frac{(u, w)}{\|w\|_A},
\]

where \( (u, \cdot) := u(w) \) denotes the duality pairing.

For any Hilbert space \( A \) and any scalar \( a > 0 \), the symbol \( aA \) denotes the space on the underlying set of the Hilbert space \( A \) equipped with the norm

\[
\|u\|_{aA}^2 := a \|u\|_A^2.
\]

For any two Hilbert spaces \( A \) and \( B \), the symbol \( A \cap B \) denotes the space on the intersection of the underlying sets, \( \{u \in A \cap B\} \), equipped with the norm

\[
\|u\|_{A \cap B}^2 := \|u\|_A^2 + \|u\|_B^2,
\]

and the symbol \( A + B \) denotes the space on the algebraic sum of the underlying sets, \( \{u_1 + u_2 : u_1 \in A, u_2 \in B\} \), equipped with the norm

\[
\|u\|_{A + B}^2 := \inf_{u_1 \in A, u_2 \in B, u = u_1 + u_2} \|u_1\|_A^2 + \|u_2\|_B^2.
\]
The spaces $A^*$, $aA$, $A \cap B$ and $A + B$ are Hilbert spaces. The fact that $A^*$ is a Hilbert space follows directly from the Riesz representation theorem, see, e.g., Theorem 1.2 in [1]. The fact that $aA$ is a Hilbert space is obvious and for the latter two see, e.g., Lemma 2.3.1 in [3].

We immediately see that the norm on $U$ can be rewritten as follows:

$$\|u\|_U = \|u\|_{H^1(\Omega) \cap L^2(\Omega)}.$$ 

To reformulate the norm on $P$, we need the following regularity assumption.

**Lemma 2** The following condition is satisfied for convex polygonal domains.

(R) Regularity of the Stokes problem. Let $f \in [L^2(\Omega)]^2$ and $g \in H^1_0(\Omega) \cap L^2_0(\Omega)$ be arbitrary but fixed and $(u, p) \in [H^1_0(\Omega)]^2 \times L^2_0(\Omega)$ be the solution of the Stokes problem, i.e., such that

$$(\nabla u, \nabla \tilde{u})_{L^2(\Omega)} + (p, \nabla \tilde{u})_{L^2(\Omega)} = (f, \tilde{u})_{L^2(\Omega)}$$

holds for all $(\tilde{u}, \tilde{p}) \in [H^1_0(\Omega)]^2 \times L^2_0(\Omega)$. Then $(u, p) \in [H^2(\Omega)]^2 \times H^1(\Omega)$ and

$$\|u\|_{H^2(\Omega)}^2 + \|p\|_{H^1(\Omega)}^2 \lesssim \|f\|_{L^2(\Omega)}^2 + \|g\|_{H^1(\Omega)}^2.$$ 

Proof Theorem 2 in [10] states (in the notation of the present paper) that provided $f \in [L^2(\Omega)]^2$, $g \in H^1(\Omega) \cap L^2_0(\Omega)$ and $\delta^{-1} g \in L^2(\Omega)$ that

$$\|u\|_{H^2(\Omega)}^2 + \|\nabla p\|_{L^2(\Omega)}^2 \lesssim \|f\|_{L^2(\Omega)}^2 + \|\nabla g\|_{L^2(\Omega)}^2 + \|\delta^{-1} g\|_{L^2(\Omega)}^2,$$

is satisfied, where $\delta : \Omega \to \mathbb{R}$ is the distance to the closest vertex. Lemma 2 in [10] states that $\|\delta^{-1} g\|_{L^2(\Omega)} \lesssim \|g\|_{H^1(\Omega)}$ is satisfied for all $g \in H^1_0(\Omega)$.

Combining these results, we obtain that

$$\|u\|_{H^2(\Omega)}^2 + \|\nabla p\|_{L^2(\Omega)}^2 \lesssim \|f\|_{L^2(\Omega)}^2 + \|g\|_{H^1(\Omega)}^2$$

is satisfied for all $f \in [L^2(\Omega)]^2$ and $g \in H^1_0(\Omega) \cap L^2_0(\Omega)$. As $p \in L^2_0(\Omega)$ was assumed, Poincaré’s inequality states that $\|p\|_{H^1(\Omega)} \lesssim \|\nabla p\|_{L^2(\Omega)}$, which finishes the proof.

Note the fact that we assume $g$ to satisfy homogeneous Dirichlet boundary conditions. This condition can be weakened but it is not possible to drop such a condition completely, cf. [10].

**Lemma 3** On domains $\Omega$, where (R) is satisfied,

$$\|p\|_{P} \approx \|p\|_{L^2(\Omega) + \beta^{-1} H^1(\Omega)}$$

holds for all $p \in L^2_0(\Omega)$. 
For a proof of this lemma, see Theorem 3.2 in [16] (this proof only needs (R) for \( f \in [L^2(\Omega)]^2 \) and \( g = 0 \)).

The discretization of problem (2) is done using standard finite element techniques. We assume to have for \( k = 0, 1, 2, \ldots \) a sequence of grids obtained by uniform refinement. On each grid level \( k \), we discretize the problem using the Galerkin approach, i.e., we have finite dimensional spaces \( X_k \subseteq X \) and consider the following problem: Find \( x_k \in X_k \) such that

\[
B(x_k, \tilde{x}_k) = F(\tilde{x}_k) \quad \text{for all } \tilde{x}_k \in X_k.
\]

Using a nodal basis, we can represent this problem in matrix-vector notation as follows:

\[
A_k \overline{x}_k = \overline{f}_k,
\]

where

\[
A_k = \begin{pmatrix} A_k & B_k^T \\ B_k & 0 \end{pmatrix}, \quad \overline{x}_k = \begin{pmatrix} u_k \\ p_k \end{pmatrix}, \quad \overline{f}_k = \begin{pmatrix} f_k \\ g_k \end{pmatrix}
\]

and the matrices \( A_k \) represents the scalar product \( a(u, \tilde{u}) := (\nabla u, \nabla \tilde{u})_{L^2(\Omega)} + \beta(u, \tilde{u})_{L^2(\Omega)} \) and \( B_k \) represents the scalar product \( b(u, \tilde{p}) := (\nabla \cdot u, \tilde{p})_{L^2(\Omega)} \).

Here and in what follows, any underlined quantity, like \( x_k \), is the representation of the corresponding non-underlined quantity, here \( x_k \), with respect to a nodal basis of the corresponding Hilbert space, here \( X_k \).

The next step is to show the discrete inf-sup condition, i.e., that

\[
\| x_k \|_X \lesssim \sup_{0 \neq \tilde{x}_k \in X_k} \frac{B(x_k, \tilde{x}_k)}{\| \tilde{x}_k \|_X} \lesssim \| x_k \|_X
\]

(A1a)

holds for all \( x_k \in X_k \).

To guarantee the discrete inf-sup condition, we have to choose a discretization which is stable for the standard Stokes problem, particularly a discretization that satisfies the weak inf-sup condition, i.e.,

\[
\sup_{0 \neq u_k \in U_k} \frac{(\nabla \cdot u_k, p_k)_{L^2(\Omega)}}{|u_k|_{L^2(\Omega)}} \gtrsim \| \nabla p_k \|_{L^2(\Omega)}
\]

should hold for all \( p_k \in P_k \). Note that this is a standard condition which guarantees that the chosen discretization is stable for the Stokes problem. In [23] it was shown that condition (6) is satisfied for the Taylor-Hood element (\( P_1 - P_2 \)-element) for polygonal domains where at least one vertex of each element is located in the interior of the domain. Here and in what follows we assume that the problem is discretized with the Taylor-Hood element and that the mesh satisfies the named condition.

Using the weak-inf sup condition (6) we can show (A1a).

**Lemma 4** If the problem is discretized using the Taylor-Hood element, condition (A1a) is satisfied.
The estimate (2.16) in [15] states that
\[ \|x_k\|_{X_{1,k}} \lesssim \sup_{0 \neq \tilde{x}_k \in X_k} \frac{B(x_k, \tilde{x}_k)}{\|\tilde{x}_k\|_{X_{1,k}}}, \]
where
\[ \|x_k\|_{X_{1,k}}^2 := \|(u_k, p_k)\|_{U_{1,k}}^2 := \|u_k\|_{H^1(\Omega)}^2 + \|p_k\|_{P_{1,k}}^2, \]
\[ \|u_k\|_{U_{1,k}}^2 := \|u_k\|_{H^1(\Omega)}^2 + \beta \|u_k\|_{L^2(\Omega)}^2 \]
and
\[ \|p_k\|_{P_{1,k}}^2 := \sup_{0 \neq w_k \in U_k} \frac{(p_k, \nabla \cdot w_k)_{L^2(\Omega)}^2}{\|w_k\|_{H^1(\Omega)}^2 + \beta \|w_k\|_{L^2(\Omega)}^2}, \]
holds for all \( x_k \in X_k \). Note that \( \|\cdot\|_{U_{1,k}} = \|\cdot\|_U \). Lemma 2.2 in [15] states that \( \|\cdot\|_{P_{1,k}} \lesssim \|\cdot\|_P \). This shows the left inequality in condition (A1a) (discrete inf-sup-condition).

Note that due to the fact that the grids are obtained by uniform refinement, the discrete subsets are nested, i.e., \( U_k \subseteq U_{k+1} \) and \( P_k \subseteq P_{k+1} \). Therefore, also \( X_k \subseteq X_{k+1} \) holds.

3 A coupled multigrid method

The problem shall be solved using a multigrid method. The abstract algorithm for solving the discretized equation (5) on grid level \( k \) reads as follows. Starting from an initial approximation \( x_k^{(0)} \), one iterate of the multigrid method is given by the following two steps:

- **Smoothing procedure:** Compute
  \[ x_k^{(0,m)} := x_k^{(0,m-1)} + \tilde{A}_k^{-1} \left( f_k - A_k x_k^{(0,m-1)} \right) \quad \text{for } m = 1, \ldots, \nu \]  
  (8)
  with \( x_k^{(0,0)} = x_k^{(0)} \). The choice of the smoother (or, in other words, of the preconditioning matrix \( \tilde{A}_k^{-1} \)) will be discussed below.

- **Coarse-grid correction:**
  - Compute the defect \( x_k^{(1)} := f_k - A_k x_k^{(0,\nu)} \) and restrict it to grid level \( k-1 \) using an restriction matrix \( I_k^{-1} \):
    \[ x_k^{(1)} := I_k^{-1} \left( f_k - A_k x_k^{(0,\nu)} \right). \]
  - Solve the coarse-grid problem
    \[ A_{k-1} x_k^{(1)} = x_k^{(1)} \]  
    approximately.

Proof: The estimate (2.16) in [15] states that
\[ \|x_k\|_{X_{1,k}} \lesssim \sup_{0 \neq \tilde{x}_k \in X_k} \frac{B(x_k, \tilde{x}_k)}{\|\tilde{x}_k\|_{X_{1,k}}}, \]
Prolongate $p_{k-1}$ to the grid level $k$ using an prolongation matrix $I^k_{k-1}$ and add the result to the previous iterate:

$$x_k^{(1)} := x_k^{(0,\nu)} + I^k_{k-1} p_{k-1}^{(1)}.$$ 

As we have assumed to have nested spaces, the intergrid-transfer matrices $I^k_{k-1}$ and $I^{k-1}_{k}$ can be chosen in a canonical way: the restriction $I^k_{k-1}$ is the standard $L^2$-projection of $X_k$ to $X_{k+1}$ and $I^{k-1}_{k} = (I^k_{k-1})^T$ is the canonical embedding.

If the problem on the coarser grid is solved exactly (two-grid method), the coarse-grid correction is given by

$$x_k^{(1)} := x_k^{(0,\nu)} + I^k_{k-1} A_{k-1}^{-1} I^{k-1}_{k} (f_k - A_k x_k^{(0,\nu)}).$$

In practice, the problem (9) is approximately solved by applying one step (V-cycle) or two steps (W-cycle) of the multigrid method, recursively. On grid level $k = 0$, the problem (9) is solved exactly.

To construct a multigrid convergence result based on Hackbusch’s splitting of the analysis into smoothing property and approximation property, we have to introduce an appropriate framework.

Convergence is shown in the following $L^2$-like norms $\| \cdot \|_k$:

$$\|x_k\|_k^2 := \|x_k\|_{L_k}^2 := (L_k x_k, x_k),$$

where

$$L_k := \left( \begin{array}{cc} (h^{-2} + \beta)M_{U,k} & h^{-2} (\beta + h^{-2})^{-1} M_{P,k} \end{array} \right),$$

(11)

where the matrices $M_{U,k}$ and $M_{P,k}$ are mass-matrices, representing the $L^2$-inner product in $U_k$ and $P_k$, respectively.

The smoothing property and the approximation property, which we will show below, read as follows.

- **Smoothing property:**

  $$\sup_{\tilde{x}_k \in X_k} \mathcal{B} \left( x_k^{(0,\nu)} - x_k^*, \tilde{x}_k \right) \leq \eta(\nu) \|x_k^{(0)} - x_k^*\|_k$$

holds for some function $\eta(\nu)$ with $\lim_{\nu \to \infty} \eta(\nu) = 0$. Here and in what follows, $x_k^* \in X_k$ is the exact solution of the discretized problem (4).

- **Approximation property:**

  $$\|x_k^{(1)} - x_k^*\|_k \leq C_A \sup_{\tilde{x}_k \in X_k} \mathcal{B} \left( x_k^{(0,\nu)} - x_k^*, \tilde{x}_k \right)$$

holds for some constant $C_A > 0$. 

[1] Stefan Takacs
It is easy to see that, if we combine both conditions, we obtain
\[ \|x_k^{(1)} - x_k^*\|_k \leq q(\nu)\|x_k^{(0)} - x_k^*\|_k, \]
where \( q(\nu) = C_A \eta(\nu) \), i.e., that the two-grid method converges for \( \nu \) large enough. The convergence of the W-cycle multigrid method can be shown under mild assumptions, see e.g. [9].

**Remark 1** The norm \( \| \cdot \|_k \) is the discrete analog of
\[ \|x_k\|_{X_{o,k}} := \|(u_k, p_k)\|_{X_{o,k}}^2 := \|u_k\|_{L^2(\Omega)}^2 + \|p_k\|_{P_{o,k}}^2. \]
\[ \|u_k\|_{U_{o,k}}^2 := h_k^{-2}\|u_k\|_{L^2(\Omega)}^2 + \beta\|u_k\|_{L^2(\Omega)}^2 \quad \text{and} \]
\[ \|p_k\|_{P_{o,k}}^2 := \sup_{0 \neq w_k \in L^2(\Omega)} \frac{h_k^{-2}(p_k, w_k)_{L^2(\Omega)}}{\|w_k\|_{L^2(\Omega)}^2 + \beta\|w_k\|_{L^2(\Omega)}^2} \quad (14) \]
\[ = h_k^{-2}(h_k^{-2} + \beta)^{-1}\|p_k\|_{L^2(\Omega)}^2. \]

Note that this norm is obtained from the norm \( \| \cdot \|_X \) by “replacing all differentials by \( h_k^{-1} \). This is a common construction principle, cf. [9].

However, in [15] the norm
\[ \|x_k\|_{X_{o,k}} := \|(u_k, p_k)\|_{X_{o,k}}^2 := \|u_k\|_{L^2(\Omega)}^2 + \|p_k\|_{P_{k}}^2, \]
\[ \|u_k\|_{U_{o,k}}^2 := h_k^{-2}\|u_k\|_{L^2(\Omega)}^2 + \beta\|u_k\|_{L^2(\Omega)}^2 \quad \text{and} \]
\[ \|p_k\|_{P_{k}}^2 := \sup_{0 \neq w_k \in U_k} \frac{(p_k, \nabla \cdot w_k)_{L^2(\Omega)}}{\|w_k\|_{H^1(\Omega)}^2 + \beta\|w_k\|_{L^2(\Omega)}^2}. \]

was used. Note that the norm for the velocity field coincides with our choice (pure \( L^2 \)) but the norm for the pressure distribution is still the original norm, as introduced in [7]. This is the reason that in the algorithms proposed in [13], a discrete pressure Poisson problem has to be solved (at least approximately) in each smoothing step.

**4 Smoother and proof of the smoothing property**

The choice of an appropriate smoother is a key issue in constructing a coupled multigrid method for an indefinite problem. In this paper, we introduce two kinds of smoothers. The first smoother is appropriate for a large class of problems including the model problem: the normal equation smoother, cf. [7], which reads as follows.
\[ \tilde{x}_k^{(0,m)} := \tilde{x}_k^{(0,m-1)} + \tau A_k^{-1} \tilde{x}_k^{(0,m-1)} - A_k \tilde{x}_k^{(0,m-1)} \]
for \( m = 1, \ldots, \nu \).

Here, a fixed \( \tau > 0 \) has to be chosen such that the spectral radius \( \rho(\tau A_k^{-1} A_k) \) is bounded away from 2 on all grid levels \( k \) and for all choices of the parameter \( \beta \).
Remark 2 The distributive smoother which was proposed in [15] for the general-
ized Stokes flow problem was a normal equation smoother. There, the nor-
mal equation was taken in another Hilbert space. Therefore, the application of
the inverse of the matrix, which represents the scalar product on the Hilbert
space, to some given vector $w_k$ (in our notation: solving the linear system
$\mathcal{L}_k x_k = w_k$), a Poisson problem has to be solved for the pressure distribution.

Using a standard inverse inequality, one can show that

$$
\|x_k\|_X \lesssim \|x_k\|_k
$$

is satisfied for all $x_k \in X_k$. Based on this result, using an eigenvalue analysis
one can show the following lemma, cf. [7].

Lemma 5 The damping parameter $\tau > 0$ can be chosen independent of grid
level $k$ and the choice of the parameter $\beta$ such that

$$
\tau \rho(\hat{A}_k^{-1} A_k) \leq 2 - \epsilon < 2,
$$

holds for some constant $\epsilon > 0$. For this choice of $\tau$, there is a constant $C_S > 0$
independent of the grid level $k$ and the choice of the parameter $\beta$, such that the smoothing property
[12] is satisfied with rate $\eta(\nu) := C_S \nu^{-1/2}$.

Certainly, the smoothing procedure [12] should be efficient-to-apply. Using
the fact, that the mass matrices $M_{U,k}$ and $M_{P,k}$ in (11) and their diagonals
are spectrally equivalent under weak assumptions, for the practical realization
of the normal equation smoother these mass matrices can be replaced by their
diagonals.

The second smoother, which we propose, is a Uzawa type smoother, cf. [19].
Here, one step of the smoother to compute $x_k^{(0,m)} = (u_k^{(0,m)}, p_k^{(0,m)})$
based on $x_k^{(0,m-1)} = (u_k^{(0,m-1)}, p_k^{(0,m-1)})$ reads as follows:

$$
\begin{align*}
\hat{u}_k^{(0,m-1/2)} &:= u_k^{(0,m-1)} + \tau \hat{A}_k^{-1} \left( \hat{f}_k - A_k u_k^{(0,m-1)} - B_k^T p_k^{(0,m-1)} \right) \\
\hat{p}_k^{(0,m-1)} &:= p_k^{(0,m-1)} - \sigma \hat{S}_k^{-1} \left( \hat{g}_k - B_k u_k^{(0,m-1/2)} \right) \\
\hat{u}_k^{(0,m)} &:= u_k^{(0,m-1)} + \tau \hat{A}_k^{-1} \left( \hat{f}_k - A_k u_k^{(0,m-1)} - B_k^T p_k^{(0,m)} \right),
\end{align*}
$$

where $\hat{A}_k$ and $\hat{S}_k$ are the (1,1)-block and the (2,2)-block of $\mathcal{L}_k$, respectively.

The smoother can be rewritten in the compact notation [8], where

$$
\hat{A}_k := \begin{pmatrix} \tau^{-1} A_k & B_k^T \\ B_k & \tau B_k \hat{A}_k^{-1} B^T - \sigma^{-1} \hat{S}_k \end{pmatrix}.
$$

Here, the smoothing property can be shown using the theory introduced in [19].
Lemma 6 Let $\tilde{A}_k$ and $\tilde{S}_k$ be the $(1,1)$-block and the $(2,2)$-block of $L_k$, respectively. Then $\tau > 0$ and $\sigma > 0$ can be chosen independent of the grid level and the choice of the parameter $\beta$ such that

$$\tau^{-1} \tilde{A}_k \geq A_k \quad \text{and} \quad \sigma^{-1} \tilde{S}_k \geq \tau B_k \tilde{A}_k^{-1} B_k^T.$$  

For this choice of $\tau$ and $\sigma$, there is a constant $C_S > 0$, independent of the grid level $k$ and the choice of the parameter $\beta$, such that the smoothing property (12) is satisfied with rate $\eta(\nu) := C_S \nu^{-1/2}$.

Proof The fact that $\tau > 0$ and $\sigma > 0$ can be chosen independent of the grid level and the choice of the parameter $\beta$ follow from standard inverse inequalities.

For the smoothing property, we apply Theorem 4 in [19] with the choice $\mathcal{K}_k := L_k^{-1/2} \tilde{A}_k L_k^{-1/2}$ and $\tilde{K}_k := L_k^{-1/2} \tilde{A}_k L_k^{-1/2}$. This immediately implies

$$\| L_k^{-1/2} A_k (I - \tilde{A}_k A_k)^{1/2} \|_{\mathcal{C}} \leq \eta(\nu) \| L_k^{-1/2} \tilde{A}_k L_k^{-1/2} \|_{\mathcal{C}}.$$  

As $\ell_k^T A_k \tilde{A}_k \ell_k = B(x_k, \tilde{x}_k) \lesssim \| x_k \|_X \| \tilde{x}_k \|_X \leq \| x_k \|_k \| \tilde{x}_k \|_k \approx \| L_k^{1/2} \tilde{A}_k \|_{\mathcal{C}} \| L_k^{1/2} \tilde{A}_k \|_{\mathcal{C}}$ holds for all $x_k, \tilde{x}_k \in X_k$, we obtain $\| L_k^{-1/2} A_k L_k^{-1/2} \|_{\mathcal{C}} \lesssim 1$, which finishes the proof.

Remark 3 Uzawa type smoothers have also been considered in Section 6.2 in [15]. As for the case of normal equation smoothers (distributive smoothers) due to the particular norm that was chosen in [15], again a discrete pressure Poisson problem has to be solved, cf. also Section 8 in [15].

5 Proof of the approximation property

The proof of the approximation property is done using the approximation theorem introduced in [21] which requires besides the conditions (A1) and (A1a) two more conditions (conditions (A3) and (A4)) involving, besides the Hilbert space $X$, two more Hilbert spaces: the weaker space $X_{-k} := (X_{-}, \| \cdot \|_{X_{-k}})$ and the stronger space $X_{+k} := (X_{+}, \| \cdot \|_{X_{+k}})$, which are chosen as follows.

As weaker space we choose $X_{-} := U_{-} \times P_{-}$, where $U_{-} := [L^2(\Omega)]^2$ and $P_{-} := [H^1_0(\Omega) \cap L^2(\Omega)]^*$. These Hilbert spaces are equipped with norms

$$\| x \|^2_{X_{-k}} := \| u \|^2_{H^1(\Omega) \cap [H^1_0(\Omega)]} + \| p \|^2_{L^2(\Omega)},$$

$$\| u \|^2_{H_{-k}} := h_k^{-2} \| u \|^2_{H^2(\Omega) \cap [H^1_0(\Omega)]},$$

$$\| p \|^2_{L_{-k}} := h_k^{-2} \| p \|^2_{[H^1_0(\Omega)]^*}.$$  

Remark 4 The idea behind the construction of the norm $\| \cdot \|_{X_{-k}}$ is to take the norm $\| \cdot \|_X$ and "replace all occurrences of $H^1$ by $h_k^{-1} L^2$ and all occurrences of $L^2$ by $h_k^{-1} H^{-1/2}$. This is different to the idea of constructing the norm $\| \cdot \|_k$, where "only the $H^1$-terms have been replaced by $h_k^{-1} L^2$."
So, in the present section, we show the approximation property in the norm
\[ \| \cdot \|_{X_{-k}}, \text{i.e.,} \]
\[ \| x_k^{(1)} - x^*_k \|_{X_{-k}} \leq \tilde{C}_A \sup_{\tilde{x}_k \in X_k} \frac{B \left( x_k^{(0,\nu)} - x^*_k, \tilde{x}_k \right)}{\| \tilde{x}_k \|_{X_{-k}}}. \tag{15} \]

We will show below that this version of the approximation property is stronger than the required estimate (13).

Note that dual spaces are \((X_-)^* := (U_-)^* \times (P_-)^*\), where \((U_-)^* = [L^2(\Omega)]^2\) and \((P_-)^* = H^1_0(\Omega) \cap L^2_0(\Omega)\), equipped with norms
\[ \| f \|_{(X_-)^*} = \| f \|_{(U_-)^*} + \| g \|_{(P_-)^*}^2, \]
\[ \| f \|_{(U_-)^*} = h_k^2 \| f \|_{L^2(\Omega) + \beta^{-1} H^2(\Omega)} \text{ and} \]
\[ \| g \|_{(P_-)^*} = h_k^2 \| g \|_{H^1_0(\Omega) \cap \beta L^2(\Omega)}. \]

As stronger space we choose \(X_+ := U_+ \times P_+\), where \(U_+ := [H^2(\Omega) \cap H^1_0(\Omega)]^2\) and \(P_+ := H^1(\Omega) \cap L^2_0(\Omega)\), equipped with norms
\[ \| x \|_{X_{-+}^k}^2 := \| (u, p) \|_{X_{-+}^k}^2 := \| u \|_{U_{-+}^k}^2 + \| p \|_{P_{-+}^k}^2, \]
\[ \| u \|_{U_{-+}^k}^2 := h_k^2 \| u \|_{H^2(\Omega) \cap \beta H^1(\Omega)} \text{ and} \]
\[ \| p \|_{P_{-+}^k}^2 := h_k^2 \| p \|_{H^1(\Omega) + \beta^{-1} H^2(\Omega)}. \]

For showing the approximation property, we need some auxiliary results. The first result is a standard approximation error estimate:

**Theorem 1** On all grid levels \(k\), the approximation error estimate
\[ \inf_{x_k \in X_k} \| x - x_k \|_X \lesssim \| x \|_{X_{-+}^k} \text{ for all } x \in X_+ \tag{A3} \]
is satisfied.

The proof of this theorem is rather easy as everything decouples. For completeness, we give a proof of this theorem in the appendix.

The next step is to show the following regularity result for the generalized Stokes problem. To do so, we have to introduce a standard regularity result for the Poisson problem. The following regularity statement holds, see, e.g., [3]:

**Lemma 7** On convex polygonal domains, the following regularity statement holds:

**(R1)** Regularity of the Poisson problem. Let \( g \in L^2(\Omega) \) and \( p \in H^1(\Omega) \cap L^2_0(\Omega) \) be such that
\[ (\nabla p, \nabla \tilde{p})_{H^1(\Omega)} = (g, \tilde{p})_{L^2(\Omega)} \text{ for all } \tilde{p} \in H^1(\Omega) \cap L^2_0(\Omega). \]

Then \( p \in H^2(\Omega) \) and \( \| p \|_{H^2(\Omega)} \lesssim \| g \|_{L^2(\Omega)}. \)
Now, we are able to formulate the regularity result for the generalized Stokes problem, which we are going to show next.

**Theorem 2** Suppose that the domain $\Omega$ is such that the regularity assumptions (R) and (R1) are satisfied. Then the following result for the generalized Stokes problem holds.

(A4) For all grid levels $k$ and all $F \in (X_-)^*$, the solution $x_F \in X$ of the problem,

$$\text{find } x \in X \text{ such that } B(x, \tilde{x}) = F(\tilde{x}) \quad \text{for all } \tilde{x} \in X, \quad (16)$$

satisfies $x_F \in X_+$ and the inequality

$$\|x_F\|_{X_{+,k}} \lesssim \|F\|_{(X_-)^*}.$$  

For showing this result, we need some preliminary results. Using the regularity assumption (R), we can show the following lemma.

**Lemma 8** Suppose that the assumption (R) is satisfied. Let $F \in (X_-)^*$. Then, $x_F$, the solution of (16), satisfies $x_F \in X_+$.

**Proof** Let $F(\tilde{y}, \tilde{p}) := (f, \tilde{u})_{L^2(\Omega)} + (g, \tilde{p})_{L^2(\Omega)}$ for $f \in [L^2(\Omega)]^2$ and $g \in H_0^1(\Omega) \cap L_b^2(\Omega)$. Rewrite the problem as follows:

$$(\nabla u_F, \nabla \tilde{u})_{L^2(\Omega)} + (p_F, \nabla \cdot \tilde{u})_{L^2(\Omega)} = (f, \tilde{u})_{L^2(\Omega)} - \beta(u_F, \tilde{u})_{L^2(\Omega)}$$

$$\nabla \cdot (u_F, \tilde{p})_{L^2(\Omega)} = (g, \tilde{p})_{L^2(\Omega)}$$

for all $\tilde{u} \in U$ and $\tilde{p} \in P$. As $f - \beta u_F \in [L^2(\Omega)]^2$ and $g \in H_0^1(\Omega) \cap L_b^2(\Omega)$, we obtain using regularity assumption (R) immediately that $x_F \in X_+$.

Note that the combination of the argument used in the proof of this lemma and condition (A1) immediately implies $\|x_F\|_{X_{+,k}} \lesssim (\beta)\|F\|_{(X_-)^*}$, where $(\beta)$ is a constant depending on $\beta$. For showing a robust estimate, we need to do some more work. So, we introduce two lemmas.

**Lemma 9** Suppose that $\Omega$ is such that the assumptions (R) and (R1) are satisfied. Let $F \in (X_-)^*$ and let $x_F = (u_F, p_F) \in X_+$ be the solution of (16). Then $p_F$ satisfies the estimate

$$\|p_F\|^2_{X_{+,k}} \lesssim \|u_F\|^2_{X_{+,k}} + \|F\|^2_{(X_-)^*}.$$  

**Proof** Let $F(\tilde{y}, \tilde{p}) := (f, \tilde{u})_{L^2(\Omega)} + (g, \tilde{p})_{L^2(\Omega)}$ for $f \in [L^2(\Omega)]^2$ and $g \in H_0^1(\Omega) \cap L_b^2(\Omega)$. Let $f_2 \in [H_0^1(\Omega)]^2$ be arbitrary but fix and set $f_1 := f - f_2 \in [L^2(\Omega)]^2$. We have that

$$(p_F, \nabla \cdot \tilde{u})_{L^2(\Omega)} = - (\nabla u_F, \nabla \tilde{u})_{L^2(\Omega)} - \beta(u_F, \tilde{u})_{L^2(\Omega)} + (f_1 + f_2, \tilde{u})_{L^2(\Omega)}$$

is satisfied for all $\tilde{u} \in [H_0^1(\Omega)]^2$. This is equivalent to

$$- (\nabla p_F, \tilde{u})_{L^2(\Omega)} = (\Delta u_F, \tilde{u})_{L^2(\Omega)} - \beta(u_F, \tilde{u})_{L^2(\Omega)} + (f_1 + f_2, \tilde{u})_{L^2(\Omega)}$$
for all \( \tilde{u} \in [H^1_0(\Omega)]^2 \). Note that \( \nabla p_\mathcal{F}, \Delta u_\mathcal{F}, u_\mathcal{F}, f_1 \) and \( f_2 \) are in \( [L^2(\Omega)]^2 \). Therefore, the above statement holds for all \( \tilde{u} \in [L^2(\Omega)]^2 \) (because \( [H^1_0(\Omega)]^2 \) is dense in \( [L^2(\Omega)]^2 \)), particularly for \( \tilde{u} := \nabla \tilde{p} \), where \( \tilde{p} \in H^1(\Omega) \cap L^2(\Omega) \). So, we obtain

\[
- (\nabla p_\mathcal{F}, \nabla \tilde{p})_{L^2(\Omega)} = (\Delta u_\mathcal{F}, \nabla \tilde{p})_{L^2(\Omega)} - \beta(u_\mathcal{F}, \nabla \tilde{p})_{L^2(\Omega)} + (f_1 + f_2, \nabla \tilde{p})_{L^2(\Omega)}
\]

for all \( \tilde{p} \in H^1(\Omega) \cap L^2(\Omega) \).

Let \( p_1 \in H^1(\Omega) \cap L^2(\Omega) \) be such that

\[
- (\nabla p_1, \nabla \tilde{p})_{L^2(\Omega)} = (\Delta u_\mathcal{F}, \nabla \tilde{p})_{L^2(\Omega)} + (f_1, \nabla \tilde{p})_{L^2(\Omega)}
\]

holds for all \( \tilde{p} \in H^1(\Omega) \cap L^2(\Omega) \).

Note that \( \Delta u_\mathcal{F} + f_1 \in [L^2(\Omega)]^2 \) and therefore the right-hand-side is a functional in \( [H^1_0(\Omega)]^* \). So, existence and uniqueness of \( p_1 \in H^1(\Omega) \cap L^2(\Omega) \) is guaranteed. Using the choice \( \tilde{p} := p_1 \), we obtain

\[
\| \nabla p_1 \|_{L^2(\Omega)} \leq \| \Delta u_\mathcal{F} \|_{L^2(\Omega)} + \| f_1 \|_{L^2(\Omega)} \leq \| u_\mathcal{F} \|_{H^2(\Omega)} + \| f_1 \|_{L^2(\Omega)}.
\]

Using Poincaré’s inequality, we obtain further

\[
\| p_1 \|_{H^1(\Omega)} \lesssim \| u_\mathcal{F} \|_{H^2(\Omega)} + \| f_1 \|_{L^2(\Omega)}.
\]

Let \( p_2 \in H^1(\Omega) \cap L^2(\Omega) \) be such that

\[
- (\nabla p_2, \nabla \tilde{p})_{L^2(\Omega)} = -\beta(u_\mathcal{F}, \nabla \tilde{p})_{L^2(\Omega)} + (f_2, \nabla \tilde{p})_{L^2(\Omega)}
\]

for all \( \tilde{p} \in H^1(\Omega) \cap L^2(\Omega) \). This implies, as \( u_\mathcal{F} \in [H^1_0(\Omega)]^2 \) and \( f_2 \in [H^1_0(\Omega)]^2 \), that

\[
- (\nabla p_2, \nabla \tilde{p})_{L^2(\Omega)} = \beta(\nabla \cdot u_\mathcal{F}, \tilde{p})_{L^2(\Omega)} - (\nabla \cdot f_2, \tilde{p})_{L^2(\Omega)}
\]

holds. As \( \beta \nabla \cdot u_\mathcal{F} - \nabla \cdot f_2 \in L^2(\Omega) \), existence and uniqueness of \( p_2 \) is guaranteed.

Condition (R1) implies moreover \( p_2 \in H^2(\Omega) \) and

\[
\| p_2 \|^2_{H^2(\Omega)} \lesssim \beta^2 \| \nabla \cdot u_\mathcal{F} \|^2_{L^2(\Omega)} + \| \nabla \cdot f_2 \|^2_{L^2(\Omega)} \lesssim \beta^2 \| u_\mathcal{F} \|^2_{H^1(\Omega)} + \| f_2 \|^2_{H^1(\Omega)}.
\]

Note that from (17), (18) and (20), we obtain

\[
(\nabla (p_1 + p_2), \nabla \tilde{p})_{L^2(\Omega)} = (\nabla p_\mathcal{F}, \nabla \tilde{p})_{L^2(\Omega)}
\]

is satisfied for all \( \tilde{p} \in H^1(\Omega) \cap L^2(\Omega) \), which implies (because \( p_\mathcal{F} \in L^2(\Omega) \) and \( p_1 + p_2 \in L^2(\Omega) \)) that \( p_\mathcal{F} = p_1 + p_2 \) is satisfied.

So, we have using (19) and (21)

\[
\| p_\mathcal{F} \|^2_{P_{\mathcal{F},h}} \leq h_k^2 \| p_\mathcal{F} \|^2_{H^1(\Omega) + \beta^{-1}H^2(\Omega)}
\]

\[
\leq \inf_{p_{q_1 + q_2}, q_1 \in H^1(\Omega) \cap L^2(\Omega), q_2 \in H^2(\Omega) \cup L^2(\Omega)} h_k^2 \| q_1 \|^2_{H^1(\Omega)} + h_k^2 \| q_2 \|^2_{\beta^{-1}H^2(\Omega)}
\]

\[
\leq h_k^2 \| p_1 \|^2_{H^1(\Omega)} + h_k^2 \| p_2 \|^2_{\beta^{-1}H^2(\Omega)}
\]

\[
\lesssim h_k^2 \| u_\mathcal{F} \|^2_{H^2(\Omega)} + h_k^2 \| u_\mathcal{F} \|^2_{\beta^{-1}H^1(\Omega)} + h_k^2 \| f_1 \|^2_{H^2(\Omega)} + h_k^2 \| f_2 \|^2_{\beta^{-1}H^2(\Omega)}
\]

\[
= h_k^2 \| u_\mathcal{F} \|^2_{H^2(\Omega)} + h_k^2 \| f_1 \|^2_{H^2(\Omega)} + h_k^2 \| f_2 \|^2_{\beta^{-1}H^1(\Omega)}.
\]
As $f_2 \in [H^2_0(\Omega)]^2$ was chosen arbitrary, we can take the infimum over all $f_2$, which finishes the proof.

**Lemma 10** Let $\mathcal{F} := (f, g) \in (X_\omega)^*$ and let $x_\mathcal{F} = (u_\mathcal{F}, p_\mathcal{F})$ be the solution of \((\mathcal{F})\). Then $u_\mathcal{F}$ satisfies the estimate

$$
\|u_\mathcal{F}\|_{X_{\omega+k}}^2 \lesssim \|\mathcal{F}\|_{(X_{\omega-k})^*} \|x_\mathcal{F}\|_{X_{\omega+k}}.
$$

**Proof** As $H_0^1(\Omega)$ is dense in $L^2(\Omega)$, for $u \in [H^2(\Omega)]^2$ the function $-\Delta u \in [L^2(\Omega)]^2$ can be approximated by some function $w^\epsilon \in [H^2_0(\Omega)]^2$ such that

$$
\|(-\Delta)u - w^\epsilon\|_{L^2(\Omega)}^2 \leq \epsilon.
$$

So, we can introduce an operator $-\Delta^\epsilon : [H^2(\Omega)]^2 \to [H^1_0(\Omega)]^2$ such that

$$
\|(-\Delta)u - (-\Delta^\epsilon)u\|_{L^2(\Omega)}^2 \leq \epsilon
$$

for all $u \in H^2(\Omega)$.

Analogously, we introduce the operator $\nabla^\epsilon : H^1(\Omega) \to [H^1_0(\Omega)]^2$ such that

$$
\|\nabla p - \nabla^\epsilon p\|_{L^2(\Omega)}^2 \leq \epsilon
$$

for all $p \in H^1(\Omega)$.

The idea of this proof is to show that for all $\epsilon > 0$ there is some $\tilde{x}^\epsilon \in X$ such that

$$
\begin{align*}
    &h^{-2}_k \|u_\mathcal{F}\|_{X_{\omega+k}}^2 - (1 + \beta^{1/2} + \beta)h^{-1}_k \|x_\mathcal{F}\|_{X_{\omega+k}} - \epsilon^2 \\
    &\lesssim \mathcal{B}(x_\mathcal{F}, \tilde{x}^\epsilon) \\
    &\lesssim h^{-2}_k \|\mathcal{F}\|_{(X_{\omega-k})^*} \|x_\mathcal{F}\|_{X_{\omega+k}} + (1 + \beta^{1/2})h^{-1}_k \|\mathcal{F}\|_{(X_{\omega-k})^*} + \epsilon^2
\end{align*}
$$

(22)

holds. The statement of the Lemma follows for $\epsilon$ approaching 0.

In the following, we show that (22) and (23) are satisfied for the choice $\tilde{x}^\epsilon := (-\Delta^\epsilon u_\mathcal{F}, \nabla^\epsilon \cdot \nabla^\epsilon p_\mathcal{F})$.

First we show (22) by estimating the individual summands of $\mathcal{B}(x_\mathcal{F}, \tilde{x}^\epsilon)$ separately.

$$
\beta(u_\mathcal{F}, -\Delta u_\mathcal{F})_{L^2(\Omega)} \geq \beta(u_\mathcal{F}, -\Delta u_\mathcal{F})_{L^2(\Omega)} - \epsilon \beta \|u_\mathcal{F}\|_{L^2(\Omega)} \\
= \beta(\nabla u_\mathcal{F}, \nabla u_\mathcal{F})_{L^2(\Omega)} - \epsilon \beta \|u_\mathcal{F}\|_{L^2(\Omega)} \geq \beta \|u_\mathcal{F}\|_{H^1(\Omega)}^2 - \epsilon \beta h^{-1}_k \|x_\mathcal{F}\|_{X_{\omega+k}}
$$

is satisfied due to $u_\mathcal{F} \in H^2(\Omega) \cap H^1_0(\Omega)$ and Friedrichs’ inequality.

For the next summand, we obtain

$$
(\nabla u_\mathcal{F}, \nabla(-\Delta^\epsilon)u_\mathcal{F})_{L^2(\Omega)} = (\Delta u_\mathcal{F}, \Delta^\epsilon u_\mathcal{F})_{L^2(\Omega)} \\
\geq (\Delta u_\mathcal{F}, \Delta u_\mathcal{F})_{L^2(\Omega)} - \epsilon \|\Delta u_\mathcal{F}\|_{L^2(\Omega)} \geq \|u_\mathcal{F}\|_{H^2(\Omega)}^2 - \epsilon \beta^{1/2} h^{-1}_k \|x_\mathcal{F}\|_{X_{\omega+k}}
$$

due to the fact that $\Delta^\epsilon$ maps into $H^1_0(\Omega)$. Moreover we use $u_\mathcal{F} \in H^2(\Omega) \cap H^1_0(\Omega)$ and Friedrichs’ inequality.
For the next two summands, we obtain

\[(\nabla \cdot u_{f}, \nabla \cdot \nabla^2 p_{f})_{L^2(\Omega)} + (\nabla \cdot (-\Delta^e)u_{f}, p_{f})_{L^2(\Omega)}\]

\[= -(\nabla \nabla \cdot u_{f}, \nabla^e p_{f})_{L^2(\Omega)} + (\Delta u_{f}, \nabla p_{f})_{L^2(\Omega)}\]

\[\geq -(\nabla \nabla \cdot u_{f}, \nabla^e p_{f})_{L^2(\Omega)} + (\Delta u_{f}, \nabla^e p_{f})_{L^2(\Omega)} - \epsilon \|\Delta^e u_{f}\|_{L^2(\Omega)}\]

\[\geq (\nabla u_{f}, \nabla^2 p_{f})_{L^2} + (\Delta u_{f}, \nabla^e p_{f})_{L^2} - \epsilon (\|\Delta u_{f}\|_{L^2} + \|\nabla^e p_{f}\|_{L^2} + \epsilon)\]

\[= (\nabla u_{f}, \nabla^e p_{f})_{L^2} + (\nabla \cdot u_{f}, \nabla^e p_{f})_{L^2} - \epsilon (\|\Delta u_{f}\|_{L^2} + \|\nabla^e p_{f}\|_{L^2} + \epsilon)\]

\[\geq -\epsilon (\|\Delta u_{f}\|_{L^2(\Omega)} + \|\nabla^e p_{f}\|_{L^2(\Omega)} + \epsilon) \geq -\epsilon (L_{\Omega}^{\beta^1/2}(\Omega) + \|\nabla^e p_{f}\|_{H^1(\Omega)} + \epsilon)\]

\[\geq -\epsilon (1 + \beta^{1/2})h_{k}^{-1} \|x_{f}\|_{X_{\alpha+\beta^1/2}} - \epsilon^2.\]

This shows (22).

The next step is to show (23) again for the choice \(\tilde{x}^e := (-\Delta^e u_{f}, \nabla \cdot \nabla^e p_{f})\).

Let \(f_2 \in [H_0^{1}(\Omega)]^2\) and \(f_1 := f - f_2\). Then

\[(f_1, -\Delta^e u_{f})_{L^2(\Omega)} \leq \|f_1\|_{L^2(\Omega)}\|u_{f}\|_{H^1(\Omega)} + \epsilon \|f_1\|_{L^2(\Omega)}\]

holds as well as

\[(f_2, -\Delta^e u_{f})_{L^2(\Omega)} \leq (\nabla f_2, \nabla u_{f})_{L^2(\Omega)} + \epsilon \|f_2\|_{L^2(\Omega)}\]

\[\leq \|f_2\|_{\beta^{-1}H^1(\Omega)}\|u_{f}\|_{\beta H^1(\Omega)} + \epsilon \beta^{1/2} \|f_2\|_{\beta^{-1}H^1(\Omega)}.\]

This implies

\[(f, -\Delta^e u_{f})_{L^2(\Omega)}\]

\[\leq \|f\|_{L^2(\Omega) + \beta^{-1}H^1(\Omega)}\|u_{f}\|_{H^1(\Omega) + \beta H^1(\Omega)} + \epsilon (1 + \beta^{1/2}) \|f\|_{L^2(\Omega)} + \beta^{-1}H^1(\Omega).\]

Let \(p_2 \in H^2(\Omega)\) and \(p_1 := p - p_2\). We have

\[(g, \nabla \cdot \nabla^e p_1)_{L^2(\Omega)} = -(\nabla g, \nabla^e p_1)_{L^2(\Omega)} \leq \|g\|_{H^1(\Omega)} \|p_1\|_{H^1(\Omega)} + \epsilon \|g\|_{H^1(\Omega)}\]

Moreover, using \(g \in H^1_0(\Omega)\), we have also

\[(g, \nabla \cdot \nabla^e p_2)_{L^2(\Omega)} = -(\nabla g, \nabla^e p_2)_{L^2(\Omega)} \leq \|g\|_{H^1(\Omega)}\]

\[= (g, \nabla \cdot \nabla^2 p_2)_{L^2(\Omega)} + \epsilon \|g\|_{H^1(\Omega)} \leq \|g\| \|\beta \nabla^2 p_2\|_{L^2(\Omega)} + \epsilon \|g\|_{H^1(\Omega)}\]

and therefore

\[(g, \nabla \cdot \nabla^e p_{f})_{L^2(\Omega)} \leq \|g\|_{H^2(\Omega) + \beta H^1(\Omega)} \|p_{f}\|_{H^1(\Omega) + \beta^{-1}H^2(\Omega)} + \epsilon \|g\|_{H^1(\Omega)}\]

\[\leq \|g\|^{\beta^1/2}_{H^2(\Omega) + \beta H^1(\Omega)} \|p_{f}\|_{H^1(\Omega) + \beta^{-1}H^2(\Omega)} + \epsilon \beta^{-1} \|F\|_{X_{\alpha+\beta^1/2}}.\]

By combining these results, we immediately obtain (23). This finishes the proof.

Now, we can show Theorem 2.
Proof (of Theorem 3). Note that Lemma 9 already states that \( x_F \in X^+_{k} \). It remains to show
\[
\|x_F\|_{X^+_{k}} \lesssim \|F\|_{(X^-_{k})^*}
\]  
(24)
Note that Lemma 9 implies immediately
\[
\|x_F\|_{X^+_{k}}^2 = \|u_F\|_{U^+_{k}}^2 + \|p_F\|_{P^+_{k}}^2 \lesssim \|u_F\|_{U^+_{k}}^2 + \|F\|_{(X^-_{k})^*}^2
\]
If we combine this result with the statement of Lemma 10, we obtain
\[
\|x_F\|_{X^+_{k}} \leq C \left( \|x_F\|_{X^-_{k}} \|F\|_{(X^-_{k})^*} + \|F\|_{(X^-_{k})^*}^2 \right)
\]
for some constant \( C > 0 \) (independent of \( k \) and \( \beta \)) which implies
\[
\|x_F\|_{X^+_{k}} \leq \frac{1}{2} \left( C + \sqrt{4C + C^2} \right) \|F\|_{(X^-_{k})^*},
\]
and further (24). This finishes the proof.

As we have shown that the conditions (A1), (A1a), (A3) and (A4) are satisfied, we can apply Theorem 4.1 in [21] and obtain.

Theorem 3 Assume that \( \Omega \) is such that (R) and (R1) are satisfied and assume that the problem is discretized using the Taylor-Hood element. Then the coarse-grid correction (10) satisfies the approximation property (15) where \( C_A \) only depends on the constants that appear in the conditions (A1), (A1a), (A3) and (A4).

Using an estimate
\[
\|x_k\|_{k} \lesssim \|x_k\|_{X^-_{k}}
\]
holds for all \( x_k \in X_k \),  
(25)
the approximation property (13), i.e.,
\[
\|x^{(1)}_k - x^*_k\|_{k} \leq C_A \sup_{x_k \in X_k} B \left( x_k^{(0,\nu)} - x^*_k, \bar{x}_k \right) \|\bar{x}_k\|_{k}
\]
is a direct consequence of (15). So, the final step is to show (25).

Lemma 11 The estimate (25) is satisfied.

Proof The analysis can be done component-wise. So it suffices to show that
\[
\|u_k\|_{U,k} \lesssim \|u_k\|_{U^-_{k}} \quad \text{and} \quad \|p_k\|_{P,k} \lesssim \|p_k\|_{P^-_{k}}
\]
(26)
is satisfied for all \( u_k \in U_k \) and all \( p_k \in P_k \).

For the proof, we define on each grid level the functions \( \psi_k : \Omega \rightarrow \mathbb{R} \) as follows. The function \( \psi_k \) is a continuous function that is linear on each element (cf. the Courant element). The function \( \psi_k \) is defined to be 0 on all vertices which are located on \( \partial \Omega \) and to be 1 on all other vertices (interior vertices).

In what follows, \( \bar{P}_k \) is the space of functions obtained by discretizing the space \( H^1(\Omega) \) using the standard Courant element. Note that – as we have used
the Taylor Hood element – the identity \( \hat{P}_k = \{ p_k + a : p_k \in P_k \text{ and } a \in \mathbb{R} \} \) holds.

First, we start with some preliminary inequalities. We note that \( \psi_k \leq 1 \) holds on the whole domain \( \Omega \) and therefore the inequality

\[
\| \psi_k p_k \|_{L^2(\Omega)}^2 \leq \| p_k \|_{L^2(\Omega)}^2
\]

holds for all \( p_k \in \hat{P}_k \). Moreover, the inequalities

\[
\| \psi_k p_k \|_{H^1(\Omega)}^2 \leq h_k^{-2} \| p_k \|_{L^2(\Omega)}^2
\]

and

\[
(p_k, \psi_k p_k)_{L^2(\Omega)} \geq (p_k, p_k)_{L^2(\Omega)}
\]

hold for all \( p_k \in \hat{P}_k \). The proofs of these two inequalities are rather technical. For completeness, we gave proofs for these inequalities in the Appendix.

Now we can show the two inequalities in (26). We obtain using the definition of \( \| \cdot \|_{U_{-1, k}} \), the choice \( \tilde{u} := u_k \in [H_0^1(\Omega)]^2 \) and a standard inverse inequality

\[
\| u_k \|_{U_{-1, k}}^2 = \sup_{0 \neq \tilde{u} \in [L^2(\Omega)]^2} \frac{(u_k, \tilde{u})_{L^2(\Omega)}}{h_k^2 \| \tilde{u} \|_{L^2(\Omega) + \beta^{-1} H_0^1(\Omega)}^2} \geq \frac{(u_k, \tilde{u})_{L^2(\Omega)}}{h_k^2 \| u_k \|_{L^2(\Omega) + \beta^{-1} H_0^1(\Omega)}^2}
\]

\[
= \inf_{w \in [H_0^1(\Omega)]^2} \frac{h_k^2 \| w - u_k \|_{L^2(\Omega)}^2}{(u_k, \tilde{u})_{L^2(\Omega)}} \geq \inf_{w \in U_k} \frac{h_k^2 \| w - u_k \|_{L^2(\Omega)}^2}{(u_k, \tilde{u})_{L^2(\Omega)}} \geq \inf_{w \in U_k} \frac{h_k^2 \| w - u_k \|_{L^2(\Omega)}^2 + \| w \|_{H^1(\Omega)}^2}{(u_k, \tilde{u})_{L^2(\Omega)}}
\]

\[
= (\beta + h_k^{-2}) \| u_k \|_{L^2(\Omega)}^2 = \| u_k \|_{U_{-1, k}}^2,
\]

i.e., the first inequality.

Note that we obtain using the definition of \( \| \cdot \|_{P_{-1, k}} \) and the fact that \( p_k \in P_k \subseteq L_0^2(\Omega) \) that

\[
\| p_k \|_{P_{-1, k}}^2 = \sup_{0 \neq q \in H_0^1(\Omega) \cap L^2(\Omega)} \frac{(p_k, q)_{L^2(\Omega)}}{h_k^2 \| q \|_{H_0^1(\Omega) + \beta L^2(\Omega)}^2}
\]

\[
= \sup_{0 \neq q \in H_0^1(\Omega) \cap L^2(\Omega)} \frac{(p_k - a, q)_{L^2(\Omega)}}{h_k^2 \| q \|_{H_0^1(\Omega) + \beta L^2(\Omega)}^2}
\]

is satisfied, where \( a \in \mathbb{R} \) is such that \( \psi_k \ast (p_k - a) \in L_0^2(\Omega) \).

By plugging in \( q := \psi_k \ast (p_k - a) \), we obtain using (27), (28) and (29) (note that \( p_k \in P_k \) implies \( p_k - a \in \hat{P}_k \) that

\[
\| p_k \|_{P_{-1, k}}^2 \geq \frac{(p_k - a, p_k - a)_{L^2(\Omega)}}{\| p_k - a \|_{L^2(\Omega)}^2 + \beta h_k^2 \| p_k - a \|_{L^2(\Omega)}^2}
\]

\[
= (1 + \beta h_k^{-2})^{-1} \| p_k - a \|_{L^2(\Omega)}^2 \geq (1 + \beta h_k^{-2})^{-1} \| p_k \|_{L^2(\Omega)}^2 = \| p_k \|_{P_{-1, k}}^2.
\]
where the last inequality is satisfied due to the fact that \( p_k \in P_k \subseteq L_2^0(\Omega) \). So, this shows the second inequality and finishes the proof.

So, we have shown the approximation property (13). So, we obtain as follows.

**Theorem 4** Assume that

- \( \Omega \) is such that the regularity assumptions (R) and (R1) are satisfied,
- the problem is discretized using the Taylor-Hood element and
- one of the smoothers proposed in this paper is used.

Then the two-grid method converges if sufficiently many smoothing steps are applied. Namely, we have

\[
\| x^{(1)}_k - x^*_k \|_k \leq q(\nu) \| x^{(0)}_k - x^*_k \|_k,
\]

with \( q(\nu) := C_S C_A \nu^{-1/2} \), where the constants \( C_A \) and \( C_S \) are independent of the grid level \( k \) and the choice of the parameter \( \beta \).

The convergence of the W-cycle multigrid method follows under weak assumptions, cf. [9].

### 6 Numerical Results

In this section, we illustrate the convergence theory presented within this paper with numerical results. The numerical experiments were done as follows.

For the numerical experiments, the domain \( \Omega \) was chosen to be the unit square \( \Omega := (0,1)^2 \). As mentioned in Section 2, the weak inf-sup-condition (6) can be shown for the Taylor-Hood element only if at least one vertex of each element is in the interior of the domain \( \Omega \). As this is not satisfied for the standard decomposition of the unit square into two triangular elements, we choose the coarsest grid level \( k = 0 \) to be a decomposition of the domain \( \Omega \) into 8 triangles, as seen in Figure 1. The grid levels \( k = 1, 2, \ldots \) were constructed by uniform refinement, i.e., every triangle was decomposed into four subtriangles.

![Fig. 1 Discretization on grid levels k = 1 and k = 2, where the squares denote the degrees of freedom of (the components of) the velocity field \( u \) and the dots denote the degrees of freedom of pressure distribution \( p \)](image)
The right-hand-side functions $f$ and $g$ have been chosen such that the solution of the problem on grid level $k$ is the $L^2$-projection of the exact solution\[ u(\xi_1, \xi_2) := \phi(\xi_1, \xi_2) \left( \frac{\xi_2}{2} - \frac{1}{2} \right) \quad \text{and} \quad p(\xi_1, \xi_2) := \phi(\xi_1, \xi_2), \]
where \[ \phi(\xi_1, \xi_2) := \max \left\{ 0, \min \left\{ 1, 2 - 4\sqrt{(\xi_1 - \frac{1}{2})^2 + (\xi_2 - \frac{1}{2})^2} \right\} \right\}, \]
into $X_k$.

For solving the discretized problem, we have used the proposed W-cycle multigrid method. To obtain a proper scaling, the following choice of matrix $L_k$ was taken for the numerical experiments\[ L_k := \begin{pmatrix} \hat{A}_k & \hat{S}_k \end{pmatrix}, \quad \text{where} \quad \hat{A}_k := \text{diag } A_k \quad \text{and} \quad \hat{S}_k := \text{diag } B_k \hat{A}_k^{-1} B_k^T. \]

Note that the matrix $L_k$, introduced above, is spectrally equivalent to the matrix $L_k$, introduced in Section 3. So, the choice introduced above is still covered by the convergence theory. The damping parameter is chosen to be $\tau = 0.35$, for all grid levels $k$ and all choices of $\beta$.

For the Uzawa type smoother, the matrices $\hat{A}_k$ and $\hat{S}_k$ have been chosen as introduced in (30), and the damping parameters have been chosen to be $\tau := \sigma := 0.8$. Note that the matrices $A_k$ and $S_k$, introduced in this section, are spectrally equivalent to the choices introduced in Section 3. So, the choice introduced above is still covered by the convergence theory.

The number of iterations and the convergence rate were measured as follows: we start with $x_k^{(0)}$ and measure the reduction of the error in each step using the norm $\| \cdot \|_k$. The iteration was stopped when the initial error was reduced by a factor of $\epsilon = 10^{-9}$. The convergence rate $q$ is the mean convergence rate in this iteration, i.e.,\[ q = \left( \frac{\| x_k^{(n)} - x_k^* \|_k}{\| x_k^{(0)} - x_k^* \|_k} \right)^{1/n}, \]
where $n$ is the number of iterations needed to reach the stopping criterion. Here, $x_k^*$ is the exact solution and $x_k^{(i)}$ is the $i$-th iterate.

In Table 1 we compare for a fixed grid level (level $k = 4$) and a fixed choice $\beta = 1$ the convergence rates for several choices of $\nu$, the number of pre- and post-smoothing steps. We see that the convergence rate behaves approximately like $\nu^{-1/2}$, if the number of smoothing steps is increased. This is consistent with the theory which guarantees the convergence rate being bounded by $C \nu^{-1/2}$. We observe that the preconditioned normal equation smoother already converges for $\nu = 1 + 1$ smoothing steps, while for the Uzawa type smoother $\nu = 3 + 3$ smoothing steps are necessary.
A robust multigrid method for the time-dependent Stokes problem

\[ \nu = 1 + 1 \quad \nu = 2 + 2 \quad \nu = 3 + 3 \quad \nu = 4 + 4 \quad \nu = 8 + 8 \quad \nu = 16 + 16 \]

| \( n \) | \( q \) | \( n \) | \( q \) | \( n \) | \( q \) | \( n \) | \( q \) | \( n \) | \( q \) |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 88    | 0.789 | 46    | 0.634 | 30    | 0.496 | 24    | 0.412 | 17    | 0.293 | 11    | 0.148 |

**Normal equation smoother**

| \( \beta = 1 \) | \( \beta = 10^2 \) | \( \beta = 10^4 \) | \( \beta = 10^6 \) | \( \beta = 10^8 \) | \( \beta = 10^{10} \) |
|----------|----------|----------|----------|----------|----------|
| \( k = 4 \) | 30        | 0.496    | 30        | 0.493    | 30        | 0.496    | 68    | 0.736 | 71    | 0.745 | 71    | 0.745 |
| \( k = 5 \) | 29        | 0.489    | 29        | 0.488    | 22        | 0.388    | 64    | 0.722 | 70    | 0.744 | 71    | 0.745 |
| \( k = 6 \) | 29        | 0.484    | 29        | 0.486    | 27        | 0.457    | 52    | 0.670 | 70    | 0.743 | 71    | 0.745 |
| \( k = 7 \) | 28        | 0.475    | 28        | 0.475    | 28        | 0.470    | 35    | 0.553 | 70    | 0.742 | 71    | 0.745 |
| \( k = 8 \) | 28        | 0.469    | 28        | 0.469    | 28        | 0.468    | 20    | 0.347 | 67    | 0.732 | 71    | 0.746 |

**Uzawa type smoother**

| \( \beta = 1 \) | \( \beta = 2 \) | \( \beta = 4 \) | \( \beta = 8 \) |
|----------|----------|----------|----------|
| \( k = 4 \) | 14        | 0.212    | 13        | 0.201    | 6        | 0.030    | 7        | 0.051    | 8        | 0.059    | 8        | 0.059    |
| \( k = 5 \) | 13        | 0.194    | 13        | 0.193    | 9        | 0.095    | 7        | 0.043    | 7        | 0.050    | 7        | 0.050    |
| \( k = 6 \) | 12        | 0.176    | 12        | 0.176    | 11       | 0.145    | 6        | 0.024    | 7        | 0.047    | 7        | 0.048    |
| \( k = 7 \) | 12        | 0.166    | 12        | 0.166    | 11       | 0.150    | 5        | 0.013    | 7        | 0.044    | 7        | 0.045    |
| \( k = 8 \) | 11        | 0.147    | 11        | 0.147    | 11       | 0.145    | 8        | 0.058    | 7        | 0.038    | 7        | 0.043    |

**Table 1** Number of iterations \( n \) and convergence rate \( q \) for the normal equation smoother and the Uzawa type smoother depending with \( \nu = \nu_{\text{pre}} + \nu_{\text{post}} \) smoothing steps on grid level \( k = 4 \) for \( \beta = 1 \)

**Table 2** Number of iterations \( n \) and convergence rate \( q \) for the normal equation smoother with \( \nu = 3 + 3 \) smoothing steps

**Table 3** Number of iterations \( n \) and convergence rate \( q \) for the Uzawa type smoother with \( \nu = 3 + 3 \) smoothing steps

In Tables 2 and 3 we compare various grid levels \( k \) and choices of the parameter \( \beta \). We have used a fixed choice of \( \nu = 3 + 3 \) smoothing steps. First we observe that, for both smoothers, the number of iterations seems to be well-bounded for all grid levels \( k \) which yields an optimal convergence behavior. Moreover, we see that the number of iterations is also well-bounded for a wide range of choices of the parameter \( \beta \), i.e., we observe also robust convergence as predicted by the convergence theory.

Comparing both kinds of smoothers, we see that the Uzawa type smoother leads to much faster convergence rates than the preconditioned normal equation smoother. Note that moreover the computational complexity of the Uzawa
type smoother (per iteration) is slightly smaller than the complexity of the normal equation smoother.

It has to be mentioned that for the model problem, also the (more efficient) V-cycle multigrid method converges with rates comparable to the convergence rates of the W-cycle multigrid method. However, the V-cycle is not covered by the convergence theory.

The numerical experiments done by the author have shown that the convergence rates can be improved slightly by adjusting the choice of the parameters to the grid levels and the choice of $\beta$. However, the main goal of this paper is to show that the proposed method also works well for fixed choices of the parameter.

7 Conclusions and Further Work

In the present paper we have proposed a coupled multigrid solver for the generalized Stokes problem where the smoothing property is needed in the scaled $L^2$-norm $\|\cdot\|_k$. This allows to construct a multigrid method where the smoother is a simple linear iteration (which consists only of divisions and the multiplication of vectors with the system matrix $A_k$). In the present paper, a preconditioned normal equation smoother and an Uzawa type smoother have been chosen but it seems possible to find also other simple smoothers which satisfy the smoothing property in the norm $\|\cdot\|_k$.

The convergence rates observed for the multigrid method proposed in the present paper are comparable with the rates observed for the methods proposed in [15]. Note that for applying the methods proposed in the named paper, it is necessary to solve a Poisson problem in each smoothing step. This is not needed for the method proposed in the present paper. The main contribution of this paper is a new way of setting up the norms where the convergence is shown in. The technique of the convergence proof that has been applied in the present paper is also extendable to the Stokes control problem, cf. [20].

Acknowledgments. The author thanks Markus Kollmann for providing code used to compute the numerical results presented in this paper. Moreover, the support of the University of Oxford and the Technische Universität Chemnitz are gratefully acknowledged.

8 Appendix

Proof (of Theorem 7). Note that it suffices to show approximation error results for the individual variables separately. Using a standard interpolation operator $\Pi_k : [H^2(\Omega)]^2 \rightarrow U_k$, we obtain for the velocity field

$$\|u - \Pi_k u\|_{L^2(\Omega)}^2 \lesssim h_k^2 \|u\|_{H^1(\Omega)}^2$$

and

$$\|u - \Pi_k u\|_{H^1(\Omega)}^2 \lesssim h_k^2 \|u\|_{H^2(\Omega)}^2,$$
for all $u \in [H^2(\Omega)]^2$ and therefore

$$\inf_{u_k \in U_k} \|u - u_k\|_{L^2}^2 \leq \|u - \Pi_k u\|_{L^2}^2 = \|u - \Pi_k u\|_{H^1(\Omega)}^2 + \beta \|u - \Pi_k u\|_{H^2_{\text{v}}(\Omega)}^2 \leq h_k^2 \left(\|u\|_{H^2_{\text{v}}(\Omega)} + \beta \|u\|_{H^1(\Omega)}\right) = \|u\|_{L^2_{\text{v},h}}^2.$$  

Also for the pressure distribution we can do a similar estimate. The estimates

$$\inf_{p_k \in P_k} \|p - p_k\|_{L^2_{\text{v}}(\Omega)}^2 \lesssim h_k^2 \|p\|_{H^1(\Omega)}^2 \quad \text{and} \quad \inf_{p_k \in P_k} \|p - p_k\|_{H^1(\Omega)}^2 \lesssim h_k^2 \|p\|_{H^2_{\text{v}}(\Omega)}^2$$

are standard approximation error results which imply

$$\inf_{p_k \in P_k} \|p - p_k\|_{T}^2 = \inf_{p_k \in P_k} \|p - p_k\|_{L^2_{\text{v}}(\Omega) + \beta^{-1}H^1(\Omega)}^2 = \inf_{p_k \in P_k} \|q_1\|_{L^2(\Omega)}^2 + \|q_2\|_{H^1(\Omega)}^2 = \inf_{p_k \in P_k} \|p - p_k\|_{L^2_{\text{v}}(\Omega)}^2 + \|p - p_k\|_{H^1(\Omega)}^2 \leq \inf_{p_k \in P_k} \|p - p_k\|_{L^2_{\text{v}}(\Omega)}^2 + \|p - p_k\|_{H^1(\Omega)}^2 \leq h_k^2 \inf_{p_k \in P_k} \|p - p_k\|_{L^2_{\text{v}}(\Omega)}^2 + \|p - p_k\|_{H^1(\Omega)}^2 = h_k^2 \inf_{p_k \in P_k} \|p - p_k\|_{H^2_{\text{v}}(\Omega) + \beta^{-1}H^2(\Omega)}^2,$$

for all $p \in H^1(\Omega) \cap L^2_{\text{v}}(\Omega)$, which finishes the proof.

Proof (of of inequality (28)). First note that

$$\|\psi_k p_k\|_{H^1(\Omega)}^2 = \|\nabla(\psi_k p_k)\|_{L^2(\Omega)}^2 + \|\psi_k p_k\|_{L^2(\Omega)}^2 \lesssim \|\nabla(\psi_k p_k)\|_{L^2(\Omega)}^2 + h_k^{-2} \|p_k\|_{L^2(\Omega)}^2$$

is satisfied due to (27) and $h_k \lesssim 1$. So, it suffices to bound $\|\nabla(\psi_k p_k)\|_{L^2(\Omega)}^2$ from above.

Assume that the mesh consists of the elements $T_i$ for $i = 1, \ldots, M_k$. First note that

$$\|p_k\|_{L^2(\Omega)}^2 = \sum_{i=1}^{M_k} \|p_k\|_{L^2(T_i)}^2 \quad \text{and} \quad \|\nabla(\psi_k p_k)\|_{L^2(\Omega)}^2 = \sum_{i=1}^{M_k} \|\nabla(\psi_k p_k)\|_{L^2(T_i)}^2$$

is satisfied.

So it suffices to show

$$\|\nabla(\psi_k p_k)\|_{L^2(T_i)}^2 \lesssim h_k^{-2} \|p_k\|_{L^2(T_i)}^2$$

for all $i = 1, \ldots, M_k$. 

Note that on all interior elements (no vertex is located on $\partial \Omega$) we have $\psi_k = 1$. So, the inequality (31) is a standard inverse inequality. So, it remains to show (31) on all elements where one or two vertices are located on $\partial \Omega$. (Note that we have assumed already in Section 2 that each element has at least one vertex in the interior of $\Omega$.)

We have

$$
\|\nabla (\psi_k(\xi)p_k(\xi))\|^2_{L^2(T_i)} = \int_{T_i} \|\nabla (\psi_k(\xi)p_k(\xi))\|^2_{L^2} \, d\xi
= \int_{T_i} \| (\nabla \psi_k(\xi))p_k(\xi) + \psi_k(\xi)(\nabla p_k(\xi))\|^2_{L^2} \, d\xi
\leq \int_{T_i} \|\nabla \psi_k(\xi)\|^2_{L^2} \|p_k(\xi)\|^2_{L^2} \, d\xi + \int_{T_i} \psi_k(\xi)^2 \|\nabla p_k(\xi)\|^2_{L^2} \, d\xi.
$$

(32)

For estimating the first summand, $\int_{T_i} \|\nabla \psi_k(\xi)\|^2_{L^2} \|p_k(\xi)\|^2_{L^2} \, d\xi$, we use the fact that $\psi_k$ is linear and therefore $\nabla \psi_k$ is constant. So, we obtain

$$
\int_{T_i} \|\nabla \psi_k(\xi)\|^2_{L^2} \|p_k(\xi)\|^2_{L^2} \, d\xi \leq \|\nabla \psi_k\|^2_{L^2} \|p_k\|^2_{L^2(\Omega)}.
$$

Note that $\psi_k$ takes the value 0 on two vertices of the element, say $P_1$ and $P_2$, and the value 1 on the third vertex $P_3$ (the case that $\psi_k$ takes the value 1 on two vertices, say $P_1$ and $P_2$ and the value 0 on the third vertex is completely analogous). It is geometrically evident that $\|\nabla \psi_k\|_{L^2}$ is equal to the reciprocal of the length of the altitude on the edge $P_1P_2$. The reciprocal of the length of the altitude is bounded from above by $h_k^{-1}$. This shows

$$
\int_{T_i} \|\nabla \psi_k(\xi)\|^2_{L^2} \|p_k(\xi)\|^2_{L^2} \, d\xi \leq h_k^{-2} \|p_k\|^2_{L^2(\Omega)}.
$$

The second summand in (32), $\int_{T_i} \psi_k(\xi)^2 \|\nabla p_k(\xi)\|^2_{L^2} \, d\xi$, can be bounded from above using $\psi_k(\xi)^2 \leq 1$ by $\|\nabla p_k\|^2_{L^2(\Omega)}$, which can be bounded from above by $h_k^{-2} \|p_k\|^2_{L^2(\Omega)}$ using a standard inverse inequality. This finishes the proof.

**Proof (of inequality (29)).** Assume that the mesh consists of the elements $T_i$ for $i = 1, \ldots, M_k$. First note that

$$(p_k, \psi_k p_k)_{L^2(\Omega)} = \sum_{i=1}^{M_k} (p_k, \psi_k p_k)_{L^2(T_i)} \quad \text{and} \quad (p_k, p_k)_{L^2(\Omega)} = \sum_{i=1}^{M_k} (p_k, p_k)_{L^2(T_i)}.$$

So, it suffices to show (29) for the individual elements $T_i$, i.e.,

$$(p_k, p_k)_{L^2(T_i)} \lesssim (p_k, \psi_k p_k)_{L^2(T_i)} \quad (33)$$

for all $p_k \in \tilde{P}_k$. Note that on all interior elements (no vertex is located on $\partial \Omega$) we have $\psi_k = 1$. So, the inequality (33) is obviously satisfied. So, it remains to show (31) on all elements where one (case 1) or two vertices (case 2) are
located on \( \partial \Omega \). (Note that we have assumed already in Section 2 that each element has at least one vertex in the interior of \( \Omega \)).

**Case 1.** The statement which we have to show reads as follows:

\[
\int_{\Gamma_i} \hat{p}_k^2(\xi) \, d\xi \lesssim \int_{\Gamma_i} \hat{\psi}_k(\xi) \, \hat{p}_k^2(\xi) \, d\xi. \tag{34}
\]

Note that, both, \( p_k(\xi) \) and \( \psi_k(\xi) \), are linear functions. These integrals can be computed using the reference element \( \Delta := \{(\xi_1, \xi_2) \in (0,1)^2 : \xi_1 + \xi_2 < 1\} \).

It is well known that there is a linear transformation \( \Phi_i : \Delta \rightarrow \Gamma_i \). Using the standard substitution rule, we obtain that (34) is equivalent to

\[
\int_{\Delta} \hat{p}_k^2(\xi) \, |\det \nabla^2 \Phi_i(\xi)| \, d\xi \lesssim \int_{\Delta} \hat{\psi}_k(\xi) \, \hat{p}_k^2(\xi) \, |\det \nabla^2 \Phi_i(\xi)| \, d\xi,
\]

where \( \hat{p}(\xi) = u(\Phi_i^{-1}(\xi)) \) and \( \hat{\psi}(\xi) = \psi(\Phi_i^{-1}(\xi)) \) and \( |\det \nabla^2 \Phi_i(\xi)| \) is the absolute value of the Jacobi determinant of the transformation. Because the transformation is linear, both, \( \hat{p} \) and \( \hat{\psi} \), are linear functions and, moreover, the Jacobi determinant is a constant. So it suffices to show

\[
\int_{\Delta} \hat{p}_k^2(\xi) \, d\xi \lesssim \int_{\Delta} \hat{\psi}_k(\xi) \, \hat{p}_k^2(\xi) \, d\xi, \tag{35}
\]

for all linear functions \( \hat{p}_k \). As mentioned above, the function \( \hat{\psi}_k \) takes the value 1 on two vertices and the value 0 on one vertex. We assume without loss of generality, that it takes the value 0 on \((0,1)\). This directly implies that \( \hat{\psi}_k(\xi) = 1 - \xi_2 \).

Using a general approach for \( \hat{p}_k \), like \( \hat{p}_k(\xi_1, \xi_2) := a_0 + a_1 \xi_1 + a_2 \xi_2 \), we can compute both integrals in (35) and obtain for \( a := (a_0, a_1, a_2)^T \) that

\[
\int_{\Delta} \hat{p}_k^2(\xi) \, d\xi = a^T F a \quad \text{with} \quad F := \frac{1}{24} \begin{pmatrix} 12 & 4 & 4 \\ 4 & 2 & 1 \\ 4 & 1 & 2 \end{pmatrix}
\]

\[
\int_{\Delta} \hat{\psi}_k(\xi) \, \hat{p}_k^2(\xi) \, d\xi = a^T G a \quad \text{with} \quad G := \frac{1}{120} \begin{pmatrix} 40 & 15 & 10 \\ 15 & 8 & 3 \\ 10 & 3 & 4 \end{pmatrix}.
\]

Obviously, we obtain

\[
\frac{\int_{\Delta} \hat{p}_k^2(\xi) \, d\xi}{\int_{\Delta} \hat{\psi}_k(\xi) \, \hat{p}_k^2(\xi) \, d\xi} \leq \lambda_{\text{max}}(G^{-1}F) = \frac{1}{3} \left( 6 + \sqrt{6} \right),
\]

where \( \lambda_{\text{max}} \) is the largest eigenvalue. This finishes the proof for Case 1.

**Case 2.** The analysis of Case 2 is analogous to the analysis of Case 1. Here, we assume that \( \hat{\psi}_k(\xi) \) takes the value 0 on the vertices \((0,0)\) and \((1,0)\) and the value 1 on the vertex \((0,1)\). So, we obtain \( \hat{\psi}_k(\xi) = \xi_2 \). Here we obtain, using the same arguments as above, that the inequality

\[
\int_{\Gamma_i} \hat{p}_k^2(\xi) \, d\xi \leq \left( 4 + \sqrt{6} \right) \int_{\Gamma_i} \hat{\psi}_k(\xi) \, \hat{p}_k^2(\xi) \, d\xi,
\]

holds, which finishes the proof.
References

1. R. Adams and J. Fournier, *Sobolev Spaces*, Academic Press, 2008, 2nd ed.
2. M. Bercovier and O. Pironneau, *Error estimates for finite element method solution of the Stokes problem in primitive variables*, Numerische Mathematik 33 (1979), 211 – 224.
3. J. Bergh and J. Lofstrom, *Interpolation Spaces, an Introduction*, Springer, Berlin, 1976.
4. D. Braess and R. Sarazin, *An efficient smoother for the Stokes problem*, Applied Numerical Mathematics 23 (1997), no. 1, 3 – 19.
5. J.H Bramble and J.E. Pasciak, *Iterative techniques for time dependent stokes problems*, Comp. Math. Appl. 33 (1997), 13–30.
6. S. Brenner, *A nonconforming multigrid method for the stationary stokes problem*, Math. Comp. 55 (1990), 411–473.
7. S.C. Brenner, *Multigrid methods for parameter dependent problems*, RAIRO, Modélisation Math. Anal. Numér 30 (1996), 265 – 297.
8. M. Dauge, *Elliptic boundary value problems on corner domains. Smoothness and asymptotics of solutions*, Lecture Notes in Mathematics, 1341. Berlin etc.: Springer-Verlag (1988).
9. W. Hackbusch, *Multi-Grid Methods and Applications*, Springer, Berlin, 1985.
10. R.B. Kellogg and J.E. Osborn, *A regularity result for the Stokes problem in a convex polygon*, Journal of Functional Analysis 21 (1976), no. 4, 397–431.
11. G. M. Kobelkov and M. A. Olshanskii, *Effective preconditioning of uzuwa type schemes for generalized stokes problem*, Numer. Math. 86 (2000), 443–470.
12. M. Larin and A. Reusken, *A comparative study of efficient iterative solvers for generalized stokes problem*, Numer. Linear Algebra Appl. 15 (2008), 13–34.
13. K. Mardal and R. Winther, *Uniform preconditioners for the time dependent stokes problem*, Numerische Mathematik 98 (2004), 305–327 (English).
14. M. Olshanskii, *Multigrid Analysis for the Time Dependent Stokes Problem*, Mathematics of Computation 81 (2012), no. 277, 57 – 79.
15. M. Olshanskii, J. Peters, and A. Reusken, *Uniform preconditioners for a parameter dependent saddle point problem with application to generalized stokes interface equations*, Numerische Mathematik (2005), 159–191.
16. S. Takacs, *A robust all-at-once multigrid method for the Stokes control problem*, 2014, submitted.
17. S. Takacs and W. Zulehner, *Convergence analysis of all-at-once multigrid methods for elliptic control problems under partial elliptic regularity*, SIAM J. on Numerical Analysis (2012), Accepted for publication.
18. Vanka S.P., *Block-implicit multigrid solution of Navier-Stokes equations in primitive variables*, Math. Comp. 65 (1986), 138 – 158.
19. R. Verfürth, *Error estimates for a mixed finite element approximation of the Stokes equations*, RAIRO 18 (1984), 175 – 182.
20. P. Wesseling and C. W. Oosterlee, *Geometric multigrid with applications to computational fluid dynamics*, J. Comput. Appl. Math. 128 (2001), 311–334.
21. W. Zulehner, *A class of smoothers for saddle point problems*, Computing, 65 (2000), 227–246.
22. *Non-standard Norms and Robust Estimates for Saddle Point Problems*, SIAM J. on Matrix Anal. & Appl 32 (2011), 536 – 560.