TECHNISCHE UNIVERSITÄT CHEMNITZ

René Schneider Gerd Wachsmuth

A-posteriori error estimation for control-constrained, linear-quadratic optimal control problems

CSC/15-04

Chemnitz Scientific Computing Preprints

Impressum:

Chemnitz Scientific Computing Preprints — ISSN 1864-0087
(1995–2005: Preprintreihe des Chemnitzer SFB393)

Herausgeber:
Professuren für
Numerische und Angewandte Mathematik
an der Fakultät für Mathematik
der Technischen Universität Chemnitz

Postanschrift:
TU Chemnitz, Fakultät für Mathematik
09107 Chemnitz
Sitz:
Reichenhainer Str. 41, 09126 Chemnitz

http://www.tu-chemnitz.de/mathematik/csc/

TECHNISCHE UNIVERSITÄT CHEMNITZ

Chemnitz Scientific Computing Preprints

René Schneider Gerd Wachsmuth

A-posteriori error estimation for control-constrained, linear-quadratic optimal control problems

CSC/15-04

CSC/15-04 ISSN 1864-0087 May 2015

Author's addresses:

René Schneider
Gerd Wachsmuth
TU Chemnitz, Fakultät für Mathematik
09107 Chemnitz, Germany

http://www.tu-chemnitz.de/mathematik/

A-posteriori error estimation for control-constrained, linear-quadratic optimal control problems

René Schneider* Gerd Wachsmuth†

April 27, 2015

We derive a-posteriori error estimates for control-constrained, linear-quadratic optimal control problems. The error is measured in a norm which is motivated by the objective. The error estimator is separated into three contributions: the error in the variational inequality (i.e., in the optimality condition for the control) and the errors in the state and adjoint equation. Hence, one can use well-established estimators for the differential equations. We show that the error estimator is reliable and efficient. We apply the error estimator to two distributed optimal control problems with distributed and boundary observation, respectively. Numerical examples exhibit a good error reduction if we use the local error contributions for an adaptive mesh refinement.

1 Introduction

We consider the a-posteriori error analysis of a control-constrained, linear-quadratic optimal control problem. To present the ideas, we use the problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|y - y_d\|^2_{L^2(\partial\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ \text{such that} \quad & -\Delta y + y = u \quad \text{in } \Omega \\ & \frac{\partial}{\partial n} y = 0 \quad \text{on } \partial\Omega \\ \text{and} \quad & u_a \le u \le u_b \end{aligned} \tag{1.1}$$

as an example, where u and y denote the unknown control and state respectively, y_d the given desired state, Ω the given PDE domain and $\partial\Omega$ its boundary. Note that the state y is only observed on the boundary $\partial\Omega$. This example and the assumptions on the data are discussed in more detail in Section 4. We emphasize that our technique is applicable to a general class of optimal control problems, see Section 2.

Our goal is as follows. Let $(\bar y, \bar u, \bar p)$ be the unique solution of the optimality system, see (2.3), where $\bar p$ is the (optimal) adjoint state. Given an approximate solution (y_h, u_h, p_h), we are interested in estimating

*TU Chemnitz, Faculty of Mathematics, Research Group Numerical Mathematics (Numerical Analysis)
†TU Chemnitz, Faculty of Mathematics, Research Group Numerical Mathematics (Partial Differential Equations)

the error in the control variable, that is, $\|\bar{u} - u_h\|_{L^2(\Omega)}$. It is, however, not clear how to bound this error by a reliable and efficient a-posteriori estimator. Instead, we estimate a combination of all quantities involved in the optimality system, that is, a (weighted) sum of $\|\bar{y} - y_h\|_{L^2(\partial\Omega)}$, $\|\bar{u} - u_h\|_{L^2(\Omega)}$, $\|\bar{p} - p_h\|_{L^2(\Omega)}$. Note that the error in the state is measured in $L^2(\partial\Omega)$, whereas the error in the control and in the adjoint state is measured in $L^2(\Omega)$. The former norm is the norm of the observation in the objective, whereas the latter one is the norm of the regularization term acting on the control.

The error estimator is separated into three contributions, which are related to the residuals in the three equations of the optimality system (2.3). That is, one term measures the defect in the variational inequality and two terms estimate the error in the state and adjoint equation. For the above example (1.1), these latter terms are

$$\|A^{-1} u_h - y_h\|_{L^2(\partial\Omega)} \qquad \text{and} \qquad \|A^{-\star} (y_d - y_h) - p_h\|_{L^2(\Omega)}, \tag{1.2}$$

where $A = (-\Delta + 1) : H^1(\Omega) \to (H^1(\Omega))'$ is the partial differential operator. Note that these two errors are estimated in the same norms in which the errors in y and p are measured, respectively. Moreover, these two error terms can now be replaced by standard error estimates. In our numerical examples, we use residual-based error estimators.

A similar approach is used in Kohls et al. [2012, 2014]. In difference to our work, the error in the state and adjoint state is measured and estimated in the energy space corresponding to the state equation. That is, in the case of problem (1.1), the errors are measured as $\|\bar{y} - y_h\|_{H^1(\Omega)}$, $\|\bar{p} - p_h\|_{H^1(\Omega)}$, and the error estimate involves the terms

$$\|A^{-1} u_h - y_h\|_{H^1(\Omega)} \qquad \text{and} \qquad \|A^{-\star} (y_d - y_h) - p_h\|_{H^1(\Omega)}.$$

Other contributions concerning a-posteriori error analysis for control-constrained optimal control problems are, e.g., Liu and Yan [2001], Hintermüller et al. [2008], Yan and Zhou [2008], see also the references therein. In these papers, however, the error in the state and adjoint state is typically measured in the energy space and only specific optimal control problems are studied. We also mention the contributions Becker et al. [2007], Vexler and Wollner [2008], in which the authors estimate the error in the objective or in a quantity of interest using the dual-weighted-residual method.

As we will demonstrate, we can get substantially better results for $\|\bar{u} - u_h\|_{L^2(\Omega)}$ using the norms in (1.2) instead of the energy norms.

The paper is organized as follows. In Section 2, we introduce an abstract control problem and derive the a-posteriori error estimates, see in particular Theorem 2.4 and Theorem 2.6. In Section 3 and Section 4 this theory is applied to two example problems similar to (1.1). Finally, we conclude and give some perspectives in Section 5.

2 Abstract linear-quadratic control problem

In this section, we discuss an abstract linear-quadratic control problem with pointwise control constraints. First, we recall some preliminary, well-known results concerning existence and optimality conditions. In Section 2.2, we construct the a-posteriori error estimator.

2.1 Preliminaries: existence and optimality conditions

The abstract optimal control problem which is discussed in this section is given by

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\,\|C\,y - z_d\|_Z^2 + \frac{\alpha}{2}\,\|u\|_{L^2(\Omega_u)}^2 \\ \text{with respect to} \quad & (y,u) \in Y \times L^2(\Omega_u), \\ \text{such that} \quad & A\,y - B\,u = 0 \\ \text{and} \quad & u_a \le u \le u_b \quad \text{a.e. in } \Omega_u. \end{aligned} \tag{2.1}$$

Here, Y is a reflexive Banach space, Z is a Hilbert space, and Ω_u is a finite measure space. In the sequel, we abbreviate $U = L^2(\Omega_u)$. The linear operator $A : Y \to Y'$ is an isomorphism and $B : L^2(\Omega_u) \to Y'$, $C : Y \to Z$ are bounded linear operators. The desired observation is $z_d \in Z$. The control bounds u_a, u_b belong to $L^2(\Omega_u)$ and satisfy $u_a \le u_b$. We set

$$U_{\text{ad}} := \{u \in L^2(\Omega_u) : u_a \le u \le u_b\}.$$

The regularization parameter α is assumed to be positive.

In the abstract problem (2.1), the operator A models the linear (partial) differential operator, B is the control operator and C the observation operator.

As examples for the measure space Ω_u, we mention (a subset of) Ω in case of distributed control or (a subset of) the boundary $\partial\Omega$ in case of boundary control, where $\Omega \subset \mathbb{R}^d$ is the domain of the partial differential equation (PDE). Also the case of a finite-dimensional control space $U = \mathbb{R}^n$ is possible if we take a finite set $\Omega_u = \{1, \ldots, n\}$ equipped with the counting measure.

Let us define the control-to-observation operator $\mathcal{S} = C\,A^{-1}\,B \in \mathcal{L}(U, Z)$, which is well defined since A is an isomorphism. We obtain the reduced problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\,\|\mathcal{S}\,u - z_d\|_Z^2 + \frac{\alpha}{2}\,\|u\|_U^2, \\ \text{with respect to} \quad & u \in U \\ \text{such that} \quad & u_a \le u \le u_b \quad \text{a.e. in } \Omega_u, \end{aligned} \tag{2.2}$$

which is equivalent to (2.1). Indeed, $\bar u$ is a solution of (2.2) if and only if $(A^{-1}\,B\,\bar u, \bar u)$ is a solution of (2.1).

Let us give some well-known results concerning the problems (2.1) and (2.2). The proofs are standard and, hence, omitted. We refer to Tröltzsch [2009] for an introduction.

Lemma 2.1. There exists a unique solution $(\bar y, \bar u) \in Y \times U$ of (2.1).

In what follows, we will identify the Hilbert spaces U and Z with their duals. Using the adjoint operators

$$B^\star : Y \to U, \qquad A^\star : Y \to Y', \qquad C^\star : Z \to Y',$$

we can write down the optimality conditions of first order of (2.1).

Lemma 2.2. Let $(\bar y, \bar u) \in Y \times U$ be the unique solution of (2.1). Then, there exists a unique adjoint state $\bar p \in Y$, such that the system

$$\bar y = A^{-1}\,B\bar u \tag{2.3a}$$

$$\bar p = A^{-\star}\,C^\star\,(z_d - C\,\bar y) \tag{2.3b}$$

$$(\alpha\,\bar u - B^\star\,\bar p,\, u - \bar u)_U \ge 0 \qquad \text{for all } u \in U_{\text{ad}} \tag{2.3c}$$

is satisfied.

Due to the convexity of problem (2.1) these conditions are also sufficient for optimality. It is well-known, that the variational inequality (2.3c) is equivalent to the pointwise projection formula

$$\bar u(x) = \operatorname{Proj}_{[u_a(x), u_b(x)]} \frac{B^\star\,\bar p(x)}{\alpha} \quad \text{for almost all } x \in \Omega_u, \tag{2.4}$$

see, e.g., [Tröltzsch, 2009, Lemma 2.26]. Here, $\operatorname{Proj}_{[a,b]}\, c$ is the projection of $c \in \mathbb{R}$ onto the interval $[a, b] \subset \mathbb{R}$.

By using the adjoint $\mathcal{S}^\star : Z \to U$ of the control-to-observation operator, we can write down the optimality conditions of first order of (2.2) in a slightly different form, which will be more convenient for the error analysis later in Section 2.2.

Lemma 2.3. Let $\bar u \in U$ be the unique solution of (2.2). Then, there exist unique $\bar z \in Z$ and $\bar\rho \in U$, such that the system

$$\bar z = \mathcal{S}\,\bar u \tag{2.5a}$$

$$\bar\rho = \mathcal{S}^\star\,(\bar z - z_d) \tag{2.5b}$$

$$(\alpha\,\bar u - \bar\rho,\, u - \bar u)_U \ge 0 \quad \text{for all } u \in U_{\text{ad}} \tag{2.5c}$$

is satisfied.

The quantity $\bar z \in Z$ is the optimal observation and $\bar\rho \in U$ is the required information on the optimal adjoint state for the variational inequality (2.3c), $\bar\rho = B^\star\,\bar p$. In the literature on optimal control of ordinary differential equations, this $\bar\rho$ is usually called switching function.

2.2 Error estimator

In this section, we want to give a (computable) error estimate which measures the distance of any triple $(z_h, u_h, \rho_h) \in Z \times U \times U$ to the unique solution $(\bar z, \bar u, \bar\rho)$ of the optimality system given in (2.5). Later in Section 3 and Section 4, this triple (z_h, u_h, ρ_h) will result from a finite element discretization.

To this end, we fix an arbitrary tuple $(\tilde u_h, \tilde\rho_h) \in U_{\text{ad}} \times U$, such that the variational inequality

$$(\alpha\,\tilde u_h - \tilde\rho_h,\, u - \tilde u_h)_U \ge 0 \qquad \text{for all } u \in U_{\text{ad}} \tag{2.6}$$

is satisfied. By using $u = \tilde u_h$ in (2.3c) and $u = \bar u$ in (2.6), we obtain

$$\alpha\,\|\tilde u_h - \bar u\|_U^2 \le (\bar\rho,\, \bar u - \tilde u_h)_U + (-\tilde\rho_h,\, \bar u - \tilde u_h)_U. \tag{2.7}$$

On the other hand, using the cosine theorem in the Hilbert space Z, we get

$$\begin{aligned}\frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - \bar z\|_Z^2 + \frac{1}{2}\,\|z_h - \bar z\|_Z^2 &= \frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - z_h\|_Z^2 + (z_h - \bar z,\, \mathcal{S}\,\tilde u_h - \bar z)_Z \\ &= \frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - z_h\|_Z^2 + \big(\mathcal{S}^\star\,(z_h - z_d) + \bar\rho,\, \tilde u_h - \bar u\big)_U.\end{aligned}$$

Together with (2.7), this leads to

$$\frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - \bar z\|_Z^2 + \frac{1}{2}\,\|z_h - \bar z\|_Z^2 + \alpha\,\|\tilde u_h - \bar u\|_U^2 \le \big(\tilde\rho_h - \mathcal{S}^\star\,(z_d - z_h),\, \tilde u_h - \bar u\big)_U + \frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - z_h\|_Z^2.$$

An application of Young's inequality implies

$$\frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - \bar z\|_Z^2 + \frac{1}{2}\,\|z_h - \bar z\|_Z^2 + \frac{\alpha}{2}\,\|\tilde u_h - \bar u\|_U^2 \le \frac{1}{2\,\alpha}\,\|\tilde\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 + \frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - z_h\|_Z^2.$$

Together with the estimate

$$\|u_h - \bar u\|_U^2 \le 2\,\|\tilde u_h - u_h\|_U^2 + 2\,\|\tilde u_h - \bar u\|_U^2$$

we arrive at

$$\frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - \bar z\|_Z^2 + \frac{1}{2}\,\|z_h - \bar z\|_Z^2 + \frac{\alpha}{4}\,\|u_h - \bar u\|_U^2 \le \frac{1}{2\,\alpha}\,\|\tilde\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 + \frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - z_h\|_Z^2 + \frac{\alpha}{2}\,\|\tilde u_h - u_h\|_U^2. \tag{2.8}$$

In this estimate, only the known quantities z_h, u_h, ρ_h and the (still) arbitrary $\tilde u_h$ appear on the right-hand side. Now, we introduce the abbreviation

$$L := \frac{\|\mathcal{S}\|_{\mathcal{L}(U,Z)}^2}{\alpha} = \frac{\|\mathcal{S}^\star\|_{\mathcal{L}(Z,U)}^2}{\alpha}. \tag{2.9}$$

The estimates

$$\frac{1}{2\,\alpha}\,\|\tilde\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 \le \frac{1}{\alpha}\,\|\tilde\rho_h - \rho_h\|_U^2 + \frac{1}{\alpha}\,\|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2,$$

$$\frac{1}{2}\,\|\mathcal{S}\,\tilde u_h - z_h\|_Z^2 \le \|\mathcal{S}\,\tilde u_h - \mathcal{S}\,u_h\|_Z^2 + \|\mathcal{S}\,u_h - z_h\|_Z^2 \le \alpha\,L\,\|\tilde u_h - u_h\|_U^2 + \|\mathcal{S}\,u_h - z_h\|_Z^2,$$

together with (2.8) imply

$$\begin{aligned}&\frac{1}{2}\,\|z_h - \bar z\|_Z^2 + \frac{\alpha}{4}\,\|u_h - \bar u\|_U^2 \\ &\quad\le \frac{1}{\alpha}\,\|\tilde\rho_h - \rho_h\|_U^2 + \frac{1}{\alpha}\,\|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 + \alpha\,(2^{-1} + L)\,\|\tilde u_h - u_h\|_U^2 + \|\mathcal{S}\,u_h - z_h\|_Z^2.\end{aligned}$$

Using additionally

$$\begin{aligned}\frac{1}{4}\,\|\rho_h - \bar\rho\|_U^2 &\le \frac{1}{2}\,\|\rho_h - \mathcal{S}^\star\,(z_h - z_d)\|_U^2 + \frac{1}{2}\,\|\mathcal{S}^\star\,(z_h - \bar z)\|_U^2 \\ &\le \frac{1}{2}\,\|\rho_h - \mathcal{S}^\star\,(z_h - z_d)\|_U^2 + \frac{L\,\alpha}{2}\,\|z_h - \bar z\|_Z^2,\end{aligned}$$

we arrive at

$$\begin{aligned}&\frac{1}{2}\,\|z_h - \bar z\|_Z^2 + \frac{\alpha}{4}\,\|u_h - \bar u\|_U^2 + \frac{1}{4\,\alpha}\|\rho_h - \bar\rho\|_U^2 \\ &\quad\le \frac{1+L}{\alpha}\,\|\tilde\rho_h - \rho_h\|_U^2 + \frac{1\frac{1}{2} + L}{\alpha}\,\|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 \\ &\qquad + \alpha\,\big(\frac{1}{2} + L\big)\,(1+L)\,\|\tilde u_h - u_h\|_U^2 + (1+L)\,\|\mathcal{S}\,u_h - z_h\|_Z^2.\end{aligned} \tag{2.10}$$

This motivates the definition of our error estimator

$$\operatorname{est}(z_h, u_h, \rho_h, \tilde u_h, \tilde\rho_h)^2 := \frac{1}{\alpha}\,\|\tilde\rho_h - \rho_h\|_U^2 + \frac{1}{\alpha}\,\|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 + \alpha\,\|\tilde u_h - u_h\|_U^2 + \|\mathcal{S}\,u_h - z_h\|_Z^2 \tag{2.11}$$

for the error

$$\operatorname{err}(z_h, u_h, \rho_h)^2 := \|z_h - \bar z\|_Z^2 + \alpha\,\|u_h - \bar u\|_U^2 + \frac{1}{\alpha}\|\rho_h - \bar\rho\|_U^2. \tag{2.12}$$

In the computations leading to (2.10), we have proven the following theorem.

Theorem 2.4. Let $(z_h, u_h, \rho_h) \in Z \times U \times U$ be arbitrary and let $(\tilde u_h, \tilde\rho_h) \in U_{\mathrm{ad}} \times U$ satisfy the variational inequality (2.6). Then, the error estimate

$$\operatorname{err}(z_h, u_h, \rho_h)^2 \le (6 + 6\,L + 4\,L^2)\operatorname{est}(z_h, u_h, \rho_h, \tilde u_h, \tilde\rho_h)^2$$

holds. The constant L was defined in (2.9).

This shows the *reliability* of our error estimator.

In what follows, we will also show that the error estimator given in (2.11) is *efficient*. It is evident, that the efficiency of the error estimator heavily relies on the choice of $(\tilde u_h, \tilde\rho_h)$. Up to now, this choice was arbitrary up to (2.6). We will now fix $(\tilde u_h, \tilde\rho_h)$ in the following assumption.

Assumption 2.5. Let $\varepsilon \ge 0$ be given. The pair $(\tilde u_h, \tilde\rho_h) \in U_{\mathrm{ad}} \times U$ satisfies the variational inequality (2.6) and

$$\frac{1}{\alpha}\|\tilde\rho_h - \rho_h\|_U^2 + \alpha\,\|\tilde u_h - u_h\|_U^2 \le \frac{1+\varepsilon}{\alpha}\|\hat\rho_h - \rho_h\|_U^2 + \alpha\,(1+\varepsilon)\,\|\hat u_h - u_h\|_U^2$$

for all solutions $(\hat u_h, \hat\rho_h) \in U_{\mathrm{ad}} \times U$ of the variational inequality

$$(\alpha\,\hat u_h - \hat\rho_h,\, u - \hat u_h) \ge 0 \qquad \text{for all } u \in U_{\mathrm{ad}}. \tag{2.13}$$

The existence and computability of a pair $(\tilde{u}_h, \tilde{\rho}_h)$ which satisfies this assumption will be shown later, see Lemma 2.7.

By using Assumption 2.5 we infer

$$\frac{1}{\alpha}\|\tilde{\rho}_h - \rho_h\|_U^2 + \alpha\,\|\tilde{u}_h - u_h\|_U^2 \le \frac{1+\varepsilon}{\alpha}\|\bar{\rho} - \rho_h\|_U^2 + \alpha\,(1+\varepsilon)\,\|\bar{u} - u_h\|_U^2,$$

since $(\bar{u}, \bar{\rho})$ solves the variational inequality. Moreover, we find

$$\begin{aligned}
\|\mathcal{S}\,u_h - z_h\|_Z^2 &\le 2\,\|\mathcal{S}\,u_h - \mathcal{S}\,\bar{u}\|_Z^2 + 2\,\|\bar{z} - z_h\|_Z^2 \le 2\,L\,\alpha\,\|u_h - \bar{u}\|_U^2 + 2\,\|\bar{z} - z_h\|_Z^2, \\
\|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 &\le 2\,\|\rho_h - \bar{\rho}\|_U^2 + 2\,\|\mathcal{S}^\star(\bar{z} - z_h)\|_U^2 \le 2\,\|\rho_h - \bar{\rho}\|_U^2 + 2\,L\,\alpha\,\|\bar{z} - z_h\|_Z^2,
\end{aligned}$$

see (2.9). Together with the previous estimate, this leads to

$$\begin{aligned}
\operatorname{est}(z_h, u_h, \rho_h, \tilde{u}_h, \tilde{\rho}_h)^2 &\le \alpha\,(1+\varepsilon)\,\|u_h - \bar{u}\|_U^2 + \frac{1+\varepsilon}{\alpha}\,\|\rho_h - \bar{\rho}\|_U^2 \\
&\qquad + 2\,\|z_h - \bar{z}\|_Z^2 + 2\,L\,\alpha\,\|u_h - \bar{u}\|_U^2 \\
&\qquad + 2\,L\,\|z_h - \bar{z}\|_Z^2 + 2\,\|\rho_h - \bar{\rho}\|_U^2 \\
&= (2 + 2\,L)\,\|z_h - \bar{z}\|_Z^2 + (1 + \varepsilon + 2\,L)\,\alpha\,\|u_h - \bar{u}\|_U^2 + \frac{3+\varepsilon}{\alpha}\,\|\rho_h - \bar{\rho}\|_U^2 \\
&\le (3 + \varepsilon + 2\,L)\,\operatorname{err}(z_h, u_h, \rho_h)^2.
\end{aligned} \tag{2.14}$$

Hence, we have shown the efficiency of our error estimate.

Theorem 2.6. Let $(z_h, u_h, \rho_h) \in Z \times U \times U$ be arbitrary and let $(\tilde{u}_h, \tilde{\rho}_h) \in U_{\text{ad}} \times U$ satisfy Assumption 2.5 for some $\varepsilon \ge 0$. Then, the efficiency estimate

$$\operatorname{est}(z_h, u_h, \rho_h, \tilde{u}_h, \tilde{\rho}_h)^2 \le (3 + \varepsilon + 2\,L)\,\operatorname{err}(z_h, u_h, \rho_h)^2$$

holds.

Let us comment on the ingredients of our error estimator. First of all, one needs to compute

$$\|\mathcal{S}\,u_h - z_h\|_Z \quad \text{and} \quad \|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U.$$

This is, of course, highly dependent on the underlying control-to-observation map. If (z_h, u_h, ρ_h) arise from a discretization of (2.2), one may utilize reliable and efficient error estimators for the corresponding discretization, see Section 3. Hence, our error estimator can build upon existing results.

The other ingredient of our error estimator is the pair $(\tilde{u}_h, \tilde{\rho}_h)$ satisfying Assumption 2.5. First, we remark that it is always possible to satisfy Assumption 2.5 with any $\varepsilon > 0$.

Lemma 2.7. Let $\varepsilon > 0$ and $(u_h, \rho_h) \in U \times U$ be arbitrary. Then, there exist $(\tilde{u}_h, \tilde{\rho}_h)$ satisfying Assumption 2.5.

Proof. We define

$$j = \inf\Big\{\frac{1}{\alpha}\|\hat{\rho}_h - \rho_h\|_U^2 + \alpha\,\|\hat{u}_h - u_h\|_U^2 : (\hat{u}_h, \hat{\rho}_h) \text{ satisfy (2.13)}\Big\}.$$

Note that this is just the distance of (u_h, ρ_h) to the set of all $(\hat{u}_h, \hat{\rho}_h)$ satisfying the variational inequality (2.13). In case $j = 0$, the pair (u_h, ρ_h) already satisfies the variational inequality (since the set of its solutions is closed) and we can choose $(\tilde{u}_h, \tilde{\rho}_h) = (u_h, \rho_h)$. Otherwise, we have $j < j\,(1+\varepsilon)$ and, hence, the existence of $(\tilde{u}_h, \tilde{\rho}_h)$ is clear.

Further, note that the choice $\varepsilon = 0$ may not be possible, since the sequence of $(\hat{u}_h, \hat{\rho}_h)$ defining j might not converge in $U \times U$.

We emphasize that this proof does not rely on the special structure of the control space U and on our admissible set U_{ad}. In fact, Lemma 2.7 remains valid for any Hilbert space U and any convex, closed $U_{\mathrm{ad}} \subset U$.

In what follows, we briefly briefly outline that for the admissible set U_{ad} under consideration one can even choose $\varepsilon = 0$. Indeed, we can utilize that solutions of the variational inequality (2.13) can be characterized by a pointwise projection similar to (2.4). The desired pair $(\tilde{u}_h, \tilde{\rho}_h)$ has to minimize the functional

$$(\hat{u}_h, \hat{\rho}_h) \mapsto \frac{1}{\alpha} \|\hat{\rho}_h - \rho_h\|_U^2 + \alpha \, \|\hat{u}_h - u_h\|_U^2$$

among all $(\hat{u}_h, \hat{\rho}_h) \in U_{\mathrm{ad}} \times U$ satisfying

$$\hat{u}_h(x) = \mathrm{Proj}_{[u_a(x), u_b(x)]} \frac{\hat{\rho}_h(x)}{\alpha} \quad \text{for almost all } x \in \Omega.$$

This is an infinite-dimensional, non-convex optimization problem. However, we can argue point-wise in order to show the existence of a global minimizer. For convenience, we rescale the ρ_h-component by α^{-1} and obtain the following problem.

$$\begin{aligned} \text{Minimize} \quad & |\hat{u}_h(x) - u_h(x)|^2 + |\alpha^{-1} \, \hat{\rho}_h(x) - \alpha^{-1} \, \rho_h(x)|^2 \\ \text{with respect to} \quad & \hat{u}_h(x), \alpha^{-1} \, \hat{\rho}_h(x) \\ \text{such that} \quad & \hat{u}_h(x) = \mathrm{Proj}_{[u_a(x), u_b(x)]} \; \alpha^{-1} \, \hat{\rho}_h(x). \end{aligned} \tag{2.15}$$

That is, we have to project the point $(u_h(x), \alpha^{-1} \, \rho_h(x))$ onto the set

$$\{(\hat{u}_h(x), \alpha^{-1} \, \hat{\rho}_h(x)) \in \mathbb{R}^2 : \hat{u}_h(x) = \mathrm{Proj}_{[u_a(x), u_b(x)]} \; \alpha^{-1} \, \hat{\rho}_h(x)\}.$$

It consists of two rays and a line segment, as depicted in Figure 2.1.

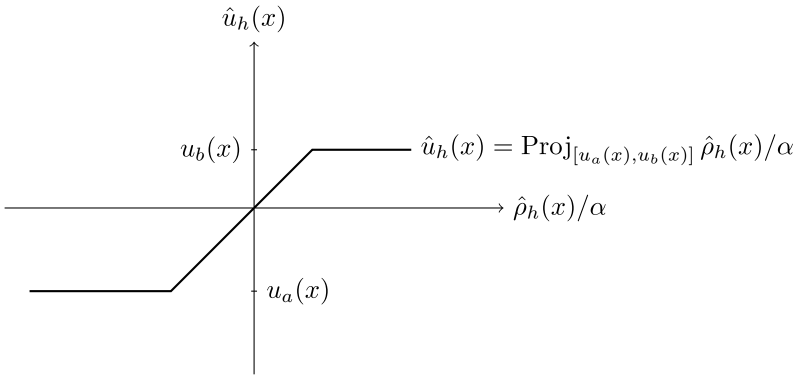


Figure 2.1: The feasible set of the problem (2.15).

Note that the problem (2.15) may have multiple global solutions. If we always select the solution $\tilde{u}_h(x), \tilde{\rho}_h(x))$ with the smallest value of $\alpha^{-1} \, \tilde{\rho}_h(x)$, the functions $\tilde{u}_h$ and $\tilde{\rho}_h$ are measurable. Since $(\mathrm{Proj}_{[u_a(x), u_b(x)]}(0), 0)$ is feasible for (2.15), we obtain the estimate

$$|\tilde{u}_h(x) - u_h(x)|^2 + |\alpha^{-1} \, \tilde{\rho}_h(x) - \alpha^{-1} \, \rho_h(x)|^2 \le |\mathrm{Proj}_{[u_a(x), u_b(x)]}(0) - u_h(x)|^2 + |0 - \alpha^{-1} \, \rho_h(x)|^2,$$

which ensures $(\tilde{u}_h, \tilde{\rho}_h) \in L^2(\Omega_u)^2 = U^2$. By construction, this pair $(\tilde{u}_h, \tilde{\rho}_h)$ satisfies Assumption 2.5 with $\varepsilon = 0$.

The numerical evaluation of

$$\|\tilde{u}_h - u_h\|_{L^2(\Omega_u)}^2 + \frac{1}{\alpha^2} \, \|\tilde{\rho}_h - \rho_h\|_{L^2(\Omega_u)}^2 = \int_{\Omega_u} |\tilde{u}_h(x) - u_h(x)|^2 + |\alpha^{-1} \, \tilde{\rho}_h(x) - \alpha^{-1} \, \rho_h(x)|^2 \, \mathrm{d}x$$

is briefly discussed in Section 3.3.

Remark 2.8. If one is interested in a better stability of the error estimator w.r.t. $\alpha \searrow 0$, one should use a slightly different error estimator. Indeed, Theorem 2.4 and Theorem 2.6 yield

$$\operatorname{err}(z_h, u_h, \rho_h)^2 \le C_1 \, (1 + L^2) \operatorname{est}(z_h, u_h, \rho_h, \tilde{u}_h, \tilde{\rho}_h)^2 \le C_2 \, (1 + L^3) \operatorname{err}(z_h, u_h, \rho_h)^2.$$

Here, C_1 and C_2 are constants, independent of all the data of our problem. Since $L = \|\mathcal{S}\|^2/\alpha$, one obtains the order $\alpha^{-3/2}$ as bound for the quotient between reliability and efficiency of our estimate.

By using the estimator

$$\widetilde{\operatorname{est}}(z_h, u_h, \rho_h, \tilde{u}_h, \tilde{\rho}_h)^2 = \frac{1}{\alpha} \, \|\tilde{\rho}_h - \rho_h\|_U^2 + \frac{1}{\alpha} \, \|\rho_h - \mathcal{S}^\star \, (z_d - z_h)\|_U^2 + (\alpha + 1) \, \|\tilde{u}_h - u_h\|_U^2 + \|\mathcal{S} \, u_h - z_h\|_Z^2$$

one can achieve

$$\operatorname{err}(z_h, u_h, \rho_h)^2 \le C_1 \, (1 + \alpha) \, \widetilde{\operatorname{est}}(z_h, u_h, \rho_h, \tilde{u}_h, \tilde{\rho}_h)^2 \le C_2 \, (1 + \alpha^2) \operatorname{err}(z_h, u_h, \rho_h)^2.$$

Here, the constants C_1, C_2 depend only on $\|\mathcal{S}\|$. Hence, the ratio of reliability and efficiency has improved to α^{-1}.

3 Distributed control of an elliptic equation

In this section, we apply the theory from Section 2 to a specific optimal control problem.

Let $\Omega \subset \mathbb{R}^2$ be a convex polygon. We consider the problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \, \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \, \|u\|_{L^2(\Omega)}^2 \\ \text{such that} \quad & -\Delta y + y = u \quad \text{in } \Omega \\ & \frac{\partial}{\partial n} y = 0 \quad \text{on } \partial\Omega \\ \text{and} \quad & u_a \le u \le u_b. \end{aligned} \tag{3.1}$$

This problem is a special case of (2.1). In fact, we have

- the state space $Y = H^1(\Omega)$, and the observation space $Z = L^2(\Omega)$
- the observation operator $C : H^1(\Omega) \to L^2(\Omega)$ is the canonical embedding, i.e.

$$(C \, y, \, v)_{L^2(\Omega)} = \langle C^\star v, \, y\rangle_{H^1(\Omega)', H^1(\Omega)} = \int_\Omega y \, v \, \mathrm{d}x \qquad \forall y \in H^1(\Omega), v \in L^2(\Omega)$$

- $A : H^1(\Omega) \to H^1(\Omega)'$ is given by $\langle A \, y, \, z\rangle = \int_\Omega \nabla y \cdot \nabla z + y \, z \, \mathrm{d}x$
- the control operator $B = C^\star$ is the canonical embedding $L^2(\Omega) \hookrightarrow H^1(\Omega)'$, see above.

It is clear that all the assumptions of Section 2 are satisfied. Since the operators B and C are just the canonical embeddings, we will not distinguish between $z = C \, y$ and y and between $\rho = B^\star \, p$ and p, respectively.

The (strong formulation of the) adjoint equation (2.3b) for (3.1) reads

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

In what follows, we apply the results of Section 2 in order to obtain a-posteriori estimates for a finite element discretization of (3.1). That is, we want to estimate the distance between the solution $(\bar{y}, \bar{u}, \bar{p})$

of (3.1) to a triple (y_h, u_h, p_h) of finite element functions. By the theory of Section 2, we obtain an estimate for the error

$$\|\bar{y} - y_h\|_{L^2(\Omega)} + \alpha \, \|\bar{u} - u_h\|_{L^2(\Omega)} + \frac{1}{\alpha} \, \|\bar{p} - p_h\|_{L^2(\Omega)}, \tag{3.2}$$

see (2.12), and we have to provide a-posteriori estimates for the error contributions in the PDEs

$$\|p_h - A^{-\star} \, (y_d - y_h)\|_{L^2(\Omega)}, \quad \|y_h - A^{-1} \, u_h\|_{L^2(\Omega)}, \tag{3.3}$$

and in the variational inequality

$$\|\tilde{p}_h - p_h\|_{L^2(\Omega)}, \quad \|\tilde{u}_h - u_h\|_{L^2(\Omega)} \tag{3.4}$$

see (2.11).

The finite element discretization is briefly introduced in Section 3.1. In Section 3.2 we recall some standard arguments leading to an $L^2(\Omega)$-error estimator for the error in the state and adjoint equation. The numerical evaluation of the error in the variational inequality is discussed in Section 3.3. Finally, we present some numerical results in Section 3.5.

3.1 Finite element discretization

We briefly introduce the required assumptions on the finite element discretization. We consider a triangulation $\mathcal{T}$ of Ω, see [Brenner and Scott, 2002, Definition 3.3.11]. In particular, we have $\bar{\Omega} = \bigcup_{K \in \mathcal{T}} K$ and the triangulation has no hanging nodes. We define the cell size

$$h_K := \operatorname{diam}(K) \qquad \text{for all } K \in \mathcal{T}$$

and the radius of the largest circle contained in K

$$\rho_K := \sup\{r > 0 : B_r(x) \subset K \text{ for some } x \in K\}.$$

Using the triangulation $\mathcal{T}$, we define the Lagrange finite elements of order k, $k \geq 1$,

$$\mathcal{P}^k := \{v : C(\bar{\Omega}) : v_{|K} \in P^k(K) \text{ for all } K \in \mathcal{T}\},$$

see [Brenner and Scott, 2002, Section 3.2]. Here, $P^k(K)$ is the space of polynomials of degree at most k on the triangle $K \in \mathcal{T}$. This space $\mathcal{P}^k$ is be used to compute approximations $(y_h, u_h, p_h) \subset (\mathcal{P}^k)^3$ of $(\bar{y}, \bar{u}, \bar{p})$, i.e. we solve the discretized optimality system

$$\langle A \, y_h, \, v_h \rangle = \langle B \, u_h, \, v_h \rangle \qquad \forall v_h \in \mathcal{P}^k \tag{3.5a}$$

$$\langle A^\star \, p_h, \, v_h \rangle = \langle C^\star (y_d - y_h), \, v_h \rangle \qquad \forall v_h \in \mathcal{P}^k \tag{3.5b}$$

$$u_h(x_i) = \operatorname{Proj}_{[u_a, u_b]} \frac{p_h(x_i)}{\alpha} \qquad \text{for all Lagrange nodes } x_i \text{ of } \mathcal{T}, \tag{3.5c}$$

compare (2.3) and (2.4).

Note that the variational inequality (2.3c) is discretized by applying the projection (2.4) at each Lagrange point only.

We denote by $\mathcal{K}$ and $\mathcal{M}$ the stiffness and mass matrix associated with (3.5). That is, $\langle \mathcal{K} \, v_h, \, w_h \rangle = \langle A \, v_h, \, w_h \rangle$ and $\langle \mathcal{M} \, v_h, \, w_h \rangle = (v_h, \, w_h)_{L^2(\Omega)}$ for all $v_h, w_h \in \mathcal{P}^k$. The discrete solutions (y_h, u_h, p_h) are obtained by solving the system

$$\mathcal{K} \, y_h = \mathcal{M} \, u_h, \tag{3.6a}$$

$$\mathcal{K} \, p_h = \mathcal{M} \, (y_d - y_h), \tag{3.6b}$$

$$u_h(x_i) = \operatorname{Proj}_{[u_a, u_b]} \frac{p_h(x_i)}{\alpha} \qquad \text{for all Lagrange nodes } x_i \text{ of } \mathcal{T}. \tag{3.6c}$$

Here, the desired state y_d is replaced by an interpolation.

3.2 Error of the FE discretization

As an ingredient for our error estimate (2.11), we need an a-posteriori error estimate for the discretized solution of a PDE in the $L^2(\Omega)$ norm, see (3.3). Such an estimate is well known and can be found in, e.g., [Verfürth, 1996, Proposition 3.8]. We will recall some standard arguments leading to such an a-posteriori estimate since we have to use similar arguments to derive an error estimate in $L^2(\partial\Omega)$ in Section 4.1.

We consider the solution $w \in H^1(\Omega)$ of the PDE

$$A\,w = F \tag{3.7}$$

and its discrete solution $w_h \in \mathcal{P}^k$ with

$$\langle A\,w_h,\, v_h\rangle = \langle F,\, v_h\rangle \qquad \text{for all } v_h \in \mathcal{P}^k. \tag{3.8}$$

Here, $F \in H^1(\Omega)'$ is given by

$$F(v) = \int_\Omega f\,v\,\mathrm{d}x + \int_{\partial\Omega} g\,v\,\mathrm{d}s,$$

with $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega)$. When applying the results to the state y_h, we will set $f = u_h$ and, similar, $f = y_d - y_h$ for the adjoint state p_h. In both cases, we set $g = 0$.

Following [Brenner and Scott, 2002, Section 9.2], we introduce $\varphi \in H^1(\Omega)$ solving the dual equation

$$\langle A\,\varphi,\, v\rangle = \int_\Omega (w - w_h)\,v\,\mathrm{d}x \qquad \text{for all } v \in H^1(\Omega). \tag{3.9}$$

Using standard arguments (integration by parts on all cells $K \in \mathcal{T}$), we arrive at

$$\begin{aligned}\|w - w_h\|^2_{L^2(\Omega)} \le \sum_{K\in\mathcal{T}} &\|f + \Delta w_h - w_h\|_{L^2(K)}\,\|\varphi - \varphi_h\|_{L^2(K)} \\ &+ \sum_{E\in\mathcal{E}(\mathcal{T})} \big\|[\![\nabla w_h]\!]_n\big\|_{L^2(E)}\,\|\varphi - \varphi_h\|_{L^2(E)}\end{aligned} \tag{3.10}$$

for arbitrary $\varphi_h \in \mathcal{P}^k$, see [Brenner and Scott, 2002, Section 9.2] for similar arguments in the case of homogeneous Dirichlet boundary conditions. Here, $\mathcal{E}(\mathcal{T})$ are the edges of the triangulation $\mathcal{T}$. The expression $[\![\nabla w_h]\!]_n$ on an edge E denotes the jump of the normal derivative in normal direction and is defined as follows

$$[\![\nabla w_h]\!]_n(x) := \begin{cases} \lim_{\varepsilon\searrow 0} n^\top\big\{\nabla w_h(x + \varepsilon\,n) - \nabla w_h(x - \varepsilon\,n)\big\} & \text{in case } E \not\subset \partial\Omega, \\ \frac{\partial}{\partial n} w_h(x) - g(x) = n^\top\,\nabla w_h(x) - g(x) & \text{in case } E \subset \partial\Omega, \end{cases} \tag{3.11}$$

where the vector n is the (outer) unit normal vector of E at point $x \in E$. Note that for edges E on the boundary, the jump term $[\![\nabla w_h]\!]_n$ is the residuum for the Neumann data.

Since the solution φ of the dual problem (3.9) satisfies

$$\|\varphi\|_{H^2(\Omega)} \le c\,\|w - w_h\|_{L^2(\Omega)},$$

see [Grisvard, 1985, Theorem 3.2.1.2], we can set $\varphi_h = \mathcal{I}\,\varphi$, where $\mathcal{I} : C(\bar\Omega) \to \mathcal{P}^k$ is the nodal interpolation. For the nodal interpolation, we have the following error estimates.

Lemma 3.1. Let $m \in (1, k+1]$ be given. Assume that $h_K/\rho_K \le \gamma$ holds for all $K \in \mathcal{T}$. Then, there is a constant $c_{k,m,\gamma}$, such that

$$\begin{aligned}\|v - \mathcal{I}\,v\|_{L^2(K)} &\le c_{k,m,\gamma}\,h_K^m\,|v|_{H^m(K)}, \\ \|v - \mathcal{I}\,v\|_{L^2(E)} &\le c_{k,m,\gamma}\,h_K^{m-1/2}\,|v|_{H^m(K)}\end{aligned}$$

for all triangles $K \in \mathcal{T}$, all edges E of K and all $v \in H^m(K)$.

Proof. The estimate on the cell is standard in case of integer m. The case of non-integer m can be found in [Feistauer, 1989, Theorem 2.19]. The proof for the estimate on the edge is similar and straightforward.

We emphasize that the constant $c_{k,m,\gamma}$ does not depend directly on the triangulation $\mathcal{T}$, but only on its chunkiness

$$\gamma_{\mathcal{T}} := \max\{h_K/\rho_K : K \in \mathcal{T}\}.$$

Using these error estimates for the interpolation with $m = 2$ in (3.10), we obtain the following theorem by standard arguments.

Theorem 3.2. Let us denote by $w \in H^1(\Omega)$ and $w_h \in \mathcal{P}^k$ the solutions of (3.7) and (3.8), respectively. Then,

$$\|w - w_h\|^2_{L^2(\Omega)} \le c \sum_{K\in\mathcal{T}} \eta_K^2, \tag{3.12}$$

where the local error indicator η_K is defined by

$$\eta_K^2 := h_K^4 \, \|f + \Delta w_h - w_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}(K)} h_K^3 \, \big\|[\![\nabla w_h]\!]_n\big\|^2_{L^2(E)}. \tag{3.13}$$

Here, $\mathcal{E}(K)$ denotes the set of all edges of the triangle K.

We emphasize that the constant c in (3.12) does not depend on the triangulation $\mathcal{T}$, but only on its chunkiness $\gamma_{\mathcal{T}}$, compare Lemma 3.1.

For problem (3.1) this error estimate is used for the state and the adjoint equation. To this end, let $(y_h, u_h, p_h) \in (\mathcal{P}^k)^3$ be given. In order to apply the error estimate (3.12), we assume that y_h solves the discretized state equation (3.5a) and p_h the discretized adjoint equation (3.5b). For each cell, we define according to Theorem 3.2 the local contributions

$$\eta^2_{K,\text{state}} := h_K^4 \, \|u_h + \Delta y_h - y_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}(K)} h_K^3 \, \big\|[\![\nabla y_h]\!]_n\big\|^2_{L^2(E)}, \tag{3.14a}$$

$$\eta^2_{K,\text{adjoint}} := h_K^4 \, \|y_d - y_h + \Delta p_h - p_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}(K)} h_K^3 \, \big\|[\![\nabla p_h]\!]_n\big\|^2_{L^2(E)}. \tag{3.14b}$$

Now, Theorem 3.2 implies that

$$\|y_h - A^{-1} u_h\|^2_{L^2(\Omega)} \le c \sum_{K\in\mathcal{T}} \eta^2_{K,\text{state}} \qquad \text{and} \tag{3.15a}$$

$$\|p_h - A^{-\star}(y_d - y_h)\|^2_{L^2(\Omega)} \le c \sum_{K\in\mathcal{T}} \eta^2_{K,\text{adjoint}} \tag{3.15b}$$

hold for some constant $c > 0$.

We remark that [Verfürth, 1998, Proposition 4.1] shows the (local) efficiency (up to higher order terms) of the error estimator for a similar problem.

3.3 Error in the variational inequality

It remains to construct $(\tilde{u}_h, \tilde{p}_h)$ satisfying Assumption 2.5. As discussed in Section 2.2, this can be done by solving (2.15) for each point $x \in \Omega$. Finally, we have to integrate

$$\eta^2_{K,\text{VI}} := \int_K \frac{1}{\alpha^2} \, |p_h - \tilde{p}_h|^2 + |u_h - \tilde{u}_h|^2 \, \mathrm{d}x \tag{3.16}$$

for each cell $K \in \mathcal{T}$. Note that Assumption 2.5 is satisfied by this construction.

Note that in general, the functions $\tilde{u}_h$ and $\tilde{p}_h$ are not piecewise polynomials due to the projection. Thus, the evaluation of the above integral requires special attention, Since the integrand may not have a high regularity, we use a quadrature rule of moderate degree and apply it on (red) subdivisions of K. Thus, (2.15) has to be solved only in the quadrature points. For the numerical experiments presented in Section 3.5 and Section 4.3, we used 2 subdivisions and a quadrature rule of degree 6.

3.4 Error estimator for the optimal control problem

Using the results from Section 3.2 and Section 3.3, we obtain an error estimator for the discretization (3.5) of problem (3.1).

Theorem 3.3. Let $(y_h, u_h, p_h) \in (\mathcal{P}^k)^3$ be given, such that the discretized state and adjoint equation (3.5a), (3.5b) are satisfied. We define the local error contribution

$$\eta_K^2 := \eta_{K,\text{state}}^2 + \frac{1}{\alpha}\,\eta_{K,\text{adjoint}}^2 + \alpha\,\eta_{K,\text{VI}}^2,$$

where $\eta_{K,\cdot}$ are defined in (3.14) and (3.16). Then, we have the error estimate

$$\|\bar{y} - y_h\|_{L^2(\Omega)}^2 + \alpha\,\|\bar{u} - u_h\|_{L^2(\Omega)}^2 + \frac{1}{\alpha}\,\|\bar{p} - p_h\|_{L^2(\Omega)}^2 \le c \sum_{K\in\mathcal{T}} \eta_K^2.$$

The proof follows from Theorem 2.4 and (3.15).

Again, the constant c does not depend directly on the triangulation $\mathcal{T}$, but only on its chunkiness $\gamma_{\mathcal{T}}$. Up to higher order terms, this error estimate is also efficient, see Theorem 2.6 and [Verfürth, 1996, Proposition 3.8], [Verfürth, 1998, Proposition 4.1].

3.5 Numerical results

We report some numerical results on the solution of (3.1). The data of the problem is given by

$$\begin{aligned} \Omega &= (0,1)^2, \quad & y_d(x,y) &= \exp(x)\,\sin(y), \\ \alpha &= 0.05, & u_a &= -3, \quad u_b = 1. \end{aligned} \tag{3.17}$$

On each mesh, we use $\mathcal{P}^k$ elements with $k \in \{1,2\}$ and solve (3.6). The solution on a fine mesh is depicted in Figure 3.1.

In the following, we will compare the results obtained by the error estimator from Theorem 3.3 with the results obtained by applying an energy-based error estimator given by

$$\eta_{K,\text{state,energy}}^2 := h_K^2\,\|u_h + \Delta y_h - y_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}(K)} h_K^1\,\big\|[\![\nabla y_h]\!]_n\big\|_{L^2(E)}^2, \tag{3.18a}$$

$$\eta_{K,\text{adjoint,energy}}^2 := h_K^2\,\|y_d - y_h + \Delta p_h - p_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}(K)} h_K^1\,\big\|[\![\nabla p_h]\!]_n\big\|_{L^2(E)}^2, \tag{3.18b}$$

$$\eta_{K,\text{energy}}^2 := \eta_{K,\text{state,energy}}^2 + \frac{1}{\alpha}\,\eta_{K,\text{adjoint,energy}}^2 + \alpha\,\eta_{K,\text{VI}}^2. \tag{3.18c}$$

This is the error estimator suggested in Kohls et al. [2014] with a slight modification in the term $\eta_{K,\text{VI}}$. The precise estimator of Kohls et al. [2014] would be obtained by setting

$$\tilde{p}_h(x) = p_h(x), \qquad \tilde{u}_h(x) = \operatorname{Proj}_{[u_a(x),u_b(x)]}\frac{p_h(x)}{\alpha}$$

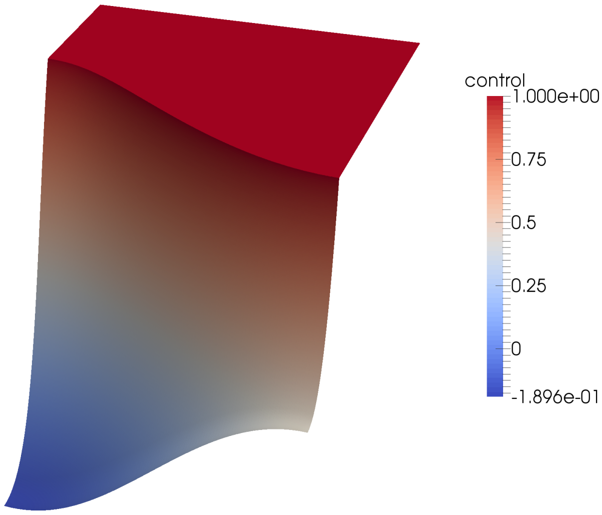


Figure 3.1: Solution of (3.1) with setting (3.17) computed on a fine mesh

for $x \in \Omega$ in (3.16).

For the marking of the cells, we use the maximum strategy, i.e., we mark all cells K satisfying

$$\eta_K^2 \geq \kappa \max_{K' \in \mathcal{T}} \eta_{K'}^2$$

for some $\kappa \in [0, 1]$. In the numerical experiments, we used $\kappa = 1/2$.

Some of the obtained meshes are shown in Figure 3.2. All four meshes result in approximately 10,000 degrees of freedom (per variable y, u and p). It can be seen that the error estimator from Theorem 3.3 is more focused on refining the interface, i.e., the boundary of the active set, whereas using the energy-based error estimator results in a finer mesh in the whole domain. By using the energy-based estimator with $\mathcal{P}^1$ elements, one does not observe a refinement at the interface. In the case of $\mathcal{P}^2$ elements, the interface is slightly refined for higher numbers of degrees of freedom (starting at roughly 50,000 degrees of freedom).

Finally, we show the total error as defined in (2.12) for both strategies in Figure 3.3. Since the exact solution is not known, the error is computed w.r.t. a solution on a fine grid with approximately 1,600,000 degrees of freedom (per variable y, u, p). From that plot it is clear that a better rate is obtained by employing the error estimator from Theorem 3.3. We emphasize that the same behaviour is observed if we only plot the error $\|u_h - \bar{u}\|_{L^2(\Omega)}$. Moreover, in case of our new estimator, the errors in all three components

$$\|\bar{y} - y_h\|_{L^2(\Omega)}, \qquad \|\bar{u} - u_h\|_{L^2(\Omega)}, \qquad \|\bar{p} - p_h\|_{L^2(\Omega)}$$

converge with the same order. In case of the energy estimator (3.18), the $L^2(\Omega)$-error in the state and adjoint converges faster.

4 Distributed control of an elliptic equation with boundary observation

In order to demonstrate the flexibility of our error estimator, we consider a variant of (3.1), in which we replace the distributed observation by an observation on the boundary.

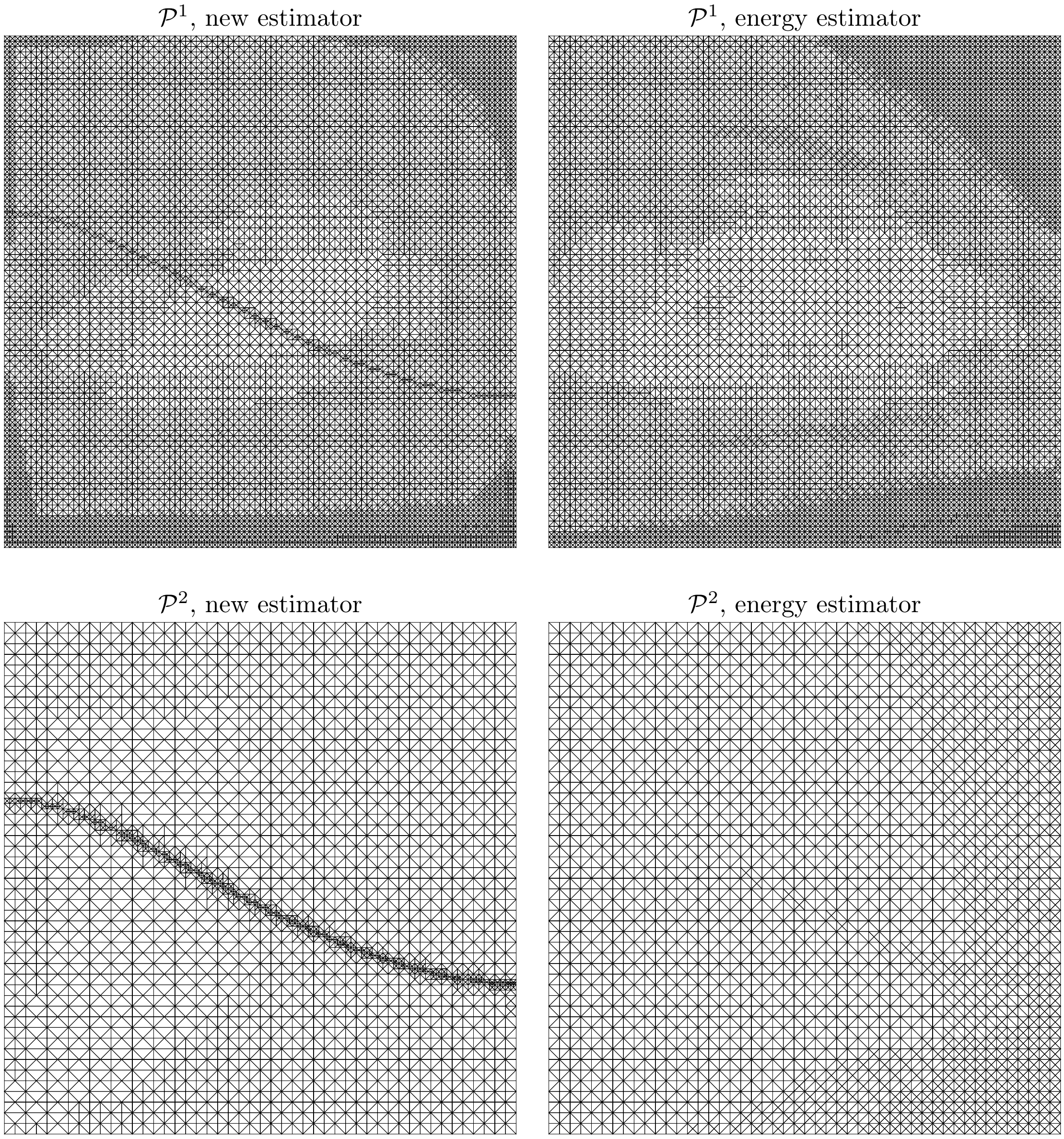


Figure 3.2: Meshes obtained by an adaptive refinement for the solution of (3.1) with $\mathcal{P}^1$ (top row) and $\mathcal{P}^2$ elements (bottom row). For the left column, we used the error estimator from Theorem 3.3 and an energy-based error estimator (3.18) for the right column.

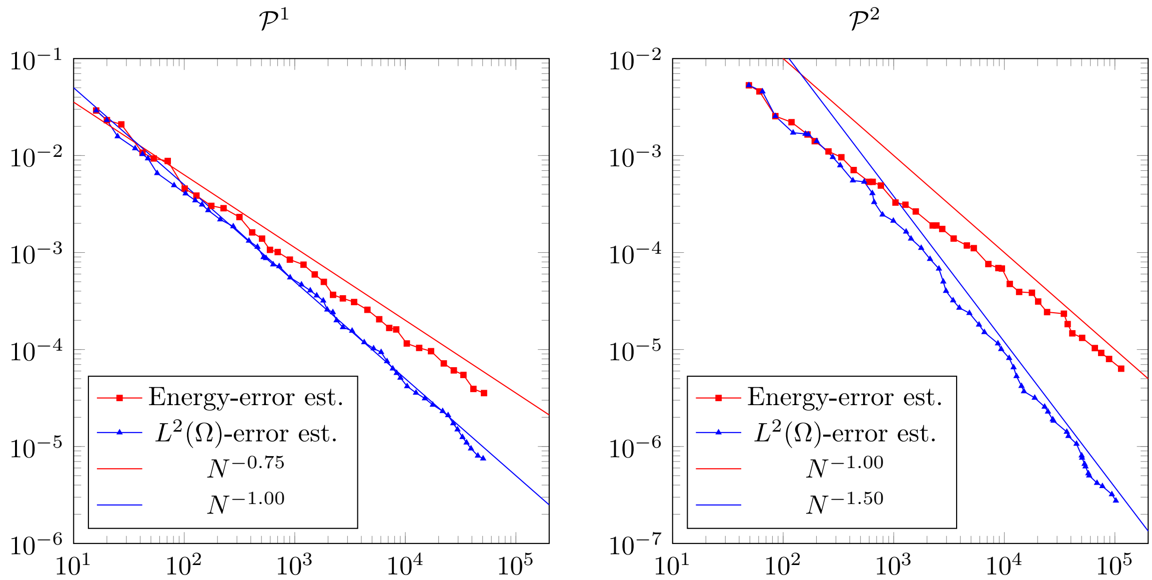


Figure 3.3: Total error (3.2) for the discretization of (3.1) with $\mathcal{P}^1$ (left) and $\mathcal{P}^2$ (right) elements for adaptive refinement based on the error estimator from Theorem 3.3 (red squares) and the energy-based error estimator (3.18) (blue triangles), respectively. The error is plotted versus the degrees of freedom N (per variable y_h, u_h, p_h). Slopes N^{-p} are provided for comparison.

Let $\Omega \subset \mathbb{R}^2$ be a convex polygon. We consider the problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|y - y_d\|^2_{L^2(\partial\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ \text{such that} \quad & -\Delta y + y = u \quad \text{in } \Omega \\ & \frac{\partial}{\partial n} y = 0 \quad \text{on } \partial\Omega \\ \text{and} \quad & u_a \le u \le u_b. \end{aligned} \tag{4.1}$$

Again, this problem is a special case of (2.1). The difference to (3.1) is that

- the observation space is $Z = L^2(\partial\Omega)$ and
- the observation operator $C : H^1(\Omega) \to L^2(\partial\Omega)$ is the trace operator, i.e.

$$(C\,y,\, v)_{L^2(\partial\Omega)} = \langle C^\star v,\, y\rangle_{H^1(\Omega)',H^1(\Omega)} = \int_\Gamma y\, v \,\mathrm{d}x \qquad \forall y \in H^1(\Omega), v \in L^2(\partial\Omega).$$

It is clear that all the assumptions of Section 2 are satisfied.

The (strong formulation of the) adjoint equation (2.3b) for (4.1) reads

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

Note that, in difference to Section 3, the difference $y_d - y$ now appears as boundary data in the adjoint equation.

Similar to Section 3, we apply the results of Section 2 in order to obtain a-posteriori estimates for a finite element discretization of (4.1). That is, we want to estimate the distance between the solution $(\bar y, \bar u, \bar p)$ of (4.1) to a triple (y_h, u_h, p_h) of finite element functions. By the theory of Section 2, we obtain an estimate for error

$$\|\bar y - y_h\|_{L^2(\partial\Omega)} + \alpha\,\|\bar u - u_h\|_{L^2(\Omega)} + \frac{1}{\alpha}\|\bar p - p_h\|_{L^2(\Omega)}. \tag{4.2}$$

The main difference to Section 3 is, that the error in the state is measured in the L^2-norm on the boundary $\partial\Omega$ and not in the L^2-norm in the domain Ω. Thus, we have to construct an a-posteriori error estimator for estimating the difference between y_h and the (continuous) solution of the state equation with right-hand side u_h in the $L^2(\partial\Omega)$-norm. To our knowledge, such an estimator is not available in the literature. This will be addressed in the next section. The error estimator for the control problem (4.1) is described in Section 4.2 and numerical results are presented in Section 4.3.

4.1 A-posteriori error estimator on the boundary

As already mentioned, we have to construct an a-posteriori error estimator for the error

$$\|A^{-1} u_h - y_h\|_{L^2(\partial\Omega)}.$$

This is due to the fact that the observation operator $C : H_0^1(\Omega) \to L^2(\partial\Omega)$ is the trace operator and, hence, the error estimate (2.11) contains

$$\|\mathcal{S}\, u_h - z_h\|_{L^2(\partial\Omega)} = \|A^{-1} u_h - y_h\|_{L^2(\partial\Omega)}.$$

The function $A^{-1} u_h$ is the solution $\tilde y_h \in H^1(\Omega)$ of the PDE

$$\langle A\, \tilde y_h,\, v\rangle = \int_\Omega u_h\, v \,\mathrm{d}x \quad \text{for all } v \in H^1(\Omega). \tag{4.3}$$

Moreover, we require that $y_h \in \mathcal{P}^k$ satisfies the discretized state equation, that is

$$\langle A\, y_h,\, v_h\rangle = \int_\Omega u_h\, v_h \,\mathrm{d}x \quad \text{for all } v_h \in \mathcal{P}^k. \tag{4.4}$$

Similar as in Section 3.2, we introduce the dual solution $\varphi \in H^1(\Omega)$ solving

$$\langle A\, \varphi,\, v\rangle = \int_{\partial\Omega} (\tilde y_h - y_h)\, v \,\mathrm{d}x \qquad \text{for all } v \in H^1(\Omega). \tag{4.5}$$

and obtain

$$\begin{aligned} \|\tilde y_h - y_h\|^2_{L^2(\partial\Omega)} \le \sum_{K\in\mathcal{T}} \|u_h + \Delta y_h - y_h\|_{L^2(K)}\, \|\varphi - \varphi_h\|_{L^2(K)} \\ + \sum_{E\in\mathcal{E}(\mathcal{T})} \big\|[\![\nabla y_h]\!]_n\big\|_{L^2(E)}\, \|\varphi - \varphi_h\|_{L^2(E)} \end{aligned} \tag{4.6}$$

for arbitrary $\varphi_h \in \mathcal{P}^k$.

In difference to the situation of Section 3.2, we cannot bound the $H^2(\Omega)$-norm of the dual solution φ by the error $\|\tilde y_h - y_h\|_{L^2(\partial\Omega)}$, since this error is the Neumann datum of φ, see (4.5).

However, from Jerison and Kenig [1981] we now that

$$\|\varphi\|_{H^{3/2}(\Omega)} \le c\, \|\tilde y_h - y_h\|_{L^2(\partial\Omega)}. \tag{4.7}$$

The same result can be obtained by using an interpolation between $s > 1/2$ and $s < 1/2$ in [Dauge, 1988, Corollary 23.5]. Moreover, using [Khoromskij and Melenk, 2003, Theorem A.1], we find

$$\|r^{1/2}\, \nabla^2 \varphi\|_{L^2(\Omega)} \le c\, \|\tilde y_h - y_h\|_{L^2(\partial\Omega)} \quad \text{and} \quad \|r^{3/2}\, \nabla^3 \varphi\|_{L^2(\Omega)} \le c\, \|\tilde y_h - y_h\|_{L^2(\partial\Omega)} \tag{4.8}$$

where $r(x) = \operatorname{dist}(x, \partial\Omega)$ is the distance to the boundary. Estimate (4.8) enables us to control higher derivatives of φ in the interior of our domain. Similar estimates are available for arbitrary high derivatives.

For simplicity of the demonstration, we now consider the case $k = 2$, i.e. the case of piecewise quadratic finite elements. Let $K \in \mathcal{T}$ be a triangle and define the distance to the boundary

$$d_K := \operatorname{dist}(K, \partial\Omega) := \min_{x \in K} \operatorname{dist}(x, \partial\Omega).$$

In case $d_K = 0$, the triangle lies at the boundary and we cannot make use of the additional regularity from (4.8). From Lemma 3.1, we obtain the estimates

$$\|\varphi - \mathcal{I}\,\varphi\|_{L^2(K)} \le c\, h_K^{3/2}\, \|\varphi\|_{H^{3/2}(K)} \quad \text{and} \quad \|\varphi - \mathcal{I}\,\varphi\|_{L^2(E)} \le c\, h_K\, \|\varphi\|_{H^{3/2}(K)},$$

where E is an edge of K.

If, however, $d_K > 0$, we can use $r \ge d_K$ on K to obtain

$$\|\nabla^2\varphi\|_{L^2(K)} \le d_K^{-1/2}\, \|r^{1/2}\, \nabla^2\varphi\|_{L^2(K)} \quad \text{and} \quad \|\nabla^3\varphi\|_{L^2(K)} \le d_K^{-3/2}\, \|r^{3/2}\, \nabla^2\varphi\|_{L^2(K)}.$$

This yields $\varphi \in H^3(K)$ and we can apply Lemma 3.1 with $m = 3/2$, $m = 2$ and $m = 3$. Hence,

$$\begin{aligned}
&\|\varphi - \mathcal{I}\,\varphi\|_{L^2(K)} + h_K^{1/2}\, \|\varphi - \mathcal{I}\,\varphi\|_{L^2(E)} \\
&\qquad \le c\, \min\{h_K^{3/2}\, \|\varphi\|_{H^{3/2}(K)}, d_K^{-1/2}\, h_K^2\, \|r^{1/2}\, \nabla^2\varphi\|_{L^2(K)}, d_K^{-3/2}\, h_K^3\, \|r^{3/2}\, \nabla^3\varphi\|_{L^2(K)}\} \\
&\qquad \le c\, \min\{h_K^{3/2}, d_K^{-1/2}\, h_K^2, d_K^{-3/2}\, h_K^3\} \\
&\qquad\qquad\qquad \cdot \max\{\|\varphi\|_{H^{3/2}(K)}, \|r^{1/2}\, \nabla^2\varphi\|_{L^2(K)}, \|r^{3/2}\, \nabla^3\varphi\|_{L^2(K)}\}.
\end{aligned}$$

Using that estimate in (4.6) and utilizing (4.7), (4.8), we finally obtain the following theorem.

Theorem 4.1. We assume $k = 2$. Let $\tilde{y}_h \in H^1(\Omega)$ and $y_h \in \mathcal{P}^2$ satisfy (4.3) and (4.4), respectively. Then,

$$\|\tilde{y}_h - y_h\|_{L^2(\partial\Omega)}^2 \le c \sum_{K \in \mathcal{T}} \eta_{K,\text{state}}^2, \tag{4.9}$$

where the local error contribution $\eta_{K,\text{state}}$ is given by

$$\begin{aligned}
\eta_{K,\text{state}}^2 = \min\{h_K^3,\ d_K^{-1}\, h_K^4,\ d_K^{-3}\, h_K^6\}& \\
\cdot \Big\{\big\|u_h + \Delta y_h - y_h\big\|_{L^2(K)}^2 + \sum_{E \in \mathcal{E}(K)} h_K^{-1}\, \big\|[\![\nabla y_h]\!]_n\big\|_{L^2(E)}^2\Big\}&.
\end{aligned} \tag{4.10}$$

Here, we used the convention that

$$\min\{h_K^3,\ d_K^{-1}\, h_K^4,\ d_K^{-3}\, h_K^6\} = h_K^3$$

in case $d_K = 0$.

As in Theorem 3.2, the constant c in (4.9) does not depend on the triangulation $\mathcal{T}$, but only on its chunkiness $\gamma_{\mathcal{T}}$, compare Lemma 3.1.

We briefly mention, how this a-posteriori estimate can be generalized to other values of the polynomial degree k. In case $k = 1$, we cannot use $m = 3$ in Lemma 3.1. Hence, the expression involving min in the local error contribution (4.10) has to be replaced by

$$\min\{h_K^3, d_K^{-1}\, h_K^4\}.$$

If we use a higher value of k, one has, similar to (4.8), also estimates for higher derivatives of φ. Thus, one can apply Lemma 3.1 with $m = 3/2$ and $m = 2, \ldots, k + 1$. Finally, one can replace the min-expression in (4.10) by

$$\min\{h_K^3, d_K^{-1}\, h_K^4, d_K^{-3}\, h_K^6, \ldots, d_K^{1-2\,k}\, h_K^{2\,k+2}\}.$$

We note that this expression can be simplified, since the middle terms are a weighted geometric mean of the first and last term:

$$d_K^{1-2\,m}\,h_K^{2\,m-2} = \Big(\big(h_K^3\big)^{\frac{2\,(m-k)}{1-2\,m}} \cdot \big(d_K^{1-2\,k}\,h_K^{2\,k+2}\big)^1 \Big)^{\frac{1-2\,m}{1-2\,k}}.$$

Hence, we have

$$\min\{h_K^3, d_K^{-1}\,h_K^4, d_K^{-3}\,h_K^6, \ldots, d_K^{1-2\,k}\,h_K^{2\,k+2}\} = \min\{h_K^3, d_K^{1-2\,k}\,h_K^{2\,k+2}\}.$$

4.2 Error estimator for the optimal control problem

In this section, we state the a-posteriori error estimator for the control problem (4.1).

Theorem 4.2. Let $(y_h, u_h, p_h) \in (\mathcal{P}^k)^3$ be given, such that the discretized state and adjoint equation are satisfied. We define the local error contribution

$$\eta_K^2 := \eta_{K,\text{state}}^2 + \frac{1}{\alpha}\,\eta_{K,\text{adjoint}}^2 + \alpha\,\eta_{K,\text{VI}}^2,$$

where $\eta_{K,\text{VI}}$ is given in (3.16) and $\eta_{K,\text{state}}$ is defined in (4.10). The contribution from the adjoint equation is

$$\eta_{K,\text{adjoint}}^2 := h_K^4\,\|\Delta p_h - p_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}(K)} h_K^3\,\big\|[\![\nabla p_h]\!]_n\big\|_{L^2(E)}^2, \tag{4.11}$$

where the adjoint Neumann data $g = y_d - y_h$ enters in the definition of the jump term (3.11). Then, we have the error estimate

$$\|y_h - \bar y\|_{L^2(\partial\Omega)}^2 + \alpha\,\|u_h - \bar u\|_{L^2(\Omega)}^2 + \frac{1}{\alpha}\,\|p_h - \bar p\|_{L^2(\Omega)}^2 \le c \sum_{K\in\mathcal{T}} \eta_K^2.$$

Proof. The contribution for the adjoint state $\eta_{K,\text{adjoint}}$ follows from Theorem 3.2. The overall error estimate follows from Theorem 2.4 and Theorem 4.1.

The efficiency of the error estimator is an open problem. It would follow from the efficiency of the $L^2(\partial\Omega)$-error estimator (4.10), which is, however, unknown. This should be addressed in future work.

4.3 Numerical results

We report some numerical results on the solution of (4.1). The data of the problem is as in (3.17), except for $u_b = 7.5$. The discrete solution is obtained as in (3.6), but the mass matrix in (3.6b) has to be replaced by a boundary mass matrix. The solution on a fine mesh is depicted in Figure 4.1.

In the following we will compare the results obtained by the error estimator from Theorem 4.2 with the results obtained by applying an energy-based error estimator. The energy-based error estimator is obtained as described in Section 3.5, i.e.,

$$\eta_{K,\text{state,energy}}^2 := h_K^2\,\|u_h + \Delta y_h - y_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}(K)} h_K^1\,\big\|[\![\nabla y_h]\!]_n\big\|_{L^2(E)}^2, \tag{4.12a}$$

$$\eta_{K,\text{adjoint,energy}}^2 := h_K^2\,\|\Delta p_h - p_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}(K)} h_K^1\,\big\|[\![\nabla p_h]\!]_n\big\|_{L^2(E)}^2, \tag{4.12b}$$

$$\eta_{K,\text{energy}}^2 := \eta_{K,\text{state,energy}}^2 + \frac{1}{\alpha}\,\eta_{K,\text{adjoint,energy}}^2 + \alpha\,\eta_{K,\text{VI}}^2. \tag{4.12c}$$

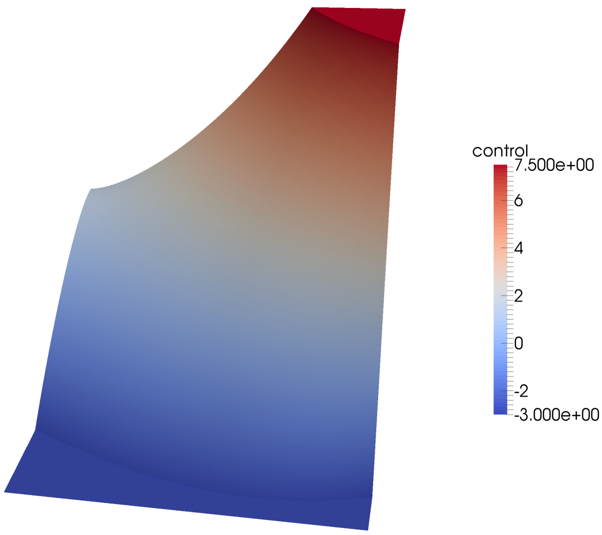


Figure 4.1: Solution of (4.1) computed on a fine mesh

As in (4.11), (3.18b) depends on the adjoint Neumann data $g = y_d - y_h$ via the definition of the jump term (3.11).

Some of the obtained meshes are shown in Figure 4.2. As in the case of distributed observation, all four meshes result in approximately 10,000 degrees of freedom (per variable). Again, the tailored error estimator from Theorem 4.2 is more focused on refining the interface, whereas using the energy-based error estimator results in a fine mesh in the vertices of Ω. The energy-based estimator does not result in a refinement of the interface in the case of $\mathcal{P}^1$ elements. A slight refinement of the interface in case of $\mathcal{P}^2$ elements is observed for higher numbers of degrees of freedom (starting at roughly 50,000 degrees of freedom).

Finally, the total error (2.12) for both strategies is shown in Figure 4.3. The error is computed w.r.t. a solution on a fine grid with approximately 1,600,000 degrees of freedom. Again, a better rate is obtained by employing the error estimator from Theorem 4.2 and the same behaviour is observed if we only plot the error $\|u_h - \bar{u}\|_{L^2(\Omega)}$. For our new estimator, the three errors

$$\|\bar{y} - y_h\|_{L^2(\partial\Omega)}, \qquad \|\bar{u} - u_h\|_{L^2(\Omega)}, \qquad \|\bar{p} - p_h\|_{L^2(\Omega)}$$

converge with the same order, whereas for the energy estimator the state and adjoint converges faster in these norms.

5 Conclusion and outlook

In this paper, we have derived a new abstract error estimator for optimal control problems. This abstract estimator was applied to formulate residual-based estimators for the finite element discretization of two optimal control problems. The novelty of the approach is that we measure the error in spaces which are motivated by the objective. Numerical examples confirm that our approach is superior to an energy-norm based approach if we use the error estimators for an adaptive refinement of the mesh.

Due to the abstract theory of Section 2, it is straightforward to extend the idea of our error estimator to other linear-quadratic problems, e.g., to Neumann boundary control problems, provided suitable error estimators for the PDE discretization with respect to the considered norms/spaces are available.

We mention that it is also possible to apply our error-estimate for post-processed controls, see Meyer and Rösch [2004], or for the variational discretization, see Hinze [2005]. In these cases, no error occurs

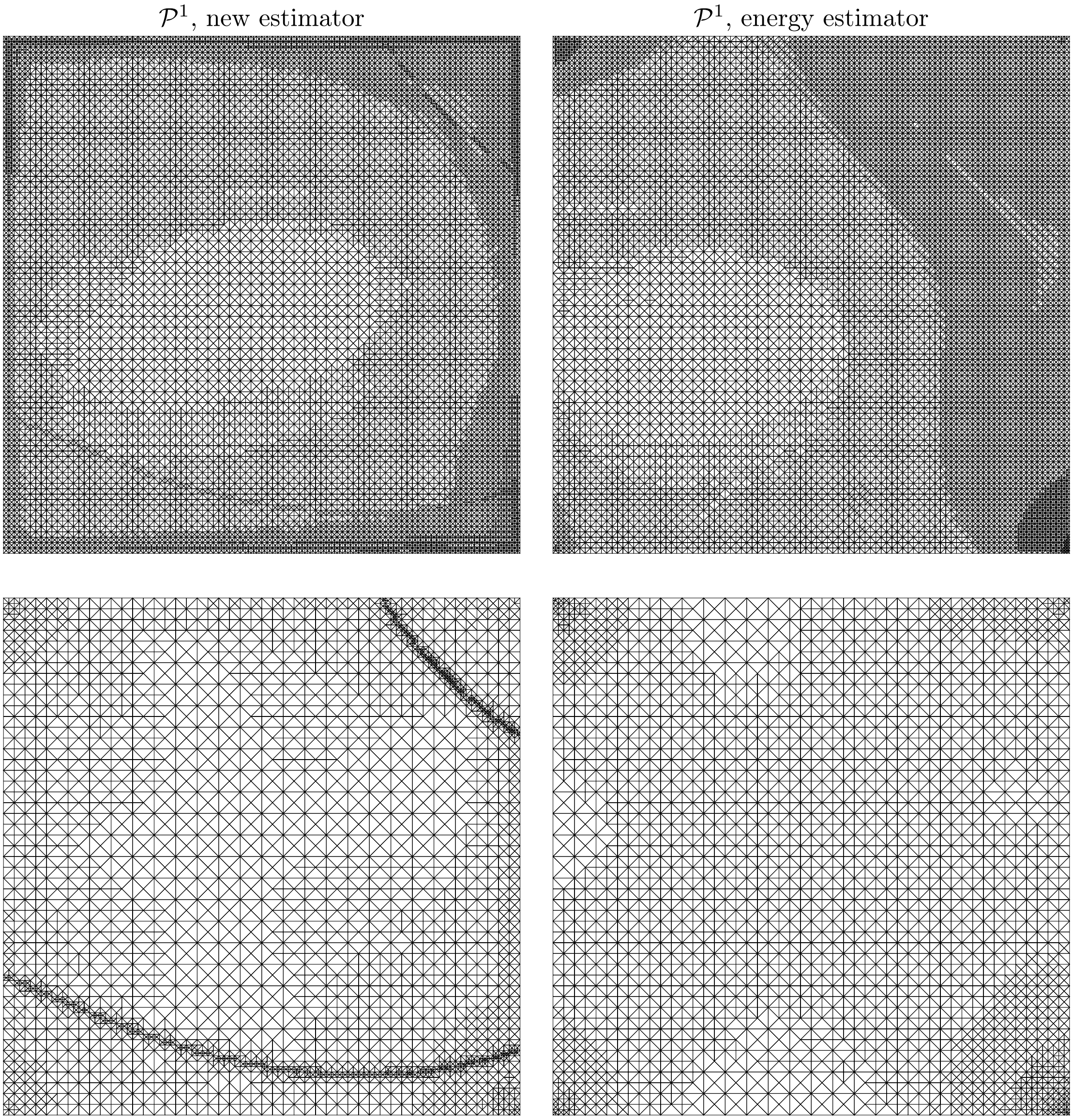
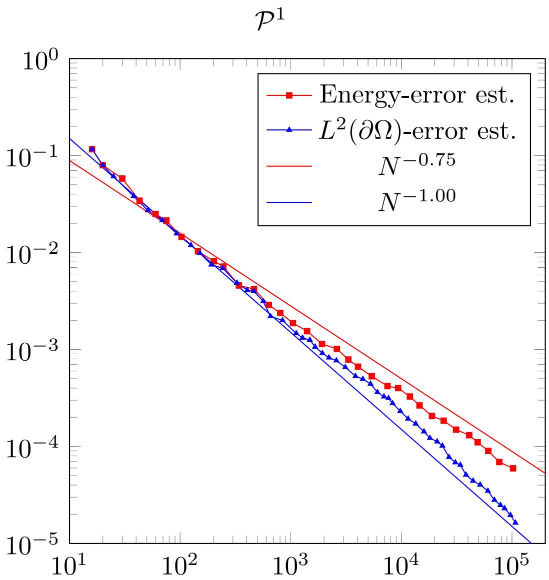


Figure 4.2: Meshes obtained by an adaptive refinement for the solution of (4.1) with $\mathcal{P}^1$ (top row) and $\mathcal{P}^2$ elements (bottom row). For the left column, we used the error estimator from Theorem 4.2 and an energy-based error estimator (4.12) for the right column.



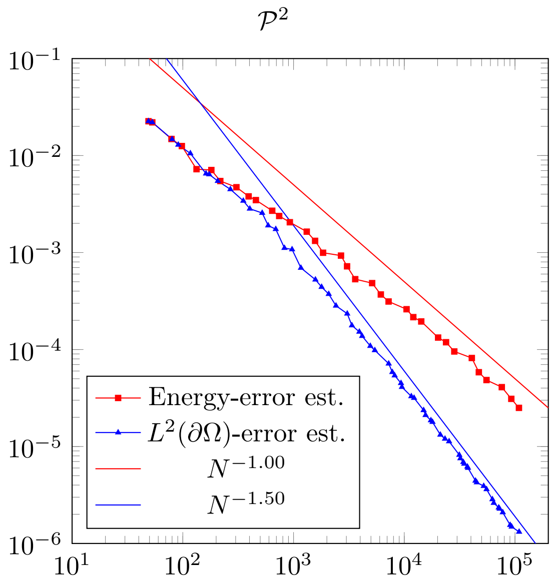


Figure 4.3: Total error (4.2) for the discretization of (4.1) with $\mathcal{P}^1$ (left) and $\mathcal{P}^2$ (right) elements for adaptive refinement based on the error estimator from Theorem 4.2 (red squares) and the energy-based error estimator (4.12) (blue triangles), respectively. The error is plotted versus the degrees of freedom N (per variable y_h, u_h, p_h). Slopes N^{-p} are provided for comparison.

in the variational inequality and the error estimator becomes

$$\frac{1}{\alpha}\, \|\rho_h - \mathcal{S}^\star\,(z_d - z_h)\|_U^2 + \|\mathcal{S}\, u_h - z_h\|_Z^2.$$

Here, u_h is the post-processed control in case of post-processing.

The extension to nonlinear optimal control problems may be the subject of further research.

References

R. Becker, M. Braack, D. Meidner, R. Rannacher, and B. Vexler. Adaptive finite element methods for PDE-constrained optimal control problems. In *Reactive flows, diffusion and transport*, pages 177–205. Springer, Berlin, 2007. doi: 10.1007/978-3-540-28396-6_8.

S. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, second edition, 2002.

M. Dauge. *Elliptic Boundary Value Problems on Corner Domains*. Number 1341 in Lecture Notes in Mathematics. Springer, 1988.

M. Feistauer. On the finite element approximation of functions with noninteger derivatives. *Numerical Functional Analysis and Optimization. An International Journal*, 10(1-2):91–110, 1989. ISSN 0163-0563. doi: 10.1080/01630568908816293.

P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.

M. Hintermüller, R. H. W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM: Control, Optimisation and Calculus of Variations*, 14:540–560, 2008.

M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Computational Optimization and Applications*, 30(1):45–61, 2005. ISSN 0926-6003. doi: 10.1007/s10589-005-4559-5.

D. S. Jerison and C. E. Kenig. The Neumann problem on Lipschitz domains. *American Mathematical Society. Bulletin. New Series*, 4(2):203–207, 1981. ISSN 0273-0979. doi: 10.1090/S0273-0979-1981-14884-9.

B. N. Khoromskij and J. M. Melenk. Boundary concentrated finite element methods. *SIAM Journal on Numerical Analysis*, 41(1):1–36, 2003. ISSN 0036-1429. doi: 10.1137/S0036142901391852.

K. Kohls, A. Rösch, and K. G. Siebert. A posteriori error estimators for control constrained optimal control problems. In *Constrained optimization and optimal control for partial differential equations*, volume 160 of *Internat. Ser. Numer. Math.*, pages 431–443. Birkhäuser/Springer Basel AG, Basel, 2012. doi: 10.1007/978-3-0348-0133-1_22.

K. Kohls, A. Rösch, and K. G. Siebert. A posteriori error analysis of optimal control problems with control constraints. *SIAM Journal on Control and Optimization*, 52(3):1832–1861, 2014. ISSN 0363-0129. doi: 10.1137/130909251.

W. Liu and N. Yan. A posteriori error estimates for distributed convex optimal control problems. *Journal Advances in Computational Mathematics*, 15:285–309, 2001. doi: 10.1023/A:1014239012739.

C. Meyer and A. Rösch. Superconvergence properties of optimal control problems. *SIAM Journal on Control and Optimization*, 43(3):970–985, 2004. ISSN 0363-0129. doi: 10.1137/S0363012903431608.

F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg, Wiesbaden, second edition, 2009.

R. Verfürth. A posteriori error estimates for nonlinear problems. L^r-estimates for finite element discretizations of elliptic equations. *RAIRO Modélisation Mathématique et Analyse Numérique*, 32(7): 817–842, 1998. ISSN 0764-583X.

R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Advances in numerical mathematics. Wiley-Teubner, 1996.

B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM Journal on Control and Optimization*, 47(1):509–534, 2008.

N. Yan and Z. Zhou. A priori and a posteriori error estimates of streamline diffusion finite element method for optimal control problem governed by convection dominated diffusion equation. *Numerical Mathematics. Theory, Methods and Applications*, 1(3):297–320, 2008. ISSN 1004-8979. doi: 10.1016/j.cam.2008.01.006.

Chemnitz Scientific Computing Preprints – ISSN 1864-0087