JOHANNES KEPLER UNIVERSITY LINZ

Institute of Computational Mathematics

A Robust Multigrid Method for Elliptic Optimal Control Problems

Joachim Schöberl
Center for Computational Engineering Science, RWTH Aachen
52062 Aachen, Germany

René Simon
Faculty of Mathematics, University Duisburg-Essen
47057 Duisburg, Germany

Walter Zulehner
Institute of Computational Mathematics, Johannes Kepler University
Altenberger Str. 69, 4040 Linz, Austria

NuMa-Report No. 2010-01 January 2010

A–4040 LINZ, Altenbergerstraße 69, Austria

Technical Reports before 1998:

1995

95-1 Hedwig Brandstetter
Was ist neu in Fortran 90? March 1995

95-2 G. Haase, B. Heise, M. Kuhn, U. Langer
Adaptive Domain Decomposition Methods for Finite and Boundary Element Equations. August 1995

95-3 Joachim Schöberl
An Automatic Mesh Generator Using Geometric Rules for Two and Three Space Dimensions. August 1995

1996

96-1 Ferdinand Kickinger
Automatic Mesh Generation for 3D Objects. February 1996

96-2 Mario Goppold, Gundolf Haase, Bodo Heise und Michael Kuhn
Preprocessing in BE/FE Domain Decomposition Methods. February 1996

96-3 Bodo Heise
A Mixed Variational Formulation for 3D Magnetostatics and its Finite Element Discretisation. February 1996

96-4 Bodo Heise und Michael Jung
Robust Parallel Newton-Multilevel Methods. February 1996

96-5 Ferdinand Kickinger
Algebraic Multigrid for Discrete Elliptic Second Order Problems. February 1996

96-6 Bodo Heise
A Mixed Variational Formulation for 3D Magnetostatics and its Finite Element Discretisation. May 1996

96-7 Michael Kuhn
Benchmarking for Boundary Element Methods. June 1996

1997

97-1 Bodo Heise, Michael Kuhn and Ulrich Langer
A Mixed Variational Formulation for 3D Magnetostatics in the Space $H(rot) \cap H(div)$ February 1997

97-2 Joachim Schöberl
Robust Multigrid Preconditioning for Parameter Dependent Problems I: The Stokes-type Case. June 1997

97-3 Ferdinand Kickinger, Sergei V. Nepomnyaschikh, Ralf Pfau, Joachim Schöberl
Numerical Estimates of Inequalities in $H^{\frac{1}{2}}$. August 1997

97-4 Joachim Schöberl
Programmbeschreibung NAOMI 2D und Algebraic Multigrid. September 1997

From 1998 to 2008 technical reports were published by SFB013. Please see
`http://www.sfb013.uni-linz.ac.at/index.php?id=reports`
From 2004 on reports were also published by RICAM. Please see
`http://www.ricam.oeaw.ac.at/publications/list/`
For a complete list of NuMa reports see
`http://www.numa.uni-linz.ac.at/Publications/List/`

A ROBUST MULTIGRID METHOD FOR ELLIPTIC OPTIMAL CONTROL PROBLEMS

JOACHIM SCHÖBERL *, RENÉ SIMON †, AND WALTER ZULEHNER ‡

Abstract. We consider the discretized optimality system of a special class of elliptic optimal control problems and propose an all-at-once multigrid method for solving this discretized system. Under standard assumptions the convergence of the multigrid method and the robustness of the convergence rates with respect to the involved parameter are shown. Numerical experiments are presented for illustrating the theoretical results.

Key words. multigrid methods, all-at-once methods, robust methods, elliptic optimal control, PDE constrained optimization

AMS subject classifications. 65N22, 65N55, 49M25

1. Introduction. In this paper we consider the following optimization problem: Minimize the cost functional $J(y, u)$, given by

$$J(y,u) = \frac{1}{2} \int_\Omega (y - y_d)^2 \, dx + \frac{\alpha}{2} \int_\Omega u^2 \, dx,$$

subject to the elliptic boundary value problem (the state equation)

$$-\Delta y + y = u \quad \text{in } \Omega, \quad \frac{\partial y}{\partial n} = 0 \quad \text{on } \Gamma,$$

for y (the state) with u (the control) on the right-hand side of the differential equation. Here Ω is a given bounded domain in $\mathbb{R}^d$, $d \in \{1, 2, 3\}$, with Lipschitz-continuous boundary Γ, Δ denotes the Laplacian operator, $\partial y/\partial n$ denotes the normal derivative, y_d is given (the desired state) and α is a given real and positive parameter. The parameter α could be seen either as a measure for the costs of the control or as a regularization parameter of the otherwise ill-posed problem.

The main purpose of this paper is to construct and analyze numerical methods which produce an approximate solution to the optimization problem of reasonable accuracy and whose computational efficiency does not significantly depend on the parameter α, in particular for small values of α.

The optimization problem stated above has been quite frequently considered in literature. It is a simple example of an elliptic optimal control problem with distributed control and distributed observation. It serves as a model case for more general elliptic optimal control problems including boundary control and/or boundary observations, see, e.g., [18], [26]. Optimal control problems belong to the class of PDE (partial differential equation) constrained optimization problems, which arise in many areas of sciences and engineering, see [17].

Such problems from applications are by some order of magnitude more challenging than the simple model problem, due to possible nonlinearities and/or additional

*Center for Computational Engineering Science, RWTH Aachen, 52062 Aachen, Germany (joachim.schoeberl@rwth-aachen.de).

†Faculty of Mathematics, University Duisburg-Essen, 47057 Duisburg, Germany (rene.simon@uni-due.de).

‡Institute of Computational Mathematics, Johannes Kepler University, 4040 Linz, Austria (zulehner@numa.uni-linz.ac.at). The work of the last two authors was supported in part by the Austrian Science Fund (FWF) under the grant SFB F013/F1309.

inequality constraints on the control and/or state. Various techniques (like linearization and modern penalty methods) have been developed to cope with such additional difficulties. These methods require to solve a sequence of simpler problems, which are (to some extend) comparable in complexity with the simple model problem, and which also typically involve parameters like the parameter α here, where robustness of the method with respect to these parameters is an important issue. So, a better understanding of methods for the simple model problem will hopefully (in the long run) also contribute to a better understanding of the design of methods for more challenging classes of problems.

Appropriate function spaces for the model problem are

$$V = H^1(\Omega) \quad \text{and} \quad \mathscr{U} = L^2(\Omega)$$

for the state y and for the control u, respectively. The state equation is considered to be given in weak form:

$$\int_\Omega (\nabla y \cdot \nabla z + y\, z)\ dx = \int_\Omega u\, z\ dx \quad \text{for all } z \in V,$$

where ∇ denotes the gradient operator and the symbol $\cdot$ denotes the Euclidean inner product in $\mathbb{R}^d$.

A particular feature of the model problem, which will be exploited in this paper, is the form of the cost functional (distributed observation) and the right-hand side of the weak form of the state equation (distributed control). The bilinear form on the left-hand side of the state equation results from the particular differential operator and boundary conditions of the state equation, which were chosen for simplicity only. The presented numerical method for solving the model problem and the analysis of the method easily carry over to more general classes of state equations as long as their weak formulation leads to a symmetric, coercive and bounded bilinear form.

The model problem has a unique solution $(y, u) \in V \times \mathscr{U}$, which is characterized by the following system (the optimality system, Karush-Kuhn-Tucker system, in short KKT system), see [18]:

- the adjoint state equation:

$$\int_\Omega y\, z\ dx + \int_\Omega (\nabla z \cdot \nabla p + z\, p)\ dx = \int_\Omega y_d\, z\ dx \quad \text{for all } z \in V, \tag{1.1}$$

  which is the weak form of the elliptic boundary value problem

$$-\Delta p + p = y - y_d \quad \text{in } \Omega, \quad \frac{\partial p}{\partial n} = 0 \quad \text{on } \Gamma,$$

  for the new unknown p (the Lagrangian multiplier or adjoint state), which lies in the function space

$$Q = H^1(\Omega);$$

- the control equation:

$$\alpha \int_\Omega u\, v\ dx - \int_\Omega v\, p\ dx = 0 \quad \text{for all } v \in \mathscr{U}, \tag{1.2}$$

  which, of course, simply means

$$\alpha\, u = p,$$

- the state equation:

$$\int_\Omega (\nabla y \cdot \nabla q + y\, q)\ dx - \int_\Omega u\, q\ dx = 0 \quad \text{for all } q \in Q = H^1(\Omega). \tag{1.3}$$

Observe that the set $Q = H^1(\Omega)$ for the Lagrangian multiplier coincides with the previously introduced set $V = H^1(\Omega)$ for the state. The reason for using different symbols is that, later on, we will see that it is advantageous to equip these identical sets with different norms.

As we have seen, our model problem allows a natural partitioning of the variables into three parts (the state y, the control u and the adjoint state p) and a corresponding partitioning of the optimality system. There are two other equivalent formulations which are frequently used as starting points for the development of methods. The particularly simple control equation allows an explicit elimination of the control u, leading to a partitioning of the remaining variables into two parts, y and p, and a correspondingly reduced system of two equations, which consists of (1.1) and (1.3) with u replaced by $\alpha^{-1}\, p$. Alternatively, since the state equation is uniquely solvable in y for any u (in short $y = y(u)$) and the adjoint state equation is uniquely solvable in p for any y (in short $p = p(y)$), the optimality system can be reduced to a single equation in u:

$$\alpha\, u = p(y(u)). \tag{1.4}$$

Starting point for discretizing the problem is either the original formulation as an optimization problem or one of the three possible formulations of the optimality system (either as a system of three equations in (y, u, p), or a system of two equations in (y, p) or a single equation in u). Following the approach, taken, e.g., in [7], we choose the reduced optimality system in (y, p). We consider the standard finite element discretization as a Galerkin method with continuous piecewise linear finite elements on a simplicial mesh for both the state y and the adjoint state p. Due to the variational form of the optimality system this corresponds exactly to the Ritz method applied to the original optimization problem with the same finite element functions for the state y and the control u, and also to the Galerkin method applied to the unreduced optimality system with the same finite element functions for the state y, the control u and the adjoint state p.

The main focus of this paper is the construction of efficient numerical methods for solving the discretized problem and a rigorous analysis of the proposed methods. Since we are interested in accurate approximations, the discretization process will result in a large-scale system, in general. We will concentrate on multigrid methods, which have been proven to be one of the fastest known methods for solving large-scale systems of discretized problems. Originally designed and analyzed for elliptic problems this class of methods has gained growing interest for PDE constrained optimization problems over the last years, see [8] and the many references contained there. There are (at least) two possible approaches: Either the multigrid idea is directly applied to the (reduced or not reduced) optimality system (all-at-once approach), see, e.g., [25], [1], [7], [22], [24], or (as the second approach) multigrid methods are applied to single equations in y and/or in p and/or in u as building blocks of the overall iterative method, see, e.g., [3], [4], [5], [6], [16], [21], [20]. The two approaches are indistinguishable if applied to the single equation (1.4), which was the starting point for a multigrid method in [14]. The method in [14] relies on available exact (or, at least, accurate) solvers for the state and the adjoint state equation with respect to y and p, respectively.

An obvious advantage of the second approach is that existing and well evaluated methods off the shelf, so to speak, can be used as building blocks. On the other hand the all-at-once approach offers a stand-alone solution, possibly well adapted to the problem class. In this paper we will focus on the all-at-once approach for the reduced optimality system in (y, p).

So far, the analysis of all-at-once multigrid methods for PDE constrained optimization problems is not as developed as for elliptic PDEs. Besides numerously reported numerical evidence on the efficiency of the multigrid methods the analytical knowledge is mainly based either on a Fourier analysis, see, e.g., [1], [7], which, strictly speaking, covers only the case of uniform meshes with special boundary conditions (and small perturbations of this situation) or on compactness arguments, see [14], [7], which, however, (at least theoretically) require that the coarsest mesh is sufficiently fine. A more rigorous discussion following the classical approach by showing the approximation property and the smoothing property, see [15], are contained in [22], based on the concept of transforming smoothers but not yet fully covering the model problem, and in [24], where the dependence of the convergence rates on the parameter α for small values of α was not satisfactory in theory as well as in practice.

In this paper we will show, as expected for multigrid methods, mesh-independent convergence of the proposed method for a hierarchy of uniformly refined meshes. Moreover, as a special feature of our approach, we were also able to additionally show robustness of the convergence rates with respect to the parameter α.

The paper is organized as follows: Section 2 contains the framework and basic estimates for the exact and the approximate solution of the problem which are needed in the subsequent multigrid convergence analysis. The multigrid method and a discussion of the approximation property are presented in Section 3, while in Section 4 the analysis is completed by discussing the smoother needed in the multigrid method and the smoothing property. Numerical experiments, which confirm and illustrate the theoretical results, are shown in Section 5. Two technical lemmas needed in Section 2 are transferred to an appendix.

2. Framework and basic estimates. The optimality system can be reformulated as a variational problem in the product space $\hat{X} = V \times \mathscr{U} \times Q$ by simply adding the equations (1.1), (1.2), (1.3): Find $(y, u, p) \in \hat{X}$ such that

$$\hat{\mathcal{B}}((y, u, p), (z, v, q)) = (y_d, z)_{L^2(\Omega)} \quad \text{for all } (z, v, q) \in \hat{X} \tag{2.1}$$

with the bilinear form

$$\begin{aligned}&\hat{\mathcal{B}}((y, u, p), (z, v, q)) = (y, z)_{L^2(\Omega)} + (z, p)_{H^1(\Omega)}\\ &+\alpha\,(u, v)_{L^2(\Omega)} - (v, p)_{L^2(\Omega)} + (y, q)_{H^1(\Omega)} - (u, q)_{L^2(\Omega)},\end{aligned}$$

written in terms of the standard inner products in $L^2(\Omega)$ and $H^1(\Omega)$:

$$(f, g)_{L^2(\Omega)} = \int_\Omega f\, g\ dx, \quad (f, g)_{H^1(\Omega)} = \int_\Omega (\nabla f \cdot \nabla g + f\, g)\ dx.$$

Well-posedness. In [21] it was shown that, for the norm on $\hat{X} = V \times \mathscr{U} \times Q$, given by

$$\|(z, v, q)\|_{\hat{X}}^2 = \|z\|_V^2 + \|v\|_{\mathscr{U}}^2 + \|q\|_Q^2$$

with

$$\|z\|_V^2 = \|z\|_{L^2(\Omega)}^2 + \alpha^{1/2} \|z\|_{H^1(\Omega)}^2, \quad \|v\|_{\mathscr{U}} = \alpha^{1/2} \|v\|_{L^2(\Omega)}$$

and

$$\|q\|_Q^2 = \alpha^{-1} \|q\|_{L^2(\Omega)}^2 + \alpha^{-1/2} \|q\|_{H^1(\Omega)}^2,$$

we have

$$\hat{c}_1 \|(y,u,p)\|_{\hat{X}} \le \sup_{0 \ne (z,v,q) \in \hat{X}} \frac{\hat{\mathcal{B}}((y,u,p),(z,v,q))}{\|(z,v,q)\|_{\hat{X}}} \le \hat{c}_2 \|(y,u,p)\|_{\hat{X}} \tag{2.2}$$

for all $(y,u,p) \in \hat{X}$ with constants $\hat{c}_1, \hat{c}_2 > 0$ independent of the parameter α.

So, the bilinear form is bounded and satisfies an inf-sup condition uniformly in α. By the theorem of Babuška-Aziz, see [2], it follows that a unique solution to the variational problem (2.1) exists and that the solution depends continuously on the data, uniformly in α.

In this paper we will use the simple form of the control equation (1.2) for explicitly eliminating the control from the optimality system. This leads to the reduced problem consisting of the original adjoint state equation and the state equation in which the control u is replaced by $\alpha^{-1} p$: Find $y \in V$ and $p \in Q$ such that

$$\begin{aligned} (y,z)_{L^2(\Omega)} + (z,p)_{H^1(\Omega)} &= (y_d,z)_{L^2(\Omega)} && \text{for all } z \in V, \\ (y,q)_{H^1(\Omega)} - \alpha^{-1}(p,q)_{L^2(\Omega)} &= 0 && \text{for all } q \in Q, \end{aligned}$$

or, equivalently, as a variational problem in the product space $X = V \times Q$: Find $(y,p) \in X$ such that

$$\mathcal{B}((y,p),(z,q)) = (y_d,z)_{L^2(\Omega)} \quad \text{for all } (z,q) \in X \tag{2.3}$$

with

$$\mathcal{B}((y,p),(z,q)) = (y,z)_{L^2(\Omega)} + (z,p)_{H^1(\Omega)} + (y,q)_{H^1(\Omega)} - \alpha^{-1} (p,q)_{L^2(\Omega)}.$$

A simple consequence of the well-posedness of the original formulation in $\hat{X}$ is

THEOREM 2.1. *There are constants $c_1, c_2 > 0$ independent of α such that*

$$c_1 \|(y,p)\|_X \le \sup_{0 \ne (z,q) \in X} \frac{\mathcal{B}((y,p),(z,q))}{\|(z,q)\|_X} \le c_2 \|(y,p)\|_X \quad \textit{for all } (y,p) \in X,$$

where the norm in X is given by

$$\|(z,q)\|_X^2 = \|z\|_V^2 + \|q\|_Q^2.$$

Proof. For $u = \alpha^{-1} p$, one easily sees that

$$\hat{\mathcal{B}}((y,u,p),(z,v,q)) = \mathcal{B}((y,p),(z,q)),$$

which implies

$$\sup_{0 \ne (z,v,q) \in \hat{X}} \frac{\hat{\mathcal{B}}((y,u,p),(z,v,q))}{\|(z,v,q)\|_{\hat{X}}} = \sup_{0 \ne (z,q) \in X} \frac{\mathcal{B}((y,p),(z,q))}{\|(z,q)\|_X}.$$

Since, for $u = \alpha^{-1} p$,

$$\|(y,p)\|_X^2 \le \|(y,u,p)\|_{\hat{X}}^2 = \|y\|_V^2 + \alpha^{-1} \|p\|_{L^2(\Omega)}^2 + \|p\|_Q^2 \le 2 \|(y,p)\|_X^2,$$

the estimates for $\mathcal{B}$ follow immediately from the corresponding estimates (2.2) for $\hat{\mathcal{B}}$ with $c_1 = \hat{c}_1$ and $c_2 = \sqrt{2}\,\hat{c}_2$. ∎

So, by the theorem of Babuška-Aziz, the variational problem (2.3) is also well-posed uniformly in α.

For now on and throughout the paper, we will use the symbol c for a generic constant, possibly different at each occasion it appears but independent of α and, later on, also independent of the mesh size h of an underlying mesh.

Regularity. For the forthcoming analysis we need the following assumption (full elliptic regularity of the state/adjoint state equation):

Assumption (R). For $f \in L^2(\Omega)$, let $y_f \in V$ solve the variational problem

$$(y_f, z)_{H^1(\Omega)} = (f, z)_{L^2(\Omega)} \quad \text{for all } z \in V.$$

Then $y_f \in H^2(\Omega)$ and there is a constant c independent of f such that

$$\|y_f\|_{H^2(\Omega)} \le c\, \|f\|_{L^2(\Omega)}.$$

This condition is, for example, satisfied if Ω is sufficiently smooth, see, e.g., [19], or if Ω is convex and polygonal/polyhedral, see, e.g, [12], [11].

So, for a right-hand side of the state equation in $(L^2(\Omega))^* \subset V^*$, the solution lies in the subspace $H^2(\Omega) \subset V$ equipped with a stronger norm and the solution can be estimated in this stronger norm by the norm of right-hand side in $(L^2(\Omega))^*$. Here, H^* denotes the dual of a Hilbert space H.

From the full elliptic regularity of the state/adjoint state equation a similar property can be derived for the reduced optimality system. For this we first introduce the subspace $X_+ \subset X$:

$$X_+ = V_+ \times Q_+ \quad \text{with } V_+ = Q_+ = H^2(\Omega),$$

and equip this set with a stronger norm, given by

$$\|(z,q)\|_{X_+}^2 = \|z\|_{V_+}^2 + \|q\|_{Q_+}^2,$$

where

$$\|z\|_{V_+}^2 = \|z\|_{H^1(\Omega)}^2 + \alpha^{1/2}\, \|z\|_{H^2(\Omega)}^2, \quad \|q\|_{Q_+}^2 = \alpha^{-1}\, \|q\|_{H^1(\Omega)}^2 + \alpha^{-1/2}\, \|q\|_{H^2(\Omega)}^2.$$

These norms are derived from the original norms in V and Q by raising the differentiation order by 1.

With these notations we have:

Theorem 2.2. *Suppose that Assumption (R) is satisfied. For $f \in L^2(\Omega)$, let (y_f, p_f) solve the variational problem*

$$\mathcal{B}((y_f, p_f), (z,q)) = (f, z)_{L^2(\Omega)} \quad \textit{for all } (z,q) \in X. \tag{2.4}$$

Then $(y_f, p_f) \in X_+$ and there is a constant c independent of α such that

$$\|(y_f, p_f)\|_{X_+} \le c\, \alpha^{-1/4}\, \|f\|_{L^2(\Omega)} \quad \textit{for all } f \in L^2(\Omega).$$

For $g \in L^2(\Omega)$, let (y_g, p_g) solve the variational problem

$$\mathcal{B}((y_g, p_g), (z, q)) = (g, q)_{L^2(\Omega)} \quad \textit{for all } (z, q) \in X.$$

Then $(y_g, p_g) \in X_+$ and there is a constant c independent of α such that

$$\|(y_g, p_g)\|_{X_+} \leq c\, \alpha^{1/4}\, \|g\|_{L^2(\Omega)} \quad \textit{for all } g \in L^2(\Omega).$$

Proof. First we consider the variational problem (2.4), which can be written in the form:

$$\begin{aligned} (p_f, z)_{H^1(\Omega)} &= (f - y_f, z)_{L^2(\Omega)} && \text{for all } z \in V, \\ (y_f, q)_{H^1(\Omega)} &= \alpha^{-1}\, (p_f, q)_{L^2(\Omega)} && \text{for all } q \in Q. \end{aligned}$$

From Assumption (R) it follows that $y_f, p_f \in H^2(\Omega)$ and that

$$\|p_f\|_{H^2(\Omega)} \leq c\, \left(\|f\|_{L^2(\Omega)} + \|y_f\|_{L^2(\Omega)}\right) \quad \text{and} \quad \|y_f\|_{H^2(\Omega)} \leq c\, \alpha^{-1}\, \|p_f\|_{L^2(\Omega)}.$$

Now

$$\|y_f\|_{L^2(\Omega)} \leq \|y_f\|_V \leq \|(y_f, p_f)\|_X, \quad \|p_f\|_{L^2(\Omega)} \leq \alpha^{1/2}\, \|p_f\|_Q \leq \alpha^{1/2}\, \|(y_f, p_f)\|_X.$$

From Theorem 2.1, we obtain

$$\|(y_f, p_f)\|_X \leq c\, \|\mathcal{F}_f\|_{X^*} = c \sup_{\in V} \frac{(f, z)_{L^2(\Omega)}}{\|z\|_V} \leq c \sup_{0 \neq z \in V} \frac{(f, z)_{L^2(\Omega)}}{\|z\|_{L^2(\Omega)}} = c\, \|f\|_{L^2(\Omega)}.$$

Therefore, it follows that

$$\|p_f\|_{H^2(\Omega)} \leq c\, \|f\|_{L^2(\Omega)} \quad \text{and} \quad \|y_f\|_{H^2(\Omega)} \leq c\, \alpha^{-1/2}\, \|f\|_{L^2(\Omega)}.$$

Together with

$$\|y_f\|_{H^1(\Omega)} \leq \alpha^{-1/4}\, \|y_f\|_V \leq \alpha^{-1/4}\, \|(y_f, p_f)\|_X \leq c\, \alpha^{-1/4}\, \|f\|_{L^2(\Omega)}$$

and

$$\|p_f\|_{H^1(\Omega)} \leq \alpha^{1/4}\, \|p_f\|_V \leq \alpha^{1/4}\, \|(y_f, p_f)\|_X \leq c\, \alpha^{1/4}\, \|f\|_{L^2(\Omega)}$$

we obtain

$$\|y_f\|_{V_+}^2 = \|y_f\|_{H^1(\Omega)}^2 + \alpha^{1/2}\, \|y_f\|_{H^2(\Omega)}^2 \leq c\, \alpha^{-1/2}\, \|f\|_{L^2(\Omega)}^2$$

and

$$\|p_f\|_{Q_+}^2 = \alpha^{-1}\, \|p_f\|_{H^1(\Omega)}^2 + \alpha^{-1/2}\, \|p_f\|_{H^2(\Omega)}^2 \leq c\, \alpha^{-1/2}\, \|f\|_{L^2(\Omega)}^2,$$

from which the first estimate immediately follows.

The corresponding estimates for (y_g, p_g) follow completely analogously. ∎

The Hilbert spaces V, Q, V_+, and Q_+ with the corresponding norms introduced above can also be written in the form

$$V = L^2(\Omega) \cap \left[\alpha^{1/4}\, H^1(\Omega)\right], \quad Q = \alpha^{-1/2}\, V$$

and

$$V_+ = H^1(\Omega) \cap \left[\alpha^{1/4} H^2(\Omega)\right], \quad Q_+ = \alpha^{-1/2}\, V_+,$$

by using the following notations:

NOTATION 1. *Let a be a positive real number and H, H_1, H_2 be Hilbert spaces. Then aH denotes the Hilbert space of all functions from H equipped with the norm*

$$\|h\|_{aH} = a\|h\|_H,$$

and $H_1 \cap H_2$ denotes the Hilbert space of all functions from the intersection of H_1 and H_2 with the norm

$$\|h\|_{H_1 \cap H_2} = \left(\|h\|_{H_1}^2 + \|h\|_{H_2}^2\right)^{1/2}.$$

The discrete problem and an error estimate. For discretizing the infinite dimensional problem in $X = V \times Q$ we replace V and Q by finite dimensional subspaces V_h and Q_h. In particular, as an example, we choose the space of piecewise linear and continuous functions on a simplicial subdivision $\mathcal{T}_h$ of the domain Ω. As usual, let h_T denote the diameter of an element $T \in \mathcal{T}_h$ and let $h = \max_{T \in \mathcal{T}_h} h_T$ denote the mesh size. Throughout the section we will assume that $\mathcal{T}_h$ is shape-regular, i.e., there is a constant σ such that

$$\frac{h_T}{\rho_T} \le \sigma \quad \text{for all } T \in \mathcal{T}_h,$$

where ρ_T denotes the diameter of the largest ball inside $\overline{T}$.

The approximate solution $(y_h, p_h) \in X_h = V_h \times Q_h$ is given by the variational problem:

$$\mathcal{B}((y_h, p_h), (z, q)) = (y_d, z)_{L^2(\Omega)} \quad \text{for all } (z, q) \in X_h. \tag{2.5}$$

A crucial element of the forthcoming multigrid convergence analysis is a discretization error estimate in an L^2-like norm for general right-hand sides in $(L^2(\Omega) \times L^2(\Omega))^*$: For this we introduce the space $X_- \supset X$:

$$X_- = V_- \times Q_- \quad \text{with } V_- = Q_- = L^2(\Omega)$$

and equip this set a weaker norm, given by

$$\|(z, q)\|_{X_-,h}^2 = \|z\|_{V_-,h}^2 + \|q\|_{Q_-,h}^2,$$

where

$$\|z\|_{V_-,h}^2 = \|z\|_{L^2(\Omega)}^2 + \alpha^{1/2}\, h^{-2}\, \|z\|_{L^2(\Omega)}^2 = (1 + \alpha^{1/2}\, h^{-2})\, \|z\|_{L^2(\Omega)}^2,$$
$$\|q\|_{Q_-,h}^2 = \alpha^{-1}\, \|q\|_{L^2(\Omega)}^2 + \alpha^{-1/2}\, h^{-2}\, \|q\|_{L^2(\Omega)}^2 = \alpha^{-1}(1 + \alpha^{1/2}\, h^{-2})\, \|q\|_{L^2(\Omega)}^2.$$

These norms are derived from the original norms in V and Q by replacing all appearing H^1-norms by L^2-norms scaled with a factor h^{-1}, where the parameter h is the mesh size of the underlying mesh. In this sense, these norms are mesh-dependent, explicitly expressed by the extra subindex h.

With these notations we have:

THEOREM 2.3. *Suppose that Assumption (R) is satisfied. For $\mathcal{F} \in (X_-)^*$, let $(y,p) \in X$ solve the variational problem*

$$\mathcal{B}((y,p),(z,q)) = \mathcal{F}(z,q) \quad \textit{for all } (z,q) \in X,$$

and let $(y_h, p_h) \in X_h$ be the approximate solution, given by

$$\mathcal{B}((y_h,p_h),(z,q)) = \mathcal{F}(z,q) \quad \textit{for all } (z,q) \in X_h.$$

Then there is a constant c independent of α and h such that, for all $\mathcal{F} \in (X_-)^*$,

$$\|(y-y_h, p-p_h)\|_{X_-,h} \le c\,\|\mathcal{F}\|_{(X_-)^*,h}. \quad \textit{with} \quad \|\mathcal{F}\|_{(X_-)^*,h} = \sup_{0\neq(z,q)\in X_-} \frac{\mathcal{F}(z,q)}{\|(z,q)\|_{X_-,h}}.$$

Proof. The proof of this estimate follows the classical line of arguments: Because of Theorem 2.1 we first estimate the discretization error in the X-norm by the approximation error (Cea's lemma):

$$\|(y-y_h, p-p_h)\|_X \le c \inf_{(z_h,q_h)\in X_h} \|(y,p)-(z_h,q_h)\|_X.$$

The approximation error can be estimated by the norm of (y,p) in $X + [hX_+]$, see Lemma 5.1 in the appendix, which can further be estimated by the norm of $\mathcal{F}$ in $(X_-)^*$, see Lemma 5.2 in the appendix. This results in the following estimate of the discretization error in the X-norm:

$$\|(y-y_h, p-p_h)\|_X \le c\,\|\mathcal{F}\|_{(X_-)^*,h} \tag{2.6}$$

for some constant c independent of α and h.

For the estimate in the weaker norm of X_- we use the classical Aubin-Nitsche duality trick, as outlined for saddle point problems in [10]: Consider the dual problem

$$\mathcal{B}((z,q),(w,s)) = \mathcal{F}^*(z,q) \quad \text{for all } (z,q) \in X$$

for some right hand side $\mathcal{F}^* \in (X_-)^*$. Assumption (R) ensures that $(w,s) \in X_+$.Then, using the Galerkin orthogonality, we have

$$\mathcal{F}^*(y-y_h, p-p_h) = \mathcal{B}((y-y_h,p-p_h),(w-w_h,s-s_h)) \quad \text{for all } (w_h,s_h) \in X_h.$$

Therefore, by using Theorem 2.1, it follows that

$$\mathcal{F}^*(y-y_h,p-p_h) \le c\,\|(y-y_h,p-p_h)\|_X \inf_{(w_h,s_h)\in X_h} \|(w-w_h,s-s_h)\|_X.$$

Using Lemma 5.1 and Lemma 5.2, see the appendix, for the second factor we obtain:

$$\mathcal{F}^*(y-y_h,p-p_h) \le c\,\|(y-y_h,p-p_h)\|_X\,\|\mathcal{F}^*\|_{(X_-)^*,h} \quad \text{for all } \mathcal{F}^* \in (X_-)^*,$$

which implies

$$\|(y-y_h,p-p_h)\|_{X_-,h} = \sup_{0\neq\mathcal{F}^*\in(X_-)^*} \frac{\mathcal{F}^*(y-y_h,p-p_h)}{\|\mathcal{F}^*\|_{(X_-)^*,h}} \le c\,\|(y-y_h,p-p_h)\|_X,$$

which, by using (2.6), completes the proof. □

3. A multigrid method and the approximation property. Next we will discuss a multigrid method for solving (2.5). We assume that there is a hierarchy of increasingly finer subdivisions $\mathcal{T}_k$ of Ω with mesh sizes h_k for each level $k = 0, 1, \ldots$, where $\mathcal{T}_{k+1}$ is obtained from $\mathcal{T}_k$ by some uniform refinement technique leading to $h_{k+1} = h_k/2$ and nested finite element spaces of continuous and piecewise linear functions:

$$V_k \times Q_k = X_k \subset X_{k+1} = V_{k+1} \times Q_{k+1}.$$

The approximate solution $(y_k, p_k) \in X_k$ at level k is given by

$$\mathcal{B}((y_k, p_k), (z, q)) = \mathcal{F}_k(z, q) \quad \text{for all } (z, q) \in X_k.$$

for the right-hand side $\mathcal{F}_k(z, q) = (y_d, z)_{L^2(\Omega)}$.

Let $(y_k^{(0)}, p_k^{(0)})$ be an approximation to (y_k, p_k). Then one iteration step of the considered multigrid method proceeds in two stages:

1. Smoothing step: Perform m steps of a (simple) iterative procedure

$$(y_k^{(0,j)}, p_k^{(0,j)}) = \mathcal{S}_k\left((y_k^{(0,j-1)}, p_k^{(0,j-1)}), \mathcal{F}_k\right) \quad \text{for } j = 1, \ldots, m,$$

starting from $(y_k^{(0)}, p_k^{(0)})$.

2. Coarse grid correction: Set

$$\widetilde{\mathcal{F}}_{k-1}(z, q) = \mathcal{F}(z, q) - \mathcal{B}\left((y_k^{(0,m)}, p_k^{(0,m)}), (z, q)\right) \quad \text{for all } (z, q) \in X_{k-1}.$$

Let $(\widetilde{t}_{k-1}, \widetilde{r}_{k-1}) \in V_{k-1} \times Q_{k-1}$ be the solution of

$$\mathcal{B}((t_{k-1}, r_{k-1}), (z, q)) = \widetilde{\mathcal{F}}_{k-1}(z, q) \quad \text{for all } (z, q) \in X_{k-1}. \tag{3.1}$$

If $k - 1 = 0$, compute the exact solution of (3.1) and set $(t_0, r_0) = (\widetilde{t}_0, \widetilde{r}_0)$.
If $k-1 > 0$, compute (t_{k-1}, r_{k-1}) by applying γ steps of the multigrid method applied to (3.1) with zero starting values.
Set the next iterate

$$(y_k^{(1)}, p_k^{(1)}) = (y_k^{(0,m)}, p_k^{(0,m)}) + (t_{k-1}, r_{k-1}).$$

The specification and properties of the iterative procedure in the smoothing step will be discussed in the next section.

In this section we concentrate on the coarse grid correction and consider first the case of a two-grid method ($k = 1$). We closely follow the line of arguments in the pioneering paper [9] on multigrid methods for parameter dependent problems. However, the dependence on the parameter is different from the situation discussed in [9] and, therefore, requires a different treatment. This is, in particular, reflected by the special norms appearing in the regularity results, see Lemma 5.2 in the appendix.

An important part of the convergence analysis is the so-called approximation property, which relates the error before and after the coarse grid correction. We have:

Theorem 3.1 (Approximation property). *Suppose that Assumption (R) is satisfied. Then, for the two-grid method, there is a constant c independent of α and the mesh size h_k such that:*

$$\left\|(y_k^{(1)}, p_k^{(1)}) - (y_k, p_k)\right\|_{0,k} \le c \left\|(y_k^{(0,m)}, p_k^{(0,m)}) - (y_k, p_k)\right\|_{2,k}$$

with the norms

$$\|(z,q)\|_{0,k} = \|(z,q)\|_{X_-,h_k} \quad \textit{and} \quad \|(w,s)\|_{2,k} = \sup_{0\neq(z,q)\in X_k} \frac{\mathcal{B}((w,s),(z,q))}{\|(z,q)\|_{0,k}}.$$

Proof. One easily sees that

$$(y_k - y_k^{(1)}, p_k - p_k^{(1)}) = (t_k, r_k) - (t_{k-1}, r_{k-1})$$

with $(t_k, r_k) = (y_k - y_k^{(0,m)}, p_k - p_k^{(0,m)}) \in X_k$ and $(t_{k-1}, r_{k-1}) \in X_{k-1}$, given by (3.1).

Observe that

$$\begin{aligned}
\mathcal{B}((t_{k-1},r_{k-1}),(z,q)) &= \mathcal{F}(z,q) - \mathcal{B}\left((y_k^{(0,m)}, p_k^{(0,m)}),(z,q)\right) \\
&= \mathcal{B}\left((y_k,p_k) - (y_k^{(0,m)}, p_k^{(0,m)}),(z,q)\right) \\
&= \mathcal{B}\left((t_k,r_k),(z,q)\right) \quad \text{for all } (z,q)\in X_{k-1}.
\end{aligned} \tag{3.2}$$

For a given functional $\mathcal{F}^* \in (X_-)^*$, let $(w,s)\in X$ satisfy

$$\mathcal{B}((z,q),(w,s)) = \mathcal{F}^*(z,q) \quad \text{for all } (z,q)\in X,$$

let $(w_k, s_k)\in X_k$ satisfy

$$\mathcal{B}((z,q),(w_k,s_k)) = \mathcal{F}^*(z,q) \quad \text{for all } (z,q)\in X_k,$$

and let $(w_{k-1}, s_{k-1})\in X_{k-1}$ satisfy

$$\mathcal{B}((z,q),(w_{k-1},s_{k-1})) = \mathcal{F}^*(z,q) \quad \text{for all } (z,q)\in X_{k-1}.$$

Then

$$\begin{aligned}
\mathcal{F}^*(t_k - t_{k-1}, r_k - r_{k-1}) &= \mathcal{B}\left((t_k,r_k) - (t_{k-1},r_{k-1}),(w_k,s_k)\right) \\
&= \mathcal{B}\left((t_k,r_k),(w_k,s_k) - (w_{k-1},s_{k-1})\right),
\end{aligned}$$

since

$$\begin{aligned}
\mathcal{B}((t_{k-1},r_{k-1}),(w_k,s_k)) &= \mathcal{F}^*(t_{k-1},r_{k-1}) = \mathcal{B}((t_{k-1},r_{k-1}),(w_{k-1},s_{k-1})) \\
&= \mathcal{B}((t_k,r_k),(w_{k-1},s_{k-1}))
\end{aligned}$$

by using (3.2). Hence

$$\mathcal{F}^*(t_k - t_{k-1}, r_k - r_{k-1}) \le \|(t_k,r_k)\|_{2,k}\, \|(w_k,s_k) - (w_{k-1},s_{k-1})\|_{0,k}.$$

By using the triangle inequality and Theorem 2.3 we obtain

$$\begin{aligned}
\|(w_k,s_k) - (w_{k-1},s_{k-1})\|_{0,k} &= \|(w_k,s_k) - (w_{k-1},s_{k-1})\|_{X_-,h_k} \\
&\le \|(w_k,s_k) - (w,s)\|_{X_-,h_k} + \|(w,s) - (w_{k-1},s_{k-1})\|_{X_-,h_k} \\
&\le \|(w_k,s_k) - (w,s)\|_{X_-,h_k} + \|(w,s) - (w_{k-1},s_{k-1})\|_{X_-,h_{k-1}} \\
&\le c\left[\|\mathcal{F}^*\|_{(X_-)^*,h_k} + \|\mathcal{F}^*\|_{(X_-)^*,h_{k-1}}\right] \le c\,\|\mathcal{F}^*\|_{(X_-)^*,h_k}.
\end{aligned}$$

This leads to

$$\mathcal{F}^*\left((t_k,r_k) - (t_{k-1},r_{k-1})\right) \le c\,\|(t_k,r_k)\|_{2,k}\,\|\mathcal{F}^*\|_{(X_-)^*,h_k}.$$

Therefore,

$$
\begin{aligned}
&\|(t_k, r_k) - (t_{k-1}, r_{k-1})\|_{0,k} = \|(t_k, r_k) - (t_{k-1}, r_{k-1})\|_{X_-,h_k} \\
&= \sup_{0 \neq \mathcal{F}^* \in (X_-)^*} \frac{\mathcal{F}^*\left((t_k, r_k) - (t_{k-1}, r_{k-1})\right)}{\|\mathcal{F}^*\|_{(X_-)^*,h_k}} \leq c\,\|(t_k, r_k)\|_{2,k},
\end{aligned}
$$

which completes the proof. □

4. The smoother and the smoothing property. Next we discuss the smoothing step in details, applied to a discrete problem of the form

$$
\mathcal{B}((y_k, p_k), (z, q)) = \mathcal{F}_k(z, q) \quad \text{for all } (z, q) \in X_k \tag{4.1}
$$

for a general right-hand side $\mathcal{F}_k \in (X_k)^*$. For this it will be more convenient to use matrix-vector notation: We choose the standard nodal basis in $V_k = Q_k$. Then, for any finite element function w from $V_k = Q_k$ the vector $\underline{w} \in \mathbb{R}^{n_k}$ denotes the vector of coefficients relative to the nodal basis. Furthermore, we introduce the matrix representation of the bilinear form $\mathcal{B}$ on $V_k \times Q_k$ by

$$
\mathcal{B}\left((w, s), (z, q)\right) = (M_k \underline{w}, \underline{z})_{\ell^2} + (K_k \underline{z}, \underline{s})_{\ell^2} + (K_k \underline{w}, \underline{q})_{\ell^2} - \alpha^{-1}\,(M_k \underline{s}, \underline{q})_{\ell^2}
$$

and the vector representation of the linear form

$$
\mathcal{F}_k(z, q) = (\underline{f}_k, \underline{z})_{\ell^2} + (\underline{g}_k, \underline{q})_{\ell^2},
$$

where $(.,.)_{\ell^2}$ denotes the Euclidean inner product. The matrices M_k and K_k are the mass matrix and the stiffness matrix representing the L^2-inner product and the H^1-inner product on $V_k = Q_k$, respectively.

In matrix-vector notation the discrete problem (4.1) can be written as

$$
\mathcal{A}_k \begin{bmatrix} \underline{y}_k \\ \underline{p}_k \end{bmatrix} = \begin{bmatrix} \underline{f}_k \\ \underline{g}_k \end{bmatrix} \quad \text{with} \quad \mathcal{A}_k = \begin{bmatrix} M_k & K_k \\ K_k & -\alpha^{-1}\,M_k \end{bmatrix}.
$$

As smoothing procedure we will use a preconditioned Richardson method:

$$
\begin{bmatrix} \underline{y}_k^{(0,j)} \\ \underline{p}_k^{(0,j)} \end{bmatrix} = \begin{bmatrix} \underline{y}_k^{(0,j-1)} \\ \underline{p}_k^{(0,j-1)} \end{bmatrix} - \hat{\mathcal{A}}_k^{-1} \left(\begin{bmatrix} \underline{f}_k \\ \underline{g}_k \end{bmatrix} - \mathcal{A}_k \begin{bmatrix} \underline{y}_k^{(0,j-1)} \\ \underline{p}_k^{(0,j-1)} \end{bmatrix} \right) \tag{4.2}
$$

with a simple preconditioner $\hat{\mathcal{A}}_k$.

For constructing $\hat{\mathcal{A}}_k$ we follow the approach of transforming smoothers, see [27]. By applying a left and a right transformation to the system matrix $\mathcal{A}_k$ we obtain the transformed system matrix

$$
\mathcal{A}_k' = \begin{bmatrix} 0 & I \\ -\alpha^{-1/2}\,I & 0 \end{bmatrix} \mathcal{A}_k \begin{bmatrix} I & 0 \\ 0 & \alpha^{1/2}\,I \end{bmatrix} = \begin{bmatrix} K_k & -\alpha^{-1/2}\,M_k \\ -\alpha^{-1/2}\,M_k & -K_k \end{bmatrix}. \tag{4.3}
$$

For this transformed matrix we use a symmetric Uzawa preconditioner:

$$
\hat{\mathcal{A}}_k' = \begin{bmatrix} \hat{K}_k & -\alpha^{-1/2}\,M_k \\ -\alpha^{-1/2}\,M_k & \alpha^{-1} M_k \hat{K}_k^{-1} M_k - \hat{S}_k \end{bmatrix} \tag{4.4}
$$

with positive definite matrices $\hat{K}_k$ and $\hat{S}_k$. By transforming back we finally obtain our preconditioner:

$$\hat{\mathcal{A}}_k = \begin{bmatrix} 0 & I \\ -\alpha^{-1/2}\, I & 0 \end{bmatrix}^{-1} \hat{\mathcal{A}}'_k \begin{bmatrix} I & 0 \\ 0 & \alpha^{1/2}\, I \end{bmatrix}^{-1} = \begin{bmatrix} M_k & \hat{S}_k - \alpha^{-1}\, M_k \hat{K}_k^{-1} M_k \\ \hat{K}_k & -\alpha^{-1}\, M_k \end{bmatrix}. \tag{4.5}$$

Performing one step of this iterative method requires first to calculate the residual:

$$\begin{bmatrix} \widetilde{\underline{f}}_k \\ \widetilde{\underline{g}}_k \end{bmatrix} = \begin{bmatrix} \underline{f}_k \\ \underline{g}_k \end{bmatrix} - \mathcal{A}_k \begin{bmatrix} \underline{y}_k^{(0,j-1)} \\ \underline{p}_k^{(0,j-1)} \end{bmatrix},$$

then to solve the linear system

$$\begin{bmatrix} M_k & \hat{S}_k - \alpha^{-1}\, M_k \hat{K}_k^{-1} M_k \\ \hat{K}_k & -\alpha^{-1}\, M_k \end{bmatrix} \begin{bmatrix} \underline{w}_k \\ \underline{s}_k \end{bmatrix} = \begin{bmatrix} \widetilde{\underline{f}}_k \\ \widetilde{\underline{g}}_k \end{bmatrix},$$

and, finally, compute the next iterate

$$\begin{bmatrix} \underline{y}_k^{(0,j)} \\ \underline{p}_k^{(0,j)} \end{bmatrix} = \begin{bmatrix} \underline{y}_k^{(0,j-1)} \\ \underline{p}_k^{(0,j-1)} \end{bmatrix} + \begin{bmatrix} \underline{w}_k \\ \underline{s}_k \end{bmatrix}.$$

The solution of the linear system for $\underline{w}_k$ and $\underline{s}_k$ from above is obtained by solving the following three linear systems successively:

$$\begin{aligned} \hat{K}_k \widetilde{\underline{w}}_k &= \widetilde{\underline{g}}_k, \\ \hat{S}_k\, \underline{s}_k &= \widetilde{\underline{f}}_k - M_k \widetilde{\underline{w}}_k, \\ \hat{K}_k \underline{w}_k &= \widetilde{\underline{g}}_k + \alpha^{-1}\, M_k \underline{s}_k. \end{aligned} \tag{4.6}$$

In order to obtain a smoothing step with low computational costs we choose diagonal matrices for $\hat{K}_k$ and $\hat{S}_k$. In particular, we consider the choice

$$\hat{K}_k = \frac{1}{\sigma_k} \operatorname{diag} K_k \tag{4.7}$$

and

$$\hat{S}_k = \frac{1}{\tau_k} \operatorname{diag} S_k \quad \text{with} \quad S_k = K_k + \alpha^{-1}\, M_k \hat{K}_k^{-1} M_k \tag{4.8}$$

with appropriate relaxation parameters $\sigma_k > 0$ and $\tau_k > 0$.

The total computational costs for one step of this smoothing iteration is dominated by the 6 matrix-vector multiplications for the computation of the residuals and the right-hand sides in (4.6).

We have

Theorem 4.1 (Smoothing property). *Assume that there are constants $\sigma > 0$ and $\tau > 0$ independent of the mesh size h_k and α such that the relaxation factors in* (4.7) *and* (4.8) *satisfy the conditions*

$$\sigma \le \sigma_k \le \frac{1}{\lambda_{max}\left((\operatorname{diag} K_k)^{-1} K_k\right)} \quad \textit{and} \quad \tau \le \tau_k \le \frac{1}{\lambda_{max}\left((\operatorname{diag} S_k)^{-1} S_k\right)}. \tag{4.9}$$

Then, for the preconditioned Richardson method (4.2) *with preconditioner* (4.5)*, there is a constant* c *independent of* h_k *and* α *such that*

$$\left\|(y_k^{(0,m)}, p_k^{(0,m)}) - (y_k, p_k)\right\|_{2,k} \le c\left(1 + \alpha^{-1/2}\, h_k^2\right)\, \eta_0(m)\, \left\|(y_k^{(0)}, p_k^{(0)}) - (y_k, p_k)\right\|_{0,k}$$

with

$$\eta_0(m) = \frac{1}{2^{m-1}}\binom{m-1}{[m/2]} \le \begin{cases} \sqrt{\frac{2}{\pi(m-1)}} & \text{for even } m, \\ \sqrt{\frac{2}{\pi m}} & \text{for odd } m. \end{cases}$$

Here $\binom{n}{\ell}$ *denotes the binomial coefficient and* $[x]$ *denotes the largest integer smaller than or equal to* x*.*

Proof. In the proof the following standard notations are used: $M \ge N$ iff $M - N$ is positive semi-definite, for symmetric matrices M and N. For a symmetric and positive definite matrix P, the norms $\|v\|_P$ and $\|M\|_P$ of a vector v and a matrix M are given by

$$\|v\|_P = \sqrt{(Pv, v)_{\ell^2}} \quad \text{and} \quad \|M\|_P = \sup_{v\neq 0}\frac{\|Mv\|_P}{\|v\|_P}.$$

With these notations we can rewrite the mesh-dependent norms introduced in the previous section using matrix-vector notation:

$$\begin{aligned}\|(z,q)\|_{0,k}^2 = \|(z,q)\|_{X_-,h_k}^2 &= (1 + \alpha^{1/2} h_k^{-2})\left(\|z\|_{L^2(\Omega)}^2 + \alpha^{-1}\,\|q\|_{L^2(\Omega)}^2\right)\\ &= (1 + \alpha^{1/2} h_k^{-2})\left((M_k \underline{z}, \underline{z})_{\ell^2} + \alpha^{-1}\,(M_k \underline{q}, \underline{q})_{\ell^2}\right)\\ &= \left(\mathcal{L}_k \begin{bmatrix}\underline{z}\\ \underline{q}\end{bmatrix}, \begin{bmatrix}\underline{z}\\ \underline{q}\end{bmatrix}\right)_{\ell^2} = \left\|\begin{bmatrix}\underline{z}\\ \underline{q}\end{bmatrix}\right\|_{\mathcal{L}_k}^2\end{aligned}$$

with

$$\mathcal{L}_k = (1 + \alpha^{1/2} h_k^{-2})\begin{bmatrix} M_k & 0\\ 0 & \alpha^{-1}\, M_k\end{bmatrix}$$

and

$$\begin{aligned}\|(w,s)\|_{2,k}^2 = \sup_{0\neq(z,q)\in X_k}\frac{\mathcal{B}((w,s),(z,q))^2}{\|(z,q)\|_{X_-,h_k}^2} &= \sup_{0\neq \underline{z},\underline{q}\in\mathbb{R}^{n_k}}\frac{\left(\mathcal{A}_k\begin{bmatrix}\underline{w}\\ \underline{s}\end{bmatrix}, \begin{bmatrix}\underline{z}\\ \underline{q}\end{bmatrix}\right)_{\ell^2}^2}{\left(\mathcal{L}_k\begin{bmatrix}\underline{z}\\ \underline{q}\end{bmatrix}, \begin{bmatrix}\underline{z}\\ \underline{q}\end{bmatrix}\right)_{\ell^2}}\\ = \left(\mathcal{L}_k^{-1}\mathcal{A}_k\begin{bmatrix}\underline{w}\\ \underline{s}\end{bmatrix}, \mathcal{A}_k\begin{bmatrix}\underline{w}\\ \underline{s}\end{bmatrix}\right)_{\ell^2} &= \left\|\mathcal{A}_k\begin{bmatrix}\underline{w}\\ \underline{s}\end{bmatrix}\right\|_{\mathcal{L}_k^{-1}}^2 = \left\|\mathcal{L}_k^{-1}\mathcal{A}_k\begin{bmatrix}\underline{w}\\ \underline{s}\end{bmatrix}\right\|_{\mathcal{L}_k}^2.\end{aligned}$$

Moreover, we have the following relation between the error $(y_k^{(0,m)}, p_k^{(0,m)}) - (y_k, p_k)$ after m steps and the initial error $(y_k^{(0)}, p_k^{(0)}) - (y_k, p_k)$ in matrix-vector notation:

$$\begin{bmatrix}\underline{y}_k^{(0,m)} - \underline{y}_k\\ \underline{p}_k^{(0,m)} - \underline{p}_k\end{bmatrix} = (\mathcal{M}_k)^m \begin{bmatrix}\underline{y}_k^{(0)} - \underline{y}_k\\ \underline{p}_k^{(0)} - \underline{p}_k\end{bmatrix}$$

with the error propagation matrix

$$\mathcal{M}_k = I - \hat{\mathcal{A}}_k^{-1} \mathcal{A}_k.$$

Then the property to be shown reads:

$$\left\| \mathcal{L}_k^{-1} \mathcal{A}_k \, (\mathcal{M}_k)^m \, \underline{e}_k \right\|_{\mathcal{L}_k} \leq c \left(1 + \alpha^{-1/2} \, h_k^2 \right) \eta_0(m) \, \|\underline{e}_k\|_{\mathcal{L}_k}$$

for all initial errors $\underline{e}_k = (\underline{y}_k^{(0)} - \underline{y}_k, \underline{p}_k^{(0)} - \underline{p}_k)^T$, i.e.:

$$\left\| \mathcal{L}_k^{-1} \mathcal{A}_k \, (\mathcal{M}_k)^m \right\|_{\mathcal{L}_k} \leq c \left(1 + \alpha^{-1/2} \, h_k^2 \right) \eta_0(m).$$

It is straight forward to show

$$\|\underline{e}_k\|_{\mathcal{L}_k} = (1 + \alpha^{1/2} \, h_k^{-2})^{1/2} \, \|\underline{e}_k'\|_{\mathcal{L}_k'}$$

and

$$\left\| \mathcal{L}_k^{-1} \mathcal{A}_k \, (\mathcal{M}_k)^m \, \underline{e}_k \right\|_{\mathcal{L}_k} = \frac{\alpha^{1/2}}{(1 + \alpha^{1/2} \, h_k^{-2})^{1/2}} \left\| (\mathcal{L}_k')^{-1} \, \mathcal{A}_k' \, (\mathcal{M}_k')^m \, \underline{e}_k' \right\|_{\mathcal{L}_k'}$$

with the transformed matrices $\mathcal{A}_k'$ and $\hat{\mathcal{A}}_k'$, see (4.3) and (4.4), and

$$\underline{e}_k' = \begin{bmatrix} I & 0 \\ 0 & \alpha^{-1/2} \, I \end{bmatrix} \underline{e}_k, \quad \mathcal{M}_k' = I - (\hat{\mathcal{A}}_k')^{-1} \mathcal{A}_k', \quad \text{and} \quad \mathcal{L}_k' = \begin{bmatrix} M_k & 0 \\ 0 & M_k \end{bmatrix}.$$

Therefore,

$$\left\| \mathcal{L}_k^{-1} \mathcal{A}_k \, (\mathcal{M}_k)^m \right\|_{\mathcal{L}_k} = \frac{\alpha^{1/2}}{1 + \alpha^{1/2} \, h_k^{-2}} \left\| (\mathcal{L}_k')^{-1} \, \mathcal{A}_k' \, (\mathcal{M}_k')^m \right\|_{\mathcal{L}_k'}.$$

The scaling parameters σ_k and τ_k are uniformly bounded from below and ensure that

$$\hat{K}_k \geq K_k \quad \text{and} \quad \hat{S}_k \geq S_k.$$

Then we know from [24] that:

$$\left\| (\mathcal{L}_k')^{-1} \, \mathcal{A}_k' \, (\mathcal{M}_k')^m \right\|_{\mathcal{L}_k'} \leq \eta_0(m) \left\| (\mathcal{L}_k')^{-1} \, \mathcal{Q}_k' \right\|_{\mathcal{L}_k'}$$

with

$$\eta_0(m) = \frac{1}{2^{m-1}} \binom{m-1}{[m/2]} \quad \text{and} \quad \mathcal{Q}_k' = \begin{bmatrix} \hat{K}_k - K_k & 0 \\ 0 & \hat{S}_k - S_k \end{bmatrix}.$$

One easily sees that

$$\left\| (\mathcal{L}_k')^{-1} \, \mathcal{Q}_k' \right\|_{\mathcal{L}_k'} = \max \left(\lambda_{\max} \left(M_k^{-1} (\hat{K}_k - K_k) \right), \lambda_{\max} \left(M_k^{-1} (\hat{S}_k - S_k) \right) \right).$$

By a simple scaling argument, one easily shows that

$$\lambda_{\max} \left(M_k^{-1} (\hat{K}_k - K_k) \right) \leq \lambda_{\max} \left(M_k^{-1} \hat{K}_k \right) \leq \frac{1}{\sigma} \, \lambda_{\max} \left(M_k^{-1} \operatorname{diag} K_k \right) \leq c \, h_k^{-2}$$

and

$$\begin{aligned}\lambda_{\max}\left(M_k^{-1}(\hat{S}_k - S_k)\right) \le \lambda_{\max}\left(M_k^{-1}\hat{S}_k\right) &\le \frac{1}{\tau}\,\lambda_{\max}\left(M_k^{-1}\operatorname{diag} S_k\right)\\ &\le c\,\left(h_k^{-2} + \alpha^{-1}\,h_k^2\right)\end{aligned}$$

for some constant c independent of k and α. Therefore, we obtain

$$\begin{aligned}\left\|\mathcal{L}_k^{-1}\mathcal{A}_k\,(\mathcal{M}_k)^m\right\|_{\mathcal{L}_k} &= \frac{\alpha^{1/2}}{1+\alpha^{1/2}\,h^{-2}}\,\left\|(\mathcal{L}_k')^{-1}\,\mathcal{A}_k'\,(\mathcal{M}_k')^m\right\|_{\mathcal{L}_k'}\\ &\le c\,\frac{\alpha^{1/2}\,\left(h_k^{-2}+\alpha^{-1}\,h_k^2\right)}{1+\alpha^{1/2}\,h_k^{-2}}\,\eta_0(m) \le c\,\left(1+\alpha^{-1/2}\,h_k^2\right)\,\eta_0(m),\end{aligned}$$

which completes the proof. $\square$

Remark 1. *For* $\sigma_k = 1/\operatorname{nnz}(K_k)$ *and* $\tau_k = 1/\operatorname{nnz}(S_k)$ *the condition* (4.9) *in Theorem 4.1 is certainly satisfied. Here* $\operatorname{nnz}(M)$ *denotes the maximum number of non-zero entries per row for the matrix* M. *In general, this parameter choice is too conservative. For the actual computation of these parameters in the numerical experiments we refer to the forthcoming section.*

By combining approximation property and smoothing property we immediately obtain the main result of this paper:

Theorem 4.2 (Two-grid convergence). *Under the assumptions of Theorem 3.1 and Theorem 4.1 there is a constant* $c > 0$ *independent of the mesh size* h_k *and* α *such that*

$$\|(y_k^{(1)}, p_k^{(1)}) - (y_k, p_k)\|_{0,k} \le c\,\frac{1+\alpha^{-1/2}\,h_k^2}{\sqrt{m}}\,\|(y_k^{(0)}, p_k^{(0)}) - (y_k, p_k)\|_{0,k}.$$

So, for fixed α and sufficiently many smoothing steps m the two-grid method converges with a convergence rate independent of h_k. If we assume that α is not too small compared to h_k, more precisely, if there is a constant c independent of h_k and α such that $\alpha \ge c\,h_k^4$, then the convergence rate is also independent of α.

Remark 2. *Similar to the estimates in the proof of Theorem 4.1 one can additionally show that*

$$\left\|(y_k^{(0,m)}, p_k^{(0,m)}) - (y_k, p_k)\right\|_{0,k} \le c\,\left(1+\alpha^{-1/2}\,h_k^2\right)\,\left\|(y_k^{(0)}, p_k^{(0)}) - (y_k, p_k)\right\|_{0,k}$$

for some constant c *independent of* h_k *and* α. *Then, by a standard perturbation argument, the analogous properties as above follow for the convergence of the multigrid method with* $\gamma \ge 2$, *in particular, for the choice* $\gamma = 2$, *the so-called W-cycle.*

Remark 3. *In [23] an extra relaxation parameter* $\omega \in (0, 2)$ *was introduced for the smoother:*

$$\begin{bmatrix} \underline{y}_k^{(0,j)} \\ \underline{p}_k^{(0,j)} \end{bmatrix} = \begin{bmatrix} \underline{y}_k^{(0,j-1)} \\ \underline{p}_k^{(0,j-1)} \end{bmatrix} - \omega\,\hat{\mathcal{A}}_k^{-1}\left(\begin{bmatrix} \underline{f}_k \\ \underline{g}_k \end{bmatrix} - \mathcal{A}_k \begin{bmatrix} \underline{y}_k^{(0,j-1)} \\ \underline{p}_k^{(0,j-1)} \end{bmatrix}\right) \tag{4.10}$$

with $\hat{\mathcal{A}}_k$ *given by* (4.5). *For this smoother the same convergence properties can be shown as in the original case (*$\omega = 1$*), based on the results from [13]. Motivated by a local mode analysis the choice* $\omega = 1.6$ *was proposed, see [23] for details.*

REMARK 4. *If we keep* (4.7) *but choose*

$$\hat{S}_k = \hat{K}_k + \alpha^{-1} M_k \hat{K}^{-1} M_k$$

instead of (4.8) *for the smoothing iteration, it easily follows by analogous arguments that the convergence rate of the multigrid method is robust with respect to* k *and* α *without any restriction. This smoothing iteration can be seen as a preconditioned Richardson method* (4.2) *with preconditioner*

$$\hat{\mathcal{A}}_k = \begin{bmatrix} M_k & \hat{K}_k \\ \hat{K}_k & -\alpha^{-1} M_k \end{bmatrix}.$$

It additionally requires in each step to solve a linear system with the well-conditioned, symmetric, and positive definite matrix $\hat{S}_k$*, which can be done by applying a few steps of a conjugate gradient method.*

5. Numerical experiments. Next we present some numerical experiments for the domain $\Omega = (0,1)^2$. The initial mesh at grid level $k = 0$ consists of 2 triangles by connecting the points $(0,0)$ and $(1,1)$. Without loss of generality we choose homogeneous data $y_d = 0$, the exact solution is then given by $y_k = 0$ and $p_k = 0$.

In all tests we used (4.10) for the smoothing step with $\omega = 1.6$ as suggested in [23]. The relaxation parameters in (4.7) and (4.8) were chosen in the following way: For grid levels $k \leq 6$ we used $\sigma_k = \lambda_k^{-1}$ and $\tau_k = \mu_k^{-1}$, where λ_k and μ_k denote the approximations to the largest eigenvalues of the eigenvalue problems $K_k v_k = \lambda\,(\operatorname{diag} K_k) v_k$ and $S_k v_k = \mu\,(\operatorname{diag} S_k) v_k$, respectively, obtained by the power method. For all grid level $k > 6$ the same parameters as for $k = 6$ were used.

The starting values $y_k^{(0)}$ and $p_k^{(0)}$ were chosen randomly. The accuracy of an approximation was measured by the $\mathcal{L}_k^{-1}$-norm of its residual, which coincides with the $\|.\|_{2,k}$-norm of the error. In the actual computation of the $\mathcal{L}_k^{-1}$-norm the mass matrix appearing in $\mathcal{L}_k$ was replaced by $h_k^2\, I$. The multigrid iteration was performed until this modified norm of the residual was reduced by a factor $\varepsilon = 10^{-10}$.

Table 5.1 shows, for each grid level k and the parameter $\alpha = 1$, the total number N of unknowns and the number of iterations depending on the number m of smoothing steps, written in the format $m = m_{\text{pre}} + m_{\text{post}}$ for m_{pre} and m_{post} pre- and post-smoothing steps, respectively.

TABLE 5.1
Number of W-cycle iterations for $\alpha = 1$

k	N	α	smoothing steps			
			1+1	2+2	4+4	8+8
4	578	1	22	12	8	6
6	8 450	1	21	13	9	7
8	132 098	1	21	13	9	7
10	2 101 250	1	21	13	9	7

Table 5.2 shows the same information as Table 5.1, however for α decreasing by a factor of 1/16 from one level to the next. So α varies from 1 to roughly $6 \cdot 10^{-8}$. Observe that h is decreasing by a factor 1/2 at the same time. So this corresponds to a relation $\alpha = \mathcal{O}(h_k^4)$.

Table 5.2
Number of W-cycle iterations for $\alpha = c\,h_k^4$

k	N	α	smoothing steps			
			1+1	2+2	4+4	8+8
4	578	1	22	12	8	6
6	8 450	2^{-8}	21	13	9	7
8	132 098	2^{-16}	21	13	9	7
10	2 101 250	2^{-24}	22	13	9	7

The most efficient performance in terms of computing time were observed for 1 pre- and 1 post-smoothing steps. On a desktop pc with an Intel Pentium D cpu (3.20 GHz) the computing time for $k = 10$ was roughly 3 seconds per multigrid iteration step.

As expected from Theorem 4.2, Table 5.1 shows that, for fixed parameter α, the convergence rates are independent of the level k and the rates improve with an increasing number of smoothing steps. Table 5.2 confirms the results contained in Theorem 4.2 about the robustness with respect to α, if $\alpha \geq c\,h_k^4$.

Remark 5. *The presented multigrid convergence analysis covers only the W-cycle. Numerical experiments showed that the V-cycle (i.e., the choice $\gamma = 1$ in the multigrid algorithm, see Section 3) also works with about the same number of iterates. The best performance was again observed for 1 pre- and 1 post-smoothing steps with a computing time of roughly 2 seconds per multigrid iteration step for $k = 10$.*

Appendix. For $\mathcal{F} \in (X_-)^*$, let $(y, p) \in X$ be the solution of the variational problem

$$\mathcal{B}((y,p),(z,q)) = \mathcal{F}(z,q) \quad \text{for all } (z,q) \in X. \tag{5.1}$$

We first introduce the concept of algebraic sum of Hilbert spaces:

Notation 2. *Let H, H_1, H_2 be Hilbert spaces with $H_1 \subset H$ and $H_2 \subset H$. Then we introduce the algebraic sum $H_1 + H_2$ as the Hilbert space of elements from the set $\{h_1 + h_2 : h_1 \in H_1,\ h_2 \in H_2\}$ equipped with the norm*

$$\|h\|_{H_1+H_2} = \inf_{\substack{h=h_1+h_2\\ h_1\in H_1,\ h_2\in H_2}} \left(\|h_1\|_{H_1}^2 + \|h_2\|_{H_2}^2\right)^{1/2}.$$

With this notation we have the following estimate for the approximation error:

Lemma 5.1. *Assume that $(y, p) \in X_+$ and that the subdivision $\mathcal{T}_h$ is shape-regular. Then there is a constant c independent of α and h such that*

$$\inf_{(z_h,q_h)\in X_h} \|(y,p) - (z_h,q_h)\|_X \leq c\,\|(y,p)\|_{X+[hX_+]}.$$

Proof. Let $y = y_1 + y_2$ and $p = p_1 + p_2$ with $y_1, p_1 \in H^1(\Omega)$ and $y_2, p_2 \in H^2(\Omega)$. Then

$$\inf_{(z_h,q_h)\in X_h} \|(y,p) - (z_h,q_h)\|_X \leq \|(y_1,p_1)\|_X + \inf_{(z_h,q_h)\in X_h} \|(y_2,p_2) - (z_h,q_h)\|_X$$
$$\leq \sqrt{2}\left(\|(y_1,p_1)\|_X^2 + \inf_{(z_h,q_h)\in X_h} \|(y_2,p_2) - (z_h,q_h)\|_X^2\right)^{1/2}.$$

By standard interpolation arguments, it easily follows for the second term:

$$\begin{aligned}
&\inf_{(z_h,q_h)\in X_h} \|(y_2,p_2)-(z_h,q_h)\|_X^2 \\
&= \inf_{(z_h,q_h)\in X_h} \Big(\|y_2-z_h\|_{L^2(\Omega)}^2 + \alpha^{1/2}\,\|y_2-z_h\|_{H^1(\Omega)}^2 \\
&\qquad\qquad + \alpha^{-1}\,\|p_2-q_h\|_{L^2(\Omega)}^2 + \alpha^{-1/2}\,\|p_2-q_h\|_{H^1(\Omega)}^2\Big) \\
&\le c\,h^2\,\Big(\|y_2\|_{H^1(\Omega)}^2 + \alpha^{1/2}\,\|y_2\|_{H^2(\Omega)}^2 + \alpha^{-1}\,\|p_2\|_{H^1(\Omega)}^2 + \alpha^{-1/2}\,\|p_2\|_{H^2(\Omega)}^2\Big) \\
&= c\,h^2\|(y_2,p_2)\|_{X_+}^2
\end{aligned}$$

for some constant c independent of α and h. By taking the infimum over all $y_1, p_1 \in H^1(\Omega)$, $y_2, p_2 \in H^2(\Omega)$ the estimate immediately follows. □

We continue by estimating the $X + [hX_+]$-norm of the exact solution (y,p) in terms of the right-hand side $\mathcal{F}$.

LEMMA 5.2. *Suppose that Assumption (R) is satisfied. Then the solution (y,p) to the variational problem* (5.1) *lies in X_+ and there is a constant c independent of α and h such that*

$$\|(y,p)\|_{X+[h\,X_+]} \le c\,\|\mathcal{F}\|_{(X_-)^*,h}.$$

Proof. Since the norm in X_- is a scaled L^2-norm, the linear functional $\mathcal{F} \in (X_-)^*$ has the form

$$\mathcal{F}(z,q) = (f,z)_{L^2(\Omega)} + (g,q)_{L^2(\Omega)} \quad \text{for some } f,g \in L^2(\Omega).$$

Therefore, $(y,p) = (y_f,p_f) + (y_g,p_g)$ with

$$\mathcal{B}((y_f,p_f),(z,q)) = (f,z)_{L^2(\Omega)} \quad \text{and} \quad \mathcal{B}((y_g,p_g),(z,q)) = (g,q)_{L^2(\Omega)}$$

for all $(z,q) \in X$.

As in the proof of Theorem 2.2 we obtain

$$\|(y_f,p_f)\|_X \le c\,\|f\|_{L^2(\Omega)} \quad \text{and} \quad \|(y_f,p_f)\|_{X_+} \le c\,\alpha^{-1/4}\,\|f\|_{L^2(\Omega)}.$$

Therefore,

$$\begin{aligned}
\|(y_f,p_f)\|_{X+[h\,X_+]} &\le \min\left(\|(y_f,p_f)\|_X, \|(y_f,p_f)\|_{h\,X_+}\right) \\
&\le c\,\min\Big(1,\alpha^{-1/4}h\Big)\,\|f\|_{L^2(\Omega)} \le \sqrt{2}\,c\,(1+\alpha^{1/2}h^{-2})^{-1/2}\,\|f\|_{L^2(\Omega)},
\end{aligned}$$

using the simple inequality

$$\min(a,b) \le \left(\frac{a^{-2}+b^{-2}}{2}\right)^{-1/2}$$

for positive real numbers a, b. With

$$\|f\|_{L^2(\Omega)} = \sup_{0\ne z\in L^2(\Omega)} \frac{(f,z)_{L^2(\Omega)}}{\|z\|_{L^2(\Omega)}} = (1+\alpha^{1/2}h^{-2})^{1/2} \sup_{0\ne z\in V_-} \frac{(f,z)_{L^2(\Omega)}}{\|z\|_{V_-,h}}$$

we obtain

$$\|(y_f, p_f)\|_{X+[h\,X_+]} \leq \sqrt{2}c \sup_{0 \neq z \in V_-} \frac{(f, z)_{L^2(\Omega)}}{\|z\|_{V_-,h}} = \sqrt{2}c\,\|\mathcal{F}_f\|_{(X_-)^*,h}$$

with $\mathcal{F}_f(z, q) = (f, z)_{L^2(\Omega)}$. Analogously, it follows that

$$\|(y_g, p_g)\|_{X+[h\,X_+]} \leq \sqrt{2}c\,\|\mathcal{F}_g\|_{(X_-)^*,h}$$

with $\mathcal{F}_g(z, q) = (g, q)_{L^2(\Omega)}$. Therefore, by the triangle inequality

$$\begin{aligned}\|(y, p)\|_{X+[h\,X_+]} &\leq \sqrt{2}c\left(\|\mathcal{F}_f\|_{(X_-)^*,h} + \|\mathcal{F}_g\|_{(X_-)^*,h}\right) \\ &\leq 2c\left(\|\mathcal{F}_f\|^2_{(X_-)^*,h} + \|\mathcal{F}_g\|^2_{(X_-)^*,h}\right)^{1/2}.\end{aligned}$$

Since

$$\|\mathcal{F}\|^2_{(X_-)^*,h} = \|\mathcal{F}_f\|^2_{(X_-)^*,h} + \|\mathcal{F}_g\|^2_{(X_-)^*,h},$$

the proof is completed. ◻

REFERENCES

[1] E. Arian and S. Ta'asan, *Multigrid one-shot methods for optimal control problems: Infinite dimensional control*, ICASE-Report 94-52, NASA Langley Research Center, Hampton VA, 1994.

[2] I. Babuška, *Error-bounds for finite element method*, Numer. Math., 16 (1971), pp. 322–333.

[3] A. Battermann and M. Heinkenschloss, *Preconditioners for Karush-Kuhn-Tucker matrices arising in the optimal control of distributed systems*, in Desch, W. (ed.) et al., Control and Estimation of Distributed Parameter Systems. Int. Ser. Numer. Math. 126, Basel: Birkhäuser, 1998, pp. 15–32.

[4] A. Battermann and E. W. Sachs, *Block preconditioners for KKT systems in PDE-governed optimal control problems*, in Hoffmann, Karl-Heinz (ed.) et al., Fast solution of discretized optimization problems. Int. Ser. Numer. Math. 138, Basel: Birkhäuser, 2001, pp. 1–18.

[5] G. Biros and O. Ghattas, *Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver*, SIAM J. Sci. Comput., 27 (2005), pp. 687–713.

[6] ———, *Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part II: The Lagrange-Newton solver and its application to optimal control of steady viscous flows*, SIAM J. Sci. Comput., 27 (2005), pp. 714–739.

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

[8] A. Borzì and V. Schulz, *Multigrid methods for PDE optimization*, SIAM Rev., 51 (2009), pp. 361–395.

[9] S. C. Brenner, *Multigrid methods for parameter dependent problems*, RAIRO, Modelisation Math. Anal. Numer., 30 (1996), pp. 265 – 297.

[10] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, New York etc.: Springer-Verlag, 1991.

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

[12] ———, *Neumann and mixed problems on curvilinear polyhedra*, Integral Equations Oper. Theory, 15 (1992), pp. 227–261.

[13] A. Ecker and W. Zulehner, *On the smoothing property of multi-grid methods in the non-symmetric case*, Numer. Linear Algebra Appl., 3 (1996), pp. 161–172.

[14] W. Hackbusch, *Fast solution of elliptic control problems*, J. Optimization Theory Appl., 31 (1980), pp. 565–581.

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

[16] S. B. Hazra and V. Schulz, *Simultaneous pseudo-timestepping for PDE-model based optimization problems*, BIT, 44 (2004), pp. 457–472.
[17] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE constraints*, Mathematical Modelling: Theory and Applications 23. Dordrecht: Springer, 2009.
[18] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Berlin-Heidelberg-New York: Springer-Verlag, 1971.
[19] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Paris: Masson, 1967.
[20] T. Rees, H. S. Dollar, and A. J. Wathen, *Optimal solvers for PDE-constrained optimization*, Tech. Report NA-08/10, Oxford University Computing Laboratory, June 2008. (To appear in SIAM J. Sci. Comput.).
[21] J. Schöberl and W. Zulehner, *Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 752 – 773.
[22] V. Schulz and G. Wittum, *Transforming smoothers for PDE constrained optimization problems*, Computing and Visualization in Science, 11 (2008), pp. 207 – 219.
[23] R. Simon, *Multigrid solvers for saddle point problems in PDE-constrained optimization*, PhD thesis, Johannes Kepler University Linz, Austria, 2008.
[24] R. Simon and W. Zulehner, *On Schwarz-type smoothers for saddle point problems with applications to PDE-constrained optimization problems*, Numer. Math., 111 (2009), pp. 445 – 468.
[25] S. Ta'asan, *"One-shot" methods for optimal control of distributed parameter systems I: The finite dimensional case*, ICASE-Report 91-2, NASA Langley Research Center, Hampton VA, 1991.
[26] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen. Theorie, Verfahren und Anwendungen*, Wiesbaden: Vieweg, 2005.
[27] G. Wittum, *Multi-grid methods for Stokes and Navier-Stokes equations. Transforming smoothers: Algorithms and numerical results*, Numer. Math., 54 (1989), pp. 543 – 563.

Latest Reports in this series

2009

[..]

2009-02	Clemens Hofreither, Ulrich Langer and Satyendra Tomar *Boundary Elements Simulation of Linear Water Waves in a Model Basin*	February 2009
2009-03	Huidong Yang and Walter Zulehner *Numerical Simulation of Fluid-Structure Interaction Problems on Hybrid Meshes with AMG*	April 2009
2009-04	Clemens Pechstein and Robert Scheichl *Analysis of FETI Methods for Multiscale PDEs - Part II: Interface Variation*	April 2009
2009-05	Alexsandr M. Matsokin and Sergey V. Nepomnyaschikh *Domain Decomposition Preconditioning for Well Models for Reservoir Problems*	June 2009
2009-06	Helmut Gfrerer *Generalized Penalty Methods for a Class of Convex Optimization Problems with Pointwise Inequality Constraints*	July 2009
2009-07	Ulrich Langer, Huidong Yang and Walter Zulehner *A Grid-enabled Solver for the Fluid-Structure Interaction (FSI) Problem*	August 2009
2009-08	Stefan Takacs and Walter Zulehner *Multigrid Methods for Elliptic Optimal Control Problems with Neumann Boundary Control*	October 2009
2009-09	Dylan M. Copeland and Ulrich Langer *Domain Decomposition Solvers for Nonlinear Multiharmonic Finite Element Equations*	November 2009
2009-10	Huidong Yang and Walter Zulehner *A Newton Based Fluid-structure Interaction (FSI) Solver with Algebraic Multigrid Methods (AMG) on Hybrid Meshes*	November 2009
2009-11	Peter Gruber, Dorothee Knees, Sergiy Nesenenko and Marita Thomas *Analytical and Numerical Aspects of Time-dependent Models with Internal Variables*	November 2009
2009-12	Clemens Pechstein and Robert Scheichl *Weighted Poincaré Inequalities and Applications in Domain Decomposition*	November 2009
2009-13	Dylan Copeland, Michael Kolmbauer and Ulrich Langer *Domain Decomposition Solvers for Frequency-Domain Finite Element Equations*	December 2009
2009-14	Clemens Pechstein *Shape-Explicit Constants for Some Boundary Integral Operators*	December 2009
2009-15	Peter G. Gruber, Johanna Kienesberger, Ulrich Langer, Joachim Schöberl and Jan Valdman *Fast Solvers and A Posteriori Error Estimates in Elastoplasticity*	December 2009

2010

2010-01	Joachim Schöberl, René Simon and Walter Zulehner *A Robust Multigrid Method for Elliptic Optimal Control Problems*	Januray 2010

From 1998 to 2008 reports were published by SFB013. Please see
`http://www.sfb013.uni-linz.ac.at/index.php?id=reports`
From 2004 on reports were also published by RICAM. Please see
`http://www.ricam.oeaw.ac.at/publications/list/`
For a complete list of NuMa reports see
`http://www.numa.uni-linz.ac.at/Publications/List/`