Boundary concentrated finite elements for optimal control problems with distributed observation

S. Beuchler, * K. Hofer *, D. Wachsmuth, † J.-E. Wurst †

September 2, 2015

Abstract

We consider the discretization of an optimal boundary control problem with distributed observation by the boundary concentrated finite element method. If the constraint is a $H^{1+\delta}(\Omega)$ regular elliptic PDE with smooth differential operator and source term, we prove for the two dimensional case that the discretization error in the L_2 norm decreases like $N^{-\delta}$, where N is the number of unknowns. Our approach is suitable for solving a wide class of problems, among them piecewise defined data and tracking functionals acting only on a subdomain of Ω. We present several numerical results.

Acknowledgment. This work was funded by the Austrian Science Fund (FWF) grant P23484-N18.

1 Introduction

We investigate a higher-order discretization technique in order to compute approximate solutions to the following optimal control problem: Minimize the functional

$$J(y,u) := \left(\frac{1}{2} \int_\Omega (y(x) - y_d(x))^2 \,\mathrm{d}x + \frac{\alpha}{2} \int_{\Gamma_\mathcal{N}} u(x)^2 \,\mathrm{d}s_x \right) \tag{P}$$

subject to the elliptic equation

$$\begin{aligned} -\nabla \cdot (D(x)\nabla y(x)) + c(x)y(x) &= f(x) && \text{in } \Omega, \\ y(x) &= 0 && \text{on } \Gamma_\mathcal{D}, \\ \partial_{n_D} y(x) &= u(x) && \text{on } \Gamma_\mathcal{N}, \end{aligned} \tag{1.1}$$

and the box constraints

$$u_a(x) \le u(x) \le u_b(x) \qquad \text{a.e. on } \Gamma_\mathcal{N}. \tag{1.2}$$

Here, the boundary control is denoted by u, while the state is denoted by y. Under the assumptions specified in chapter 2 this problem admits a unique solution u^*.

*Institut für Numerische Simulation, Universität Bonn, 53115 Bonn, Germany
S.Beuchler: `beuchler@ins.uni-bonn.de`
K. Hofer: `hofer@ins.uni-bonn.de`
†Institut für Mathematik, Universität Würzburg, 97074 Würzburg, Germany
D. Wachsmuth: `daniel.wachsmuth@mathematik.uni-wuerzburg.de`
J.-E. Wurst: `jan-eric.wurst@mathematik.uni-wuerzburg.de`

Let us comment on existing approximation results for Neumann boundary control problems for elliptic PDEs. Depending on the discretization scheme, one can derive error estimates of the type

$$\|u^* - u_h^*\|_{L_2(\Gamma_{\mathcal{N}})} \leq Ch^s, \tag{1.3}$$

where h indicates the mesh size, and u_h^* denotes the approximation of u^*.
For a piecewise constant approximation u_h^* of u^* in the case of a convex domain, Casas, Mateos, Tröltzsch [9] proved this estimate with $s = 1$. A piecewise linear discretization yields (1.3) with $s = 3/2 - \epsilon$, as proved by Casas, Mateos [8]. Hinze, Matthes [16] established $s = 3/2$ for the variational discretization concept and also provide an L^∞ estimate with $s = 2$ and the additional factor $|\log h|$ for smooth domains. We also mention Mateos, Rösch [22], who prove approximation results with $s \in [1, 2]$, depending on the angles of a (possibly non-convex) domain Ω. In the convex case, the rate of $s = 2 - 1/p$ is shown if the optimal state y is in $W_p^2(\Omega)$. In the non-convex case and $D(x) \equiv I$, convergence rates $s = 1/2 + \pi/\omega$ are obtained, where ω is the largest inner angle of the domain. Apel, Pfefferer and Rösch [2] showed how the order $s = 3/2$ can be obtained for non-convex domains by using sufficiently graded meshes. All these results were obtained for finite elements with fixed polynomial degree $p = 1$ or $p = 2$ for the discretization of the elliptic equation.
If solutions of (1.1) are sufficiently smooth, then higher-order polynomials can be used to approximate these solutions efficiently. Depending on the smoothness of the approximated variable, h-refinement (refining elements) or p (increasing local polynomial degree) is applied leading to fast convergence. A detailed description of h and hp finite elements is given in the monographs [12, 18, 23, 26].
Let us briefly report on available literature on p- and hp-approximation of optimal control problems. Spectral methods on $\Omega = (-1, 1)^d$ were investigated by Chen, Yi, Liu [11]. Adaptive hp-methods were analyzed by Chen, Lin [10], and Gong, Liu, Yan [14]. In all these references distributed control problems with integral control constraint $\int_\Omega u \,\mathrm{d}x \geq 0$ were considered. Here, the regularity of the optimal control is not restricted by the constraint, whereas pointwise constraints of the type (1.2) restrict the regularity to $u \in W_p^1(\Omega)$ in the distributed control case.
A special case of the hp-finite element method (short: hp-FEM) is the boundary concentrated finite element method (short: BC-FEM) which was introduced by Melenk and Khoromskij in [19]. An application of this method to optimal control problems was investigated in our earlier work [7], where only boundary observation of the state was allowed, i.e. the functional $J(y, u) = \frac{1}{2}\|y - y_d\|^2_{L_2(\Gamma_{\mathcal{N}})} + \frac{\alpha}{2}\|u\|^2_{L_2(\Gamma_{\mathcal{N}})}$ was considered. In the present article, we extend these results to domain observation. This extension poses problems insofar, as the term $y - y_d$ appears as a source term in the adjoint equation. While the boundary data may be non-smooth in BC-FEM, the source terms acting on Ω have to satisfy bounds on each derivative, which is proven to be true for this setting. Moreover, we improve approximation results in the $L_2(\Gamma)$-norm, which enhances former results in [7].
For BC-FEM, the number of unknowns N behaves like $h^{-(d-1)}$, where h denotes the mesh size on the boundary and d the dimension of the space. The classical h-FEM typically has h^{-d} unknowns. This implies - for a $H^2(\Omega)$ regular elliptic PDE and $d = 2$ - a reduction of the discretization error in $L_2(\Gamma_{\mathcal{N}})$ by N^{-1} if BC-FEM is applied, whereas the application of the standard h-FEM only leads to a reduction of the discretization error by $N^{-3/4}$ (see also Remark 3.13). Apel, Pfefferer, Rösch showed in [3] that for suitably graded meshes, one can obtain

$$\|u^* - u_h^*\|_{L_2(\Gamma_{\mathcal{N}})} \leq Ch^2|\ln h|^{3/2}$$

even for non-convex domains. This puts graded h-FEM into more competitive position, additionally because less regularity on the source terms (only $L_2(\Omega)$) is needed. Their result is, however, restricted to the case of the Laplacian where the exact structure of the singularities at the vertices of Ω is known, while our results remain valid for general elliptic operators.

The paper is organized as follows: In chapter 2 we rigorously define the model problem and necessary function spaces. We also make assumptions to guarantee $H^{1+\delta}(\Omega)$ regularity with respect to the PDE constraint. As we will discretize the problem with hp finite elements, we introduce so-called countably normed spaces that capture the blowup of derivatives near the boundary $\partial\Omega$. Existence and uniqueness of a solution to the control problem follows by classical arguments.
In an appendix, we describe the construction of an interpolation operator for boundary concentrated meshes with hanging nodes. This extends previously known results [19]. It allows to use meshes composed of quadrilaterals, and thus this construction has importance in its own right.
In chapter 3 we describe the discretization of the model problem by boundary concentrated finite elements. We establish a sharp error estimate for the optimal state y^* and its approximation y_h^*, i.e. $\|y^* - y_h^*\|_{L_2(\Gamma_\mathcal{N})} \leq Ch^{\delta+1/2}$ and derive estimates on the control. This is mainly achieved through the regularity properties of the primal and adjoint equation and the approximation properties on a boundary concentrated mesh. We also improve the approximation result in [7] from $s = \delta$ to $s = \delta + 1/2$. Finally, the problem setting for piecewise analytic data and subdomain observation is investigated. In order to obtain the same approximation properties, it is necessary to guarantee the same regularity in countably normed spaces. Adapting the weighting function for those spaces, which basically means applying the BC-FEM locally on domains of analyticity and the observation subdomain, yields the same error estimate as for the standard setting, namely $s = \delta$ in (1.3).
In chapter 4 we report on the order of convergence observed by solving test problems and compare them to uniform mesh refinement.

2 Model problem and optimality conditions

2.1 Standing assumptions, notation, and function spaces

Let $\Omega \subset \mathbb{R}^2$ be an open and bounded domain with polygonal boundary $\partial\Omega = \Gamma$. The boundary Γ consists of two parts $\Gamma_\mathcal{D}$ and $\Gamma_\mathcal{N}$ with $\overline{\Gamma_\mathcal{D}} \cup \overline{\Gamma_\mathcal{N}} = \Gamma$ and $\Gamma_\mathcal{D} \cap \Gamma_\mathcal{N} = \emptyset$.
On $\Gamma_\mathcal{D}$ Dirichlet boundary conditions are prescribed, whereas the control acts through Neumann boundary conditions on $\Gamma_\mathcal{N}$. The open edges of the polygonal boundary Γ are denoted by Γ^j for $j = 1, \ldots, J$ such that

$$\Gamma = \bigcup_{j=1}^{J} \overline{\Gamma^j}, \quad \Gamma^i \cap \Gamma^j = \emptyset \quad \forall i \neq j.$$

Furthermore, it is assumed that the type of boundary condition does not change on one edge, that means

$$\Gamma^j \cap \Gamma_\mathcal{D} \neq \emptyset \qquad \Leftrightarrow \qquad \Gamma^j \cap \Gamma_\mathcal{N} = \emptyset.$$

The space of square integrable functions v with finite norm

$$\|v\|_{L_2(\Omega)} = \left(\int_\Omega |v(x)|^2 \,\mathrm{d}x \right)^{1/2}$$

is denoted by $L_2(\Omega)$. The space $H^s(\Omega)$ is defined for fractional $s > 0$ with $s =: \lfloor s \rfloor + \sigma =: m + \sigma$ as the Sobolev-Slobodeckij space of functions

$$\{ v \in L_2(\Omega) \mid D^\alpha v \in L_2(\Omega) \quad \forall |\alpha| \leq m \}$$

with finite norm

$$\|v\|_{H^s(\Omega)} := \left(\int_\Omega \sum_{|\alpha| \leq m} |D^\alpha v|^2 \,\mathrm{d}x + \int_\Omega \int_\Omega \sum_{|\alpha| = m} \frac{|D^\alpha v(x) - D^\alpha v(y)|^2}{|x - y|^{2+2\sigma}} \,\mathrm{d}x \,\mathrm{d}y \right)^{1/2}.$$

Using local coordinate systems and a finite number of open sets $\mathcal{O}_i$ $i = 1, \ldots, m$ that cover Γ, the norm H^s on Γ is defined. For details we refer to [25]. It is well known that there is a bounded trace operator

$$T : H^s(\Omega) \to H^{s-1/2}(\Gamma), \quad v \mapsto v|_\Gamma$$

if $s \leq 1$ and $s - 1/2 > 0$ is not integer, see [1, 15, 28].
The constraint (1.1) is referred to as primal equation and is understood in the weak sense with respect to the function space

$$H^1_{\Gamma_\mathcal{D}}(\Omega) := \{v \in H^1(\Omega) : v|_{\Gamma_\mathcal{D}} = 0\}.$$

We denote by ∂_{n_D} the co-normal derivative $D\nabla \cdot n$ with respect to the matrix valued differential operator D. The set of admissible controls U_{ad} whose elements satisfy (1.2) reads

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

To ensure unique solvability and smoothness the following assumptions are made, which are used throughout the paper.

Assumption 1. *Let the the data D, c be analytic in $\bar{\Omega}$ and satisfy*

$$\|\nabla^p D\|_{L_\infty(\Omega)} + \|\nabla^p c\|_{L_\infty(\Omega)} \leq C_d \gamma_d^p p! \qquad \forall p \in \mathbb{N}_0$$

for $C_d, \gamma_d > 0$. Furthermore, let $D(x)$ be symmetric and positive definite in $\overline{\Omega}$, i.e. there exists a $D_0 > 0$ such that for all $x \in \overline{\Omega}$ we have $\xi^\top D(x)\xi > D_0|\xi|^2$ for arbitrary $\xi \in \mathbb{R}^2$. Let $c(x) \geq 0$ for all $x \in \bar{\Omega}$ and $c(x) \geq c_0 > 0$ if $meas(\Gamma_\mathcal{D}) = 0$ In addition, it holds $\alpha > 0$, u_a, $u_b \in H^{1/2}(\Gamma_\mathcal{N})$ with $u_a \leq u_b$ a.e. on $\Gamma_\mathcal{N}$, and $f, y_d \in L_2(\Omega)$.

Let us note that the analyticity assumption will be of importance in chapter 3. Existence and uniqueness of solutions to (1.1) and (P) can be proven under much weaker conditions on the data.

Remark 2.1. *Since analytic functions on Ω belong to the space $L_2(\Omega)$ [26], the functional J in* (P) *is well-defined.*

2.2 Elliptic equation

Due to Assumption 1 above, the Lemma of Lax-Milgram [21, page 92] yields the classical existence and uniqueness result for the weak solution to (1.1).

Theorem 2.2. *For each $f \in L_2(\Omega)$ and $u \in L_2(\Gamma_\mathcal{N})$ there exists a uniquely determined solution $y \in H^1_{\Gamma_\mathcal{D}}(\Omega)$ of* (1.1) *satisfying*

$$\|y\|_{H^1(\Omega)} \leq C(\|f\|_{L_2(\Omega)} + \|u\|_{L_2(\Gamma_\mathcal{N})})$$

with some $C > 0$ independent of u and f.

In order to derive discretization error estimates, we need additional regularity of solutions to (1.1). We summarize related requirements in the following standing assumption.

Assumption 2. *There is a constant $C > 0$ such that for $f \in L_2(\Omega)$ and $u \in L_2(\Gamma_\mathcal{N})$ the solution y to* (1.1) *is in $H^{3/2}(\Omega)$ and satisfies*

$$\|y\|_{H^{3/2}(\Omega)} \leq C(\|f\|_{L_2(\Omega)} + \|u\|_{L_2(\Gamma_\mathcal{N})}).$$

Additionally, there exists $\delta \in [1/2, 1]$ such that for $f \in L_2(\Omega)$ and $u \in H^{1/2}(\Gamma_\mathcal{N})$ the solution y to (1.1) *is in $H^{1+\delta}(\Omega)$ and satisfies*

$$\|y\|_{H^{1+\delta}(\Omega)} \leq C(\|f\|_{L_2(\Omega)} + \|u\|_{H^{1/2}(\Gamma_\mathcal{N})}).$$

Let us comment on the fulfillment of Assumption 2 in the case of the Laplace operator ($D(x) \equiv I$). The regularity $y \in H^{3/2}(\Omega)$ in the case of pure Neumann boundary conditions ($\Gamma_{\mathcal{D}} = \emptyset$) is due to [17]. If Ω is convex, and if the opening angle ω of corners, where the boundary condition changes, satisfies $\omega \leq \frac{\pi}{2}$, then the solution y to (1.1) belongs to $H^2(\Omega)$, see [15]. Hence, Assumption 2 is satisfied with $\delta = 1$.
If Ω is a polygonal domain, it is well known, that solutions can be expanded in a regular part belonging to $H^2(\Omega)$ and singular functions with lower regularity. If no assumptions on the opening angle ω at each corner are made, the solution still is in $H^{5/4}(\Omega)$. Stipulating $\omega < \pi$ at corners with changing boundary conditions results in $y \in H^{3/2}(\Omega)$, see [20]. Again, Assumption 2 is satisfied with $\delta = 1/2$.

2.3 Regularity in countably normed spaces

Let $r(x) := \operatorname{dist}(x, \Gamma)$ denote the distance of $x \in \Omega$ to the boundary. Let us introduce for $\beta \geq 0$ the weighted norm

$$\|u\|^2_{H^2_\beta(\Omega)} := \|u\|^2_{H^1(\Omega)} + \|r^\beta \nabla^2 u\|^2_{L_2(\Omega)}.$$

The space $H^2_\beta(\Omega)$ is then defined as the closure of $C^\infty(\bar{\Omega})$ with respect to this weighted norm $\|.\|_{H^2_\beta(\Omega)}$. In addition, we introduce the countably normed space $B^2_\beta(C, \gamma)$ with $C, \gamma > 0$, see [4, 5], by

$$B^2_\beta(C,\gamma) := \{v \in H^2_\beta(\Omega) \mid \|v\|_{H^2_\beta(\Omega)} \leq C,\ \|r^{p+\beta}\nabla^{p+2} v\|_{L_2(\Omega)} \leq C\gamma^p p! \quad \forall p \in \mathbb{N}\}.$$

This space is suitable for capturing the regularity properties of an elliptic equation with analytical data. The source terms in the elliptic equation are later assumed to be in the set

$$B^0_\beta(C,\gamma) := \{v \in L_2(\Omega) \mid \|v\|_{L_2(\Omega)} \leq C,\ \|r^{p+\beta}\nabla^{p} v\|_{L_2(\Omega)} \leq C\gamma^p p! \quad \forall p \in \mathbb{N}\}.$$

We recall a result from [19]:

Theorem 2.3. *Let Assumptions 1 and 2 hold. Let $f \in B^0_{1-\delta}(C_f, \gamma_f)$ be given with $C_f, \gamma_f > 0$. If $y \in H^{1+\delta}(\Omega)$ is a solution of the differential equation*

$$-\nabla \cdot (D(x)\nabla y) + c(x) y = f(x) \qquad \text{in } \Omega,$$

then there exist constants $C, \gamma > 0$ that depend only on $\Omega, C_d, C_f, \gamma_d, \gamma_f$ and δ such that

$$\|r^{p+1-\delta}\nabla^{p+2} y\|_{L_2(\Omega)} \leq C\gamma^p p! \left(C_f + \|y\|_{H^{1+\delta}(\Omega)}\right) \qquad \forall p \in \mathbb{N}_0,$$

which implies $y \in B^2_{1-\delta}\left(C(C_f + \|y\|_{H^{1+\delta}(\Omega)}),\ \gamma\right)$.

Proof. By closely inspecting the technical proof of [19, Theorem A.1], which builds on [23], one can see that the assumptions on f are sufficient for obtaining the theorem. □

2.4 Existence of solution and optimality conditions

Since the primal equation (1.1) is uniquely solvable for each $u \in U_{ad}$ with affine-linear and continuous mapping $u \mapsto y$, the problem (P) constitutes a convex optimization problem. Hence, existence and uniqueness follow by classical arguments.

Theorem 2.4. *The optimal control problem* (P) *admits a unique optimal control u^* with associated optimal state y^*.*

Proof. For the proof we refer to [27]. □

The solution (y^*, u^*) is uniquely characterized by the following first-order necessary and sufficient optimality conditions (see [27]).

Theorem 2.5. *The pair $(y^*, u^*) \in H^1_{\Gamma_\mathcal{D}}(\Omega) \times U_{ad}$ is a solution to Problem* (P) *if and only if there exists $q^* \in H^1_{\Gamma_\mathcal{D}}(\Omega)$ such that the state equation* (1.1)*, the adjoint equation*

$$\begin{aligned} -\nabla \cdot (D(x)\nabla q^*(x)) + c(x)q^*(x) &= y^*(x) - y_d(x) && \text{in } \Omega, \\ q^*(x) &= 0 && \text{on } \Gamma_\mathcal{D}, \\ \partial_{n_D} q^*(x) &= 0 && \text{on } \Gamma_\mathcal{N}, \end{aligned} \tag{2.1}$$

and the variational inequality

$$\langle \alpha u^* + q^*, u - u^* \rangle_{L_2(\Gamma_\mathcal{N})} \geq 0 \qquad \forall u \in U_{ad} \tag{2.2}$$

are satisfied. Here $\langle \cdot, \cdot \rangle_{L_2(\Gamma_\mathcal{N})}$ denotes the inner product over the given space.

The variational inequality (2.2) is equivalent to

$$u^*(x) = P_{U_{ad}} \left(-\frac{1}{\alpha} q^*|_{\Gamma_\mathcal{N}}(x) \right) \quad \text{a.e. on } \Gamma_\mathcal{N} \tag{2.3}$$

where $P_{U_{ad}}$ denotes the L_2-projection onto the convex set U_{ad}, cf. [27].
The projection representation (2.3) implies that the optimal control inherits regularity from the trace of the adjoint state. This allows to conclude higher regularity of the solution of (P).

Theorem 2.6. *Let $f, y_d \in L_2(\Omega)$ be given. Let (y^*, u^*, q^*) satisfy the necessary optimality conditions of Theorem 2.5. Then $(y^*, u^*, q^*) \in H^{1+\delta}(\Omega) \times H^{1/2}(\Gamma_\mathcal{N}) \times H^{1+\delta}(\Omega)$.*

Proof. The projection representation (2.3) implies the regularity $u^* \in H^{1/2}(\Gamma_\mathcal{N})$, which is proven in [3, eq. (4.10)]. By assumption 2 we get the regularity $y^* \in H^{1+\delta}(\Omega)$ for the optimal state. As the right-hand side of the adjoint problem (2.1) is in $L_2(\Omega)$, we obtain $q^* \in H^{1+\delta}(\Omega)$ as well. □

3 Discretization

Since it is generally not possible to solve problem (P) analytically, it is discretized with finite elements. For the state and adjoint variable the boundary concentrated finite element method is used, which has been introduced by Khoromskij and Melenk [19]. The control is treated according to the concept of variational discretization from Hinze [16]. The boundary concentrated finite element method - an *hp*- finite element method - is described in the following.

3.1 Boundary concentrated finite elements

Let us consider an affine triangulation τ of Ω, i.e. each element $K \in \tau$ is the image of the reference square/triangle $\hat{K}$ under an affine mapping F_K. Moreover, $\bar{\Omega} = \bigcup_{K \in \tau} \bar{K}$. The triangulation may be irregular, i.e. we allow hanging nodes. These nodes have to lie in the middle of a coarse edge. Such a discretization belongs to the category of 1-irregular meshes (see [6]). We also stipulate that elements with hanging nodes have positive distance from the boundary $\partial\Omega$. This simplifies the construction of an interpolation operator.

Definition 3.1 (**γ-shape-regular**)**.** *An affine triangulation τ of Ω where h_K denotes the diameter of the element K and the mapping F_K satisfies the inequality*

$$h_K^{-1} \|F'_K\|_{L_\infty(K)} + h_K \|(F'_K)^{-1}\|_{L_\infty(K)} \leq \gamma \qquad \forall\, K \in \tau$$

is called γ-shape-regular. Here, F'_K denotes the Jacobian of F_K.

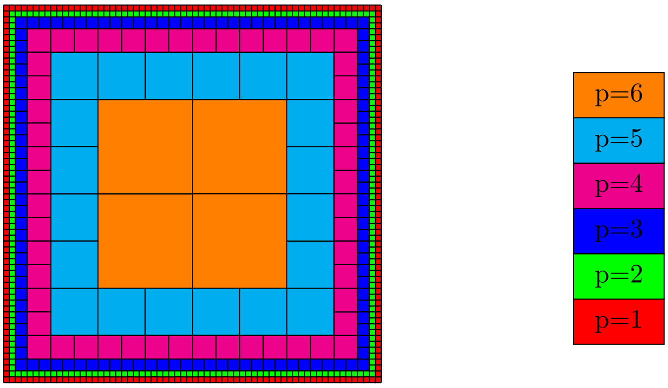

Figure 1: Boundary concentrated mesh with hanging nodes where only the elements on the boundary have the lowest polynomial degree.

The boundary concentrated finite element method uses so-called geometric meshes (see Figure 3.1), which are defined in the following.

Definition 3.2. (geometric mesh) *Let τ be a γ-shape-regular triangulation as defined above and $h := \min_{\bar{K}\cap\Gamma\neq\emptyset}\{h_K\}$ be a measure for the mesh-size at the boundary. τ is called a geometric mesh, if there exist constants $c_1, c_2 > 0$ such that for all $K \in \tau$:*

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 $c_1 \inf_{x\in K} \operatorname{dist}(x, \Gamma) \leq h_K \leq c_2 \sup_{x\in K} \operatorname{dist}(x, \Gamma)$.*

Here, condition 1 implies that elements at the boundary are a quasiuniform mesh of size h.
The hp-FEM allows different polynomial degrees on elements K and edges e. The polynomial degrees $p_K \in \mathbb{N}$ on elements $K \in \tau$ are collected in the polynomial degree vector $\mathbf{p} := (p_K)_{K\in\tau}$. The polynomial degree p_e on the edges is defined by

$$p_e := \min\{p_K | e \text{ is an edge of element } K\}, \tag{3.1}$$

i.e. the polynomial degree of the edge is given by the minimum of the polynomial degrees of the neighboring elements. Furthermore one uses the vector

$$\mathbf{p}(K) := (p_{e_1}, p_{e_2}, p_{e_3}, (p_{e_4},)p_K) \tag{3.2}$$

which contains the polynomial distribution for a triangle/square $K \in \tau$. The definition of a linear degree vector, which rules the polynomial degree distribution for geometric meshes, shall be given next.

Definition 3.3 (linear degree vector). *Let τ be a geometric mesh on Ω with boundary mesh size h. A polynomial degree vector $\mathbf{p} = (p_K)_{K\in\tau}$ is said to be a linear degree vector with slope $\alpha > 0$ if there exist constants c_1, $c_2 > 0$ such that*

$$1 + \alpha c_1 \log \frac{h_K}{h} \leq p_K \leq 1 + \alpha c_2 \log \frac{h_K}{h}\,. \tag{3.3}$$

For a typical polynomial degree distribution see Figure 3.1.
Now, we have everything at hand to define the hp-FEM spaces:

Definition 3.4. *Let τ_l be a geometric mesh on Ω with boundary mesh size h_l 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 $K \in \tau_l$ let $\mathbf{p}(K)$ be given by (3.2). Then we set*

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

where

$$\mathcal{P}_{\mathbf{p}(K)}(\hat{K}) := \{u \in \mathcal{P}_{p_K}(\hat{K}) \;:\; u|_{\hat{e}_i} \in \mathcal{P}_{p_{e_i}}, i = 1, 2, 3(, 4)\},$$

with $\hat{e}_1$, $\hat{e}_2$, $\hat{e}_3$ $(,\hat{e}_4)$ being the corresponding edges of the reference triangle/square $\hat{K}$.

The following lemma gives a quantitative estimate on the number of degrees of freedom.

Lemma 3.5. *[19, Proposition 2.7] Let τ 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 the polynomial degree p_e on the edge fulfills* (3.1) *and for all $K \in \tau$, $\mathbf{p}(K)$ is given by* (3.2). *Then it holds*

$$dim(\mathbb{V}_{l,\Gamma_{\mathcal{D}}}^{\mathbf{p}}) \sim h^{-1}.$$

Please note that classical h-FEM with uniform refined meshes yields number of degrees of freedom of the order of h^{-2} for two-dimensional domains.

3.1.1 Variational formulation

Let us now consider the primal problem (1.1). Its weak formulation reads

$$\text{find } y \in H^1_{\Gamma_{\mathcal{D}}}(\Omega): \qquad a(y, v) = \langle f, v\rangle_{L_2(\Omega)} + \langle u, v\rangle_{L_2(\Gamma_{\mathcal{N}})} \quad \forall v \in H^1_{\Gamma_{\mathcal{D}}}(\Omega) \tag{3.4}$$

with the bilinear-form

$$a(y, v) = \int_\Omega D(x)\nabla y(x) \cdot \nabla v(x)\,\mathrm{d}x + \int_\Omega c(x)y(x)v(x)\,\mathrm{d}x.$$

Replacing the space $H^1_{\Gamma_{\mathcal{D}}}(\Omega)$ by its discrete analogon $\mathbb{V}_{l,\Gamma_{\mathcal{D}}}^{\mathbf{p}}$ we obtain the discrete state equation:

$$\text{find } y_h \in \mathbb{V}_{l,\Gamma_{\mathcal{D}}}^{\mathbf{p}}: \qquad a(y_h, v_h) = \langle f, v_h\rangle_{L_2(\Omega)} + \langle u, v_h\rangle_{L_2(\Gamma_{\mathcal{N}})} \quad \forall v \in \mathbb{V}_{l,\Gamma_{\mathcal{D}}}^{\mathbf{p}}. \tag{3.5}$$

The variational formulation of the dual equation and its discretization is defined analogously. For the accuracy of the discretization we have the following H^1-error estimate.

Lemma 3.6. *Let τ be a geometric mesh on Ω with mesh size h, $\mathbf{p}$ a linear degree vector with slope α. Suppose Assumptions 1 and 2 are satisfied. Let $y \in H^{1+\delta}(\Omega)$ for some $\delta \in (0, 1]$ be a solution to the state equation* (3.4) *with data $u \in L_2(\Gamma_{\mathcal{N}})$ and $f \in B^0_{1-\delta}(C_f, \gamma_f)$, $C_f, \gamma_f > 0$. Then for sufficiently large α there is $C > 0$ independent of h such that*

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

holds.

The proof is given in the appendix, see Theorem A.12.

3.2 Error estimates on the boundary

In this section, we prove optimal error estimates for the error $\|q^* - q_h^*\|_{L_2(\Gamma_\mathcal{N})}$ when using the BC-FEM discretization. Such estimates are not available in the literature to the best of our knowledge. We will employ them below in Lemma 3.11. Let us mentions that these improved estimates will lead in addition to better overall estimates for control problems with boundary observation, see subsection 3.4 below, and thus improve our previous results [7].

Theorem 3.7. *Let τ be a geometric mesh on Ω with mesh size h, $\mathbf{p}$ a linear degree vector with slope α. Let Assumptions 1, 2 be satisfied and $y \in H^{1+\delta}(\Omega)$ be the weak solution of*

$$\begin{aligned} -\nabla \cdot (D(x)\nabla y(x)) + c(x)y(x) &= f(x) && \text{in } \Omega, \\ y(x) &= 0 && \text{on } \Gamma_\mathcal{D}, \\ \partial_{n_D} y(x) &= u(x) && \text{on } \Gamma_\mathcal{N}, \end{aligned}$$

with $u \in H^{1/2}(\Gamma_\mathcal{N})$ and $f \in B^0_{1-\delta}(C_f, \gamma_f)$ with $C_f, \gamma_f > 0$. Moreover, let $y_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$ be the solution of the corresponding discretized problem. If α is sufficiently large then there is a constant $C > 0$ independent of h and u such that it holds

$$\|y - y_h\|_{L_2(\Gamma_\mathcal{N})} \le C h^{\delta + \frac{1}{2}} \left(C_f + \|y\|_{H^{1+\delta}(\Omega)} \right).$$

Proof. We proof this result by the standard Nitsche trick. Let z denote the solution of the dual problem

$$\begin{aligned} -\nabla \cdot (D(x)\nabla z(x)) + c(x)z(x) &= 0 && \text{in } \Omega, \\ z(x) &= 0 && \text{on } \Gamma_\mathcal{D}, \\ \partial_{n_D}(x) &= y(x) - y_h(x) && \text{on } \Gamma_\mathcal{N}, \end{aligned}$$

with $z_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$ being its BC-FEM approximation.
Then it holds by Galerkin orthogonality

$$\|y - y_h\|^2_{L_2(\Gamma_\mathcal{N})} = a(z, y - y_h) = a(z - z_h,\, y - y_h) = a(z - z_h,\, y - I_h y),$$

where I_h is the BC-FEM interpolation operator from Theorem A.10. According to Theorem 2.3 the solution y satisfies

$$\|r^{p+1-\delta}\nabla^{p+2} y\|_{L_2(\Omega)} \le C_y \gamma_y^p p! \left(C_f + \|y\|_{H^{1+\delta}(\Omega)} \right) \qquad \forall p \in \mathbb{N}_0$$

with $C_y, \gamma_y > 0$ independent of u. By Theorem A.10 we obtain the following interpolation error estimate

$$\|y - I_h y\|_{H^1(\Omega)} \le C\, C_y \left(C_f + \|y\|_{H^{1+\delta}(\Omega)} \right) h^\delta$$

for sufficiently large α. The solution z of the dual problem satisfies

$$\|r^{p+1-\delta}\nabla^{p+2} z\|_{L_2(\Omega)} \le C_z \gamma_z^p p! \|z\|_{H^{1+\delta}(\Omega)} \qquad \forall p \in \mathbb{N}_0$$

with $\delta = \frac{1}{2}$ and $C_z, \gamma_z > 0$ independent of $y - y_h$, cf. Theorem 2.3.
With the same arguments as above and applying Cea's Lemma as well as Theorem A.10 we conclude

$$\|z - z_h\|_{H^1(\Omega)} \le C \|z - I_h z\|_{H^1(\Omega)} \le C\, C_z \|z\|_{H^{3/2}(\Omega)} h^{1/2}$$

with $C_z > 0$ independent of $y - y_h$ and sufficiently large α, where the choice of α is independent of $y - y_h$ and thus independent of the discretization.

Using Assumption 2 to estimate $\|z\|_{H^{3/2}(\Omega)}$ we get

$$\|z - z_h\|_{H^1(\Omega)} \leq C\,C_z \|z\|_{H^{3/2}(\Omega)} h^{\frac{1}{2}} \leq C\,C_z\, h^{\frac{1}{2}} \|y - y_h\|_{L_2(\Gamma_\mathcal{N})}.$$

Hence, we obtain

$$\begin{aligned}\|y - y_h\|^2_{L_2(\Gamma_\mathcal{N})} &\leq C\|z - z_h\|_{H^1(\Omega)} \|y - I_h y\|_{H^1(\Omega)} \\ &\leq C h^{\delta + \frac{1}{2}} \left(C_f + \|y\|_{H^{1+\delta}(\Omega)}\right) \|y - y_h\|_{L_2(\Gamma_\mathcal{N})},\end{aligned}$$

which ends the proof. □

Let us show numerically that the estimate obtained in Theorem 3.7 is sharp. To this end, we consider the following example. $y^*(r,\phi) := r^{3/2}\cos(3/2\phi)$, where (r,ϕ) denote the two dimensional polar coordinates. This harmonic function is the unique solution to

$$\begin{aligned} -\Delta y + y &= y^* && \text{in } \Omega \\ \partial_n y &= \partial_n y^* && \text{on } \Gamma_\mathcal{N} := \partial\Omega \end{aligned}$$

on the L-shaped domain $\Omega := (-1,1)^2 \setminus (0,1) \times (0,-1)$. The maximal regularity of y^* is $H^{5/3-\epsilon}(\Omega)$. Theorem 3.7 yields a convergence rate $\mathcal{O}(h^{2/3+1/2-\epsilon}) = \mathcal{O}(h^{7/6-\epsilon})$ for the L_2-error on the boundary. We compute the experimental error of convergence EOC step by step, when refining a mesh of size h_1 to h_2. The EOC for the variable y and norm $L_2(\Gamma_\mathcal{N})$ is defined as

$$\mathrm{EOC}(y, L_2(\Gamma_\mathcal{N})) := \frac{\ln\|y^* - y^*_{h_1}\|_{L_2(\Gamma_\mathcal{N})} - \ln\|y^* - y^*_{h_2}\|_{L_2(\Gamma_\mathcal{N})}}{\ln(h_1) - \ln(h_2)} \tag{3.6}$$

The expected convergence rate is confirmed by the numerical test, and can be observed in Table 1.

h	$\|y^* - y^*_h\|_{L_2(\Gamma_2)}$	$\mathrm{EOC}(y, L_2(\Gamma_2))$
1	$7.93 \cdot 10^{-2}$	-
0.5	$3.59 \cdot 10^{-2}$	1.14
0.25	$1.64 \cdot 10^{-2}$	1.13
0.125	$7.31 \cdot 10^{-3}$	1.16
0.0625	$3.28 \cdot 10^{-3}$	1.16
0.0312	$1.47 \cdot 10^{-3}$	1.16
0.0156	$6.57 \cdot 10^{-4}$	1.16

Table 1: Convergence rates for irregular example

3.3 Error estimates for the optimal control problem

The discretization of (P) now reads: Minimize

$$J(y_h, u_h) \text{ subject to (3.5) and } u_h \in U_{ad}. \tag{P$_h$}$$

The discretized problem is of similar structure as problem (P). Standard techniques as in [27] yield a unique solution (y^*_h, u^*_h), which is characterized by the following optimality conditions.

Theorem 3.8. *A pair $(y^*_h, u^*_h) \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}} \times U_{ad}$ is a solution of* (P$_h$) *if and only if it fulfills* (3.5) *and there exists $q_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$ satisfying*

$$a(q^*_h, v_h) = \langle y^*_h - y_d, v_h\rangle_{L_2(\Omega)} \quad \forall v_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$$

and

$$u_h^*(x) = P_{U_{ad}}\left(-\frac{1}{\alpha} q_h^*|_{\Gamma_\mathcal{N}}(x)\right) \quad \textit{a.e. on } \Gamma_\mathcal{N}.$$

Proof. For the proof we refer to [16, Section 3]. □

Remark 3.9. *Since the pointwise projection of a grid function on an admissible interval* $[u_a, u_b]$ *is not a finite element function in general,* u_h *is not a finite element function. Therefore* u_h^* *may have kinks that are not along the mesh edges. If* u_a *and* u_b *can be represented as finite element functions (for example if* u_a *and* u_b *are constant), then* u_h^* *can be represented by a finite number of parameters.*

In order to establish an estimator of the discretization error, let us introduce two auxiliary functions $y^h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$ and $q^h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$ as unique solutions of the discrete boundary value problem

$$a(y^h, v_h) = \langle f, v_h\rangle_{L_2(\Omega)} + \langle u^*, v_h\rangle_{L_2(\Gamma_\mathcal{N})} \qquad \forall v_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}, \tag{3.7}$$

$$a(q^h, v_h) = \langle y^* - y_d, v_h\rangle_{L_2(\Omega)} \qquad \forall v_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}. \tag{3.8}$$

These functions are discrete approximations of the optimal state y^* and adjoint state q^*, respectively.

Theorem 3.10. *[16, Theorem 3.1] Let* (u^*, y^*, q^*) *and* (u_h^*, y_h^*, q_h^*) *be the solutions to the optimal control problem* (P) *and* (P_h)*, respectively. Let* y^h *and* q^h *be the solutions to* (3.7) *and* (3.8)*. Then the following inequality*

$$\alpha\|u^* - u_h^*\|^2_{L_2(\Gamma_\mathcal{N})} + \|y^* - y_h^*\|^2_{L_2(\Omega)} \leq \frac{1}{\alpha}\|q^* - q^h\|^2_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^h\|^2_{L_2(\Omega)} \tag{3.9}$$

is satisfied.

This shows that the discretization error depends on the smoothness of the optimal state and adjoint state. In what follows we will now analyze the two error contributions $\|y^* - y^h\|^2_{L_2(\Omega)}$ and $\|q^* - q^h\|^2_{L_2(\Gamma_\mathcal{N})}$.

Lemma 3.11. *Let Assumption 1 and 2 hold. Assume* $f, y_d \in B^0_{1-\delta}(C_f, \gamma_f)$ *with* $C_f, \gamma_f > 0$*. Let* τ *be a geometric mesh on* Ω *with mesh size* h*,* $\mathbf{p}$ *a linear degree vector with sufficiently large slope* α*. Then there is* $C > 0$ *independent of* h *such that*

$$\|q^* - q^h\|_{L_2(\Gamma)} \leq C\ h^{\delta + \frac{1}{2}}$$

holds.

Proof. First, we will prove that $q^* \in B^2_{1-\delta}(C, \gamma_q)$ for some $C_q, \gamma_q > 0$. In order to achieve this regularity, we need to prove that the source term $y - y_d$ fulfills the assumptions of Theorem 2.3, i.e. $y - y_d \in B^0_{1-\delta}(C, \gamma)$ for some $C, \gamma > 0$. By Theorem 2.3 we already have $y^* \in B^2_{1-\delta}(C_y, \gamma_y)$ with $C_y, \gamma_y > 0$, which implies

$$\|r^{p+1-\delta}\nabla^{p+2} y\|_{L_2(\Omega)} \leq C_y \gamma_y^p p! \|y\|_{H^{1+\delta}(\Omega)} \qquad \forall p \in \mathbb{N}_0. \tag{3.10}$$

Due to the definition of r it holds $0 \leq r \leq \mathrm{diam}(\Omega)$. Thus it holds for $p \in \mathbb{N}_0$

$$\begin{aligned}\|r^{p+3-\delta}\nabla^{p+2} y\|_{L_2(\Omega)} &\leq \mathrm{diam}(\Omega)^2 \|r^{p+1-\delta}\nabla^{p+2} y\|_{L_2(\Omega)} \\ &\leq \mathrm{diam}(\Omega)^2\, C_y \gamma_y^p p! \|y\|_{H^{1+\delta}(\Omega)}, \\ &\leq \mathrm{diam}(\Omega)^2 \gamma_y^{-2}\, C_y \gamma_y^{p+2} (p+2)! \|y\|_{H^{1+\delta}(\Omega)},\end{aligned}$$

which proves

$$\|r^{p+1-\delta}\nabla^p y\|_{L_2(\Omega)} \leq \operatorname{diam}(\Omega)^2\gamma_y^{-2}\, C_y\gamma_y^p p!\|y\|_{H^{1+\delta}(\Omega)}$$

for all $p > 1$. The cases $p = 0$ and $p = 1$ follow from

$$\|r^{1-\delta}y\|_{L_2(\Omega)} \leq \operatorname{diam}(\Omega)^{1-\delta}\|y\|_{H^1(\Omega)} \leq \operatorname{diam}(\Omega)^{1-\delta}\|y\|_{H^{1+\delta}(\Omega)},$$
$$\|r^{2-\delta}\nabla y\|_{L_2(\Omega)} \leq \operatorname{diam}(\Omega)^{2-\delta}\|y\|_{H^1(\Omega)} \leq \operatorname{diam}(\Omega)^{2-\delta}\|y\|_{H^{1+\delta}(\Omega)}.$$

Hence $y - y_d \in B^0_{1-\delta}(C,\gamma)$ for some $C, \gamma > 0$, and Theorem 2.3 can be used to conclude $q^* \in B^2_{1-\delta}(C_q,\gamma_q)$ with $C_q, \gamma_q > 0$. Theorem 3.7 yields the claim. □

Therewith the main approximation result can now be formulated:

Theorem 3.12. *Let Assumptions 1, 2 hold and $f, y_d \in B^0_{1-\delta}(C_f,\gamma_f)$ with $C_f, \gamma_f > 0$. Let τ be a geometric mesh on Ω with mesh size h, $\mathbf{p}$ a linear degree vector with sufficiently large slope α. Let (u^*, y^*, q^*) and (u^*_h, y^*_h, q^*_h) be the solutions of the optimal control problem* (P) *and its discretized version* (P_h) *with the corresponding states and adjoint states. The solution of the boundary value problem* (1.1) *shall be $H^{1+\delta}$-regular with $\delta \in (0, 1)$, that means $y^*, q^* \in H^{1+\delta}(\Omega)$. Then there exists a constant $C > 0$ independent of h and it holds*

$$\|u^* - u^*_h\|_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^*_h\|_{L_2(\Omega)} \leq Ch^\delta.$$

Proof. With Theorem 3.10 it holds

$$\alpha\|u^* - u^*_h\|^2_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^*_h\|^2_{L_2(\Omega)} \leq \frac{1}{\alpha}\|q^* - q^h\|^2_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^h\|^2_{L_2(\Omega)}.$$

Exploiting the approximation properties of Lemma 3.6 and Lemma 3.11 with a sufficiently small mesh size ($h < 1$) yields

$$\begin{aligned}\alpha\|u^* - u^*_h\|^2_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^*_h\|^2_{L_2(\Omega)} &\leq \left(\frac{1}{\alpha}h + 1\right)\left(Ch^\delta\right)^2\\ &\leq \left(\frac{1}{\alpha} + 1\right)\left(Ch^\delta\right)^2 \leq \tilde{C}h^{2\delta},\end{aligned}$$

which is the desired estimate. □

Remark 3.13. *Let N be the dimension of the finite element space $\mathbb{V}^{\mathbf{p}}_{l,\Gamma_\mathcal{D}}$ and h the mesh size of the boundary Γ. Lemma 3.5 implies $h \sim N^{-1}$, therefore the result of Theorem 3.12 can be read as*

$$\sqrt{\alpha}\|u^* - u^*_h\|_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^*_h\|_{L_2(\Omega)} \leq CN^{-\delta}$$

if the problem is $H^{1+\delta}$-regular. In h-FEM the dimension of the approximation space grows as $N \sim h^{-2}$, which would lead to (combined with an estimate of [8])

$$\sqrt{\alpha}\|u^* - u^*_h\|_{L_2(\Gamma_\mathcal{N})} + \|y^* - y^*_h\|_{L_2(\Omega)} \leq CN^{-\frac{3}{4}\delta}$$

for an $H^{1+\delta}$-regular problem (see [7] for a similar comparison). Hence for an H^2-regular problem the error reduces for BC-FEM as $\mathcal{O}(N^{-1})$ and for h-FEM as $\mathcal{O}(N^{-3/4})$. Therefore - if the dimension N of the approximation space for the boundary value problem (P) *is fixed - the discretization by boundary concentrated finite elements gives a smaller error for a N being large enough. Using graded meshes and the results of [3] would yield competitive results for the Laplacian.*

Remark 3.14. *Our numerical examples with known solution (chapter 4) show an error reduction of $h^{2\delta}$. In order to prove such a convergence rate, one would need optimal estimates of the $L_2(\Omega)$-error, which are not available to the best of our knowledge.*
The main reason is that the Aubin-Nitsche trick does not work for the BC-FEM, since no error estimate of the type $\|y - y_h\|_{H^1(\Omega)} \leq Ch^\delta \|f\|_{L_2(\Omega)}$ is available for solutions of the elliptic partial differential equation (1.1) with right-hand side f and $u = 0$. The best currently available L_2-estimate was proven by Eibner and Melenk in [13]. They show that for every compact $\Omega' \subset\subset \Omega$ there exists $\delta' \in [0, \delta]$ such that for all elements $K \subset\subset \Omega'$ the error estimate $\|y-y_h\|_{L_2(K)} \leq Ch^{\delta+\delta'}$ holds. However, δ' depends on Ω', and it is unclear under which conditions $\delta = \delta'$ can be proven.

3.4 Boundary observation

As already indicated, we can improve the error estimate of [7] with Theorem 3.7 in the case of boundary observation. Consider the minimization of

$$J(u, y) = \frac{1}{2}\|y - y_d\|^2_{L_2(\Gamma_\mathcal{N})} + \frac{\alpha}{2}\|u\|^2_{L_2(\Gamma_\mathcal{N})}$$

subject to the same constraints as our model problem. The proof for the error in the optimal control [7, Theorem 3.8,3.9] is similar to the one given above. Using the new results of Theorem 3.7 one can prove

$$\|u^* - u_h^*\|_{L_2(\Gamma_\mathcal{N})} + \|y^* - y_h^*\|_{L_2(\Gamma_\mathcal{N})} \leq Ch^{\delta+1/2}. \tag{3.11}$$

This enhances the previous results of [7], which only had the convergence rate h^δ.

3.5 Piecewise analytic data and subdomain observation

The model problem (P) tries to drive the state y to a desired state y_d. Different from [7], where the domain of observation is Γ and y_d thus appears in the Neumann boundary condition of the adjoint problem, we deal with $y_d : \Omega \to \mathbb{R}$.

It is possible to generalize the above theory even further, to include piecewise analytic data and/or subdomain observation. To this end, let Ω consist of $S \in \mathbb{N}$ disjoint subsets Ω_i with Lipschitz boundaries $\partial\Omega_i$ such that $\bar{\Omega} = \cup_{i=1}^S \bar{\Omega}_i$. Let $\tilde{r}(x) := \text{dist}(x, \cup_{i=1}^S \partial\Omega_i)$. We stipulate that y_d be analytic on Ω_i for all $i = 1, \dots, S$ and

$$\|\tilde{r}^{p+1-\delta}\nabla^p y_d\|_{L_2(\Omega_i)} \leq C_i \gamma_i^p p! \quad \forall p \in \mathbb{N}_0.$$

Let q be the solution of the adjoint problem (2.1), then the restriction $q_i := q|_{\Omega_i}$ solves

$$\begin{aligned} -\nabla \cdot (D(x)\nabla q_i(x)) + c(x)q_i(x) &= y_i(x) - y_{d,i}(x) && \text{in } \Omega_i, \\ q_i(x) &= 0 && \text{on } \Gamma_\mathcal{D} \cap \partial\Omega_i, \\ \partial_{n_D} q_i(x) &= 0 && \text{on } \Gamma_\mathcal{N} \cap \partial\Omega_i, \\ q_i(x) &= q(x) && \text{on } \partial\Omega_i \backslash (\Gamma_\mathcal{D} \cup \Gamma_\mathcal{N}). \end{aligned}$$

Recall that $y \in B^2_{1-\delta}(\Omega)$ due to Theorem 2.3 and the arguments in the proof of Lemma 3.11. As $\tilde{r}(x) \leq r(x)$ for arbitrary $x \in \Omega$ and $\Omega_i \subset \Omega$, we find $y \in \tilde{B}^2_{1-\delta}(\Omega_i) \cap \tilde{B}^2_{1-\delta}(\Omega)$, where the tilde indicates that $\tilde{r}$ is the weighting function for the derivatives in the countably normed space (see section 2.3). We can apply Theorem 2.3 locally on Ω_i and get $q_i \in \tilde{B}^2_{1-\delta}(\Omega_i)$, as well as $q \in \tilde{B}^2_{1-\delta}(\Omega)$.

By building a geometric mesh on Ω that does h-refinement not only near $\partial\Omega$ but also $\partial\Omega_i$, it is possible to apply Lemma 3.6 for y (and 3.11 for q respectively). This yields again

$$\|y^* - y^h\|_{L_2(\Omega)} \leq Ch^\delta, \quad \|q^* - q^h\|_{L_2(\Gamma_\mathcal{N})} \leq Ch^{\delta+1/2}$$

and the approximation result for the optimal control and state (Theorem 3.12) remain valid.
Observing y only on a subdomain $\Omega_s \subsetneq \Omega$ means minimizing

$$\frac{1}{2}\|\chi_{\Omega_s}(y - y_d)\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Gamma_\mathcal{N})}$$

with $\chi_{\Omega_s} : \Omega \to \{0, 1\}$ being the characteristic function of Ω_s. Then the right hand side of the adjoint problem is given by $\chi_{\Omega_s}(y-y_d)$, which is piecewise analytic, and can be treated as sketched above.
If we deal with both subdomain observation and piecewise analytic data (possibly in f, too), we need to pass to a partition of Ω which is compatible for all domains of analyticity.

4 Numerical examples

Let us now report about the results of our numerical experiments.

4.1 Problem description

The optimal control problem is given by

$$\min J(y, u) := \frac{1}{2}\|y - y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Gamma_\mathcal{N})} + \int_{\Gamma_\mathcal{N}} e_q y \, \mathrm{d}s_x,$$

subject to the constraints

$$\begin{aligned} -\nabla \cdot (D(x)\nabla y(x)) &= f(x) && \text{in } \Omega, \\ \partial_{n_D} y(x) &= u(x) + e_y(x) && \text{on } \Gamma_\mathcal{N}, \\ y(x) &= 0 && \text{on } \Gamma_\mathcal{D}, \end{aligned}$$

where u is the Neumann control and $(y, u) \in H^1(\Omega) \times L_2(\Gamma_\mathcal{N})$. The inhomogeneities $e_y, e_q \in H^{1/2}(\Gamma_\mathcal{N})$ are introduced for the construction of a test example with known solution (y^*, u^*). Let us remark that the error estimates developed above are still valid. The proof of the result of theorem 3.10 is not affected by these affine inhomogeneities. Moreover, the interior regularity of the solutions y^* and q^* is not affected by these boundary data. The set of admissible data for u is given by

$$u_a \leq u \leq u_b. \tag{4.1}$$

The adjoint equation reads

$$\begin{aligned} -\nabla \cdot (D(x)\nabla q(x)) &= y(x) - y_d(x) && \text{in } \Omega, \\ \partial_{n_D} q(x) &= e_q(x) && \text{on } \Gamma_\mathcal{N}, \\ q(x) &= 0 && \text{on } \Gamma_\mathcal{D}. \end{aligned}$$

As already mentioned the state y and the adjoint q are discretized by means of the boundary-concentrated FEM, whereas the control u is implicitly discretized via variational discretization from the projection formula

$$u^* = P_{[u_a,u_b]}\left(-\frac{1}{\alpha}q^*(x)|_{\Gamma_\mathcal{N}}\right),$$

see e.g. Section 3.3.

4.2 Example 1 - Problem with known solution

We set

$$\alpha = 1, \quad u_a \equiv 1, \quad u_b \equiv 6.$$

and $\Omega = (0,1)^2$ with $\Gamma_{\mathcal{D}} = \{x_1 = 0\} \cup \{x_2 = 0\}$ and $\Gamma_{\mathcal{N}} = \Gamma \setminus \Gamma_{\mathcal{D}}$.
The optimal adjoint for $D \equiv 1$ is given by

$$q^*(x) = -x_1 x_2^2 e^{x_1+x_2}.$$

Therefore, the optimal control reads

$$u^*(x) = P_{[u_a,u_b]}\left(x_1 x_2^2 e^{x_1+x_2}|_{\Gamma_{\mathcal{N}}}\right).$$

The Laplacian of the adjoint is

$$-\Delta q^*(x) = \underbrace{(2x_2^2 + 2x_1)e^{x_1+x_2}}_{=:-y_d(x)} + \underbrace{(4x_1x_2 + 2x_1x_2^2)e^{x_1+x_2}}_{=:y^*(x)}.$$

The normal derivative of q is

$$e_q(x) = \begin{cases} -2x_2^2 e^{1+x_2} & \text{for } x_1 = 1 \\ -3x_1 e^{1+x_1} & \text{for } x_2 = 1 \end{cases}$$

whereas the boundary term e_y is given by

$$e_y(x) = -u(x) + \frac{\partial y^*}{\partial n}(x)$$

$$\frac{\partial y}{\partial n} = \begin{cases} (8x_2 + 4x_2^2)e^{1+x_2} & \text{for } x_1 = 1 \\ 14x_1 e^{x_1+1} & \text{for } x_2 = 1. \end{cases}$$

The right hand side f reads

$$f(x) = -(4x_1x_2^2 + 16x_1x_2 + 4x_2^2 + 8x_2 + 12x_1)e^{x_1+x_2}.$$

First, the order of convergence for h-FEM with uniform refinement and BC-FEM shall be compared. Therefore, the $L_2(\Omega)$-error in y and $L_2(\Gamma_{\mathcal{N}})$-error in q, u are plotted against the mesh-width $1/h$ in Figure 2.

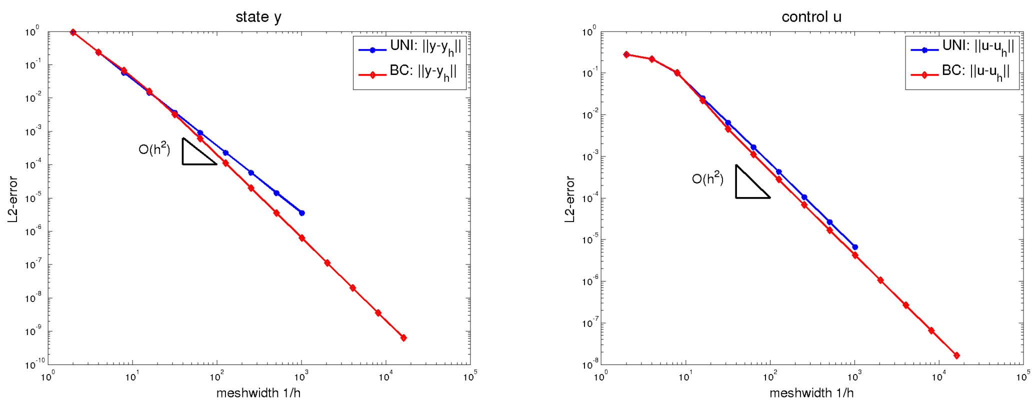

Figure 2: Error for uniform refinement with $p = 1$ and for a boundary concentrated mesh in dependence of the mesh-width for the control and the state

The plots show an $\mathcal{O}(h^2)$ order of convergence for h-FEM and BC-FEM, respectively. Theoretically we would expect $\mathcal{O}(h)$ from the theory for BC-FEM, which is a gap between theory and numerics (see Remark 3.14). Moreover, the error for the state in BC-FEM decreases slightly faster to zero than for uniform refinement. Figure 3 displays the discretization error for control and state with respect to the numbers of unknowns N. Since $N \sim h^{-1}$ for BC-FEM and $N \sim h^{-2}$ for h-FEM, the convergence for BC-FEM is much faster than for h-FEM. For instance if the error tolerance is fixed to $\varepsilon \approx 10^{-6}$, about 2% of the unknowns of h-FEM are required in order to obtain this accuracy, see the left picture of Figure 3. Furthermore, it becomes possible to solve optimal control problems with higher accuracy. Summarizing, the experiments show the advantages of BC-FEM, if the optimal control problem has to be solved quite accurately.

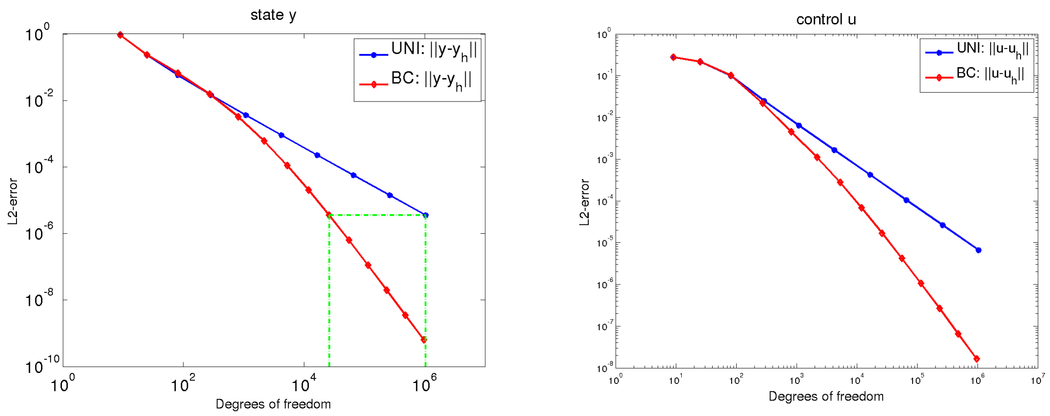

Figure 3: Comparison of the L_2 error in dependence of the degrees of freedom.

4.3 Example 2 - Domain observation

We choose

$$\alpha = 0.01, \quad u_a \equiv -10, \quad u_b \equiv 10, \quad e_y \equiv e_q \equiv 0$$

and

$$D(x) = \operatorname{diag}(2 + \sin(2\pi x_1 x_2)), \quad f \equiv 5, \quad y_d \equiv 5.$$

The domain $\Omega = (0,1)^2$ has the boundaries $\Gamma_{\mathcal{N}} = \{x_1 = 0\} \cup \{x_2 = 0\}$ and $\Gamma_{\mathcal{D}} = \Gamma \setminus \Gamma_{\mathcal{N}}$. As the error of the control behaves like the error in the adjoint variable, we only show the convergence history for the state y and adjoint q. As the solution (y^*, u^*) is unknown, the errors are computed with respect to the solution on the finest mesh.

h	$\|y_h^* - y^*\|_{L_2(\Omega)}$	$\mathrm{EOC}(y, L_2(\Omega))$	$\|q_h^* - q^*\|_{L_2(\Gamma_{\mathcal{N}})}$	$\mathrm{EOC}(q, L_2(\Gamma_{\mathcal{N}}))$
0.25	$4.34 \cdot 10^{-2}$	-	$7.22 \cdot 10^{-3}$	-
0.125	$1.25 \cdot 10^{-2}$	1.79	$2.13 \cdot 10^{-3}$	1.76
0.0625	$2.65 \cdot 10^{-3}$	2.24	$4.79 \cdot 10^{-4}$	2.15
0.0312	$4.75 \cdot 10^{-4}$	2.48	$9.57 \cdot 10^{-5}$	2.32
0.0156	$8.94 \cdot 10^{-5}$	2.41	$2.00 \cdot 10^{-5}$	2.26
0.00781	$1.67 \cdot 10^{-5}$	2.42	$4.32 \cdot 10^{-6}$	2.21
0.00391	$2.86 \cdot 10^{-6}$	2.55	$8.97 \cdot 10^{-7}$	2.27
0.00195	-	-	-	-

Table 2: Convergence rates for boundary concentrated mesh refinement.

The table shows that the experimental order of convergence lies in $[2, 2.6]$ if the first levels (where $h \in [0.125, 1]$) are neglected. Those mesh-widths correspond to relatively coarse meshes where the influence of some constants can shadow the asymptotic convergence. The state variable y converges faster compared to the adjoint variable q. Again, the observed convergence is higher than the result we proved.

4.4 Example 3 - Subdomain observation

Similar as before,

$$\alpha = 0.1, \quad u_a \equiv -0.8, \quad u_b \equiv -0.2, \quad e_y \equiv e_q \equiv 0.$$

Here, we set the domain $\Omega = (-2, 2)^2 \supset \Omega_s = (-1, 1)^2$ and Neumann boundary $\Gamma_{\mathcal{N}}$ on the line $\{x_2 = 2\}$ and $\Gamma_{\mathcal{D}} = \Gamma \setminus \Gamma_{\mathcal{N}}$. We choose

$$D \equiv 1, \quad f \equiv 1, \quad y_d = \chi_{\Omega_s}.$$

The numerical results are as follows.

h	$\|y_h^* - y^*\|_{L_2(\Omega)}$	EOC$(y, L_2(\Omega))$	$\|q_h^* - q^*\|_{L_2(\Gamma_{\mathcal{N}})}$	EOC$(q, L_2(\Gamma_{\mathcal{N}}))$
1	$2.16 \cdot 10^{-1}$	-	$2.22 \cdot 10^{-2}$	-
0.5	$5.57 \cdot 10^{-2}$	1.96	$6.23 \cdot 10^{-3}$	1.83
0.25	$1.44 \cdot 10^{-2}$	1.95	$1.67 \cdot 10^{-3}$	1.9
0.125	$3.18 \cdot 10^{-3}$	2.18	$3.59 \cdot 10^{-4}$	2.22
0.0625	$5.83 \cdot 10^{-4}$	2.45	$4.93 \cdot 10^{-5}$	2.86
0.0312	$1.06 \cdot 10^{-4}$	2.46	$6.94 \cdot 10^{-6}$	2.83
0.0156	$1.94 \cdot 10^{-5}$	2.45	$1.12 \cdot 10^{-6}$	2.63
0.00781	$3.68 \cdot 10^{-6}$	2.40	$2.23 \cdot 10^{-7}$	2.33
0.00391	$7.34 \cdot 10^{-7}$	2.33	$5.12 \cdot 10^{-8}$	2.12
0.000488	-	-	-	-

Table 3: Convergence rates for boundary concentrated mesh refinement.

The numerical behavior is similar to the previous example, i.e. we observe higher convergence rates than we were able to prove. While the EOC stays again in $[2, 2.5]$ for the state y, the adjoint variable shows rates up to 2.83 and thus converges faster.
A mesh we used for Example 3 can be seen in Figure 4. It is concentrated near the boundaries of the domain Ω and the observation domain Ω_s.

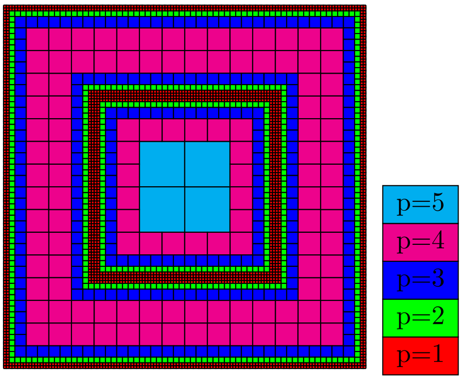

Figure 4: Admissible discretization for subdomain observation.

A Appendix

We will construct a BC-fem interpolation operator. Since we allow hanging nodes, these results generalize [19]. For the interpolation error we obtain approximation results comparable to those obtained in [19] for regular meshes.

A.1 Estimates of local element size and polynomial degrees

In the interpolation estimates below, it will be important to have comparable element size and element polynomial degree for neighboring elements. For meshes without hanging nodes, we have the following result from [24, Lemma 2.3], its extension to meshes with hanging nodes as used here is straightforward.

Lemma A.1. *Let τ be a γ-shape-regular mesh. Then there exists a constant $C(\gamma)$ such that for two neighboring elements K, K' with $\bar{K} \cap \bar{K}' \neq \emptyset$ there holds*

$$C(\gamma)^{-1} h_K \leq h_{K'} \leq C(\gamma) h_K. \tag{A.1}$$

Theorem A.2. *Let τ be a geometric mesh with a linear polynomial degree vector and slope α. Then there is a constant $C(\alpha)$ depending on γ such that for two neighboring elements K, K' with $\bar{K} \cap \bar{K}' \neq \emptyset$ it holds*

$$C(\alpha)^{-1} p_K \leq p_{K'} \leq C(\alpha) p_K.$$

Moreover, $C(\alpha) \in \mathcal{O}(\alpha)$.

Proof. The constants c_1, c_2 defining the linear degree vector naturally satisfy $c_2 > c_1$, cf. Definition 3.3. Using the properties of the linear degree vector and Lemma A.1 we can estimate

$$\begin{aligned}
p_{K'} &\leq 1 + \alpha c_2 \log(h_{K'}/h) \\
&\leq 1 + \alpha c_2 \log(C(\gamma) h_K/h) \\
&\leq 1 + \alpha c_2 \log(h_K/h) + \alpha c_2 \log(C(\gamma)) \\
&\leq c_2 c_1^{-1} (1 + \alpha c_1 \log(h_K/h) + \alpha c_2 \log(C(\gamma))) \\
&\leq c_2 c_1^{-1} (p_K + p_K \alpha c_2 \log(C(\gamma))) \\
&\leq c_2 c_1^{-1} (1 + \alpha c_2 \log(C(\gamma))) p_K.
\end{aligned}$$

The same computation yields a bound of p_K from above. This proves the claim with $C(\alpha) := \frac{c_2}{c_1}(1 + \alpha c_2 \log(C(\gamma)))$. □

A.2 Extension and projection operators

The reference element we have in mind is the square $[-1, 1]^2$, but we will keep the notation relatively neutral to make the results applicable to triangles as well. The index i is taken from $\{1, 2, 3(, 4)\}$. We take the reference element $\hat{K}$ and the space $Q_p(\hat{K}) := span\{x^i y^j \mid 0 \leq i, j \leq p\}$. Triangles would require the space $span\{x^i y^j \mid 0 \leq i + j \leq p\}$.
As our mesh will have hanging nodes, we assume that each edge e_i of the reference element has an associated polynomial degree $p_i := p_{e_i}$ (see (3.1)) with $p_i \leq p$. The constructed approximant will lie in

$$P_{\mathbf{p}(K)}(\hat{K}) := \{f \in Q_p(\hat{K}) \mid \deg(f|_{e_i}) = p_i, \quad \deg(f) \leq p\}. \tag{A.2}$$

We first need an extension operator acting from $\partial\hat{K}$ to $\hat{K}$ (see [23, Lemma 3.2.3]).

Lemma A.3. *Let $f \in C(\partial\hat{K})$ be a polynomial of degree p_i on the i-th edge of the reference element for all i. There exists a linear extension mapping $E : C(\partial\hat{K}) \to P_{\mathbf{p}(K)}(\hat{K})$ with the following properties*

$$(Ef)|_{e_i} = f \tag{A.3}$$

$$\| Ef \|_{L_\infty(\hat{K})} + p^{-2} \| \nabla Ef \|_{L_\infty(\hat{K})} \leq c \| f \|_{L_\infty(\partial\hat{K})} \tag{A.4}$$

Proof. We prove this only in the case of $\hat{K}$ being the reference square. The extension to triangular $\hat{K}$ is straightforward, see e.g. [23, Lemma 3.2.3].
By subtracting a bilinear function from f we can assume that it vanishes on the vertices of the reference element. For each $f_i := f|_{e_i}$ we construct an extension $E_i(f_i) \in P_{\mathbf{p}(K)}(\hat{K})$ which is zero at all other edges e_j, $j \neq i$.
Let us demonstrate the construction of $E_i(f_i)$ for e_1, $e_1 := \{(x,y) \in \mathbb{R}^2 \mid x \in [-1,1], y = -1\}$. Here we define $E_1(f_1) := \frac{1-y}{2} f(x)$. Analogously we define the extension from the edges e_i, $i > 1$. This way we get an extension $F := E(f) := \sum_i E_i(f_i)$.
With the inverse estimate $\| \nabla F \|_{L_\infty(\hat{K})} \leq cp^2 \| F \|_{L_\infty(\hat{K})}$ ([26, Theorem 4.76]) with $p \geq p_i$ we only need to show $\| F \|_{L_\infty(\hat{K})} \leq c \| f \|_{L_\infty(\partial\hat{K})}$. This is a trivial estimate: $\|E_1(f_1)\|_{L_\infty(\hat{K})} \leq \|f_1\|_{L^\infty(e_1)}$, as $\frac{1-y}{2} \leq 1$ on $\hat{K}$.
In the case that f does not vanish in the vertices let us denote by F_0 the bilinear interpolation of f that is exact in the vertices. Then we set $Ef := F_0 + \sum_i E_i(f_i - F_0)$. It is now easy to argue that the extension fulfills the claim. □

A.3 Construction of the bc-fem interpolation operator on meshes with hanging nodes

In the following, u will denote just a function and not the control variable as before. The aim of this section is to construct an interpolant on the reference element. It is desired to interpolate a function u living on the physical domain Ω by pulling it back to the reference element for each element of the finite element discretization τ.
The constructed interpolator will be needed for elements in the interior of Ω. There, we need to distinguish between elements possessing a hanging node or not.
At first, we will construct the interpolator for elements without hanging nodes. The following theorem is similar to [19, Lemma 2.9]. We give a proof here in order to track the dependence of the constants on the parameter α of the linear degree vector.
In the sequel we will denote by $GL(q,f)$ the one-dimensional Gauss-Lobatto interpolation of degree q for the function f on an edge.

Theorem A.4. *Let $\hat{K}$ be the reference element. Let u be a function on Ω whose pull back $\hat{u} = u \circ F_K$ is analytic on $\bar{\hat{K}}$ and satisfies*

$$\| \nabla^{q+2} \hat{u} \|_{L_\infty(\hat{K})} \leq C_u \gamma_u^q q!, \quad q = 0,1,2,\dots$$

Then there exists an interpolant $I(u) \in P_{\mathbf{p}(K)}(\hat{K})$ such that

1. $I(\hat{u})|_{e_i} = GL(p_i, \hat{u}|_{e_i})$,
2. $\| I(\hat{u}) - \hat{u} \|_{W^{1,\infty}(\hat{K})} \leq C_\alpha C_u e^{-bp_m}$,

where $b > 0$ depends on γ_u, and $C_\alpha > 0$ depends on γ_u and α with $C_\alpha = \mathcal{O}(\alpha^6)$ for $\alpha \to \infty$. Here, p_m denotes the minimal polynomial degree is defined by $p_m := \min_i\{p_i\}$ and naturally $p_m \leq p_i \leq p$ with p being the degree of the image of I, i.e. $P_{\mathbf{p}(K)}(\hat{K})$.

Proof. We restrict $\hat{u} \in C(\hat{K})$ to the boundary $\partial\hat{K}$ and define the piecewise Gauss-Lobatto interpolation operator

$$\begin{aligned} i : C(\partial\hat{K}) &\to \{f \in C(\partial\hat{K}) \mid f|_{e_i} \text{ is polynom with degree } p_i\}, \\ i(\hat{u})(x) &= GL(p_i, \hat{u}|_{e_i})(x) \quad \forall x \in \partial\hat{K}. \end{aligned}$$

Let us define the finite-dimensional subspace

$$V := \{u \in P_{\mathbf{p}(K)}(\hat{K}) \ : \ \hat{u}|_{\partial\hat{K}} = 0\}.$$

Since V is finite-dimensional, there is a linear and bounded projection operator $\Pi : P_{\mathbf{p}(K)}(\hat{K}) \to V$ with $\|\Pi\|_{\mathcal{L}(C(\bar{\hat{K}}),C(\bar{\hat{K}}))} \leq \sqrt{\mathrm{dim}V}$, confer [23, Theorem A.4.1]. Since $V \subset P_{\mathbf{p}(K)} \subset Q_p(\hat{K})$, we have $\dim(V) \leq (p+1)^2$, which shows $\|\Pi\|_{\mathcal{L}(C(\bar{\hat{K}}),C(\bar{\hat{K}}))} \leq p+1$.
The interpolation operator I is now defined by

$$I(\hat{u}) := E(i(\hat{u})) + \Pi(\hat{u} - E(i(\hat{u})))$$

with the extension operator E from Lemma A.3. By construction, the first property is fulfilled. If $\hat{u} \in P_{\mathbf{p}(K)}(\hat{K})$ it follows that $i(\hat{u}) = \hat{u}|_{\partial\hat{K}}$ and therefore $\hat{u} - E(i(\hat{u})) \in V$. Thus, I interpolates functions of $P_{\mathbf{p}(K)}(\hat{K})$ exactly.
Let $\hat{u} \in C(\bar{\hat{K}})$ be given. Let us first estimate the norm of I by

$$\begin{aligned} \| I(\hat{u}) \|_{L_\infty(\hat{K})} &\leq c\| i(\hat{u}) \|_{L_\infty(\partial\hat{K})} + (p+1)\| \hat{u} - E(i(\hat{u})) \|_{L_\infty(\hat{K})} \\ &\leq c(1+\ln p)\| \hat{u} \|_{L_\infty(\partial\hat{K})} + (p+1)\| \hat{u} \|_{L_\infty(\hat{K})} \\ &\quad + c(1+\ln p)(p+1)\| \hat{u} \|_{L_\infty(\partial\hat{K})}, \end{aligned}$$

where we used [23, Lemma 3.2.1] to bound the Gauss-Lobatto-interpolation operator i. Exploiting $\| \hat{u} \|_{L_\infty(\partial\hat{K})} \leq \| \hat{u} \|_{L_\infty(\hat{K})}$ for $\hat{u} \in C(\bar{\hat{K}})$ yields the estimate

$$\| I(\hat{u}) \|_{L_\infty(\hat{K})} \leq C_I p(1+\ln p)\| \hat{u} \|_{L_\infty(\hat{K})}.$$

Regarding approximation properties, it now follows with arbitrary $v \in P_{\mathbf{p}(K)}(\hat{K})$ and using $v = Iv$

$$\begin{aligned} \| \hat{u} - I(\hat{u}) \|_{L_\infty(\hat{K})} &= \| (\hat{u} - v) - I(\hat{u} - v) \|_{L_\infty(\hat{K})} \\ &\leq (1 + C_I p(1+\ln p))\| \hat{u} - v \|_{L_\infty(\hat{K})}. \end{aligned}$$

In order to achieve an approximation property in $W^{1,\infty}(\hat{K})$, we need to estimate the first derivatives of $\hat{u} - I(\hat{u})$:

$$\begin{aligned} &\| \nabla(\hat{u} - I(\hat{u})) \|_{L_\infty(\hat{K})} \\ &\quad = \| \nabla((\hat{u} - v) - I(\hat{u} - v)) \|_{L_\infty(\hat{K})} \\ &\quad \leq \| \nabla(\hat{u} - v) \|_{L_\infty(\hat{K})} + \| \nabla(I(\hat{u}) - v) \|_{L_\infty(\hat{K})} \\ &\quad \leq \| \nabla(\hat{u} - v) \|_{L_\infty(\hat{K})} + Cp^2\| (I(\hat{u}) - v) \|_{L_\infty(\hat{K})} \\ &\quad \leq \| \nabla(\hat{u} - v) \|_{L_\infty(\hat{K})} + Cp^2(\| (I(\hat{u}) - \hat{u}) \|_{L_\infty(\hat{K})} + \| \hat{u} - v \|_{L_\infty(\hat{K})}) \\ &\quad \leq \| \nabla(\hat{u} - v) \|_{L_\infty(\hat{K})} + Cp^2(2 + C_I p(1+\ln p))\| \hat{u} - v \|_{L_\infty(\hat{K})}. \end{aligned}$$

In the last two estimates, we can pass to the infimum because v was arbitrary, which shows

$$\begin{aligned}\|\hat{u}-I(\hat{u})\|_{L_\infty(\hat{K})} &\leq \hat{C}_1 p(1+\ln p) \inf_{v\in P_{\mathbf{p}(K)}(\hat{K})} \|\hat{u}-v\|_{L_\infty(\hat{K})},\\ \|\nabla(\hat{u}-I(\hat{u}))\|_{L_\infty(\hat{K})} &\leq \inf_{v\in P_{\mathbf{p}(K)}(\hat{K})} \{\|\nabla(\hat{u}-v)\|_{L_\infty(\hat{K})}\\ &\qquad + \hat{C}_2 p^3(1+\ln p)\|\hat{u}-v\|_{L_\infty(\hat{K})}\}.\end{aligned}$$

Relying on best approximation results in the space $P_{\mathbf{p}(K)}(\hat{K})$, we have [23, Theorem 3.2.19]

$$\begin{aligned}\inf_{v\in P_p(\hat{K})} \|\hat{u}-v\|_{L_\infty(\hat{K})} &\leq CC_u e^{-b' p_m},\\ \inf_{v\in P_p(\hat{K})} \|\nabla(\hat{u}-v)\|_{L_\infty(\hat{K})} &\leq CC_u e^{-b' p_m}\end{aligned}$$

with constants C, b' depending both on γ_u.
Collecting the estimates above, we obtain

$$\begin{aligned}\|I(\hat{u})-\hat{u}\|_{W^{1,\infty}(\hat{K})} &\leq \hat{C}_1 p(1+\ln p) CC_u e^{-b' p_m}\\ &\qquad + CC_u e^{-b' p_m} + \hat{C}_2 p^3(1+\ln p)\hat{C}_1 p(1+\ln p) CC_u e^{-b' p_m}.\end{aligned}$$

We have from Theorem A.2 that $C(\alpha)^{-1} p_{K'} \leq p_K \leq C(\alpha) p_{K'}$ for two neighboring elements K, K'. Hence, we can bound $p \leq C(\alpha) p_m$ because the minimal polynomial degree is determined by at least one neighbor. This way we get

$$\|I(\hat{u})-\hat{u}\|_{W^{1,\infty}(\hat{K})} \leq \hat{C}_3 C(\alpha)^6 p_m^6 CC_u e^{-b' p_m} \tag{A.5}$$

Absorbing p_m^6 by decreasing the constant b' yields

$$\|I(\hat{u})-\hat{u}\|_{W^{1,\infty}(\hat{K})} \leq C_\alpha C_u e^{-b p_m}$$

with C_α depending on α, γ_u and b on γ_u, and $C_\alpha = \mathcal{O}(\alpha^6)$ for $\alpha \to \infty$. □

Remark A.5. *We cannot avoid the constant $p(1+\ln p)$ in the estimates of $\|\hat{u}-I(\hat{u})\|_{L_\infty(\hat{K})}$ and $\|\nabla(\hat{u}-I(\hat{u}))\|_{L_\infty(\hat{K})}$ as we allow different polynomial degrees in the interior and on the edges of elements.*

In the second step, we will construct an interpolation operator that can deal with hanging nodes. To begin with, we cite an one-dimensional interpolation result of [23, Lemma 3.2.6].

Lemma A.6. *Let u be analytic on the interval $[-1,1]$ and satisfy for some C_u, γ_u*

$$\|\nabla^{q+2} u\|_{L_\infty(I)} \leq C_u \gamma_u^q q! \quad q = 0,1,2,\ldots$$

There are constants $C, b > 0$ depending on γ_u such that $GL(q,u)$ satisfies for $p = 1,2,\ldots$

$$\|u - GL(p,u)\|_{W^{1,\infty}(I)} \leq CC_u e^{-bp}.$$

Proof. In [23, Lemma 3.2.6], the estimate $\|u - GL(p,u)\|_{W^{1,\infty}(I)} \leq \kappa C_u \left(\frac{1}{1+\sigma}\right)^{p+1}$ is proven with $\kappa, \sigma > 0$ depending on γ_u. With $C = \kappa(1+\sigma)^{-1}$ and $b = \ln(1+\sigma)$ we obtain the desired estimate. □

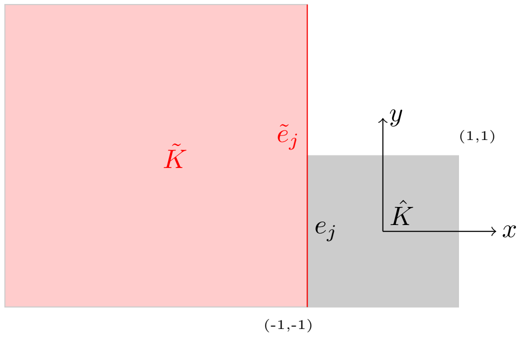

Figure 5: reference element with hanging node and its neighbor (possibly distorted).

Let us describe now the construction of an interpolator on elements with hanging nodes. Depending on the position of the hanging nodes, we prolong the local edge e_j to the full coarse edge $\tilde{e}_j$, with j from $\{1, 2, 3(, 4)\}$. An exemplary situation is depicted in Figure 5.

Theorem A.7 (hanging nodes). *Let $\hat{K}$ be the reference element. Let u be a function on Ω whose pull back $\hat{u}$ is analytic on $\bar{\hat{K}}$ and satisfies*

$$\| \nabla^{q+2} \hat{u} \|_{L_\infty(\hat{K})} \leq C_u \gamma_u^q q!, \quad q = 0, 1, 2, \ldots \tag{A.6}$$

Let the indices i represent the free edges, whereas j denotes constrained edges due to the existence of hanging nodes. If additionally it holds

$$\| \nabla^{q+2} \hat{u} \|_{L_\infty(\tilde{e}_j)} \leq C_u \gamma_u^q q!, \quad q = 0, 1, 2, \ldots \tag{A.7}$$

with $C_2, \gamma_2 > 0$ then there exists an interpolant $I(\hat{u}) \in P_{\mathbf{p}(K)}(\hat{K})$ such that

1. $\tilde{I}(\hat{u})|_{e_i} = GL(p_i, \hat{u}|_{e_i})$,
2. $\tilde{I}(\hat{u})|_{e_j} = GL(p_j, \hat{u}|_{\tilde{e}_j})|_{e_j}$.
3. $\| \tilde{I}(\hat{u}) - \hat{u} \|_{W^{1,\infty}(\hat{K})} \leq \tilde{C}(\alpha) C_u e^{-bp_m}$

where b depends on γ_u, γ_2. The constant $\tilde{C}(\alpha)$ is at most $\mathcal{O}(\alpha^6)$ for $\alpha \to \infty$.

Let us comment on the impact of Theorem A.7. Due to property 1. and 2., it is possible to construct a complete interpolant in an element by element fashion. Together with Theorem A.4 it is guaranteed, that the resulting interpolant is continuous across each edge and therefore the global interpolant lies in the conforming finite element space $\mathbb{V}^{\mathbf{p}}_{l,\Gamma_{\mathcal{D}}}$. This is possible as the definition of the finite element space enforces that the polynomial degree on a constrained edge coincides with the polynomial degree on the corresponding coarse edge.

Proof. We define the piecewise Gauss-Lobatto interpolation operator as

$$\tilde{i} : C(\partial\hat{K} \cup \bigcup_j \tilde{e}_j) \to \{f \in C(\partial\hat{K}) : \ f_i|_{e_i} \text{ is polynomial of degree } p_i\},$$

$$\tilde{i}(\hat{u})(x) = GL(p_i, \hat{u}|_{e_i})(x), \quad x \in e_i,$$
$$\tilde{i}(\hat{u})(x) = GL(p_j, \hat{u}|_{\tilde{e}_j})|_{e_j}, \quad x \in e_j.$$

The function $\hat{u} = u \circ F_K$ can also be evaluated at points outside of $\hat{K}$ since the mapping F_K is analytic. Thus, the Gauss-Lobatto interpolation on $\tilde{e}$ is well defined.
With the operators defined in the proof of Theorem A.4 we define the interpolation operator as

$$\tilde{I} = E(\tilde{i}(\hat{u})) - \Pi(\hat{u} - E(i(\hat{u}))).$$

We compute

$$\| \tilde{I}(\hat{u}) - \hat{u} \|_{W^{1,\infty}(\hat{K})} \leq \| I(\hat{u}) - \hat{u} \|_{W^{1,\infty}(\hat{K})} + \| I(\hat{u}) - \tilde{I}(\hat{u}) \|_{W^{1,\infty}(\hat{K})},$$

where I is given by Theorem A.4. The first addend is bounded by $C_\alpha C_u e^{-bp_m}$ due to Theorem A.4. So we only need to estimate the second one. Using Lemma A.3 we find

$$\begin{aligned}
&\| I(\hat{u}) - \tilde{I}(\hat{u}) \|_{W^{1,\infty}(\hat{K})} \\
&\quad = \| E(i(u)) - E(\tilde{i}(u)) \|_{W^{1,\infty}(\hat{K})} \\
&\quad \leq cp^2 \| i(\hat{u}) - \tilde{i}(\hat{u}) \|_{L_\infty(\partial \hat{K})} \\
&\quad = cp^2 \| \sum_j GL(p_j, \hat{u}|_{e_j}) - GL(p_j, \hat{u}|_{\tilde{e}_j})|_{e_j} \|_{L_\infty(e_j)} \\
&\quad \leq cp^2 \sum_j \left(\| GL(p_j, \hat{u}|_{e_j}) - \hat{u} \|_{L_\infty(e_j)} + \| GL(p_j, \hat{u}|_{\tilde{e}_j})|_{e_j} - \hat{u} \|_{L_\infty(e_j)} \right).
\end{aligned}$$

The first addends are bounded due to (A.6) and Lemma A.6.

$$\sum_j \| GL(p_j, \hat{u}|_{e_j}) - \hat{u} \|_{L_\infty(e_j)} \leq \sum_j CC_u e^{-b_1 p_j} \leq 4CC_u e^{-b_1 p_m}. \tag{A.8}$$

If we use an affine mapping from $\tilde{e}_j$ to $[-1, 1]$, the prerequisite (A.7) transforms into

$$\| \nabla^{q+2} \hat{u} \|_{L_\infty(-1,1)} \leq C_u (2\gamma_u)^q q!, \quad q = 0, 1, 2, \dots.$$

Using again Lemma A.6 we find

$$\begin{aligned}
\sum_j \| GL(p_j, \hat{u}|_{\tilde{e}_j})|_{e_j} - \hat{u} \|_{L_\infty(e_j)} &\leq \sum_j \| GL(p_j, \hat{u}|_{\tilde{e}_j}) - \hat{u} \|_{L_\infty(\tilde{e}_j)} \\
&\leq \sum_j CC_u e^{-b_2 p_j} \leq 4CC_u e^{-b_2 p_m}.
\end{aligned}$$

With $cp^2 \leq cC(\alpha)^2$, the final estimate reads

$$\begin{aligned}
\| \tilde{I}(\hat{u}) - \hat{u} \|_{W^{1,\infty}(\hat{K})} &\leq C_\alpha C_u e^{-bp_m} + cC(\alpha)^2 (CC_u e^{-b_1 p_m} + CC_u e^{-b_2 p_m}) \\
&\leq \tilde{C}_\alpha C_u e^{-\tilde{b} p_m}.
\end{aligned}$$

with $\tilde{C}_\alpha$ depending on α, γ_u and $\tilde{b}$ on γ_u. As $C_\alpha \in \mathcal{O}(\alpha^6)$, it follows $\tilde{C}_\alpha \in \mathcal{O}(\alpha^6)$. □

Remark A.8. *Note that the interpolation operator projects $\hat{u} - E(i(\hat{u}))$ instead of $\hat{u} - E(\tilde{i}(\hat{u}))$ onto the subspace V of polynomials vanishing at the boundary of the element. This simplifies the interpolation error estimates.*

A.4 Best-approximation and discretization error estimates

First we establish an easy lemma to conveniently check the prerequisites of Theorem A.4 and A.7.

Lemma A.9. *Let u be a function on Ω that satisfies*

$$\| \nabla^q u \|_{L_2(\Omega)} \leq C_u \gamma_u^q q!, \quad q = 0, 1, 2, ... \tag{A.9}$$

Then u is analytic on $\bar{\Omega}$ and scaling constants $C_s, c_s > 0$ exist such that

$$\| \nabla^q u \|_{C(\bar{\Omega})} \leq C_s C_u (c_s \gamma_u)^q q!, \quad q = 0, 1, 2, ...$$

Proof. For an arbitrary but fixed q, we have $\nabla^q u \in H^2(\Omega)$. A Sobolev embedding implies

$$\| \nabla^q u \|_{C(\bar{\Omega})} \leq C \| \nabla^q u \|_{H^2(\Omega)}.$$

Estimating each derivative of u appearing in the $H^2(\Omega)$ norm separately with (A.9) yields

$$\| \nabla^q u \|_{C(\bar{\Omega})} \leq C(1 + \gamma_u + \gamma_u^2) C_u \gamma_u^q (q + 2)!.$$

Choosing $C_s := 2C(1 + \gamma_u + \gamma_u^2)$ and $c_s = 6$, which implies $c_s^q \geq (q + 2)(q + 1)$ for $q \geq 1$, proves the estimate, which in turn gives analyticity of u on $\bar{\Omega}$. □

The proof of the following theorem is inspired by [19, Proposition 2.10].

Theorem A.10. *Let τ be a $\gamma-$shape-regular geometric mesh with the properties of section A.1. Let $u \in B^2_{1-\delta}(C_u, \gamma_u)$ for some $\delta \in (0, 1]$ and $\mathbf{p}$ the linear degree vector with slope α. Then for sufficiently large α it holds*

$$inf \{ \| u - v \|_{H^1(\Omega)} \mid v \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_{\mathcal{D}}} \} \leq C \, C_u \, h^\delta.$$

Here, C depends on Ω, γ_u, α and the shape regularity constant γ, but not on C_u. The choice of α depends on all these constants as well but not on C_u.

We want to construct the interpolant element by element. On elements abutting the boundary we will use the linear interpolant because the linear degree vector does not allow larger polynomial degrees on elements of size h.
For elements not abutting the boundary we want to take advantage of the increased polynomial degree to achieve good approximation quality. The previous error estimates of the interpolants, however, depend on the minimal polynomial degree p_m which is determined by at least one neighbor element. To guarantee that the neighbor's polynomial degree (and thus p_m) can be increased sufficiently, we introduce a second layer of elements near the boundary.

Proof. Overall we distinguish the following cases:

1. Elements K collected in τ_b abutting the boundary, i.e. $\bar{K} \cap \partial\Omega \neq \emptyset$.
2. Elements in the 'second' layer near the boundary, i.e. $K \in \tau$ such that $\bar{K} \cap \partial\Omega = \emptyset$ and $\exists K' \in \tau$ with $\bar{K} \cap \bar{K}' \neq \emptyset$, $\bar{K}' \cap \partial\Omega \neq \emptyset$. These elements are collected in τ_s.
3. Elements without hanging nodes which do not belong to $\tau_b \cup \tau_s$. They are collected in τ_f.
4. Elements that do not fall into the previous categories, i.e. elements with hanging nodes which do not belong to $\tau_b \cup \tau_s$. They form the set τ_c.

Let $u \in B^2_{1-\delta}(C_u,\gamma_u)$. For an element K we define the constant C_K by

$$C_K^2 = \sum_{q=0}^{\infty} \frac{1}{(2\gamma_u)^{2q}(q!)^2} \| r^{q+1-\delta} \nabla^{q+2} u \|^2_{L_2(K)}.$$

It holds

$$\| r^{q+1-\delta}\nabla^{q+2} u \|_{L_2(K)} \le C_K (2\gamma_u)^q q!, \tag{A.10}$$

$$\sum_{K\in\tau} C_K^2 \le \frac{4}{3} C_u^2. \tag{A.11}$$

Additionally, we define

$$\tilde{C}_K^2 := C_K^2 + \sum_{K' : \bar{K}\cap\bar{K}' \neq \emptyset} C_{K'}^2,$$

which implies $\sum_{K\in\tau} \tilde{C}_K^2 \le \frac{16}{3} C_u^2$.
We construct an interpolant $u_h \in \mathbb{V}^{\mathbf{p}}_{l,\Gamma_{\mathcal{D}}}$ of u for each element K falling into one of the four categories above. In the following, the index q will always be from $\mathbb{N}\cup\{0\}$.

1. $K \in \tau_b$. Let I_{lin} denote the linear or bilinear interpolation. We set $u_h|_K := I_{lin} u|_K$. We use [19, Appendix B.4] and the property 1. of Definition 3.2 to obtain

$$\| u - u_h \|_{H^1(K)} \le \| u - I_{lin}(u) \|_{H^1(K)} \le C h_K^\delta \| r^{1-\delta}\nabla^2 u \|_{L_2(K)} \le C h^\delta C_K.$$

3. $K \in \tau_f$. The pullback $\hat{u}$ of u on $\hat{K}$ satisfies

$$\begin{aligned} \| \nabla^{q+2}\hat{u} \|_{L_2(\hat{K})} &\le C h_K^{q+1} \| \nabla^{q+2} u \|_{L_2(K)} \\ &\le C h_K^{q+1} \| r^{q+1-\delta}\nabla^{q+2} u \|_{L_2(K)} \frac{1}{\inf_{x\in K} r(x)^{q+1-\delta}} \end{aligned} \tag{A.12}$$

As $r(x)$ for $x \in K$ is bounded from below by the diameter of the largest inscribed circle of a neighboring element, γ-shape-regularity yields

$$\inf_{x\in K} r(x) \ge \tilde{c}(\gamma) h_K$$

for a $\tilde{c}(\gamma) > 0$. Consequently,

$$\| \nabla^{q+2}\hat{u} \|_{L_2(\hat{K})} \le C C_K h_K^\delta (2\tilde{c}\gamma_u)^q q!.$$

where C is possibly rescaled by $\tilde{c}(\gamma)$.
We set $u_h|_K := I(\hat{u}) \circ F_K^{-1}$, where I is given by Theorem A.4. Due to Lemma A.9 we can apply Theorem A.4 and get

$$\| u - u_h \|_{H^1(K)} \le C_\alpha C C_K h_K^\delta e^{-b p_{m,K}}$$

with b, C_α given by Theorem A.4 depending on γ_u but not on C_u and K. Using

$$p_m = p_{K'} \ge c\alpha \ln(h_{K'}/h)$$

for a neighbor K' of element K, we arrive at

$$\| u - u_h \|_{H^1(K)} \le C_\alpha C C_K h_{K'}^{\delta-\alpha b} h^{\alpha b}.$$

Using $h_{K'} \ge ch$ yields

$$h_{K'}^{\delta-\alpha b} h^{\alpha b} \le h^{\min\{\delta,\alpha b\}}.$$

4. $K \in \tau_c$. We set $\hat{K} := F_K^{-1}(K)$ and denote the edges of $\hat{K}$ that possess a hanging node by e_j, $j \in \{1, \dots, 4\}$. The coarse edge that contains e_j is denoted by $\tilde{e}_j$ in reference coordinates. Let K_j denote the neighboring element of K that contains the same hanging node, i.e. $\bar{K}_j \cap F_K(\tilde{e}_j) \neq \emptyset$, and set $\hat{K}_j := F_K^{-1}(K_j)$. For an illustration see Figure 6.

   In order to apply Theorem A.7, we have to estimate L^∞-norms of the pullback on the extended edge $\tilde{e}_j$. With the properties of the elements in τ_c we obtain

$$\| \nabla^{q+2}\hat{u} \|_{L_\infty(\tilde{e}_j)} \leq \| \nabla^{q+2}\hat{u} \|_{C(\tilde{e}_j)} \leq C\| \nabla^{q+2}\hat{u} \|_{C(\hat{E}_j)}$$

   with $\hat{E}_j = \operatorname{int}\operatorname{conv}(\hat{K} \cup \hat{e}_j) \subset \hat{K} \cup \hat{K}_j$. Let us emphasize that the constant C depends on $\tilde{E}_j$ but not on $\tilde{K}_j$. Hence, C is independent of K_j, and thus it is independent of the mesh.

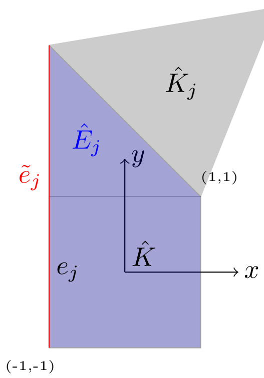

Figure 6: reference element $\hat{K}$ enlarged to $\hat{E}_j$ to handle a hanging node.

   Since $h_{K_j'}$ and h_K are comparable due to Lemma A.1, we obtain analogously to (A.12)

$$\| \nabla^{q+2}\hat{u} \|_{L_2(\hat{K}_j)} \leq C C_{K_j} h_K^\delta (2\gamma_u)^q q! \tag{A.13}$$

   with a possibly larger constant C independent of K, K_j. The two estimates (A.12) and (A.13) yield

$$\| \nabla^{q+2}u \|_{L_2(\hat{E}_j)} \leq C(C_{K_j} + C_K) h_K^\delta (2\gamma_u)^q q!$$

   and Lemma A.9 shows that the prerequisites for Theorem A.7 are fulfilled. So we set $u_h|_K := \tilde{I}(\hat{u}) \circ F_K^{-1}$, with $\tilde{I}$ given by Theorem A.7. The result of this theorem yields

$$\begin{aligned} \| u - u_h \|_{H^1(K)} &\leq \tilde{C}_\alpha C \left(C_K + \sum_{K' : \bar{K} \cap \bar{K}' \neq \emptyset} C_{K'} \right) h_K^\delta e^{-bp_{m,K}} \\ &= \tilde{C}_\alpha C \tilde{C}_K h_K^\delta e^{-bp_{m,K}}. \end{aligned}$$

   Arguing as in the case $K \in \tau_f$, we find

$$\| u - u_h \|_{H^1(K)} \leq \tilde{C}_\alpha C \tilde{C}_K h^{\min\{\delta, \alpha b\}}.$$

2. $K \in \tau_s$. Here, we set $u_h|_K := I(\hat{u}) \circ F_K^{-1}$ if K has no hanging nodes or $u_h|_K := \tilde{I}(\hat{u}) \circ F_K^{-1}$ otherwise. Analogously as in the cases $K \in \tau_f$, $K \in \tau_c$, we obtain

$$\| u - u_h \|_{H^1(K)} \leq \tilde{C}_\alpha C \tilde{C}_K h_K^\delta e^{-bp_{m,K}}.$$

However, we cannot apply $p_m \geq \alpha \ln(h_K/h)$ because $p_m = 1$, and thus p_m is fixed and cannot be increased. In geometric meshes, the element size h_K is proportional to the size of a neighboring element. In the second layer, there is a neighbor abutting the boundary, so we find $C(\gamma)^{-1}h \leq h_K \leq C(\gamma)c_2 h$. Thus, we obtain for a possibly adapted C

$$\| u - u_h \|_{H^1(K)} \leq \tilde{C}_\alpha C \tilde{C}_K h^\delta.$$

Overall we now estimate

$$\sum_{K\in\tau} \| u - u_h \|^2_{H^1(K)} \leq C^2 \Big(\sum_{K\in\tau_b} C_K^2 h^{2\delta} + \tilde{C}_\alpha^2 \sum_{K\in\tau_s} \tilde{C}_K^2 h^{2\delta} + C_\alpha^2 \sum_{K\in\tau_f} C_K^2 h^{2\min\{\delta,\alpha b\}} + \tilde{C}_\alpha^2 \sum_{K\in\tau_c} \tilde{C}_K^2 h^{2\min\{\delta,\alpha b\}} \Big).$$

Since b is independent of α, we can choose α large enough to obtain

$$\sum_{K\in\tau} \| u - u_h \|^2_{H^1(K)} \leq C^2 \; C_u^2 \; h^{2\delta}.$$

By construction u_h is a continuous function on $\bar{\Omega}$. Thus, it holds $u_h \in H^1(\Omega)$ and

$$\| u - u_h \|_{H^1(\Omega)} \leq C \; C_u \; h^\delta.$$

□

Remark A.11. *The proof only works for affine linear or bilinear mappings F_K. The reason is that prolonged edges of the reference element have to be straight lines under F_K, so that in global coordinates they coincide with the coarse edges. Together with the property that hanging nodes are in the middle of a coarse edge, the described procedure and usage of interpolation operators works.*

Theorem A.12 (Lemma 3.6)**.** *Let τ be a geometric mesh on Ω with mesh size h, $\mathbf{p}$ a linear degree vector with slope α. Suppose Assumptions 1 and 2 are satisfied. Let $y \in H^{1+\delta}(\Omega)$ for some $\delta \in (0,1]$ be a solution to the state equation (3.4) with data $u \in L_2(\Gamma_\mathcal{N})$ and $f \in B^0_{1-\delta}(C_f, \gamma_f)$, $C_f, \gamma_f > 0$. Then for sufficiently large α there is $C > 0$ independent of h such that*

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

holds.

Proof. According to Theorem 2.3 the solution also lies in $B^2_{1-\delta}(C_y, \gamma_y)$ for some constants $C_y, \gamma_y > 0$. In view of the best approximation properties of the FE solution (Cea's lemma) and the approximation quality from Theorem A.10 the proof is complete. □

References

[1] R.A. Adams. *Sobolev Spaces.* Academic Press, New York, London, 1975.

[2] T. Apel, J. Pfefferer, and A. Rösch. Finite element error estimates for Neumann boundary control problems on graded meshes. *Comput. Optim. Appl.*, 52(1):3–28, 2012.

[3] T. Apel, J. Pfefferer, and A. Rösch. Finite element error estimates on the boundary with application to optimal control. *Math. Comp.*, 84(291):33–70, 2015.

[4] I. Babuška and B. Guo. Regularity of the Solution of Elliptic Problems with Piecewise Analytic Data. Part I: Boundary Value Problems for Linear Elliptic Equation of Second Order. *SIAM J. Math. Anal.*, 19(1):172–203, 1988.

[5] I. Babuška and B. Guo. Regularity of the Solution of Elliptic Problems with Piecewise Analytic Data. Part II: The Trace Spaces and Application to the Boundary Value Problems with Nonhomogeneous Boundary Conditions. *SIAM J. Math. Anal.*, 20(4):763–781, 1989.

[6] R. E. Bank, A. H. Sherman, and A. Weiser. Some refinement algorithms and data structures for regular local mesh refinement. *Scientific Computing, Applications of Mathematics and Computing to the Physical Sciences*, 1:3–17, 1983.

[7] S. Beuchler, C. Pechstein, and D. Wachsmuth. Boundary concentrated finite elements for optimal boundary control problems of elliptic PDEs. *Comput. Optim. Appl.*, 51(2):883–908, 2012.

[8] E. Casas and M. Mateos. Error estimates for the numerical approximation of Neumann control problems. *Comput. Optim. Appl.*, 39(3):265–295, 2008.

[9] 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.

[10] Y. Chen and Y. Lin. A posteriori error estimates for *hp* finite element solutions of convex optimal control problems. *J. Comput. Appl. Math.*, 235(12):3435–3454, 2011.

[11] Y. Chen, N. Yi, and W. Liu. A Legendre-Galerkin spectral method for optimal control problems governed by elliptic equations. *SIAM J. Numer. Anal.*, 46(5):2254–2275, 2008.

[12] L. Demkowicz. *Computing with hp-adaptive Finite Elements.* CRC Press Taylor Francis, Boca Raton, London, New York, 2006.

[13] T. Eibner and J.M. Melenk. A local error analysis of the boundary-concentrated *hp*-FEM. *IMA J. Numer. Anal.*, 27(1):752–778, 2007.

[14] W. Gong, W. Liu, and N. Yan. A posteriori error estimates of *hp*-FEM for optimal control problems. *Int. J. Numer. Anal. Model.*, 8(1):48–69, 2011.

[15] P. Grisvard. *Elliptic Problems in Nonsmooth Domains.* Pitman, Boston–London–Melbourne, 1985.

[16] M. Hinze and U. Matthes. A note on variational discretization of elliptic neumann boundary control. *Control & Cybernetics*, 38:577–591, 2009.

[17] D. S. Jerison and C. E. Kenig. The Neumann problem on Lipschitz domains. *Bull. Amer. Math. Soc. (N.S.)*, 4(2):203–207, 1981.

[18] G.E.M. Karniadakis and S.J. Sherwin. *Spectral/hp Element Methods for Computational Fluid Dynamics.* Numerical Mathematics and Scientific Computation Series. Oxford University Press, 1999.

[19] B. N. Khoromskij and J. M. Melenk. Boundary Concentrated Finite Element Methods. *SIAM J. Numer. Anal.*, 41:1–36, 2003.

[20] A. Kufner and A.-M. Sändig. *Some applications of weighted Sobolev spaces.* BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1987.

[21] K.Yosida. *Functional Analysis.* Springer-Verlag, Berlin, Göttingen, Heidelberg, 1965.

[22] M. Mateos and A. Rösch. On saturation effects in the Neumann boundary control of elliptic optimal control problems. *Comput. Optim. Appl.*, 49(2):359–378, 2011.

[23] J. M. Melenk. *hp-Finite Element Methods for Singular Perturbations*, volume 1796 of *Lecture Notes in Mathematics.* Springer Berlin / Heidelberg, 2002.

[24] J. M. Melenk. *hp*-interpolation of nonsmooth functions and an application to *hp*-a posteriori error estimation. *SIAM J. Numer. Anal.*, 43(1):127–155, 2005.

[25] J. Nečas. *Les méthodes directes en théorie des équations elliptiques.* Academia, Prague, 1967.

[26] C. Schwab. *p- and hp- Finite Element Methods - Theory and Applications in Solid and Fluid Mechanics.* Numerical Mathematics and Scientific Computation. Clarendon Press, Oxford, 1998.

[27] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics.* AMS, Providence, RI, 2010.

[28] J. Wloka. *Funktionalanalysis und Anwendungen.* Walter de Gruyter, Berlin, 1971.