

Boundary concentrated finite elements for optimal boundary control problems of elliptic PDEs

S. Beuchler, C. Pechstein, D. Wachsmuth

RICAM-Report 2010-04

Boundary concentrated finite elements for optimal boundary control problems of elliptic PDEs

Sven Beuchler ^{*} Clemens Pechstein [†] Daniel Wachsmuth [‡]

June 9, 2010

Abstract

We investigate the discretization of optimal boundary control problems for elliptic equations by the boundary concentrated finite element method. We prove that the discretization error $\|u^* - u_h^*\|_{L^2(\Gamma)}$ decreases like N^{-1} , where N is the total number of unknowns. This makes the proposed method favorable in comparison to the h -version of the finite element method, where the discretization error behaves like $N^{-3/4}$. Moreover, we present an algorithm that solves the discretized problem in almost optimal complexity. The paper is complemented with numerical results.

Keywords: optimal boundary control, boundary concentrated finite element method, discretization error estimates

1 Introduction

This paper deals with the numerical solution of the following optimal boundary control problem: Minimize the functional $J(y, u)$ given by

$$J(y, u) = \frac{1}{2} \int_{\Gamma_2} (y - y_d)^2 \, dS + \frac{\alpha}{2} \int_{\Gamma_2} u^2 \, dS, \quad (1.1)$$

where (y, u) solves

$$\begin{aligned} -\nabla \cdot (D\nabla y) + cy &= f && \text{in } \Omega, \\ y &= 0 && \text{on } \Gamma_1, \\ D \frac{\partial y}{\partial n} &= u && \text{on } \Gamma_2, \end{aligned} \quad (1.2)$$

and u obeys the control constraints

$$u_a \leq u \leq u_b \text{ a.e. on } \Gamma_2. \quad (1.3)$$

Here, the control is denoted by u , while the solution y of (1.2) is the corresponding state. We consider the two-dimensional case $\Omega \subset \mathbb{R}^2$. For the precise statement of the assumptions on the

^{*}Johann Radon Institute for Computational and Applied Mathematics Austrian Academy of Sciences, Altenberger Straße 69, 4040 Linz, Austria, sven.beuchler@ricam.oeaw.ac.at

[†]Institute of Computational Mathematics, Johannes Kepler University, Altenberger Straße 69, 4040 Linz, Austria, clemens.pechstein@numa.uni-linz.ac.at

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

data of this problem we refer to Section 2.1. Under these assumption, the optimal control problem is uniquely solvable, where we denote the optimal control by u^* with associated optimal state y^* .

In order to compute approximations of the solution (y^*, u^*) a finite-dimensional optimization problem has to be derived. The numerical solution of (1.1)-(1.3) requires discretization methods for the approximate solution of (1.2). Among other methods, the finite element method (FEM) of low order, i.e. the h -version of the FEM, is a very powerful method for the discretization of (1.2), see e.g. [4, 8]. Applied to optimal control problems of the form (1.1), let us comment on available error estimates of the type

$$\|u^* - u_h^*\|_{L_2(\Gamma)} = \mathcal{O}(h^s)$$

for the h -version of the FEM in the H^2 -regular case. Casas, Mateos, and Tröltzsch [6] proved the order $s = 1$ for piecewise constant discretization of the control for general boundary control problems with smooth data. The case of variational discretization of the control was treated by Casas and Mateos [5] as well as Hinze and Matthes [19], resulting in the optimal order $s = 3/2$. Since the number of unknowns N is proportional to h^{-d} , this implies for the two-dimensional case $d = 2$ that the error behaves at best like $\|u^* - u_h^*\|_{L_2(\Gamma)} = \mathcal{O}(N^{-3/4})$. The improved order of $s = 2 - \varepsilon$ was recently shown by Mateos and Rösch [24] for problems with domain observation, e.g. with $J_\Omega(y, u) = \frac{1}{2}\|y - y_d\|_{L_2(\Omega)}^2 + \frac{\alpha}{2}\|u\|_{L_2(\Gamma)}^2$. This is in accordance to the optimal order $s = 3/2$ for boundary observation, since typically the estimate for $\|y - y_h\|_{L_2(\Gamma)}$ is worse by a factor of $h^{-1/2}$ compared to $\|y - y_h\|_{L_2(\Omega)}$, which reduces the maximal achievable convergence order. We note that all the results mentioned above were obtained for quasi-uniform meshes with decreasing mesh size h and for a fixed polynomial degree of $p = 1$ or 2 . For related work on error estimates for optimal Dirichlet boundary control we refer to [7, 9, 25, 26].

For the discretization of boundary value problems of the form (1.2) with smooth solutions, spectral methods (see e.g. [20]) and finite elements of higher order (p -version, see e.g. [10, 11, 31, 33] and the references therein) have become more and more popular. In contrast to h -FEM, in the p -version of the FEM the polynomial degree p is increased and the mesh-size h is kept constant. Both ideas, mesh refinement and increasing the polynomial degree, are combined in the so-called hp -version of the FEM. The advantage of the p -version and hp -version in comparison to the h -version is that the discrete solution converges faster to the exact solution with respect to the number of unknowns N (of course provided that the solution is sufficiently smooth). Boundary concentrated finite element methods (BC-FEM) belong to the family of hp -FEM. In the BC-FEM, the meshes are refined geometrically to the boundary whereas the polynomial degree p is increased linearly to the interior in order to enrich the finite element approximation space, see Section 3.1 for a detailed description. The BC-FEM has been introduced and investigated by Khoromskij and Melenk [22], and allows us to solve problems of the form (1.2) with smooth coefficients as accurate as the corresponding standard finite element methods (FEM) but with a significantly smaller number of unknowns. Indeed, the total number of unknowns in the BC-FEM is $\mathcal{O}(h^{-(d-1)})$ compared to $\mathcal{O}(h^{-d})$ for the standard h -FEM, where h denotes the average discretization parameter of the boundary, and d is the spatial dimension. For a related early work, we refer to Yserentant [35].

In this paper, we will use the boundary concentrated finite element method (BC-FEM) to obtain a discretization of the elliptic equation (1.2). The controls are then treated by the variational discretization concept, as introduced by Hinze [18]. There the control is not discretized but given as the projection of a discrete adjoint state. Here, we prove an convergence order of $s = 1$ for the BC-FEM discretization, see Theorem 3.10. Since $h = \mathcal{O}(N^{-1})$ for the BC-FEM, this gives an error reduction of $\|u^* - u_h^*\|_{L_2(\Gamma)} = \mathcal{O}(N^{-1})$. Moreover, numerical experiments indicate an error reduction of $\mathcal{O}(N^{-3/2})$. On the contrary, the best known results in the literature for standard finite element discretizations yield a rate of $\|u_h^* - u^*\|_{L_2(\Gamma)} = \mathcal{O}(N^{-3/4})$ as mentioned above. This clearly favors the use of BC-FEM in case of boundary control problems with smooth data. For details, see Remark 3.12 and the numerical results in Section 5.

In addition, this paper presents fast solvers for the finite-dimensional optimization problem. Due to the special structure of the matrices in the BC-FEM discretization, this method allows for a direct solver of the resulting optimization problem in almost optimal complexity, see below Section 4.2. For the h -version direct solvers of the discrete problem with optimal or almost optimal complexity are not available. Here, one has to resort to the application of preconditioned iterative solvers for saddle point problems, see e.g. [16, 32, 30], which lead to algorithms with optimal arithmetical complexity.

The rest of the paper is organized as follows. In section 2, we introduce the optimal control problem and show existence and uniqueness of its solution. Section 3 is devoted to the discretization of the optimal control problem and to error estimates. In section 4, we consider algorithmic aspects of the numerical solution. In particular, we present an algorithm that solves the discretized problem in quasi-optimal computational complexity. Section 5 contains two numerical examples that support our theoretical results and compare BC-FEM against standard h -FEM. We conclude with a short discussion about possible generalizations in section 6.

Throughout this paper, the L_2 -norm of the trace of a function $v : \Omega \mapsto \mathbb{R}$ at Γ is simply denoted by $\|v\|_{L_2(\Gamma)}$.

2 Model Problem and Optimality Conditions

2.1 Model Problem

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain with $\Gamma_1 \cup \Gamma_2 = \partial\Omega$, $\Gamma_1 \cap \Gamma_2 = \emptyset$, $\text{meas}(\Gamma_1) > 0$, $\text{meas}(\Gamma_2) > 0$. Let us introduce the elliptic equation

$$\begin{aligned} -\nabla \cdot (D(x)\nabla y(x)) + c(x)y(x) &= f(x) && \text{in } \Omega, \\ y(x) &= 0 && \text{on } \Gamma_1, \\ D(x)\frac{\partial y}{\partial n}(x) &= u(x) && \text{on } \Gamma_2, \end{aligned} \tag{2.1}$$

which has to be solved for $y \in H^1(\Omega)$. Now, we consider the control problem of minimizing

$$J(u, y) = \frac{1}{2} \int_{\Gamma_2} (y(x) - y_d(x))^2 \, dS + \frac{\alpha}{2} \int_{\Gamma_2} u^2(x) \, dS \tag{2.2}$$

over all $(y, u) \in H^1(\Omega) \times L_2(\Gamma_2)$ subject to the elliptic equation (2.1) and the control constraints

$$u_a \leq u \leq u_b \quad \text{a.e. on } \Gamma_2. \tag{2.3}$$

Let us denote the optimal control problem (2.1)–(2.3) by (P). The set of admissible controls for (P) is given by

$$U_{ad} = \{u \in L_2(\Gamma_2) : u_a \leq u \leq u_b \quad \text{a.e. on } \Gamma_2\}.$$

Concerning the data of the state equation (2.1), we make the following smoothness assumption on the data.

Assumption 2.1. *The functions $D, c, f : \bar{\Omega} \mapsto \mathbb{R}$ are bounded and analytic. Moreover, we assume that $D(x) \geq D_0 > 0$ and $c(x) \geq 0$ for all $x \in \bar{\Omega}$.*

In order to keep the notation as simple as possible we restrict ourselves to homogeneous Dirichlet boundary conditions on Γ_1 . The results can be generalized to inhomogeneous Dirichlet boundary

conditions of the form $u = g$ on Γ_1 with $g \in H^{1/2}(\Gamma_1)$ in the second equation in (2.1).

In order to prove existence of solutions to (P) as well as a-priori discretization error estimates, we take the following assumptions on the data of the optimization problem.

Assumption 2.2. *We have $\alpha > 0$, $y_d \in L_2(\Gamma_2)$, and $u_a, u_b \in H^{1/2}(\Gamma_2)$ with $u_a \leq u_b$ a.e. on Γ_2 .*

We assume in the sequel that both Assumptions 2.1 and 2.2 are satisfied.

2.2 Existence of solutions and optimality conditions

Due to convexity, the problem under consideration is uniquely solvable.

Theorem 2.3. *The optimal control problem (P) admits a unique optimal control u^* .*

Moreover, the solution can be characterized by the following necessary optimality conditions. These conditions are also sufficient since the optimal control problem is convex.

Theorem 2.4. *Let u^* be the solution of (P) with associated state y^* . Then there exists an adjoint state $p^* \in H^1(\Omega)$ such that the adjoint equation*

$$\begin{aligned} -\nabla \cdot (D(x)\nabla p^*(x)) + c(x)p^*(x) &= 0 && \text{in } \Omega, \\ p^*(x) &= 0 && \text{on } \Gamma_1, \\ D(x)\frac{\partial p^*}{\partial n}(x) &= y^*(x) - y_d(x) && \text{on } \Gamma_2 \end{aligned} \quad (2.4)$$

and the variational inequality

$$\langle \alpha u^* + p^*, w - u^* \rangle_{L_2(\Gamma_2)} \geq 0 \quad \forall w \in U_{ad} \quad (2.5)$$

are satisfied. Moreover, the following pointwise representation of the optimal control holds

$$u^*(x) = \mathcal{P}_{[u_a, u_b]} \left(-\frac{1}{\alpha} p^*|_{\Gamma_2}(x) \right) \quad \text{a.e. on } \Gamma_2. \quad (2.6)$$

Here, $\mathcal{P}_{[u_a, u_b]}(w)$ denotes the projection of $w \in \mathbb{R}$ on the interval $[u_a, u_b]$.

For proofs of both theorems, we refer to [34].

The projection representation (2.6) implies that the optimal control is as regular as the trace of the adjoint state. Hence, $u^* \in H^{1/2}(\Gamma_2)$ if $u_a, u_b \in H^{1/2}(\Gamma_2)$. Additionally, one has the following regularity result for convex domains Ω .

Theorem 2.5. *Let Ω be a convex domain. Then we have the regularity $(y^*, u^*, p^*) \in H^2(\Omega) \times H^2(\Omega)$.*

Proof. As already mentioned, the projection representation (2.6) implies the regularity $u^* \in H^{1/2}(\Gamma_2)$. Then by standard regularity results, the optimal state y^* is in $H^2(\Omega)$. Hence, the Neumann data $y_d - y^*$ of the adjoint equation is in $H^{1/2}(\Gamma_2)$ thanks to Assumption 2.2, which in turn gives the regularity $p^* \in H^2(\Omega)$. \square

3 Discretization, error estimates

Since it is in general not possible to solve (P) analytically, we discretize (P). Following [19], we consider a sequence of problems where the state y is the solution of a finite element discretization of (2.1). Here, we use the boundary concentrated finite element method, which has been introduced by Khoromskij and Melenk, [22]. This method is described in the next subsection.

3.1 Boundary concentrated finite elements

We restrict our considerations to γ -shape-regular triangulations \mathcal{T} of Ω consisting of affine triangles, i.e., each element $K \in \mathcal{T}$ is the image $F_K(\hat{K})$ of the reference triangle \hat{K} , where the mapping F_K satisfies the inequality

$$h_K^{-1} \|F'_K\|_{L^\infty(K)} + h_K \|(F'_K)^{-1}\|_{L^\infty(K)} \leq \gamma \quad \forall K \in \mathcal{T}.$$

Here h_K denotes the diameter of the element K . Moreover, we assume \mathcal{T} to be a geometric mesh with boundary mesh size h (see also Figure 1).

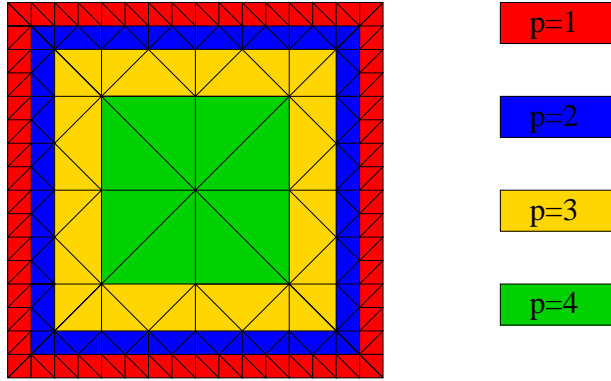


Figure 1: Boundary concentrated triangular finite element mesh and corresponding polynomial degree.

The precise description is given in the following definition.

Definition 3.1. (geometric mesh) *There exist constants $c_1, c_2 > 0$ such that for all $K \in \mathcal{T}$:*

1. *if $\overline{K} \cap \partial\Omega \neq \emptyset$, then $h \leq h_K \leq c_2 h$,*
2. *if $\overline{K} \cap \partial\Omega = \emptyset$, then $c_1 \inf_{\mathbf{x} \in K} \text{dist}(\mathbf{x}, \partial\Omega) \leq h_K \leq c_2 \sup_{\mathbf{x} \in K} \text{dist}(\mathbf{x}, \partial\Omega)$.*

In order to define hp -FEM spaces on the mesh \mathcal{T} , we associate a polynomial degree $p_K \in \mathbb{N}$ with each element $K \in \mathcal{T}$ and collect these p_K in the polynomial degree vector $\mathbf{p} := (p_K)_{K \in \mathcal{T}}$. Furthermore we associate a polynomial degree

$$p_e := \min \{p_K \mid e \text{ is an edge of element } K\} \quad (3.1)$$

with each edge e of the triangulation. We denote the vector containing the polynomial distribution of the triangle $K \in \mathcal{T}$ with edges $\{e_i \mid i = 1, 2, 3\}$ by

$$\mathbf{p}(K) := (p_{e_1}, p_{e_2}, p_{e_3}, p_K). \quad (3.2)$$

In conjunction with geometric meshes, a particularly useful polynomial degree distribution is the linear degree vector.

Definition 3.2 (linear degree vector). *Let \mathcal{T} be a geometric mesh on Ω with boundary mesh size h . A polynomial degree vector $\mathbf{p} = (p_K)_{K \in \mathcal{T}}$ is said to be a linear degree vector with slope $\alpha > 0$ if there exist constants $c_1, c_2 > 0$ such that*

$$1 + \alpha c_1 \log \frac{h_K}{h} \leq p_K \leq 1 + \alpha c_2 \log \frac{h_K}{h}. \quad (3.3)$$

Now we are in the position to define our hp -FEM spaces.

Definition 3.3 (hp -FEM spaces). *Let \mathcal{T}_l be a geometric mesh on Ω with boundary mesh size h_l . Let \mathbf{p} be a linear degree vector. Furthermore, for all edges e let p_e be given by (3.1) and for all $K \in \mathcal{T}$ let $\mathbf{p}(K)$ be given by (3.2). Then we set*

$$\begin{aligned}\mathbb{W}_l^{\mathbf{P}} &= \mathbb{W}^{\mathbf{P}}(\Omega, \mathcal{T}_l) := \{u \in H^1(\Omega) \mid u \circ F_K \in \mathcal{P}_{\mathbf{p}(K)}(\hat{K}) \quad \forall K \in \mathcal{T}_l\}, \\ \mathbb{W}_{l, \Gamma_1}^{\mathbf{P}} &:= \mathbb{W}_l^{\mathbf{P}} \cap H_{\Gamma_1}^1(\Omega),\end{aligned}$$

where

$$\mathcal{P}_{\mathbf{p}(K)}(\hat{K}) := \{u \in \mathcal{P}_{p_K}(\hat{K}) \mid u \in \mathcal{P}_{p_{e_i}}, i = 1, 2, 3\},$$

with $\hat{e}_1, \hat{e}_2, \hat{e}_3$ being the corresponding edges of the reference triangle \hat{K} .

A typical polynomial degree distribution is displayed in Figure 1. For $y, v \in H_{\Gamma_1}^1(\Omega)$, we introduce

$$\begin{aligned}a(y, v) &= \int_{\Omega} (\nabla y(x) \cdot D(x) \nabla v(x) + c(x) y(x) v(x)) \, dx, \\ l_u(v) &= \int_{\Omega} f(x) v(x) \, dx + \int_{\Gamma_2} u(x) v(x) \, dS.\end{aligned}$$

Now, the FE-discretization of equation (2.1) reads as follows.

Problem 3.4 (FE-approximation). *For given $u \in L_2(\Gamma_2)$ find $y_h \in \mathbb{W}_{l, \Gamma_1}^{\mathbf{P}}$ such that*

$$a(y_h, v_h) = l_u(v_h) \quad \forall v_h \in \mathbb{W}_{l, \Gamma_1}^{\mathbf{P}}. \quad (3.4)$$

To solve Problem 3.4 numerically, we equip the space $\mathbb{W}_{l, \Gamma_1}^{\mathbf{P}}$ with the basis

$$\Phi = \Phi_V \cup \Phi_E \cup \Phi_I = \{\phi_i \mid i = 1, \dots, N_b\}, \quad (3.5)$$

that is organized in the following way.

- The functions $\Phi_V = \{\phi_i \mid i = 1, \dots, N_V\}$ are the usual nodal basis functions associated to each vertex of the mesh (that is not on the Dirichlet boundary).
- The functions $\Phi_E = \{\phi_i \mid i = N_V + 1, \dots, N_E\}$ are edge bubble functions. Each such function is associated to an edge e_k of the mesh and supported by those two elements that have the edge e_k in common.
- The remaining functions $\Phi_I = \{\phi_i \mid i = N_E + 1, \dots, N_b\}$ are interior bubble functions. Each of them is associated to one element of the mesh and is supported on that element only.

Examples of concrete high order basis functions for triangular elements can be constructed with the aid of Jacobi polynomials, see e.g. [2, 3, 12, 20, 27]. Now, the solution of Problem (3.4) is equivalent to solve the system of linear algebraic equations

$$K \underline{y} = \underline{f}, \quad (3.6)$$

where

$$K = [a(\phi_j, \phi_k)]_{k, j=1}^{N_b}, \quad \underline{f} = [l_u(\phi_j)]_{j=1}^{N_b}, \quad \underline{y} = [y_k]_{k=1}^{N_b}.$$

In comparison to the h -version of the FEM, the boundary concentrated finite element method has much less unknowns. The following result (together with an elementary proof) can be found in [22, Proposition 2.7].

Lemma 3.5. *Let \mathcal{T} be a geometric mesh on Ω with boundary mesh size h , and let \mathbf{p} be a linear degree vector. Furthermore, for all edges e let p_e be given by (3.1), and for all $K \in \mathcal{T}$ let $\mathbf{p}(K)$ be given by (3.2). Then, $N_b = \dim(\mathbb{V}_{l,\Gamma_1}^{\mathbf{p}}) \sim h^{-1}$.*

The approximation properties of the boundary concentrated hp -FEM, are summarized in the following theorem.

Theorem 3.6. *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain and let \mathcal{T} be a geometric mesh on Ω with boundary mesh size h . Let \mathbf{p} be a linear degree vector with its slope α being sufficiently large. Let y and y_h be the solutions of (2.1) and (3.4), respectively. Then, if $y \in H^{1+\delta}(\Omega)$ for some $\delta \in (0, 1)$, we have*

$$\|y - y_h\|_{H^1(\Omega)} \leq C h^\delta.$$

Proof. The result follows from [22, Theorem 2.13]. \square

The theorem states that H^2 -regular problems have almost the same approximation properties in the H^1 -norm as an h -version of the FEM. Local error estimates in the L^2 norm have been developed in [13]. To the knowledge of the authors, global L^2 error estimates are not available.

For the optimal control problem (P), error estimates in the $L_2(\Gamma_2)$ norm are required.

Corollary 3.7. *Let the assumptions of Theorem 3.6 hold. If $y \in H^{1+\delta}(\Omega)$ for some $\delta \in (0, 1)$, then*

$$\|y - y_h\|_{L_2(\Gamma)} \leq C h^\delta.$$

Proof. The result follows from Theorem 3.6 and the fact that the trace operator is bounded from $H^1(\Omega)$ to $L_2(\Gamma)$, cf. e.g. [5]. \square

Remark 3.8. *In comparison to Theorem 3.6, the above corollary does not give the full approximation order $h^{\frac{3}{2}\delta}$. This is caused by the fact that right hand side of the error estimate of Theorem 3.6 is not of the form $C h^\delta \|f\|_{L_2}$. Therefore, the Aubin-Nitsche trick is not applicable. Much weaker local $L_2(\Omega)$ error estimates are available, see [13]. Numerical experiments indicate an approximation order of $h^{2\delta}$ in the $L_2(\Omega)$ norm, see also [13].*

3.2 Error estimates for the optimal control problem

The discretized optimal control problem is defined as the minimization of $J(y_h, u_h)$, given by (2.2), subject to the discretized equation (3.4), and the control constraints (2.3). This problem will be called (P_h) in the sequel. Its solution is denoted by u_h^* with discrete optimal state y_h^* .

Similarly as for the continuous problem (P), one can show that (P_h) admits a unique solution. Moreover, there exists a discrete adjoint state p_h^* that fulfills a discretized version of the adjoint equation (2.4).

All in all, the discrete solution $(y_h^*, p_h^*, u_h^*) \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{p}} \times \mathbb{V}_{l,\Gamma_1}^{\mathbf{p}} \times L_2(\Gamma_2)$ is given as the solution of the following first-order necessary optimality system, consisting of the discrete state equation

$$a(y_h^*, v_h) = \int_{\Omega} f(x) v_h(x) dx + \int_{\Gamma_2} u_h^*(x) v_h(x) dS = l_{u_h^*}(v_h) \quad \forall v_h \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{p}}, \quad (3.7)$$

the discrete adjoint equation

$$a(v_h, p_h^*) = \int_{\Gamma_2} v_h(x) (y_h^*(x) - y_d(x)) dS =: \langle v_h, y_h^* - y_d \rangle_{L_2(\Gamma_2)} \quad \forall v_h \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{p}}, \quad (3.8)$$

and the projection equation

$$u_h^* = \mathcal{P}_{[u_a, u_b]} \left(-\frac{1}{\alpha} p_h^*|_{\Gamma_2} \right). \quad (3.9)$$

The latter equation is equivalent to the variational inequality

$$\langle \alpha u_h^* + p_h^*, w - u_h^* \rangle_{L_2(\Gamma_2)} \geq 0 \quad \forall w \in U_{ad}. \quad (3.10)$$

Relation (3.9) implies that the discrete optimal control u_h^* is not a finite element function in general, since it is the pointwise projection of a grid function on the admissible interval $[u_a, u_b]$. Hence, u_h^* may have kinks that are not along the mesh edges.

Let us remark that the non-smooth system (3.7)–(3.9) can be solved by means of a semi-smooth Newton method, see below Section 4.1 below.

We are now in the position to formulate our approximation error estimates. To this end, we introduce the functions $y^h \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{P}}$ and $p^h \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{P}}$ (notice the superscript h) as the unique solutions of the discrete boundary value problems

$$a(y^h, v_h) = l_{u^*}(v_h) \quad \forall v_h \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{P}}, \quad (3.11)$$

$$a(v_h, p^h) = \langle v_h, y^* - y_d \rangle_{L_2(\Gamma_2)} \quad \forall v_h \in \mathbb{V}_{l,\Gamma_1}^{\mathbf{P}}, \quad (3.12)$$

for the state and adjoint state, respectively.

Theorem 3.9. *Let (u^*, y^*, p^*) and (u_h^*, y_h^*, p_h^*) be the solutions of the optimal control problems (P) and (P_h) , respectively. Moreover, let y^h and p^h be the solutions of (3.11) and (3.12), respectively. Then it holds*

$$\alpha \|u^* - u_h^*\|_{L_2(\Gamma_2)}^2 + \|y^* - y_h^*\|_{L_2(\Gamma_2)}^2 \leq \frac{1}{\alpha} \|p^* - p^h\|_{L_2(\Gamma)}^2 + \|y^* - y^h\|_{L_2(\Gamma)}^2 \quad (3.13)$$

Proof. The idea of the proof was given in [19] for a different functional.

Since the set of admissible controls is the same for the discretized problem as for the continuous one, the controls u^* and u_h^* are admissible for both problems. Hence, we can test (2.5) with u_h^* and (3.10) with u^* . This gives the two inequalities

$$\begin{aligned} \langle \alpha u_h^* + p_h^*, u^* - u_h^* \rangle_{L_2(\Gamma_2)} &\geq 0, \\ \langle \alpha u^* + p^*, u_h^* - u^* \rangle_{L_2(\Gamma_2)} &\geq 0. \end{aligned}$$

The summation of both inequalities, the weak formulation of (2.1), and equation (3.11) imply the estimate

$$\begin{aligned} \alpha \|u^* - u_h^*\|_{L_2(\Gamma_2)}^2 &\leq \langle u_h^* - u^*, p^* - p_h^* \rangle_{L_2(\Gamma_2)} \\ &= \langle u_h^* - u^*, p^* - p^h \rangle_{L_2(\Gamma_2)} + \langle u_h^* - u^*, p^h - p_h^* \rangle_{L_2(\Gamma_2)} \\ &\leq \|u^* - u_h^*\|_{L_2(\Gamma_2)} \|p^* - p^h\|_{L_2(\Gamma_2)} + a(y_h^* - y^h, p^h - p_h^*). \end{aligned}$$

Next, we use the relations (3.12) and (3.8) with $v_h = y_h - y^h$ and the elementary inequality $2(a-b)(b-c) \leq (a-c)^2 + (b-c)^2$ to obtain

$$\begin{aligned} a(y_h^* - y^h, p^h - p_h^*) &= a(y_h^* - y^h, p^h) - a(y_h^* - y^h, p_h^*) \\ &= \langle y^* - y_d, y_h^* - y^h \rangle_{L_2(\Gamma_2)} - \langle y_h^* - y_d, y_h^* - y^h \rangle_{L_2(\Gamma_2)} \\ &= \langle y^* - y_h^*, y_h^* - y^h \rangle_{L_2(\Gamma_2)} = -\frac{1}{2} \|y^* - y_h^*\|_{L_2(\Gamma_2)}^2 + \frac{1}{2} \|y^* - y^h\|_{L_2(\Gamma_2)}^2. \end{aligned}$$

Inserting this in the above estimate gives

$$\begin{aligned} \alpha \|u^* - u_h^*\|_{L_2(\Gamma_2)}^2 &\leq \|u^* - u_h^*\|_{L_2(\Gamma_2)} \|p^* - p^h\|_{L_2(\Gamma_2)} \\ &\quad - \frac{1}{2} \|y^* - y_h^*\|_{L_2(\Gamma_2)}^2 + \frac{1}{2} \|y^* - y^h\|_{L_2(\Gamma_2)}^2. \end{aligned}$$

Finally, Young's inequality $ab \leq \frac{1}{2\varepsilon} a^2 + \frac{\varepsilon}{2} b^2$ is applied to the first part of the right hand side of the previous estimate with $\varepsilon = \alpha$. This yields the estimates

$$\|u^* - u_h^*\|_{L_2(\Gamma_2)} \|p^* - p^h\|_{L_2(\Gamma_2)} \leq \frac{1}{2\alpha} \|p^* - p^h\|_{L_2(\Gamma_2)}^2 + \frac{\alpha}{2} \|u^* - u_h^*\|_{L_2(\Gamma_2)}^2$$

and

$$\begin{aligned} \alpha \|u^* - u_h^*\|_{L_2(\Gamma_2)}^2 &\leq \frac{1}{2\alpha} \|p^* - p^h\|_{L_2(\Gamma_2)}^2 + \frac{\alpha}{2} \|u^* - u_h^*\|_{L_2(\Gamma_2)}^2 - \\ &\quad - \frac{1}{2} \|y^* - y_h^*\|_{L_2(\Gamma_2)}^2 + \frac{1}{2} \|y^* - y^h\|_{L_2(\Gamma_2)}^2, \end{aligned}$$

which immediately implies the desired estimate. \square

We are now in the position to formulate our main approximation result.

Theorem 3.10. *Let (u^*, y^*, p^*) and (u_h^*, y_h^*, p_h^*) , be the solutions of the optimal control problems (P) , (P_h) with corresponding states and adjoint states, respectively. Let the solution of the boundary value problem (2.1) be $H^{1+\delta}$ -regular with $\delta \in (0, 1)$, that is $y^*, p^* \in H^{1+\delta}(\Omega)$. Then, there exists a generic constant C , independent of the discretization parameter h , such that*

$$\sqrt{\alpha} \|u^* - u_h^*\|_{L_2(\Gamma_2)} + \|y^* - y_h^*\|_{L_2(\Gamma_2)} \leq C h^\delta. \quad (3.14)$$

Proof. The result is a consequence of Theorem 3.9 and the approximation properties of the finite element spaces, cf. Theorem 3.7 and [4]. \square

Remark 3.11. *In the case of an H^2 -regular problem, cf. Theorem 2.5, we are able to prove a convergence order of h^δ with $\delta \rightarrow 1$ for the L_2 -norms of the errors of the control and state. For the h -version of the FEM, Casas and Mateos have proved a convergence order of $h^{3/2}$, see [5]. In our numerical experiments, we observed a convergence order of $h^{3/2}$ as well, but due to the non-applicability of the Aubin-Nitsche trick, this convergence rate can not be proven, see also the discussion in Remark 3.8 above.*

Remark 3.12. *Let N be the dimension of the finite element space $\mathbb{W}_{l,\Gamma_1}^{\mathbf{P}}$, and let h denote the meshsize on the boundary Γ . Since Lemma 3.5 implies $h \sim N^{-1}$, the result of Theorem 3.10 can be read as*

$$\sqrt{\alpha} \|u^* - u_h^*\|_{L_2(\Gamma_2)} + \|y^* - y_h^*\|_{L_2(\Gamma_2)} \leq C N^{-\delta}$$

if the problem is $H^{1+\delta}$ regular. This is much better than in the h -version FEM. In the latter case, the dimension N of the approximation space grows as $N \sim h^{-2}$. Combined with the estimate of [5], this gives

$$\sqrt{\alpha} \|u^* - u_h^*\|_{L_2(\Gamma_2)} + \|y^* - y_h^*\|_{L_2(\Gamma_2)} \leq C N^{-\frac{3}{4}\delta}$$

for an $H^{1+\delta}$ -regular problem, cf. [5]. For H^2 -regular problems, the error reduces as almost $\mathcal{O}(N^{-1})$ and as $\mathcal{O}(N^{-3/4})$ for the BC-FEM and the uniform h -FEM, respectively. Therefore, if the dimension N of the approximation space for the boundary value problems (P) , (2.4), is fixed, the discretization by boundary concentrated finite elements gives a smaller error for N being large enough.

Remark 3.13. *Instead of the variational discretization method above, one could also discretize the controls with piecewise constant functions. With standard arguments, see e.g. [6], one obtains the following a-priori error estimate*

$$\sqrt{\alpha} \|u^* - u_h^*\|_{L_2(\Gamma_2)} + \|y^* - y_h^*\|_{L_2(\Gamma_2)} \leq C h^\delta.$$

That is, the convergence order is the same as for the variational discretization, compare Theorem 3.10. However, this result gives the best possible convergence order, which is confirmed by our numerical experiments, where we got $\|u^* - u_h^*\|_{L_2(\Gamma_2)} = \mathcal{O}(h)$. Hence, the convergence order cannot be improved if finer L_2 -error estimates are available for the BC-FEM. Moreover, the numerical results indicate that the piecewise constant discretization yields to worse convergence rates than the variational discretization.

4 Algorithmic aspects and implementation

4.1 Solution of the discretized problem

In order to solve the discretized optimization problem (P_h) , one has to compute the numerical solution of the coupled problem (3.7)–(3.9), e.g.

$$\begin{aligned} a(y_h, v_h) &= \langle f, v_h \rangle_{L_2(\Omega)} + \langle u_h, v_h \rangle_{L_2(\Gamma_2)} & \forall v_h \in \mathbb{V}_{l,0}^{\mathbf{P}}, \\ a(v_h, p_h) &= \langle v_h, (y_h - y_d) \rangle_{L_2(\Gamma_2)} & \forall v_h \in \mathbb{V}_{l,0}^{\mathbf{P}}, \\ u_h &= \mathcal{P}_{[u_a, u_b]} \left(-\frac{1}{\alpha} p_h \right). \end{aligned} \quad (4.1)$$

for (u_h, y_h, p_h) . For a given u_h , it is possible to compute $y_h(u_h)$ from the first equation. If one inserts this into the second equation of (4.1), one can compute $p_h(y_h(u_h)) = p_h(u_h)$. Finally, this is inserted into the last equation. In the end, this gives a single equation for the optimal control u_h , e.g.

$$F(u_h) := u_h - \mathcal{P}_{[u_a, u_b]} \left(-\frac{1}{\alpha} p_h(u_h) \right) = 0. \quad (4.2)$$

This nonlinear and non-smooth equation is solved by a semi-smooth Newton-method. There, the non-smooth projection \mathcal{P} is linearized. We will sketch here only the algorithmic realization, for a convergence analysis we refer to [19].

Let an iterate u_h^k be given. Then the next iterate u_h^{k+1} is defined as the solution of the semi-smooth Newton step

$$F(u_h^k) + G(u_h^k)(u_h^{k+1} - u_h^k) = 0,$$

where $G(u_h^k) : L_2(\Gamma_2) \rightarrow L_2(\Gamma_2)$ is the element of the Newton derivative of F at u_h^k given by

$$G(u_h^k)(u_h^{k+1} - u_h^k) = (u_h^{k+1} - u_h^k) - d \cdot \left(-\frac{1}{\alpha} (p_h(u_h^{k+1}) - p_h(u_h^k)) \right)$$

with $d \in L^\infty(\Gamma_2)$ defined by

$$d(x) := \begin{cases} 0 & \text{if } -\frac{1}{\alpha} p_h(u_h^k)(x) \geq u_b(x) \\ 1 & \text{if } -\frac{1}{\alpha} p_h(u_h^k)(x) \in (u_a(x), u_b(x)) \\ 0 & \text{if } -\frac{1}{\alpha} p_h(u_h^k)(x) \leq u_a(x). \end{cases}$$

It turns out that u_h^{k+1} is the solution of the linear equation

$$u_h^{k+1}(x) = \begin{cases} u_b(x) & \text{if } -\frac{1}{\alpha} p_h(u_h^k)(x) \geq u_b(x) \\ -\frac{1}{\alpha} p_h(u_h^{k+1})(x) & \text{if } -\frac{1}{\alpha} p_h(u_h^k)(x) \in (u_a(x), u_b(x)) \\ u_a(x) & \text{if } -\frac{1}{\alpha} p_h(u_h^k)(x) \leq u_a(x). \end{cases} \quad (4.3)$$

for $x \in \Gamma_2$ a.e. In order to write this as an operator equation, let us introduce the following sets, where the control constraints are active or inactive:

$$\begin{aligned}\mathcal{A}_b(u_h) &= \left\{ x \in \Gamma_2 : -\frac{1}{\alpha} p_h(u_h^k)(x) \geq u_b(x) \right\}, \\ \mathcal{I}(u_h) &= \left\{ x \in \Gamma_2 : -\frac{1}{\alpha} p_h(u_h^k)(x) \in (u_a(x), u_b(x)) \right\}, \\ \mathcal{A}_a(u_h) &= \left\{ x \in \Gamma_2 : -\frac{1}{\alpha} p_h(u_h^k)(x) \leq u_a(x) \right\}.\end{aligned}$$

Additionally, let us define for a set $M \subset \Gamma_2$ the multiplication operator with the characteristic function of M by

$$E_M : L_2(\Gamma_2) \rightarrow L_2(\Gamma_2), \quad E_M(v) = \chi_M v.$$

Obviously, we have by (4.3) $u_h^{k+1}|_{\mathcal{A}_a(u_h^k)} = u_a|_{\mathcal{A}_a(u_h^k)}$ and $u_h^{k+1}|_{\mathcal{A}_b(u_h^k)} = u_b|_{\mathcal{A}_b(u_h^k)}$, hence only $u_{h,\mathcal{I}}^{k+1} := u_h^{k+1}|_{\mathcal{I}(u_h^k)}$ has to be determined by a linear system. Then (4.3) can be written equivalently as the linear equation

$$E_{\mathcal{I}(u_h^k)} \left(u_h^{k+1} + \frac{1}{\alpha} p_h(u_h^{k+1}) \right) = 0. \quad (4.4)$$

In order to write this equation as a linear equation in $u_{h,\mathcal{I}}^{k+1}$ let us introduce $z_h(u_h)$ and $q_h(u_h)$ as the solutions of the linear equations

$$\begin{aligned}a(z_h, v_h) &= \langle u_h, v_h \rangle_{L_2(\Gamma_2)} \quad \forall v_h \in \mathbb{V}_{l,0}^{\mathbf{P}}, \\ a(v_h, q_h) &= \langle v_h, z_h \rangle_{L_2(\Gamma_2)} \quad \forall v_h \in \mathbb{V}_{l,0}^{\mathbf{P}},\end{aligned}$$

which are similar to (4.1) but do not contain the affine linear terms f and y_d . Then $u_{h,\mathcal{I}}^{k+1}$ solves the system

$$u_{h,\mathcal{I}}^{k+1} + \frac{1}{\alpha} E_{\mathcal{I}(u_h^k)} q_h(u_{h,\mathcal{I}}^{k+1}) = -\frac{1}{\alpha} E_{\mathcal{I}(u_h^k)} p_h \left(E_{\mathcal{A}_a(u_h^k)} u_a + E_{\mathcal{A}_b(u_h^k)} u_b \right). \quad (4.5)$$

However, the solution $u_{h,\mathcal{I}}^{k+1}$ of this equation is not a finite element function in general, since it is the truncation of a finite element function $u_{h,\mathcal{I}}^{k+1} = E_{\mathcal{I}(u_h^k)} \tilde{u}_h^{k+1}$, which makes it difficult to handle numerically. However, it is easy to see that the finite element function \tilde{u}_h^{k+1} can be computed as the solution of

$$E_{\mathcal{I}(u_h^k)} \tilde{u}_h^{k+1} + \frac{1}{\alpha} E_{\mathcal{I}(u_h^k)} q_h(E_{\mathcal{I}(u_h^k)} \tilde{u}_h^{k+1}) = -\frac{1}{\alpha} E_{\mathcal{I}(u_h^k)} p_h \left(E_{\mathcal{A}_a(u_h^k)} u_a + E_{\mathcal{A}_b(u_h^k)} u_b \right). \quad (4.6)$$

To make this equation accessible for the numerical realization, we test it with a function $w_h \in U_h := (\text{span}(\Phi_v) \cap H_{\Gamma_1}^1) |_{\Gamma_2}$. Doing so, we obtain the following finite-dimensional system for the coefficient vector $\underline{\tilde{u}_h^{k+1}}$ that defines \tilde{u}_h^{k+1}

$$\left(M_{\mathcal{I}} + \frac{1}{\alpha} M_{\Gamma_2, \mathcal{I}}^\top K^{-1} M_{\Gamma_2} K^{-1} M_{\Gamma_2, \mathcal{I}} \right) \underline{\tilde{u}_h^{k+1}} = -\frac{1}{\alpha} M_{\Gamma_2, \mathcal{I}}^\top K^{-1} M_{\Gamma_2} (K^{-1} \underline{f} - \underline{y}_d) \quad (4.7)$$

with

$$M_{\mathcal{I}} = \left[\int_{\mathcal{I}(u_h^k)} \phi_i \phi_j \, dS \right]_{i,j=1}^{N_V}, \quad M_{\Gamma_2, \mathcal{I}} = \left[\int_{\mathcal{I}(u_h^k)} \phi_i \phi_j \, dS \right]_{i,j=1}^{N_b, N_V}, \quad M_{\Gamma_2} = \left[\int_{\Gamma_2} \phi_i \phi_j \, dS \right]_{i,j=1}^{N_b},$$

$$\underline{f} = \left[\int_{\Omega} f \phi_j \, dx + \int_{\mathcal{A}_a(u_h^k)} u_a \phi_j \, dS + \int_{\mathcal{A}_b(u_h^k)} u_b \phi_j \, dS \right]_{j=1}^{N_b}, \quad \underline{y_d} = \left[\int_{\Gamma_2} y_d \phi_j \, dS \right]_{j=1}^{N_b}.$$

If the solution \tilde{u}_h^{k+1} of this system is computed, then one can compute u_h^{k+1} and $p_h(u_h^{k+1})$ as

$$u_h^{k+1} = E_{\mathcal{A}_a(u_h^k)} u_a + E_{\mathcal{I}(u_h^k)} \tilde{u}_h^{k+1} + E_{\mathcal{A}_b(u_h^k)} u_b, \quad \text{and} \quad (4.8)$$

$$\underline{p_h(u_h^{k+1})} = K^{-1} M_{\Gamma_2} \left(K^{-1} \left(\underline{f} + M_{\Gamma_2, \mathcal{I}} \tilde{u}_h^{k+1} \right) - \underline{y_d} \right). \quad (4.9)$$

Only the adjoint state $p_h(u_h^{k+1})$ is used to define the active sets for the next iteration, hence kinks or discontinuities of the control iterates u_h^{k+1} will not be accumulated during the iterations.

Since the operator on the left-hand side of (4.6) is positive definite on $L_2(\mathcal{I}(u_h^k))$, the matrix on the left-hand side of (4.7) is positive definite on the span of all basis functions whose support has non-empty intersection with the inactive set $\mathcal{I}(u_h^k)$. That means, one can solve the finite-dimensional system (4.7) by the conjugate gradient (CG) method with diagonal preconditioner, [17].

In order to obtain a fast algorithm, the fast generation of the matrix K and the multiplications by the matrix K^{-1} (see (3.6)) in each iteration step of the CG-method have to be performed efficiently. The fast generation of a stiffness matrix in hp -FEM is much more advanced than for h -FEM. In hp -FEM, there are a lot of fast summation techniques available, see [10, 11, 20]. They guarantee the generation of K in almost optimal complexity, [13]. The solution of systems $K\underline{y} = \underline{f}$ is explained in the following subsection.

4.2 A fast direct solver for the boundary concentrated FEM

In this subsection, we review a fast direct solution method for the system $K\underline{y} = \underline{f}$ (3.6) which has been presented in [21]. The solver is based on the Cholesky decomposition of a permuted sparse matrix PKP^\top . The permutation matrix P can be chosen such that the decomposition of PKP^\top is much sparser than that of K , which results in a faster solution.

In order to find a good permutation matrix P , i.e., to find a good numbering of the basis functions, we use the method of nested dissection, cf. [14, 15]. In the following we describe that method for the (simplified) case $p = 1$ where we have only vertex-based functions. The set of nodes \mathcal{N} is split into three disjoint subsets, \mathcal{N}_1 , \mathcal{N}_2 and \mathcal{N}_0 , such that

- \mathcal{N}_0 is the separator of \mathcal{N}_1 , \mathcal{N}_2 , i.e. no finite element $\mathcal{T}_j \in T$ exists that has vertices both in \mathcal{N}_1 and \mathcal{N}_2 (see Figure 2 for an example).

The separator \mathcal{N}_0 is chosen such that its cardinality $|\mathcal{N}_0|$ is as small as possible. We now *renumber* the node indices, such that the nodes in \mathcal{N}_1 come first, then the nodes in \mathcal{N}_2 , and finally the nodes in \mathcal{N}_0 . Correspondingly, we get a permuted stiffness matrix \tilde{K} of the form

$$\tilde{K} = \begin{bmatrix} \tilde{K}_{11} & \mathbf{0} & \tilde{K}_{10} \\ \mathbf{0} & \tilde{K}_{22} & \tilde{K}_{20} \\ \tilde{K}_{10}^\top & \tilde{K}_{20}^\top & \tilde{K}_{00} \end{bmatrix}.$$

Due to the separator property of \mathcal{N}_0 , there is no coupling between the nodes in \mathcal{N}_1 and \mathcal{N}_2 . Next, this renumbering strategy is applied to the sets \mathcal{N}_1 and \mathcal{N}_2 recursively. The construction of the separators is described in [21] and can be straightforwardly extended to the case $p > 1$. Khoromskij and Melenk give the following result on the computational cost and storage.

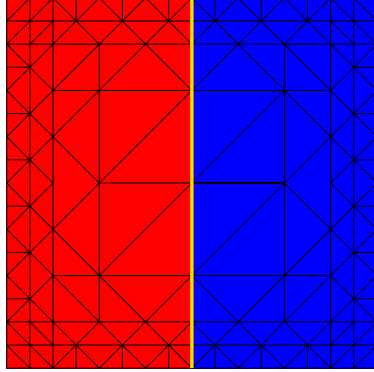


Figure 2: Separator of the nodes of a boundary concentrated finite element mesh. The separator, e.g. the nodes in \mathcal{N}_0 is colored yellow, the nodes in \mathcal{N}_1 and \mathcal{N}_2 are colored, blue and red, respectively.

Theorem 4.1. *Let P be the permutation matrix corresponding to a renumbering $[\phi_{i_1}, \dots, \phi_{i_{N_b}}]$ of the basis functions $[\Phi]$ based on nested dissection and let K be the original stiffness matrix (3.6).*

Then, the Cholesky decomposition of the matrix PKP^\top can be computed in $\mathcal{O}(N_b \log^8 N_b)$ floating point operations, where N_b denotes the size of the matrix K . The Cholesky decomposition requires a storage of $\mathcal{O}(N_b \log^3 N_b)$.

Proof. The proof has been presented in [21, Theorem 3.18]. □

Summarizing, a multiplication with the matrix in (4.7) can be performed in $\mathcal{O}(N_b \log^8 N_b)$ operations. Therefore, if the number of CG-iterations for the solution of (4.7) is bounded, the total cost for one Newton-step in (4.2) is $\mathcal{O}(N_b \log^8 N_b)$ flops.

Similarly, one can solve the system (4.6) in each step of the semi-smooth Newton method. There, a linear system of the form

$$\underbrace{\begin{bmatrix} E_k & \mathbf{0} & M_{\Gamma_2, \mathcal{I}}^\top \\ -M_{\Gamma_2, \mathcal{I}} & K & \mathbf{0} \\ \mathbf{0} & -M_{\Gamma_2} & K \end{bmatrix}}_{=: \mathcal{T}} \begin{bmatrix} \underline{u}_h \\ \underline{y}_h \\ \underline{p}_h \end{bmatrix} = \begin{bmatrix} E_{\mathcal{A}_a} u_a + E_{\mathcal{A}_b} u_b \\ \underline{f} \\ \underline{y}_d \end{bmatrix} \quad (4.10)$$

has to be solved for $[\underline{u}_h|_{\mathcal{I}}, \underline{y}_h, \underline{p}_h]^\top$, where

$$E_k = \alpha M_{\mathcal{I}} + E_{\mathcal{A}_a} + E_{\mathcal{A}_b}.$$

The renumbering in Theorem 4.1 by nested dissection can also be extended to the system (4.10), since all involved matrices have the same (or a similar) sparsity pattern as K . Therefore, we denote the basis functions by ϕ_j^u , ϕ_j^y and ϕ_j^p for the control u , the state y and the adjoint state p , respectively. For a given node j of the mesh, we introduce the row vector \underline{z}_j corresponding to all basis functions associated to this node, i.e. $\underline{z}_j = [\phi_j^u, \phi_j^y, \phi_j^p]$ if j is a node on the boundary, and $\underline{z}_j = [\phi_j^y, \phi_j^p]$ else. Based on the numbering of the basis functions in Theorem 4.1, the basis functions are ordered $[\underline{z}_{i_1}, \dots, \underline{z}_{i_{N_b}}]$. The corresponding permutation is denoted by \mathcal{P} . Then, the LU-decomposition of the matrix $\mathcal{P}\mathcal{T}\mathcal{P}^\top$ requires $\mathcal{O}(N_b \log^8 N_b)$ floating point operations. In comparison to the previous approach, the decomposition of the matrix \mathcal{T} has to be computed at each step of the semi-smooth Newton method, since the matrices E_k and $M_{\Gamma_2, \mathcal{I}}$ change and need to be recomputed in each step.

Remark 4.2. We arranged the matrix \mathcal{T} in (4.10) such that the diagonal blocks are symmetric positive definite, which prevents expensive pivot searching in the method of [21]. Writing this matrix T as a symmetric saddle point matrix would have produced a zero $(3, 3)$ -block, which would have made pivot search necessary and the resulting direct solver would not have optimal complexity anymore.

5 Numerical Experiments

In this section we present two examples. The first one demonstrates the correctness of BC-FEM and its efficiency compared to the h -FEM for a problem with known analytical solution, the second one is a problem with non-polynomial solution. The results are in accordance with our theory.

5.1 A problem with known analytical solution

Our first example is taken from [5]. Note that the setting is different from our theoretical setting. In particular, the deviation from the desired state is measured over the *domain* instead of a part of the boundary. Let $\Omega = [0, 1]^2$ and $\Gamma = \partial\Omega$. We consider the problem of minimizing

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

subject to $(y, u) \in H^1(\Omega) \times L_2(\Gamma)$ and $0 \leq u(x) \leq 1$ for all $x \in \Gamma$ a.e., and the linear state equation

$$\begin{aligned} -\Delta y(x) + c(x) y(x) &= e_1(x) & \forall x \in \Omega, \\ \frac{\partial y}{\partial n}(x) + y(x) &= e_2(x) + u(x) & \forall x \in \Gamma. \end{aligned}$$

The data is specified as follows,

$$\begin{aligned} \alpha &= 1, \\ c(x_1, x_2) &= 1 + x_1^2 - x_2^2, \\ e_1(x_1, x_2) &= -2 + (1 + x_1^2 - x_2^2)(1 + 2x_1^2 + x_1 x_2 - x_2^2), \\ e_2(x_1, x_2) &= \begin{cases} 1 - x_1 + 2x_1^2 - x_2^3 & \text{on } \Gamma_1 \\ 7 + 2x_2 - x_2^2 - \min(8(x_2 - 0.5)^2 + 0.58, 1) & \text{on } \Gamma_2 \\ -2 + 2x_1 + x_1^2 & \text{on } \Gamma_3 \\ 1 - x_2 - x_2^2 & \text{on } \Gamma_4 \end{cases} \\ e_u(x_1, x_2) &= \begin{cases} 1 - x_1^3 & \text{on } \Gamma_1 \\ -1 - \min(8(x_2 - 0.5)^2 + 0.58, 1 - 16x_2(x_2 - \xi_1^*)(x_2 - \xi_2^*)(x_2 - 1)) & \text{on } \Gamma_2 \\ -1 - x_2^2 & \text{on } \Gamma_3 \\ -1 + x_2(1 - x_2) & \text{on } \Gamma_4 \end{cases} \\ e_y(x_1, x_2) &= 1 \\ y_d(x_1, x_2) &= x_1^2 + x_1 x_2 \end{aligned}$$

where $\xi_{1,2}^* = \frac{1}{2} \pm \frac{\sqrt{21}}{20}$ and Γ_1 to Γ_4 are the four sides of the square, starting from the bottom side and going counterclockwise. The exact solution is given by

$$\begin{aligned} u^*(x_1, x_2) &= \begin{cases} x_1^3 & \text{on } \Gamma_1 \\ \min(8(x_2 - 0.5)^2 + 0.58, 1) & \text{on } \Gamma_2 \\ x_1^2 & \text{on } \Gamma_3 \\ 0 & \text{on } \Gamma_4 \end{cases} \\ y^*(x_1, x_2) &= 1 + 2x_1^2 + x_1 x_2 - x_2^2. \end{aligned}$$

The exact adjoint state is given by $p^*(x_1, x_2) = 1$.

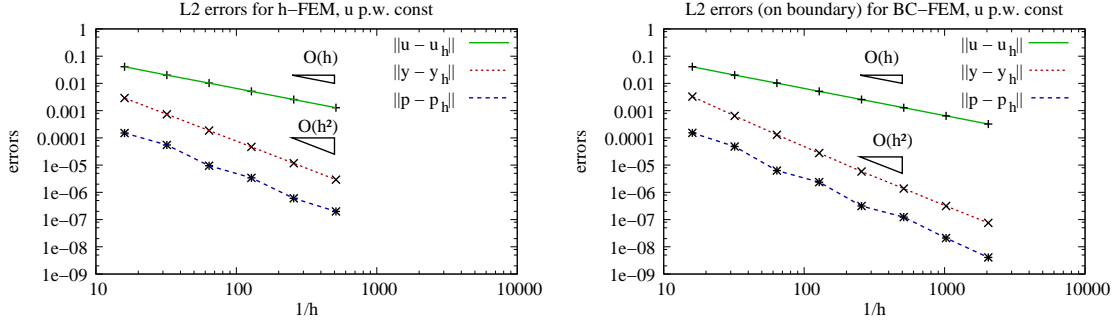


Figure 3: Piecewise constant control: errors of control, state, and adjoint state in the $L^2(\Gamma)$ -norm plotted against $1/h$, where h is the boundary mesh size. *Left: h-FEM. Right: BC-FEM.*

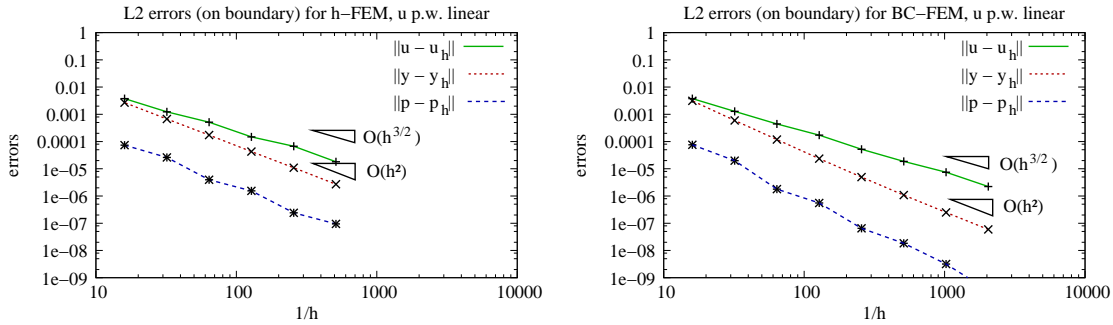


Figure 4: Variational discretization: $L^2(\Gamma)$ -errors of control, state, and adjoint state plotted against $1/h$, where h is the boundary mesh size. *Left: h-FEM. Right: BC-FEM.*

We discretize the problem with both h -FEM and BC-FEM. For the h -FEM we use a uniform triangulation of the domain. The BC-FEM is applied on a geometrically refined mesh with a polynomial degree vector with slope one, cf. Fig. 1 and Definition 3.2. Moreover, we compare two types of control discretizations: piecewise constant (as analyzed in e.g. [6]) and variational discretization, which we have analyzed in connection with BC-FEM. In order to solve the linear systems (determined from stiffness and mass matrices), we have used PARDISO [28, 29].

Let us report on the results for the two approaches for the discretization of the control.

- (a) *Discontinuous piecewise constant control.* In Fig. 3, we display the $L^2(\Gamma)$ -errors of the control, the state, and the adjoint state with respect to the known exact solution. Note that in this figure and in all the following ones, we use a double-logarithmic scale. In particular, we see that for both FE discretizations,

$$\|u_h - u^*\|_{L^2(\Gamma)} = \mathcal{O}(h), \quad \|y_h - y^*\|_{L^2(\Gamma)} = \mathcal{O}(h^2), \quad \|p_h - p^*\|_{L^2(\Gamma)} = \mathcal{O}(h^2).$$

This means, asymptotically, the h -FEM and the BC-FEM have the same convergence order for our optimal control problem. For a comparison of the two methods, we have plotted the control error against the degrees of freedom used for the state y_h in Fig. 5, left. We see that the BC-FEM outperforms the h -FEM.

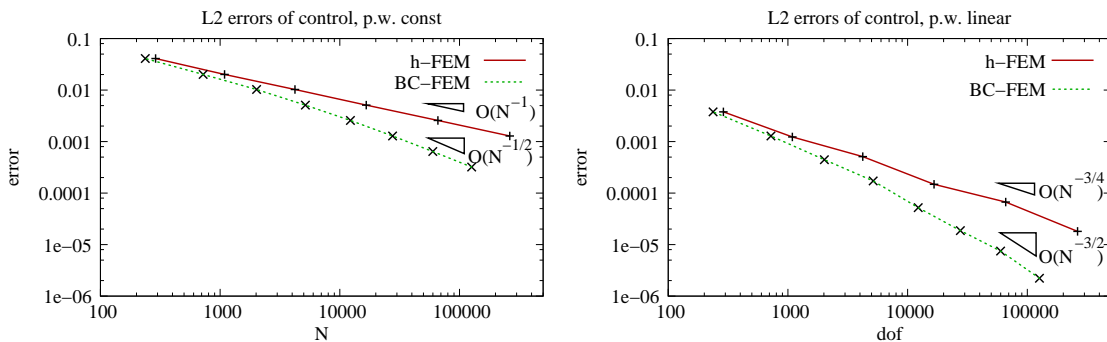


Figure 5: $L^2(\Gamma)$ -errors of control vs. number of degrees of freedom for the state for the h -FEM and BC-FEM. *Left*: piecewise constant control. *Right*: variational discretization.

- (b) *Variational discretization*. Only state and adjoint state are discretized, the control is discretized implicitly as explained in section 3.2. In Fig. 4, we display the L^2 -errors compared with the exact solution. There, for both FE discretizations,

$$\|u_h - u^*\|_{L_2(\Gamma)} = \mathcal{O}(h^{3/2}), \quad \|y_h - y^*\|_{L_2(\Gamma)} = \mathcal{O}(h^2), \quad \|p_h - p^*\|_{L_2(\Gamma)} = \mathcal{O}(h^2).$$

Again, a comparison of both methods with respect to the control error and the degrees of freedom (see Fig. 5, right) shows that the BC-FEM outperforms the h -FEM.

5.2 A problem problem with non-polynomial solution

In contrast to the previous example, the following one fits to our theoretical setting. Again, we consider the unit square $\Omega = (0, 1)^2$ with $\Gamma = \partial\Omega$. We wish to minimize the functional

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

subject to $(y, u) \in H^1(\Omega) \times L_2(\Gamma)$ and $-1 \leq u(x) \leq 0.8$ for all $x \in \Gamma$ a.e., and the linear state equation

$$\begin{aligned} -\Delta y(x) + y(x) &= 0 & \forall x \in \Omega, \\ \frac{\partial y}{\partial n}(x) &= u(x) & \forall x \in \Gamma, \end{aligned}$$

where $\alpha = 0.01$ and

$$y_d(x_1, x_2) = \sqrt{x_1 + x_2} \quad \text{for } (x_1, x_2) \in \Gamma.$$

In the following we use the implicit discretization of the control (as in Case (b) of the previous example). Fig. 6 shows plots of the approximants u_h along the boundary Γ for various refinement levels. In that figure the boundary is parameterized counterclockwise (arc length) starting at the origin.

With the h -FEM, we compute the solution for 8 refinement levels and use that as our reference solution. In Fig. 7, left, we plot the L^2 -error of u over boundary compared to the reference solution for h -FEM and BC-FEM for various refinement levels. For both methods we get the same order of convergence, namely

$$\|u_h - u_{\text{ref}}\|_{L^2(\Gamma)} = \mathcal{O}(h^2).$$

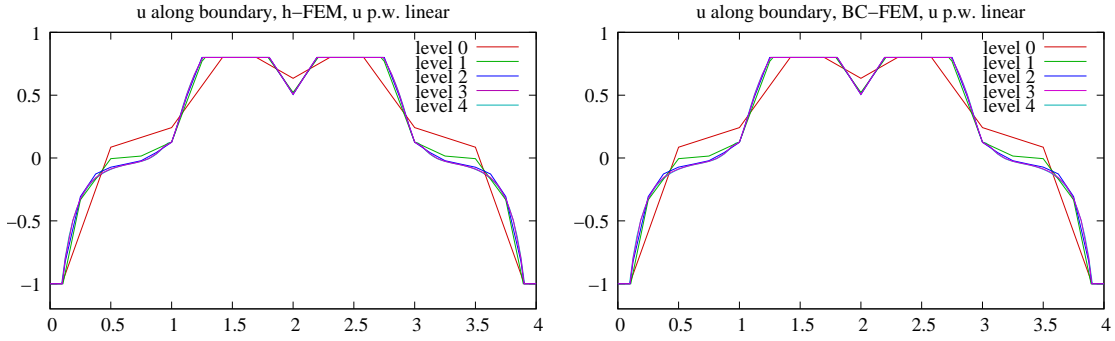


Figure 6: Control approximants u_h for various refinement levels, plotted *along* the boundary Γ . *Left*: h -FEM. *Right*: BC-FEM.

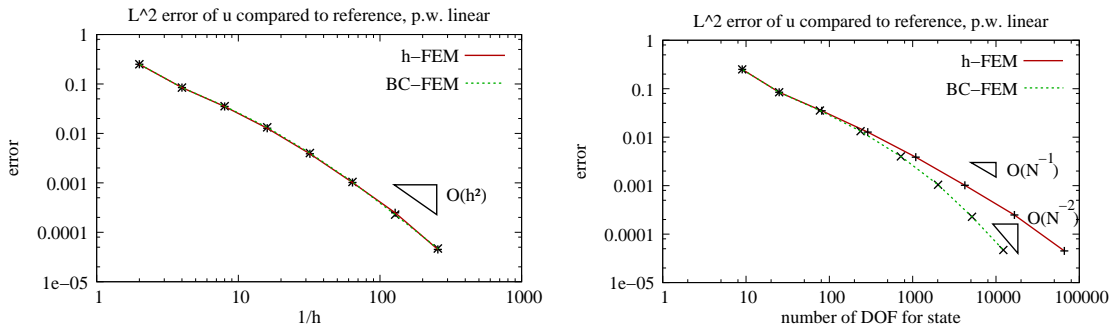


Figure 7: $L^2(\Gamma)$ -error of control with respect to reference solution: h -FEM and BC-FEM. *Left*: error vs. $1/h$. *Right*: error vs. number of degrees of freedom used for the state equation.

This indicates that the actual error behaves as $\mathcal{O}(h^2)$, which is better than the theoretically ensured rate of $\mathcal{O}(h^{3/2})$ for the h -FEM, cf. [5], and the (probably pessimistic) rate of $\mathcal{O}(h)$ for the BC-FEM, cf. Theorem 3.10. For the h -FEM, the error behaves as $\mathcal{O}(N^{-1})$, whereas it behaves as $\mathcal{O}(N^{-2})$ for the BC-FEM, cf. Fig. 7, right.

6 Conclusions and Generalizations

To the knowledge of the authors, this is the first paper that uses finite elements of high polynomial degree p for an optimal control problem. We proved that the dimension N of the approximation space can be decreased by the usage of BC-FEM. This accelerates the numerical algorithms and makes it possible to compute the solution with higher accuracy. The convergence estimates of Theorem 3.10 require smooth data for the diffusion coefficient D and the right-hand side f in (2.1). By straightforward arguments, these results can be generalized to piecewise smooth diffusion coefficients. Then, the interface concentrated FEM has to be used instead of BC-FEM for the discretization of the state equation (2.1) and of the adjoint equation (2.4), see [1, 23].

However, the convergence estimate results are limited to the cases where approximation results for BC-FEM are available. Nevertheless, it should be possible to generalize the results to other problem classes, e.g.

- the three-dimensional case,
- minimizing a functional in the domain, e.g. $\int_{\Omega}(y - y_d)^2 + \alpha \int_{\Gamma_2} u^2$ instead of (1.1),

if the corresponding approximation results are provided. In all two cases, this is expected to be very technical.

Another open question is the proof of the numerical observed approximation order $\mathcal{O}(h^{3/2})$, which requires new techniques for L_2 -error estimates.

Acknowledgements

This work was supported by the Austrian Academy of Sciences and by the FWF projects P20121-N18 and P21564-N18.

References

- [1] S. Beuchler, T. Eibner, and U. Langer. Primal and dual interface concentrated iterative substructuring methods. *SIAM J. Numer. Anal.*, 46(6):2818–2842, 2008.
- [2] S. Beuchler and V. Pillwein. Shape functions for tetrahedral p -fem using integrated Jacobi polynomials. *Computing*, 80:345–375, 2007.
- [3] S. Beuchler and J. Schöberl. New shape functions for triangular p -FEM using integrated Jacobi polynomials. *Numer. Math.*, 103(3):339–366, 2006.
- [4] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer Verlag, New York, 1994.
- [5] E. Casas and M. Mateos. Error estimates for the numerical approximation of Neumann control problems. *Comput. Optim. Appl.*, 39(3):265–295, 2008.
- [6] E. Casas, M. Mateos, and F. Tröltzsch. Error estimates for the numerical approximation of boundary semilinear elliptic control problems. *Comput. Optim. Appl.*, 31(2):193–219, 2005.
- [7] E. Casas and J.-P. Raymond. Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. *SIAM J. Control Optim.*, 45(5):1586–1611 (electronic), 2006.

- [8] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.
- [9] K. Deckelnick, A. Günther, and M. Hinze. Finite element approximation of Dirichlet boundary control for elliptic PDEs on two- and three-dimensional curved domains. *SIAM J. Control Optim.*, 48(4):2798–2819, 2009.
- [10] L. Demkowicz. *Computing with hp Finite Elements*. CRC Press, Taylor and Francis, 2006.
- [11] L. Demkowicz, J. Kurtz, D. Pardo, M. Paszyński, W. Rachowicz, and A. Zdunek. *Computing with hp-adaptive finite elements. Vol. 2*. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2008. Frontiers: three dimensional elliptic and Maxwell problems with applications.
- [12] M. Dubiner. Spectral methods on triangles and other domains. *J. Sci. Computing*, 6:345, 1991.
- [13] T. Eibner. *Randkonzentrierte und adaptive hp-FEM*. PhD thesis, TU Chemnitz, 2006.
- [14] A. George. Nested dissection of a regular finite element mesh. *SIAM J. Numer. Anal.*, 10:345–363, 1973.
- [15] A. George and J. W.-H. Liu. *Computer solution of large sparse positive definite systems*. Prentice-Hall Inc. Englewood Cliffs. New Jersey, 1981.
- [16] R. Herzog and E. Sachs. Preconditioned conjugate gradient method for optimal control problems with control and state constraints. Technical Report 088, SPP1253, 2009.
- [17] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, 49:409–436, 1952.
- [18] M. Hinze. A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.*, 30(1):45–61, 2005.
- [19] M. Hinze and U. Matthes. A note on variational discretization of elliptic neumann boundary control. *Contr. Cyben.*, 2009. to appear.
- [20] G. M. Karniadakis and S. J. Sherwin. *Spectral/hp element methods for CFD*. Oxford University Press. Oxford, 1999.
- [21] B. N. Khoromskij and J. M. Melenk. An efficient direct solver for the boundary concentrated FEM in 2D. *Computing*, 69(2):91–117, 2002.
- [22] B. N. Khoromskij and J. M. Melenk. Boundary concentrated finite element methods. *SIAM J. Numer. Anal.*, 41(1):1–36, 2003.
- [23] U. Langer and C. Pechstein. All-floating coupled data-sparse boundary and interface-concentrated finite element tearing and interconnecting methods. *Comput. Vis. Sci.*, 11(4-6):307–317, 2008.
- [24] M. Mateos and A. Rösch. On saturation effects in the Neumann boundary control of elliptic optimal control problems. *Comput. Optim. Appl.*, 2009. online first, DOI 10.1007/s10589-009-9299-5.
- [25] S. May, R. Rannacher, and B. Vexler. Error analysis for a finite element approximation of elliptic Dirichlet boundary control problem. 2008.
- [26] G. Of, T. X. Phan, and O. Steinbach. Boundary element methods for Dirichlet boundary control problems. Bericht 2010/1, Institut für Numerische Mathematik, TU Graz, 2010.
- [27] S. A. Orszag. Spectral methods for problems in complex geometries. *J. Comp. Phys.*, pages 37–80, 1980.
- [28] O. Schenk and K. Gärtner. Solving unsymmetric sparse systems of linear equations with PARDISO. *Journal of Future Generation Computer Systems*, 20(3):475–487, 2004.
- [29] O. Schenk and K. Gärtner. On fast factorization pivoting methods for sparse symmetric indefinite systems. *Elec. Trans. Numer. Anal.*, 23:158–179, 2006.
- [30] 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(3):752–773 (electronic), 2007.

- [31] C. Schwab. *p- and hp-finite element methods. Theory and applications in solid and fluid mechanics.* Clarendon Press, Oxford, 1998.
- [32] R. Simon and W. Zulehner. On Schwarz-type smoothers for saddle point problems with applications to PDE-constrained optimization problems. *Numer. Math.*, 111(3):445–468, 2009.
- [33] P. Solin, K. Segeth, and I. Dolezel. *Higher-Order Finite Element Methods.* Chapman and Hall, CRC Press, 2003.
- [34] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen.* Vieweg, Wiesbaden, 2005.
- [35] H. Yserentant. Coarse grid spaces for domains with a complicated boundary. *Numer. Algorithms*, 21(1-4):387–392, 1999. Numerical methods for partial differential equations (Marrakech, 1998).