Computational Optimization and Applications (2020) 76:155–199
https://doi.org/10.1007/s10589-020-00171-5

Error estimates for the finite element approximation of bilinear boundary control problems

Max Winkler[1]

Received: 11 January 2019 / Published online: 7 February 2020

Abstract
In this article a special class of nonlinear optimal control problems involving a bilinear term in the boundary condition is studied. These kind of problems arise for instance in the identification of an unknown space-dependent Robin coefficient from a given measurement of the state, or when the Robin coefficient can be controlled in order to reach a desired state. Necessary and sufficient optimality conditions are derived and several discretization approaches for the numerical solution of the optimal control problem are investigated. Considered are both a full discretization and the postprocessing approach meaning that we compute an improved control by a pointwise evaluation of the first-order optimality condition. For both approaches finite element error estimates are shown and the validity of these results is confirmed by numerical experiments.

Keywords Bilinear boundary control · Identification of Robin parameter · Finite element error estimates · Postprocessing approach

1 Introduction

This paper is concerned with bilinear boundary control problems of the form

$$J(y,u) := \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Gamma)} \to \min!$$

subject to

$$\begin{aligned} -\Delta y + y &= f \quad \text{in} \quad \Omega, \\ \partial_n y + u\, y &= g \quad \text{on} \quad \Gamma, \\ u \in U_{ad} &:= \{v \in L^2(\Gamma) : u_a \le u \le u_b \quad \text{a.e. on} \quad \Gamma\}, \end{aligned}$$

✉ Max Winkler
max.winkler@mathematik.tu-chemnitz.de

1 Faculty of Mathematics, Technische Universität Chemnitz, Chemnitz, Germany

where $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, is a bounded domain, $\alpha > 0$ is the regularization parameter, $y_d \in L^2(\Omega)$ is a desired state and $0 \le u_a < u_b$ are the control bounds.

As an application of bilinear boundary control problems we mentioned the identification of an unknown Robin coefficient from a given measurement y_d of the state quantity. This is for instance of interest in the modeling of stem cell division processes [16, 17], where u is the unknown parameter describing the chemical reactions between proteins from the cell interior and the cell cortex. For further applications, u can be interpreted as a heat-exchange coefficient in thermodynamics or as a quantity for corrosion damage in electrostatics. There are many publications dealing with the identification of the Robin coefficient, see for instance [12, 23, 31, 34]. Only a few papers use an optimal control approach similar to the one considered in the present article. We mention [22, 25], where the parabolic version of our model problem is considered. The authors prove convergence of a finite element approximation but no convergence rate is established. A similar problem is discussed in [21], dealing with the recovery of the Robin parameter in a variational inequality.

The aim of the present paper is to derive necessary and sufficient optimality conditions for the optimal control problem and to investigate several numerical approximations regarding convergence towards a local solution. This complements a previous contribution of Kröner and Vexler [27] where the distributed control case, meaning that the bilinear term $u\,y$ appears in the differential equation, is discussed. The main results in their article are error estimates for the approximate controls in the $L^2(\Omega)$-norm for several finite element approximations. To be more precise, the convergence rate 1 is shown for piecewise constant and 3/2 for piecewise linear approximations for the control. Moreover, advanced discretization concepts like the postprocessing approach [32] and the variational discretization [24] are investigated which allow an improvement up to a convergence rate of 2. It is the purpose of the present article to extend the results to the case of bilinear boundary control.

The numerical analysis of boundary control problems is usually more difficult than for distributed control problems as the adjoint control-to-state operator maps onto some Sobolev/Lebesgue space defined on the boundary. As a consequence, error estimates for the traces of finite element solutions have to be proved, more precisely, in the $L^2(\Gamma)$ -norm. Here, we consider two different discretization approaches. The first one is a full discretization using piecewise linear finite elements for the states and piecewise constant functions on the boundary for the control approximation. Under the assumption that the domain has a Lipschitz boundary we show that the discrete optimal control converges with the optimal rate 1. To show this result we exploit the local coercivity of the objective, best-approximation properties of the control space and suboptimal error estimates for the state and adjoint equation. In order to obtain a more accurate solution we also investigate the postprocessing approach where an improved control is computed by a pointwise application of the first-order optimality condition to the discrete state variables. For this approach we have to assume more regularity for the exact solution and thus, we restrict our considerations to two-dimensional domains with sufficiently smooth boundary. Under this assumption we show the optimal convergence rate of $2 - \varepsilon$ with arbitrary $\varepsilon > 0$ which is the rate one would also expect in the case of linear quadratic boundary control problems and smooth solutions [3, 4, 33] (even with $h^{-\varepsilon}$ replaced by $|\ln h|$, where h is the maximal element diameter of the finite element mesh). The proof relies on the non-expansivity of the projection onto the feasible set as well as sharp error estimates for the state and adjoint state in $L^2(\Gamma)$. To obtain estimates in these norms superconvergence properties of the midpoint interpolant, finite element error estimates for the Ritz projection in $L^2(\Gamma)$ and a supercloseness result between the midpoint interpolant of the exact and the discrete solution are exploited. To show the $L^2(\Gamma)$-norm error estimate we will, as we consider smooth solutions, derive a maximum norm estimate. To the best of the author's knowledge these results are not available in the literature for problems with Robin boundary conditions. Based on the ideas from [18] we formulate the missing proof.

We moreover note that the setting discussed here does not fit into the well-known framework of the semilinear optimal control problems discussed e. g. in [5, 9, 11, 29], as these contributions deal with nonlinearities depending solely on the state variable. However, many techniques can be reused for the problem considered here. The only publication where more general nonlinearities depending both on the state and the control variable is, to the best of the author's knowledge, [35]. Therein optimality conditions are discussed but there is no theory on the numerical analysis of approximation methods for this problem class available yet. However, we think that the consideration of bilinear control problems may serve as a starting point for the investigation of a more general class of nonlinear optimal control problems.

The article is structured as follows. In Sect. 2 we discuss the solubility of the state equation and regularity results for its solution. In Sect. 3 we analyze the optimal control problem. In particular, necessary and sufficient optimality conditions are investigated. Section 4 is devoted to the finite element discretization of the state equation, where we show finite element error estimates required for the numerical analysis of the optimal control problem later. The discretization of the optimal control problem is considered in Sect. 5. In particular, we discuss convergence rates for the numerical solution obtained by a full discretization of the optimal control problem as well as for an improved control obtained by a postprocessing step. The latter result requires some auxiliary results that we discuss in the appendix. To be more precise, a maximum norm error estimate for the finite element solution of an elliptic equation with Robin boundary conditions is needed. A proof is given in Appendix 1. Moreover, a proof of local error estimates for the midpoint interpolant and the $L^2(\Gamma)$ projection onto piecewise constant functions on the boundary is needed. To the best of the author's knowledge these results are not available in the literature in case of domains with curved boundaries. Thus, we discuss these auxiliary results in Appendix 2. Finally, we will compare the theoretical results with numerical experiments in Sect. 6.

2 Analysis of the state equation

We consider the boundary value problem

$$-\Delta y + y = f \quad \text{in} \quad \Omega, \qquad \partial_n y + u\, y = g \quad \text{on} \quad \Gamma,$$

on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, with data $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$. The corresponding weak formulation reads

$$\text{Find } y \in H^1(\Omega) : \qquad a_u(y,v) = F(v) \qquad \forall v \in H^1(\Omega), \tag{1}$$

with

$$\begin{aligned} a_u(y,v) &:= (\nabla y, \nabla v)_{L^2(\Omega)} + (y,v)_{L^2(\Omega)} + (u\, y, v)_{L^2(\Gamma)}, \\ F(v) &:= (f,v)_{L^2(\Omega)} + (g,v)_{L^2(\Gamma)}. \end{aligned}$$

First, we show an existence and uniqueness result for (1). Therefore, we introduce a decomposition of the control into positive and negative parts $u^+, u^- \in L^2_+(\Gamma) := \{v \in L^2(\Gamma) : v \geq 0 \text{ a. e. on } \Gamma\}$ such that $u = u^+ - u^-$. The following result then relies on the Lax–Milgram–Lemma. However, an assumption on the coefficient u is required.

Lemma 1 *Assume that $u \in L^2(\Gamma)$ satisfies*

$$\|u^-\|_{L^2(\Gamma)} < \frac{1}{c_*^2} \tag{2}$$

with the constant c^* which is due to the estimate $\|v\|_{L^4(\Gamma)} \leq c^* \|v\|_{H^1(\Omega)}$. Then, the solution y of (1) *belongs to $H^1(\Omega)$ and satisfies the a priori estimate*

$$\|y\|_{H^1(\Omega)} \leq \frac{1}{\gamma_u} \left(\|f\|_{H^1(\Omega)^*} + \|g\|_{H^{-1/2}(\Gamma)} \right)$$

with $\gamma_u := 1 - c_*^2 \, \|u^-\|_{L^2(\Gamma)} > 0$.

Proof The boundedness of a_u follows directly from the Cauchy–Schwarz inequality and the continuity of the trace operator $\tau : H^1(\Omega) \to L^4(\Gamma)$. This implies

$$\begin{aligned} a(y,z) &\leq \|y\|_{H^1(\Omega)} \, \|z\|_{H^1(\Omega)} + \|u\|_{L^2(\Gamma)} \, \|y\|_{L^4(\Gamma)} \, \|z\|_{L^4(\Gamma)} \\ &\leq \left(1 + c_*^2 \, \|u\|_{L^2(\Gamma)}\right) \|y\|_{H^1(\Omega)} \, \|z\|_{H^1(\Omega)}. \end{aligned}$$

To show the coercivity we take into account the decomposition $u = u^+ - u^-$ to get

$$a(y,y) \geq \|y\|^2_{H^1(\Omega)} - \int_\Gamma u^- y^2 \geq \left(1 - c_*^2 \, \|u^-\|_{L^2(\Gamma)}\right) \|y\|^2_{H^1(\Omega)}.$$

Here, the assumption (2) will ensure the coercivity. An application of the Lax–Milgram Lemma leads to the desired result. □

Note that $\{v \in L^2(\Gamma) : \|v^-\|_{L^2(\Omega)} < c_*^{-2}\}$ is an open subset of $L^2(\Gamma)$. This is the key idea which allows us to avoid the two-norm discrepancy for the optimal control problem as we will see that the reduced objective functional is differentiable with respect to the $L^2(\Gamma)$-topology. In the following we will hide the dependency of the estimates on $\|u^-\|_{L^2(\Gamma)}$ and thus γ_u in the generic constant as we impose positive control bounds in the considered optimal control problem.

Later, we will frequently make use of the following Lipschitz estimate.

Lemma 2 *If $u_1, u_2 \in L^2(\Gamma)$ satisfy the assumption (2), the corresponding states $y_1, y_2 \in H^1(\Omega)$ solving*

$$a_{u_i}(y_i, v) = (f_i, v)_{L^2(\Omega)} + (g_i, v)_{L^2(\Gamma)} \quad \forall v \in H^1(\Omega),\ i = 1, 2,$$

fulfill the estimate

$$\|y_1 - y_2\|_{H^1(\Omega)} \le \big(\|u_1 - u_2\|_{L^2(\Gamma)} \|y_2\|_{H^1(\Omega)} \\ + \|f_1 - f_2\|_{H^1(\Omega)^*} + \|g_1 - g_2\|_{H^{-1/2}(\Gamma)} \big).$$

Proof Subtracting the variational formulations for y_1 and y_2 from each other leads to

$$(\nabla(y_1 - y_2), \nabla v)_{L^2(\Omega)} + (y_1 - y_2, v)_{L^2(\Omega)} + (u_1\,(y_1 - y_2), v)_{L^2(\Gamma)} \\ = (f_1 - f_2, v)_{L^2(\Omega)} + (g_1 - g_2, v)_{L^2(\Gamma)} + ((u_2 - u_1)\, y_2, v)_{L^2(\Gamma)}.$$

The result follows from Lemma 1 and the continuity of the product mapping from $L^2(\Gamma) \times H^{1/2}(\Gamma)$ to $H^{-1/2}(\Gamma)$, see [20, Theorem 1.4.4.2]. □

In the following theorem we collect some regularity results for the solution of (1).

Lemma 3 *Let $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, be a bounded Lipschitz domain. By $y \in H^1(\Omega)$ we denote the solution of (1). The following a priori estimates are valid, under the assumption that the input data possess the regularity demanded by the right-hand side:*

(a) *If $r > 2n/(1+n)$ and $p > 2$ for $n = 2$ and $p \ge 4$ for $n = 3$, then*

$$\|y\|_{H^{3/2}(\Omega)} + \|y\|_{H^1(\Gamma)} \le c\big(1 + \|u\|_{L^p(\Gamma)}\big)\big(\|f\|_{L^r(\Omega)} + \|g\|_{L^2(\Gamma)}\big).$$

(b) If $r > n/2,\ s > n - 1,$ *and $p \ge 2$ for $n = 2$ and* $p > 8/3$ for $n = 3$, *then*

$$\|y\|_{C(\overline{\Omega})} \le c\big(1 + \|u\|_{L^p(\Gamma)}\big)^2 \big(\|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Gamma)}\big).$$

(c) *Furthermore, if Ω is a convex polygonal/polyhedral domain, or possesses a boundary which is of class $C^{1,1}$, there holds*

$$\|y\|_{H^2(\Omega)} \le c\big(1 + \|u\|_{H^{1/2}(\Gamma)}\big)^2 \big(\|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\Gamma)}\big).$$

Proof (a)
In [15, Theorem 1.12] it is shown that the problem

$$-\Delta y = F \ \text{ in } \ \Omega, \qquad \partial_n y = G \ \text{ on } \ \Gamma$$

possesses a solution in $H^{3/2}(\Omega)$ provided that $F \in H^{s-2}(\Omega)$ for some $s \in (3/2, 2]$ and $G \in L^2(\Gamma)$, as well as $\int_\Omega F + \int_\Gamma G = 0$. The solubility condition is satisfied in our situation with $F = f - y$ and $G = g - u\,y$ and becomes clear when testing (1) with $v \equiv 1$. The regularity required for F follows from the embedding

$f \in L^r(\Omega) \hookrightarrow H^{-1/2+\varepsilon}(\Omega)$ for sufficiently small $\varepsilon > 0$. Moreover, the Hölder inequality and the continuity of the trace operator $\tau : H^1(\Omega) \to L^q(\Gamma)$ for $q < \infty$ ($n = 2$) or $q \le 4$ ($n = 3$) imply $\|u\,y\|_{L^2(\Gamma)} \le c\,\|u\|_{L^p(\Gamma)}\,\|y\|_{H^1(\Omega)}$, from which we conclude $G \in L^2(\Gamma)$. From [15, Theorem 1.12] and Lemma 1 we then obtain

$$\begin{aligned}\|y\|_{H^{3/2}(\Omega)} &\le c\left(\|F\|_{L^r(\Omega)} + \|G\|_{L^2(\Gamma)} + \left|\int_\Omega y(x)\mathrm{d}x\right|\right)\\ &\le c\big(1 + \|u\|_{L^p(\Gamma)}\big)\big(\|f\|_{L^r(\Omega)} + \|g\|_{L^2(\Gamma)}\big).\end{aligned} \tag{3}$$

It remains to show the $H^1(\Gamma)$-norm estimate. We split the solution into the parts y_f and y_g solving

$$\begin{aligned} -\Delta y_f + y_f &= f & -\Delta y_g + y_g &= 0 & &\text{in } \Omega,\\ \partial_n y_f &= 0 & \partial_n y_g &= g - uy & &\text{on } \Gamma. \end{aligned}$$

Using [19, Theorem 5.4] we directly deduce

$$\|y_g\|_{H^1(\Gamma)} \le c\,\|g - u\,y\|_{L^2(\Gamma)} \le c\big(\|g\|_{L^2(\Gamma)} + \|u\|_{L^p(\Gamma)}\,\|y\|_{H^1(\Omega)}\big)$$

and Lemma 1 leads to the desired estimate for y_g. For the function y_f, we get the desired estimate by an application of a trace theorem and the a priori estimate (3) which can in case of $g \equiv 0$ be improved to

$$\|y_f\|_{H^1(\Gamma)} \le c\,\|y_f\|_{H^{3/2+\varepsilon}(\Omega)} \le c\,\|f\|_{L^r(\Omega)},$$

provided that $\varepsilon > 0$ is sufficiently small. The validity of the second step can be confirmed by means of [15, Theorem 1.12] and [14, Theorem 23.3]. The decomposition $y = y_f + y_g$ and the estimates shown above imply the desired estimate in the $H^1(\Gamma)$ -norm.

(b) We prove the result for the case $n = 3$. The two-dimensional case follows from the same arguments. From [8, Theorem 3.1] it is known that the solution of (1) belongs to $C(\overline{\Omega})$ if $f \in L^r(\Omega)$, $r > n/2$, and $g - uy \in L^s(\Gamma)$, $s > n - 1$. The latter assumption can be concluded from the Hölder inequality, a Sobolev embedding and a trace theorem, which implies

$$\|u\,y\|_{L^s(\Gamma)} \le c\,\|u\|_{L^p(\Gamma)}\,\|y\|_{L^8(\Gamma)} \le c\,\|u\|_{L^p(\Gamma)}\,\|y\|_{H^{5/4+\varepsilon}(\Omega)}$$

for $1/p + 1/8 = 1/(2 + \varepsilon)$. A simple computation shows that $p > 8/3$ and $s = 2 + \varepsilon$ with $\varepsilon > 0$ sufficiently small guarantee the validity of the previous steps. It remains to show $y \in H^{5/4+\varepsilon}(\Omega)$. This can be deduced from [14, Theorem 23.3] where the a priori estimate

$$\|y\|_{H^{5/4+\varepsilon}(\Omega)} \le c\big(\|f\|_{H^{3/4-\varepsilon}(\Omega)^*} + \|g - u\,y\|_{H^{-1/4+\varepsilon}(\Gamma)}\big) \tag{4}$$

is stated. The regularity demanded by the right-hand side of (4) is confirmed with the embeddings $f \in L^r(\Omega) \hookrightarrow H^{3/4-\varepsilon}(\Omega)^*$ and $g \in L^s(\Gamma) \hookrightarrow H^{-1/4+\varepsilon}(\Gamma)$. Moreover, there holds $\|u\,y\|_{H^{-1/4+\varepsilon}(\Gamma)} \le c\,\|u\|_{L^p(\Gamma)}\,\|y\|_{L^4(\Gamma)}$, see [20, Theorem 1.4.4.2]. Collecting up the arguments above leads to

$$\begin{aligned}\|y\|_{C(\overline{\Omega})} &\le c\big(\|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Gamma)} + \|u\|_{L^p(\Gamma)}\|y\|_{H^{5/4+\varepsilon}(\Omega)}\big)\\ &\le c\big(1 + \|u\|_{L^p(\Gamma)}\big)^2\big(\|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Omega)} + \|y\|_{H^1(\Omega)}\big)\end{aligned}$$

and the assertion follows after insertion of the a priori estimate from Lemma 1.

(c) With an embedding we deduce from the assumption that $u \in L^4(\Gamma)$. Hence, (4) is applicable which implies $y \in H^{3/4}(\Gamma)$ and thus, $u\,y \in H^{1/2}(\Gamma)$, see [20, Theorem 1.4.4.2]. The $H^2(\Omega)$-regularity of y then follows from a shift theorem applied to the equation with boundary conditions $\partial_n y = g - uy \in H^{1/2}(\Gamma)$ on Γ, see [20, Theorem 2.4.2.7] (for domains with smooth boundary) or [20, Theorem 4.4.3.8] (for convex polygonal domains). □

3 The optimal control problem

Due to the well-posedness of the state equation we may introduce the control-to-state operator $S : U_{ad} \to H^1(\Omega)$ defined by $S(u) := y$, with y solving (1). This allows to reformulate the optimal control problem introduced in Sect. 1 and we arrive at

$$j(u) := \frac{1}{2}\|S(u) - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Gamma)} \to \min! \tag{5}$$

subject to $u \in U_{ad} := \{v \in L^2(\Gamma) : u_a \le v \le u_b \text{ a. e. on } \Gamma\}$. Here, $\alpha > 0$ is the regularization parameter, $y_d \in L^2(\Omega)$ the desired state and $0 < u_a < u_b$ the control bounds. Our aim is to derive necessary and sufficient optimality conditions as well as regularity results for local solutions. Note, that the operator S is non-affine and consequently, j is non-convex. The existence of at least one local solution can be concluded from standard arguments [39], taking into account that for a minimizing sequence $\{u_n\} \subset L^q(\Gamma)$, $q \in (2, \infty)$, the corresponding states converge strongly in $L^p(\Gamma)$ for each $p < 4$, which is due the compact embedding $H^1(\Omega) \hookrightarrow L^p(\Gamma)$.

3.1 Optimality conditions

To derive optimality conditions differentiability properties of the (implicitly defined) operator S are of interest.

Lemma 4 *The operator $S : U_{ad} \to H^1(\Omega)$ is infinitely many times Fréchet differentiable with respect to the $L^2(\Gamma)$-topology. The first derivative $\delta y := S'(u)\delta u$ is the weak solution of the tangent equation*

$$\begin{cases} -\Delta\delta y + \delta y = 0 & \text{in } \Omega,\\ \partial_n \delta y + u\,\delta y = -\delta u\, y & \text{on } \Gamma.\end{cases} \tag{6}$$

Proof The result follows from an application of the implicit function theorem to the operator $e : H^1(\Omega) \times U \to H^1(\Omega)^*$ with $U := \{v \in L^2(\Gamma) : v \text{ fulfills } (2)\}$ defined by

$$e(y,u)v := (\nabla y, \nabla v)_{L^2(\Omega)} + (y,v)_{L^2(\Omega)} + (u\,y,v)_{L^2(\Gamma)} - (f,v)_{L^2(\Omega)} - (g,v)_{L^2(\Omega)},$$

whose roots are solutions of (1). We choose $\delta y \in H^1(\Omega)$, $\delta u \in U$ such that $u + \delta u \in U$ (note that U is an open subset of $L^2(\Gamma)$). First, we confirm that the linear operator $e'(y,u) : H^1(\Omega) \times U \to H^1(\Omega)^*$ defined by

$$e'(y,u)(\delta y, \delta u) := (\nabla \delta y, \nabla \cdot)_{L^2(\Omega)} + (\delta y, \cdot)_{L^2(\Omega)} + (u\,\delta y + y\,\delta u, \cdot)_{L^2(\Gamma)} \tag{7}$$

is the Fréchet-derivative of e. This is a consequence of

$$e(y+\delta y, u+\delta u) - e(y,u) = e'(y,u)(\delta y, \delta u) + (\delta u\,\delta y, \cdot)_{L^2(\Gamma)}$$

and the fact that the remainder term satisfies

$$\begin{aligned}
\|\delta u\,\delta y\|_{H^1(\Omega)^*} &= \sup_{\varphi \in H^1(\Omega)} \frac{(\delta u\,\delta y, \varphi)_{L^2(\Gamma)}}{\|\varphi\|_{H^1(\Omega)}} \\
&\le c \sup_{\varphi \in H^1(\Omega)} \frac{\|\varphi\|_{L^4(\Gamma)}}{\|\varphi\|_{H^1(\Omega)}} \|\delta u\|_{L^2(\Gamma)} \|\delta y\|_{L^4(\Gamma)} \\
&\le c\,\|\delta u\|_{L^2(\Gamma)} \|\delta y\|_{H^1(\Omega)} \le c\Big(\|\delta u\|^2_{L^2(\Gamma)} + \|\delta y\|^2_{H^1(\Omega)}\Big) \\
&= o(\|(\delta y, \delta u)\|_{H^1(\Omega)\times L^2(\Gamma)}),
\end{aligned} \tag{8}$$

where we applied the generalized Hölder inequality and $H^1(\Omega) \hookrightarrow L^4(\Gamma)$. The second Fréchet derivative

$$e'' : H^1(\Omega) \times U \to \mathcal{L}((H^1(\Omega) \times L^2(\Gamma))^2, H^1(\Omega)^*)$$

is given by

$$e''(y,u)(\delta y, \delta u)(\tau y, \tau u) := (\tau u\,\delta y + \delta u\,\tau y, \cdot)_{L^2(\Gamma)}$$

and the mapping $(y,u) \mapsto e''(y,u)$ is continuous. The derivatives of order $n \ge 3$ vanish. Hence, $e : H^1(\Omega) \times U \to H^1(\Omega)^*$ is of class C^∞.

Finally, due to Lemma 1 we conclude that the linear mapping

$$\delta y \mapsto e_y(y,u)\delta y = (\nabla \delta y, \nabla \cdot)_{L^2(\Omega)} + (\delta y, \cdot)_{L^2(\Omega)} + (u\,\delta y, \cdot)_{L^2(\Gamma)} \in H^1(\Omega)^*$$

is bijective. The implicit function theorem implies the assertion and the derivative $\delta y := S'(u)\delta u$ is given by $e'(y,u)(\delta y, \delta u) = 0$. This corresponds to the weak formulation of (6). □

From the chain rule and Lemma 4 we directly conclude the following differentiability result:

Lemma 5 *The functional $j : U_{ad} \to \mathbb{R}$ is infinitely many times Fréchet differentiable with respect to the $L^2(\Gamma)$-topology and the first derivative is given by*

$$\left\langle j'(u), v \right\rangle = (S(u) - y_d, S'(u)v)_{L^2(\Omega)} + \alpha\,(u, v)_{L^2(\Gamma)}, \qquad v \in L^2(\Gamma). \tag{9}$$

The derivative of j can be simplified exploiting a precise representation of the adjoint $S'(u)^* : H^1(\Omega)^* \to L^2(\Gamma)$ of the linearized control-to-state operator $S'(u)$. In order to compute this, we introduce the the adjoint state $p \in H^1(\Omega)$ as the weak solution of the adjoint equation

$$\begin{cases} -\Delta p + p = y - y_d & \text{in } \Omega, \\ \partial_n p + u\,p = 0 & \text{on } \Gamma. \end{cases} \tag{10}$$

Testing the variational problems for (10) and (6) with $\delta y := S'(u)\delta u$ and p, respectively, leads to the relation

$$(y - y_d, \delta y)_{L^2(\Omega)} = -(y\,p, \delta u)_{L^2(\Gamma)},$$

which implies

$$S'(u)^*(y - y_d) := -[y\,p]_\Gamma. \tag{11}$$

In the following we denote the control-to-adjoint mapping $Z : L^2(\Gamma) \to H^1(\Omega)$ defined by $u \mapsto Z(u) := p$ via (10) with $y = S(u)$. Finally, we are able to formulate the necessary optimality condition

$$\left\langle j'(u), v - u \right\rangle = (S(u) - y_d, S'(u)(v - u))_{L^2(\Omega)} + \alpha\,(u, v - u)_{L^2(\Gamma)} \geq 0 \quad \forall v \in U_{ad}$$

and with (11) we get the equivalent representation

$$\left\langle j'(u), v - u \right\rangle = (\alpha\,u - S(u)\,Z(u), v - u)_{L^2(\Gamma)} \geq 0 \quad \forall v \in U_{ad}.$$

Taking into account the definitions of S and Z we can write this variational inequality in the form

$$\begin{array}{lll} -\Delta y + y = f & -\Delta p + p = y - y_d & \text{in } \Omega, \\ \partial_n y + u\,y = g & \partial_n p + u\,p = 0 & \text{on } \Gamma, \\ \quad (\alpha\,u - y\,p, v - u)_{L^2(\Gamma)} \geq 0 & & \text{for all } v \in U_{ad}. \end{array} \tag{12}$$

The latter inequality is equivalent to the projection formula

$$u = \Pi_{ad}\left(\frac{1}{\alpha}[y\,p]_\Gamma\right) \tag{13}$$

with Π_{ad} the $L^2(\Gamma)$-projection onto U_{ad}.

As the problem (5) is not convex, we have to investigate second-order sufficient conditions. To obtain the Hessian of j we apply the product rule and get

$$j''(u)(\delta u, \tau u) = \big(\alpha\,\tau u + S'(u)\tau u\,Z(u) + S(u)\,Z'(u)\tau u, \delta u\big)_{L^2(\Gamma)}. \tag{14}$$

The function $\tau p = Z'(u)\tau u \in H^1(\Omega)$ is the weak solution of the *"dual for Hessian"-equation*

$$
\begin{cases} -\Delta \tau p + \tau p = \tau y & \text{in } \Omega, \\ \partial_n \tau p + u\, \tau p = -\tau u\, p & \text{on } \Gamma, \end{cases} \tag{15}
$$

where $\tau y = S'(u)\tau u$. As in the proof of Lemma 3 this follows from the implicit function theorem. Note that also further representations of the Hessian are possible. For instance, a direct application of the product rule to (9) yields

$$
\begin{aligned}
j''(u)(\delta u, \tau u) &= \alpha(\delta u, \tau u)_{L^2(\Gamma)} + (S'(u)\tau u, S'(u)\delta u)_{L^2(\Omega)} \\
&\quad + (S(u) - y_d, S''(u)(\delta u, \tau u))_{L^2(\Omega)} \\
&= \alpha(\delta u, \tau u)_{L^2(\Gamma)} + (\delta y, \tau y)_{L^2(\Omega)} + (y - y_d, \delta\tau y)_{L^2(\Omega)},
\end{aligned}
$$

with $y = S(u)$, $\delta y = S'(u)\delta y$, $\tau y = S'(u)\tau u$ and $\delta\tau y = S''(u)(\delta u, \tau u)$. The latter relation means that $\delta\tau y \in H^1(\Omega)$ is the weak solution of

$$
\begin{cases} -\Delta \delta\tau y + \delta\tau y = 0 & \text{in } \Omega, \\ \partial_n \delta\tau y + u\, \delta\tau y = -\delta u\, \tau y - \tau u\, \delta y & \text{on } \Gamma. \end{cases}
$$

Moreover, due to the definition of p and $\delta\tau y$ there holds the relation $(y - y_d, \delta\tau y)_{L^2(\Omega)} = -(p, \delta u \tau y + \tau u\, \delta y)_{L^2(\Gamma)}$ and as a consequence, we can further simplify the representation of the Hessian and obtain

$$
j''(u)(\delta u, \tau u) = \alpha(\delta u, \tau u)_{L^2(\Gamma)} + (\delta y, \tau y)_{L^2(\Omega)} - (p, \delta u\, \tau y + \tau u\, \delta y)_{L^2(\Gamma)}.
$$

Next, we derive some stability and Lipschitz properties of S, Z, S' and Z'. As the following results require different assumptions on f, y_d and g we simply assume the most restrictive ones, this is,

$$
f, y_d \in L^\infty(\Omega), \qquad g \in H^{1/2}(\Gamma).
$$

Moreover, we will hide the dependency on these quantities in the generic constant to simplify the notation.

Lemma 6 *Let $u \in L^2(\Gamma)$ satisfy the assumption (2). The control-to-state operator S satisfies the following inequalities:*

$$
\begin{aligned}
\|S(u)\|_{H^1(\Omega)} &\le c, \\
\|S(u)\|_{H^{3/2}(\Omega)} + \|S(u)\|_{H^1(\Gamma)} &\le c\,(1 + \|u\|_{L^{p_1}(\Gamma)}), \\
\|S(u)\|_{L^\infty(\Omega)} &\le c\,(1 + \|u\|_{L^{p_2}(\Gamma)})^2,
\end{aligned}
$$

with $p_1 > 2$ and $p_2 \ge 2$ for $n = 2$, and $p_1 \ge 4$ and $p_2 > 8/3$ for $n = 3$. The estimates remain valid when replacing the operator S by the control-to-adjoint operator Z.

Proof The inequalities for S are a direct consequence of Lemmata 1 and 3. The inequalities for Z can be derived with similar arguments, but the right-hand side of the adjoint equation involves the corresponding state $S(u)$. However, in all cases the norms of $S(u) - y_d$ can be bounded by $c\,(1 + \|S(u)\|_{H^1(\Omega)}) \le c$. □

Lemma 7 *Given are $u, \delta u \in L^2(\Gamma)$ and it is assumed that u satisfies* (2). *Then, the following stability estimates hold true:*

$$\|S'(u)\delta u\|_{H^1(\Omega)} \le c\,\|\delta u\|_{L^2(\Gamma)},$$
$$\|S'(u)\delta u\|_{H^{3/2}(\Omega)} \le c\big(1 + \|u\|_{L^p(\Gamma)}\big)^3 \|\delta u\|_{L^2(\Gamma)},$$

with $p > 2$ for $n = 2$ and $p \ge 4$ for $n = 3$. The estimates remain valid when replacing S' by Z'.

Proof In the following we write $y := S(u)$ and $\delta y = S'(u)\delta u$. The stability in $H^1(\Omega)$ follows directly from Lemma 1 and the estimate

$$\|\delta u\, y\|_{H^{-1/2}(\Gamma)} = \sup_{\substack{\varphi \in H^{1/2}(\Gamma)\\ \varphi \not\equiv 0}} \frac{(\delta u\, y, \varphi)_{L^2(\Gamma)}}{\|\varphi\|_{H^{1/2}(\Gamma)}} \le c\,\|\delta u\|_{L^2(\Gamma)}\,\|y\|_{H^1(\Omega)}, \tag{16}$$

which follows from the same arguments used already in (8). The boundedness of $y := S(u)$ in $H^1(\Omega)$ can be found in the previous lemma. The estimate in the $H^{3/2}(\Omega)$ -norm follows analogously with Lemma 3a) and

$$\|y\,\delta u\|_{L^2(\Gamma)} \le c\,\|y\|_{L^\infty(\Omega)}\,\|\delta u\|_{L^2(\Gamma)}$$

and the stability in $L^\infty(\Omega)$ proved in Lemma 6.

The estimates for Z' are deduced with similar techniques. With the a priori estimate from Lemma 3a) and the embedding $H^1(\Omega) \hookrightarrow L^r(\Omega)$ which holds for $r < \infty$ ($n = 2$) or $r \le 6$ ($n = 3$) we get

$$\begin{aligned}\|Z'(u)\delta u\|_{H^{3/2}(\Omega)} &\le c\big(1 + \|u\|_{L^p(\Gamma)}\big)\big(\|p\,\delta u\|_{L^2(\Gamma)} + \|\delta y\|_{H^1(\Omega)}\big)\\ &\le c\big(1 + \|u\|_{L^p(\Gamma)}\big)\big(1 + \|p\|_{L^\infty(\Gamma)}\big)\|\delta u\|_{L^2(\Gamma)}\end{aligned}$$

with $p = Z(u)$. The stability of Z in $L^\infty(\Omega)$ is discussed in the previous lemma. □

Lemma 8 *Let $u, v \in L^2(\Gamma)$ satisfy assumption* (2). *Then, the following Lipschitz estimates hold:*

$$\|S(u) - S(v)\|_{H^1(\Omega)} \le c\,\|u - v\|_{L^2(\Gamma)},$$
$$\|S'(u)\delta u - S'(v)\delta u\|_{H^1(\Omega)} \le c\,\|u - v\|_{L^2(\Gamma)}\|\delta u\|_{L^2(\Gamma)}.$$

The estimates are also valid when replacing S by Z and Z by Z'.

Proof The estimates for S and S' follow directly from Lemma 2 and the stability estimates for S and S' in $H^1(\Omega)$ proved in the Lemmata 6 and 7. The Lipschitz estimate for Z is proved in a similar way. In this case one has to apply the Lipschitz estimate shown for S to the term $\|S(u) - S(v)\|_{H^1(\Omega)}$ appearing due to the differences in the right-hand sides. With the same idea we show the Lipschitz estimate for Z'. Using again Lemma 2 we get

$$\|Z'(u)\delta u - Z'(v)\delta u\|_{H^1(\Omega)} \le c\Big(\|u-v\|_{L^2(\Gamma)}\,\|Z'(u)\delta u\|_{H^1(\Omega)}$$
$$+\,\|S'(u)\delta u - S'(v)\delta v\|_{H^1(\Omega)} + \|\delta u\,(Z(u)-Z(v))\|_{H^{-1/2}(\Gamma)}\Big).$$

It remains to bound the three terms on the right-hand side. To this end, we apply Lemma 7 to the first term, the Lipschitz estimate for $S'(\cdot)\delta u$ to the second term, and the multiplication rule (16) with $y = Z(u) - Z(v)$ as well as the Lipschitz estimate for Z to the third term. □

As the optimal control problem is non-convex we have to deal with local solutions. For some local solution $\bar{u} \in U_{ad}$ we require the following second-order sufficient condition:

Assumption 1 (*SSC*) The objective functional is locally convex near the local solution $\bar{u}$, i. e., a constant $\delta > 0$ exists such that

$$j''(\bar{u})(v,v) \ge \delta\,\|v\|^2_{L^2(\Gamma)} \qquad \forall v \in L^2(\Gamma). \tag{17}$$

With standard arguments and the estimate we will prove below in Corollary 1 one can show that each function $\bar{u} \in U_{ad}$ fulfilling the first-order necessary condition (12) and the second-order sufficient condition (17) is indeed a local solution and satisfies the quadratic growth condition

$$j(\bar{u}) \le j(u) - \gamma\,\|u-\bar{u}\|^2_{L^2(\Gamma)} \qquad \forall u \in B_\tau(\bar{u}),$$

with certain constants $\gamma, \tau > 0$. Note that there are weaker assumptions which are sufficient for local minima, for instance one could formulate (17) for all directions v from a critical cone. However, with this assumption the convergence proof for the postprocessing approach presented in Sect. 5.3 requires some more careful investigations, in particular the construction of a modified interpolant onto U_{ad}. One possible solution for this issue can be found in [29].

Later, we will require the following Lipschitz estimate for the Hessian of j.

Lemma 9 *Let $u, v \in L^2(\Gamma)$ fulfilling (2) be given. Then, the Lipschitz-estimate*

$$\left|j''(u)(\delta u,\delta u) - j''(v)(\delta u,\delta u)\right| \le c\,\|\delta u\|^2_{L^2(\Gamma)}\,\|u-v\|_{L^2(\Gamma)}.$$

is valid for all $\delta u \in L^2(\Gamma)$.

Proof To shorten the notation we write $y_u = S(u)$, $p_u = Z(u)$, $\delta y_u = S'(u)\delta u$ and $\delta p_u = Z'(u)\delta u$. From the representation (14) we obtain

$$\left|j''(v)(\delta u,\delta u) - j''(u)(\delta u,\delta u)\right|$$
$$\le \left|(p_u\,\delta y_u - p_v\,\delta y_v + y_u\,\delta p_u - y_v\,\delta p_v, \delta u)_{L^2(\Gamma)}\right|.$$

We estimate the right-hand side using the Cauchy–Schwarz inequality, the embedding $H^1(\Omega) \hookrightarrow L^4(\Gamma)$ and the Lipschitz estimates from Lemma 8 as well as the a priori estimates from Lemmata 6 and 7. This implies

$$\begin{aligned}
&\left|(p_u\, \delta y_u - p_v\, \delta y_v, \delta u)_{L^2(\Gamma)}\right| \\
&\quad \le c\big(\|p_u - p_v\|_{H^1(\Omega)}\, \|\delta y_u\|_{H^1(\Omega)} + \|\delta y_u - \delta y_v\|_{H^1(\Omega)}\, \|p_v\|_{H^1(\Omega)}\big)\, \|\delta u\|_{L^2(\Gamma)} \\
&\quad \le c\, \|u - v\|_{L^2(\Gamma)}\, \|\delta u\|^2_{L^2(\Gamma)}.
\end{aligned}$$

With similar arguments we deduce

$$\begin{aligned}
&\left|\big(y_u\, \delta p_u - y_v\, \delta p_v, \delta u\big)_{L^2(\Gamma)}\right| \\
&\quad \le c\big(\|y_u - y_v\|_{H^1(\Omega)}\, \|\delta p_u\|_{H^1(\Omega)} + \|\delta p_u - \delta p_v\|_{H^1(\Omega)}\, \|y_v\|_{H^1(\Omega)}\big)\, \|\delta u\|_{L^2(\Gamma)} \\
&\quad \le c\, \|u - v\|_{L^2(\Gamma)}\, \|\delta u\|^2_{L^2(\Gamma)},
\end{aligned}$$

and conclude the assertion. □

Corollary 1 *Let $\bar u \in U_{ad}$ be a local solution of (5) satisfying Assumption 1. Then, some $\varepsilon > 0$ exists such that the inequality*

$$j''(u)(\delta u, \delta u) \ge \frac{\delta}{2}\|\delta u\|^2_{L^2(\Gamma)}$$

is valid for all $\delta u \in L^2(\Gamma)$ and $u \in L^2(\Gamma)$ with $\|u - \bar u\|_{L^2(\Gamma)} \le \varepsilon$.

Proof The assertion follows immediately from the previous lemma. For further details we refer to [27, Lemma 2.23]. □

In the next Lemma we will collect some basic regularity results for the solution of (5).

Lemma 10 *Let $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, be a Lipschitz domain. Each local solution $\bar u \in U_{ad}$ of (5) and the corresponding states $\bar y = S(\bar u)$, $\bar p = Z(\bar u)$ satisfy*

$$\bar u \in H^1(\Gamma) \cap L^\infty(\Gamma), \qquad \bar y, \bar p \in H^{3/2}(\Omega) \cap H^1(\Gamma) \cap C(\overline{\Omega}).$$

Proof All regularity result, except $\bar u \in H^1(\Gamma)$, follow directly from Lemma 3. To show $\bar u \in H^1(\Gamma)$ we apply the product rule

$$\|\bar y\, \bar p\|_{H^1(\Gamma)} \le c\big(\|\bar y\|_{H^1(\Gamma)}\, \|\bar p\|_{L^\infty(\Omega)} + \|\bar y\|_{L^\infty(\Omega)}\, \|\bar p\|_{H^1(\Gamma)}\big) \le c$$

and confirm $\bar y\, \bar p \in H^1(\Gamma)$. The desired result then follows after an application of the Stampacchia-Lemma, [26, p. 50], to the projection formula (13). The fact that the Stampacchia-Lemma is also valid on the boundary Γ is discussed in [28, Lemma 2.8] and [30, Lemma 3.3]. □

Under additional assumptions on the geometry of Ω we can show even higher regularity. This is needed for the postprocessing approach studied in Sect. 5.3 where we will show almost quadratic convergence of the control approximations.

Lemma 11 *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a $C^{1,1}$-boundary Γ. Then, there holds*

$$\bar{u} \in W^{1,q}(\Gamma) \cap H^{2-1/q}(\tilde{\Gamma}), \qquad \bar{y}, \bar{p} \in W^{2,q}(\Omega), \qquad q < \infty,$$

for all $\tilde{\Gamma} \subset\subset \mathcal{A}$ or $\tilde{\Gamma} \subset\subset \mathcal{I}$, where $\mathcal{A} := \{x \in \Gamma : u(x) \in \{u_a, u_b\}\}$ and $\mathcal{I} := \Gamma \setminus \mathcal{A}$ denote the active and inactive set, respectively.

Proof With the regularity results obtained already in Lemma 10, in particular $\bar{u} \in H^{1/2}(\Gamma)$, and Lemma 3c) we conclude $\bar{y}, \bar{p} \in H^2(\Omega) \hookrightarrow W^{1,q}(\Gamma)$ for all $q < (1, \infty)$ and a further application of the multiplication rule yields $\bar{y}\bar{p} \in W^{1,q}(\Gamma)$. From (13) we conclude the property $\bar{u} \in W^{1,q}(\Gamma)$. Furthermore, we confirm that $\bar{u}\,\bar{y}, \bar{u}\,\bar{p} \in W^{1-1/q,q}(\Gamma)$ and a standard shift theorem for the Neumann problem, compare also the technique used in the proof of Lemma 3a), results in $\bar{y}, \bar{p} \in W^{2,q}(\Omega)$. Repeating the arguments above, i. e., using the multiplication rule and the projection formula, we obtain $\bar{u} \in W^{2-1/q,q}(\tilde{\Gamma}) \hookrightarrow H^{2-1/q}(\tilde{\Gamma})$. □

We chose the assumptions of the previous lemma in such a way that the regularity is only restricted due to the projection formula. Of course, when the control bounds are never active we could further improve the regularity results.

4 Finite element approximation of the state equation

This section is devoted to the finite element approximation of the variational problem (1). While the results from the previous sections are valid for arbitrary Lipschitz domains (unless otherwise explicitly assumed), we have to assume more smoothness of the boundary Γ in order to establish our discretization results:

(A1) The domain $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, possesses a Lipschitz continuous boundary Γ which is piecewise C^1.

This definition includes arbitrary (possibly non-convex) polygonal or polyhedral domains. Indeed, the regularity of solutions is in this case also restricted by corner and edge singularities. However, for the first convergence result we require only $H^{3/2}(\Omega) \cap H^1(\Gamma)$-regularity of the solution. Later, we want to investigate improved discretization techniques for which more regularity is needed. Then, we will use a stronger assumption on the domain.

First, we introduce shape-regular triangulations $\{\mathcal{T}_h\}_{h>0}$ of Ω consisting of triangles ($n = 2$) or tetrahedra ($n = 3$). The elements T may have curved edges/faces such that the property

$$\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} \overline{T}$$

is valid for an arbitrary domain Ω. Moreover, we assume that the triangulations are feasible in the sense of Ciarlet [13].

The mesh parameter $h > 0$ is the maximal element diameter

$$h = \max_{T\in\mathcal{T}_h} h_T, \quad h_T := \mathrm{diam}(T).$$

The family of meshes $\{\mathcal{T}_h\}_{h>0}$ is assumed to be quasi-uniform, this means some $\kappa > 0$ independent of h exists such that each element $T \in \mathcal{T}_h$ contains a ball with radius ρ_T satisfying the estimate $\frac{\rho_T}{h} \geq \kappa$. Each triangulation $\mathcal{T}_h$ of Ω induces also a triangulation $\mathcal{E}_h$ of the boundary Γ.

By $F_T : \hat{T} \to T$ we denote the transformations from the reference triangle or tetrahedron $\hat{T}$ to the world element $T \in \mathcal{T}_h$. The transformations F_T may be non-affine for elements with curved faces. Here, we consider transformations of the form

$$F_T = \tilde{F}_T + \Phi_T,$$

with some affine function $\tilde{F}_T(\hat{x}) = \tilde{B}_T\hat{x} + \tilde{b}_T$, $\tilde{B}_T \in \mathbb{R}^{n\times n}$, $\tilde{b} \in \mathbb{R}^n$, chosen in such a way that if T is a curved boundary element, $\tilde{T} = \tilde{F}_T(\hat{T})$ is an n-simplex whose vertices coincide with the vertices of T. The assumed shape-regularity implies $\|\tilde{B}_T\| \leq c\, h_T$ and $\|\tilde{B}_T^{-1}\| \leq h_T^{-1}$, see [13, Theorem 15.2].

To guarantee the validity of interpolation error estimates we assume:

(A2) The triangulations $\mathcal{T}_h$ are regular of order 2 in the sense of [6], this is, for all sufficiently small $h > 0$ there holds

$$\sup_{\hat{x}\in\hat{T}} \|D\Phi_T(\hat{x}) \cdot \tilde{B}_T^{-1}\| \leq c < 1, \qquad \sup_{\hat{x}\in\hat{T}} \|D^2\Phi_T(\hat{x})\| \leq ch^2, \tag{18}$$

for all $T \in \mathcal{T}_h$.

There are multiple strategies to construct the mappings F_T satisfying these assumptions and we refer the reader for instance to [6, 37, 41]. Therein, it is assumed that Γ is piecewise C^3, only in the second reference C^4 is required.

The trial and test space is defined by

$$V_h := \{v_h \in C(\overline{\Omega}) : v_h = \hat{v}_h \circ F_T^{-1},\ \hat{v}_h \in \mathcal{P}_1(\hat{T}) \text{ for all } T \in \mathcal{T}_h\}.$$

Next, we introduce an interpolation operator onto V_h. We partly use the quasi-interpolant proposed by Bernardi [6], but use a modification for boundary nodes as in [36], see also [1, 2]. To each interior node x_i, $i = 1, \ldots, N^{\mathrm{in}}$, of $\mathcal{T}_h$, we associate an element $\sigma_i := T \in \mathcal{T}_h$ with $x_i \in T$. For the boundary nodes x_i, $i = N^{\mathrm{in}} + 1, \ldots, N$, we define instead $\sigma_i := E \in \mathcal{E}_h$ with $x_i \in E$. Instead of using nodal values as for the Lagrange interpolant, we use the nodal values of some regularized function computed by an L^2-projection over σ_i. The interpolation operator $\Pi_h : W^{1,1}(\Omega) \to V_h$ is defined as follows. For each node $i = 1, \ldots, N$ we define a local L^2-projection $\hat{\pi}_i$ onto $\mathcal{P}_1(\sigma_i)$ by

$$\int_{\hat{\sigma}_i} (\hat{\pi}_i(u) - u\circ F_i)\, \hat{q} = 0 \qquad \forall \hat{q} \in \mathcal{P}_1,$$

with F_i the transformation from the reference element $\hat{T}$ $(i = 1, \dots, N^{\text{in}})$ or $\hat{E}$ $(i = N^{\text{in}} + 1, \dots, N)$ onto σ_i. The interpolation operator is defined by

$$\Pi_h v(x) = \sum_{i=1}^{N} (\hat{\pi}_i(u)\circ F_i^{-1})(x_i)\, \varphi_i(x),$$

where $\{\varphi_i\}_{i=1,\dots,N}$ is the nodal basis of V_h. Note, that due to the modification for boundary nodes, this operator is only applicable to $W^{1,1}(\Omega)$-functions. The desired interpolation properties remain valid. In particular, for each $T \in \mathcal{T}_h$, there holds

$$\|u - \Pi_h u\|_{H^m(T)} \le c h^{\ell - m} \|u\|_{H^\ell(S_T)}, \quad m \le \ell \le 2,\ \ell \ge 1, \tag{19}$$

where S_T is the patch of elements adjacent to T, see [6, Theorem 4.1], [36, Theorem 4.1]. Due to the special choice of the patches σ_i for the boundary nodes we get similar interpolation error estimates on the boundary elements $E \in \mathcal{E}_h$, this is,

$$\|u - \Pi_h u\|_{H^m(E)} \le c h^{\ell - m} \|u\|_{H^\ell(S_E)}, \quad m \le \ell \le 2, \tag{20}$$

with the patch S_E of that boundary elements $E' \in \mathcal{E}_h$ that touch E. The proof follows from the same arguments as in [36, Theorem 4.1].

The finite element solutions of (1) are characterized by the variational formulations

$$\text{Find } y_h \in V_h : \quad a_u(y_h, v_h) = F(v_h) \qquad \forall v_h \in V_h. \tag{21}$$

As in the continuous case one can show that (21) possesses a unique solution for each $h > 0$.

With the usual arguments we can derive an error estimate for the approximation error in the energy-norm.

Lemma 12 *Assume that* ***(A1)*** *and* ***(A2)*** *are satisfied and that the solution y of* (1) *belongs to $H^s(\Omega)$ with some $s \in [1, 2]$. Then, there holds the error estimate*

$$\|y - y_h\|_{H^1(\Omega)} \le c\, h^{s-1}\, \|y\|_{H^s(\Omega)}. \tag{22}$$

Proof The proof follows from the Céa-Lemma and the interpolation error estimates (19). □

Of particular interest are error estimates on the boundary. This is required in order to derive error estimates for boundary control problems. To this end, we prove first a suboptimal result which is valid for arbitrary Lipschitz domains Ω.

Lemma 13 *Let the assumptions **(A1)** and **(A2)** be satisfied. It is assumed that the solution y of (1) belongs to $H^{3/2}(\Omega)$. Moreover, the parameter u fulfills (2) and belongs to $L^p(\Gamma)$ with $p > 2$ for $n = 2$ and $p \geq 4$ for $n = 3$. Then, the error estimate*

$$\|y - y_h\|_{L^2(\Gamma)} \leq c\, h\big(1 + \|u\|_{L^p(\Gamma)}\big)\|y\|_{H^{3/2}(\Omega)} \leq c\, h\big(1 + \|u\|_{L^p(\Gamma)}\big)^2$$

holds, for all $h > 0$.

Proof We introduce the dual problem

$$\text{Find } w \in H^1(\Omega) : \quad a_u(v, w) = (y - y_h, v)_{L^2(\Gamma)} \qquad \forall v \in H^1(\Omega)$$

and obtain with the typical arguments of the Aubin-Nitsche trick

$$\begin{aligned}\|y - y_h\|^2_{L^2(\Gamma)} &\leq c\, \|y - y_h\|_{H^1(\Omega)}\, \|w - \Pi_h w\|_{H^1(\Omega)}\\ &\leq c\, h\, \|w\|_{H^{3/2}(\Omega)}\, \|y\|_{H^{3/2}(\Omega)}.\end{aligned}$$

The last step is an application of Lemma 12 and the interpolation error estimate (19). The regularity required for the dual solution w can be deduced from Lemma 3 with $f \equiv 0$ and $g = y - y_h$. Taking into account the a priori estimate

$$\|w\|_{H^{3/2}(\Omega)} \leq c\big(1 + \|u\|_{L^p(\Gamma)}\big)\|y - y_h\|_{L^2(\Gamma)}$$

we conclude the assertion. □

If the solution is more regular, we can also show a higher convergence rate. In this case we will use the Hölder inequality and a trace theorem to obtain $\|y - y_h\|_{L^2(\Gamma)} \leq \|y - y_h\|_{L^\infty(\Omega)}$, and insert the following result.

Theorem 2 *Consider a planar domain domain $\Omega \in \mathbb{R}^2$. Let $u \in H^{1/2}(\Gamma)$ with $u \geq 0$ a. e., and assume that **(A1)** and **(A2)** are satisfied. Assume that the solution y of (1) belongs to $y \in W^{2,q}(\Omega)$ with $q \in [2, \infty)$. Then, the error estimate*

$$\|y - y_h\|_{L^\infty(\Omega)} \leq c\, h^{2-2/q}\, |\ln h|\, \|y\|_{W^{2,q}(\Omega)}$$

is valid.

The proof requires rather technical arguments and is postponed to the appendix.

5 The discrete optimal control problem

In the following we investigate the discretized optimal control problem:

$$\text{Find } u_h \in U_h^{ad} : \quad J_h(y_h, u_h) := \frac{1}{2}\|y_h - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u_h\|^2_{L^2(\Gamma)} \to \min! \tag{23}$$

subject to

$$y_h \in V_h, \quad a_{u_h}(y_h, v_h) = F(v_h) \qquad \forall v_h \in V_h.$$

The reduced objective functional is denoted by $j_h(u_h) := J_h(S_h(u_h), u_h)$. We use piecewise linear finite elements to approximate the state y, i. e., the space V_h is defined as in the previous section. The controls are sought in the space of piecewise constant functions,

$$U_h^{ad} := \{w_h \in L^\infty(\Gamma) : w_h|_E \in \mathcal{P}_0 \quad \forall E \in \mathcal{E}_h\} \cap U_{ad},$$

where $\mathcal{E}_h$ is the triangulation of the boundary induced by $\mathcal{T}_h$.

As in the continuous case we can derive a first-order necessary optimality condition which reads

$$\begin{cases} a_{u_h}(y_h, v_h) = F(v_h) & \text{for all } v_h \in V_h, \\ a_{u_h}(v_h, p_h) = (y_h - y_d, v_h)_{L^2(\Omega)} & \text{for all } v_h \in V_h, \\ (\alpha\, u_h - y_h p_h, w_h - u_h)_{L^2(\Gamma)} \geq 0 & \text{for all } w_h \in U_h^{ad}. \end{cases} \tag{24}$$

The discrete control-to-state operator $S_h : L^2(\Gamma) \to V_h$ and the discrete control-to-adjoint operator $Z_h : L^2(\Gamma) \to V_h$ are defined by $y_h = S_h(u)$ and $p_h = Z_h(u)$ with

$$\begin{aligned} a_u(y_h, v_h) &= F(v_h) && \forall v_h \in V_h, \\ a_u(v_h, p_h) &= (y_h - y_d, v_h)_{L^2(\Omega)} && \forall v_h \in V_h. \end{aligned}$$

Analogous to the continuous case we compute the first and second derivatives of j_h and obtain

$$j_h'(u)\delta u = \left(\alpha\, u - S_h(u)\, Z_h(u), \delta u\right)_{L^2(\Gamma)} \tag{25}$$

and

$$j_h''(u)(\delta u, \tau u) = \left(\alpha\, \tau u - S_h(u)\, Z_h'(u)\tau u - S_h'(u)\tau u\, Z_h(u), \delta u\right)_{L^2(\Gamma)}, \tag{26}$$

where $\tau y_h = S_h'(u)\tau u \in V_h$ and $\tau p_h = Z_h'(u)\tau u \in V_h$ are the solutions of

$$\begin{aligned} a_u(\tau y_h, v_h) &= -(\tau u\, y_h, v_h)_{L^2(\Gamma)} && \text{for all } v_h \in V_h, \\ a_u(v_h, \tau p_h) &= (\tau y_h, v_h)_{L^2(\Omega)} - (\tau u\, p_h, v_h)_{L^2(\Gamma)} && \text{for all } v_h \in V_h, \end{aligned}$$

with $y_h = S_h(u)$ and $p_h = Z_h(u)$. These are the discretized versions of the equations (6) and (15). The first-order optimality condition reads in the short form

$$(\alpha\, u_h - S_h(u_h)\, Z_h(u_h), w_h - u_h)_{L^2(\Gamma)} \geq 0 \qquad \text{for all } w_h \in U_h^{ad}. \tag{27}$$

5.1 Properties of the discrete control-to-state/adjoint operator

In Sect. 3 we have derived several stability and Lipschitz properties for the operators S, Z, S' and Z'. Here, we will derive the discrete analogues that are needed in the following. Throughout this section we assume that **(A1)** and **(A2)** are fulfilled.

Lemma 14 *There hold the following properties:*

$$\|S_h(u)\|_{H^1(\Gamma)} \le c\big(1+\|u\|_{L^{p_1}(\Gamma)}\big)^2,$$
$$\|S_h(u)\|_{L^\infty(\Omega)} \le c\big(1+\|u\|_{L^{p_2}(\Gamma)}\big)^2,$$

for $p_1, p_2 > 2$ *for* $n = 2$ *and* $p_1 \ge 4$, $p_2 > 4$ *for* $n = 3$. *These estimates remain valid when replacing* S_h *by* Z_h.

Proof We start with the estimate in the $H^1(\Gamma)$-norm. With the triangle inequality and an inverse estimate we obtain

$$\begin{aligned}\|S_h(u)\|_{H^1(\Gamma)} &\le c\big(\|S(u)-S_h(u)\|_{H^1(\Gamma)} + \|S(u)\|_{H^1(\Gamma)}\big)\\ &\le c\,\big(\|S(u)-\Pi_h S(u)\|_{H^1(\Gamma)} + h^{-1}\|S(u)-\Pi_h S(u)\|_{L^2(\Gamma)}\\ &\quad + h^{-1}\,\|S(u)-S_h(u)\|_{L^2(\Gamma)} + \|S(u)\|_{H^1(\Gamma)}\big).\end{aligned}$$

The first two terms are bounded by the last one due to (20) and it remains to apply the stability estimate from Lemma 6. For the third term we apply the error estimate from Lemma 13. This implies the first estimate.

We prove the maximum norm estimate only for the case $n = 3$. In the following, we write $y_h := S_h(u)$. We introduce the function $\tilde y \in H^1(\Omega)$ solving the problem

$$-\Delta\tilde y + \tilde y = f \ \text{ in } \ \Omega, \qquad \partial_n \tilde y = g - u\,y_h \ \text{ on } \ \Gamma.$$

Obviously, y_h is the Neumann Ritz-projection of $\tilde y$, i. e.,

$$a^{\mathrm N}(y_h-\tilde y, v_h) = \int_\Omega \big(\nabla(y_h-\tilde y)\cdot\nabla v_h + (y_h-\tilde y)\,v_h\big) = 0 \quad \text{for all } \ v_h\in V_h.$$

Let $x^*\in\bar T^*$ with $T^*\in\mathcal T_h$ be the point where $|y_h|$ attains its maximum. With an inverse inequality and the Hölder inequality we get

$$\begin{aligned}\|y_h\|_{L^\infty(\Omega)} &= |y_h(x^*)| \le c\,|T^*|^{-1}\,\|y_h\|_{L^1(T^*)}\\ &\le c\big(|T^*|^{-1}\,\|\tilde y - y_h\|_{L^1(T^*)} + \|\tilde y\|_{L^\infty(T^*)}\big)\\ &= c\,(\delta^h, \tilde y - y_h)_{L^2(\Omega)} + c\,\|\tilde y\|_{L^\infty(\Omega)},\end{aligned} \tag{28}$$

where δ^h is a regularized delta function defined by $\delta^h(x) = |T^*|^{-1}\,\mathrm{sgn}\,(\tilde y(x) - y_h(x))$ if $x\in T^*$ and $\delta^h(x) = 0$ otherwise. The second term on the right-hand side can be treated with the arguments used already in the proof of Lemma 3b), namely

$$\|\tilde y\|_{L^\infty(\Omega)} \le c\big(\|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Gamma)} + \|u\,y_h\|_{L^s(\Gamma)}\big)$$

with $r > 3/2$ and $s = 2 + \varepsilon$ with $\varepsilon > 0$ sufficiently small such that the following arguments remain valid. We estimate the last term with the Hölder inequality for $p_2 = 4\,(2+\varepsilon)/(2-\varepsilon)$ and $p' = 4$ (note that $1/p_2 + 1/p' = 1/s$) and the embedding $H^1(\Omega) \hookrightarrow L^4(\Gamma)$. This yields

$$\|u\,y_h\|_{L^s(\Gamma)} \le c\,\|u\|_{L^{p_2}(\Gamma)}\,\|y_h\|_{L^4(\Gamma)} \le c\,\|u\|_{L^{p_2}(\Gamma)}\,\|y_h\|_{H^1(\Omega)}.$$

It remains to exploit stability of S_h in the $H^1(\Omega)$-norm to conclude

$$\|\tilde y\|_{L^\infty(\Omega)} \le c\,(1 + \|u\|_{L^{p_2}(\Gamma)}). \tag{29}$$

The estimate for the first term on the right-hand side of (28) is based on the ideas from [40, Section 3.6]. First, we introduce a regularized Green's function $g^h \in H^1(\Omega)$ solving the variational problem $a^{\mathrm N}(z, g^h) = (\delta^h, z)_{L^2(\Omega)}$ for all $z \in H^1(\Omega)$. The Neumann Ritz-projection of g^h is denoted by g^h_h. Using the Galerkin orthogonality we obtain

$$\begin{aligned}(\delta^h, \tilde y - y_h) &= a^{\mathrm N}(\tilde y - y_h, g^h) = a^{\mathrm N}(\tilde y - \Pi_h\tilde y, g^h - g^h_h)\\ &\le c\,h^{1/2}\,\|\tilde y\|_{H^{3/2}(\Omega)}\,\|g^h\|_{H^1(\Omega)},\end{aligned} \tag{30}$$

where the last step follows form the stability of the Ritz projection and the interpolation error estimate (19). To bound the $H^1(\Omega)$-norm of g^h we apply the ellipticity of $a^{\mathrm N}$, the definition of g^h, the Hölder inequality and an embedding to arrive at

$$\begin{aligned}c\,\|g^h\|^2_{H^1(\Omega)} &\le a^{\mathrm N}(g^h, g^h) = (\delta^h, g^h)_{L^2(\Omega)}\\ &\le c\,\|\delta^h\|_{L^{6/5}(\Omega)}\,\|g^h\|_{L^6(\Omega)} \le c\,h^{-1/2}\,\|g^h\|_{H^1(\Omega)}.\end{aligned}$$

The last step follows from the property $\|\delta^h\|_{L^{6/5}(\Omega)} \le c\,|T^*|^{-1/6} \le c\,h^{-1/2}$ that can be confirmed with a simple computation. Insertion into (30) and taking into account (28) and (29) yields the desired stability estimate.

The estimates for Z_h follow in a similar way. One just has to replace f by $S_h(u) - y_d$ and the result follows from the estimates proved already for $S_h(u)$. □

Lemma 15 *Assume that $u, v \in L^2(\Gamma)$ satisfy the assumption (2). Then, the Lipschitz estimate*

$$\|S_h(u) - S_h(v)\|_{H^1(\Omega)} \le c\,\|u - v\|_{L^2(\Gamma)}$$

holds.

Proof The proof follows with the same arguments as in the continuous case, see Lemmata 2 and 8. □

Next, we discuss some error estimates for the approximation of the control-to-state and control-to-adjoint operator. While error estimates for S_h and Z_h are a direct consequence of Lemma 12, the results for the linearized operators S_h' and Z_h' require some more effort as for instance $S'(u)\delta u - S_h'(u)\delta u$ does not fulfill the Galerkin orthogonality.

Lemma 16 *For each $u \in U_{ad}$ and $\delta u \in L^2(\Gamma)$ the error estimates*

$$\|S(u) - S_h(u)\|_{H^1(\Omega)} \le c\, h^{1/2}\, (1 + \|u\|_{L^p(\Gamma)}),$$
$$\|S'(u)\delta u - S_h'(u)\delta u\|_{H^1(\Omega)} \le c\, h^{1/2}\, (1 + \|u\|_{L^p(\Gamma)})^3\, \|\delta u\|_{L^2(\Gamma)}$$

are valid for $p > 2$ for $n = 2$ and $p \ge 4$ for $n = 3$. The results are also valid when replacing S and S_h by Z and Z_h, as well as S' and S_h' by Z' and Z_h', respectively.

Proof The first estimate is just a combination of the Lemmata 6 and 12. To show the estimate for the linearized operators we introduce again the abbreviations $y := S(u)$, $y_h := S_h(u)$, $\delta y := S'(u)\delta u$ and $\delta y_h := S_h'(u)\delta u$. Moreover, define the auxiliary function $\delta\tilde{y}_h \in V_h$ as the solution of

$$a_u(\delta\tilde{y}_h, v_h) = (y\,\delta u, v_h)_{L^2(\Gamma)} \qquad \forall v_h \in V_h.$$

This function fulfills the Galerkin orthogonality, i. e., $a_u(\delta y - \delta\tilde{y}_h, v_h) = 0$ for all $v_h \in V_h$. Hence, we obtain with Lemma 12 and the Lipschitz-property from Lemma 2 (note that this Lemma is also valid for the discrete solutions)

$$\begin{aligned}\|\delta y - \delta y_h\|_{H^1(\Omega)} &\le c\big(\|\delta y - \delta\tilde{y}_h\|_{H^1(\Omega)} + \|\delta\tilde{y}_h - \delta y_h\|_{H^1(\Omega)}\big)\\ &\le c\big(h^{1/2}\, \|\delta y\|_{H^{3/2}(\Omega)} + \|\delta u\,(y - y_h)\|_{H^{-1/2}(\Gamma)}\big).\end{aligned}$$

For the first term we simply insert the second estimate from Lemma 7. The second term on the right-hand side is further estimated by means of [20, Theorem 1.4.4.2] and a trace theorem which yield

$$\|\delta u\,(y - y_h)\|_{H^{-1/2}(\Gamma)} \le c\, \|\delta u\|_{L^2(\Gamma)}\, \|y - y_h\|_{H^1(\Omega)},$$

and the assertion follows after an application of the estimate shown already for $S(u) - S_h(u)$. The estimates for Z and Z' follow with similar arguments. ◻

5.2 Convergence of the fully discrete solutions

Throughout this subsection we assume that the properties **(A1)** and **(A2)** are fulfilled. These assumptions are needed to guarantee the required regularity of the solution and the validity of interpolation error estimates.

As the solutions of both the continuous and discrete optimal control problem (5) and (23), respectively, are not unique we have to construct a sequence of discrete local solutions converging towards a continuous one. The first question which arises is whether such a sequence exists. To this end, we introduce a localized problem

$$j_h(u_h) \to \min! \quad \text{s. t. } u_h \in U_h^{ad} \cap B_\varepsilon(\bar u), \tag{31}$$

where $\bar u \in U_{ad}$ is a fixed local solution of (5) fulfilling Assumption 1 and $B_\varepsilon(\bar u)$ is the $L^2(\Gamma)$-ball with radius ε around $\bar u$. The parameter $\varepsilon > 0$ is arbitrary but sufficiently small. First, we show that this problem possesses a unique local solution which would immediately follow if we could show that the coercivity discussed in Corollary 1 is transferred to the discrete case. The following arguments are similar to the investigations in [11], in particular Theorem 4.4 and 4.5 therein.

Lemma 17 *Let $\bar u \in U_{ad}$ be a local solution of (5). Assume that $\varepsilon > 0$ and $h > 0$ are sufficiently small. Then, the inequality*

$$j_h''(u)\delta u^2 \geq \frac{\delta}{4}\|\delta u\|^2_{L^2(\Gamma)}$$

is valid for all u satisfying $\|u - \bar u\|_{L^2(\Gamma)} \leq \varepsilon$.

Proof With the explicit representations of j'' and j_h'' from (14) and (26), respectively, and Corollary 1, we obtain

$$\begin{aligned}\frac{\delta}{2}\|\delta u\|^2_{L^2(\Gamma)} &\leq \left(j''(u)\delta u^2 - j_h''(u)\delta u^2\right) + j_h''(u)\delta u^2 \\ &\leq \Big(\|y_h\,\delta p_h - y\,\delta p\|_{L^2(\Gamma)} + \|\delta y_h\,p_h - \delta y\,p\|_{L^2(\Gamma)}\Big)\,\|\delta u\|_{L^2(\Gamma)} + j_h''(u)\delta u^2,\end{aligned} \tag{32}$$

with $y = S(u)$, $p = Z(u)$, $\delta y = S'(u)\delta u$ and $\delta p = Z'(u)\delta u$, and the discrete analogues $y_h = S_h(u)$, $p_h = Z_h(u)$, $\delta y_h = S_h'(u)\delta u$ and $\delta p_h = Z_h'(u)\delta u$. It remains to bound the two norms in parentheses appropriately. Therefore, we apply the triangle inequality, the stability properties for S', S_h, Z' and Z_h from Lemmata 6, 7 and 14 as well as the error estimates from Lemma 16. Note that the control bounds provide the regularity for u that is required for these estimates. As a consequence we obtain

$$\begin{aligned}&\|y_h\,\delta p_h - y\,\delta p\|_{L^2(\Gamma)} \\ &\quad\leq c\big(\|y - y_h\|_{H^1(\Omega)}\,\|\delta p\|_{H^1(\Omega)} + \|\delta p - \delta p_h\|_{H^1(\Omega)}\,\|y_h\|_{H^1(\Omega)}\big) \\ &\quad\leq c\,h^{1/2}\,\|\delta u\|_{L^2(\Gamma)}.\end{aligned}$$

With similar arguments we can show

$$\begin{aligned}&\|\delta y_h\,p_h - \delta y\,p\|_{L^2(\Gamma)} \\ &\quad\leq c\big(\|\delta y - \delta y_h\|_{H^1(\Omega)}\,\|p_h\|_{H^1(\Omega)} + \|p - p_h\|_{H^1(\Omega)}\,\|\delta y\|_{H^1(\Omega)}\big) \\ &\quad\leq c\,h^{1/2}\,\|\delta u\|_{L^2(\Gamma)}.\end{aligned}$$

The previous two estimates together with (32) imply

$$\frac{\delta}{2}\,\|\delta u\|^2_{L^2(\Gamma)} \leq c\,h^{1/2}\,\|\delta u\|^2_{L^2(\Gamma)} + j_h''(\bar u)\delta u^2.$$

Choosing h sufficiently small such that $c\,h^{1/2} \le \frac{\delta}{4}$ leads to the assertion. □

Theorem 3 *Let $\bar{u} \in U_{ad}$ be a local solution of (5) satisfying Assumption 1. Assume that $\varepsilon > 0$ and $h_0 > 0$ are sufficiently small. Then, the auxiliary problem (31) possesses a unique solution for each $h \le h_0$ denoted by $\bar{u}_h^\varepsilon$, and there holds*

$$\lim_{h\to 0} \|\bar{u} - \bar{u}_h^\varepsilon\|_{L^2(\Gamma)} = 0.$$

Proof The existence of at least one solution of (31) follows immediately from the compactness and non-emptyness of $U_h^{ad} \cap B_\varepsilon(\bar{u})$. Note that the $L^2(\Gamma)$-projection $Q_h\bar{u}$ of $\bar{u}$ onto U_h, defined in (81) in Appendix 2, belongs to $U_h^{ad} \cap B_\varepsilon(\bar{u})$ provided that $h > 0$ is sufficiently small. This means that the feasible set is not empty. Due to Lemma 17 this solution is unique.

Moreover, the family $\{\bar{u}_h^\varepsilon\}_{h\le h_0}$ is bounded and hence, a weakly convergent sequence $\{\bar{u}_{h_k}^\varepsilon\}_{k\in\mathbb{N}}$ with $h_k \searrow 0$ exists. The weak limit is denoted by $\tilde{u} \in L^2(\Gamma)$ and from the convexity of the feasible set we deduce $\tilde{u} \in U_h^{ad} \cap B_\varepsilon(\bar{u})$. Without loss of generality it is assumed that $\bar{u}_h^\varepsilon \rightharpoonup \tilde{u}$ in $L^2(\Gamma)$ as $h \searrow 0$.

Next, we show that $\tilde{u}$ is a local minimum of the continuous problem. First, we show the convergence of the corresponding states which follows with the arguments from [10]. First, we employ the triangle inequality to get

$$\|S(\tilde{u}) - S_h(\bar{u}_h^\varepsilon)\|_{H^1(\Omega)} \le c\big(\|S(\tilde{u}) - S_h(\tilde{u})\|_{H^1(\Omega)} + \|S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon)\|_{H^1(\Omega)}\big). \quad (33)$$

For the first term on the right-hand side we exploit convergence of the finite element method proved in Lemma 16 which yields

$$\|S(\tilde{u}) - S_h(\tilde{u})\|_{H^1(\Omega)} \to 0, \quad h \searrow 0.$$

With similar arguments as in the proof of Lemma 2 we moreover deduce

$$\begin{aligned}\|S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon)\|^2_{H^1(\Omega)} &= -(\tilde{u}\,S_h(\tilde{u}) - \bar{u}_h^\varepsilon\,S_h(\bar{u}_h^\varepsilon), S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon))_{L^2(\Gamma)} \\ &= -((\tilde{u} - \bar{u}_h^\varepsilon)\,S_h(\tilde{u}), S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon))_{L^2(\Gamma)} \\ &\quad - \int_\Gamma \bar{u}_h^\varepsilon\,(S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon))^2.\end{aligned}$$

The integral term on the right-hand side is non-negative due to the lower control bounds $\bar{u}_h^\varepsilon \ge u_a \ge 0$. We can bound the first term on the right-hand side with the Cauchy–Schwarz inequality and the multiplication rule from [20, Theorem 1.4.4.2] which provides

$$\begin{aligned}&\|S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon)\|^2_{H^1(\Omega)} \\ &\quad\le \|\tilde{u} - \bar{u}_h^\varepsilon\|_{H^{-s}(\Gamma)}\,\|S_h(\tilde{u})\|_{H^1(\Gamma)}\,\|S_h(\tilde{u}) - S_h(\bar{u}_h^\varepsilon)\|_{H^1(\Omega)}\end{aligned}$$

for arbitrary $s \in (0, 1/2)$. Note that there holds $\|\tilde{u} - \bar{u}_h^\varepsilon\|_{H^{-s}(\Gamma)} \to 0$ for $h \searrow 0$ due to the compact embedding $L^2(\Gamma) \hookrightarrow H^{-s}(\Gamma)$, $s > 0$. It remains to bound the second factor on the right-hand side by an application of Lemma 14 and to divide the whole estimate by the third factor. After insertion of this estimate into (33) we obtain the strong convergence of the states, this is,

$$\|S(\tilde{u}) - S_h(\bar{u}_h^\varepsilon)\|_{H^1(\Omega)} \to 0 \quad \text{for} \quad h \searrow 0. \tag{34}$$

Next, we show that $\tilde{u}$ is a local solution of the continuous problem (5). To this end we exploit (34) and the lower semi-continuity of the norm map to arrive at

$$j(\tilde{u}) \le \liminf_{h\searrow 0} j_h(\bar{u}_h^\varepsilon) \le \limsup_{h\searrow 0} j_h(\bar{u}_h^\varepsilon) \le \limsup_{h\searrow 0} j_h(Q_h\bar{u}) \le j(\bar{u}). \tag{35}$$

The second to last step follows from the optimality of $\bar{u}_h^\varepsilon$ for (31) and the admissibility of the $L^2(\Gamma)$-projection $Q_h\bar{u}$ for sufficiently small $h > 0$. The last step follows from the strong convergence of the $L^2(\Gamma)$-projection Q_h in $L^2(\Gamma)$. Note that this implies $\lim_{h\searrow 0} \|S_h(Q_h\bar{u}) - S(\bar{u})\|_{L^2(\Omega)} = 0$. Due to Assumption 1 the solution $\bar{u}$ is unique within $B_\varepsilon(\bar{u})$ when $\varepsilon > 0$ is sufficiently small. This implies $\tilde{u} = \bar{u}$. Note that all "$\le$" signs in (35) then turn to "=" signs.

To conclude the strong convergence of the sequence $\{\bar{u}_h^\varepsilon\}_{h>0}$ we show additionally the convergence of the norms. This follows from (35) and the strong convergence of the states from which we infer

$$\begin{aligned}\frac{\alpha}{2}\lim_{h\searrow 0} \|\bar{u}_h^\varepsilon\|^2_{L^2(\Gamma)} &= \lim_{h\searrow 0}\left(j_h(\bar{u}_h^\varepsilon) - \frac{1}{2}\|S_h(\bar{u}_h^\varepsilon) - y_d\|^2_{L^2(\Omega)}\right)\\ &= j(\bar{u}) - \frac{1}{2}\|S(\bar{u}) - y_d\|^2_{L^2(\Omega)} = \frac{\alpha}{2}\|\bar{u}\|^2_{L^2(\Gamma)}.\end{aligned}$$

□

The previous lemma guarantees that every local solution $\bar{u} \in U_{ad}$ satisfying the second-order sufficient condition in Assumption 1 can be approximated by a sequence of local solutions of the discretized problems (31). Due to $\bar{u}_h^\varepsilon \in B_\varepsilon(\bar{u})$ and $\bar{u}_h^\varepsilon \to \bar{u}$ for $h \searrow 0$ (i. e., the constraint $\bar{u}_h^\varepsilon \in B_\varepsilon(\bar{u})$ is never active), the functions $\bar{u}_h^\varepsilon$ are local solutions of the discrete problems (23) provided that $h > 0$ is small enough. Hence, we neglect the superscript ε in the following and denote by $\bar{u}_h$ the sequence of discrete local solutions converging to the local solution $\bar{u}$.

Next, we show linear convergence of the sequence $\bar{u}_h$.

Theorem 4 *Let $\bar{u} \in U_{ad}$ be a local solution of (5) which fulfills Assumption 1, and $\{\bar{u}_h\}_{h>0}$ are local solutions of (23) with $\bar{u}_h \to \bar{u}$ for $h \searrow 0$. Then, the error estimate*

$$\|\bar{u} - \bar{u}_h\|_{L^2(\Gamma)} \le \frac{c}{\sqrt{\delta}} h$$

holds.

Proof Let $\xi = \bar{u} + t(\bar{u}_h - \bar{u})$ with $t \in (0, 1)$. From Corollary 1 we obtain for sufficiently small h the estimate

$$\frac{\delta}{2}\|\bar u-\bar u_h\|^2_{L^2(\Gamma)} \le j''(\xi)(\bar u-\bar u_h)^2$$
$$= j'(\bar u)(\bar u-\bar u_h) - j'(\bar u_h)(\bar u-\bar u_h),$$

where the last step follows from the mean value theorem for some $t \in (0,1)$. Next, we confirm with the first-order optimality conditions that

$$j'(\bar u)(\bar u-\bar u_h) \le 0 \le j_h'(\bar u_h)(Q_h\bar u-\bar u_h)$$

with the $L^2(\Gamma)$ projection Q_h onto U_h. Note that the property $Q_h\bar u \in U_{ad}$ is trivially satisfied. Insertion into the inequality above leads to

$$\frac{\delta}{2}\|\bar u-\bar u_h\|^2_{L^2(\Gamma)} \le (j_h'(\bar u_h)-j'(\bar u_h))(Q_h\bar u-\bar u_h) - j'(\bar u_h)(\bar u-Q_h\bar u). \tag{36}$$

An estimate for the second part follows from orthogonality of the $L^2(\Gamma)$-projection, this is,

$$\begin{aligned} j'(\bar u_h)(\bar u-Q_h\bar u) &= (\alpha\,\bar u_h - S(\bar u_h)\,Z(\bar u_h), \bar u-Q_h\bar u)_{L^2(\Gamma)} \\ &= (Q_h(S(\bar u_h)\,Z(\bar u_h)) - S(\bar u_h)\,Z(\bar u_h), \bar u-Q_h\bar u)_{L^2(\Gamma)} \\ &\le c\,h^2\,\|S(\bar u_h)\,Z(\bar u_h)\|_{H^1(\Gamma)}\,\|\bar u\|_{H^1(\Gamma)}. \end{aligned} \tag{37}$$

Furthermore, we exploit the Leibniz rule and the stability properties for S and Z from Lemma 6 to obtain

$$\begin{aligned} \|S(\bar u_h)\,Z(\bar u_h)\|_{H^1(\Gamma)} \le c\Big(&\|S(\bar u_h)\|_{H^1(\Gamma)}\,\|Z(\bar u_h)\|_{L^\infty(\Omega)} \\ &+ \|S(\bar u_h)\|_{L^\infty(\Omega)}\,\|Z(\bar u_h)\|_{H^1(\Gamma)}\Big) \le c. \end{aligned} \tag{38}$$

Next, we discuss the first term on the right-hand side of (36). Insertion of the definition of j_h' and j' and the stability of Q_h yield

$$\begin{aligned} &(j_h'(\bar u_h)-j'(\bar u_h))(Q_h\bar u-\bar u_h) \\ &\quad= (S(\bar u_h)\,Z(\bar u_h) - S_h(\bar u_h)\,Z_h(\bar u_h), Q_h(\bar u-\bar u_h))_{L^2(\Gamma)} \\ &\quad\le c\,\Big(\|Z(\bar u_h)\|_{L^\infty(\Gamma)}\,\|S(\bar u_h)-S_h(\bar u_h)\|_{L^2(\Gamma)} \\ &\qquad+ \|S_h(\bar u_h)\|_{L^\infty(\Gamma)}\,\|Z(\bar u_h)-Z_h(\bar u_h)\|_{L^2(\Gamma)}\Big)\,\|\bar u-\bar u_h\|_{L^2(\Gamma)} \\ &\quad\le c\,h\big(\|Z(\bar u_h)\|_{L^\infty(\Gamma)}\,\|S(\bar u_h)\|_{H^{3/2}(\Omega)} + \|S_h(\bar u_h)\|_{L^\infty(\Gamma)}\,\|Z(\bar u_h)\|_{H^{3/2}(\Omega)}\big) \\ &\qquad\times\|\bar u-\bar u_h\|_{L^2(\Gamma)}. \end{aligned}$$

In the last step we inserted the finite element error estimates from Lemma 13. Exploiting also the stability estimates from Lemmata 6 and 14 we obtain

$$(j_h'(\bar u_h)-j'(\bar u_h))(Q_h\bar u-\bar u_h) \le c\,h\,\|\bar u-\bar u_h\|_{L^2(\Gamma)}.$$

Together with (36), (37) and (38) we arrive at the assertion. □

5.3 Postprocessing approach

In this section we consider the so-called postprocessing approach introduced in [33]. The basic idea is to compute an “improved” control $\tilde{u}_h$ by a pointwise evaluation of the projection formula, i. e.,

$$\tilde{u}_h := \Pi_{\mathrm{ad}}\Big(\frac{1}{\alpha}[\bar{y}_h\bar{p}_h]_\Gamma\Big), \tag{39}$$

where $\bar{y}_h$ and $\bar{p}_h$ is the discrete state and adjoint state, respectively, obtained by the full discretization approach discussed in Sect. 5.2. As we require higher regularity of the exact solution in order to observe a higher convergence rate than for the full discretization approach, we replace **(A1)** by the stronger assumption

(A1’) The domain Ω is planar and its boundary is globally C^3.

The most technical part of convergence proofs for this approach is the proof of L^2-norm estimates for the state variables. This is usually done by considering the following three terms separately:

$$\begin{aligned}\|\bar{y}-\bar{y}_h\|_{L^2(\Gamma)} \le c\Big(&\|\bar{y}-S_h(\bar{u})\|_{L^2(\Gamma)} + \|S_h(\bar{u})-S_h(R_h\bar{u})\|_{L^2(\Gamma)}\\ &+\|S_h(R_h\bar{u})-\bar{y}_h\|_{L^2(\Gamma)}\Big).\end{aligned} \tag{40}$$

In [33] $R_h : C(\Gamma) \to U_h$ is chosen as the midpoint interpolant. We will construct and investigate such an operator in Appendix 1. Note that a definition of a midpoint interpolant on curved elements is not straight-forward. The first term on the right-hand side of (40) is a finite element error in the $L^2(\Gamma)$-norm. We collect the required estimates in the following Lemma.

Lemma 18 *For all $q < \infty$ there hold the estimates*

$$\begin{aligned}\|\bar{y}-S_h(\bar{u})\|_{L^2(\Gamma)} &\le c\,h^{2-2/q}\,|\ln h|\,\|\bar{y}\|_{W^{2,q}(\Omega)}\\ \|\bar{p}-Z_h(\bar{u})\|_{L^2(\Gamma)} &\le c\,h^{2-2/q}\,|\ln h|\,\big(\|\bar{p}\|_{W^{2,q}(\Omega)}+\|\bar{y}\|_{H^2(\Omega)}\big).\end{aligned}$$

Proof The first estimate follows from the Hölder inequality and the maximum norm estimate derived in Theorem 2. The second estimate requires an intermediate step. We denote by $p^h(\bar{u}) \in V_h$ the solution of the equation

$$a_{\bar{u}}(p^h(\bar{u}), v_h) = (S(\bar{u}) - y_d, v_h)_{L^2(\Omega)} \qquad \forall v_h \in V_h.$$

As $p^h(\bar{u})$ is the Ritz-projection of $\bar{p}$ we can apply Theorem 2 again and obtain

$$\|\bar{p}-p^h(\bar{u})\|_{L^2(\Gamma)} \le c\,h^{2-2/q}\,|\ln h|\,\|\bar{p}\|_{W^{2,q}(\Gamma)}.$$

To show an estimate for the error between $p^h(\bar{u})$ and $Z_h(\bar{u})$ we test the equations defining both functions by $v_h = p^h(\bar{u}) - Z_h(\bar{u})$, compare the proof of Lemma 2. Together with the non-negativity of $\bar{u}$ we obtain

$$\begin{aligned}
&\|p^h(\bar{u}) - Z_h(\bar{u})\|^2_{H^1(\Omega)} \\
&\quad = -\int_\Gamma \bar{u}\,(p^h(\bar{u}) - Z_h(\bar{u}))^2 + (S(\bar{u}) - S_h(\bar{u}), p^h(\bar{u}) - Z_h(\bar{u}))_{L^2(\Omega)} \\
&\quad \le c\,h^2\,\|y\|_{H^2(\Omega)}\,\|p^h(\bar{u}) - Z_h(\bar{u})\|_{H^1(\Omega)}.
\end{aligned}$$

The last step follows from the estimate $\|S(\bar{u}) - S_h(\bar{u})\|_{L^2(\Omega)} \le c\,h^2\,\|S(\bar{u})\|_{H^2(\Omega)}$ which is a consequence of the Aubin-Nitsche trick. With the triangle inequality we conclude the desired estimate for the discrete control-to-adjoint operator. □

To obtain an optimal error estimate for the second term we need an additional assumption which is used in all contributions studying the postprocessing approach. To this end, define the subsets $\mathcal{K}_2 := \cup\{\bar{E} : E \in \mathcal{E}_h,\ E \subset \mathcal{A},\ \text{ or } E \subset \mathcal{I}\}$ and $\mathcal{K}_1 := \Gamma \setminus \mathcal{K}_2$. In the following we will assume that $\mathcal{K}_1$ satisfies

$$|\mathcal{K}_1| \le c\,h. \tag{41}$$

The idea of this assumption is, that the control can only switch between active and inactive set on $\mathcal{K}_1$. Only due to these switching points the regularity of the control is reduced, see also Lemma 11. One can in general expect that this happens at finitely many points and thus, the assumption (41) is not very restrictive.

As an intermediate result required to prove estimates for $Z_h(\bar{u}) - Z_h(R_h\bar{u})$ in $L^2(\Gamma)$, we need an estimate for $S_h(\bar{u}) - S_h(R_h\bar{u})$ in $L^2(\Omega)$.

Lemma 19 *For all $q < \infty$ there holds the estimate*

$$\|S_h(\bar{u}) - S_h(R_h\bar{u})\|_{L^2(\Omega)} \le c\,h^{2-2/q}\big(1 + \|\bar{u}\|_{W^{1,q}(\Gamma)} + \|\bar{u}\|_{H^{2-1/q}(\mathcal{K}_2)}\big).$$

Proof To shorten the notation we write $e_h := S_h(\bar{u}) - S_h(R_h\bar{u})$. Moreover, we introduce the function $w \in H^1(\Omega)$ solving the equation

$$a_{\bar{u}}(v, w) = (e_h, v)_{L^2(\Omega)} \qquad \forall v \in H^1(\Omega). \tag{42}$$

This implies

$$\|e_h\|^2_{L^2(\Omega)} = a_{\bar{u}}(e_h, w - \Pi_h w) + a_{\bar{u}}(e_h, \Pi_h w). \tag{43}$$

Next, we discuss both terms on the right-hand side separately. The first one is treated with the Cauchy–Schwarz inequality and the interpolation error estimate (19). These arguments lead to

$$a_{\bar{u}}(e_h, w - \Pi_h w) \le c\,h\,\|e_h\|_{H^1(\Omega)}\,\|e_h\|_{L^2(\Omega)}.$$

The $H^1(\Omega)$-norm of e_h is further estimated by the Lipschitz property from Lemma 15 and the interpolation error estimate for the midpoint interpolant from Lemma 25. This yields

$$\|e_h\|_{H^1(\Omega)} \le c\,\|\bar{u} - R_h\bar{u}\|_{L^2(\Gamma)} \le c\,h\,\|\bar{u}\|_{W^{1,q}(\Gamma)}$$

for all $q \ge 2$.

Insertion into the estimate above taking into account the stability estimates from Lemma 14 yields

$$a_{\bar{u}}(e_h, w - \Pi_h w) \le ch^2 \|\bar{u}\|_{W^{1,q}(\Gamma)} \|e_h\|_{L^2(\Omega)}. \tag{44}$$

Next, we consider the second term on the right-hand side of (43). After a reformulation by means of the definition of S_h we get

$$\begin{aligned} a_{\bar{u}}(e_h, \Pi_h w) &= a_{\bar{u}}(S_h(\bar{u}), \Pi_h w) - a_{\bar{u}}(S_h(R_h\bar{u}), \Pi_h w) \\ &= a_{R_h\bar{u}}(S_h(R_h\bar{u}), \Pi_h w) - a_{\bar{u}}(S_h(R_h\bar{u}), \Pi_h w) \\ &= ((R_h\bar{u} - \bar{u})\, S_h(R_h\bar{u}), \Pi_h w)_{L^2(\Gamma)}. \end{aligned} \tag{45}$$

We can further estimate this term with the interpolation error estimate from Lemma 27

$$\begin{aligned} ((R_h\bar{u} - \bar{u})\, S_h(R_h\bar{u}), \Pi_h w)_{L^2(\Gamma)} &\le c\,h^2\,\|\bar{u}\|_{H^1(\Gamma)}\,\|S_h(R_h\bar{u})\,\Pi_h w\|_{H^1(\Gamma)} \\ &\quad + \|S_h(R_h\bar{u})\|_{L^\infty(\Gamma)}\,\|\Pi_h w\|_{L^\infty(\Gamma)} \sum_{E\in\mathcal{E}_h} \left|\int_E (\bar{u} - R_h\bar{u})\right| \\ &\le c\left(h^2 + \sum_{E\in\mathcal{E}_h} \left|\int_E (\bar{u} - R_h\bar{u})\right|\right)\left(1 + \|\bar{u}\|_{H^1(\Gamma)}\right)\|S_h(R_h\bar{u})\|_{H^1(\Gamma)}\,\|\Pi_h w\|_{H^1(\Gamma)}. \end{aligned} \tag{46}$$

The last step follows from the embedding $H^1(\Gamma) \hookrightarrow L^\infty(\Gamma)$ and the multiplication rule $\|u\,v\|_{H^1(\Gamma)} \le c\,\|u\|_{H^1(\Gamma)}\,\|v\|_{H^1(\Gamma)}$, see [20, Theorem 1.4.4.2]. Both properties are only fulfilled in case of $n = 2$.

Let us discuss the terms on the right-hand side separately. For elements $E \subset \mathcal{K}_1$ we can exploit the assumption (41) which provides the estimate $\sum_{E\subset\mathcal{K}_1} |E| \le c\,h$ and the second interpolation error estimate from Lemma 25 to arrive at

$$\sum_{\substack{E\in\mathcal{E}_h \\ E\subset\mathcal{K}_1}} \left|\int_E (\bar{u} - R_h\bar{u})\right| \le c\,h\,\|\bar{u} - R_h\bar{u}\|_{L^\infty(\Gamma)} \le c\,h^{2-1/q}\,\|\nabla\bar{u}\|_{L^q(\Gamma)}.$$

On elements $E \subset \mathcal{K}_2$ the control has higher regularity, namely $\bar{u} \in H^{2-1/q}(E)$. To this end, we show by interpolation arguments in Banach spaces, see e. g. [7, Section 14.3], that the two estimates from Lemma 25 (the second one with $r = 1$ and $q = 2$) also imply

$$\int_E (v - R_h v) \le c\,h^{5/2-1/q}\,\|v\|_{H^{2-1/q}(E)}, \quad v \in H^{2-1/q}(\Gamma),\ q \in [1, \infty]. \tag{47}$$

As a consequence, we deduce

$$\sum_{\substack{E\in\mathcal{E}_h\\ E\subset\mathcal{K}_2}} \left|\int_E (\bar u - R_h\bar u)\right| \le c\, h^{5/2-1/q}\, \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)} \Big(\sum_{\substack{E\in\mathcal{E}_h\\ E\subset\mathcal{K}_2}} 1\Big)^{1/2}$$
$$\le c\, h^{2-1/q}\, \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}.$$

The remaining terms on the right-hand side of (46) can be treated with stability estimates for S_h (see Lemma 14) and R_h, the estimate $\|\Pi_h w\|_{H^1(\Gamma)} \le c\,\|w\|_{H^1(\Gamma)}$ stated in (20) and the a priori estimate $\|w\|_{H^1(\Gamma)} \le c\,\|e_h\|_{L^2(\Omega)}$ from Lemma 3a). Insertion of the previous estimates into (46) yields

$$\begin{aligned}&((R_h\bar u - \bar u)\, S_h(R_h\bar u), \Pi_h w)_{L^2(\Gamma)}\\ &\quad\le c\, h^{2-1/q}\big(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\big)\|e_h\|_{L^2(\Omega)}.\end{aligned} \tag{48}$$

Note that we hide the of lower-order norms of $\bar u$ in the generic constant as these quantities may be estimated by means of the control bounds u_a and u_b. Insertion of (44), (45) and (48) into (43) and dividing by $\|e_h\|_{L^2(\Omega)}$ implies the assertion. □

Lemma 20 *Under the assumption* (41) *the estimates*

$$\|S_h(\bar u) - S_h(R_h\bar u)\|_{L^2(\Gamma)} \le c\, h^{2-1/q}\, \big(1 + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)} + \|\bar u\|_{W^{1,q}(\Gamma)}\big),$$
$$\|Z_h(\bar u) - Z_h(R_h\bar u)\|_{L^2(\Gamma)} \le c\, h^{2-1/q}\, \big(1 + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)} + \|\bar u\|_{W^{1,q}(\Gamma)}\big)$$

are valid for arbitrary $q \in [2,\infty)$.

Proof We will only prove the second estimate as the first one follows from the same technique and is even easier as the right-hand sides of the equations defining $S_h(\bar u)$ and $S_h(R_h\bar u)$ coincide. This is not the case for the control-to-adjoint operator.

To shorten the notation we write $e_h := Z_h(\bar u) - Z_h(R_h\bar u)$. As in the previous lemma we rewrite the error by a duality argument using a dual problem similar to (42) with solution $w \in H^1(\Omega)$, more precisely,

$$a_{\bar u}(v, w) = (e_h, v)_{L^2(\Gamma)} \qquad \forall v \in H^1(\Omega).$$

This yields

$$\|e_h\|^2_{L^2(\Gamma)} = a_{\bar u}(e_h, w - \Pi_h w) + a_{\bar u}(e_h, \Pi_h w). \tag{49}$$

We rewrite the second expression in (49) and get analogous to (45)

$$\begin{aligned}&a_{\bar u}(e_h, \Pi_h w)\\ &\quad= a_{\bar u}(Z_h(\bar u), \Pi_h w) \pm a_{R_h\bar u}(Z_h(R_h\bar u), \Pi_h w) - a_{\bar u}(Z_h(R_h\bar u), \Pi_h w)\\ &\quad= (S_h(\bar u) - S_h(R_h\bar u), \Pi_h w)_{L^2(\Omega)} + ((R_h\bar u - \bar u)\, Z_h(R_h\bar u), \Pi_h w)_{L^2(\Gamma)}.\end{aligned} \tag{50}$$

Note that the first term would not appear when deriving estimates for S_h instead of Z_h as the equations defining $S_h(\bar u)$ and $S_h(R_h\bar u)$ have the same right-hand side.

The first term can be treated with the Cauchy–Schwarz inequality, Lemma 19 and the estimate $\|\Pi_h w\|_{L^2(\Omega)} \le c\,\|w\|_{H^1(\Omega)} \le c\,\|e_h\|_{L^2(\Gamma)}$ which can be deduced from (19) and Lemma 1 with $g = e_h$. These ideas lead to

$$\begin{aligned}(S_h(\bar u) - S_h(R_h\bar u), \Pi_h w)_{L^2(\Omega)}&\\ \le c\,h^{2-1/q}\left(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\right)\|e_h\|_{L^2(\Gamma)}.&\end{aligned} \tag{51}$$

For the second term on the right-hand side of (50) we apply the same steps as for (48) with the only modification that the a priori estimate $\|w\|_{H^1(\Gamma)} \le c\,\|e_h\|_{L^2(\Gamma)}$ from Lemma 3a) has to be employed. From this we infer

$$\begin{aligned}&((R_h\bar u - \bar u)\,Z_h(R_h\bar u), \Pi_h w)_{L^2(\Gamma)}\\ &\quad\le c\,h^{2-1/q}\left(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\right)\|Z_h(R_h\bar u)\|_{H^1(\Gamma)}\,\|\Pi_h w\|_{H^1(\Gamma)}\\ &\quad\le c\,h^{2-1/q}\left(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\right)\|e_h\|_{L^2(\Gamma)}.\end{aligned} \tag{52}$$

In the last step we used the boundedness of $Z_h(R_h\bar u)$, see Lemma 14. Insertion of (51) and (52) into (50) leads to

$$a_{\bar u}(e_h, \Pi_h w) \le c\,h^{2-1/q}\left(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\right)\|e_h\|_{L^2(\Gamma)}. \tag{53}$$

It remains to discuss the first term on the right-hand side of (49). We obtain with the boundedness of $a_{\bar u}$, the interpolation error estimate (19) and Lemma 3a)

$$\begin{aligned}a_{\bar u}(e_h, w - \Pi_h w) &\le c\,h^{1/2}\,\|e_h\|_{H^1(\Omega)}\,\|w\|_{H^{3/2}(\Omega)}\\ &\le c\,h^{1/2}\,\|e_h\|_{H^1(\Omega)}\,\|e_h\|_{L^2(\Gamma)}.\end{aligned} \tag{54}$$

An estimate for the expression $\|e_h\|_{H^1(\Omega)}$ follows from the equality

$$\|e_h\|^2_{H^1(\Omega)} + (\bar u\,Z_h(\bar u) - R_h\bar u\,Z_h(R_h\bar u), e_h)_{L^2(\Gamma)} = (S_h(\bar u) - S_h(R_h\bar u), e_h)_{L^2(\Omega)}$$

which can be deduced by subtracting the equations for $Z_h(\bar u)$ and $Z_h(R_h\bar u)$ from each other. Rearranging the terms yields

$$\begin{aligned}\|e_h\|^2_{H^1(\Omega)} \le&((R_h\bar u - \bar u)\,Z_h(R_h\bar u), e_h)_{L^2(\Gamma)}\\ &- (\bar u\,e_h, e_h)_{L^2(\Gamma)} + \|S_h(\bar u) - S_h(R_h\bar u)\|_{L^2(\Omega)}\|e_h\|_{L^2(\Omega)}.\end{aligned}$$

The second term on the right-hand side can be bounded by zero as $\bar u \ge 0$. An estimate for the last term is proved in Lemma 19. For the first term we apply the estimate (52) with $\Pi_h w$ replaced by e_h. All together, we obtain

$$\begin{aligned}\|e_h\|^2_{H^1(\Omega)} \le c\,h^{2-1/q}&\left(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\right)\\ &\times\left(\|e_h\|_{H^1(\Gamma)} + \|e_h\|_{H^1(\Omega)}\right).\end{aligned} \tag{55}$$

Moreover, with an inverse inequality and a trace theorem we get

$$\|e_h\|_{H^1(\Gamma)} \le c\,h^{-1/2}\,\|e_h\|_{H^{1/2}(\Gamma)} \le c\,h^{-1/2}\,\|e_h\|_{H^1(\Omega)}.$$

Consequently, we deduce from (55)

$$\|e_h\|_{H^1(\Omega)} \le c\,h^{3/2-1/q}\big(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\big).$$

Insertion into (54) leads to

$$a_{\bar u}(e_h, w - \Pi_h w) \le c\,h^{2-1/q}\,\big(1 + \|\bar u\|_{W^{1,q}(\Gamma)} + \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}\big)\|e_h\|_{L^2(\Gamma)}.$$

Together with (53) and (49) we conclude the desired estimate for Z_h. □

Lemma 21 *Under the assumption* (41) *there holds the estimate*

$$\|R_h\bar u - \bar u_h\|_{L^2(\Gamma)} \le c\,h^{2-2/q}\,|\ln h|$$

with

$$c = c\big(\|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}, \|\bar u\|_{W^{1,q}(\Gamma)}, \|\bar y\|_{W^{2,q}(\Omega)}, \|\bar p\|_{W^{2,q}(\Omega)}\big).$$

Proof We observe that each function $\xi := t\,R_h\bar u + (1-t)\,\bar u_h$ for $t \in [0,1]$ satisfies

$$\|\bar u - \xi\|_{L^2(\Gamma)} \le t\|\bar u - R_h\bar u\|_{L^2(\Gamma)} + (1-t)\|\bar u - \bar u_h\|_{L^2(\Gamma)} < \varepsilon,$$

for arbitrary $\varepsilon > 0$ provided that h is sufficiently small. This follows from the convergence of the midpoint interpolant, see Lemma 25, and convergence of $\bar u_h$ towards $\bar u$, see Theorem 3. Hence, with the coercivity of j_h'' proved in Lemma 17 and the mean value theorem we conclude

$$\begin{aligned}\frac{\delta}{4}\,\|R_h\bar u - \bar u_h\|^2_{L^2(\Gamma)} &\le j_h''(\xi)(R_h\bar u - \bar u_h)^2\\ &= j_h'(R_h\bar u)(R_h\bar u - \bar u_h) - j_h'(\bar u_h)(R_h\bar u - \bar u_h).\end{aligned}$$

For the latter term we exploit the discrete optimality condition and the fact that the continuous optimality condition holds even pointwise. This implies the inequality

$$j_h'(\bar u_h)(R_h\bar u - \bar u_h) \ge 0 \ge (\alpha\,R_h\bar u - R_h(\bar y\,\bar p), R_h\bar u - \bar u_h)_{L^2(\Gamma)}.$$

Insertion into the estimate above implies

$$\begin{aligned}&\frac{\delta}{4}\|R_h\bar u - \bar u_h\|^2_{L^2(\Gamma)}\\ &\quad\le \big(R_h(\bar y\,\bar p) - \bar y\,\bar p + \bar y\,\bar p - S_h(R_h\bar u)\,Z_h(R_h\bar u), R_h\bar u - \bar u_h\big)_{L^2(\Gamma)}.\end{aligned} \tag{56}$$

The right-hand side can be decomposed into two parts. With appropriate intermediate functions we obtain for the latter one

$$
\begin{aligned}
&(\bar y\,\bar p - S_h(R_h\bar u)\,Z_h(R_h\bar u), R_h\bar u - \bar u_h)_{L^2(\Gamma)} \\
&\quad = ((\bar y - S_h(R_h\bar u))\,\bar p + S_h(R_h\bar u)\,(\bar p - Z_h(R_h\bar u)), R_h\bar u - \bar u_h)_{L^2(\Gamma)} \\
&\quad \le c\,\big(\|\bar y - S_h(R_h\bar u)\|_{L^2(\Gamma)}\,\|\bar p\|_{L^\infty(\Gamma)} \\
&\qquad + \|\bar p - Z_h(R_h\bar u)\|_{L^2(\Gamma)}\,\|S_h(R_h\bar u)\|_{L^\infty(\Gamma)}\big)\,\|R_h\bar u - \bar u_h\|_{L^2(\Gamma)}.
\end{aligned}
$$

Moreover, we apply the triangle inequality and the estimates from Lemmata 18 and 20 to deduce

$$
\begin{aligned}
\|\bar y - S_h(R_h\bar u)\|_{L^2(\Gamma)} &\le \|\bar y - S_h(\bar u)\|_{L^2(\Gamma)} + \|S_h(\bar u) - S_h(R_h\bar u)\|_{L^2(\Gamma)} \\
&\le c\,h^{2-2/q}\,|\ln h|.
\end{aligned}
$$

Analogously, one can derive an estimate for the term $\|\bar p - Z_h(R_h\bar u)\|_{L^2(\Gamma)}$. Moreover, we apply Lemmata 6 and 14 to bound the norms of $p = Z(\bar u)$ and $S_h(R_h\bar u)$, respectively. All together we obtain the estimate

$$
(\bar y\,\bar p - S_h(R_h\bar u)\,Z_h(R_h\bar u), R_h\bar u - \bar u_h)_\Gamma \le c\,h^{2-2/q}\,|\ln h|\,\|R_h\bar u - \bar u_h\|_{L^2(\Gamma)}. \tag{57}
$$

Next we discuss that part of (56) which involves the term $R_h(\bar y\,\bar p) - \bar y\,\bar p$ in the first argument. Here, we again use the interpolation error estimate (47) exploiting regularity in fractional-order Sobolev spaces and obtain

$$
\begin{aligned}
(R_h(\bar y\,\bar p) - \bar y\,\bar p, R_h\bar u - \bar u_h)_{L^2(\Gamma)} &= \sum_{E\in\mathcal{E}_h} [R_h\bar u - \bar u_h]_E \int_E \big(\bar y\,\bar p - R_h(\bar y\,\bar p)\big) \\
&\le c \sum_{E\in\mathcal{E}_h} h^{1/2}\,[R_h\bar u - \bar u_h]_E\,h^{2-1/q}\,\|\bar y\,\bar p\|_{H^{2-1/q}(E)} \\
&\le c\,h^{2-1/q}\,\|R_h\bar u - \bar u_h\|_{L^2(\Gamma)}\,\|\bar y\,\bar p\|_{H^{2-1/q}(\Gamma)}.
\end{aligned} \tag{58}
$$

With [20, Theorem 1.4.4.2] and a trace theorem we conclude

$$
\|\bar y\,\bar p\|_{H^{2-1/q}(\Gamma)} \le c\,\|\bar y\|_{W^{2-1/q,q}(\Gamma)}\,\|\bar p\|_{W^{2-1/q,q}(\Gamma)} \le c\,\|\bar y\|_{W^{2,q}(\Omega)}\,\|\bar p\|_{W^{2,q}(\Omega)}.
$$

Insertion of the estimates (57) and (58) into (56), and dividing the resulting estimate by $\|R_h u - u_h\|_{L^2(\Gamma)}$, leads to the desired result. □

Now we are in the position to state the main result of this section.

Theorem 5 *Let* $(\bar y, \bar u, \bar p)$ *be a local solution of* (12) *satisfying the assumption* (41). *Moreover, let* $\{\bar u_h\}_{h>0}$ *be a sequence of local solutions of* (27) *such that for sufficiently small* $\varepsilon, h_0 > 0$ *the property*

$$
\|\bar u - \bar u_h\|_{L^2(\Gamma)} < \varepsilon \qquad \forall h < h_0
$$

holds. Then, the error estimate

$$
\|\bar u - \tilde u_h\|_{L^2(\Gamma)} \le c\,h^{2-2/q}\,|\ln h|
$$

is satisfied with $c = c(\|\bar u\|_{W^{1,q}(\Gamma)}, \|\bar u\|_{H^{2-1/q}(\mathcal{K}_2)}, \|\bar y\|_{W^{2,q}(\Omega)}, \|\bar p\|_{W^{2,q}(\Omega)})$.

Proof With the projection formulas (13) and (39), respectively, the non-expansivity of the operator Π_{ad} and the triangle inequality we obtain

$$\begin{aligned}\|\bar u - \tilde u_h\|_{L^2(\Gamma)} &\le c\,\Big\|\Pi_{ad}\Big(\frac{1}{\alpha}\,\bar y\,\bar p\Big) - \Pi_{ad}\Big(\frac{1}{\alpha}\,\bar y_h\,\bar p_h\Big)\Big\|_{L^2(\Gamma)} \\ &\le \frac{c}{\alpha}\,\big(\|\bar y - \bar y_h\|_{L^2(\Gamma)}\,\|\bar p\|_{L^\infty(\Omega)} + \|\bar y_h\|_{L^\infty(\Omega)}\,\|\bar p - \bar p_h\|_{L^2(\Gamma)}\big).\end{aligned}$$

The assertion follows after insertion of (40) together with the estimates obtained in Lemmata 18, 20 and 21, as well as the stability estimates of Z and S_h from Lemmata 3 and 14, respectively. □

6 Numerical experiments

It is the purpose of this last section to confirm the theoretical results by numerical experiments. To this end, we reformulate the discrete optimality condition (27) and use the equivalent projection formula

$$u_h = \Pi_{ad}\Big(\frac{1}{\alpha}R_h^{\mathrm{Simp}}(S_h(u_h)\,Z_h(u_h))\Big). \tag{59}$$

Here, $R_h^{\mathrm{Simp}} : C(\Gamma) \to U_h$ is a projection operator based on the Simpson rule, this is,

$$[R_h^{\mathrm{Simp}}(v)]_E = \frac{1}{6}\big(v(x_{E_1}) + 4v(x_E) + v(x_{E_2})\big),$$

where x_{E_1} and x_{E_2} are the endpoints of the boundary edge $E \in \mathcal{E}_h$ and x_E its midpoint. The numerical solution of (59) is computed by a semismooth Newton-method.

The input data of the considered benchmark problem is chosen as follows. The computational domain is the unit square $\Omega := (0,1)^2$. We define the exact Robin parameter $\tilde u$ by

$$\tilde u(x_1,x_2) := \begin{cases} \max(-0.01,\ 1-30(x_1-0.5)^2), & \text{if } x_1 = 0, \\ -0.01, & \text{otherwise}, \end{cases}$$

and use the desired state $y_d = S_h(\tilde u)$ and the right-hand side $f \equiv 0$. Moreover, the regularization parameter $\alpha = 10^{-2}$ and the control bounds $u_a = 0$, $u_b = \infty$ are used.

We compute the numerical solution of our benchmark problem on a sequence of meshes starting with $\mathcal{T}_{h_0}$, $h_0 = \sqrt{2}$, consisting of two rectangular triangles only. The remaining grids $\mathcal{T}_{h_i}$, $i = 1, 2, \ldots$, are obtained by a double bisection through the longest edge of each element applied to the previous mesh. This guarantees $h_i = \frac{1}{2}h_{i-1}$. In order to compute the discretization error we use the solution on the mesh $\mathcal{T}_{h_{11}}$ as an approximation of the exact solution, this means,

$$\|\bar u - \bar u_{h_i}\|_{L^2(\Gamma)} \approx \|\bar u_{h_{11}} - \bar u_{h_i}\|_{L^2(\Gamma)}, \quad i = 0, 1, \ldots, 10.$$

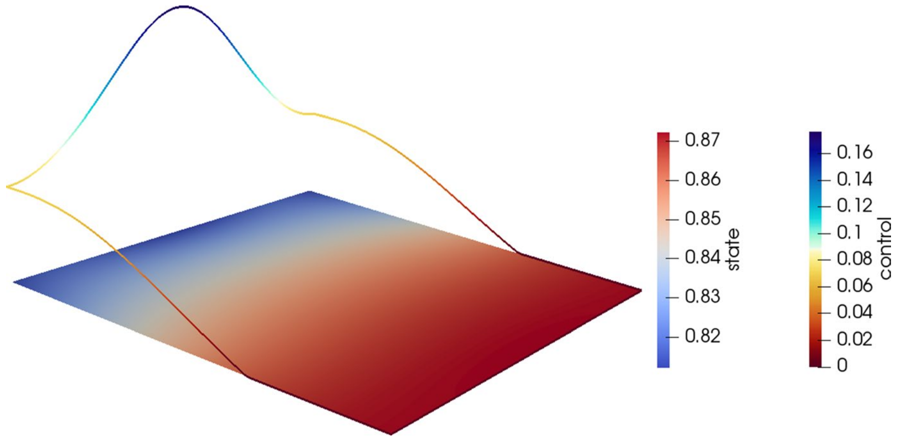

Fig. 1 Optimal state (surface) and the optimal control (boundary curve) for the benchmark problem

Table 1 Experimentally computed errors for the full discretization and the postprocessing approach with the corresponding experimental convergence rates (in parentheses)

i	DOF	BD DOF	$\|u - u_{h_i}\|_{L^2(\Gamma)}$ (eoc)	$\|u - \tilde{u}_{h_i}\|_{L^2(\Gamma)}$ (eoc)
3	113	32	1.60e−2 (1.06)	1.81e−2 (1.15)
4	353	64	5.81e−3 (1.46)	4.43e−3 (2.03)
5	1217	128	2.56e−3 (1.18)	1.03e−3 (2.11)
6	4481	256	1.24e−3 (1.05)	1.65e−4 (2.64)
7	17,153	512	6.17e−4 (1.00)	7.52e−5 (1.13)
8	67,073	1024	3.06e−4 (1.01)	1.79e−5 (2.07)
9	265,217	2048	1.49e−4 (1.04)	4.31e−6 (2.05)
10	1,054,716	4096	6.67e−5 (1.16)	8.50e−7 (2.34)

Analogously, we compute the error for the approximation obtained by the post-processing strategy. However, in this case the exact solution is approximated by $\bar{u} \approx \Pi_{ad}(\frac{1}{\alpha}\bar{y}_{h_{11}}\bar{p}_{h_{11}})$. The error norms $\|\Pi_{ad}(\frac{1}{\alpha}\bar{y}_{h_{11}}\bar{p}_{h_{11}}) - \Pi_{ad}(\frac{1}{\alpha}\bar{y}_{h_i}\bar{p}_{h_i})\|_{L^2(\Gamma)}$, $i = 0, \dots, 11$, are computed element-wise by the Simpson quadrature formula with the modification that elements E are split at those points where $\bar{y}_{h_i}\bar{p}_{h_i}$ or $\bar{y}_{h_{11}}\bar{p}_{h_{11}}$ change its sign.

The optimal control and corresponding state of our benchmark problem is illustrated in Fig. 1 and the measured discretization errors as well as the experimentally computed convergence rates are summarized in Table 1. As we have proven in Theorem 4 the numerical solutions obtained by a full discretization using a piecewise constant control approximation converge with the optimal convergence rate 1. Moreover, it is confirmed that the solution obtained with a postprocessing step, see Theorem 5, converges with order $2 - \varepsilon$, $\varepsilon > 0$. Note that we actually proved the results for the case that the boundary is smooth which is indeed not the case in our example. However, the corner singularities contained in the solution are for a 90°-corner comparatively mild so that the regularity results from Lemma 11 remain valid.

Acknowledgements Open Access funding provided by Projekt DEAL.

Appendix 1: Proof of Theorem 2

The proof of the maximum norm estimate presented in Theorem 2 follows basically from the arguments of [18, 38]. For the convenience of the reader we want to repeat the proof as the result of Theorem 2 is, for our specific situation, not directly available in the literature. The novelty of the present proof is that it includes curved elements as well as Robin boundary conditions. In the aforementioned articles, a representation of the error term based on a regularized Dirac function is used. This function forms the right-hand side of a dual problem whose solution is an approximation of Green's function. The main difficulty is to bound this solution in appropriate norms.

To this end, we denote by $T^* \in \mathcal{T}_h$ the element where $|y - y_h|$ attains its maximum. The regularized Dirac function is defined by $\delta^h(x) := |T^*|^{-1} \operatorname{sgn}(y(x) - y_h(x))$ if $x \in T^*$, and $\delta^h(x) := 0$ if $x \notin T^*$. The corresponding Green's function denoted by g^h solves the problem

$$-\Delta g^h + g^h = \delta^h \quad \text{in} \quad \Omega, \qquad \partial_n g^h + u\, g^h = 0 \quad \text{on} \quad \Gamma. \tag{60}$$

The Dirac function satisfies the properties

$$\|\delta^h\|_{L^1(\Omega)} \le c, \qquad \|\delta^h\|_{L^2(\Omega)} \le c\, h^{-1}. \tag{61}$$

We start our considerations with some a priori estimates for the solution g^h.

Lemma 22 *The following a priori estimates hold:*

$$(i) \quad \|g^h\|_{H^1(\Omega)} \le c\,|\ln h|^{1/2} \quad (ii)\ \|g^h\|_{H^2(\Omega)} \quad \le c\, h^{-1}$$
$$(iii) \quad \|g^h\|_{L^\infty(\Omega)} \le c\,|\ln h|$$

Proof (ii) To show the estimate in the $H^2(\Omega)$-norm we apply the a priori estimate from Lemma 3c) and $\|\delta^h\|_{L^2(\Omega)} \le ch^{-1}$.

(i) The weak form of (60) and the property (61) imply

$$\begin{aligned} \gamma_u\, \|g^h\|^2_{H^1(\Omega)} &\le a_u(g^h, g^h) = (\delta^h, g^h)_{L^2(\Omega)} \le c\, \|g^h\|_{L^\infty(\Omega)} \\ &\le c\big(\|g^h - g_h^h\|_{L^\infty(\Omega)} + |\ln h|^{1/2}\, \|g_h^h\|_{H^1(\Omega)}\big), \end{aligned} \tag{62}$$

where the discrete Sobolev inequality was applied in the last step. The function $g_h^h \in V_h$ is the Ritz-projection of g^h and satisfies the usual stability estimate

$$\|g_h^h\|_{H^1(\Omega)} \le c\,\|g^h\|_{H^1(\Omega)}. \tag{63}$$

Next, we derive a suboptimal error estimate for the finite-element error in the $L^\infty(\Omega)$ -norm. Using an inverse inequality, estimates for the interpolant Π_h from (19), the Aubin-Nitsche trick and the a priori estimate shown already in the $H^2(\Omega)$-norm we deduce

$$\begin{aligned}\|g^h - g_h^h\|_{L^\infty(\Omega)} &\le \|g^h - \Pi_h g^h\|_{L^\infty(\Omega)} + c\,h^{-1}\big(\|g^h - \Pi_h g^h\|_{L^2(\Omega)} + \|g^h - g_h^h\|_{L^2(\Omega)}\big)\\ &\le c\,h\,\|g^h\|_{H^2(\Omega)} \le c.\end{aligned} \tag{64}$$

Note that we hide the dependency on u, or more precisely on $\|u\|_{H^{1/2}(\Omega)}$ and lower-order norms, in the generic constant to simplify the notation. Insertion of (63) and (64) into (62) yields with Young's inequality

$$\gamma_u\,\|g^h\|^2_{H^1(\Omega)} \le c\,|\ln h| + \frac{1}{2}\,\gamma_u\,\|g^h\|^2_{H^1(\Omega)}.$$

The desired estimate follows form a kick-back-argument.

(iii) The $L^\infty(\Omega)$-estimate follows directly from (62), (63) and (64) using the inequality (i). □

Next, we show an a priori estimate for g^h in a weighted norm. This is the key idea which allows us to bound second derivatives by a logarithmic factor only. The weight function we will use is defined by

$$\sigma(x) := \sqrt{|x - x^*|_2^2 + c\,h^2},$$

with $x^* := \arg\max_{x \in T^*} |y - y_h|(x)$. This function satisfies

$$\|\sigma^{-1}\|_{L^2(\Omega)} \le c\,|\ln h|^{1/2}, \qquad \|\sigma^{-1}\|_{L^2(\Gamma)} \le c\,h^{-1/2}, \tag{65}$$

which follows from a simple computation. With this weight function at hand we can prove the following regularity result:

Lemma 23 *Assume that* $u \in H^{1/2}(\Gamma)$. *There holds the estimate*

$$\|\sigma\,\nabla^2 g^h\|_{L^2(\Omega)} \le c\,|\ln h|^{1/2}.$$

Proof We introduce the functions $\xi_i := |x_i - x_i^*|$, $i = 1, 2$, which allow us to write

$$\|\sigma\,\nabla^2 g^h\|^2_{L^2(\Omega)} = \sum_{i=1}^{2} \|\xi_i\,\nabla^2 g^h\|^2_{L^2(\Omega)} + c\,h^2\,\|\nabla^2 g^h\|^2_{L^2(\Omega)}. \tag{66}$$

With the reverse product rule we obtain

$$\|\xi_i\,\nabla^2 g^h\|^2_{L^2(\Omega)} \le \|\nabla^2(\xi_i\,g^h)\|^2_{L^2(\Omega)} + \|\nabla g^h\|^2_{L^2(\Omega)}. \tag{67}$$

Moreover, we easily confirm that $\xi_i\,g^h$ is the solution of the problem

$$\begin{aligned} -\Delta(\xi_i\,g^h) + \xi_i\,g^h &= -2\,\frac{\partial g^h}{\partial x_i} + \xi_i\,\delta^h && \text{in } \Omega,\\ \partial_n(\xi_i\,g^h) + u\,\xi_i\,g^h &= g^h\,n_i && \text{on } \Gamma, \end{aligned}$$

where n_i is the ith component of the outer unit normal vector on Γ. Lemma 3c) using the property $\|\xi_i\,\delta^h\|_{L^2(\Omega)} \le c$, which follows from a simple computation, leads to

$$\begin{aligned} \|\nabla^2(\xi_i\,g^h)\|_{L^2(\Omega)} &\le c\big(1 + \|u\|_{H^{1/2}(\Gamma)}\big)\big(1 + \|\nabla g^h\|_{L^2(\Omega)} + \|g^h\|_{H^{1/2}(\Gamma)}\big)\\ &\le c\big(1 + \|g^h\|_{H^1(\Omega)}\big). \end{aligned} \tag{68}$$

Insertion into (67) and using Lemma 22(i) leads to

$$\|\xi_i\,\nabla^2 g^h\|_{L^2(\Omega)} \le c\,|\ln h|^{1/2}, \quad i = 1, 2.$$

An estimate for the second term on the right-hand side of (66) is derived in Lemma 22(ii). ☐

Next, we derive some error estimates for the approximation g^h_h in several norms.

Lemma 24 *Assume that $u \in H^{1/2}(\Gamma)$. Then, there hold the error estimates*

$$\begin{aligned} h^{-1}\,\|g^h - g^h_h\|_{L^2(\Omega)} + \|\nabla(g^h - g^h_h)\|_{L^2(\Omega)} &\le c,\\ \|\sigma\nabla^2(g^h - g^h_h)\|_{L^2_{pw}(\Omega)} &\le c\,|\ln h|^{1/2}. \end{aligned}$$

Proof The first estimate follows directly from the $H^1(\Omega)$-error estimate stated in Lemma 12 and the Aubin-Nitsche trick. Moreover, the a priori estimate for the $H^2(\Omega)$-norm of g^h from Lemma 22 has to be exploited.

In the second estimate one observes that the discrete function g^h_h would vanish except on curved elements (note that g^h_h is affine on the reference element only, but not on T). With the transformation result [6, Lemma 2.3] we obtain

$$\begin{aligned} \|\nabla^2(g^h - g^h_h)\|_{L^2(T)} &\le c\,|T|^{1/2}\,h^{-2}\sum_{k=0}^{2} h^{4-2k}\,\|\hat\nabla^k(\hat g^h - \hat g^h_h)\|_{L^2(\hat T)}\\ &\le c\big(\|\nabla^2 g^h\|_{L^2(T)} + h\,\|g^h - g^h_h\|_{H^1(T)}\big), \end{aligned}$$

where $g^h = \hat g^h \circ F_T^{-1}$, $g^h_h = \hat g^h_h \circ F_T^{-1}$. Taking into account $\inf_{x\in T}\sigma(x) \sim \sup_{x\in T}\sigma(x)$, which holds due to the assumed shape-regularity, and $|\sigma(x)| \le c$ for all $x \in \Omega$, we obtain

$$\|\sigma\,\nabla^2(g^h - g^h_h)\|_{L^2_{pw}(\Omega)} \le c\big(\|\sigma\,\nabla^2 g^h\|_{L^2(\Omega)} + h\,\|g^h - g^h_h\|_{H^1(\Omega)}\big).$$

The first term has been discussed in the previous lemma and the last term has been considered in the present Lemma already. □

Now we are in the position to prove Theorem 2.

Proof With an inverse inequality and the Hölder inequality, the definition of δ^h and a maximum norm estimate for the interpolant Π_h, see e. g. [6, Theorem 4.1], we obtain

$$\begin{aligned} \|y - y_h\|_{L^\infty(\Omega)} &= \|y - y_h\|_{L^\infty(T^*)} \\ &\le c\big(\|y - \Pi_h y\|_{L^\infty(T^*)} + |T^*|^{-1}\,\|\Pi_h y - y_h\|_{L^1(T^*)}\big) \\ &\le c\big(\|y - \Pi_h y\|_{L^\infty(T^*)} + (\delta^h, y - y_h)_{L^2(\Omega)}\big) \\ &\le c\big(h^{2-2/q}\,\|y\|_{W^{2,q}(\Omega)} + a_u(y - y_h, g^h)\big), \end{aligned} \tag{69}$$

where $g_h^h \in V_h$ denotes the Ritz projection of g^h.

For the latter part on the right-hand side of (69) we get with the Galerkin orthogonality, the Hölder inequality, a trace theorem for the boundary integral term as well as $\|u\|_{L^\infty(\Gamma)} \le c$

$$\begin{aligned} a_u(y - y_h, g^h) &= a_u(y - \Pi_h y, g^h - g_h^h) \\ &\le c\,\|y - \Pi_h y\|_{W^{1,\infty}(\Omega)}\,\|g^h - g_h^h\|_{W^{1,1}(\Omega)}. \end{aligned} \tag{70}$$

An estimate for the interpolation error is deduced in [6]. The $L^1(\Omega)$-norms can be replaced by weighted $L^2(\Omega)$-norms involving the weighting function σ. Taking into account the properties (65) we obtain

$$\|g^h - g_h^h\|_{W^{1,1}(\Omega)} \le c\,|\ln h|^{1/2}\,\big(\|\sigma\,\nabla(g^h - g_h^h)\|_{L^2(\Omega)} + \|g^h - g_h^h\|_{L^2(\Omega)}\big). \tag{71}$$

In the following we will show that the expressions on the right-hand side of (71) are bounded by $c\,h\,|\ln h|^{1/2}$. Therefore, we apply the reverse product rule and get

$$\begin{aligned} &\|\sigma\,\nabla(g^h - g_h^h)\|^2_{L^2(\Omega)} \\ &\quad = (\nabla(\sigma^2\,(g^h - g_h^h)), \nabla(g^h - g_h^h))_{L^2(\Omega)} - ((g^h - g_h^h)\,\nabla\sigma^2, \nabla(g^h - g_h^h))_{L^2(\Omega)}. \end{aligned}$$

From this we conclude

$$\begin{aligned} \Theta^2 &:= \|\sigma\nabla(g^h - g_h^h)\|^2_{L^2(\Omega)} + \|g^h - g_h^h\|^2_{L^2(\Omega)} \\ &\le a_u(\sigma^2(g^h - g_h^h), g^h - g_h^h) - ((g^h - g_h^h)\nabla\sigma^2, \nabla(g^h - g_h^h))_{L^2(\Omega)}. \end{aligned} \tag{72}$$

Here, we exploited that $(u\,\sigma^2\,(g^h - g_h^h), g^h - g_h^h)_{L^2(\Gamma)} \ge 0$ due to $u \ge u_a \ge 0$. Next, we introduce the abbreviation $z := \sigma^2\,(g^h - g_h^h)$. The Galerkin orthogonality of $g^h - g_h^h$, Young's inequality and the trace theorem taking into account $|\sigma| + |\nabla\sigma| \le c$ yield

$$\|\sigma\,v\|_{L^2(\Gamma)} \le c\big(\|\sigma\,v\|_{L^2(\Omega)} + \|\nabla(\sigma\,v)\|_{L^2(\Omega)}\big) \le c\big(\|v\|_{L^2(\Omega)} + \|\sigma\,\nabla v\|_{L^2(\Omega)}\big)$$

and thus,

$$
\begin{aligned}
a_u(\sigma^2(g^h - g_h^h), g^h - g_h^h) &= a_u(z - \Pi_h z, g^h - g_h^h) \\
&\le \frac{1}{4}\,\Theta^2 + c\Big(\|\sigma^{-1}\,\nabla(z - \Pi_h z)\|^2_{L^2(\Omega)} \\
&\quad + \|\sigma^{-1}\,(z - \Pi_h z)\|^2_{L^2(\Omega)} + \|\sigma^{-1}\,u\,(z - \Pi_h z)\|^2_{L^2(\Gamma)} \Big).
\end{aligned} \tag{73}
$$

Next, we derive local interpolation error estimates. In the following we use the notation $\underline{\sigma}_T := \inf_{x\in T} \sigma(x)$ and $\overline{\sigma}_T := \sup_{x\in T} \sigma(x)$. Due to the assumed shape-regularity there holds $\underline{\sigma}_T \sim \overline{\sigma}_T$ for all $T \in \mathcal{T}_h$, and hence,

$$
\begin{aligned}
&\|\sigma^{-1}\,\nabla(z - \Pi_h z)\|_{L^2(T)} + h^{-1}\,\|\sigma^{-1}\,(z - \Pi_h z)\|_{L^2(T)} \\
&\quad \le c\,\underline{\sigma}_T^{-1}\,h\,\|\sigma^2\,(g^h - g_h^h)\|_{H^2(S_T)},
\end{aligned}
$$

where S_T is the patch of all elements adjacent to T (note that Π_h is a quasi-interpolant). The Leibniz rule and the properties $|\nabla\sigma^2| \le \sigma$ and $|\nabla^2\sigma^2| \le c$ imply

$$
\begin{aligned}
\|\sigma^2\,(g^h - g_h^h)\|_{H^2(S_T)} \le &\|g^h - g_h^h\|_{L^2(S_T)} + \|\sigma\,\nabla(g^h - g_h^h)\|_{L^2(S_T)} \\
&+ \|\sigma^2\,\nabla^2(g^h - g_h^h)\|_{L^2(S_T)}.
\end{aligned}
$$

Next, we combine the two estimates above and take into account the properties $h\,\underline{\sigma}_T^{-1} \le c$ and $\overline{\sigma}_T \sim \overline{\sigma}_{S_T}$ which follows from the assumed quasi-uniformity. Summation over all $T \in \mathcal{T}_h$ and an application of Lemma 24 yields

$$
\begin{aligned}
&\|\sigma^{-1}\nabla(z - \Pi_h z)\|_{L^2(\Omega)} + h^{-1}\,\|\sigma^{-1}(z - \Pi_h z)\|_{L^2(\Omega)} \\
&\quad \le c\Big(\|g^h - g_h^h\|_{L^2(\Omega)} + h\,\|\nabla(g^h - g_h^h)\|_{L^2(\Omega)} + h\,\|\sigma\,\nabla^2(g^h - g_h^h)\|_{L^2_{pw}(\Omega)} \Big) \\
&\quad \le c\,h\,|\ln h|^{1/2}.
\end{aligned} \tag{74}
$$

It remains to discuss the third term on the right-hand side of (73). With interpolation error estimates for Π_h on the boundary, compare also (20), and $u \in L^\infty(\Gamma)$ we obtain

$$
\begin{aligned}
\|\sigma^{-1}\,u\,(z - \Pi_h z)\|_{L^2(E)} &\le c\,h\,\underline{\sigma}_E^{-1}\,\|\nabla z\|_{L^2(S_E)} \\
&\le c\,h\big(\|g^h - g_h^h\|_{L^2(E)} + \|\sigma\,\nabla(g^h - g_h^h)\|_{L^2(S_E)} \big),
\end{aligned} \tag{75}
$$

where we exploited the product rule and the property $\nabla\sigma^2 \le 2\sigma\mathbf{1}$ in the last step. With a trace theorem and Lemma 24 we conclude

$$
\|g^h - g_h^h\|_{L^2(\Gamma)} \le c\,\|g^h - g_h^h\|_{H^1(\Omega)} \le c,
$$

and with a multiplicative trace theorem, Young's inequality, the product rule and the estimates from Lemma 24 we obtain

$$\begin{aligned}\|\sigma\,\nabla(g^h-g_h^h)\|_{L^2(\Gamma)} &\le c\Big(\|\sigma\,\nabla(g^h-g_h^h)\|_{L^2(\Omega)}+\|\nabla(\sigma\,\nabla(g^h-g_h^h))\|_{L^2_{pw}(\Omega)}\Big)\\ &\le c\Big(\|\nabla(g^h-g_h^h))\|_{L^2(\Omega)}+\|\sigma\,\nabla^2(g^h-g_h^h)\|_{L^2_{pw}(\Omega)}\Big)\\ &\le c\,|\ln h|^{1/2}.\end{aligned}$$

The estimate (75) then simplifies to

$$\|\sigma^{-1}\,(z-\Pi_h z)\|_{L^2(\Gamma)}\le c\,h\,|\ln h|^{1/2}. \tag{76}$$

Insertion of (74) and (76) into (73) leads to the estimate

$$a(\sigma^2(g^h-g_h^h),g^h-g_h^h)\le\frac{1}{4}\Theta^2+c\,h^2\,|\ln h|. \tag{77}$$

It remains to show an estimate for the second term on the right-hand side of (72). Due to $|\nabla\sigma^2|\le 2\sigma\mathbf{1}$, Young's inequality and the $L^2(\Omega)$-error estimate from Lemma 24 we get

$$\begin{aligned}((g^h-g_h^h)\nabla\sigma^2,\nabla(g^h-g_h^h))_{L^2(\Omega)} &\le c\,\|g^h-g_h^h\|^2_{L^2(\Omega)}+\frac{1}{4}\,\|\sigma\,\nabla(g^h-g_h^h)\|^2_{L^2(\Omega)}\\ &\le c\,h^2+\frac{1}{4}\,\Theta^2.\end{aligned} \tag{78}$$

Insertion of (77) and (78) into (72) yields

$$\Theta^2\le\frac{1}{2}\,\Theta^2+c\,h^2\,|\ln h| \tag{79}$$

and with a kick-back-argument we conclude $\Theta^2=c\,h^2\,|\ln h|$. Finally, we collect up the previous estimates. To this end, we insert (79) into (71), the resulting estimate into (70) and this into (69). □

Appendix 2: Local estimates for the midpoint interpolant and the $L^2(\Gamma)$-projection

To the best of the author's knowledge there are no error estimates for the midpoint interpolant defined on a curved boundary available in the literature. Thus, we prove the following Lemmata which are needed in the proof of Lemma 20.

Consider a single boundary element $E\subset\bar T$ with corresponding element $T\in\mathcal{T}_h$. A parametrization of the boundary element is given by $E:=\{\gamma_E(\xi):=F_T(\xi,0),\ \xi\in(0,1)\}$ when assuming that the edge of $\hat T$ with endpoints (0, 0), (1, 0) is mapped onto E. In the following we denote the length of a boundary element $E\in\mathcal{E}_h$ by $L_E=\int_0^1|\dot\gamma_E(\xi)|\,\mathrm{d}\xi$.

Lemma 25 *For each function $u : \Gamma \to \mathbb{R}$ there exists some piecewise constant function $R_h u \in U_h$ satisfying the local estimates*

$$\left|\int_E (u - R_h u)\right| \le c\, h^{5/2}\big(\|\nabla u\|_{L^2(E)} + \|\nabla^2 u\|_{L^2(E)}\big),$$
$$\|u - R_h u\|_{L^r(E)} \le c\, h^{1+1/r-1/q}\, \|\nabla u\|_{L^q(E)}, \quad r \in [1,\infty],\ q \in (1,\infty],$$

for all $E \in \mathcal{E}_h$, provided that u possesses the regularity demanded by the right-hand side.

Proof Let us first construct a suitable interpolation operator. To obtain the desired second-order accuracy we have to guarantee that the property $\int_E p = \int_E R_h p$ holds for all functions $p(\gamma_E(\xi)) = \hat{p}(\xi)$ with some first-order polynomial $\hat{p}(\xi) := a + b\,\xi$. The transformation to $\hat{E} := (0,1) \times \{0\}$ yields

$$\int_E (p(x) - R_h p)\mathrm{d}s_x = \int_0^1 (\hat{p}(\xi) - \hat{p}(\xi_E))\, |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi = b \int_0^1 (\xi - \xi_E)\, |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi = 0.$$

The latter step holds true when choosing

$$\xi_E := \frac{1}{L_E} \int_0^1 \xi\, |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi$$

To this end, we define our operator by $R_h u|_E := (u \circ \gamma_E)(\xi_E)$. Obviously, the definition of R_h depends on the transformations F_T.

To show the interpolation error estimates we apply the property $\int_E (p - R_h p) = 0$ for arbitrary p satisfying $\hat{p} = p \circ \gamma_E \in \mathcal{P}_1$, the stability of the interpolant $\hat{R}_h \hat{u} = \hat{u}(\xi_E)$, the properties (18) of the transformation F_T and the Bramble–Hilbert Lemma. This yields

$$\begin{aligned}\int_E (u - R_h u)\mathrm{d}s_x &= \int_0^1 (I - \hat{R}_h)(\hat{u} - \hat{p})(\xi)\, |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi \\ &\le c\, h\, \|\hat{u} - \hat{p}\|_{L^\infty(\hat{E})} \le c\, h\, \|\partial_{\xi\xi}\hat{u}\|_{L^2(\hat{E})}.\end{aligned}$$

For the transformation back to the world element E we apply the chain rule

$$\partial_{\xi\xi}\hat{u}(\xi) = \dot{\gamma}_E(\xi)^\top\, \nabla^2 u(\gamma_E(\xi))\, \dot{\gamma}_E(\xi) + \ddot{\gamma}_E(\xi)^\top\, \nabla u(\gamma_E(\xi))$$

and the properties (18) to arrive at

$$\begin{aligned}\|\partial_{\xi\xi}\hat{u}\|_{L^2(\hat{E})} &\le c\, h^2 \left(\max_\xi |\dot{\gamma}_E(\xi)|^{-1} \sum_{|\alpha|=1}^{2} \int_0^1 (D^\alpha u(\gamma_E(\xi)))^2\, |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi \right)^{1/2} \\ &\le c\, h^2 \left(\min_\xi |\dot{\gamma}_E(\xi)| \right)^{-1/2} \big(\|\nabla u\|_{L^2(E)} + \|\nabla^2 u\|_{L^2(E)}\big).\end{aligned}$$

Finally, the norm of $\dot{\gamma}_E$ can be bounded by means of

$$|\dot{\gamma}_E(\xi)|^{-1} = |DF_T(\xi,0)(1,0)^\top|^{-1} \le \left(\min_{|x|=1} |DF_T(\xi,0)x|\right)^{-1} = \|DF_T(\xi,0)^{-1}\|. \tag{80}$$

Note, that the last step is valid for the spectral norm only.

An application of Lemma 2.2 from [6] which provides $\sup_{\hat{x}\in\hat{T}} \|DF_T(\hat{x})^{-1}\| \le ch^{-1}$ leads to the first estimate.

The second estimate follows with similar arguments. For an arbitrary constant $\hat{p}$ we then obtain

$$\begin{aligned}\|u - R_h u\|_{L^r(E)} &\le c\, h^{1/r}\, \|(I - \hat{R}_h)(\hat{u} - \hat{p})\|_{L^r(\hat{E})} \\ &\le c\, h^{1/r}\, |\hat{u}|_{W^{1,q}(\hat{E})} \le c\, h^{1+1/r-1/q}\, \|\nabla u\|_{L^q(E)}.\end{aligned}$$

□

A further operator that is needed in Sect. 3 is the $L^2(\Gamma)$-projection onto U_h. In case of curved boundaries, this operator reads

$$[Q_h v]|_E := \frac{1}{L_E} \int_0^1 v(\gamma_E(\xi))\, |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi$$

for each $E \in \mathcal{E}_h$. Again, this definition depends on the parametrizations γ_E of the boundary elements $E \in \mathcal{E}_h$. By a simple computation one can show that this definition implies the orthogonality property

$$(u - Q_h u, v_h)_{L^2(\Gamma)} = 0 \qquad \forall v_h \in U_h. \tag{81}$$

With similar arguments as in the previous lemma we obtain the following local estimate which is standard in case of a boundary consisting of straight edges.

Lemma 26 *Assume that $u \in H^1(\Gamma)$. Then the estimate*

$$\|u - Q_h u\|_{L^2(E)} \le c\, h\, \|\nabla u\|_{L^2(E)}$$

is fulfilled for all $E \in \mathcal{E}_h$.

Proof We introduce a further projection onto U_h, namely $[\tilde{Q}_h u]|_E := \int_0^1 u(\gamma_E(\xi))\, \mathrm{d}\xi$. Using (81), the transformation to the reference element as in the previous lemma and the Poincaré inequality we obtain

$$\begin{aligned}\|u - Q_h u\|^2_{L^2(E)} &\le \|u - \tilde{Q}_h u\|^2_{L^2(E)} \\ &= \int_0^1 \left(u(\gamma_E(\xi)) - \int_0^1 u(\gamma_E(\xi'))\, \mathrm{d}\xi'\right)^2 |\dot{\gamma}_E(\xi)|\, \mathrm{d}\xi \\ &\le c\, h\, \|\partial_\xi u(\gamma_E(\cdot))\|^2_{L^2(0,1)} \le c\, h^2\, \|\nabla u\|^2_{L^2(E)},\end{aligned}$$

where the last step is a consequence of the chain rule $\partial_\xi u(\gamma_E(\xi)) = \nabla u(\gamma_E(\xi)) \cdot \dot{\gamma}_E(\xi)$ and $|\dot{\gamma}_E(\xi)| \sim h$ for all $\xi \in (0,1)$, see also (80). □

We conclude this section with an estimate for an expression which is need in Lemma 20.

Lemma 27 *Assume that the functions u and v belong to* $\in H^1(\Gamma)$. *Then the inequality*

$$(u - R_h u, v)_{L^2(\Gamma)} \le c\, h^2\, \|\nabla u\|_{L^2(\Gamma)}\, \|\nabla v\|_{L^2(\Gamma)} + c\, \|v\|_{L^\infty(\Gamma)} \sum_{E\in\mathcal{E}_h} \left|\int_E (u - R_h u)\right|$$

is valid.

Proof First, we split the term under consideration using the $L^2(\Gamma)$-projection onto U_h and obtain

$$(u - R_h u, v)_{L^2(\Gamma)} = (u - Q_h u, v - Q_h v)_{L^2(\Gamma)} + (Q_h(u - R_h u), v)_{L^2(\Gamma)}.$$

The first term on the right-hand side can be treated with the local estimate from Lemma 26 which yields

$$(u - Q_h u, v - Q_h v)_{L^2(\Gamma)} \le c\, h^2\, \|\nabla u\|_{L^2(\Gamma)}\, \|\nabla v\|_{L^2(\Gamma)}.$$

For the second term we exploit the definition of Q_h and R_h on the reference element. For each $E \in \mathcal{E}_h$ we then obtain

$$\begin{aligned}\|Q_h(u - R_h u)\|_{L^1(E)} &= \int_0^1 \left|\frac{1}{L_E}\int_0^1 (u(\gamma_E(\xi)) - u(\gamma_E(\xi_E)))\, |\dot\gamma_E(\xi)|\, \mathrm{d}\xi\right| |\dot\gamma_E(\xi')|\, \mathrm{d}\xi' \\ &= \left|\int_0^1 (u(\gamma_E(\xi)) - u(\gamma_E(\xi_E)))\, |\dot\gamma_E(\xi)|\, \mathrm{d}\xi\right| = \left|\int_E (u - R_h u)\right|,\end{aligned}$$

where we used $\int_0^1 |\dot\gamma_E(\xi')|\, \mathrm{d}\xi' = L_E$ in the second step. Consequently, we obtain

$$(Q_h(u - R_h u), v)_{L^2(\Gamma)} \le c\, \|v\|_{L^\infty(\Gamma)} \sum_{E\in\mathcal{E}_h} \left|\int_E (u - R_h u)\right|$$

and conclude the assertion. □

References

1. Apel, Th: Anisotropic Finite Elements: Local Estimates and Applications. Teubner, Stuttgart (1999)
2. Apel, Th, Lombardi, A.L., Winkler, M.: Anisotropic mesh refinement in polyhedral domains: error estimates with data in $L^2(\Omega)$. ESAIM Math. Model. Numer. Anal. **48**(4), 1117–1145 (2014). https://doi.org/10.1051/m2an/2013134
3. Apel, Th, Pfefferer, J., Rösch, A.: Finite element error estimates on the boundary with application to optimal control. Math. Comp. **84**, 33–70 (2015). https://doi.org/10.1090/S0025-5718-2014-02862-7

4. Apel, Th, Pfefferer, J., Winkler, M.: Error estimates for the postprocessing approach applied to Neumann boundary control problems in polyhedral domains. IMA J. Numer. Anal. **38**(4), 1984–2025 (2018). https://doi.org/10.1093/imanum/drx059
5. Arada, N., Casas, E., Tröltzsch, F.: Error estimates for the numerical approximation of a semilinear elliptic control problem. Comput. Optim. Appl. **23**(2), 201–229 (2002). https://doi.org/10.1023/A:1020576801966
6. Bernardi, C.: Optimal finite-element interpolation on curved domains. SIAM J. Numer. Anal. **26**(5), 1212–1240 (1989). https://doi.org/10.1137/0726068
7. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd ed. Texts in Applied Mathematics. Springer, New York (2008)
8. Casas, E.: Boundary control of semilinear elliptic equations with pointwise state constraints. SIAM J. Control Optim. **31**(4), 993–1006 (1993). https://doi.org/10.1137/0331044
9. Casas, E., Mateos, M.: Error estimates for the numerical approximation of Neumann control problems. Comput. Optim. Appl. **39**(3), 265–295 (2008). https://doi.org/10.1007/s10589-007-9056-6
10. Casas, E., Mateos, M.: Uniform convergence of the FEM. Applications to state constrained control problems. Comput. Appl. Math. **21**(1), 67–100 (2002)
11. Casas, E., Mateos, M., Tröltzsch, F.: Error estimates for the numerical approximation of boundary semilinear elliptic control problems. Comput. Optim. Appl. **31**(2), 193–219 (2005). https://doi.org/10.1007/s10589-005-2180-2
12. Chaabane, S., Jaoua, M.: Identification of Robin coefficients by the means of boundary measurements. Inverse Probl. **15**(6), 1425–1438 (1999)
13. Ciarlet, P.G.: Basic error estimates for elliptic problems. In: Ciarlet, P.G., Lions, J.L. (eds.) Finite Element Methods, vol. 2. Handbook of Numerical Analysis, pp. 17–352. Elsevier, North-Holland (1991)
14. Dauge, M.: Elliptic Boundary Value Problems on Corner Domains. Springer, Berlin (1988). https://doi.org/10.1007/BFb0086682
15. Dhamo, V.: Optimal Boundary Control of Quasilinear Elliptic Partial Diffierential Equations: Theory and Numerical Analysis. PhD thesis. TU Berlin (2012)
16. Egger, H., et al.: Analysis and numerical solution of coupled volume-surface reaction–diffiusion systems with application to cell biology. Appl. Math. Comput. **336**, 351–367 (2018). https://doi.org/10.1016/j.amc.2018.04.031. ISSN: 0096-3003
17. Fellner, K., Rosenberger, S., Tang, B.Q.: Quasi-steady-state approximation and numerical simulation for a volume-surface reaction–diffiusion system. Commun. Math. Sci. **14**(6), 1553–1580 (2016). https://doi.org/10.4310/cms.2016.v14.n6.a5
18. Frehse, J., Rannacher, R.: Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente. Bonn. Math. Schr. **89**, 92–114 (1976)
19. Gesztesy, F., Mitrea, M.: A description of all self-adjoint extensions of the Laplacian and Krein-type resolvent formulas on non-smooth domains. J. Anal. Math. **113**, 53–172 (2011). https://doi.org/10.1007/s11854-011-0002-2
20. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Pitman, Boston (1985)
21. Gwinner, J.: On two-coefficient identification in elliptic variational inequalities. Optimization **67**(7), 1017–1030 (2018). https://doi.org/10.1080/02331934.2018.1446955
22. Hào, D.N., Thanh, P.X., Lesnic, D.: Determination of the heat transfer coefficients in transient heat conduction. IOP Inverse Probl. (2013). https://doi.org/10.1088/0266-5611/29/9/095020
23. Hetmaniok, E., et al.: Identification of the heat transfer coefficient in the two-dimensional model of binary alloy solidification. Heat Mass Transf. **53**(5), 1657–1666 (2017). https://doi.org/10.1007/s00231-016-1923-1
24. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. Comput. Optim. Appl. **30**(1), 45–61 (2005). https://doi.org/10.1007/s10589-005-4559-5
25. Jin, B., Lu, X.: Numerical identification of a Robin coefficient in parabolic problems. Math. Comp. **81**(279), 1369–1398 (2012). https://doi.org/10.1090/S0025-5718-2012-02559-2
26. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications, Vol. 88. Pure and Applied Mathematics. Academic Press, New York (1980)
27. Kröner, A., Vexler, B.: A priori error estimates for elliptic optimal control problems with a bilinear state equation. J. Comput. Appl. Math. **230**(2), 781–802 (2009). https://doi.org/10.1016/j.cam.2009.01.023

28. Krumbiegel, K., Meyer, C., Rösch, A.: A priori error analysis for linear quadratic elliptic Neumann boundary control problems with control and state constraints. SIAM J. Control Optim. **48**(8), 5108–5142 (2010). https://doi.org/10.1137/090746148. ISSN: 0363-0129
29. Krumbiegel, K., Pfefferer, J.: Superconvergence for Neumann boundary control problems governed by semilinear elliptic equations. Comput. Optim. Appl. **61**(2), 373–408 (2015). https://doi.org/10.1007/s10589-014-9718-0
30. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in L ^{2} for a class of evolution equations. SIAM J. Control Optim. **46**(5), 1726–1753 (2007). https://doi.org/10.1137/060670110. ISSN: 0363-0129
31. Liu, J., Nakamura, G.: Recovering the boundary corrosion from electrical potential distribution using partial boundary data. Inverse Probl. Imaging **11**(3), 521–538 (2017). https://doi.org/10.3934/ipi.2017024
32. Mateos, M., Rösch, A.: On saturation effects in the Neumann boundary control of elliptic optimal control problems. Comput. Optim. Appl. **49**(2), 359–378 (2011). https://doi.org/10.1007/s10589-009-9299-5
33. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. SIAM J. Control Optim. **43**(3), 970–985 (2004). https://doi.org/10.1137/S0363012903431608
34. Mohebbi, F., Sellier, M.: Identification of space- and temperature-dependent heat transfer coefficient. Int. J. Therm. Sci. **128**, 28–37 (2018). https://doi.org/10.1016/j.ijthermalsci.2018.02.007
35. Rösch, A., Tröltzsch, F.: An optimal control problem arising from the identification of nonlinear heat transfer laws. Pol. Acad. Sci. Comm. Autom. Control Robot. Arch. Control Sci. **1**(3–4), 183–195 (1992)
36. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comp. **54**(190), 483–493 (1990). https://doi.org/10.2307/2008497
37. Scott, R.: Finite Element Techniques for Curved Boundaries. PhD thesis. MIT (1973)
38. Scott, R.: Optimal L^∞ estimates for the finite element method on irregular meshes. Math. Comp. **30**, 681–697 (1976). https://doi.org/10.2307/2005390
39. Tröltzsch, F.: Optimal Control of Partial Diffierential Equations: Theory, Methods, and Applications. Graduate Studies in Mathematics. American Mathematical Society (2010). ISBN: 978-0-82-184904-0
40. Winkler, G.: Control Constrained Optimal Control Problems in Non-convex Three Dimensional Polyhedral Domains. PhD Thesis. TU Chemnitz (2008)
41. Zlamal, M.: Curved elements in the finite element method. I. SIAM J. Numer. Anal. **10**, 229–240 (1973). https://doi.org/10.1137/0710022