Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

Convergence analysis of all-at-once multigrid methods for elliptic control problems under partial elliptic regularity

Stefan Takacs Walter Zulehner

DK-Report No. 2012-08 06 2012

A–4040 LINZ, ALTENBERGERSTRASSE 69, AUSTRIA

Supported by
Austrian Science Fund (FWF) Upper Austria

DK sponsors:

- **Johannes Kepler University Linz** (JKU)
- **Austrian Science Fund** (FWF)
- **Upper Austria**

CONVERGENCE ANALYSIS OF ALL-AT-ONCE MULTIGRID METHODS FOR ELLIPTIC CONTROL PROBLEMS UNDER PARTIAL ELLIPTIC REGULARITY*

STEFAN TAKACS† AND WALTER ZULEHNER‡

Abstract. In this paper we consider the convergence theory for an all-at-once multigrid method for a distributed optimal control problem. Such an analysis has been recently done, see [11]. Here, we give a new proof which is based on a more straight-forward approach. The main benefit of this new approach is the possibility to extend the analysis to domains where full elliptic regularity, i.e., H^2-regularity for the Poisson problem, cannot be guaranteed.

Key words. PDE-constrained optimization, all-at-once multigrid methods, reduced regularity

AMS subject classifications. 65N55 35Q93 49J20

1. Introduction. The main goal of this paper is the presentation of a systematic approach for the construction and the analysis of all-at-once multigrid methods for parameter-dependent saddle point problems. The analysis is similar to [11], where the following model problem was considered:

Find $(y, u) \in H^1(\Omega) \times L^2(\Omega)$ such that

$$J(y,u) = \|y - y_D\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}$$

is minimized subject to the state equation

$$-\Delta y + y = u \text{ in } \Omega \qquad \text{and} \qquad \frac{\partial y}{\partial n} = 0 \text{ on } \partial\Omega. \tag{1.1}$$

Here, $\Omega \subseteq \mathbb{R}^d$ (with $d \in \{2,3\}$) is a bounded polygonal or polyhedral domain, $L^2(\Omega)$ 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. $y_D \in L^2(\Omega)$ is a given function (desired state). The given parameter $\alpha > 0$ is, depending on the interpretation, a cost parameter or a regularization parameter. For simplicity we restrict ourselves to this model problem. The generalization to other elliptic state equations is straight-forward.

In [11] an all-at-once multigrid method was introduced. It was proven that this method converges with rates bounded uniformly in the parameter α using the following regularity assumption on the solution of the state equation.

(R) *Full elliptic regularity:* There is a constant $C_R > 0$ such that for every $u \in L^2(\Omega)$, the solution y of (1.1) satisfies

$$y \in H^2(\Omega) \qquad \text{and} \qquad \|y\|_{H^2(\Omega)} \leq C_R \|u\|_{L^2(\Omega)}.$$

Such a regularity assumption can be guaranteed, e.g., for domains with a sufficiently smooth boundary (see, e.g., [10]) or polygonal or polyhedral domains which are convex (see, e.g., [5] or [6]). For domains with reentrant corners, condition **(R)** is typically not satisfied.

In this paper, we give a convergence proof under the weaker assumption.

*The research was funded by the Austrian Science Fund (FWF): W1214-N15, project DK12.

†Doctoral Program Computational Mathematics, Johannes Kepler University Linz, Austria, (stefan.takacs@dk-compmath.jku.at)

‡Institute of Computational Mathematics, Johannes Kepler University Linz, Austria, (zulehner@numa.uni-linz.ac.at)

(R') *Partial elliptic regularity:* For some $s \in [0,1)$, there is a constant $C_R > 0$ such that for every $u \in (H^s(\Omega))^*$, the solution y of (1.1) satisfies

$$y \in H^{2-s}(\Omega) \qquad \text{and} \qquad \|y\|_{H^{2-s}(\Omega)} \le C_R \|u\|_{(H^s(\Omega))^*}.$$

Here, $H^s(\Omega)$ is a standard fractional order Sobolev space, cf. [1], and $(H^s(\Omega))^*$ is its dual space. This regularity assumption is satisfied for polygonal domains with reentrant corners: If Ω is a non-convex polygonal domain with $\omega > \pi$ being the largest angle of the boundary polygonal (measured from inside), assumption **(R')** is satisfied for all s with

$$1 - \frac{\pi}{\omega} < s < 1,$$

cf. Remark 2.4.6 in [7].

For fixed α, it is straight-forward how to extend the convergence proof to the case of partial regularity, following the line of arguments in [8]. These bounds are not robust in the parameter α and deteriorate for α approaching 0. Existing robust convergence results for all-at-once multigrid methods, like [3, 11], were based on full elliptic regularity. The authors are not aware of a straight-forward extension of these results to the case of partial elliptic regularity.

In the present paper, we will give a robust convergence proof under partial regularity. For the proof we will make use of parameter-dependent norms which will be constructed using the concept of Hilbert space interpolation, cf. [1, 4].

This paper is organized as follows. In Section 2 we present a standard multigrid framework. In Section 3, we recall standard smoothers and show that they satisfy the smoothing property. The convergence analysis for the case of partial elliptic regularity is carried out in Section 4. The numerical results, we present in Section 5, illustrate the convergence result. Concluding remarks can be found in Section 6.

2. The general framework.

In this section we recall a standard multigrid framework for solving the model problem. We derive the optimality system and its discretization in Subsection 2.1. In Subsection 2.2 we introduce all-at-once multigrid methods.

2.1. Optimality systems.

The solution of the model problem is characterized by the Karush-Kuhn-Tucker-system (KKT-system). This system depends on the state y, the control u and the Lagrange multiplier, say p, associated with the constraint, here the state equation. This system immediately implies that $u = \alpha^{-1} p$ holds, which allows to eliminate the control u and leads to the following reduced KKT-system. Find $(y,p) \in H^1(\Omega) \times H^1(\Omega)$ such that

$$\begin{array}{rcrcl} (y,\widetilde{y})_{L^2(\Omega)} & + & (p,\widetilde{y})_{H^1(\Omega)} & = & (y_D,\widetilde{y})_{L^2(\Omega)} \\ (y,\widetilde{p})_{H^1(\Omega)} & - & \alpha^{-1}(p,\widetilde{p})_{L^2(\Omega)} & = & 0 \end{array}$$

holds for all $(\widetilde{y},\widetilde{p}) \in H^1(\Omega) \times H^1(\Omega)$. For the details, see, e.g., [11].

We obtain a variational problem for $x = (y,p)$ in the product space $X = Y \times P$, which can be written as follows. Find $x \in X$ such that

$$\mathcal{B}(x,\widetilde{x}) = \mathcal{F}(\widetilde{x}) \qquad \text{for all } \widetilde{x} \in X, \tag{2.1}$$

where

$$\begin{aligned} \mathcal{B}((y,p),(\widetilde{y},\widetilde{p})) &:= (y,\widetilde{y})_{L^2(\Omega)} + (p,\widetilde{y})_{H^1(\Omega)} + (y,\widetilde{p})_{H^1(\Omega)} - \alpha^{-1}(p,\widetilde{p})_{L^2(\Omega)}, \\ \mathcal{F}(\widetilde{y},\widetilde{p}) &:= (y_D,\widetilde{y})_{L^2(\Omega)}. \end{aligned}$$

Existence and uniqueness of the solution is guaranteed by the following condition.

(A1) *There are constants* $\underline{C} > 0$ *and* $\overline{C}$ *such that*

$$\underline{C}\|x\|_X \le \sup_{0 \ne \widetilde{x} \in X} \frac{\mathcal{B}(x, \widetilde{x})}{\|\widetilde{x}\|_X} \le \overline{C}\|x\|_X$$

holds for all $x \in X$.

In [11] and [14] it was shown that this condition is satisfied with constants $\underline{C}$ and $\overline{C}$ independent of α for $X := H^1(\Omega) \times H^1(\Omega)$ equipped with the parameter-dependent norm

$$\|x\|_X := (\|y\|_Y^2 + \|p\|_P^2)^{1/2},$$

where

$$\|y\|_Y := (\|y\|^2_{L^2(\Omega)} + \alpha^{1/2}\|y\|^2_{H^1(\Omega)})^{1/2} \text{ and}$$
$$\|p\|_P := (\alpha^{-1}\|p\|^2_{L^2(\Omega)} + \alpha^{-1/2}\|p\|^2_{H^1(\Omega)})^{1/2}.$$

NOTATION 2.1. *For Hilbert spaces* A_1 *and* A_2, $A_1 \cap A_2$ *denotes the Hilbert space of all elements from the intersection of* A_1 *and* A_2 *with norm*

$$\|\cdot\|_{A_1 \cap A_2} := (\|\cdot\|^2_{A_1} + \|\cdot\|^2_{A_2})^{1/2}.$$

For a scalar $\gamma > 0$ *and a Hilbert space* A, γA *denotes the Hilbert space of all elements from* A *with norm*

$$\|\cdot\|_{\gamma A} := \gamma\|\cdot\|_A.$$

Using this notation, we can rewrite the norms introduced above as follows

$$\|y\|_Y = \|y\|_{L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega)} \qquad \text{and} \qquad \|p\|_P = \|p\|_{\alpha^{-1/2}L^2(\Omega)\cap\alpha^{-1/4}H^1(\Omega)}.$$

A standard way to discretize (2.1) is the Galerkin principle. Let $X_k \subset X$, $k = 0, 1, 2, \ldots$, be a sequence of finite-dimensional subspaces. For simplicity, we restrict ourselves to the case of nested spaces, i.e., we assume that $X_k \subseteq X_{k+1}$. The Galerkin approximation of (2.1) is given by: Find $x_k \in X_k$ such that

$$\mathcal{B}(x_k, \widetilde{x}_k) = \mathcal{F}(\widetilde{x}_k) \qquad \text{for all } \widetilde{x}_k \in X_k. \tag{2.2}$$

Here, the spaces X_k are product spaces, too, i.e., $X_k = Y_k \times P_k$. For case of the model problem, $Y_k = P_k$ is a reasonable choice. For this case the following condition was shown in [11].

(A1a) *There are constants* $\underline{C}_D > 0$ *and* $\overline{C}_D$ *such that for all grid levels* k

$$\underline{C}_D\|x_k\|_X \le \sup_{0 \ne \widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_X} \le \overline{C}_D\|x_k\|_X$$

holds for all $x_k \in X_k$.

This guarantees existence and uniqueness of the solution of the discretized problem. As for condition **(A1)**, the constants $\underline{C}_D$ and $\overline{C}_D$ are independent of α.

The variational problem (2.2) reads in matrix-vector notation as follows

$$\mathcal{A}_k \, \underline{x}_k = \underline{f}_k, \qquad \text{where} \qquad \mathcal{A}_k := \begin{pmatrix} M_k & K_k \\ K_k & -\alpha^{-1}M_k \end{pmatrix} \tag{2.3}$$

and M_k and K_k are the standard mass and stiffness matrices, respectively. The matrix $\mathcal{A}_k$ is symmetric and indefinite. Here and in what follows, any underlined quantity, like $\underline{x}_k$, denotes the coefficient vector of the corresponding finite element function, here $x_k \in X_k$, with respect to a basis of X_k.

2.2. All-at-once multigrid methods. In this subsection, we introduce a standard all-at-once multigrid framework for solving the discretized equation (2.3) on grid level k. Starting from an initial approximation $\underline{x}_k^{(0)}$ one iterate of the multigrid method is given by the following two steps:

- *Smoothing procedure:* Compute
$$\underline{x}_k^{(0,m)} := \underline{x}_k^{(0,m-1)} + \hat{\mathcal{A}}_k^{-1}(\underline{f}_k - \mathcal{A}_k\, \underline{x}_k^{(0,m-1)}) \qquad \text{for } m = 1, \ldots, \nu$$
with $\underline{x}_k^{(0,0)} = \underline{x}_k^{(0)}$.
- *Coarse-grid correction:*
  - Compute the defect $\underline{r}_k^{(1)} := (\underline{f}_k - \mathcal{A}_k\, \underline{x}_k^{(0,\nu)})$ and restrict it to grid level $k-1$ using an restriction matrix I_k^{k-1}:
$$\underline{r}_{k-1}^{(1)} := I_k^{k-1}(\underline{f}_k - \mathcal{A}_k\, \underline{x}_k^{(0,\nu)}).$$
  - Solve the coarse-grid problem
$$\mathcal{A}_{k-1}\, \underline{p}_{k-1}^{(1)} = \underline{r}_{k-1}^{(1)} \tag{2.4}$$
approximatively.
  - Prolongate $\underline{p}_{k-1}$ to the grid level k using a prolongation matrix I_{k-1}^k and add the result to the previous iterate:
$$\underline{x}_k^{(1)} := \underline{x}_k^{(0,\nu)} + I_{k-1}^k \underline{p}_{k-1}^{(1)}.$$

As we have assumed nested spaces, the intergrid-transfer matrices I_{k-1}^k and I_k^{k-1} can be chosen in a canonical way: I_{k-1}^k is the canonical embedding and the restriction I_k^{k-1} is its transpose. The choice of $\hat{\mathcal{A}}_k^{-1}$ will be discussed in the next subsection.

If the problem on the coarser grid is solved exactly (two-grid method), the coarse-grid correction is given by
$$\underline{x}_k^{(1)} := \underline{x}_k^{(0,\nu)} + I_{k-1}^k \mathcal{A}_{k-1}^{-1} I_k^{k-1}(\underline{f}_k - \mathcal{A}_k\, \underline{x}_k^{(0,\nu)}). \tag{2.5}$$
In practice the problem (2.4) is approximatively solved by applying one step (V-cycle) or two steps (W-cycle) of the multigrid method, recursively. On grid level $k = 0$ the problem (2.4) is solved exactly.

To show a multigrid convergence result based on Hackbusch's splitting of the analysis in smoothing property and approximation property we have to introduce an appropriate norm.

For each $s \in [0, 1)$ we need a linear function space $X_-^s \supseteq X$ equipped with a family of norms $\|\cdot\|_{X_{-,k}^s}$, which depend on the grid level k and on the choice of the parameter α. The choice of the parameter α will always be clear from the context, so only the dependence on s and k is denoted explicitly. Note that the function space X_-^s, equipped with any of the norms $\|\cdot\|_{X_{-,k}^s}$, is a Hilbert space, denoted by $X_{-,k}^s := (X_-^s, \|\cdot\|_{X_{-,k}^s})$.

Due to Hackbusch, the analysis is based on the following properties.

- *Smoothing property:*
$$\sup_{\widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,\nu)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X_{-,k}^s}} \leq \eta(\nu) \|x_k^{(0)} - x_k\|_{X_{-,k}^s} \tag{2.6}$$
holds for some function $\eta(\nu)$ with $\lim_{\nu\to\infty} \eta(\nu) = 0$.

- *Approximation property:*

$$\|x_k^{(1)} - x_k\|_{X^s_{-,k}} \le C_A \sup_{\widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,\nu)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X^s_{-,k}}} \tag{2.7}$$

  holds for some constant $C_A > 0$.

It is easy to see that, if we combine both conditions, we obtain

$$\|x_k^{(1)} - x_k\|_{X^s_{-,k}} \le q(\nu)\|x_k^{(0)} - x_k\|_{X^s_{-,k}},$$

where $q(\nu) = C_A\eta(\nu)$, i.e., that the two-grid method converges for ν large enough. The convergence of the W-cycle multigrid method can be shown under mild assumptions, see e.g. [8].

For full elliptic regularity, which corresponds to $s = 0$ in our notation, such a linear space $X^0_- = X_-$ equipped with (mesh-dependent) norms $\|\cdot\|_{X^0_{-,k}} = \|\cdot\|_{X_{-,k}}$ has already been introduced, see [11]. We will recall the choice of the space X_- and the norms $\|\cdot\|_{X_{-,k}}$, which leads to the Hilbert spaces $X_{-,k} := (X_-, \|\cdot\|_{X_{-,k}})$, in the next section. Now, for general $s \in (0, 1)$, we propose the following choice of the Hilbert spaces $X^s_{-,k}$:

$$X^s_{-,k} := [X_{-,k}, X]_s,$$

based on the following notation.

NOTATION 2.2. *For each $\theta \in (0, 1)$, the interpolant of two Hilbert spaces A_1 and A_2, defined using the real K-method, cf. Theorem 15.1 in [9], itself is a Hilbert space, denoted by*

$$[A_1, A_2]_\theta.$$

This choice corresponds exactly to the strategy in [8] for analyzing the multigrid method for the Poisson problem, where (in our notation) $X^s_- = H^s(\Omega)$ was chosen.

For further reference we recall three results on Hilbert space interpolation, cf. [1, 4].

Applied to standard Sobolev spaces $H^m(\Omega)$ and $H^n(\Omega)$, we obtain

$$[H^m(\Omega), H^n(\Omega)]_\theta = H^{(1-\theta)m+\theta n}(\Omega),$$

i.e., a standard Sobolev space. For scaled Hilbert spaces, the interpolation is the (weighted) geometric mean, i.e., for Hilbert spaces A_1 and A_2 and scalars $\gamma_1 > 0$ and $\gamma_2 > 0$ we obtain

$$[\gamma_1 A_1, \gamma_2 A_2]_\theta = \gamma_1^{1-\theta}\gamma_2^\theta [A_1, A_1]_\theta. \tag{2.8}$$

One main result on interpolation spaces is the interpolation theorem, cf. Theorem 3.2.23 in [4], which reads as follows.

THEOREM 2.3 (Interpolation Theorem). *Let A_1, A_2, B_1 and B_2 be Hilbert spaces and let $T : A_1 + A_2 := \{a_1 + a_2 \;:\; a_1 \in A_1, a_2 \in A_2\} \to B_1 + B_2$ be an operator such that there are constants C_1 and C_2 such that for all $a_1 \in A_1$ and $a_2 \in A_2$, the mapping properties $Ta_1 \in B_1$ and $Ta_2 \in B_2$ and the bounds*

$$\|Ta_1\|_{B_1} \le C_1\|a_1\|_{A_1} \qquad \textit{and} \qquad \|Ta_2\|_{B_2} \le C_2\|a_2\|_{A_2}$$

are satisfied.

Then for all $a \in [A_1, A_2]_\theta$, we obtain $Ta \in [B_1, B_2]_\theta$ and

$$\|Ta\|_{[B_1,B_2]_\theta} \leq C(\theta) C_1^{1-\theta} C_2^{\theta} \|a\|_{[A_1,A_2]_\theta},$$

where $C(\theta)$ only depends on θ.

3. Smoothers. The choice of an appropriate smoother is a key issue in constructing an all-at-once multigrid method. Here, we recall two classes of smoothers which are appropriate for the model problem.

The first class are *normal equation smoothers*, cf. [3], given by

$$\underline{x}_k^{(0,m)} := \underline{x}_k^{(0,m-1)} + \tau \underbrace{\mathcal{L}_k^{-1} \mathcal{A}_k \mathcal{L}_k^{-1}}_{\hat{\mathcal{A}}_k^{-1} :=} (\underline{f}_k - \mathcal{A}_k \, \underline{x}_k^{(0,m-1)}) \qquad \text{for } m = 1, \ldots, \nu.$$

Here, $\mathcal{L}_k$ is the matrix representing the scalar product $(\cdot,\cdot)_{X_{-,k}}$, i.e.,

$$(\mathcal{L}_k \underline{x}_k, \underline{\widetilde{x}}_k)_{\ell^2} = (x_k, \widetilde{x}_k)_{X_{-,k}} \qquad \text{for all } x_k, \widetilde{x}_k \in X_k, \tag{3.1}$$

where $(\cdot,\cdot)_{\ell^2}$ is the Euclidean inner product. The damping parameter $\tau > 0$ is chosen such that $\tau < 2/\rho(\hat{\mathcal{A}}_k^{-1}\mathcal{A}_k)$, where $\rho(M)$ denotes the spectral radius of a matrix M. Note that the definition of the smoother is independent of the regularity parameter s.

For robust convergence, we need to choose τ independent of the grid level k and the parameter α. This is possible if the norm $\|\cdot\|_{X_{-,k}}$ is chosen such that the following condition is satisfied.

(A2) *There is a constant C_M such that*

$$\|x_k\|_X \leq C_M \|x_k\|_{X_{-,k}} \qquad \text{for all } x_k \in X_k$$

holds on all grid levels k and for all choices of the parameter α.

Certainly, this iterative procedure should be efficient-to-apply. This is satisfied if $\|\cdot\|_{X_{-,k}}$ is a scaled L^2-norm. The norm $\|\cdot\|_{X_{-,k}}$ can be constructed by replacing each occurrence of $\|\cdot\|_{H^1(\Omega)}$ in the norm $\|\cdot\|_X$ by $h_k^{-1}\|\cdot\|_{L^2(\Omega)}$. This leads to

$$\|x\|_{X_{-,k}} := (\|y\|_{Y_{-,k}}^2 + \|p\|_{P_{-,k}}^2)^{1/2}, \tag{3.2}$$

where

$$\|y\|_{Y_{-,k}} := (1 + \alpha^{1/2} h_k^{-2})^{1/2} \|y\|_{L^2(\Omega)} \text{ and}$$
$$\|p\|_{P_{-,k}} := \alpha^{-1/2} (1 + \alpha^{1/2} h_k^{-2})^{1/2} \|p\|_{L^2(\Omega)}.$$

Using a standard inverse inequality, one can show that condition **(A2)** is satisfied for this choice.

For the case $s = 0$, the smoothing property of the preconditioned normal equation smoother was shown in [3]. We recall the result:

LEMMA 3.1. *Assume that* **(A1a)** *and* **(A2)** *hold. Then $\tau > 0$ can be chosen independent of grid level k and the choice of the parameter α such that*

$$\tau\, \rho(\hat{\mathcal{A}}_k^{-1}\mathcal{A}_k) \leq \rho_{max} < 2$$

holds. For this choice of τ, there is a constant $C_S > 0$ independent of grid level k and choice of the parameter α such that the smoothing rate satisfies

$$\eta(\nu) := C_S \nu^{-1/2},$$

where the constant $C_S > 0$ is independent of grid level k and parameter α.

REMARK 3.2. *The matrix $\mathcal{L}_k$ is a 2-by-2 block-diagonal matrix, where each block is a scaled mass matrix. Due to the fact that the mass matrix is spectrally equivalent to its diagonal, also $\mathcal{L}_k$ is spectrally equivalent to its diagonal. Therefore, if we replace the matrix $\mathcal{L}_k$ by*

$$\hat{\mathcal{L}}_k := \begin{pmatrix} \operatorname{diag}(M_k + \alpha^{1/2} K_k) & \\ & \alpha^{-1} \operatorname{diag}(M_k + \alpha^{1/2} K_k) \end{pmatrix},$$

the smoothing property for the corresponding iterative method is still satisfied.

It remains to show the smoothing property in the norm $\|\cdot\|_{X^s_{-,k}}$, which will be done below.

The second class of smoothers, that we consider for the model problem, is the class of *collective iteration schemes*. Such methods have been proposed, e.g., in [13] or [2]. One method of this class is the collective Richardson iteration, which is given by

$$\underline{x}_k^{(0,m)} := \underline{x}_k^{(0,m-1)} + \tau \underbrace{\begin{pmatrix} m_k I & k_k I \\ k_k I & -\alpha^{-1} m_k I \end{pmatrix}}_{\hat{\mathcal{A}}_k :=}^{-1} (\underline{f}_k - \mathcal{A}_k\, \underline{x}_k^{(0,m-1)}) \quad \text{for } m = 1, \dots, \nu,$$

where m_k and k_k are the largest entries of the diagonals of M_k and K_k, respectively, and the parameter $\tau \in (0,1)$ is chosen independent of grid level k and parameter α. In [12] the smoothing property was shown. We recall the result:

LEMMA 3.3. *For $\tau \in (0,1)$ chosen independent of grid level k and parameter α, the collective Richardson iteration satisfies the smoothing property with*

$$\eta(\nu) := C_S \nu^{-1/2},$$

where the constant $C_S > 0$ is independent of grid level k and parameter α.

The proof in [12] fixes the choice of the norm $\|\cdot\|_{X_{-,k}}$ (up to constants) to the norm specified in (3.2).

We have mentioned above that for both smoothers the smoothing property is satisfied for the norm $\|\cdot\|_{X_{-,k}}$, i.e., for $s = 0$. We can use interpolation theory to carry over this smoothing result to the case $s \in (0,1)$.

LEMMA 3.4. *Assume that $\mathcal{A}_k$ is symmetric and that the smoother is given by*

$$\underline{x}_k^{(0,m)} := \underline{x}_k^{(0,m-1)} + \hat{\mathcal{A}}_k^{-1}(\underline{f}_k - \mathcal{A}_k\, \underline{x}_k^{(0,m-1)}) \qquad \textit{for } m = 1, \dots, \nu,$$

where $\hat{\mathcal{A}}_k$ is a symmetric matrix. Assume that this smoother satisfies the smoothing property in the norm $\|\cdot\|_{X_{-,k}}$, i.e.,

$$\sup_{\widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,\nu)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X_{-,k}}} \le \eta(\nu) \|x_k^{(0)} - x_k\|_{X_{-,k}}$$

holds. Moreover, assume that condition **(A1a)** *holds and the smoother is power-bounded, i.e., that there is a constant C_B such that*

$$\|x_k^{(0,m)} - x_k\|_{X_{-,k}} \le C_B \|x_k^{(0)} - x_k\|_{X_{-,k}} \tag{3.3}$$

holds for all $m \in \mathbb{N}$.

Then for all $s \in (0,1)$ *the smoother satisfies the smoothing property also in the norm* $\|\cdot\|_{X^s_{-,k}}$*, i.e., there is a constant* $\widetilde{C}_S$*, depending only on* s*,* $\underline{C}_D$ *and* $\overline{C}_D$ *and* C_B*, such that*

$$\sup_{\widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,\nu)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X^s_{-,k}}} \leq \widetilde{C}_S \eta(\nu)^{1-s} \|x_k^{(0)} - x_k\|_{X^s_{-,k}} \tag{3.4}$$

is satisfied.

Proof. The proof is done using interpolation. By assumption, we know that the smoothing property

$$\sup_{\widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,\nu)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X_{-,k}}} \leq \eta(\nu) \|x_k^{(0)} - x_k\|_{X_{-,k}}$$

is satisfied. We will also show that there is a constant $C > 0$ such that

$$\sup_{\widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,\nu)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X}} \leq C \|x_k^{(0)} - x_k\|_{X} \tag{3.5}$$

holds. Then the interpolation theorem (Theorem 2.3) immediately implies (3.4).

In order to show (3.5), we reformulate the condition in matrix-vector notation:

$$\|\mathcal{A}_k (I - \tau \hat{\mathcal{A}}_k^{-1} \mathcal{A}_k)^\nu \underline{r}_k\|_{\mathcal{Q}_k^{-1}} \leq C \|\underline{r}_k\|_{\mathcal{Q}_k}$$

has to be shown for all $\underline{r}_k := \underline{x}_k^{(0)} - \underline{x}_k$. Here, the matrix $\mathcal{Q}_k$ represents the scalar product $(\cdot,\cdot)_X$ on X_k. In other words, we have to show that the spectral norm of

$$\mathcal{P}_k := \mathcal{Q}_k^{-1/2} \mathcal{A}_k (I - \tau \hat{\mathcal{A}}_k^{-1} \mathcal{A}_k)^\nu \mathcal{Q}_k^{-1/2}$$

is bounded by a constant. This matrix is symmetric, so we have

$$\begin{aligned} \|\mathcal{P}_k\|_{\ell^2} &= \rho(\mathcal{P}_k) \\ &= \rho(\mathcal{L}_k^{1/2} \mathcal{Q}_k^{-1} \mathcal{A}_k (I - \tau \hat{\mathcal{A}}_k^{-1} \mathcal{A}_k)^\nu \mathcal{L}_k^{-1/2}) \\ &\leq \|\mathcal{L}_k^{1/2} \mathcal{Q}_k^{-1} \mathcal{A}_k \mathcal{L}_k^{-1/2}\|_{\ell^2} \|\mathcal{L}_k^{1/2} (I - \tau \hat{\mathcal{A}}_k^{-1} \mathcal{A}_k)^\nu \mathcal{L}_k^{-1/2}\|_{\ell^2}, \end{aligned}$$

where $\|\cdot\|_{\ell^2}$ is the spectral norm and the matrix $\mathcal{L}_k$ is as defined in (3.1). Here, the second factor can be bounded from above by C_B using condition (3.3). The first factor can be bounded from above by two times the numerical radius, where the numerical radius $r(M)$ of a matrix M is given by

$$r(M) := \sup_{0 \neq \underline{x}_k \in \mathbb{R}^{N_k}} \frac{|(M\underline{x}_k, \underline{x}_k)_{\ell^2}|}{|(\underline{x}_k, \underline{x}_k)_{\ell^2}|}.$$

We obtain

$$
\begin{aligned}
\|\mathcal{L}_k^{1/2}\mathcal{Q}_k^{-1}\mathcal{A}_k\mathcal{L}_k^{-1/2}\|_{\ell^2} &\le 2r(\mathcal{L}_k^{1/2}\mathcal{Q}_k^{-1}\mathcal{A}_k\mathcal{L}_k^{-1/2}) \\
&\le 2 \sup_{0\neq \underline{x}_k\in\mathbb{R}^{N_k}} \frac{|(\mathcal{L}_k^{1/2}\mathcal{Q}_k^{-1}\mathcal{A}_k\mathcal{L}_k^{-1/2}\underline{x}_k,\underline{x}_k)_{\ell^2}|}{|(\underline{x}_k,\underline{x}_k)_{\ell^2}|} \\
&= 2 \sup_{0\neq \underline{x}_k\in\mathbb{R}^{N_k}} \frac{|(\mathcal{Q}_k^{-1/2}\mathcal{L}_k\mathcal{Q}_k^{-1}\mathcal{A}_k\mathcal{Q}_k^{-1/2}\underline{x}_k,\underline{x}_k)_{\ell^2}|}{|(\mathcal{Q}_k^{-1/2}\mathcal{L}_k\mathcal{Q}_k^{-1/2}\underline{x}_k,\underline{x}_k)_{\ell^2}|} \\
&\le 2 \sup_{0\neq \underline{x}_k,\underline{y}_k\in\mathbb{R}^{N_k}} \frac{|(\mathcal{Q}_k^{-1/2}\mathcal{A}_k\mathcal{Q}_k^{-1/2}\underline{x}_k,\underline{y}_k)_{\ell^2}|}{|(\underline{x}_k,\underline{y}_k)_{\ell^2}|} \le 2\|\mathcal{Q}_k^{-1/2}\mathcal{A}_k\mathcal{Q}_k^{-1/2}\|_{\ell^2} \\
&= 2 \sup_{0\neq \underline{x}_k\in\mathbb{R}^{N_k}} \frac{|(\mathcal{Q}_k^{-1/2}\mathcal{A}_k\mathcal{Q}_k^{-1/2}\underline{x}_k,\underline{x}_k)_{\ell^2}|}{(\underline{x}_k,\underline{x}_k)_{\ell^2}} = 2 \sup_{0\neq \underline{x}_k\in\mathbb{R}^{N_k}} \frac{|(\mathcal{A}_k\underline{x}_k,\underline{x}_k)_{\ell^2}|}{(\mathcal{Q}_k\underline{x}_k,\underline{x}_k)_{\ell^2}} \\
&= 2 \sup_{0\neq \underline{x}_k\in\mathbb{R}^{N_k}} \frac{|\mathcal{B}(x_k,x_k)|}{\|x_k\|_X^2} \le 2\overline{C}_D,
\end{aligned}
$$

where $\overline{C}_D$ is the constant in **(A1a)**. This shows (3.5) which finishes the proof. ∎

The conditions of Lemma 3.4 are satisfied for both smoothers discussed above: The symmetry of $\hat{\mathcal{A}}_k$ is in both cases obvious. Power boundedness can be shown for the preconditioned normal equation smoother with an easy eigenvalue analysis. For the collective Richardson smoother, power boundedness was shown in [12].

4. Approximation property. The proof of the approximation property follows classical convergence proofs, cf. [8]. We have already mentioned that Hackbusch's proof of the approximation property for the Poisson problem uses the spaces $X_-^s = H^s(\Omega)$ and $X = H^1(\Omega)$. In his proof a H^s-approximation result is constructed using an Aubin-Nitsche duality trick. For this purpose, he uses that the regularity assumption **(R')** states that the solution of the Poisson problem is a function in $H^{2-s}(\Omega)$.

This motivates to introduce, beside X_-^s and X, also a third space X_+^s (which plays the same role as $H^{2-s}(\Omega)$ for the Poisson equation) equipped with appropriate norms $\|\cdot\|_{X_{+,k}^s}$. In the next theorem we collect all the assumptions we need for a proof of the approximation property. Then it will become more transparent how to choose the Hilbert space $X_{+,k}^s := (X_+^s, \|\cdot\|_{X_{+,k}^s})$.

We closely follow the line of arguments in [8] for the Poisson equation.

As mentioned above, Hackbusch needs the regularity assumption **(R')**. For our proof we need an analogous assumption which reads as follows.

(A3) *There is a constant $C_R > 0$ such that for all grid levels k, all $\mathcal{F} \in (X_{-,k}^s)^*$ the solution $x \in X$ of the problem, find $x \in X$ such that*

$$\mathcal{B}(x,\widetilde{x}) = \mathcal{F}(\widetilde{x}) \qquad \text{for all } \widetilde{x} \in X, \tag{4.1}$$

satisfies $x \in X_+^s$ and the bound

$$\|x\|_{X_{+,k}^s} \le \hat{C}_R \|\mathcal{F}\|_{(X_{-,k}^s)^*}.$$

Besides the regularity assumption, Hackbusch needs in his proof an approximation error estimate. In our framework, we need the following analogous assumption.

(A4) *There is a constant $C_I > 0$ such that for all grid levels k and all $x \in X_+$ the approximation error result*

$$\inf_{x_k \in X_k} \|x - x_k\|_X \leq C_I \|x\|_{X^s_{+,k}}$$

is satisfied.

With these conditions, we have

THEOREM 4.1. *Let for $k = 0, 1, 2, \dots$ the symmetric matrices $\mathcal{A}_k$ be obtained by discretizing problem* (2.1) *using a sequence of finite-dimensional nested subspaces $X_{k-1} \subseteq X_k \subset X$. Assume that there are Hilbert spaces $X^s_+ \subseteq X \subseteq X^s_-$ with mesh-dependent norms $\|\cdot\|_{X^s_{+,k}}$, $\|\cdot\|_X$ and $\|\cdot\|_{X^s_{-,k}}$ such that the conditions* **(A1)**, **(A1a)**, **(A3)** *and* **(A4)** *hold.*

Then the coarse-grid correction (2.5) *satisfies the* approximation property (2.7) *with a constant C_A, only depending on $\underline{C}$, $\overline{C}$, $\underline{C}_D$, $\overline{C}_D$, C_I and $\hat{C}_R$*

To some extend this is a known result in literature. For sake of self-containedness and specially tailored to our notation, we give a proof of this theorem in the Appendix.

4.1. Convergence analysis under full elliptic regularity. Before we start discussing the convergence analysis under partial elliptic regularity, we have to recall the results for the case of full regularity.

First note that in Subsection 2.1, we have already seen that conditions **(A1)** and **(A1a)** are satisfied for the model problem.

We choose the norm $\|\cdot\|_{X^0_{-,k}} = \|\cdot\|_{X_{-,k}}$ as introduced in (3.2), which guarantees the smoothing property and condition **(A2)**.

Already for $s = 0$ (full elliptic regularity) the choice of the linear space $X^0_+ = X_+$ equipped with norms $\|\cdot\|_{X_{+,k}}$ is not straight-forward. The basic idea is to choose the Hilbert space $X_{+,k} := (X_+, \|\cdot\|_{X_{+,k}})$ such that the Hilbert space X is the interpolant at $\theta = 1/2$ of the Hilbert spaces $X_{-,k}$ and $X_{+,k}$, i.e.,

$$X = [X_{-,k}, X_{+,k}]_{1/2}. \tag{4.2}$$

This is satisfied for the choice $X_+ := H^2(\Omega) \times H^2(\Omega)$ with associated mesh-dependent norms

$$\|x\|_{X_{+,k}} := (\|y\|^2_{Y_{+,k}} + \|p\|^2_{P_{+,k}})^{1/2},$$

where

$$\|y\|_{Y_{+,k}} := (1 + \alpha^{1/2} h_k^{-2})^{-1/2} (\|y\|^2_{L^2(\Omega)} + \alpha \|y\|^2_{H^2(\Omega)})^{1/2} \text{ and}$$
$$\|p\|_{P_{+,k}} := \alpha^{-1} (1 + \alpha^{1/2} h_k^{-2})^{-1/2} (\|p\|^2_{L^2(\Omega)} + \alpha \|p\|^2_{H^2(\Omega)})^{1/2}.$$

The combination of (2.8) and Lemma 6.1, introduced in the Appendix, immediately implies that this choice of the Hilbert space $X_{+,k} = (X_+, \|\cdot\|_{X_{+,k}})$ guarantees (4.2).

REMARK 4.2. *The choice of $X_{+,k}$ in this paper is different to the choice used in [11]. Note that the choice in the present paper is closer to the classical proofs for the full elliptic regularity case, cf. [8], where* (4.2) *is also satisfied.*

With this choice of $X_{+,k}$ we now have

LEMMA 4.3. *Condition* **(A4)** *is satisfied for $s = 0$.*

Proof. Note that $X_+ = Y_+ \times P_+$ and $X = Y \times P$. Therefore, we can do the analysis for y and p separately. As the analysis for p is completely analogous, we consider y only.

Using standard finite element theory, we know that there is an interpolation operator $\Pi_k : H^1(\Omega) \to Y_k$ such that

$$\|y - \Pi_k y\|_{H^l(\Omega)} \leq C h_k^{(n-m)/2} \|y\|_{H^n(\Omega)} \qquad \text{for all } y \in H^n(\Omega) \cap H^1(\Omega) \tag{4.3}$$

for $0 \leq m \leq n \leq 2$. Using the fact that

$$\|\cdot\|_{\alpha^{1/4} H^1(\Omega)} = \|\cdot\|_{[L^2(\Omega), \alpha^{1/2} H^2(\Omega)]_{1/2}} \leq (\|\cdot\|^2_{L^2(\Omega)} + \alpha \|\cdot\|^2_{H^2(\Omega)})^{1/2}$$

holds, we can show – analogously to Lemma 2.3 in [11] – using the definition the involved norms $\|\cdot\|_Y$ and $\|\cdot\|_{Y_{+,k}}$ that

$$\|y - \Pi_k y\|_Y \leq C \|y\|_{Y_{+,k}} \qquad \text{for all } y \in Y_+ \tag{4.4}$$

holds. □

And, finally, one can show the following result.

Lemma 4.4. *In the framework of this section, condition* **(A3)** *is satisfied for* $s = 0$ *if condition* **(R)** *is satisfied.*

The proof is postponed until the next subsection, where the general case $s \in [0, 1)$ is considered.

4.2. Convergence analysis under partial elliptic regularity. The next step is the construction of the Hilbert space $X^s_{+,k}$ for $s \in (0, 1)$. Here, we use the observation (4.2) and follow classical results. We have already used $X^s_{-,k} = [X_{-,k}, X]_s$ for the construction of $X^s_{-,k}$. Observe, that in Hackbusch's proof, $X^s_{+,k} = H^{2-s}(\Omega) = [H^2(\Omega), H^1(\Omega)]_s = [X_{+,k}, X]_s$ is satisfied, which again is used as construction principle. So, for the case of the model problem, we choose the Hilbert space $X^s_{+,k}$ as follows:

$$X^s_{+,k} := [X_{+,k}, X]_s. \tag{4.5}$$

Note that, analogously to (4.2), the identity

$$X = [X^s_{-,k}, X^s_{+,k}]_{1/2}$$

is satisfied, cf. the reiteration theorem (Theorem 3.2.20, Corollary 3.2.17 in [4]).

Lemma 4.5. *The Hilbert spaces* $X^s_{-,k}$ *and* $X^s_{+,k}$*, introduced in* (3.2) *and* (4.5)*, are the linear spaces*

$$X^s_- = H^s(\Omega) \times H^s(\Omega) \qquad \text{and} \qquad X^s_+ = H^{2-s}(\Omega) \times H^{2-s}(\Omega).$$

equipped with norms which are equivalent (with constants independent of k and α) to the following norms

$$\|x\|_{X^s_{-,k}} = (\|y\|^2_{Y^s_{-,k}} + \|p\|^2_{P^s_{-,k}})^{1/2},$$

where

$$\|y\|_{Y^s_{-,k}} = (1 + \alpha^{1/2} h_k^{-2})^{(1-s)/2} (\|y\|^2_{L^2(\Omega)} + \alpha^{s/2} \|y\|^2_{H^s(\Omega)})^{1/2} \text{ and}$$
$$\|p\|_{P^s_{-,k}} = \alpha^{-1} (1 + \alpha^{1/2} h_k^{-2})^{(1-s)/2} (\|p\|^2_{L^2(\Omega)} + \alpha^{s/2} \|p\|^2_{H^s(\Omega)})^{1/2}.$$

and

$$\|x\|_{X^s_{+,k}} = (\|y\|^2_{Y^s_{+,k}} + \|p\|^2_{P^s_{+,k}})^{1/2},$$

where

$$\|y\|_{Y^s_{+,k}} = (1+\alpha^{1/2}h_k^{-2})^{-(1-s)/2}(\|y\|^2_{L^2(\Omega)} + \alpha^{(2-s)/2}\|y\|^2_{H^{2-s}(\Omega)})^{1/2} \text{ and}$$
$$\|p\|_{P^s_{+,k}} = \alpha^{-1}(1+\alpha^{1/2}h_k^{-2})^{-(1-s)/2}(\|p\|^2_{L^2(\Omega)} + \alpha^{(2-s)/2}\|p\|^2_{H^{2-s}(\Omega)})^{1/2}.$$

Proof. First note that $X^s_{-,k}$ and $X^s_{+,k}$, defined by (3.2) and (4.5), have product structure. Therefore, it suffices to discuss the y-part and the p-part separately.

First we consider only the norms in $Y^s_{-,k}$ and $Y^s_{+,k}$.

Using the identity (2.8), the reiteration theorem (Theorem 3.2.20, Corollary 3.2.17 in [4]) and Lemma 6.1, we obtain

$$\begin{aligned}\|y\|_{Y^s_{+,k}} &= \|y\|_{[Y_{+,k},Y]_s} \sim \|y\|_{[Y_{+,k},Y_{-,k}]_{s/2}}\\ &= (1+\alpha^{1/2}h_k^{-2})^{-(1-s)/2}\|y\|_{[L^2(\Omega)\cap\alpha^{1/2}H^2(\Omega),L^2(\Omega)]_{s/2}}\\ &\sim (1+\alpha^{1/2}h_k^{-2})^{(s-1)/2}(\|y\|^2_{L^2(\Omega)} + \alpha^{(2-s)/2}\|y\|^2_{H^{2-s}(\Omega)})^{1/2},\end{aligned}$$

where $\sim$ denotes the equivalence of norms, where the constants are independent of h_k and α.

We can show using the same arguments

$$\begin{aligned}\|y\|_{Y^s_{-,k}} &= \|y\|_{[(Y_{-,k}),Y]_s} = \|y\|_{[(1+\alpha^{1/2}h_k^{-2})^{1/2}L^2(\Omega),L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega)]_s}\\ &\sim (1+\alpha^{1/2}h_k^{-2})^{(1-s)/2}(\|y\|^2_{L^2(\Omega)} + \alpha^{s/2}\|y\|_{H^s(\Omega)})^{1/2}.\end{aligned}$$

The same can be done with the norms in $P^s_{+,k}$ and $P^s_{-,k}$. ∎

As mentioned, we use Theorem 4.1 to show the approximation property. Already in Subsection 2.1, we have mentioned that conditions **(A1)** and **(A1a)**, which are independent of the choice of s, are satisfied. We still have to show conditions **(A4)** and **(A3)**.

LEMMA 4.6. *Condition* **(A4)** *is satisfied for all* $s \in [0,1]$.

Proof. This proof is done using interpolation. Condition **(A4)** was shown for $s = 0$ in Lemma 4.3.

For $s = 1$, we can show using (4.3) that

$$\|y - \Pi_k y\|_Y \le C\|y\|_Y \tag{4.6}$$

holds for all $y \in Y$.

Using (4.4) and (4.6), the interpolation theorem (Theorem 2.3) implies that

$$c_s\, \|y-\Pi_k y\|_Y = \|y-\Pi_k y\|_{[Y,Y]_s} \le C\|y\|_{[Y_{+,k},Y]_s} = C\|y\|_{Y^s_{+,k}}$$

holds for all $y \in Y^s_+$ with $c_s = \sqrt{1/(2s(1-s))}$, which finishes the proof for $s \in (0,1)$. ∎

LEMMA 4.7. *Assume that regularity assumption* **(R')** *is satisfied for some* $s \in [0,1)$. *Then condition* **(A3)** *holds with a constant* $\hat{C}_R$ *independent of the grid level* k *and the choice of the parameter* α.

For the proof of this lemma, we need the following lemma, whose proof can be found in the Appendix.

LEMMA 4.8. *Assume that regularity assumption* **(R')** *is satisfied for some* $s \in [0,1)$. *Let* $f \in (H^s(\Omega))^*$. *Consider the problem: Find* $y_f \in H^1(\Omega)$ *such that*

$$(y_f,\widetilde{y})_{L^2(\Omega)} + \alpha^{1/2}(y_f,\widetilde{y})_{H^1(\Omega)} = \langle f,\widetilde{y}\rangle \tag{4.7}$$

is satisfied for all $\widetilde{y} \in H^1(\Omega)$*. Then there is a constant* $\widetilde{C}_R > 0$ *such that for all* $\alpha > 0$ *the solution* y_f *satisfies* $y_f \in H^{2-s}(\Omega)$ *and*

$$\|y_f\|_{Y^s_{+,k}} \leq \widetilde{C}_R \|f\|_{(Y^s_{-,k})^*}$$

holds.

Proof. *(of Lemma 4.7)* This proof is organized as follows. In step 1, we show that for $\mathcal{F} \in (X^s_-)^*$ the solution x of (4.1) satisfies $x \in X^s_+$. In step 2, we will see that the estimate in **(A3)** can be rewritten as a stability estimate. In step 3, we will see that **(R')** implies a stability estimate for a parameter-dependent elliptic problem. In step 4, we will show that this stability estimate implies, following the guidelines of [14] the stability estimate to be shown due to step 3.

Step 1. Considering the two lines of the KKT-system (2.1) separately, and keeping in mind that $y \in H^1(\Omega) \subseteq (H^s(\Omega))^*$ and $p \in H^1(\Omega) \subseteq (H^s(\Omega))^*$, we obtain using **(R')** that $y \in H^{2-s}(\Omega) = Y^s_+$ and $p \in H^{2-s}(\Omega) = P^s_+$ is satisfied.

Step 2. For showing

$$\|x\|_{X^s_{+,k}} \leq \hat{C}_R \|\mathcal{F}\|_{(X^s_{-,k})^*},$$

we reformulate this condition as stability estimate (inf-sup condition):

$$\|x\|_{X^s_{+,k}} \leq \hat{C}_R \sup_{0 \neq \widetilde{x} \in X} \frac{\mathcal{B}(x, \widetilde{x})}{\|\widetilde{x}\|_{X^s_{-,k}}} \qquad \text{for all } x \in X^s_+. \tag{4.8}$$

Using the representation of the norms, as introduced in Lemma 4.5, (4.8) can be rewritten as follows

$$(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2} \leq \hat{C}_R \sup_{0 \neq (\widetilde{y}, \widetilde{p}) \in X} \frac{\mathcal{B}((y,p),(\widetilde{y},\widetilde{p}))}{(\|\widetilde{y}\|^2_{\widetilde{Y}^s_-} + \|\widetilde{p}\|^2_{\widetilde{P}^s_-})^{1/2}} \tag{4.9}$$

where

$$\|\cdot\|_{\widetilde{Y}^s_-} := \|\cdot\|_{L^2(\Omega) \cap \alpha^{s/4} H^s(\Omega)}, \qquad \|\cdot\|_{\widetilde{P}^s_-} := \|\cdot\|_{\alpha^{-1/2} L^2(\Omega) \cap \alpha^{(s-2)/4} H^s(\Omega)},$$
$$\|\cdot\|_{\widetilde{Y}^s_+} := \|\cdot\|_{L^2(\Omega) \cap \alpha^{(2-s)/4} H^{2-s}(\Omega)}, \qquad \|\cdot\|_{\widetilde{P}^s_+} := \|\cdot\|_{\alpha^{-1/2} L^2(\Omega) \cap \alpha^{-s/4} H^{2-s}(\Omega)}.$$

Step 3. Analogously to the case above, also the the regularity statement of Lemma 4.8 can be rewritten as an inf-sup condition.

$$\begin{aligned}
\|y\|_{\widetilde{Y}^s_+} &\leq \widetilde{C}_R \sup_{0 \neq \widetilde{y} \in H^1(\Omega)} \frac{(y, \widetilde{y})_{L^2(\Omega)} + \alpha^{1/2} (y, \widetilde{y})_{H^1(\Omega)}}{\|\widetilde{y}\|_{\widetilde{Y}^s_-}} \\
&\leq \widetilde{C}_R \left(\sup_{0 \neq \widetilde{y} \in H^1(\Omega)} \frac{(y, \widetilde{y})_{L^2(\Omega)}}{\|\widetilde{y}\|_{\widetilde{Y}^s_-}} + \sup_{0 \neq \widetilde{p} \in H^1(\Omega)} \frac{(y, \widetilde{p})_{H^1(\Omega)}}{\|\widetilde{p}\|_{\widetilde{P}^s_-}} \right)
\end{aligned} \tag{4.10}$$

holds for all $y \in H^{2-s}(\Omega)$. We can show completely analogously that also

$$\|p\|_{\widetilde{P}^s_+} \leq \widetilde{C}_R \left(\sup_{0 \neq \widetilde{y} \in H^1(\Omega)} \frac{(p, \widetilde{y})_{H^1(\Omega)}}{\|\widetilde{y}\|_{\widetilde{Y}^s_-}} + \sup_{0 \neq \widetilde{p} \in H^1(\Omega)} \frac{\alpha^{-1}(p, \widetilde{p})_{L^2(\Omega)}}{\|\widetilde{p}\|_{\widetilde{P}^s_-}} \right) \tag{4.11}$$

holds for all $p \in H^{2-s}(\Omega)$.

Step 4. Now we show that (4.10) and (4.11) imply (4.9). The proof follows the lines of the proof of the stability estimate Theorem 2.3 in [14].

Note that $\|\cdot\|_{L^2(\Omega)} \leq \|\cdot\|_{\widetilde{Y}^s_-} \leq \|\cdot\|_{\widetilde{Y}^s_+}$ and $\|\cdot\|_{L^2(\Omega)} \leq \|\cdot\|_{\widetilde{P}^s_-} \leq \|\cdot\|_{\widetilde{P}^s_+}$ are satisfied.

Now

$$
\begin{aligned}
&2 \sup_{0\neq\widetilde{x}\in X} \frac{\mathcal{B}(x,\widetilde{x})}{(\|\widetilde{y}\|^2_{\widetilde{Y}^s_-} + \|\widetilde{p}\|^2_{\widetilde{P}^s_-})^{1/2}} \\
&\quad \geq \sup_{0\neq\widetilde{y}\in Y} \frac{(y,\widetilde{y})_{L^2(\Omega)} + (\widetilde{y},p)_{H^1(\Omega)}}{\|\widetilde{y}\|_{\widetilde{Y}^s_-}} + \sup_{0\neq\widetilde{p}\in P} \frac{(y,\widetilde{p})_{H^1(\Omega)} - \alpha^{-1}(p,\widetilde{p})_{L^2(\Omega)}}{\|\widetilde{p}\|_{\widetilde{P}^s_-}} \\
&\quad \geq \left(\sup_{0\neq\widetilde{y}\in Y} \frac{(\widetilde{y},p)^2_{H^1(\Omega)}}{\|\widetilde{y}\|^2_{\widetilde{Y}^s_-}} + \sup_{0\neq\widetilde{p}\in P} \frac{(y,\widetilde{p})^2_{H^1(\Omega)}}{\|\widetilde{p}\|^2_{\widetilde{P}^s_-}} \right)^{1/2} \\
&\qquad - \left(\sup_{0\neq\widetilde{y}\in Y} \frac{(y,\widetilde{y})^2_{L^2(\Omega)}}{\|\widetilde{y}\|^2_{\widetilde{Y}^s_-}} + \sup_{0\neq\widetilde{p}\in P} \frac{\alpha^{-2}(p,\widetilde{p})^2_{L^2(\Omega)}}{\|\widetilde{p}\|^2_{\widetilde{P}^s_-}} \right)^{1/2} \\
&\quad = (\xi - \eta)(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2},
\end{aligned}
$$

where

$$
\eta := \frac{1}{(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2}} \left(\sup_{0\neq\widetilde{y}\in Y} \frac{(y,\widetilde{y})^2_{L^2(\Omega)}}{\|\widetilde{y}\|^2_{\widetilde{Y}^s_-}} + \sup_{0\neq\widetilde{p}\in P} \frac{\alpha^{-2}(p,\widetilde{p})^2_{L^2(\Omega)}}{\|\widetilde{p}\|^2_{\widetilde{P}^s_-}} \right)^{1/2},
$$

$$
\xi := \frac{1}{(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2}} \left(\sup_{0\neq\widetilde{y}\in Y} \frac{(\widetilde{y},p)^2_{H^1(\Omega)}}{\|\widetilde{y}\|^2_{\widetilde{Y}^s_-}} + \sup_{0\neq\widetilde{p}\in P} \frac{(y,\widetilde{p})^2_{H^1(\Omega)}}{\|\widetilde{p}\|^2_{\widetilde{P}^s_-}} \right)^{1/2}.
$$

A second bound can be constructed as follows

$$
\begin{aligned}
&\sup_{0\neq\widetilde{x}\in X} \frac{\mathcal{B}(x,\widetilde{x})}{(\|\widetilde{y}\|^2_{\widetilde{Y}^s_-} + \|\widetilde{p}\|^2_{\widetilde{P}^s_-})^{1/2}} \geq \frac{\mathcal{B}((y,p),(y,-p))}{(\|y\|^2_{\widetilde{Y}^s_-} + \|-p\|^2_{\widetilde{P}^s_-})^{1/2}} = \frac{(y,y)_{L^2(\Omega)} + \alpha^{-1}(p,p)_{L^2(\Omega)}}{(\|y\|^2_{\widetilde{Y}^s_-} + \|p\|^2_{\widetilde{P}^s_-})^{1/2}} \\
&\quad \geq \frac{(y,y)_{L^2(\Omega)} + \alpha^{-1}(p,p)_{L^2(\Omega)}}{(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2}} = \frac{\sup_{0\neq\widetilde{y}\in Y} \frac{(y,\widetilde{y})^2_{L^2(\Omega)}}{(\widetilde{y},\widetilde{y})_{L^2(\Omega)}} + \sup_{0\neq\widetilde{p}\in P} \frac{\alpha^{-2}(p,\widetilde{p})^2_{L^2(\Omega)}}{\alpha^{-1}(\widetilde{p},\widetilde{p})_{L^2(\Omega)}}}{(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2}} \\
&\quad \geq \frac{\sup_{0\neq\widetilde{y}\in Y} \frac{(y,\widetilde{y})^2_{L^2(\Omega)}}{\|\widetilde{y}\|^2_{\widetilde{Y}^s_-}} + \sup_{0\neq\widetilde{p}\in P} \frac{\alpha^{-2}(p,\widetilde{p})^2_{L^2(\Omega)}}{\|\widetilde{p}\|^2_{\widetilde{P}^s_-}}}{(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2}} = \eta^2(\|y\|^2_{\widetilde{Y}^s_+} + \|p\|^2_{\widetilde{P}^s_+})^{1/2}.
\end{aligned}
$$

The inequalities (4.10) and (4.11) imply $\xi + \eta \geq \widetilde{C}_R > 0$. In the same way as in [14], it follows that there is an upper bound for the constant in (4.9) only depending on $\widetilde{C}_R$. ∎

As we have shown **(A1)**, **(A1a)**, **(A3)** and **(A4)**, we conclude using Theorem 4.1 as follows.

COROLLARY 4.9. *Assume that regularity assumption* **(R')** *is satisfied for some* $s \in [0, 1)$. *Then the approximation property holds with a constant* C_A *independent of the grid level* k *and the choice of the parameter* α.

As mentioned, the combination of approximation property and smoothing property shows the convergence of the two-grid method.

In Subsection 4.2 we have shown the approximation property. In Section 3 we have shown that the preconditioned normal equation smoother and the collective Richardson smoother satisfy the smoothing property. This allows to conclude as follows.

COROLLARY 4.10. *Assume that regularity assumption* **(R')** *is satisfied for some* $s \in [0, 1)$. *Furthermore, assume that the preconditioned normal equation smoother or the collective Richardson smoother is applied.*

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

$$\|x_k^{(1)} - x_k\|_{X^s_{-,k}} \leq q(\nu)\|x_k^{(0)} - x_k\|_{X^s_{-,k}},$$

with $q(\nu) = C_T \nu^{(1-s)/2}$, *where the constant* C_T *is independent of the grid level* k *and of the choice of the parameter* α. *The constant* C_T *may depend on* s.

The convergence of the W-cycle multigrid method follows under weak assumptions.

5. Numerical Results. In this section, we present numerical results to illustrate the convergence theory presented in this paper. The domain Ω was chosen to be the L-shaped domain $\Omega := (0,2)^2 \backslash [1,2)^2$ and the unit square $\Omega := (0,1)^2$. For the L-shaped domain assumption **(R')** holds for $s > \frac{1}{3}$ and for the unit square assumption **(R)** is satisfied.

On the coarsest grid level $k = 0$ the discretization of the unit square was done by decomposing the square into two triangles by connecting the points $(0,0)$ and $(1,1)$. The L-shaped domain was discretized analogously into 6 triangles. The grid levels $k = 1, 2, \ldots$ were in both cases constructed by uniform refinement, i.e., every triangle was decomposed into four subtriangles.

For the simulation a W-cycle multigrid method with ν pre- and ν post-smoothing steps was used. The number of iterations and convergence rates were measured as follows: we start with an random initial error and measure the reduction of the error in each step using the norm $\|\cdot\|_{X_{-,k}}$. The iteration was stopped when the initial error was reduced by a factor of $\epsilon = 10^{-6}$. The convergence rates q is the mean convergence rate in this iteration, i.e.,

$$q = \left(\frac{\|x_k^{(n)} - x_k\|_{X_{-,k}}}{\|x_k^{(0)} - x_k\|_{X_{-,k}}} \right)^{1/n},$$

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 5.1 we compare the convergence rates for the L-shaped domain and for the unit square. We observe that the convergence rates are comparable, i.e., the multigrid method does not suffer from the lack of full regularity. The numerical tests have been done for the preconditioned normal equation smoother (damped with $\tau = 7/16$), the collective Jacobi iteration (damped with $\tau = 3/4$) and the collective Gauss Seidel iteration (without damping). Here, the preconditioned normal equation smoother and the collective Richardson smoother are covered by the convergence theory. Due to the fact that the grid is uniform, collective Jacobi smoother and collective Richardson smoother are practically identical. The collective Gauss Seidel

	$\nu = 1+1$		$\nu = 2+2$		$\nu = 4+4$	
	n	q	n	q	n	q
Ω is unit square						
Preconditioned normal equation	48	0.75	24	0.56	14	0.35
Collective Jacobi smoother	13	0.33	8	0.16	6	0.08
Collective Gauss Seidel smoother	7	0.10	5	0.05	4	0.02
Ω is L-shaped domain						
Preconditioned normal equation	49	0.75	25	0.57	14	0.36
Collective Jacobi smoother	14	0.35	8	0.17	6	0.08
Collective Gauss Seidel smoother	7	0.11	5	0.06	4	0.03

TABLE 5.1
Number of iterations n and convergence rate q on grid level $k = 5$, $\alpha = 1$

	$\alpha = 1$		$\alpha = 10^{-4}$		$\alpha = 10^{-8}$		$\alpha = 10^{-12}$	
	n	q	n	q	n	q	n	q
$k = 5$	49	0.75	50	0.76	41	0.71	50	0.76
$k = 6$	48	0.75	49	0.75	51	0.76	53	0.77
$k = 7$	49	0.75	49	0.75	55	0.77	56	0.78
$k = 8$	49	0.75	49	0.75	51	0.76	44	0.73
$k = 9$	49	0.75	49	0.75	49	0.75	44	0.73

TABLE 5.2
L-shaped domain: Number of iterations n and convergence rate q for preconditioned normal equation smoother *for $\tau = 7/16$ and $\nu = \nu_{pre} + \nu_{post} = 1+1$ smoothing steps*

smoother is not covered by the theory but a quite natural alternative to the collective Jacobi smoother.

In Table 5.1 we observe that the collective point smoothers are faster than the normal equation smoothers. Moreover, we see that also for the L-shaped domain the convergence rates decay faster than $\nu^{-1/2}$ for increasing values of ν, although theory predicts a decay of $\nu^{-1/2}$ only for the full elliptic regularity case.

In Tables 5.2 and 5.3, we see moreover that the convergence rates are robust in the grid level k and in the choice of the parameter α.

Although not covered by the analysis, numerical experiments show that also the V-cycle converges with rates comparable with the convergence rates of the W-cycle method for the model problem.

6. Conclusions. In this paper we gave a convergence proof for an elliptic distributed control model problem which is slightly different to the proof given in [11]. This proof of the present paper has the advantage that it also holds for to domains where full elliptic regularity cannot be guaranteed, which includes non-convex polygonal domains. The generalization of the presented work to other elliptic differential operators is obvious.

Appendix. *Proof. (of Theorem 4.1)* The details of this proof follow Theorems 2.5 and 3.1 in [11].

	$\alpha = 1$		$\alpha = 10^{-4}$		$\alpha = 10^{-8}$		$\alpha = 10^{-12}$	
	n	q	n	q	n	q	n	q
$k = 5$	14	0.35	13	0.34	9	0.21	13	0.33
$k = 6$	13	0.34	13	0.34	12	0.29	13	0.33
$k = 7$	13	0.34	13	0.34	13	0.33	13	0.33
$k = 8$	13	0.34	13	0.34	13	0.34	11	0.26
$k = 9$	13	0.34	13	0.34	13	0.34	11	0.26

TABLE 5.3

L-shaped domain: Number of iterations n and convergence rate q for collective Jacobi smoother *for $\tau = 3/4$ and $\nu = \nu_{pre} + \nu_{post} = 1 + 1$ smoothing steps*

In this proof, for sake of simplicity C is a generic constant that only depends on $\underline{C}$, $\overline{C}$, $\underline{C}_D$, $\overline{C}_D$, C_I and C_R.

Let $x \in X$ and $x_k \in X_k$ be such that

$$\begin{aligned} \mathcal{B}(x, \widetilde{x}) &= \mathcal{F}(\widetilde{x}) \qquad &&\text{for all } \widetilde{x} \in X, \\ \mathcal{B}(x_k, \widetilde{x}_k) &= \mathcal{F}(\widetilde{x}_k) \qquad &&\text{for all } \widetilde{x}_k \in X_k. \end{aligned}$$

First we show that

$$\|x - x_k\|_{X^s_{-,k}} \le C \sup_{0 \ne \widetilde{x} \in X^s_{-,k}} \frac{\mathcal{F}(\widetilde{x})}{\|\widetilde{x}\|_{X^s_{-,k}}} \tag{6.1}$$

holds. The proof of this estimate follows the classical line of arguments: Because of **(A1)** and **(A1a)**, we can estimate the discretization error in the X-norm by the approximation error:

$$\|x - x_k\|_X \le C \inf_{\widetilde{x}_k \in X_k} \|x - \widetilde{x}_k\|_X .$$

Using **(A4)** and **(A3)** we obtain further

$$\|x - x_k\|_X \le C \|\mathcal{F}\|_{(X^s_{-,k})^*} .$$

For the estimate in the norm $\| \cdot \|_{X^s_{-,k}}$, we use the Aubin-Nitsche duality trick: For every (arbitrarily but fixed) $\mathcal{F}^* \in (X^s_{-,k})^*$, we consider the following problem: Find $\hat{x}_{\mathcal{F}^*} \in X$ such that

$$\mathcal{B}(\widetilde{x}, \hat{x}_{\mathcal{F}^*}) = \mathcal{F}^*(\widetilde{x}) \qquad \text{for all } \widetilde{x} \in X.$$

Using Galerkin orthogonality, we obtain

$$\mathcal{F}^*(x - x_k) = \mathcal{B}(x - x_k, \hat{x}_{\mathcal{F}^*}) = \mathcal{B}(x - x_k, \hat{x}_{\mathcal{F}^*} - \hat{x}_k)$$

for all $\hat{x}_k \in X_k$. Using **(A1)** and **(A1a)**, we obtain

$$\mathcal{F}^*(x - x_k) \le C \|x - x_k\|_X \inf_{\hat{x}_k \in X_k} \|\hat{x}_{\mathcal{F}^*} - \hat{x}_k\|_X .$$

As above we obtain

$$\mathcal{F}^*(x - x_k) \le C \|x - x_k\|_X \|\mathcal{F}^*\|_{(X^s_{-,k})^*} ,$$

which implies (as we may choose $\mathcal{F}^*$ arbitrarily)

$$\|x - x_k\|_{X^s_{-,k}} = \sup_{0 \neq \mathcal{F}^* \in (X^s_{-,k})^*} \frac{\mathcal{F}^*(x - x_k)}{\|\mathcal{F}^*\|_{(X^s_{-,k})^*}} \leq C \|x - x_k\|_X,$$

which shows (6.1). Now we may show the approximation property

$$\|x_k^{(1)} - x_k\|_{X^s_{-,k}} \leq C_A \sup_{0 \neq \widetilde{x}_k \in X_k} \frac{\mathcal{B}(x_k^{(0,m)} - x_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X^s_{-,k}}}.$$

One easily sees that $x_k - x_k^{(1)} = t_k - t_{k-1}$, where $t_k := x_k - x_k^{(0,m)}$ and the function $t_{k-1} \in X_{k-1}$ is given by the formula for the coarse-grid correction step in variational formulation

$$\mathcal{B}(t_{k-1}, \widetilde{x}_{k-1}) = \mathcal{F}(\widetilde{x}_{k-1}) - \mathcal{B}(x_k^{(0,m)}, \widetilde{x}_{k-1}) \qquad \text{for all } \widetilde{x}_{k-1} \in X_{k-1}.$$

We observe that

$$\mathcal{B}(t_{k-1}, \widetilde{x}_{k-1}) = \mathcal{F}(\widetilde{x}_{k-1}) - \mathcal{B}(x_k^{(0,m)}, \widetilde{x}_{k-1}) = \mathcal{B}(x_k - x_k^{(0,m)}, \widetilde{x}_{k-1}) = \mathcal{B}(t_k, \widetilde{x}_{k-1}) \tag{6.2}$$

for all $\widetilde{x}_{k-1} \in X_{k-1}$. For a given $\mathcal{F}^* \in (X^s_{-,k})^*$, let $\hat{x} \in X$, $\hat{x}_k \in X_k$ and $\hat{x}_{k-1} \in X_{k-1}$ satisfy

$$\begin{aligned} \mathcal{B}(\widetilde{x}, \hat{x}) &= \mathcal{F}^*(\widetilde{x}) && \text{for all } \widetilde{x} \in X, \\ \mathcal{B}(\widetilde{x}_k, \hat{x}_k) &= \mathcal{F}^*(\widetilde{x}_k) && \text{for all } \widetilde{x}_k \in X_k, \\ \mathcal{B}(\widetilde{x}_{k-1}, \hat{x}_{k-1}) &= \mathcal{F}^*(\widetilde{x}_{k-1}) && \text{for all } \widetilde{x}_{k-1} \in X_{k-1}. \end{aligned}$$

Then

$$\mathcal{F}^*(t_k - t_{k-1}) = \mathcal{B}(t_k - t_{k-1}, \hat{x}_k) = \mathcal{B}(t_k, \hat{x}_k - \hat{x}_{k-1})$$

since

$$\mathcal{B}(t_{k-1}, \hat{x}_k) = \mathcal{F}^*(t_{k-1}) = \mathcal{B}(t_{k-1}, \hat{x}_{k-1}) = \mathcal{B}(t_k, \hat{x}_{k-1})$$

using (6.2). Hence

$$\mathcal{F}^*(t_k - t_{k-1}) \leq \sup_{0 \neq \widetilde{x}_k \in X_k} \frac{\mathcal{B}(t_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X^s_{-,k}}} \|\mathcal{F}^*\|_{(X^s_{-,k})^*}.$$

Therefore,

$$\|t_k - t_{k-1}\|_{X^s_{-,k}} = \sup_{0 \neq \mathcal{F}^* \in (X^s_{-,k})^*} \frac{\mathcal{F}^*(t_k - t_{k-1})}{\|\mathcal{F}^*\|_{(X^s_{-,k})^*}} \leq C \sup_{0 \neq \widetilde{x}_k \in X_k} \frac{\mathcal{B}(t_k, \widetilde{x}_k)}{\|\widetilde{x}_k\|_{X^s_{-,k}}},$$

which completes the proof. ∎

LEMMA 6.1. *For all Banach spaces A_1 and A_2, the Banach spaces $A_1 \cap [A_2, A_1]_\theta$ and $[A_1 \cap A_2, A_1]_\theta$ are equal and have equivalent norms. The constants, describing the equivalence, only depend on the choice of θ.*

Proof. In this proof, $C > 0$ is a generic constant which is independent of k and α but which may depend on θ.

First note that

$$\|u\|_{[A_1\cap A_2,A_1]_\theta} \geq C\|u\|_{[A_1,A_1]_\theta\cap[A_2,A_1]_\theta} \geq C\|u\|_{A_1\cap[A_2,A_1]_\theta}$$

follows directly from the monotonicity of the interpolation.

So it remains to show $\|u\|_{[A_1\cap A_2,A_1]_\theta} \leq C\|u\|_{A_1\cap[A_2,A_1]_\theta}$. Let $u \in [A_1\cap A_2]_\theta \cap A_1$. The definition of the norms on the interpolation spaces (real K-method, cf. [9]) and elementary relations yield

$$\begin{aligned}
\|u\|^2_{[A_1\cap A_2,A_1]_\theta} &= \int_0^\infty t^{-2\theta-1} \inf_{u_1+u_2=u} \left(\|u_1\|^2_{A_1} + \|u_1\|^2_{A_2} + t^2\|u_2\|^2_{A_1}\right) \mathrm{d}t \\
&= \int_0^1 t^{-2\theta-1} \inf_{u_1+u_2=u} \left(\|u_1\|^2_{A_1} + \|u_1\|^2_{A_2} + t^2\|u_2\|^2_{A_1}\right) \mathrm{d}t \\
&\quad + \int_1^\infty t^{-2\theta-1} \inf_{u_1+u_2=u} \left(\|u_1\|^2_{A_1} + \|u_1\|^2_{A_2} + t^2\|u_2\|^2_{A_1}\right) \mathrm{d}t.
\end{aligned}$$

By replacing the infimum by a particular choice, using the triangular inequality and by computing the integrals, we obtain

$$\begin{aligned}
\|u\|^2_{[A_1\cap A_2,A_1]_\theta} &\leq \int_0^1 t^{-2\theta-1} t^2 \|u\|^2_{A_1} \mathrm{d}t \\
&\quad + \int_1^\infty t^{-2\theta-1} \inf_{u_1+u_2=u} \left((\|u\|_{A_1} + \|u_2\|_{A_1})^2 + \|u_1\|^2_{A_2} + t^2\|u_2\|^2_{A_1}\right) \mathrm{d}t \\
&\leq \frac{1}{2-2\theta}\|u\|^2_{A_1} + \frac{1}{\theta}\|u\|^2_{A_1} + 2\int_1^\infty t^{-2\theta-1} \left(\inf_{u_1+u_2=u} \|u_1\|^2_{A_2} + (1+t)^2\|u_2\|^2_{A_1}\right) \mathrm{d}t.
\end{aligned}$$

By a variable transformation and again using the definition of the norms on the interpolation spaces, we obtain that further

$$\begin{aligned}
\|u\|^2_{[A_1\cap A_2,A_1]_\theta} &\leq \frac{1}{(1-\theta)\theta}\|u\|^2_{A_1} \\
&\quad + 2\left(\frac{1}{2}\right)^{-2\theta-1} \int_1^\infty (1+t)^{-2\theta-1} \inf_{u_1+u_2=u} \left(\|u_1\|^2_{A_2} + (1+t)^2\|u_2\|^2_{A_1}\right) \mathrm{d}t \\
&= \frac{1}{(1-\theta)\theta}\|u\|^2_{A_1} + 2^{2\theta+2} \int_2^\infty t^{-2\theta-1} \inf_{u_1+u_2=u} \left(\|u_1\|^2_{A_2} + t^2\|u_2\|^2_{A_1}\right) \mathrm{d}t \\
&\leq \frac{1}{(1-\theta)\theta}\|u\|^2_{A_1} + 2^{2\theta+2}\|u\|^2_{[A_2,A_1]_\theta} \leq C(\theta)^2\|u\|^2_{[A_2,A_1]_\theta\cap A_1}
\end{aligned}$$

holds, which finishes the proof for $C(\theta) = \max\left\{(1-\theta)^{-1/2}\theta^{-1/2}, 2^{\theta+1}\right\}$. ∎

Proof. (of Lemma 4.8) We have to show that $y_f \in H^{2-s}(\Omega)$ and

$$\|y_f\|_{L^2(\Omega)\cap\alpha^{(2-s)/4}H^{2-s}(\Omega)} \leq \widetilde{C}_R\|f\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*}.$$

As $y_f \in H^1(\Omega)$, we have using $y_f \equiv (y_f,\cdot)_{L^2(\Omega)}$ that $y_f \in (H^s(\Omega))^*$. Therefore, if we consider the problem, find $y \in Y$ such that

$$(y,\widetilde{y})_{H^1(\Omega)} = \langle \alpha^{-1/2}(f-y_f), \widetilde{y}\rangle \qquad \text{holds for all } \widetilde{y} \in H^1(\Omega),$$

that regularity assumption **(R')** states $y_f \in H^{2-s}(\Omega)$ and

$$\|y_f\|_{H^{2-s}(\Omega)} \leq C_R(\|f\|_{(H^s(\Omega))^*} + \alpha^{-1/2}\|y_f\|_{(H^s(\Omega))^*}).$$

This can be bounded from above and we obtain

$$\|y_f\|_{H^{2-s}(\Omega)} \le C(\alpha)\|f\|_{(H^s(\Omega))^*},$$

where $C(\alpha)$ is some constant that may depend on α.

Now, in a second step we construct a result that is robust in α. Let $f \in L^2(\Omega)$ be arbitrarily but fixed.

We may consider the following formulation of the problem: Find $y \in H^1(\Omega)$ such that

$$(y,\widetilde{y})_{H^1(\Omega)} = (\alpha^{-1/2}(f-y_f),\widetilde{y})_{L^2(\Omega)} \qquad \text{holds for all } \widetilde{y} \in H^1(\Omega).$$

Lax-Milgram Theorem (applied directly to the energy norm $\|\cdot\|_{H^1(\Omega)}$) shows that the solution $y_f \in H^1(\Omega)$ satisfies

$$\|y_f\|_{H^1(\Omega)} = \alpha^{-1/2}\|f-y_f\|_{(H^1(\Omega))^*} \tag{6.3}$$

and the regularity assumption **(R')** shows that $y_f \in H^{2-s}(\Omega)$ and

$$\|y_f\|_{H^{2-s}(\Omega)} \le C_R\alpha^{-1/2}\|f-y_f\|_{(H^s(\Omega))^*}. \tag{6.4}$$

We may also consider the following formulation of the problem: Find $y \in H^1(\Omega)$ such that

$$(y,\widetilde{y})_{L^2(\Omega)} + \alpha^{1/2}(y,\widetilde{y})_{H^1(\Omega)} = (f,\widetilde{y})_{L^2(\Omega)} \qquad \text{holds for all } \widetilde{y} \in H^1(\Omega). \tag{6.5}$$

The Lax-Milgram theorem (applied for the energy norm $\|\cdot\|_{L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega)}$) shows that the solution y_f satisfies

$$\|y_f\|_{L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega)} = \|f\|_{(L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega))^*}. \tag{6.6}$$

The combination of (6.3) and (6.6) shows:

$$\|f-y_f\|_{(\alpha^{1/4}H^1(\Omega))^*} \le C\|f\|_{(L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega))^*}. \tag{6.7}$$

If we choose $\widetilde{y} = y = y_f$ in (6.5), we obtain using $\alpha^{1/2}(y_f,\widetilde{y})_{H^1(\Omega)} \ge 0$ that $\|y_f\|^2_{L^2(\Omega)} \le \|f\|_{L^2(\Omega)}\|y_f\|_{L^2(\Omega)}$ and therefore $\|y_f\|_{L^2(\Omega)} \le \|f\|_{L^2(\Omega)}$and therefore

$$\|f-y_f\|_{L^2(\Omega)} \le C\|f\|_{L^2(\Omega)}. \tag{6.8}$$

The combination of (6.7), (6.8), the interpolation theorem (Theorem 2.3) and Lemma 6.1 shows:

$$\|f-y_f\|_{(\alpha^{s/4}H^s(\Omega))^*} \le C\|f\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*},$$

which reads, if combined with (6.4), as follows:

$$\|y_f\|_{\alpha^{(2-s)/4}H^{2-s}(\Omega)} \le C\|f\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*}. \tag{6.9}$$

The equation (6.6) implies $\|y_f\|_{L^2(\Omega)} \le \|f\|_{(L^2(\Omega)\cap\alpha^{1/4}H^1(\Omega))^*}$,which shows using the fact $\|\cdot\|_{[A,B]_s} \le \|\cdot\|_{A\cap B}$ that $\|y_f\|_{L^2(\Omega)} \le \|f\|_{(L^2(\Omega)\cap\alpha^{1/4}H^s(\Omega))^*}$,which can be combined with (6.9) to the desired result:

$$\|y_f\|_{L^2(\Omega)\cap\alpha^{(2-s)/4}H^{2-s}(\Omega)} \le C\|f\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*}$$

for all $f \in L^2(\Omega)$.

Due to the fact that $L^2(\Omega)$ is dense in $(H^s(\Omega))^*$ we have for $f_0 \in (H^s(\Omega))^*$ and $f_\epsilon \in L^2(\Omega)$ with $\|f_0 - f_\epsilon\|_{(H^s(\Omega))^*} \leq \epsilon$ that

$$\begin{aligned}
&\|y_{f_0}\|_{L^2(\Omega)\cap\alpha^{(2-s)/4}H^{2-s}(\Omega)} \\
&\qquad \leq \|y_{f_\epsilon}\|_{L^2(\Omega)\cap\alpha^{(2-s)/4}H^{2-s}(\Omega)} + \|y_{f_\epsilon} - y_{f_0}\|_{L^2(\Omega)\cap\alpha^{(2-s)/4}H^{2-s}(\Omega)} \\
&\qquad \leq C\|f_\epsilon\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*} + C(\alpha)\|f_\epsilon - f_0\|_{(H^s(\Omega))^*} \\
&\qquad \leq C\|f_0\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*} + (1 + C(\alpha))\|f_\epsilon - f_0\|_{(H^s(\Omega))^*} \\
&\qquad \leq C\|f_0\|_{(L^2(\Omega)\cap\alpha^{s/4}H^s(\Omega))^*} + (1 + C(\alpha))\epsilon
\end{aligned}$$

holds, which shows the desired result for $\epsilon \to 0$. Here, y_{f_0} and y_{f_ϵ} are the solutions of the variational problem (4.7) for right-hand-sides f_0 and f_ϵ, respectively. ∎

REFERENCES

[1] R. Adams and J. Fournier. *Sobolev Spaces*. Academic Press, 2008. 2nd ed.

[2] A. Borzi, K. Kunisch, and D.Y. Kwak. Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system. *SIAM J. on Control and Optimization*, 41(5):1477 – 1497, 2003.

[3] S.C. Brenner. Multigrid methods for parameter dependent problems. *RAIRO, Modélisation Math. Anal. Numér*, 30:265 – 297, 1996.

[4] P. Butzer and H. Berens. *Semi-Groups of Operators and Approximation*. Springer-Verlag, Berlin Heidelberg New York, 1967.

[5] M. Dauge. Elliptic boundary value problems on corner domains. Smoothness and asymptotics of solutions. *Lecture Notes in Mathematics, 1341. Berlin etc.: Springer-Verlag*, 1988.

[6] M. Dauge. Neumann and mixed problems on curvilinear polyhedra. *Integral Equations Oper. Theory*, 15:227 – 261, 1992.

[7] Grisvard, P. *Singularities in Boundary Value Problems*. Springer-Verlag, Berlin Heidelberg New York London, 1992.

[8] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer, Berlin, 1985.

[9] J.L. Lions and Magenes. *Non-Homogeneous Boundary Value Problems and Applications*. Berlin-Heidelberg-New York: Springer-Verlag, 1972.

[10] J. Necas. *Les méthodes directes en théorie des équations elliptiques*. Masson, Paris, 1967.

[11] J. Schöberl, R. Simon, and W. Zulehner. A Robust Multigrid Method for Elliptic Optimal Control Problems. *SIAM J. on Numerical Analysis*, 49:1482 – 1503, 2011.

[12] S. Takacs and W. Zulehner. Convergence Analysis of Multigrid Methods with Collective Point Smoothers for Optimal Control Problems. *Computing and Visualization in Science*, 14(3):131–141, 2011.

[13] U. Trottenberg, C. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London, 2001.

[14] W. Zulehner. Non-standard Norms and Robust Estimates for Saddle Point Problems. *SIAM J. on Matrix Anal. & Appl*, 32:536 – 560, 2011.

Technical Reports of the Doctoral Program
"Computational Mathematics"

2012

2012-01 M.T. Khan: *Formal Semantics of* MiniMaple January 2012. Eds.: W. Schreiner, F. Winkler

2012-02 M. Kollmann, W. Zulehner: *A Robust Preconditioner for Distributed Optimal Control for Stokes Flow with Control Constraints* January 2012. Eds.: U. Langer, R. Ramlau

2012-03 W. Krendl, V. Simoncini, W. Zulehner: *Stability Estimates and Structural Spectral Properties of Saddle Point Problems* February 2012. Eds.: U. Langer, V. Pillwein

2012-04 V. Pillwein, S. Takacs: *A local Fourier convergence analysis of a multigrid method using symbolic computation* April 2012. Eds.: M. Kauers, W. Zulehner

2012-05 I. Georgieva, C. Hofreither: *Tomographic Reconstruction of Harmonic Functions* April 2012. Eds.: U. Langer, V. Pillwein

2012-06 M.T. Khan: *Formal Semantics of a Specification Language for* MiniMaple April 2012. Eds.: W. Schreiner, F. Winkler

2012-07 M. Borges-Quintana, M.A. Borges-Trenard, I. Márquez-Corbella and E. Martínez-Moro: *Computing coset leaders and leader codewords of binary codes* May 2012. Eds.: F. Winkler, V. Pillwein

2012-08 S. Takacs and W. Zulehner: *Convergence analysis of all-at-once multigrid methods for elliptic control problems under partial elliptic regularity* June 2012. Eds.: U. Langer, R. Ramlau

The complete list since 2009 can be found at
`https://www.dk-compmath.jku.at/publications/`

Doctoral Program
“Computational Mathematics”

Director:

Prof. Dr. Peter Paule
Research Institute for Symbolic Computation

Deputy Director:

Prof. Dr. Bert Jüttler
Institute of Applied Geometry

Address:

Johannes Kepler University Linz
Doctoral Program “Computational Mathematics”
Altenbergerstr. 69
A-4040 Linz
Austria
Tel.: ++43 732-2468-6840

E-Mail:

office@dk-compmath.jku.at

Homepage:

http://www.dk-compmath.jku.at