Technische Universität Graz

Robust discretization and solvers for elliptic optimal control problems with energy regularization

U. Langer, O. Steinbach, H. Yang

**Berichte aus dem
Institut für Angewandte Mathematik**

Technische Universität Graz

Robust discretization and solvers for elliptic optimal
control problems with energy regularization

U. Langer, O. Steinbach, H. Yang

Berichte aus dem
Institut für Angewandte Mathematik

Robust discretization and solvers for elliptic optimal control problems with energy regularization

Ulrich Langer[*],   Olaf Steinbach[†],   Huidong Yang[‡]

Abstract

We consider the finite element discretization and the iterative solution of singularly perturbed elliptic reaction-diffusion equations in three-dimensional computational domains. These equations arise from the optimality conditions for elliptic distributed optimal control problems with energy regularization that were recently studied by M. Neumüller and O. Steinbach (2020). We provide quasi-optimal a priori finite element error estimates which depend both on the mesh size h and on the regularization parameter ϱ. The choice $\varrho = h^2$ ensures optimal convergence which only depends on the regularity of the target function. For the iterative solution, we employ an algebraic multigrid preconditioner and a balancing domain decomposition by constraints (BDDC) preconditioner. We numerically study robustness and efficiency of the proposed algebraic preconditioners with respect to the mesh size h, the regularization parameter ϱ, and the number of subdomains (cores) p. Furthermore, we investigate the parallel performance of the BDDC preconditioned conjugate gradient solver.

Keywords: Elliptic optimal control problems, energy regularization, finite element discretization, a priori error estimates, fast solvers

2010 MSC: 49K20, 49M41, 35J25, 65N12, 65N30, 65N55

1   Introduction

In the recent work [22], M. Neumüller and O. Steinbach have investigated regularization error estimates for the solution of distributed optimal control problems in energy spaces on the basis of the following model problem: Minimize the cost functional

$$\mathcal{J}(u_\varrho, z_\varrho) = \frac{1}{2} \int_\Omega [u_\varrho(x) - \overline{u}(x)]^2 \, dx + \frac{\varrho}{2} \|z_\varrho\|^2_{H^{-1}(\Omega)} \tag{1}$$

with respect to the state $u_\varrho \in H_0^1(\Omega)$ and the control $z_\varrho \in H^{-1}(\Omega)$ subject to the constraints

$$-\Delta u_\varrho = z_\varrho \quad \text{in } \Omega, \quad u_\varrho = 0 \quad \text{on } \Gamma, \tag{2}$$

where $\overline{u} \in L^2(\Omega)$ represents the given desired state (target), and $\varrho \in \mathbb{R}_+$ denotes the regularization parameter. Here, the computational domain $\Omega \subset \mathbb{R}^3$ is assumed

---
[*]Institute for Computational Mathematics, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria, Email: ulanger@numa.uni-linz..ac.at

[†]Institut für Angewandte Mathematik, Technische Universität Graz, Steyrergaße 30, 8010 Graz, Austria, Email: o.steinbach@tugraz.at

[‡]Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenberger Straße 69, 4040 Linz, Austria, Email: huidong.yang@ricam.oeaw.ac.at

to be a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$. Note that

$$\|z_\varrho\|_{H^{-1}(\Omega)} = \|\nabla u_\varrho\|_{L^2(\Omega)}.$$

The associated adjoint state equation reads: Find the adjoint state $p_\varrho \in H_0^1(\Omega)$ such that

$$-\Delta p_\varrho = u_\varrho - \overline{u} \quad \text{in } \Omega, \quad p_\varrho = 0 \quad \text{on } \Gamma. \tag{3}$$

Further, the associated gradient equation is given by the relation

$$p_\varrho + \varrho\, u_\varrho = 0 \quad \text{in } \Omega. \tag{4}$$

Using the gradient equation (4) for eliminating the adjoint state p_ϱ from (3), we finally obtain the singularly perturbed reaction-diffusion equation

$$-\varrho\Delta u_\varrho + u_\varrho = \overline{u} \quad \text{in } \Omega, \quad u_\varrho = 0 \quad \text{on } \Gamma, \tag{5}$$

for defining the state u_ϱ. Once the state u_ϱ is computed from (5), the adjoint state p_ϱ is given by (4), and the control $z_\varrho = -\Delta u_\varrho = \varrho^{-1}(\overline{u} - u_\varrho)$ by (2) and (5).

The variational formulation of the Dirichlet boundary value problem (5) is to find $u_\varrho \in H_0^1(\Omega)$ such that

$$\varrho \int_\Omega \nabla u_\varrho(x) \cdot \nabla v(x)\, dx + \int_\Omega u_\varrho(x)\, v(x)\, dx = \int_\Omega \overline{u}(x)\, v(x)\, dx \tag{6}$$

is satisfied for all $v \in H_0^1(\Omega)$. When assuming some regularity of the given target $\overline{u} \in H_0^s(\Omega) := [L^2(\Omega), H_0^1(\Omega)]_s$ for $s \in [0,1]$ or $\overline{u} \in H_0^1(\Omega) \cap H^s(\Omega)$ for $s \in (1,2]$, the following error estimate with respect to the regularization parameter ϱ has been shown in [22]:

$$\|u_\varrho - \overline{u}\|_{L^2(\Omega)} \leq c\, \varrho^{s/2}\, \|\overline{u}\|_{H^s(\Omega)}. \tag{7}$$

Moreover, in the case $\overline{u} \in H_0^1(\Omega) \cap H^s(\Omega)$ for $s \in (1,2]$, there also holds

$$\|\nabla(u_\varrho - \overline{u})\|_{L^2(\Omega)} \leq c\, \varrho^{(s-1)/2}\, \|\overline{u}\|_{H^s(\Omega)}. \tag{8}$$

For $\varrho \to 0$, the boundary value problem (5) belongs to a class of singularly perturbed problems, see, e.g., [7]. For robust numerical methods to treat such problems, we refer to the monograph [8].

In this paper, we investigate the errors $\|u_{\varrho h} - \overline{u}\|_{L^2(\Omega)}$ and $\|\nabla(u_{\varrho h} - \overline{u})\|_{L^2(\Omega)}$ of the finite element approximation $u_{\varrho h}$ to the target $\overline{u}$ in terms of both the regularization parameter ϱ and the discretization parameter h. The error estimate is based on the estimates (7) and (8), and well-known finite element discretization error estimates. It turns out that the choice $\varrho = h^2$ gives the optimal convergence rate $O(h^2)$ for continuous, piecewise linear finite elements. Moreover, we consider iterative methods for the solution of the linear system of algebraic equations that arise from the Galerkin finite element discretization of the variational problem (6), which are robust with respect to the mesh size h and the regularization parameter ϱ. Such solvers have been addressed in many works. For example, geometric multigrid methods, which are robust with respect to the mesh size h and the parameter ϱ in the case of standard Galerkin discretization methods, were analyzed in [25], based on the convergence rate $\min\{1, h^2/\varrho\}$ in the L^2 norm as provided in [28]. The parameter independent contraction number of the multigrid method is achieved by a combination of such a deteriorate approximation property with an improved smoothing property. Robust and optimal algebraic multilevel (AMLI) methods for reaction-diffusion type problems were studied in [14]. Therein, a uniformly convergent AMLI method with optimal complexity was shown using the constant in the strengthened Cauchy-Bunyakowski-Schwarz inequality computed for both mass

2

and stiffness matrices. Robust block-structured preconditioning on both piecewise uniform meshes and graded meshes with boundary layers has been considered in [19, 23] in comparison with standard robust multigrid preconditioners. The new block diagonal preconditioner was based on the partitioning of degrees of freedom into those on the corners, edges, and interior points, respectively. The perturbation parameter independent condition number of the preconditioned linear system using diagonal and incomplete Cholesky preconditioning methods for singularly perturbed problems on layer-adapted meshes has been recently shown in [24]. Considering the adaptive finite element discretization on simplicial meshes with local refinements, we employ an algebraic multigrid (AMG) preconditioner [4, 5, 27] and the balancing domain decomposition by constraints (BDDC) preconditioner [6, 20, 21] for solving the discrete system. We make performance studies of the preconditioned conjugate gradient (PCG) method that uses these AMG and BDDC preconditioners. In particular, the BDDC preconditioned conjugate gradient solver shows excellent strong scalability.

The reminder of the paper is organized as follows: Section 2 deals with the finite element discretization of the equation (5). In Section 3, we describe both the AMG and BDDC preconditioners that are used in solving the system of finite element equations. Numerical results are presented and discussed in Section 4. Finally, some conclusions are drawn in Section 5.

2   Discretization

To perform the finite element discretization of the variational form (6), we introduce conforming finite element spaces $V_h \subset H_0^1(\Omega)$. Particularly, we use the standard finite element space $V_h = S_h^1(\Omega_h) \cap H_0^1(\Omega)$ spanned by continuous and piecewise linear basis functions. These functions are defined with respect to some admissible decomposition $\mathcal{T}_h(\Omega)$ of the domain Ω into shape regular simplicial finite elements τ_ℓ, and are zero on Γ. Here, $\overline{\Omega}_h = \bigcup_\ell \overline{\tau}_\ell$, and h denotes a suitable mesh-size parameter, see, e.g., [3, 29]. Then the finite element approximation of (6) is to find $u_{\varrho h} \in V_h$ such that

$$\varrho \int_\Omega \nabla u_{\varrho h}(x) \cdot \nabla v_h(x)\, dx + \int_\Omega u_{\varrho h}(x)\, v_h(x)\, dx = \int_\Omega \overline{u}(x)\, v_h(x)\, dx \qquad (9)$$

for all test functions $v_h \in V_h$.

In [22], the convergence of u_ϱ towards the target $\overline{u}$, and the error estimates (7) and (8) were proved. In the numerical experiments, presented in [22], u_ϱ was computed on a very fine grid to minimize the influence of the discretization. Now we are going to investigate the effects of the finite element discretization, and the final aim is to provide estimates of the errors $\|u_{\varrho h} - \overline{u}\|_{L^2(\Omega)}$ and $\|\nabla(u_{\varrho h} - \overline{u})\|_{L^2(\Omega)}$ in terms of both ϱ and h for sufficiently smooth target functions $\overline{u}$.

Theorem 1. *Let us assume that $\Omega \subset \mathbb{R}^3$ is convex, and that the target function satisfies $\overline{u} \in H_0^1(\Omega) \cap H^2(\Omega)$. Then there are positive constants C_0, C_1, C_2 and C_3 such that the error estimates*

$$\|u_{\varrho h} - \overline{u}\|_{L^2(\Omega)}^2 \leq \left(C_0 h^4 + C_1 \varrho h^2 + C_2 \varrho^2 \right) \|\overline{u}\|_{H^2(\Omega)}^2 \qquad (10)$$

and

$$\|\nabla(u_{\varrho h} - \overline{u})\|_{L^2(\Omega)}^2 \leq \left(C_0 h^4 \varrho^{-1} + C_1 h^2 + C_3 \varrho \right) \|\overline{u}\|_{H^2(\Omega)}^2 \qquad (11)$$

hold. For the choice $\varrho = h^2$, we therefore have

$$\|u_{\varrho h} - \overline{u}\|_{L^2(\Omega)} \leq c\, h^2\, \|\overline{u}\|_{H^2(\Omega)} \quad and \quad \|\nabla(u_{\varrho h} - \overline{u})\|_{L^2(\Omega)} \leq \tilde{c}\, h\, \|\overline{u}\|_{H^2(\Omega)}.$$

3

Proof. When subtracting the Galerkin variational formulation (9) from (6) for test functions $v_h \in V_h$ this gives the Galerkin orthogonality

$$\varrho \int_\Omega \nabla(u_\varrho - u_{\varrho h}) \cdot \nabla v_h \, dx + \int_\Omega (u_\varrho - u_{\varrho h}) \, v_h \, dx = 0 \quad \forall v_h \in V_h.$$

Therefore,

$$\varrho \|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}^2 + \|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)}^2$$
$$= \varrho \int_\Omega \nabla(u_\varrho - u_{\varrho h}) \cdot \nabla(u_\varrho - u_{\varrho h}) \, dx + \int_\Omega (u_\varrho - u_{\varrho h})(u_\varrho - u_{\varrho h}) \, dx$$
$$= \varrho \int_\Omega \nabla(u_\varrho - u_{\varrho h}) \cdot \nabla(u_\varrho - v_h) \, dx + \int_\Omega (u_\varrho - u_{\varrho h})(u_\varrho - v_h) \, dx$$
$$\leq \varrho \|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)} \|\nabla(u_\varrho - v_h)\|_{L^2(\Omega)} + \|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)} \|u_\varrho - v_h\|_{L^2(\Omega)}$$
$$\leq \sqrt{\varrho \|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}^2 + \|u_\varrho - u_{\varrho,h}\|_{L^2(\Omega)}^2}$$
$$\cdot \sqrt{\varrho \|\nabla(u_\varrho - v_h)\|_{L^2(\Omega)}^2 + \|u_\varrho - v_h\|_{L^2(\Omega)}^2}$$

follows, i.e., we have Cea's estimate

$$\varrho \|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}^2 + \|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)}^2 \leq \varrho \|\nabla(u_\varrho - v_h)\|_{L^2(\Omega)}^2 + \|u_\varrho - v_h\|_{L^2(\Omega)}^2$$

for all $v_h \in V_h$. Inserting the Lagrangian interpolation $v_h = I_h u_\varrho$ of u_ϱ, and using standard interpolation error estimates, we arrive at the error estimate

$$\varrho \|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}^2 + \|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)}^2 \quad \leq \quad c_1 \, \varrho \, h^2 \, |u_\varrho|_{H^2(\Omega)}^2 + c_2 \, h^4 \, |u_\varrho|_{H^2(\Omega)}^2$$
$$= \quad \left[c_1 \, \varrho + c_2 \, h^2\right] h^2 \, |u_\varrho|_{H^2(\Omega)}^2, \qquad (12)$$

where the positive constants c_1 and c_2 are nothing but the constants in the H^1 and L^2 interpolation error estimates; see, e.g., [3, 29]. Note that, under the assumptions made, we have $u_\varrho \in H_0^1(\Omega) \cap H^2(\Omega)$. Due to

$$-\varrho \Delta u_\varrho = \overline{u} - u_\varrho,$$

we now conclude

$$|u_\varrho|_{H^2(\Omega)}^2 \leq c_3 \, \frac{1}{\varrho^2} \, \|\overline{u} - u_\varrho\|_{L^2(\Omega)}^2 \qquad (13)$$

with the positive H^2 coercivity constant c_3. Since we assume $\overline{u} \in H_0^1(\Omega) \cap H^2(\Omega)$, we can use (7) for $s = 2$, i.e.,

$$\|\overline{u} - u_\varrho\|_{L^2(\Omega)} \leq c_4 \, \varrho \, \|\overline{u}\|_{H^2(\Omega)}. \qquad (14)$$

For less regular $\overline{u}$, we have a reduced order in ϱ; see Theorem 3.2 in [22] as well as (7). Combining (13) and (14), we get

$$|u_\varrho|_{H^2(\Omega)}^2 \leq c_3 c_4^2 \, \|\overline{u}\|_{H^2(\Omega)}^2. \qquad (15)$$

Inserting (15) into (12) this gives

$$\varrho \|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}^2 + \|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)}^2 \leq c_3 c_4^2 \left[c_1 \, \varrho + c_2 \, h^2\right] h^2 \, \|\overline{u}\|_{H^2(\Omega)}^2.$$

Using the triangle inequality, and again (14), we arrive at

$$\|u_{\varrho h} - \overline{u}\|_{L^2(\Omega)}^2 \quad \leq \quad \left[\|u_{\varrho h} - u_\varrho\|_{L^2(\Omega)} + \|u_\varrho - \overline{u}\|_{L^2(\Omega)}\right]^2$$
$$\leq \quad 2 \left[\|u_{\varrho h} - u_\varrho\|_{L^2(\Omega)}^2 + \|u_\varrho - \overline{u}\|_{L^2(\Omega)}^2\right]$$
$$\leq \quad 2 \, c_3 c_4^2 \left[c_1 \, \varrho + c_2 \, h^2\right] h^2 \, \|\overline{u}\|_{H^2(\Omega)}^2 + 2 c_4^2 \, \varrho^2 \, \|\overline{u}\|_{H^2(\Omega)}^2,$$

4

that is nothing but (12) with $C_0 = 2c_2c_3c_4^2$, $C_1 = 2c_1c_3c_4^2$, and $C_2 = 2c_4^2$.

Moreover, we also have

$$\|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}^2 \leq c_3 c_4^2 \left[c_1 \varrho + c_2 h^2\right] h^2 \varrho^{-1} \|\overline{u}\|_{H^2(\Omega)}^2.$$

Now, using (8) for $s = 2$, we finally obtain

$$
\begin{aligned}
\|\nabla(u_{\varrho h} - \overline{u})\|_{L^2(\Omega)}^2 &\leq \left[\|\nabla(u_{\varrho h} - u_\varrho)\|_{L^2(\Omega)} + \|\nabla(u_\varrho - \overline{u})\|_{L^2(\Omega)}\right]^2 \\
&\leq 2 \left[\|\nabla(u_{\varrho h} - u_\varrho)\|_{L^2(\Omega)}^2 + \|\nabla(u_\varrho - \overline{u})\|_{L^2(\Omega)}^2\right] \\
&\leq 2\,c_3 c_4^2 \left[c_1 \varrho + c_2 h^2\right] h^2 \varrho^{-1} \|\overline{u}\|_{H^2(\Omega)}^2 + 2c_5^2 \varrho \|\overline{u}\|_{H^2(\Omega)}^2,
\end{aligned}
$$

which is (11) with $C_3 = 2c_5^2$. □

The finite element space $V_h = \mathrm{span}\{\varphi_1, \ldots, \varphi_{N_h}\}$ is spanned by the standard nodal Courant basis functions $\{\varphi_1, \ldots, \varphi_{N_h}\}$. Once the basis is chosen, the finite element scheme (9) is equivalent to the system of finite element equations

$$A\underline{u} = \underline{f}, \tag{16}$$

where the system matrix $A = \varrho K + M$ consists of the scaled stiffness matrix K and the mass matrix M, respectively. The matrix entries and the entries of the right-hand side $\underline{f}$ are defined by

$$K_{ij} = \int_\Omega \nabla\varphi_j \cdot \nabla\varphi_i \, dx, \quad M_{ij} = \int_\Omega \varphi_j \, \varphi_i \, dx, \quad f_i = \int_\Omega \overline{u} \, \varphi_i \, dx$$

for $i, j = 1, \ldots, N_h$. The solution $\underline{u} = (u_j) \in \mathbb{R}^{N_h}$ of (16) contains the unknown coefficients of the finite element solution

$$u_{\varrho h} = \sum_{j=1}^{N_h} u_j \varphi_j \in V_h$$

of (9). Robust and fast iterative solvers for (16) will be considered in the next section.

3 AMG and BDDC Preconditioners

In this section, we present two precondioners that lead to robust and efficient solvers for the system (16) of finite element equations when applying a preconditioned conjugate gradient algorithm. The BDDC preconditioner is especially suited for parallel computers.

3.1 The AMG preconditioner

In constrast to geometric multigrid methods, e.g, [10], usually relying on underlying hierarchical meshes, algebraic multigrid methods [4, 5, 27, 34] are more flexible with respect to complex geometries, adaptive mesh refinements, and so on. We refer to [9] for a comparison between these two methods. In this work, the particular AMG preconditioner developed in [13] in combination with the conjugate gradient (CG) method is adopted to solve the discrete system. This AMG method was originally developed for second-order elliptic problems. It has been extended to act as a robust preconditioner to systems arising from the discretization of coupled vector field problems, e.g., to fluid and elasticity problems in fluid-structure interaction

solvers [17, 35]. More recently, in [31], by choosing a proper smoother [26], it was also used as a preconditioner for the non-symmetric and positive definite system that arises from the Petrov-Galerkin space-time finite element discretization of the heat equation [30].

In this AMG method, the coarsening strategy is based on a simple red-black colouring algorithm [13, Algorithm 2]: Pick up an unmarked degree of freedom (Dof), and mark it black (coarse); all neighboring Dofs are marked in red (fine); repeat this procedure until all Dofs are visited. Afterwards, the fine Dofs (red) are interpolated by an average over their neighboring coarse Dofs (black), which defines a linear interpolation operator P. The coarse grid operator is constructed by Galerkin's method, i.e., $A_c = P^\top A P$. Assume that we apply m pre- and post-smoothing steps, the iteration operator for the two grid method is given by

$$M_{amg} := \mathcal{S}^m (I - P A_c^{-1} P^\top A) \mathcal{S}^m,$$

where $\mathcal{S}$ is the iteration matrix for a smoothing step. For such a symmetric and positive definite system, one sweep of the symmetric Gauss–Seidel method is used, that is, one forward on the down cycle together with one backward on the up cycle. Then, we may formulate the two-grid preconditioner P_{AMG} as follows (with $m = 1$):

$$P_{AMG}^{-1} := \mathcal{S} P A_c^{-1} P^\top \mathcal{S}. \tag{17}$$

Note that, in the coarse grid correction step, one may apply AMG recursively; see [4, 5, 27], i.e., we replace A_c^{-1} by AMG iterations starting with a zero initial guess. In connection with optimal control problems as considered in this paper, we are especially interested in small $\varrho \in (0, 1]$. When $\varrho \to 0$, the mass matrix M becomes dominant in A. For the mass matrix, the condition number is $\mathcal{O}(1)$, which even makes the system easier to solve as observed from our numerical tests.

3.2 The BDDC preconditioner

For the (two-level) BDDC preconditioner, we first make a reordering of the system (16), i.e.,

$$\tilde{A}\underline{\tilde{u}} = \tilde{f}, \tag{18}$$

where

$$\tilde{A} := \begin{bmatrix} \tilde{A}_{II} & \tilde{A}_{IC} \\ \tilde{A}_{CI} & \tilde{A}_{CC} \end{bmatrix}, \ \underline{\tilde{u}} := \begin{bmatrix} \underline{\tilde{u}}_I \\ \underline{\tilde{u}}_C \end{bmatrix}, \ \tilde{f} := \begin{bmatrix} \tilde{f}_I \\ \tilde{f}_C \end{bmatrix}, \ \text{and} \ \tilde{A}_{II} = \text{diag}\left[\tilde{A}_{II}^1, ..., \tilde{A}_{II}^p \right].$$

Here, p represents the number of polyhedral subdomains Ω_i, as well as the number of cores. These subdomains Ω_i are obtained from the graph partitioning tool [12] and a non-overlapping domain decomposition of Ω [32]. Following common notations, the degrees of freedom in (18) are respectively decomposed into internal (I) and interface (C) Dofs. Then, we arrive at the following Schur complement system which is defined on the interface Γ_C:

$$S_C \underline{\tilde{u}}_C = \underline{g}, \tag{19}$$

where

$$S_C := \tilde{A}_{CC} - \tilde{A}_{CI} \tilde{A}_{II}^{-1} \tilde{A}_{IC}, \quad \underline{g} := \underline{\tilde{f}}_C - \tilde{A}_{CI} \tilde{A}_{II}^{-1} \underline{\tilde{f}}_I.$$

Now, following similar notations as in [6, 11, 20, 21], the BDDC preconditioner P_{BDDC} for (19) is formed as

$$P_{BDDC}^{-1} = R_C^\top (T_{sub} + T_0) R_C,$$

where the operator R_C represents the direct sum of restriction operators R_C^i that map the global interface vector to its components on a local interface $\Gamma_i := \partial \Omega_i \cap \Gamma_C$ with a proper scaling. Furthermore, the coarse level correction operator T_0 is defined by

$$T_0 = \Phi(\Phi^\top S_C \Phi)^{-1} \Phi^\top, \tag{20}$$

where $\Phi = \left[(\Phi^1)^\top, \ldots, (\Phi^N)^\top \right]^\top$ is the matrix of the coarse level basis functions. Each basis function matrix Φ^i on the subdomain interface Γ_i is computed by solving the augmented system:

$$\begin{bmatrix} S_C^i & (C^i)^\top \\ C^i & 0 \end{bmatrix} \begin{bmatrix} \Phi^i \\ \Lambda^i \end{bmatrix} = \begin{bmatrix} 0 \\ R_\Pi^i \end{bmatrix}, \tag{21}$$

where S_C^i is nothing but the local Schur complement, C^i denotes the given primal constraints of the subdomain Ω_i, and each column of Λ^i contains the vector of Lagrange multipliers. The number of columns of each Φ^i is equal to the number of global coarse level degrees of freedom, usually living on the subdomain corners, and/or interface edges, and/or faces. Moreover, the restriction operator R_Π^i maps the global interface vector in the continuous primal variable space on the coarse level to its component on Γ_i. We mention that these corners/edges/faces on the coarse space may be characterized as objects from a pure algebraic manner; see, e.g., [1]. We have recently used this abstract method in [18] for constructing BDDC preconditioners in solving the linear system that arises from the space-time finite element discretization of parabolic equations [15].

The subdomain correction operator T_{sub} is then defined as

$$T_{sub} = \sum_{i=1}^N \begin{bmatrix} (R_C^i)^\top & 0 \end{bmatrix} \begin{bmatrix} S_C^i & (C^i)^\top \\ C^i & 0 \end{bmatrix}^{-1} \begin{bmatrix} R_C^i \\ 0 \end{bmatrix},$$

with vanishing primal variables on all the coarse levels. Here, the restriction operator R_C^i maps global interface vectors to their components on Γ_i. Note that this two-level method can be extended to a multi-level method when the coarse problem becomes too large; see, e.g., [2, 36]. This means that we may apply the BDDC method to the inversion of $\Phi^\top S_C \Phi$ in the coarse problem (20) recursively.

4 Numerical experiments

In all numerical examples presented in this section, we consider the unit cube $\Omega = (0,1)^3$ as computational domain. In the first example, we choose a smooth target, whereas a discontinuous target is used in the second example. The first example is covered by Theorem 1, the second not. Finally, we consider an example with a smooth target that violates the homogeneous Dirichlet boundary conditions. The system (16) of finite element equations is solved by a preconditioned conjugate gradient (PCG) method with both AMG and BDDC preconditioners. The PCG iteration is stopped when the relative residual error of the preconditioned system reaches $\varepsilon = 10^{-8}$. In the AMG method, we have applied one V-cycle multigrid preconditioner with one pre- and post-smoothing step, respectively. We run the AMG preconditioned CG on the shared memory supercomputer MACH-2[1]. The BDDC preconditioned CG is performed on the high-performance distributed memory computing cluster RADON-1[2].

---

[1]https://www3.risc.jku.at/projects/mach2/
[2]https://www.ricam.oeaw.ac.at/hpc/

4.1 Smooth target

In the first example, we consider the smooth function

$$\bar{u}(x_1, x_2, x_3) = -\varrho \Delta u_\varrho + u_\varrho = \sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)$$

as target, which was computed from the manufactured solution

$$u_\varrho(x_1, x_2, x_3) = (3\varrho\pi^2 + 1)^{-1}\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)$$

of the reduced optimality equation (5). Then the adjoint state

$$p_\varrho(x_1, x_2, x_3) = -\frac{u_\varrho(x_1, x_2, x_3)}{\varrho(3\varrho\pi^2 + 1)} = -\frac{1}{\varrho(3\varrho\pi^2 + 1)}\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)$$

results from the gradient equation, and the optimal control

$$z_\varrho(x_1, x_2, x_3) = -\Delta u_\varrho(x_1, x_2, x_3) = \frac{3\pi^2}{3\varrho\pi^2 + 1}\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)$$

from the state equation. To check the convergence in the $L^2(\Omega)$ and $H^1(\Omega)$ norms, we first compute the discretization errors $\|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)}$ and $\|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}$ for finer and finer mesh sizes h and 6 different values of the regularization parameter ϱ; see Table 1 and Table 2, respectively. Since the given solution u_ϱ is smooth, we observe optimal convergence rates with respect to the mesh size h. Furthermore, the convergence does not deteriorate with respect to the regularization parameter ϱ. In Table 3, we numerically study the convergence of the approximation $u_{\varrho h}$ to the target $\bar{u}$ in the $L^2(\Omega)$ norm for decreasing ϱ by choosing $h = \varrho^{1/2}$ as predicted by Theorem 1. From Table 3, we clearly see a second-order convergence. We further observe a second-order convergence in the $H^{-1}(\Omega)$ norm; see also [22, Table 5].

h \ ϱ	10^0	eoc	10^{-2}	eoc	10^{-4}	eoc
2^{-2}	3.1863e−3		3.7710e−2		3.9112e−2	
2^{-3}	9.0214e−4	1.82	8.9461e−3	2.08	8.5855e−3	2.19
2^{-4}	2.2924e−4	1.98	2.1970e−3	2.03	2.0370e−3	2.08
2^{-5}	5.7222e−5	2.00	5.5283e−4	1.99	5.0177e−4	2.02
2^{-6}	1.4301e−5	2.00	1.4011e−4	1.98	1.2727e−4	1.98
2^{-7}	3.5796e−6	2.00	3.5473e−5	1.98	3.2775e−5	1.96
2^{-8}	8.9633e−7	2.00	8.9566e−6	1.99	8.4684e−6	1.95

h \ ϱ	10^{-6}	eoc	10^{-8}	eoc	10^{-10}	eoc
2^{-2}	3.9223e−2		3.9224e−2		3.9224e−2	
2^{-3}	8.6079e−3	2.19	8.6082e−3	2.19	8.6082e−3	2.19
2^{-4}	2.0355e−3	2.08	2.0356e−3	2.08	2.0356e−3	2.08
2^{-5}	4.9380e−4	2.04	4.9381e−4	2.04	4.9381e−4	2.04
2^{-6}	1.2136e−4	2.02	1.2133e−4	2.03	1.2133e−4	2.03
2^{-7}	3.0081e−5	2.01	3.0021e−5	2.01	3.0021e−5	2.01
2^{-8}	7.5374e−6	2.00	7.4604e−6	2.01	7.4619e−6	2.01

Table 1: Example 4.1: Error $\|u_\varrho - u_{\varrho h}\|_{L^2(\Omega)}$ with respect to h (columns), and fixed, but different ϱ (rows).

Table 4 provides the number of AMG preconditioned CG iterations, and the total coarsening and solving time measured in second (s) with respect to h and ϱ. From these numerical results, it is clear to see the relative robustness of the AMG preconditioner with respect to the mesh refinement and the decreasing regularization

ϱ \ h	10^0	eoc	10^{-2}	eoc	10^{-4}	eoc
2^{-2}	3.2364e−2		8.2832e−1		1.1721e−0	
2^{-3}	1.71297e−2	0.92	4.1447e−1	1.00	5.5195e−1	1.09
2^{-4}	8.6228e−3	0.99	2.0490e−1	1.02	2.6817e−1	1.04
2^{-5}	4.3068e−3	1.00	1.0187e−1	1.01	1.3232e−1	1.02
2^{-6}	2.1527e−3	1.00	5.0859e−2	1.00	6.5832e−2	1.00
2^{-7}	1.0771e−3	1.00	2.5439e−2	1.00	3.2889e−2	1.00
2^{-8}	5.3905e−4	1.00	1.2731e−2	1.00	1.6453e−2	1.00

ϱ \ h	10^{-6}	eoc	10^{-8}	eoc	10^{-10}	eoc
2^{-2}	1.1781e−0		1.1781e−0		1.1781e−0	
2^{-3}	5.5443e−1	1.09	5.5446e−1	1.09	5.5446e−1	1.09
2^{-4}	2.7032e−1	1.04	2.7036e−1	1.04	2.7036e−1	1.04
2^{-5}	1.3392e−1	1.01	1.3397e−1	1.01	1.3397e−1	1.01
2^{-6}	6.6689e−2	1.01	6.6787e−2	1.00	6.6789e−2	1.00
2^{-7}	3.3237e−2	1.00	3.3367e−2	1.00	3.3369e−2	1.00
2^{-8}	1.6566e−2	1.00	1.6683e−2	1.00	1.6687e−2	1.00

Table 2: Example 4.1: Error $\|\nabla(u_\varrho - u_{\varrho h})\|_{L^2(\Omega)}$ with respect to h (columns), and fixed, but different ϱ (rows).

ϱ	$\|\bar{u} - u_{\varrho h}\|^2_{H^{-1}(\Omega)}$	eoc	$\|\bar{u} - u_{\varrho h}\|^2_{L^2(\Omega)}$	eoc
10^0	3.54886e−3		1.16966e−1	
10^{-1}	2.11939e−3	0.22	6.98529e−2	0.22
10^{-2}	1.97952e−4	1.03	6.52427e−3	1.03
10^{-3}	3.13686e−6	1.80	1.03388e−4	1.80
10^{-4}	3.30578e−8	1.98	1.08962e−6	1.98

Table 3: Example 4.1: Errors $\|\bar{u} - u_{\varrho h}\|^2_{H^{-1}(\Omega)}$ and $\|\bar{u} - u_{\varrho h}\|^2_{L^2(\Omega)}$ for $\bar{u} \in H_0^1(\Omega) \cap H^2(\Omega)$.

parameter ϱ. The computational time scales well with respect to the number of unknowns. In Table 5, we study the robustness with respect to ϱ and parallel performance (strong scaling) of the BDDC preconditioned CG solver. So, we fix the total number of unknowns to $\#Dofs = 2,146,689$ that is related to $h = 1/128$. We observe that the numbers of BDDC iterations are rather stable with respect to the number of subdomains p, and the varying regularization parameter ϱ. Further, we observe perfect strong scalability (checking each row). Over-scaling from 32 to 64 cores may result from heterogeneity of the cluster, larger memory consumption for small number of cores, or the shared node resources with other users.

4.2 Discontinuous target

In the second example, we consider a discontinuous target function

$$\bar{u}(x) = \begin{cases} 1 & \text{for } x \in (0.25, 0.75)^3 \\ 0 & \text{for } x \in \Omega \setminus (0.25, 0.75)^3, \end{cases} \tag{22}$$

that is similar to the first example in [22]. In this case, we solve the equation on adaptively refined meshes that are driven by a residual based error indicator [33]. This leads to a lot of local refinements near the interface where the target is discontinuous. The energy regularization leads to a control that concentrates

h ϱ	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$
10^0	8 (0.01 s)	10 (0.06 s)	12 (0.72 s)	14 (8.29 s)	17 (127.06 s)	20 (1420.73 s)
10^{-2}	6 (0.01 s)	8 (0.05 s)	9 (0.56 s)	11 (6.83 s)	13 (94.28 s)	15 (910.77 s)
10^{-4}	5 (0.00 s)	4 (0.04 s)	4 (0.38 s)	5 (4.51 s)	7 (65.85 s)	8 (646.14 s)
10^{-6}	5 (0.00 s)	5 (0.04 s)	5 (0.41 s)	5 (4.46 s)	4 (52.90 s)	4 (499.10 s)
10^{-8}	5 (0.00 s)	5 (0.04 s)	5 (0.41 s)	5 (4.46 s)	5 (57.59 s)	5 (530.97 s)
10^{-10}	5 (0.00 s)	5 (0.04 s)	5 (0.41 s)	5 (4.43 s)	5 (57.04 s)	5 (560.33 s)
10^{-12}	5 (0.00 s)	5 (0.04 s)	5 (0.41 s)	5 (4.42 s)	5 (54.63 s)	5 (603.02 s)

Table 4: Example 4.1: Number of AMG preconditioned CG iterations, total coarsening and solving time measured in second (s), with respect to h and ϱ.

p ϱ	32	64	128	256
10^0	34 (46.43 s)	34 (16.04 s)	35 (5.92 s)	36 (2.98 s)
10^{-2}	32 (43.84 s)	32 (15.20 s)	33 (5.58 s)	34 (2.70 s)
10^{-4}	27 (37.20 s)	27 (13.12 s)	28 (4.85 s)	30 (2.39 s)
10^{-6}	20 (28.01 s)	20 (9.77 s)	22 (3.83 s)	22 (1.87 s)
10^{-8}	20 (28.15 s)	20 (9.75 s)	21 (3.65 s)	21 (1.76 s)
10^{-10}	20 (28.02 s)	20 (9.69 s)	21 (3.63 s)	21 (1.74 s)
10^{-12}	20 (28.01 s)	20 (9.71 s)	21 (3.69 s)	21 (1.67 s)

Table 5: Example 4.1: Number of iterations, and time measured in second (s) for the BDDC preconditioned CG solver, where $h = 1/128$ and $\#Dofs = 2,146,689$.

on the interface. As an illustration, we visualize the adaptive mesh, the state $u_{\varrho h}$, and the control $z_{\varrho h}$ in Fig. 1, which are comparable to the results in [22, Fig. 3]. Moreover, we provide the convergence of the approximation $u_{\varrho h}$ to the target $\bar{u}$ with respect to the regularization parameter ϱ in Table 6, i.e., an order $\sqrt{1.5}$ for the approximation in $H^{-1}(\Omega)$ and $\sqrt{0.5}$ in $L^2(\Omega)$. In this case, since $\bar{u} \in H^{1/2-\varepsilon}(\Omega)$, we obtain reduced convergence of the finite element approximation to the target; see also [22, Table 1]. Due to adaptive refinements for different regularization parameters $\varrho = 10^k$, $k = 0, -1, \ldots, -6$, the final meshes may contain different number of degrees of freedom. In Fig. 2, we visualize the number of AMG preconditioned CG iterations (left plot) and the computational time in seconds (right plot) for different values of ϱ during the adaptive refinement procedure. We observe that the AMG preconditioned CG iterations stay in a small range between 5 and 25; the computational time scales well with respect to the number of unknowns. To compare the performance of BDDC preconditioners with respect to an adaptive refinement, we select similar numbers of Dofs for all regularization parameters, i.e., #Dofs=2,210,254 ($\varrho = 10^0$), #Dofs=2,278,661 ($\varrho = 10^{-1}$), #Dofs=2,537,773 ($\varrho = 10^{-2}$), #Dofs=1,853,354 ($\varrho = 10^{-3}$), #Dofs=1,301,825 ($\varrho = 10^{-4}$), #Dofs=1,910,829 ($\varrho = 10^{-5}$), and #Dofs=1,895,056 ($\varrho = 10^{-6}$). In Table 7, we present the number of BDDC preconditioned CG iterations, and the corresponding solving time, measured in second (s), with respect to the number of subdomains p and the regularization parameter ϱ. For such adaptive meshes, we observe relatively good performance in both the number of BDDC preconditioned CG iterations and the computational time, as well as good strong scalability with respect to the number of subdomains p.

4.3 Smooth target with non-zero boundary conditions

In the third example, we use the smooth target

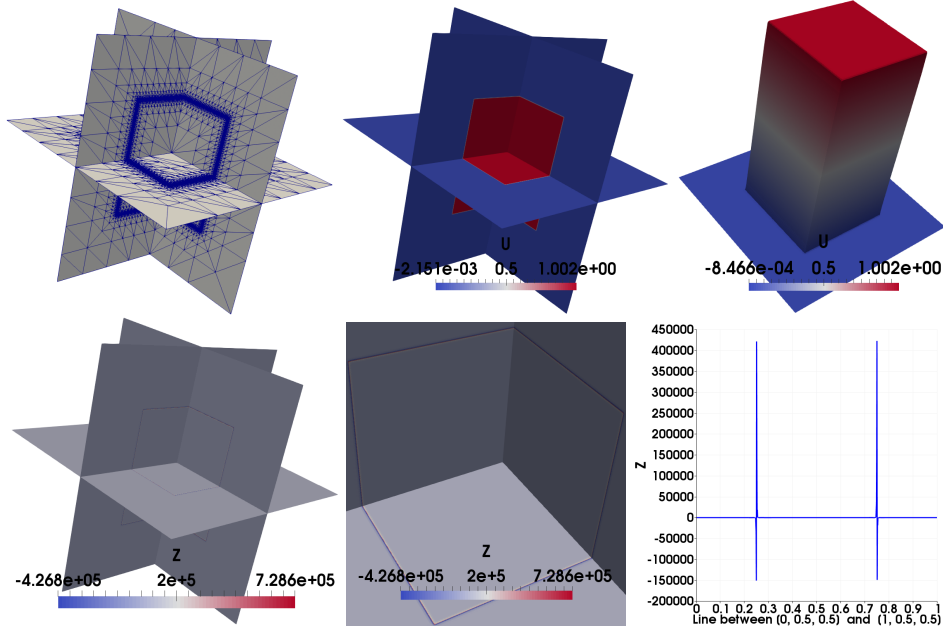$$\bar{u}(x_1, x_2, x_3) = 1 + \sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)$$

10

Figure 1: Example 4.2: Adaptive meshes at the 21th refinement step with $1,895,056$ Dofs, $u_{\varrho h}$ in Ω and on the cutting plane $x_3 = 0.5$ (up), control in Ω, near the interface, and on the line between $[0, 0.5, 0.5]$ and $[1, 0.5, 0.5]$ (down), $\varrho = 10^{-6}$.

ϱ	$\|\bar{u} - u_{\varrho h}\|^2_{H^{-1}(\Omega)}$	eoc	$\|\bar{u} - u_{\varrho h}\|^2_{L^2(\Omega)}$	eoc
10^0	2.1308e−3		1.2018e−1	
10^{-1}	1.3401e−3	0.20	9.0747e−2	0.12
10^{-2}	1.9101e−4	0.85	3.6563e−2	0.39
10^{-3}	9.0010e−6	1.33	1.1894e−2	0.49
10^{-4}	3.1821e−7	1.45	3.8028e−3	0.50
10^{-5}	1.0474e−8	1.48	1.2075e−3	0.50
10^{-6}	3.3491e−10	1.50	3.7988e−4	0.50

Table 6: Example 4.2: Errors $\|\bar{u} - u_{\varrho h}\|^2_{H^{-1}(\Omega)}$ and $\|\bar{u} - u_{\varrho h}\|^2_{L^2(\Omega)}$ of the approximations $u_{\varrho h}$ to the target $\bar{u} \in H^{1/2-\varepsilon}(\Omega)$.

considered in [22]. This target violates the homogeneous Dirichlet boundary conditions. As in the previous example, we solve the equation on adaptively refined meshes, which results in many local refinements near the boundary, where the control concentrates. For an illustration, we visualize the adaptive mesh, the state $u_{\varrho h}$, and the control $z_{\varrho h}$ in Fig. 3. Further, we observe the convergence order $\sqrt{1.5}$ for the approximation in $H^{-1}(\Omega)$, and $\sqrt{0.5}$ in $L^2(\Omega)$ with respect to ϱ; see Table 8. Since $\bar{u} \in C^\infty(\overline{\Omega})$, but $\bar{u} \notin H_0^1(\Omega)$, we only expect a reduced order of convergence for this example; see also [22, Table 7]. In Fig. 4, for each ϱ, we visualize the number of AMG preconditioned CG iterations (left plot) and the computational time in seconds (right plot) during the adaptive refinement levels. The AMG preconditioned CG iterations are in a small range between 1 and 20. The computational time scales well with respect to the number of unknowns. To see the performance of the BDDC preconditioner with respect to an adaptive refinement, we select similar numbers of Dofs for all regularization parameters, i.e., #Dofs=2,146,491 ($\varrho = 10^0$), #Dofs=2,146,689 ($\varrho = 10^{-1}$), #Dofs=2,146,575 ($\varrho = 10^{-2}$), #Dofs=2,654,801 ($\varrho = 10^{-3}$), #Dofs=1,997,688 ($\varrho = 10^{-4}$), #Dofs=3,688,105 ($\varrho = 10^{-5}$), and
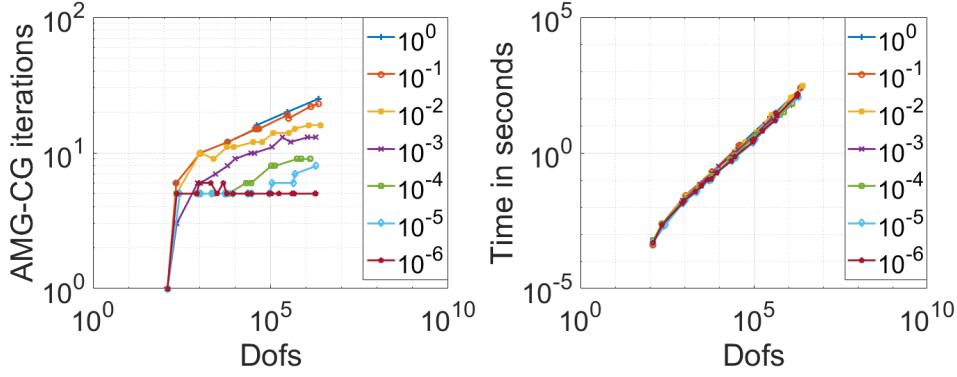
11

Figure 2: Example 4.2: Number of AMG preconditioned CG iterations (left) and computational time in seconds (right) with respect to different $\varrho \in \{10^0, 10^{-1}, ...10^{-6}\}$.

ϱ \ p	32	64	128	256	#Dofs
10^0	37 (60.77 s)	46 (24.34 s)	42 (8.20 s)	52 (5.02 s)	$2,210,254$
10^{-1}	42 (66.85 s)	45 (24.85 s)	46 (9.28 s)	43 (4.64 s)	$2,278,661$
10^{-2}	34 (60.06 s)	38 (23.86 s)	42 (9.75 s)	43 (4.79 s)	$2,537,773$
10^{-3}	35 (40.92 s)	39 (15.35 s)	35 (5.12 s)	40 (3.23 s)	$1,853,354$
10^{-4}	33 (16.79 s)	31 (6.30 s)	39 (3.17 s)	36 (1.53 s)	$1,301,825$
10^{-5}	33 (17.63 s)	33 (7.65 s)	33 (3.28 s)	38 (1.88 s)	$1,910,829$
10^{-6}	36 (18.30 s)	37 (8.42 s)	35 (3.31 s)	44 (1.93 s)	$1,895,056$

Table 7: Example 4.2: Number of BDDC preconditioned CG iterations, solving time measured in second (s) with respect to the number of subdomains p (= number of cores), regularization parameter ϱ, and different adaptive meshes (#Dofs).

#Dofs=3,676,447 ($\varrho = 10^{-6}$). Table 9 presents the number of BDDC preconditioned CG iterations, and solving times measured in second (s) with respect to the number of subdomains p and the regularization parameter ϱ. We again observe a relatively good performance in both the number of BDDC preconditioned CG iterations and the computational time, as well as good strong scalability with respect to the number of subdomains p.

ϱ	$\|\bar{u} - u_{\varrho h}\|^2_{H^{-1}(\Omega)}$	eoc	$\|\bar{u} - u_{\varrho h}\|^2_{L^2(\Omega)}$	eoc
10^0	3.5241e−2		1.5613e−0	
10^{-1}	2.1590e−2	0.21	1.0842e−0	0.16
10^{-2}	2.5430e−3	0.93	3.2593e−1	0.52
10^{-3}	8.8234e−5	1.46	9.5243e−2	0.53
10^{-4}	2.7302e−6	1.51	3.0292e−2	0.50
10^{-5}	8.5525e−8	1.50	9.6709e−3	0.50

Table 8: Example 4.3: Errors $\|\bar{u} - u_{\varrho h}\|^2_{H^{-1}(\Omega)}$ and $\|\bar{u} - u_{\varrho h}\|^2_{L^2(\Omega)}$ of the approximations $u_{\varrho h}$ to the target $\bar{u} \in C^\infty(\overline{\Omega}, \bar{u} \notin H^1_0(\Omega)$.
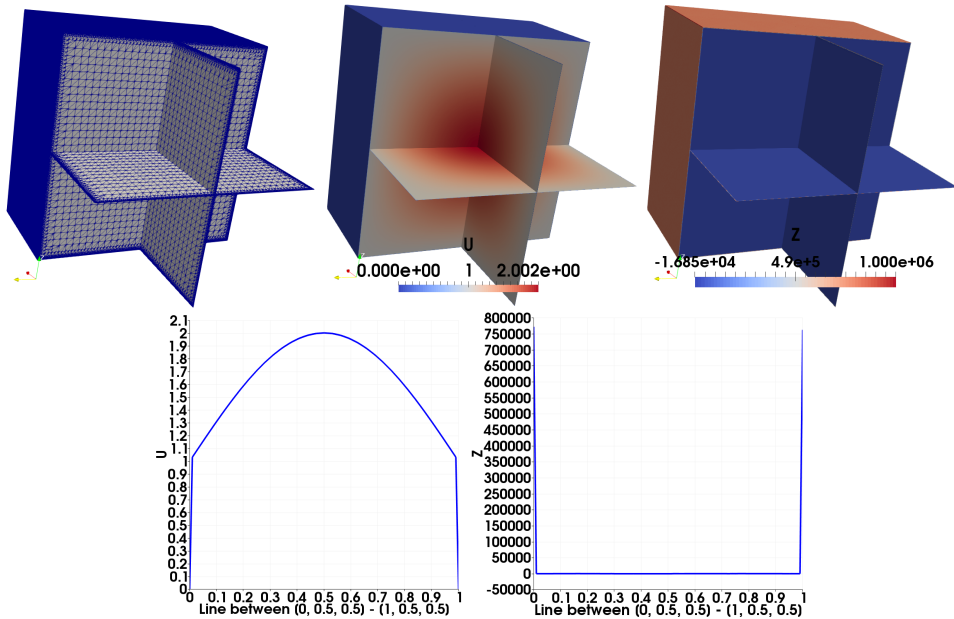
Figure 3: Example 4.3: Adaptive meshes at the 12th refinement step with $3, 688, 306$ Dofs, $u_{\varrho h}$ in Ω and $z_{\varrho h}$ in Ω (up); $u_{\varrho h}$ and $z_{\varrho h}$ along the line between $[0, 0.5, 0.5]$ and $[1, 0.5, 0.5]$ (down), $\varrho = 10^{-6}$.

5 Conclusions

Using estimates for the error between the exact solution u_{ϱ} of the optimal control problem (1) and the target $\overline{u}$ in terms of the regularization parameter ϱ, derived in [22], we have investigated the error between the finite element solution $u_{\varrho h}$ and the target $\overline{u}$ in terms of ϱ and h. Furthermore, we have studied AMG and BDDC preconditioned CG solvers for (adaptive) finite element equations arising from the elliptic optimal control problem with energy regularization. Both preconditioners have shown good performance with respect to (adaptive) mesh refinements and a (decreasing) regularization parameter. Moreover, the BDDC preconditioner has shown its strong scalability with respect to the number of subdomains on a distributed memory computer.

While in this paper we have considered a distributed control problem subject to the Poisson equation, a related analysis can be done in the case of a parabolic heat equation. These results will be published elsewhere, but see [16] for a space-time discretization approach.

Acknowledgments

References

[1] S. Badia, A. Martin, and J. Principe. Implementation and scalability analysis of balancing domain decomposition methods. *Arch. Computat. Methods Eng.*, 20:239–262, 2013.
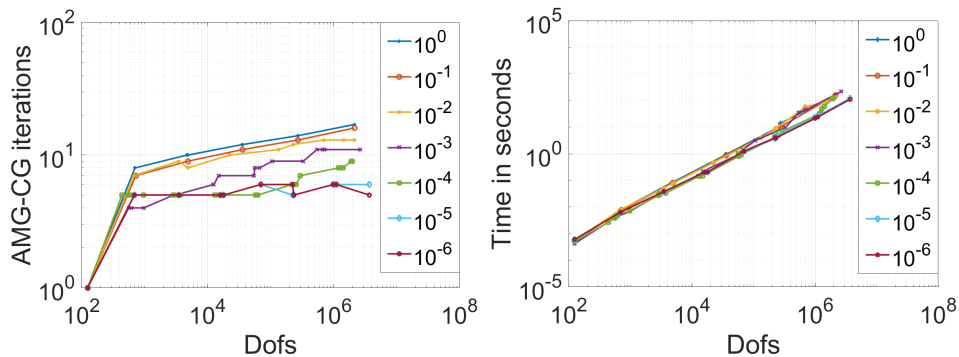
13

Figure 4: Example 4.3: Number of AMG preconditioned CG iterations (left) and computational time in seconds (right) with respect to different $\varrho \in \{10^0, 10^{-1}, ... 10^{-6}\}$.

ϱ \ p	32	64	128	256	#Dofs
10^0	33 (45.61 s)	32 (15.37 s)	32 (5.62 s)	33 (2.78 s)	$2,146,491$
10^{-1}	30 (41.51 s)	31 (15.16 s)	32 (5.71 s)	32 (2.84 s)	$2,146,689$
10^{-2}	30 (42.13 s)	31 (15.46 s)	30 (5.14 s)	32 (2.59 s)	$2,146,575$
10^{-3}	29 (32.46 s)	34 (14.30 s)	32 (7.20 s)	39 (3.27 s)	$2,654,801$
10^{-4}	29 (15.87 s)	31 (6.98 s)	31 (2.94 s)	29 (1.40 s)	$1,997,688$
10^{-5}	30 (21.03 s)	30 (9.22 s)	33 (4.34 s)	37 (2.73 s)	$3,688,105$
10^{-6}	31 (22.57 s)	33 (10.06 s)	31 (4.05 s)	38 (2.57 s)	$3,676,447$

Table 9: Example 4.3: Number of BDDC preconditioned CG iterations, solving time measured in second (s) with respect to the number of subdomains p (= number of cores), regularization parameter ϱ, and different adaptive meshes (#Dofs).

[2] S. Badia, A. F. Martín, and J. Principe. Multilevel balancing domain decomposition at extreme scales. *SIAM J. Sci. Comput.*, 38(1):C22–C52, 2016.

[3] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, Cambridge, 2007.

[4] A. Brandt, S. McCormick, and J. Ruge. Algebraic multigrid (AMG) for sparse matrix equations. In D. Evans, editor, *Sparsity and its Applications*, pages 257–284. Cambridge University Press, Cambridge, 1985.

[5] W. Briggs, V. Henson, and S. McCormick. *A Multigrid Tutorial*. SIAM, Philadelphia, second edition, 2000.

[6] C. Dohrmann. A preconditioner for substructuring based on constrained energy minimization. *SIAM J. Sci. Comput.*, 25(1):246–258, 2003.

[7] H. Goering, A. Felgenhauer, G. Lube, H. Roos, and L. Tobiska. *Singularly Perturbed Differential Equations*. Akademie Verlag, Berlin, 1983.

[8] H. Goering, M. Stynes, and L. Tobiska. *Robust Numerical Methods for Singularly Perturbed Differential Equations*. Springer Verlag, Berlin, 2008.

[9] G. Haase and U. Langer. Multigrid methods: from geometrical to algebraic versions. In A. Bourlioux, M. Gander, and G. Sabidussi, editors, *Modern Methods in Scientific Computing and Applications*, pages 103–153. Springer Netherlands, Dordrecht, 2002.

[10] W. Hackbusch. *Multi-Grid Methods and Applications.* Springer, Heidelberg, 1985.

[11] L. Jing and O. B. Widlund. FETI-DP, BDDC, and block Cholesky methods. *Int. J. Numer. Meth. Engng.*, 66(2):250–271, 2006.

[12] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.*, 20(1):359–392, 1998.

[13] F. Kickinger. Algebraic multi-grid for discrete elliptic second-order problems. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, pages 157–172, Berlin, Heidelberg, 1998. Springer Verlag.

[14] J. Kraus and M. Wolfmayr. On the robustness and optimality of algebraic multilevel methods for reaction-diffusion type problems. *Comput. Visual Sci.*, 16:15–32, 2013.

[15] U. Langer, M. Neumüller, and A. Schafelner. Space-time finite element methods for parabolic evolution problems with variable coefficients. In T. Apel, U. Langer, A. Meyer, and O. Steinbach, editors, *Advanced Finite Element Methods with Applications: Selected Papers from the 30th Chemnitz Finite Element Symposium 2017*, pages 247–275. Springer International Publishing, Cham, 2019.

[16] U. Langer, O. Steinbach, F. Tröltzsch, and H. Yang. Space-time finite element discretization of parabolic optimal control problems with energy regularization. *SIAM J. Numer. Anal.*, 2020. accepted.

[17] U. Langer and H. Yang. Robust and efficient monolithic fluid-structure-interaction solvers. *Int. J. Numer. Meth. Engng.*, 108(4):303–325, 2016.

[18] U. Langer and H. Yang. BDDC preconditioners for a space-time finite element discretization of parabolic problems. In R. Haynes, S. MacLachlan, X. Cai, L. Halpern, H. Kim, A. Klawonn, and O. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XXV*, pages 367–374, Cham, 2020. Springer.

[19] S. MacLachlan and N. Madden. Robust solution of singularly perturbed problems using multigrid methods. *SIAM J. Sci. Comput.*, 35(5):A2225–A2254, 2013.

[20] J. Mandel and C. Dohrmann. Convergence of a balancing domain decomposition by constraints and energy minimization. *Numer. Linear Algebra Appl.*, 10(7):639–659, 2003.

[21] J. Mandel, C. Dohrmann, and R. Tezaur. An algebraic theory for primal and dual substructuring methods by constraints. *Appl. Numer. Math.*, 54(2):167–193, 2005.

[22] M. Neumüller and O. Steinbach. Regularization error estimates for distributed control problems in energy spaces. *Math. Meth. Appl. Sci.*, 2020. published online.

[23] T. A. Nhan, S. MacLachlan, and N. Madden. Boundary layer preconditioners for finite-element discretizations of singularly perturbed reaction-diffusion problems. *Numer. Algor.*, 79:281–310, 2018.

[24] T. A. Nhan and N. Madden. An analysis of diagonal and incomplete Cholesky preconditioners for singularly perturbed problems on layer-adapted meshes. *J. Appl. Math. Computing*, 65:245–272, 2021.

[25] M. Olshanskii and A. Reusken. On the convergence of a multigrid method for linear reaction-diffusion problems. *Computing*, 65:193–202, 2000.

[26] C. Popa. Algebraic multigrid smoothing property of kaczmarz's relaxation for general rectangular linear systems. *Electron. Trans. Numer. Anal.*, 29:150–162, 2007-2008.

[27] J. Ruge and K. Stüben. Algebraic multigrid. In S. McCormick, editor, *Multigrid Methods*, chapter 4, pages 73–130. SIAM, Philadelphia, 1987.

[28] A. H. Schatz and L. B. Wahlbin. On the finite element method for singularly perturbed reaction-diffusion problems in two and one dimensions. *Math. Comp.*, 40(161):47–89, 1983.

[29] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary Elements*. Springer, 2008.

[30] O. Steinbach. Space-time finite element methods for parabolic problems. *Comput. Methods Appl. Math.*, 15:551–566, 2015.

[31] O. Steinbach and H. Yang. Comparison of algebraic multigrid methods for an adaptive space-time finite-element discretization of the heat equation in 3d and 4d. *Numer. Linear Algebra Appl.*, 25(3):e2143, 2018.

[32] A. Toselli and O. B. Widlund. *Domain Decomposition Methods — Algorithms and Theory*. Springer, Berlin, Heidelberg, 2005.

[33] R. Verfürth. *A posteriori error esimtation techniques for finite element methods*. Oxford University Press, Oxford, 2013.

[34] J. Xu and L. Zikatanov. Algebraic multigrid methods. *Acta Numer.*, 26:591–721, 2017.

[35] H. Yang and W. Zulehner. Numerical simulation of fluid-structure interaction problems on hybrid meshes with algebraic multigrid methods. *J. Comput. Appl. Math.*, 235(18):5367–5379, 2011.

[36] S. Zampini and X. Tu. Multilevel balancing domain decomposition by constraints deluxe algorithms with adaptive coarse spaces for flow in porous media. *SIAM J. Sci. Comput.*, 39(4):A1389–A1415, 2017.