

**Universität
Stuttgart**

**Fachbereich
Mathematik**

**A Posteriori Error Analysis of Optimal Control
Problems with Control Constraints**

Kristina Kohls, Arnd Rösch, Kunibert G. Siebert

Preprint 2013/001

Fachbereich Mathematik
Fakultät Mathematik und Physik
Universität Stuttgart
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: preprints@mathematik.uni-stuttgart.de
WWW: <http://www.mathematik.uni-stuttgart.de/preprints>

ISSN **1613-8309**

© Alle Rechte vorbehalten. Nachdruck nur mit Genehmigung des Autors.
L^AT_EX-Style: Winfried Geis, Thomas Merkle

A POSTERIORI ERROR ANALYSIS OF OPTIMAL CONTROL PROBLEMS WITH CONTROL CONSTRAINTS

KRISTINA KOHLS, ARND RÖSCH, AND KUNIBERT G. SIEBERT

ABSTRACT. We derive a unifying framework for the a posteriori error analysis of control constrained linear-quadratic optimal control problems. We consider finite element discretizations with discretized and non-discretized control. A fundamental error equivalence drastically simplifies the a posteriori error analysis for optimal control problems. It basically remains to apply error estimators for the linear state and adjoint problem. We give several examples, including stabilized discretizations, and investigate the quality of the estimators and the performance of the adaptive iteration by selected numerical experiments.

1. INTRODUCTION

Many optimization processes in science and engineering lead to optimal control problems where the sought state is a solution of a partial differential equation (PDE). Control and state may be subject to further constraints. The complexity of such problems requires sophisticated techniques for an efficient numerical approximation of the true solution. One particular method are adaptive finite element discretizations, which are typically based on the following iteration:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \quad (1.1)$$

The module **SOLVE** solves the optimal control problem in a finite element space that is defined over a given grid. The module **ESTIMATE** computes local error indicators that constitute an estimator for the true error. Based on information of the indicators the module **MARK** selects a subset of elements subject to refinement. This is executed by **REFINE** returning a suitable refinement of the current grid. In this paper we focus on the module **ESTIMATE**, i. e., the construction of *reliable* and *efficient* a posteriori error estimators for the optimal control problem in terms of the discrete solution and given data.

Starting with the seminal paper [BR78] by Babuška and Rheinboldt we nowadays have access to a well-established theory for the a posteriori error analysis of finite element discretizations of elliptic problems. This includes general inf-sup stable problems, various kinds of discretizations, and different types of error estimators. We refer to the monographs [AO00, BS01, BR03, Ver96] and the references therein.

In contrast to this, a posteriori error analysis for constrained optimal control problems is quite young. It was initiated by Liu and Yan at the beginning of this century [LY01]. Compared to the vast amount of literature about a posteriori error estimators for linear problems the existing results for optimal control problems are

2000 *Mathematics Subject Classification.* Primary 65K10, 65N30, 65N12.

Key words and phrases. Finite elements, adaptive method, a posteriori error analysis, optimal control, control constraints.

This work was supported in part via DFG research grant SI-814/5-1 “*Design and Analysis of Adaptive Finite Element Discretizations for Optimal Control Problems*” within the priority program SPP 1523 “*Optimization with Partial Differential Equations*”.

rather limited. We would like to mention [BBM⁺07, BV09, dLRMV08, HHIK08, HH08, HH10, HK10, LY02, LLY07, VW08, YZ08].

The a posteriori error analysis for constrained optimal control problems is inevitably more technical and complicated due to its intrinsic nonlinear character. Looking into existing papers one gets the following impression. On the one hand, many arguments are similar or sometimes even exactly the same as in the linear case. On the other hand, any new PDE constraint seems to require a new analysis without directly employing existing results from the linear theory. Moreover, the used techniques do not use the *residual* of the optimal control problem. This in turn does not allow to use current techniques for proving convergence of the adaptive iteration (1.1); compare with [MSV08, Sie11].

This assessment was the starting point for the presented research, which we have already sketched in [KRS12]. The key achievement is a unifying framework for the a posteriori error analysis of linear-quadratic optimal control problems with control constraints. The a posteriori error analysis then basically reduces to sum up reliable and efficient estimators for the *linear* state and adjoint problem. This procedure has at least two important advantages. Firstly, error estimation for optimal control problems can now take full advantage of the rich toolbox of existing estimators for linear problems. This includes the standard residual estimator as used in [CKNS08], variations of it [AO00, Ver96, VV09], the hierarchical estimator [SV07, Vee02, Ver96], estimators based on the solution of local problems [MNS03], estimators using recovering techniques [ZZ87, CB02], or the equilibrated residual estimators employing the famous Prager&Synge theorem [Bra07, BPS09]. Many of them have not yet been used for optimal control problems. Secondly, replacing the residual of the optimal control problem by the residuals of the linear subproblems is then the key to prove convergence of the adaptive iteration (1.1) for the optimal control problem [Koh13].

The basis of our analysis is a fundamental error equivalence of the solution of the optimal control problem to solutions of the linear state and adjoint problem. This equivalence is discretization independent and solely relies on the continuous problem being well-posed. The required properties of the continuous problem and the error equivalence is topic of §2. The discretization independent error equivalence then in turn includes the standard discretization [Trö10], the variational discretization by Hinze [Hin05], as well as stabilized discretizations like the SDFEM [RST08]. Required properties of the discretization and basic assumptions on the estimators for the linear subproblems are listed in §3. We then construct an estimator for the optimal control problem and prove reliability and efficiency.

The abstract framework is then put into life. In §4 we consider a reaction diffusion problem with boundary control and discretized control space. We employ the standard residual estimator and the hierarchical estimator for the linear subproblems. The Oseen equations with distributed control and a stabilized discretization with non-discretized control space is topic of §5. The derivation of a residual estimator exemplifies how much the abstract framework simplifies the error analysis. We conclude by two numerical experiments in §6. Resorting to the problem and the estimators from §4 we investigate the estimator quality and the performance of the adaptive method. Both estimators show adequate numerical effectivity indices and they are not sensitive to changes in the active and inactive sets. The adaptive method reveals an optimal adaptive decay rate also for a singular solution.

2. THE OPTIMAL CONTROL PROBLEM AND BASIC ERROR EQUIVALENCE

We state the abstract control constrained optimal control problem together with assumptions on the structure of the objective, spaces, and bilinear form for the constraint. We then prove the fundamental error equivalence.

2.1. The Optimal Control Problem. For given $f \in \mathbb{Y}^*$ we consider optimal control problems of the form

$$\min_{u \in \mathbb{U}^{\text{ad}}} \mathcal{J}[u, y] \quad \text{subject to} \quad \mathcal{B}[y, v] = \langle f + u, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y}, \quad (2.1)$$

where u is the control, and y the corresponding state. We have a particular interest in bilinear forms \mathcal{B} that arise from the variational formulation of PDEs in the following setting.

Assumptions on data of (2.1). We suppose that the control space $(\mathbb{U}, \langle \cdot, \cdot \rangle_{\mathbb{U}})$ as well as the state space $(\mathbb{Y}, \langle \cdot, \cdot \rangle_{\mathbb{Y}})$ are both L_2 -based Hilbert spaces. The state space is defined over some bounded domain $\Omega \subset \mathbb{R}^d$ and the control space is defined over some subset $\Gamma \subset \bar{\Omega}$, which allows us to simultaneously treat distributed controls, i. e., $\Gamma \subset \Omega$ with positive d -dimensional Lebesgue measure, and boundary controls, i. e., $\Gamma \subset \partial\Omega$ with positive $(d-1)$ -dimensional Hausdorff measure. We assume that the square of the induced norms with induced norms $\|\cdot\|_{\mathbb{Y}} = \|\cdot\|_{\mathbb{Y}(\Omega)}$ and $\|\cdot\|_{\mathbb{U}} = \|\cdot\|_{\mathbb{U}(\Gamma)}$ are additive, i. e., for any measurable subsets $\omega_1, \omega_2 \subset \Omega$ with $|\omega_1 \cap \omega_2| = 0$ and $v \in \mathbb{Y}$ we have $\|v\|_{\mathbb{Y}(\omega_1 \cup \omega_2)}^2 = \|v\|_{\mathbb{Y}(\omega_1)}^2 + \|v\|_{\mathbb{Y}(\omega_2)}^2$, and, similarly for \mathbb{U} .

The convex, closed and non-empty set $\mathbb{U}^{\text{ad}} \subset \mathbb{U}$ denotes the set of admissible controls and the spaces \mathbb{Y} and \mathbb{U} are connected via the embeddings $\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$ in the following sense: \mathbb{Y}^* is the (topological) dual space of \mathbb{Y} , $v \in \mathbb{Y}$ implies $v \in \mathbb{U}$ with $\|v\|_{\mathbb{U}} \leq \gamma \|v\|_{\mathbb{Y}}$ and $u \in \mathbb{U}$ implies $u \in \mathbb{Y}^*$ by

$$\langle u, v \rangle := \langle u, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} = \langle u, v \rangle_{\mathbb{U}} \quad \forall v \in \mathbb{Y}.$$

Technically there might be some inclusion operators, like traces, involved. For ease of presentation they are omitted in the notation.

We assume that $\mathcal{B}: \mathbb{Y} \times \mathbb{Y} \rightarrow \mathbb{R}$ is a continuous bilinear form, with continuity constant $\|\mathcal{B}\|$, satisfying the inf-sup condition

$$\inf_{\substack{v \in \mathbb{Y} \\ \|v\|_{\mathbb{Y}}=1}} \sup_{\substack{w \in \mathbb{Y} \\ \|w\|_{\mathbb{Y}}=1}} \mathcal{B}[v, w] = \inf_{\substack{w \in \mathbb{Y} \\ \|w\|_{\mathbb{Y}}=1}} \sup_{\substack{v \in \mathbb{Y} \\ \|v\|_{\mathbb{Y}}=1}} \mathcal{B}[v, w] = \beta > 0. \quad (2.2)$$

For the definition of the objective \mathcal{J} we consider a Fréchet differentiable, quadratic and strictly convex functional $\psi: \mathbb{Y} \rightarrow \mathbb{R}$. We suppose that ψ' is locally Lipschitz continuous with constant L , i. e., $\|\psi'(y) - \psi'(\bar{y})\|_{\mathbb{Y}^*(\omega)} \leq L\|y - \bar{y}\|_{\mathbb{Y}(\omega)}$ for all $y, \bar{y} \in \mathbb{Y}$ and $\omega \subset \Omega$. Given a cost parameter $\alpha > 0$ for the control we finally define

$$\mathcal{J}[u, y] := \psi(y) + \frac{\alpha}{2} \|u\|_{\mathbb{U}}^2.$$

For the discretization we additionally suppose that Ω and Γ can be meshed; the precise assumptions are stated in §3.1. In §4 and §5 we give typical examples of PDEs with corresponding Hilbert spaces \mathbb{Y} and bilinear forms \mathcal{B} , control spaces \mathbb{U} and sets of admissible controls \mathbb{U}^{ad} , as well as objective functionals ψ for the desired state.

Existence and uniqueness. The inf-sup condition (2.2) is equivalent to unique solvability of the *state* and *adjoint equation*

$$y \in \mathbb{Y}: \quad \mathcal{B}[y, v] = \langle q, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y} \quad (2.3a)$$

$$p \in \mathbb{Y}: \quad \mathcal{B}[v, p] = \langle q, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y} \quad (2.3b)$$

for any $q \in \mathbb{Y}^*$ [Neč62, Theorem 5.3]; compare also [NSV09, §2.3] for a more detailed discussion of the inf-sup theory. This particularly implies the existence of solution

operators $S, S^*: \mathbb{Y}^* \rightarrow \mathbb{Y}$ of the state and the adjoint equation. This means, for given $q \in \mathbb{Y}^*$ the unique solutions of (2.3a) and (2.3b) are $y = S(q)$ and $p = S^*(q)$, respectively. Besides that

$$\|S\|_{\mathcal{L}(\mathbb{Y}^*, \mathbb{Y})} \leq \beta^{-1} \quad \text{and} \quad \|S^*\|_{\mathcal{L}(\mathbb{Y}^*, \mathbb{Y})} \leq \beta^{-1}. \quad (2.4)$$

The assumptions on the embedding $\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$ imply for any $u \in \mathbb{U}$ the bound $\|u\|_{\mathbb{Y}^*} \leq \gamma \|u\|_{\mathbb{U}}$, which results in

$$\|S\|_{\mathcal{L}(\mathbb{U}, \mathbb{Y})} \leq \beta^{-1} \gamma. \quad (2.5)$$

Utilizing the solution operator S , we introduce the reduced cost functional $\hat{\mathcal{J}}[u] := \mathcal{J}[u, S(f + u)]$. The assumptions on ψ , $\alpha > 0$, and the inf-sup condition (2.2) imply that $\hat{\mathcal{J}}[u]$ is bounded from below and strictly convex. If \mathbb{U}^{ad} is not bounded $\hat{\mathcal{J}}[u]$ is radially unbounded. In combination with the fact that \mathbb{U}^{ad} is non-empty, convex and closed the following theorem emerges; compare with [Lio71, Trö10].

Theorem 2.1 (Existence and Uniqueness). *The constrained optimal control problem (2.1) has a unique solution $(\hat{u}, \hat{y}) = (\hat{u}, S(f + \hat{u})) \in \mathbb{U}^{\text{ad}} \times \mathbb{Y}$.*

Introducing the adjoint state $\hat{p} = S^ \psi'(\hat{y}) \in \mathbb{Y}$, the triplet $(\hat{u}, \hat{y}, \hat{p}) \in \mathbb{U}^{\text{ad}} \times \mathbb{Y} \times \mathbb{Y}$ is characterized as the unique solution of the first order optimality system*

$$\text{state equation:} \quad \mathcal{B}[\hat{y}, v] = \langle f + \hat{u}, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y}, \quad (2.6a)$$

$$\text{adjoint equation:} \quad \mathcal{B}[v, \hat{p}] = \langle \psi'(\hat{y}), v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y}, \quad (2.6b)$$

$$\text{gradient equation:} \quad \hat{u} = \Pi(\hat{p}), \quad (2.6c)$$

where $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$ denotes for given p the best approximation of $-\frac{1}{\alpha}p$ in \mathbb{U}^{ad} , i. e., $\Pi(p)$ is uniquely determined by

$$\langle \alpha \Pi(p) + p, \Pi(p) - u \rangle_{\mathbb{U}} \leq 0 \quad \forall u \in \mathbb{U}^{\text{ad}}. \quad (2.7)$$

2.2. Basic Error Equivalence. In this section we derive the basic error equivalence that is fundamental in the a posteriori error analysis of conforming discretizations of (2.6).

In order to simplify the presentation we use $a \lesssim b$ for $a \leq Cb$ with a constant that may depend on data constants $\alpha, \beta, \gamma, \|\mathcal{B}\|$, and L . We write $a \approx b$ for $a \lesssim b$ and $b \lesssim a$. This brings us in the position to state the main result of this section.

Theorem 2.2 (Error Equivalence). *If $(u, y, p) \in \mathbb{U} \times \mathbb{Y} \times \mathbb{Y}$ is arbitrarily chosen and $(\hat{u}, \hat{y}, \hat{p}) \in \mathbb{U}^{\text{ad}} \times \mathbb{Y} \times \mathbb{Y}$ is the solution to the optimality system (2.6) then*

$$\|u - \hat{u}\|_{\mathbb{U}} + \|y - \hat{y}\|_{\mathbb{Y}} + \|p - \hat{p}\|_{\mathbb{Y}} \approx \|u - \Pi(p)\|_{\mathbb{U}} + \|y - S(f + u)\|_{\mathbb{Y}} + \|p - S^* \psi'(y)\|_{\mathbb{Y}},$$

where $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$ is the nonlinear projection operator of (2.6c), $S, S^*: \mathbb{Y}^* \rightarrow \mathbb{Y}$ are the solution operators of the linear state (2.6a) and adjoint (2.6b) equation, respectively. The constants hidden in “ \approx ” are explicitly stated below in the proof.

We shall prove this theorem in several steps below. We would like to stress that no special properties of (u, y, p) are required in Theorem 2.2. As a consequence, stabilized discretizations for the approximation of $(\hat{u}, \hat{y}, \hat{p})$ are included in our theory. Compare with the applications in §4 and §5.

We start the proof of Theorem 2.2 by estimating the error in the state.

Lemma 2.3. *We have*

$$\|y - \hat{y}\|_{\mathbb{Y}} \leq \frac{\gamma}{\beta} \|u - \hat{u}\|_{\mathbb{U}} + \|y - S(f + u)\|_{\mathbb{Y}},$$

$$\|y - S(f + u)\|_{\mathbb{Y}} \leq \frac{\gamma}{\beta} \|u - \hat{u}\|_{\mathbb{U}} + \|y - \hat{y}\|_{\mathbb{Y}}.$$

Proof. From $\hat{y} = S(f + \hat{u})$ we deduce with (2.5)

$$\|S(f + u) - \hat{y}\|_{\mathbb{Y}} = \|S(u - \hat{u})\|_{\mathbb{Y}} \leq \beta^{-1}\gamma\|u - \hat{u}\|_{\mathbb{U}}.$$

This implies the first claim by

$$\|y - \hat{y}\|_{\mathbb{Y}} \leq \|y - S(f + u)\|_{\mathbb{Y}} + \|S(f + u) - \hat{y}\|_{\mathbb{Y}} \leq \|y - S(f + u)\|_{\mathbb{Y}} + \beta^{-1}\gamma\|u - \hat{u}\|_{\mathbb{U}}.$$

The second bounds follows similarly by interchanging \hat{y} and $S(f + u)$. \square

A bound of the same type can be derived for the adjoint state.

Lemma 2.4. *We have*

$$\begin{aligned} \|p - \hat{p}\|_{\mathbb{Y}} &\leq \frac{\gamma L}{\beta^2}\|u - \hat{u}\|_{\mathbb{U}} + \frac{L}{\beta}\|y - S(f + u)\|_{\mathbb{Y}} + \|p - S^*\psi'(y)\|_{\mathbb{Y}}, \\ \|p - S^*\psi'(y)\|_{\mathbb{Y}} &\leq \frac{L}{\beta}\|y - \hat{y}\|_{\mathbb{Y}} + \|p - \hat{p}\|_{\mathbb{Y}}. \end{aligned}$$

Proof. The adjoint state is $\hat{p} = S^*(\psi'(\hat{y}))$. Using similar arguments as before we obtain by (2.4) with the Lipschitz constant L of ψ'

$$\|S^*\psi'(y) - \hat{p}\|_{\mathbb{Y}} = \|S^*(\psi'(y) - \psi'(\hat{y}))\|_{\mathbb{Y}} \leq \beta^{-1}\|\psi'(y) - \psi'(\hat{y})\|_{\mathbb{Y}^*} \leq \beta^{-1}L\|y - \hat{y}\|_{\mathbb{Y}}.$$

Combining this with Lemma 2.3 gives

$$\begin{aligned} \|p - \hat{p}\|_{\mathbb{Y}} &\leq \|p - S^*\psi'(y)\|_{\mathbb{Y}} + \|S^*\psi'(y) - \hat{p}\|_{\mathbb{Y}} \\ &\leq \|p - S^*\psi'(y)\|_{\mathbb{Y}} + \beta^{-1}L(\|y - S(f + u)\|_{\mathbb{Y}} + \beta^{-1}\gamma\|u - \hat{u}\|_{\mathbb{U}}), \end{aligned}$$

which proves the first bound. The second one is shown similarly by

$$\|p - S^*\psi'(y)\|_{\mathbb{Y}} \leq \|p - \hat{p}\|_{\mathbb{Y}} + \|S^*(\psi'(\hat{y}) - \psi'(y))\|_{\mathbb{Y}} \leq \|p - \hat{p}\|_{\mathbb{Y}} + \beta^{-1}L\|y - \hat{y}\|_{\mathbb{Y}}. \quad \square$$

Main work has to be done for the error in the control.

Proposition 2.5. *We have*

$$\begin{aligned} \|u - \hat{u}\|_{\mathbb{U}} &\leq \left(1 + \frac{\gamma^2 L}{\alpha\beta^2}\right)\|u - \Pi(p)\|_{\mathbb{U}} + \frac{\gamma L}{\alpha\beta}\|y - S(f + u)\|_{\mathbb{Y}} + \frac{\gamma}{\alpha}\|p - S^*\psi'(y)\|_{\mathbb{Y}}, \\ \|u - \Pi(p)\|_{\mathbb{U}} &\leq \|u - \hat{u}\|_{\mathbb{U}} + \frac{\gamma}{\alpha}\|p - \hat{p}\|_{\mathbb{Y}}. \end{aligned}$$

Proof. We proceed in several steps.

\square We introduce the auxiliary control $\bar{u} = \Pi(p)$ and prove first

$$\|\bar{u} - \hat{u}\|_{\mathbb{U}} \leq \frac{\gamma L}{\alpha\beta}\|y - S(f + \bar{u})\|_{\mathbb{Y}} + \frac{\gamma}{\alpha}\|p - S^*\psi'(y)\|_{\mathbb{Y}}. \quad (2.8)$$

Recalling $\hat{u} = \Pi(\hat{p})$ and $\bar{u} = \Pi(p)$, the definition (2.7) of Π implies

$$\langle \alpha\bar{u} + p, \bar{u} - \hat{u} \rangle_{\mathbb{U}} \leq 0 \quad \text{and} \quad \langle \alpha\hat{u} + \hat{p}, \hat{u} - \bar{u} \rangle_{\mathbb{U}} \leq 0.$$

From this we conclude with $\hat{p} = S^*(\psi'(\hat{y}))$

$$\begin{aligned} \alpha\|\bar{u} - \hat{u}\|_{\mathbb{U}}^2 &= \langle \alpha\bar{u} + p, \bar{u} - \hat{u} \rangle_{\mathbb{U}} + \langle \alpha\hat{u} + \hat{p}, \hat{u} - \bar{u} \rangle_{\mathbb{U}} + \langle p - \hat{p}, \hat{u} - \bar{u} \rangle_{\mathbb{U}} \\ &\leq \langle p - \hat{p}, \hat{u} - \bar{u} \rangle_{\mathbb{U}} \\ &= \langle p - S^*\psi'(y), \hat{u} - \bar{u} \rangle_{\mathbb{U}} + \langle S^*(\psi'(y) - \psi'(\hat{y})), \hat{u} - \bar{u} \rangle_{\mathbb{U}}. \end{aligned}$$

The first term on the right hand side we bound by

$$\langle p - S^*\psi'(y), \hat{u} - \bar{u} \rangle_{\mathbb{U}} \leq \|p - S^*\psi'(y)\|_{\mathbb{U}}\|\hat{u} - \bar{u}\|_{\mathbb{U}} \leq \gamma\|p - S^*\psi'(y)\|_{\mathbb{Y}}\|\hat{u} - \bar{u}\|_{\mathbb{U}}.$$

Employing the definition of S and S^* , recalling $\hat{y} = S(f + \hat{u})$, and using the auxiliary state $\tilde{y} := S(f + \bar{u})$ we proceed with the second term by

$$\begin{aligned} \langle \hat{u} - \bar{u}, S^*(\psi'(y) - \psi'(\hat{y})) \rangle_{\mathbb{U}} &= \mathcal{B}[S(\hat{u} - \bar{u}), S^*(\psi'(y) - \psi'(\hat{y}))] \\ &= \langle \psi'(y) - \psi'(\hat{y}), S(\hat{u} - \bar{u}) \rangle_{\mathbb{Y}^* \times \mathbb{Y}} = \langle \psi'(y) - \psi'(\hat{y}), \hat{y} - \tilde{y} \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \\ &= \langle \psi'(y) - \psi'(\tilde{y}), \hat{y} - \tilde{y} \rangle_{\mathbb{Y}^* \times \mathbb{Y}} + \langle \psi'(\tilde{y}) - \psi'(\hat{y}), \hat{y} - \tilde{y} \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \\ &\leq \gamma \beta^{-1} L \|y - \tilde{y}\|_{\mathbb{Y}} \|\bar{u} - \hat{u}\|_{\mathbb{U}} = \gamma \beta^{-1} L \|y - S(f + \bar{u})\|_{\mathbb{Y}} \|\bar{u} - \hat{u}\|_{\mathbb{U}} \end{aligned}$$

since the convexity of ψ yields $\langle \psi'(\tilde{y}) - \psi'(\hat{y}), \hat{y} - \tilde{y} \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \leq 0$. Combining these steps we arrive at

$$\alpha \|\hat{u} - \bar{u}\|_{\mathbb{U}}^2 \leq \gamma \|\bar{u} - \hat{u}\|_{\mathbb{U}} (\beta^{-1} L \|y - S(f + \bar{u})\|_{\mathbb{Y}} + \|p - S^* \psi'(y)\|_{\mathbb{Y}}),$$

which implies (2.8).

2 We start with

$$\|y - S(f + \bar{u})\|_{\mathbb{Y}} \leq \|y - S(f + u)\|_{\mathbb{Y}} + \beta^{-1} \gamma \|u - \bar{u}\|_{\mathbb{U}}.$$

Since $\bar{u} = \Pi(p)$ we conclude the first claim from (2.8) by

$$\begin{aligned} \|u - \hat{u}\|_{\mathbb{U}} &\leq \|u - \bar{u}\|_{\mathbb{U}} + \|\bar{u} - \hat{u}\|_{\mathbb{U}} \\ &\leq \|u - \Pi(p)\|_{\mathbb{U}} + \frac{\gamma L}{\alpha \beta} \|y - S(f + \bar{u})\|_{\mathbb{Y}} + \frac{\gamma}{\alpha} \|p - S^* \psi'(y)\|_{\mathbb{Y}} \\ &\leq \left(1 + \frac{\gamma^2 L}{\alpha \beta^2}\right) \|u - \Pi(p)\|_{\mathbb{U}} + \frac{\gamma L}{\alpha \beta} \|y - S(f + u)\|_{\mathbb{Y}} + \frac{\gamma}{\alpha} \|p - S^* \psi'(y)\|_{\mathbb{Y}}. \end{aligned}$$

3 The Lipschitz constant of Π is α^{-1} , which can be seen from

$$\begin{aligned} \alpha \|\Pi(u_1) - \Pi(u_2)\|_{\mathbb{U}}^2 &= \langle \alpha \Pi(u_1) + u_1, \Pi(u_1) - \Pi(u_2) \rangle_{\mathbb{U}} \\ &\quad + \langle \alpha \Pi(u_2) + u_2, \Pi(u_2) - \Pi(u_1) \rangle_{\mathbb{U}} \\ &\quad + \langle u_1 - u_2, \Pi(u_2) - \Pi(u_1) \rangle_{\mathbb{U}} \\ &\leq \|u_1 - u_2\|_{\mathbb{U}} \|\Pi(u_2) - \Pi(u_1)\|_{\mathbb{U}} \end{aligned}$$

for all $u_1, u_2 \in \mathbb{U}$ by (2.7). This immediately yields

$$\|u - \Pi(p)\|_{\mathbb{U}} \leq \|u - \hat{u}\|_{\mathbb{U}} + \|\Pi(\hat{p}) - \Pi(p)\|_{\mathbb{U}} \leq \|u - \hat{u}\|_{\mathbb{U}} + \alpha^{-1} \gamma \|p - \hat{p}\|_{\mathbb{Y}},$$

and proves the second claim. □

The proof of Theorem 2.2 is now a direct consequence of these auxiliary results.

Proof of Theorem 2.2. Combining the first estimates of Lemmas 2.3, 2.4, and Proposition 2.5 gives the upper bound

$$\begin{bmatrix} \|u - \hat{u}\|_{\mathbb{U}} \\ \|y - \hat{y}\|_{\mathbb{Y}} \\ \|p - \hat{p}\|_{\mathbb{Y}} \end{bmatrix} \leq \begin{bmatrix} 1 + \frac{\gamma^2 L}{\alpha \beta^2} & \frac{\gamma L}{\alpha \beta} & \frac{\gamma}{\alpha} \\ \frac{\gamma}{\beta} + \frac{\gamma^3 L}{\alpha \beta^3} & 1 + \frac{\gamma^2 L}{\alpha \beta^2} & \frac{\gamma^2}{\alpha \beta} \\ \frac{L \gamma}{\beta^2} + \frac{\gamma^3 L^2}{\alpha \beta^4} & \frac{L}{\beta} + \frac{\gamma^2 L^2}{\alpha \beta^3} & 1 + \frac{L \gamma^2}{\alpha \beta^2} \end{bmatrix} \begin{bmatrix} \|u - \Pi(p)\|_{\mathbb{U}} \\ \|y - S(f + u)\|_{\mathbb{Y}} \\ \|p - S^* \psi'(y)\|_{\mathbb{Y}} \end{bmatrix}.$$

The respective second estimates immediately yield the lower bound

$$\begin{bmatrix} \|u - \Pi(p)\|_{\mathbb{U}} \\ \|y - S(f + u)\|_{\mathbb{Y}} \\ \|p - S^* \psi'(y)\|_{\mathbb{Y}} \end{bmatrix} \leq \begin{bmatrix} 1 & 0 & \frac{\gamma}{\alpha} \\ \frac{\gamma}{\beta} & 1 & 0 \\ 0 & \frac{L}{\beta} & 1 \end{bmatrix} \begin{bmatrix} \|u - \hat{u}\|_{\mathbb{U}} \\ \|y - \hat{y}\|_{\mathbb{Y}} \\ \|p - \hat{p}\|_{\mathbb{Y}} \end{bmatrix}.$$

This finishes the proof. □

3. THE DISCRETE PROBLEM AND A POSTERIORI ERROR CONTROL

We state the discretized optimal control problem, and derive the a posteriori error estimator from the basic error equivalence in Theorem 2.2.

3.1. Discretized Optimal Control Problem. We discretize the problem by finite elements and consider two variants. In the first variant we only discretize the state space and work in the continuous control space. Alternatively, we discretize both state and control space. For both approaches we shall use the same notation for the discrete solutions $(\hat{U}, \hat{Y}, \hat{P})$ and only refer to the specific discretization when needed.

Let \mathcal{T} be a conforming, exact and shape-regular triangulation of Ω that is locally quasi-uniform. Suppose that $\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}$ is a conforming finite element space over \mathcal{T} . To also allow for stabilized discretizations we let $\mathcal{B}_{\mathcal{T}}: \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T}) \rightarrow \mathbb{R}$ and $\langle \cdot, \cdot \rangle_{\mathcal{T}}$ be a *discrete* continuous bilinear form and a *discrete* duality pairing on $\mathbb{Y}(\mathcal{T})^* \times \mathbb{Y}(\mathcal{T})$. We suppose that $\mathcal{B}_{\mathcal{T}}$ satisfies the discrete inf-sup condition

$$\inf_{\substack{V \in \mathbb{Y}(\mathcal{T}) \\ \|V\|_{\mathcal{T}}=1}} \sup_{\substack{W \in \mathbb{Y}(\mathcal{T}) \\ \|W\|_{\mathcal{T}}=1}} \mathcal{B}_{\mathcal{T}}[V, W] = \beta(\mathcal{T}) > 0 \quad \text{or} \quad \inf_{\substack{W \in \mathbb{Y}(\mathcal{T}) \\ \|W\|_{\mathcal{T}}=1}} \sup_{\substack{V \in \mathbb{Y}(\mathcal{T}) \\ \|V\|_{\mathcal{T}}=1}} \mathcal{B}_{\mathcal{T}}[V, W] = \beta(\mathcal{T}) > 0,$$

where $\|\cdot\|_{\mathcal{T}}$ is a suitable mesh-dependent norm. In case on non-stabilized discretization, i. e., $\mathcal{B}_{\mathcal{T}} = \mathcal{B}$, we have $\|\cdot\|_{\mathcal{T}} = \|\cdot\|_{\mathbb{Y}}$. Otherwise, $\|\cdot\|_{\mathcal{T}}$ is chosen accordingly to the stabilization and it is typically stronger than $\|\cdot\|_{\mathbb{Y}}$. Any single inf-sup condition implies the other one, thanks to $\mathbb{Y}(\mathcal{T})$ being finite dimensional. Consequently, the inf-sup theory implies that for any $q \in \mathbb{Y}(\mathcal{T})^*$ there exist unique solutions to the discrete state and adjoint equation

$$Y \in \mathbb{Y}(\mathcal{T}) : \quad \mathcal{B}_{\mathcal{T}}[Y, V] = \langle q, V \rangle_{\mathcal{T}} \quad \forall V \in \mathbb{Y}(\mathcal{T}), \quad (3.1a)$$

$$P \in \mathbb{Y}(\mathcal{T}) : \quad \mathcal{B}_{\mathcal{T}}[V, P] = \langle q, V \rangle_{\mathcal{T}} \quad \forall V \in \mathbb{Y}(\mathcal{T}). \quad (3.1b)$$

In the a posteriori error analysis only the *continuous* inf-sup constant β from (2.2) enters but not the *discrete* inf-sup constant $\beta(\mathcal{T})$. Nevertheless, *stable discretizations* with a uniform lower bound on the discrete inf-sup constant $\beta(\mathcal{T}) \geq \underline{\beta} > 0$ are essential for optimal order a priori error estimates [Bab71] and convergence of adaptive methods [MSV08, Sie11].

Non-discretized control. We suppose that the nonlinear projection operator $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$ introduced in (2.7) is computable for discrete functions $V \in \mathbb{Y}(\mathcal{T})$. This assumption allows us to pose the following discrete optimal control problem:

$$\min_{u \in \mathbb{U}^{\text{ad}}} \mathcal{J}[u, Y] \quad \text{s. t.} \quad \mathcal{B}_{\mathcal{T}}[Y, V] = \langle f + u, V \rangle_{\mathcal{T}} \quad \forall V \in \mathbb{Y}(\mathcal{T}). \quad (3.2)$$

Discretized control. In this approach we additionally replace the continuous control space \mathbb{U} by some finite element space $\mathbb{U}(\mathcal{G})$ over a conforming, exact and shape-regular triangulation \mathcal{G} of Γ . Assuming that $\mathbb{U}^{\text{ad}}(\mathcal{G}) := \mathbb{U}^{\text{ad}} \cap \mathbb{U}(\mathcal{G})$ is nonempty, we consider the discretized optimal control problem

$$\min_{U \in \mathbb{U}^{\text{ad}}(\mathcal{G})} \mathcal{J}[U, Y] \quad \text{s. t.} \quad \mathcal{B}_{\mathcal{T}}[Y, V] = \langle f + U, V \rangle_{\mathcal{T}} \quad \forall V \in \mathbb{Y}(\mathcal{T}). \quad (3.3)$$

In this approach the discrete projection operator $\Pi_{\mathcal{G}}: \mathbb{Y}(\mathcal{T}) \rightarrow \mathbb{U}^{\text{ad}}(\mathcal{G})$ with

$$\langle \alpha \Pi_{\mathcal{G}}(P) + P, \Pi_{\mathcal{G}}(P) - U \rangle_{\mathbb{U}} \leq 0 \quad \forall U \in \mathbb{U}^{\text{ad}}(\mathcal{G}) \quad (3.4)$$

replaces the continuous projection $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$. The grids \mathcal{T} and \mathcal{G} are typically connected in the following way. In case of distributed controls $\mathcal{G} \subseteq \mathcal{T}$. In case of boundary controls \mathcal{G} is the trace of \mathcal{T} on Γ . This allows for an easy transfer of information between the finite element spaces $\mathbb{Y}(\mathcal{T})$ and $\mathbb{U}(\mathcal{G})$ as well as an exact and cheap computation of $\langle U, V \rangle_{\mathcal{T}}$ or $\Pi_{\mathcal{G}}(P)$.

First order optimality system. Using the same arguments as in §2.1 we deduce existence and uniqueness of a discrete solution $(\hat{U}, \hat{Y}) \in \mathbb{U}^{\text{ad}} \times \mathbb{Y}(\mathcal{T})$ to (3.2), and $(\hat{U}, \hat{Y}) \in \mathbb{U}^{\text{ad}}(\mathcal{G}) \times \mathbb{Y}(\mathcal{T})$ to (3.3), respectively.

Introducing the discrete adjoint state, the discrete optimal solution $(\hat{U}, \hat{Y}, \hat{P})$ is characterized by the discrete first order optimality system

$$\hat{Y} \in \mathbb{Y}(\mathcal{T}) : \quad \mathcal{B}_{\mathcal{T}}[\hat{Y}, V] = \langle f + \hat{U}, V \rangle_{\mathcal{T}} \quad \forall V \in \mathbb{Y}(\mathcal{T}), \quad (3.5a)$$

$$\hat{P} \in \mathbb{Y}(\mathcal{T}) : \quad \mathcal{B}_{\mathcal{T}}[V, \hat{P}] = \langle \psi'(\hat{Y}), V \rangle_{\mathcal{T}} \quad \forall V \in \mathbb{Y}(\mathcal{T}) \quad (3.5b)$$

and

$$\hat{U} = \Pi(\hat{P}) \in \mathbb{U}^{\text{ad}} \quad \text{or} \quad \hat{U} = \Pi_{\mathcal{G}}(\hat{P}) \in \mathbb{U}^{\text{ad}}(\mathcal{G}) \quad (3.5c)$$

in case of non-discretized control or discretized control, respectively. Note, that in the first case the ‘discrete’ optimal control \hat{U} is not a finite element function in $\mathbb{U}(\mathcal{G})$.

In order to simplify the presentation and to highlight the important arguments we pose the following simplifying assumptions.

(1) We do not commit variational crimes in the assemblage of the discrete optimality system (3.5). This is, all integrals and dualities are evaluated exactly for given data and discrete functions. This assumption is essential since rigorously estimating quadrature errors a posteriori is not possible for general data.

(2) We suppose that we can compute the exact solutions to the discretized optimal control problems (3.2) or (3.3). In practice, one solves the corresponding first order optimality system (3.5) with a semi-smooth Newton (SSN) method; compare with [Hin05, HV09] for non-discretized control and [BIK99, HIK02, KR02] for discretized control. In case of the optimal control problem under consideration the SSN method is equivalent to a primal dual active set strategy. Both methods are iterative solvers that are terminated in practice if the actual iterate is sufficiently close to the exact discrete solution. Assuming that the solver provides information about the distance of the computed solution to the exact discrete solution in an appropriate norm, it is possible to also carry out the a posteriori error analysis for an inexact discrete solution. The inexactness gives rise to an additional consistency error that can be estimated with the information from the solver; compare with [Ver96].

3.2. A Posteriori Error Control. For deriving an estimator for the optimal control problem (2.1) we combine Theorem 2.2 with estimators

$$\mathcal{E}_y(Y, q; \mathcal{T}) = \left(\sum_{T \in \mathcal{T}} \mathcal{E}_y^2(Y, q; T) \right)^{1/2} \quad \text{and} \quad \mathcal{E}_p(P, q; \mathcal{T}) = \left(\sum_{T \in \mathcal{T}} \mathcal{E}_p^2(P, q; T) \right)^{1/2}$$

for the linear problems (2.3a) and (2.3b). We denote by osc_y and osc_p the typical oscillation terms that have to be present in the lower bound. For any subset $\mathcal{T}' \subseteq \mathcal{T}$ we set

$$\text{osc}_y(Y, q; \mathcal{T}') = \left(\sum_{T \in \mathcal{T}'} \text{osc}_y^2(Y, q; T) \right)^{1/2}$$

and similarly for osc_p . For $T \in \mathcal{T}$ we let $\mathcal{N}_{\mathcal{T}}(T) = \{T' \in \mathcal{T} \mid T' \cap T \neq \emptyset\}$ be the set of its direct neighbors, and $\omega_{\mathcal{T}}(T) = \bigcup_{T' \in \mathcal{N}_{\mathcal{T}}(T)} T'$ the local patch around T .

After these preparations we are ready to precisely pose the assumptions on the estimators for the linear sub-problems. Hereafter, the constant hidden in ‘ \lesssim ’ may additionally depend on the shape-regularity of \mathcal{T} but not on a particular source term of (2.3a) and (2.3b).

Assumption 3.1 (Estimators for the Linear Problems). We suppose that \mathcal{E}_y and \mathcal{E}_p have the following properties:

- (1) **Reliability:** The estimators \mathcal{E}_y and \mathcal{E}_p provide an upper bound for the true error, i. e., for any $q_1 \in f + \mathbb{U}$ and $q_2 \in \text{rg}(\psi')$ and the Galerkin approximations $Y_{q_1}, P_{q_2} \in \mathbb{Y}(\mathcal{T})$ of (3.1a) and (3.1b), respectively, we have

$$\|Y_{q_1} - Sq_1\|_{\mathbb{Y}} \lesssim \mathcal{E}_y(Y_{q_1}, q_1; \mathcal{T}), \quad (3.6a)$$

$$\|P_{q_2} - S^*q_2\|_{\mathbb{Y}} \lesssim \mathcal{E}_p(P_{q_2}, q_2; \mathcal{T}). \quad (3.6b)$$

- (2) **Local Efficiency:** The indicators of \mathcal{E}_y and \mathcal{E}_p are local lower bounds for the true error up to oscillation, i. e., for any $Y, P \in \mathbb{Y}(\mathcal{T})$ and any $q_1 \in f + \mathbb{U}$ and $q_2 \in \text{rg}(\psi')$ we have

$$\mathcal{E}_y(Y, q_1; T) \lesssim \|Y - Sq_1\|_{\mathbb{Y}(\omega_T(T))} + \text{osc}_y(Y, q_1; \mathcal{N}_T(T)), \quad (3.7a)$$

$$\mathcal{E}_p(P, q_2; T) \lesssim \|P - S^*q_2\|_{\mathbb{Y}(\omega_T(T))} + \text{osc}_p(P, q_2; \mathcal{N}_T(T)). \quad (3.7b)$$

- (3) **Lipschitz Continuity of Indicators:** The indicators of \mathcal{E}_y and \mathcal{E}_p are Lipschitz continuous with respect to their second arguments, i. e., for $Y, P \in \mathbb{Y}(\mathcal{T})$, $q_1, \bar{q}_1 \in f + \mathbb{U}$, and $q_2, \bar{q}_2 \in \text{rg}(\psi')$ we have

$$|\mathcal{E}_y(Y, q_1; T) - \mathcal{E}_y(Y, \bar{q}_1; T)| \lesssim \|q_1 - \bar{q}_1\|_{\mathbb{U}(\Gamma \cap T)}, \quad (3.8a)$$

$$|\mathcal{E}_p(P, q_2; T) - \mathcal{E}_p(P, \bar{q}_2; T)| \lesssim L\|y - \bar{y}\|_{\mathbb{Y}(T)}, \quad (3.8b)$$

where $y, \bar{y} \in \mathbb{Y}$ are such that $\psi'(y) = q_2$ and $\psi'(\bar{y}) = \bar{q}_2$.

We would like to remark that asking for $q_1 \in f + \mathbb{U}$ and $q_2 \in \text{rg}(\psi')$ instead of $q_1, q_2 \in \mathbb{Y}^*$ in the definition of \mathcal{E}_y and \mathcal{E}_p induces additional regularity of the source terms that is often necessary to derive a specific estimator, for instance the residual estimator.

The estimator for the error in the control is constructed from the indicators $\mathcal{E}_u(U, p; T) = \|U - \Pi(p)\|_{\mathbb{U}(\Gamma \cap T)}$ with the convention $\|U - \Pi(p)\|_{\mathbb{U}(\emptyset)} = 0$, and we set

$$\|U - \Pi(p)\|_{\mathbb{U}}^2 = \mathcal{E}_u^2(U, p; \mathcal{T}) = \sum_{T \in \mathcal{T}} \mathcal{E}_u^2(U, p; T).$$

This brings us in the position to prove the main result of the paper.

Theorem 3.2 (A Posteriori Error Control). *Let $(\hat{u}, \hat{y}, \hat{p})$ be the true solution of (2.1), let $(\hat{U}, \hat{Y}, \hat{P})$ be the discrete solution either of (3.2) or (3.3), and suppose that \mathcal{E}_y and \mathcal{E}_p satisfy Assumption 3.1. Then*

$$\mathcal{E}_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; \mathcal{T}) := \mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T}) + \mathcal{E}_y(\hat{Y}, f + \hat{U}; \mathcal{T}) + \mathcal{E}_p(\hat{P}, \psi'(\hat{Y}); \mathcal{T})$$

is an estimator for the optimal control problem, which is reliable, i. e.,

$$\|\hat{U} - \hat{u}\|_{\mathbb{U}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \mathcal{E}_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; \mathcal{T}),$$

and globally efficient, i. e.,

$$\begin{aligned} \mathcal{E}_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; \mathcal{T}) &\lesssim \|\hat{U} - \hat{u}\|_{\mathbb{U}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} \\ &\quad + \text{osc}_y(\hat{Y}, f + \hat{u}, \mathcal{T}) + \text{osc}_p(\hat{P}, \psi'(\hat{y}); \mathcal{T}). \end{aligned}$$

In case of non-discretized control (3.2) we have $\mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T}) = 0$ and the estimator \mathcal{E}_{ocp} is locally efficient, i. e., for all $T \in \mathcal{T}$ we have

$$\begin{aligned} \mathcal{E}_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; T) &\lesssim \|\hat{U} - \hat{u}\|_{\mathbb{U}(\Gamma \cap T)} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}(\omega_T(T))} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}(\omega_T(T))} \\ &\quad + \text{osc}_y(\hat{Y}, f + \hat{u}, \mathcal{N}_T(T)) + \text{osc}_p(\hat{P}, \psi'(\hat{y}); \mathcal{N}_T(T)). \end{aligned}$$

Proof. \square Reliability is a direct consequence of Theorem 2.2 applied to the discrete solution $(\hat{U}, \hat{Y}, \hat{P})$ in combination with the Assumption 3.1 (1). For non-discretized control we have $\hat{U} = \Pi(\hat{P})$. This implies $\mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T}) = \|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}} = 0$.

[2] Looking at the contribution \mathcal{E}_y we derive from Assumption 3.1 (3) and (2)

$$\begin{aligned}\mathcal{E}_y(\hat{Y}, f + \hat{U}; T) &\lesssim \mathcal{E}_y(\hat{Y}, f + \hat{u}; T) + \|\hat{U} - \hat{u}\|_{\mathbb{U}(\Gamma \cap T)} \\ &\lesssim \|\hat{U} - \hat{u}\|_{\mathbb{U}(\Gamma \cap T)} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}(\omega_{\mathcal{T}}(T))} + \text{osc}_y(\hat{Y}, f + \hat{u}; \mathcal{N}_{\mathcal{T}}(T)).\end{aligned}$$

Arguing similarly for \mathcal{E}_p we derive

$$\begin{aligned}\mathcal{E}_p(\hat{P}, \psi'(\hat{Y}); T) &\lesssim \mathcal{E}_p(\hat{P}, \psi'(\hat{y}); T) + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}(T)} \\ &\lesssim \|\hat{P} - \hat{p}\|_{\mathbb{Y}(\omega_{\mathcal{T}}(T))} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}(T)} + \text{osc}_p(\hat{P}, \psi'(\hat{y}); \mathcal{N}_{\mathcal{T}}(T)).\end{aligned}$$

This shows local efficiency in case of non-discretized control.

[3] In the proof of Proposition 2.5 [3] we have shown that Π is globally Lipschitz-continuous with constant α^{-1} . This yields

$$\mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T}) \leq \|\hat{U} - \Pi(\hat{p})\|_{\mathbb{U}} + \|\Pi(\hat{p}) - \Pi(\hat{P})\|_{\mathbb{U}} \leq \|\hat{U} - \hat{u}\|_{\mathbb{U}} + \alpha^{-1}\gamma\|\hat{P} - \hat{p}\|_{\mathbb{Y}},$$

utilizing the (global) embedding $\mathbb{Y} \hookrightarrow \mathbb{U}$.

[4] Local quasi-uniformity of \mathcal{T} implies

$$\max_{T \in \mathcal{T}} \#\{T' \in \mathcal{T} \mid |\mathcal{N}_{\mathcal{T}}(T') \cap \mathcal{N}_{\mathcal{T}}(T)| > 0\} \lesssim 1.$$

Additivity of $\|\cdot\|_{\mathbb{Y}}^2$ and osc_y^2 then in turn yields

$$\sum_{T \in \mathcal{T}} \|v\|_{\mathbb{Y}(\omega_{\mathcal{T}}(T))}^2 \lesssim \|v\|_{\mathbb{Y}}^2 \quad \text{and} \quad \sum_{T \in \mathcal{T}} \text{osc}_y^2(Y, q; \mathcal{N}_{\mathcal{T}}(T)) \lesssim \text{osc}_y^2(Y, q; \mathcal{T}).$$

This is also true for $\|\cdot\|_{\mathbb{U}}^2$ and osc_p^2 . Consequently, summing up the indicators of step [2] in combination with the bound \mathcal{E}_u yields global efficiency in any case. \square

We conclude this section by some remarks about the control error indicator \mathcal{E}_u in case of discretized control.

Remark 3.3 (Efficiency of \mathcal{E}_u). (1) In case of discretized control the contribution \mathcal{E}_u does not vanish in general. For typically used discretizations and projections Π the term $\mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T})$ is a computable quantity. However, this term is often estimated further to simplify computations; compare for instance with [HHIK08, LY01]. In general this leads to a non-efficient contribution; compare with Remark 6.1 below. (2) The proof of Theorem 3.2 reveals that the contributions \mathcal{E}_y and \mathcal{E}_p are *locally* efficient. The problematic contribution for \mathcal{E}_{ocp} being locally efficient is \mathcal{E}_u . This term is only locally efficient if Π and the embedding $\mathbb{Y} \hookrightarrow \mathbb{U}$ have the following local properties:

- The continuous projection operator $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$ is *locally Lipschitz*, i. e.,

$$\|\Pi(u) - \Pi(\bar{u})\|_{\mathbb{U}(\Gamma \cap T)} \lesssim \|u - \bar{u}\|_{\mathbb{U}(\Gamma \cap T)} \quad \forall T \in \mathcal{T}, u, \bar{u} \in \mathbb{U}.$$

- The embedding $\mathbb{Y} \hookrightarrow \mathbb{U}$ is *locally uniform*, i. e.,

$$\|y\|_{\mathbb{U}(\Gamma \cap T)} \lesssim \|y\|_{\mathbb{Y}(T)} \quad \forall T \in \mathcal{T}, y \in \mathbb{Y}. \quad (3.9)$$

In this case local efficiency of \mathcal{E}_u and, consequently, of \mathcal{E}_{ocp} is evident.

The assumption on Π is for instance fulfilled for box constraints. The embedding $\mathbb{Y} \hookrightarrow \mathbb{U}$ is locally uniform in case of distributed controls but not in case of boundary controls. In the latter case, the bound in (3.9) requires a scaled trace inequality, which is not locally uniform. Consequently, local efficiency of \mathcal{E}_{ocp} is here only granted for non-discretized control.

4. A DIFFUSION-REACTION PROBLEM WITH BOUNDARY CONTROL

In this section we demonstrate how the general framework derived in the previous sections can be used to easily derive a posteriori error estimators. The particular problem introduced here will be used in the numerical experiments presented in §6.

4.1. Data of the Continuous Problem and Discretization. We state data of the continuous problem in terms of the general framework and then describe its discretization.

Data of the continuous problem. We consider the following diffusion-reaction problem with boundary control:

$$-\Delta y + y = f_2 \quad \text{in } \Omega, \quad \nabla y \cdot \mathbf{n} = \begin{cases} u + f_1 & \text{on } \Gamma, \\ 0 & \text{on } \partial\Omega \setminus \Gamma, \end{cases}$$

where $\Gamma \subset \partial\Omega$ has positive Hausdorff measure, and $f_1 \in L_2(\Gamma)$, $f_2 \in L_2(\Omega)$ is given data.

The state space for this example is $\mathbb{Y} = H^1(\Omega)$ with norm $\|\cdot\|_{\mathbb{Y}} = \|\cdot\|_{H^1(\Omega)}$, and the control space is $\mathbb{U} = L_2(\Gamma)$ with norm $\|\cdot\|_{\mathbb{U}} = \|\cdot\|_{L_2(\Gamma)}$. The embedding $\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$ is naturally given by the trace operator and γ equals its norm.

The bilinear form \mathcal{B} for the weak formulation of the PDE is given by

$$\mathcal{B}[y, v] := \int_{\Omega} \nabla v \cdot \nabla y + v y \, dV = \langle y, v \rangle_{H^1(\Omega)} = \langle y, v \rangle_{\mathbb{Y}}.$$

Consequently, \mathcal{B} is continuous and coercive on \mathbb{Y} with $\|\mathcal{B}\| = \beta = 1$. Moreover, setting $\langle f, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} := \langle f_1, v \rangle_{L_2(\Gamma)} + \langle f_2, v \rangle_{L_2(\Omega)}$ we see $f \in \mathbb{Y}^* = (H^1(\Omega))^*$.

We use box constraints, i. e., for given $a, b \in \mathbb{U}(\mathcal{T}_0)$ with $a \leq b$ we let

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid a \leq u \leq b \text{ on } \Gamma\} \neq \emptyset.$$

The continuous projection operator $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$ is therefore given for almost all $x \in \Gamma$ by

$$\Pi(u)(x) = \begin{cases} a(x) & \text{if } -\frac{1}{\alpha}u(x) \leq a(x), \\ -\frac{1}{\alpha}u(x) & \text{if } -\frac{1}{\alpha}u(x) \in [a(x), b(x)], \\ b(x) & \text{if } -\frac{1}{\alpha}u(x) \geq b(x). \end{cases}$$

Finally, for given $g \in \mathbb{U}$ and desired state $y_d \in L_2(\Omega)$ we define the objective

$$\mathcal{J}[u, y] := \psi(y) + \frac{\alpha}{2} \|u\|_{\mathbb{U}}^2 := \frac{1}{2} \|y - y_d\|_{L_2(\Omega)}^2 + \langle g, y \rangle_{L_2(\Gamma)} + \frac{\alpha}{2} \|u\|_{\mathbb{U}}^2.$$

The Fréchet derivative of ψ is given by $\psi'(y) = \langle y - y_d, \cdot \rangle_{L_2(\Omega)} + \langle g, \cdot \rangle_{L_2(\Gamma)} \in \mathbb{Y}^*$. It is locally Lipschitz continuous with constant $L = 1$.

For applying the abstract frame work we need to provide error estimators for the following problems. For given $u \in \mathbb{U}$ solve for $y = S(f + u)$, i. e.,

$$y \in \mathbb{Y} : \quad \mathcal{B}[y, v] = \langle f_1 + u, v \rangle_{L_2(\Gamma)} + \langle f_2, v \rangle_{L_2(\Omega)} \quad \forall v \in \mathbb{Y}, \quad (4.1a)$$

and for given $y \in \mathbb{Y}$ solve for $p = S^* \psi'(y)$, i. e.,

$$p \in \mathbb{Y} : \quad \mathcal{B}[v, p] = \langle y - y_d, v \rangle_{L_2(\Omega)} + \langle g, v \rangle_{L_2(\Gamma)} \quad \forall v \in \mathbb{Y}. \quad (4.1b)$$

Since \mathcal{B} is symmetric and the right hand sides have the same structure, an estimator for (4.1a) is also an estimator for (4.1b).

Discretization. We assume that the macro-triangulation \mathcal{T}_0 meshes Ω exactly and in such a way that $\mathcal{G}_0 := \{T \cap \Gamma \mid T \in \mathcal{T}_0\}$ meshes Γ exactly. We concentrate on the case of discretized control and use piecewise polynomial spaces of degree $\ell_y \geq 1$, $\ell_u \geq 0$ such that

$$\begin{aligned} \mathbb{Y}(\mathcal{T}) &:= \{V \in C^0(\Omega) \mid V|_T \in \mathbb{P}_{\ell_y}(T) \, \forall T \in \mathcal{T}\}, \\ \mathbb{U}(\mathcal{G}) &:= \{V \in L_2(\Gamma) \mid V|_E \in \mathbb{P}_{\ell_u}(E) \, \forall E \in \mathcal{G}\}. \end{aligned}$$

The assumption $a, b \in \mathbb{U}(\mathcal{G}_0)$ implies $\mathbb{U}^{\text{ad}}(\mathcal{G}) = \mathbb{U}(\mathcal{G}) \cap \mathbb{U}^{\text{ad}} \neq \emptyset$. We choose to use $\mathcal{B}_{\mathcal{T}} = \mathcal{B}$ and $\langle \cdot, \cdot \rangle_{\mathcal{T}} = \langle \cdot, \cdot \rangle_{\mathbb{Y}^* \times \mathbb{Y}}$ without any stabilization. The bilinear-form \mathcal{B} is coercive on $\mathbb{Y}(\mathcal{T}) \subset H^1(\Omega)$, which yields the discrete inf-sup condition with

$\beta_{\mathcal{T}} = \beta = 1$. Let $(\hat{U}, \hat{Y}, \hat{P}) \in \mathbb{U}^{\text{ad}}(\mathcal{G}) \times \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T})$ be the solution to the discrete first order optimality system (3.5) in this setting.

4.2. A Posteriori Error Control. We accept $\mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T}) = \|\hat{U} - \Pi(\hat{P})\|_{L_2(\Gamma)}$ as a computable quantity. Therefore Theorem 3.2 yields a reliable and globally efficient estimator \mathcal{E}_{ocp} for the overall error if we provide estimators $\mathcal{E}_y, \mathcal{E}_p$ for the linear subproblems (4.1) that satisfy Assumption 3.1. We give two examples.

Example 4.1 (Residual Estimator). The idea of the residual estimator is to decompose the residual into a regular part in $L_2(\Omega)$ and a singular part in $L_2(\Sigma)$, where Σ is the union of all inter-element sides. The residual estimator is then a sum of properly scaled L_2 norms of the two contributions. The indicators on $T \in \mathcal{T}$ for (4.1a) and (4.1b) read:

$$\begin{aligned} \mathcal{E}_y(\hat{Y}, \hat{U} + f; T) &= h_T \|\Delta \hat{Y} + \hat{Y} - f_2\|_{L_2(T)} \\ &\quad + h_T^{1/2} \|\llbracket \hat{Y} \rrbracket\|_{L_2(\partial T \setminus \Gamma)} + h_T^{1/2} \|\llbracket \hat{Y} \rrbracket - (\hat{U} + f_1)\|_{L_2(\partial T \cap \Gamma)}, \\ \mathcal{E}_p(\hat{P}, \psi'(\hat{Y}); T) &= h_T \|\Delta \hat{P} + \hat{P} - (\hat{Y} - y_d)\|_{L_2(T)} \\ &\quad + h_T^{1/2} \|\llbracket \hat{P} \rrbracket\|_{L_2(\partial T \setminus \partial \Gamma)} + h_T^{1/2} \|\llbracket \hat{P} \rrbracket - g\|_{L_2(\partial T \cap \Gamma)}, \end{aligned}$$

where $h_T = |T|^{1/d}$ is the local mesh-size and $\llbracket \cdot \rrbracket$ denotes the flux of the normal derivative on inter-element sides and the normal derivative for a boundary side.

The estimators \mathcal{E}_y and \mathcal{E}_u satisfy Assumption 3.1: Reliability and local efficiency of the estimator is well-known; compare for instance with [Ver96, Proposition 1.5]. Lipschitz-continuity of the indicators in the second component is a direct consequence of the triangle inequality.

Considering distributed control, [LY01] and [HHIK08] have derived the same estimator contributions (up to appropriate modifications accounting for distributed controls). Only the indicator \mathcal{E}_u for the control error differs; compare with Remark 6.1 below. Being one of the first results concerning error estimation for optimal control problems, their respective proofs are more involved, because the a posteriori error analysis mixes with the abstract error equivalence of Theorem 2.2.

Example 4.2 (Hierarchical Estimator). We restrict ourselves to the case $\ell_y = 1$ and $\ell_u = 0$. The idea of hierarchical estimators is based upon evaluating the residual with sufficiently many discrete functions of an enriched space $\mathbb{Y}(\mathcal{T})' \supsetneq \mathbb{Y}(\mathcal{T})$. Suitable functions are side bubble functions that are either higher order finite elements on the same grid or linear finite elements on a refined mesh. For given $(y, u + f)$ the residual $\text{Res}_y(y, u + f) \in \mathbb{Y}^*$ of the primal problem (4.1a) is

$$\langle \text{Res}_y(y, u + f), v \rangle := \mathcal{B}[y, v] - \langle f_1 + u, v \rangle_{L_2(\Gamma)} - \langle f_2, v \rangle_{L_2(\Omega)} \quad \forall v \in \mathbb{Y},$$

and for given (p, y) the residual $\text{Res}_p(p, \psi'(y)) \in \mathbb{Y}^*$ of the adjoint problem (4.1b) is

$$\langle \text{Res}_p(p, \psi'(y)), v \rangle := \mathcal{B}[v, p] - \langle y - y_d, v \rangle_{L_2(\Omega)} - \langle g, v \rangle_{L_2(\Gamma)} \quad \forall v \in \mathbb{Y}.$$

We denote by \mathcal{S} the set of all sides of \mathcal{T} . If $E \in \mathcal{S}$ is a boundary side there exists a unique $T \in \mathcal{T}$ with $E \subset T$ and we set $\omega_E := T$. Otherwise E is an inter-element side of $T, T' \in \mathcal{T}$ and we set $\omega_E := T \cup T'$. In any case we let z_E be the barycenter of S . We then consider an enrichment $\mathbb{Y}(\mathcal{T})' \supsetneq \mathbb{Y}(\mathcal{T})$ that provides for any $E \in \mathcal{S}$ a function $\Phi_E \in \mathbb{Y}(\mathcal{T})' \setminus \mathbb{Y}(\mathcal{T})$ with

$$\Phi_E(z_E) > 0, \quad \text{supp}(\Phi_E) \subset \omega_E, \quad \text{and} \quad \|\Phi_E\|_{H^1(\Omega)} = 1.$$

We let $h_E := \text{diam}(E)$. For $E \subset \partial\Omega$ we define $\text{osc}_{\mathcal{T}}(g; E) := h_E^{1/2} \|g - \bar{g}\|_{L_2(E)}$ with the mean value $\bar{g} = |E|^{-1} \int_E g \, dA$. For an interior side E we define $\text{osc}_{\mathcal{T}}(g; E) := 0$.

The side-oriented indicators on $E \in \mathcal{S}$ for (4.1a) and (4.1b) read:

$$\begin{aligned}\mathcal{E}_y^2(\hat{Y}, \hat{U} + f; E) &:= |\langle \text{Res}_y(\hat{Y}, \hat{U} + f), \Phi_E \rangle|^2 \\ &\quad + h_E^2 \|\hat{Y} - f_2\|_{L_2(\omega_E)}^2 + \text{osc}_{\mathcal{T}}^2(f_1; E) \\ \mathcal{E}_p^2(\hat{P}, \psi'(\hat{Y}); E) &:= |\langle \text{Res}_p(\hat{P}, \psi'(\hat{Y})), \Phi_E \rangle|^2 \\ &\quad + h_E^2 \|\hat{P} - (\hat{Y} - y_d)\|_{L_2(\omega_E)}^2 + \text{osc}_{\mathcal{T}}^2(g; E).\end{aligned}$$

One easily obtains element-based indicators on $T \in \mathcal{T}$ by summing up the contributions of all sides $E \subset T$.

Reliability and efficiency can be shown as follows. The indicators of the hierarchical estimator are equivalent to the indicators of the residual estimator considered in Example 4.1; compare with [KS11, §2.2.3] or [Ver96, §4]. A direct proof of reliability and efficiency without using the equivalence to the residual estimator can be found in [SV07, Vee02]. Lipschitz-continuity follows directly of Lipschitz-continuity of the residuals $\text{Res}_y, \text{Res}_p$, and the L_2 norms. Consequently, the estimators \mathcal{E}_y and \mathcal{E}_p comply with Assumption 3.1.

Using an enrichment $\mathbb{Y}(\mathcal{T})' \supsetneq \mathbb{Y}(\mathcal{T})$ that additionally provides for any element $T \in \mathcal{T}$ a function $\Phi_T \in \mathbb{Y}(\mathcal{T})' \setminus \mathbb{Y}(\mathcal{T})$ one can replace the terms $h_E \|\hat{Y} - f_2\|_{L_2(\omega_E)}$ $h_E \|\hat{P} - (\hat{Y} - y_d)\|_{L_2(\omega_E)}$ by suitable oscillation terms that are generically of higher order; compare with [KS11, §2.3.3].

To our best knowledge the hierarchical estimator has not been considered before for optimal-control problems.

5. THE OSEEN PROBLEM WITH DISTRIBUTED CONTROL AND STABILIZED DISCRETIZATION

We next demonstrate how the derived framework can be used to easily derive error estimators for the optimal control problem (2.1) with stabilized discretizations of the state and adjoint problem.

5.1. Data of the Continuous Problem and Discretization. We state the continuous problem and then a streamline diffusion discretization.

Data of the continuous problem. We consider the Oseen problem with distributed control:

$$-\Delta \mathbf{y} + [\nabla \mathbf{y}] \mathbf{b} + \nabla q = \mathbf{f} + \mathbf{u} \quad \text{in } \Omega, \quad \text{div } \mathbf{y} = 0 \quad \text{in } \Omega, \quad \mathbf{y} = 0 \quad \text{on } \partial\Omega,$$

where $\mathbf{b} \in L_\infty(\Omega; \mathbb{R}^d)$ with $\text{div } \mathbf{b} = 0$ and $\mathbf{f} \in L_2(\Omega; \mathbb{R}^d)$ is given data. Here, the state $y = (\mathbf{y}, q)$ is the velocity and the pressure of the fluid.

The state space is $\mathbb{Y} = H_0^1(\Omega; \mathbb{R}^d) \times \{q \in L_2(\Omega) \mid \langle q, 1 \rangle_{L_2(\Omega)} = 0\}$. We use the norm $\|y\|_{\mathbb{Y}(\Omega)}^2 = \|\nabla \mathbf{y}\|_{L_2(\Omega)}^2 + \|q\|_{L_2(\Omega)}^2$ for $y = (\mathbf{y}, q) \in \mathbb{Y}$. The space \mathbb{Y} is a Hilbert space with respect to $\|\cdot\|_{\mathbb{Y}}$ thanks to the Friedrichs estimate $\|\mathbf{y}\|_{L_2(\Omega)} \leq C_F \|\nabla \mathbf{y}\|_{L_2(\Omega)}$ on $H_0^1(\Omega; \mathbb{R}^d)$. The control space is $\mathbb{U} = L_2(\Omega; \mathbb{R}^d)$ with the natural norm $\|\cdot\|_{\mathbb{U}(\Omega)} = \|\cdot\|_{L_2(\Omega)}$. The embedding $\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$ is given by the inclusion $H_0^1(\Omega; \mathbb{R}^d) \subset L_2(\Omega; \mathbb{R}^d)$ and $\gamma \leq C_F$.

The bilinear form \mathcal{B} for the weak formulation of the PDE is

$$\mathcal{B}[y, v] = \mathcal{B}[(\mathbf{y}, q), (\mathbf{v}, r)] := \int_{\Omega} \nabla \mathbf{v} : \nabla \mathbf{y} + \mathbf{v} \cdot [\nabla \mathbf{y}] \mathbf{b} - \text{div } \mathbf{v} q + \text{div } \mathbf{y} r \, dV.$$

It is continuous with $\|\mathcal{B}\| \leq 3 + C_F \|\mathbf{b}\|_{L_\infty(\Omega)}$. Moreover, \mathcal{B} satisfies the inf-sup condition (2.2) thanks to $\text{div } \mathbf{b} = 0$, the Friedrichs inequality, and the solvability of the divergence equation [Gal94, Theorem III.3.1].

We use norm constraints, i. e., for given $R > 0$ and a norm $|\cdot|$ on \mathbb{R}^d we set

$$\mathbb{U}^{\text{ad}} = \{\mathbf{u} \in \mathbb{U} \mid |\mathbf{u}| \leq R \text{ in } \Omega\} \neq \emptyset.$$

The continuous projection operator $\Pi: \mathbb{U} \rightarrow \mathbb{U}^{\text{ad}}$ is therefore given for almost all $x \in \Omega$ by

$$\Pi(\mathbf{u})(x) = \begin{cases} -\frac{1}{\alpha} \mathbf{u}(x) & \text{if } \frac{1}{\alpha} |\mathbf{u}(x)| \leq R, \\ -R \frac{\mathbf{u}(x)}{|\mathbf{u}(x)|} & \text{else.} \end{cases}$$

Aiming at minimizing variations in the velocity \mathbf{y} at a low pressure q and recalling $\|\mathbf{y}\|_{\mathbb{Y}}^2 = \|(\mathbf{y}, q)\|_{\mathbb{Y}}^2 = \|\nabla \mathbf{y}\|_{L_2(\Omega)}^2 + \|q\|_{L_2(\Omega)}^2$, we define the objective by

$$\mathcal{J}[\mathbf{u}, y] = \psi(y) + \frac{\alpha}{2} \|\mathbf{u}\|_{\mathbb{U}}^2 := \frac{1}{2} \|\mathbf{y}\|_{\mathbb{Y}}^2 + \frac{\alpha}{2} \|\mathbf{u}\|_{\mathbb{U}}^2.$$

The functional $\psi = \frac{1}{2} \|\cdot\|_{\mathbb{Y}}^2$ is Fréchet differentiable, and ψ' is locally Lipschitz continuous with constant $L = 1$.

For applying the abstract framework we need to provide error estimators for the following problems. For given $\mathbf{u} \in \mathbb{U}$ solve for $y = S(\mathbf{f} + \mathbf{u})$, i. e.,

$$y = (\mathbf{y}, q) \in \mathbb{Y} : \quad \mathcal{B}[y, v] = \int_{\Omega} (\mathbf{f} + \mathbf{u}) \cdot \mathbf{v} \, dV \quad \forall v = (\mathbf{v}, r) \in \mathbb{Y}, \quad (5.1a)$$

and for given $y = (\mathbf{y}, q) \in \mathbb{Y}$ solve for $p = S^* \psi'(y)$, i. e.,

$$p = (\mathbf{p}, s) \in \mathbb{Y} : \quad \mathcal{B}[v, p] = \int_{\Omega} \nabla \mathbf{y} : \nabla \mathbf{v} + q r \, dV \quad \forall v = (\mathbf{v}, r) \in \mathbb{Y}. \quad (5.1b)$$

Since \mathcal{B} is non-symmetric and the right hand sides differ, we need different estimators for (5.1a) and (5.1b).

Discretization. In this example we consider non-discretized control. We let \mathcal{T} be an exact, conforming and shape-regular triangulation of Ω and suppose that the discrete state space $\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}$ of continuous, piecewise polynomials of fixed degree $(\ell_{\mathbf{y}}, \ell_q)$ for velocity and pressure. For the stabilization of the advection derivative $[\nabla \cdot] \mathbf{b}$ we use a streamline-diffusion finite element method (SDFEM); compare for instance with [RST08, Chap. IV]. For the SDFEM we use the discrete bilinear form $\mathcal{B}_{\mathcal{T}}[\cdot, \cdot] = \mathcal{B}[\cdot, \cdot] + \mathcal{B}_{\text{stab}}[\cdot, \cdot]$ and the discrete duality pairing $\langle \cdot, \cdot \rangle_{\mathcal{T}} = \langle \cdot, \cdot \rangle_{\mathbb{Y}^* \times \mathbb{Y}} + \langle \cdot, \cdot \rangle_{\text{stab}}$ with suitable stabilization terms that are defined next.

For parameters $\delta > 0$ and $\mu > 0$, the stabilized part of the bilinear form is defined as

$$\begin{aligned} \mathcal{B}_{\text{stab}}[Y, V] &:= \delta \sum_{T \in \mathcal{T}} h_T^2 \int_T ([\nabla \mathbf{V}] \mathbf{b} + \nabla R) \cdot (-\Delta \mathbf{Y} + [\nabla \mathbf{Y}] \mathbf{b} + \nabla Q) \, dV \\ &\quad + \mu \int_{\Omega} \operatorname{div} \mathbf{V} \operatorname{div} \mathbf{Y} \, dV. \end{aligned}$$

The discrete state equation reads: for given $\mathbf{u} \in \mathbb{U}$ solve for $Y = (\mathbf{Y}, Q) \in \mathbb{Y}(\mathcal{T})$ such that

$$\mathcal{B}_{\mathcal{T}}[Y, V] = \int_{\Omega} (\mathbf{f} + \mathbf{u}) \cdot \mathbf{V} \, dV + \langle \mathbf{f} + \mathbf{u}, V \rangle_{\text{stab}, y} \quad (5.2a)$$

for all $V = (\mathbf{V}, R) \in \mathbb{Y}(\mathcal{T})$, where

$$\langle \mathbf{f} + \mathbf{u}, V \rangle_{\text{stab}, y} = \delta \sum_{T \in \mathcal{T}} h_T^2 \int_T (\mathbf{f} + \mathbf{u}) \cdot ([\nabla \mathbf{V}] \mathbf{b} + \nabla R) \, dV.$$

The discrete adjoint problem then reads: for given $Y = (\mathbf{Y}, Q)$ solve for $P = (\mathbf{P}, S)$ such that

$$\mathcal{B}_{\mathcal{T}}[V, P] = \int_{\Omega} \nabla \mathbf{Y} : \nabla \mathbf{V} + Q R \, dV + \langle Y, V \rangle_{\text{stab}, p} \quad (5.2b)$$

for all $V = (\mathbf{V}, R) \in \mathbb{Y}(\mathcal{T})$ with

$$\langle Y, V \rangle_{\text{stab}, p} = \delta \sum_{T \in \mathcal{T}} h_T^2 \int_T \Delta \mathbf{Y} \cdot ([\nabla \mathbf{V}] \mathbf{b} - \nabla R) \, dV + \mu \int_{\Omega} Q \operatorname{div} \mathbf{V} \, dV.$$

In [RST08, Chap. IV, Lemma 3.3] it has been shown that for sufficiently small $\delta > 0$ the discrete bilinear form $\mathcal{B}_{\mathcal{T}}$ satisfies a discrete inf-sup condition with respect to the mesh-dependent norm $\|\cdot\|_{\mathcal{T}}^2 = \|\cdot\|_{\mathbb{Y}}^2 + \|\cdot\|_{\text{stab}}^2$, where

$$\|V\|_{\text{stab}}^2 = \|(\mathbf{V}, R)\|_{\text{stab}}^2 := \delta \sum_{T \in \mathcal{T}} h_T^2 \|[\nabla \mathbf{V}] \mathbf{b} + \nabla R\|_{L_2(T)}^2 + \mu \|\operatorname{div} \mathbf{V}\|_{L_2(\Omega)}^2.$$

Using the inverse estimate $h_T \|\nabla R\|_{L_2(T)} \lesssim \|R\|_{L_2(T)}$, in combination with the bound $\|\operatorname{div} \mathbf{V}\|_{L_2(\Omega)} \leq \|\nabla \mathbf{V}\|_{L_2(\Omega)}$ we see $\|V\|_{\text{stab}} \leq C(\|\mathbf{b}\|_{L_{\infty}(\Omega)}, \delta, \mu) \|V\|_{\mathbb{Y}}$. The discrete inf-sup constant $\beta(\mathcal{T})$ is bounded from below by $\underline{\beta} > 0$, which solely depends on δ, μ and the shape regularity of \mathcal{T} . Consequently, both (5.2a) and (5.2b) are uniquely solvable. On top of this, the mesh-dependent norm $\|\cdot\|_{\mathcal{T}}$ is used to prove a priori error estimates for the SDFEM discretization (5.2a) of (5.1a); compare with [RST08, Chap. IV, Theorem 3.5].

Remark 5.1 (Stable Discretization). Choosing $\ell_{\mathbf{y}} \geq 2$ and $\ell_q = \ell_{\mathbf{y}} - 1$ gives the Taylor-Hood-Element, which is known to be a stable discretization for the Oseen problem. The SDFEM not only stabilizes the advection derivative $[\nabla \cdot] \mathbf{b}$ but gives at the same time a stable discretization not depending on the chosen polynomial degree for velocity and pressure. In particular, $\ell_{\mathbf{y}} = \ell_q = 1$ is included, which is advantageous in optimal control problems for an easy computation of the projection operator $\Pi: \mathbb{Y}(\mathcal{T}) \rightarrow \mathbb{U}^{\text{ad}}$.

Remark 5.2 (Adjoint Problem). The adjoint problem (5.1b) is the weak form of

$$-\Delta \mathbf{p} - [\nabla \mathbf{p}] \mathbf{b} + \nabla r = -\Delta \mathbf{y} \quad \text{in } \Omega, \quad \operatorname{div} \mathbf{p} = q \quad \text{in } \Omega, \quad \mathbf{p} = \mathbf{0} \quad \text{on } \partial\Omega.$$

This problem is well-posed, since $y = (\mathbf{y}, q) \in \mathbb{Y}$ implies $-\Delta \mathbf{y} \in H^{-1}(\Omega; \mathbb{R}^d)$ and

$$0 = \langle q, 1 \rangle_{L_2(\Omega)} = \int_{\Omega} q \, dV = \int_{\Omega} \operatorname{div} \mathbf{p} \, dV = \int_{\partial\Omega} \mathbf{p} \cdot \mathbf{n} \, dA = 0,$$

i. e., the source q for $\operatorname{div} \mathbf{p}$ is compatible with the homogeneous Dirichlet boundary conditions for \mathbf{p} . The same arguments apply to the discrete adjoint problem (5.2b).

It is worth noticing, that (5.2b) is not well-posed for general $y = (\mathbf{y}, q)$ but requires the regularity $\Delta Y \in L_2(T)$ for all $T \in \mathcal{T}$.

5.2. A Posteriori Error Analysis. For non-discretized control we have $\mathcal{E}_u = 0$ by Theorem 3.2. Consequently, the sum of estimators $\mathcal{E}_y, \mathcal{E}_p$ for the linear problems (5.1) constitutes an estimator for the optimal control problem. Such estimators are available for standard discretizations of the Oseen problem, like the Mini-Element, or the Taylor-Hood-Element; compare for instance with [Ver89]. Up to now, estimators for the SDFEM have not been considered for the Oseen problem. We next construct estimators \mathcal{E}_y and \mathcal{E}_p and sketch the proofs of the required properties.

State equation. For ease of notation we write $\mathbf{g} = \mathbf{f} + \mathbf{u}$ and define for $T \in \mathcal{T}$ the indicators

$$\begin{aligned} \mathcal{E}_y^2(Y, \mathbf{g}; T) &:= h_T^2 \| -\Delta \mathbf{Y} + [\nabla \mathbf{Y}] \mathbf{b} + \nabla Q - \mathbf{g} \|_{L_2(T)}^2 \\ &\quad + h_T \| [\mathbf{Y}] \|_{L_2(\partial T)}^2 + \| \operatorname{div} \mathbf{Y} \|_{L_2(T)}^2. \end{aligned}$$

These indicators constitutes an efficient and reliable estimator for the SDFEM discretization of the state equation.

Proposition 5.3. *For given $\mathbf{u} \in \mathbb{U}$ let $y \in \mathbb{Y}$ be the solution to (5.1a) and $Y \in \mathbb{Y}(\mathcal{T})$ be the SDFEM approximation given by (5.2a). Then we have the global upper bound*

$$\|Y - y\|_{\mathbb{Y}}^2 \lesssim \mathcal{E}_y^2(Y, \mathbf{g}; \mathcal{T}) = \sum_{T \in \mathcal{T}} \mathcal{E}_y^2(Y, \mathbf{g}; T).$$

For any $\bar{Y} = (\bar{\mathbf{y}}, \bar{q}) \in \mathbb{Y}(\mathcal{T})$ we have the local lower bound

$$\mathcal{E}_y^2(\bar{Y}, \mathbf{g}; T) \lesssim \|\bar{Y} - y\|_{\omega_T(T)}^2 + \text{osc}_y^2(\bar{Y}, \mathbf{g}; \mathcal{N}_T(T)) \quad \forall T \in \mathcal{T}$$

with $\text{osc}_y^2(\bar{Y}, \mathbf{g}; T) = h_T^2 \|\mathbf{g} - \mathbf{g}_T\|_{L_2(T)}^2 + h_T^2 \|\mathbf{b} - \mathbf{b}_T\|_{L_2(T)}^2$, where \mathbf{g}_T and \mathbf{b}_T are suitable finite dimensional approximations of \mathbf{g} and \mathbf{b} on T .

Proof. [1] Starting point of the analysis is the residual

$$\text{Res}_y(\bar{Y}, \mathbf{g}) := \mathcal{B}[\bar{Y}, \cdot] - \langle \mathbf{g}, \cdot \rangle \in \mathbb{Y}^*,$$

for $\bar{Y} \in \mathbb{Y}(\mathcal{T})$. Well-posedness of the continuous problem (5.1a) yields for the solution y the error equivalence

$$\beta \|\bar{Y} - y\|_{\mathbb{Y}} \leq \|\text{Res}_y(\bar{Y}, \mathbf{g})\|_{\mathbb{Y}^*} \leq \|\mathcal{B}\| \|\bar{Y} - y\|_{\mathbb{Y}}$$

and it remains to estimate $\|\text{Res}_y(\bar{Y}, \mathbf{g})\|_{\mathbb{Y}^*}$.

[2] For the SDFEM solution Y we have

$$\langle \text{Res}_y(Y, \mathbf{g}), V \rangle_{\mathbb{Y}^* \times \mathbb{Y}} = -\mathcal{B}_{\text{stab}}[Y, V] + \langle \mathbf{g}, V \rangle_{\text{stab}, y} =: -\langle \text{Res}_{\text{stab}}(Y, \mathbf{g}), V \rangle_{\mathbb{Y}^* \times \mathbb{Y}}$$

for all $V \in \mathbb{Y}(\mathcal{T})$, which replaces Galerkin-orthogonality. For given $v = (\mathbf{v}, r) \in \mathbb{Y}$ we let $V = (I_{\mathcal{T}}\mathbf{v}, 0) \in \mathbb{Y}(\mathcal{T})$, where $I_{\mathcal{T}}$ denotes the Clément or Scott-Zhang interpolation operator into the discrete velocity space [Clé75, SZ90]. This in turn yields

$$\begin{aligned} \|\text{Res}_y(Y, \mathbf{g})\|_{\mathbb{Y}^*} &\leq \sup_{\substack{v \in \mathbb{Y} \\ \|v\|_{\mathbb{Y}}=1}} \langle \text{Res}_y(Y, \mathbf{g}), v - V \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \\ &\quad + C \sup_{\substack{W = (\mathbf{W}, 0) \in \mathbb{Y}(\mathcal{T}) \\ \|(\mathbf{W}, 0)\|_{\mathbb{Y}}=1}} \langle \text{Res}_{\text{stab}}(Y, \mathbf{g}), W \rangle, \end{aligned}$$

thanks to $\|V\|_{\mathbb{Y}} \leq C\|v\|_{\mathbb{Y}}$.

[3] For Stokes, i. e., $\mathbf{b} \equiv 0$, Verfürth has shown

$$\sup_{\|v\|_{\mathbb{Y}}=1} \langle \text{Res}_y(Y, \mathbf{g}), v - V \rangle_{\mathbb{Y}^* \times \mathbb{Y}} \lesssim \mathcal{E}_y(Y, \mathbf{g}; \mathcal{T})$$

where the constant hidden in ' \lesssim ' only depends on properties of $I_{\mathcal{T}}$ [Ver89, Theorem 3.1]. The generalization of this bound to Oseen, i. e., $\mathbf{b} \neq 0$, is straight forward. For the stabilization term we easily deduce for $W = (\mathbf{W}, 0)$ by the definition of \mathcal{E}_y and $\|\cdot\|_{\text{stab}}$ the bound

$$\begin{aligned} \langle \text{Res}_{\text{stab}}(Y, \mathbf{g}), W \rangle &= \delta \sum_{T \in \mathcal{T}} h_T^2 \int_T [\nabla \mathbf{W}] \mathbf{b} \cdot (-\Delta \mathbf{Y} + [\nabla \mathbf{Y}] \mathbf{b} + \nabla Q - \mathbf{g}) dV \\ &\quad + \mu \int_{\Omega} \text{div } \mathbf{W} \text{div } \mathbf{Y} dV \\ &\leq \sqrt{\max\{\delta, \mu\}} \mathcal{E}_y(Y, \mathbf{g}; \mathcal{T}) (\|W\|_{\text{stab}} + \|\text{div } \mathbf{W}\|_{L_2(\Omega)}). \end{aligned}$$

Since $\|W\|_{\text{stab}} + \|\text{div } \mathbf{W}\|_{L_2(\Omega)} \lesssim \|W\|_{\mathbb{Y}}$ we have shown the upper bound.

[4] The lower bound is the straight forward generalization of the lower bound in [Ver89, Theorem 3.1] from Stokes to Oseen. \square

Adjoint equation. We define for $T \in \mathcal{T}$ the indicators

$$\begin{aligned} \mathcal{E}_p^2(P, \psi'(Y); T) &:= h_T^2 \|\Delta(\mathbf{P} - \mathbf{Y}) - [\nabla \mathbf{P}] \mathbf{b} + \nabla S\|_{L_2(T)}^2 \\ &\quad + h_T \|\llbracket \mathbf{P} - \mathbf{Y} \rrbracket\|_{L_2(\partial T)}^2 + \|\text{div } \mathbf{P} - \mathbf{Q}\|_{L_2(T)}^2, \end{aligned}$$

which constitute an efficient and reliable estimator for the SDFEM discretization of the adjoint equation.

Proposition 5.4. *For given $Y \in \mathbb{Y}(\mathcal{T})$ let $p \in \mathbb{Y}$ be the solution to (5.1b) and $P \in \mathbb{Y}(\mathcal{T})$ be the SDFEM approximation given by (5.2b). Then we have the global upper bound*

$$\|P - p\|_{\mathbb{Y}}^2 \lesssim \mathcal{E}_p^2(P, \psi'(Y); \mathcal{T}) = \sum_{T \in \mathcal{T}} \mathcal{E}_p^2(P, \psi'(Y); T).$$

For any $\bar{P} = (\bar{\mathbf{P}}, \bar{S}) \in \mathbb{Y}(\mathcal{T})$ we have the local lower bound

$$\mathcal{E}_p^2(\bar{P}, \psi'(Y); T) \lesssim \|\bar{P} - p\|_{\omega_T(T)}^2 + \text{osc}_p^2(\bar{P}, \psi'(Y); \mathcal{N}_T(T)) \quad \forall T \in \mathcal{T}$$

with $\text{osc}_p^2(\bar{P}, \psi'(Y); T) = h_T^2 \|\mathbf{b} - \mathbf{b}_T\|_{L_2(T)}^2$.

Proof. Follows the proof to Proposition 5.3 based on the residual of the adjoint problem $\text{Res}_p(\bar{P}, \psi'(Y)) := \mathcal{B}[\cdot, \bar{P}] - \langle \psi'(Y), \cdot \rangle \in \mathbb{Y}^*$. \square

Remark 5.5 (Control Error). If $(\hat{U}, \hat{Y}, \hat{P})$ is the discrete solution of the optimal control problem (3.2), we know from Theorem 3.2 that

$$\mathcal{E}_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; \mathcal{T}) = \mathcal{E}_y(\hat{Y}, \mathbf{f} + \hat{U}; \mathcal{T}) + \mathcal{E}_p(\hat{P}, \psi'(\hat{Y}); \mathcal{T})$$

is a reliable and locally efficient estimator for (2.1). For non-discretized control there is no explicit contribution for the control error $\|\hat{U} - \hat{u}\|_{L_2(\Omega)}$. However, the total error is influenced by the distance of \hat{U} to \hat{u} .

Considering the case $\mathbf{f} \equiv 0$ we see that the oscillation term osc_y in the lower bound becomes $h_T \|\hat{u} - \hat{u}_T\|_{L_2(T)}$ with a suitable finite dimensional approximation \hat{u}_T of \hat{u} . This term encodes approximability of \hat{u} by discrete functions over \mathcal{T} in H^{-1} . Consequently, approximability of \hat{u} enters implicitly in the estimator \mathcal{E}_{ocp} by the argument \hat{U} of \mathcal{E}_y .

Remark 5.6. In [HYZ09] an a posteriori error estimator for a scalar advection diffusion problem with a simpler objective ψ is derived. Comparing our proof for the more complex Oseen problem to the proof in [HYZ09] one realizes how much the abstract framework of §2 and §3 simplifies the a posteriori error analysis, in particular for non-standard discretizations.



FIGURE 5.1. Domains and initial triangulation for Experiment 1 (left) and Experiment 2 (right). The domain Γ for the boundary control is indicated by bold lines.

6. NUMERICAL EXPERIMENTS

We conclude the article by two numerical experiments in order to examine the quality of the derived estimators and the performance of the standard adaptive loop (1.1). For this we resort to the diffusion-reaction problem with discretized boundary control from §4 in two space dimensions. We use piecewise linear finite elements for the state/adjoint state, and piecewise constant elements for the control. The cost parameter α is set to 1. We have implemented an adaptive solver for the optimal control problem in the adaptive finite element toolbox ALBERTA [SS05, SSH⁺].

The nonlinear system has been solved by a variant of the primal-dual active set strategy [HIK02].

6.1. Experiment 1: The Residual Estimator and Marking Strategies. This experiment is designed to investigate the quality of the residual estimator of Example 4.1 and the performance of the adaptive method with different marking strategies in case of a singular solution.

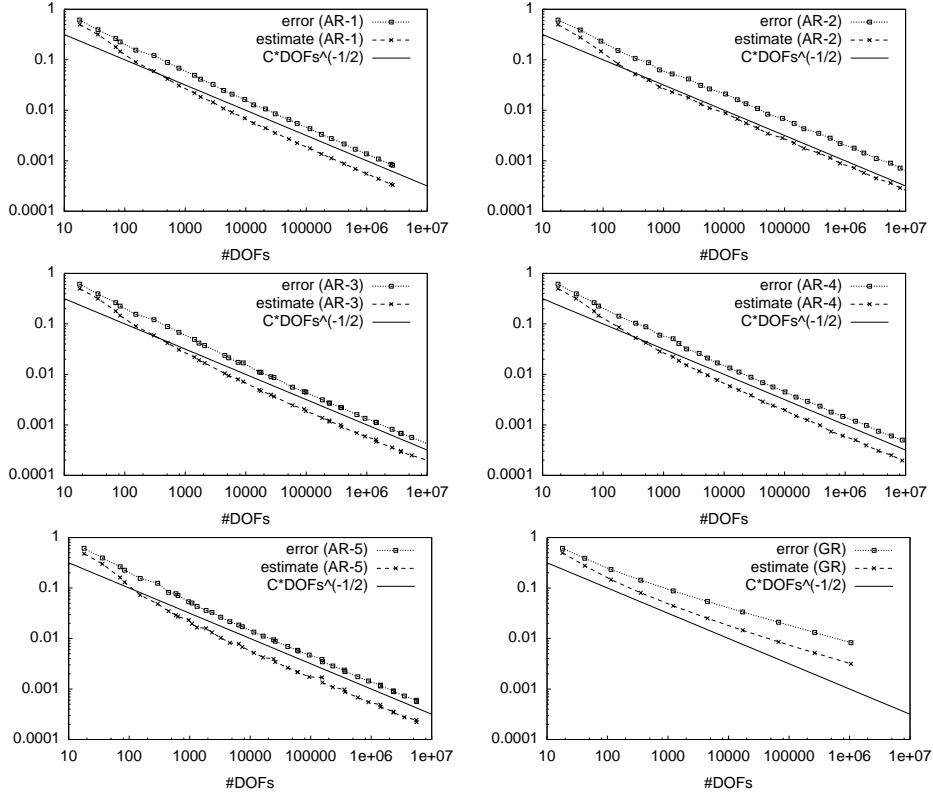


FIGURE 6.1. Experiment 1: Error and Estimator vs DOFs for the various strategies. The optimal adaptive decay rate $\#DOFs^{-1/2}$ is indicated by the solid line.

Data of the problem and true solution. We use the example considered in [MR11] on the L-shaped domain, which is given as $\Omega = (-1, 1)^2 \setminus ([0, 1] \times (-1, 0])$ with $\Gamma = \partial\Omega$; compare with Fig. 5.1. We use constant box-constraints $a = -0.5$, $b = +0.5$ and denote for $s \in \mathbb{R}$ by $\text{Proj}_a^b(s)$ the best-approximation of s in $[a, b]$. Remaining data of the problem is chosen as follows (with x from the respective domain):

$$\begin{aligned} f_1(x) &:= -\text{Proj}_{-0.5}^{+0.5} \left(-|x|^{\frac{2}{3}} \cos\left(\frac{2}{3}\phi(x)\right) \right), \\ f_2(x) &:= 0, \\ y_d(x) &:= -|x|^{\frac{2}{3}} \cos\left(\frac{2}{3}\phi(x)\right), \\ g(x) &:= \frac{2}{3}|x|^{-\frac{4}{3}} \left(-\cos\left(\frac{2}{3}\phi(x)\right) \cdot x + x^\perp \sin\left(\frac{2}{3}\phi(x)\right) \right) \cdot n_\Gamma(x). \end{aligned}$$

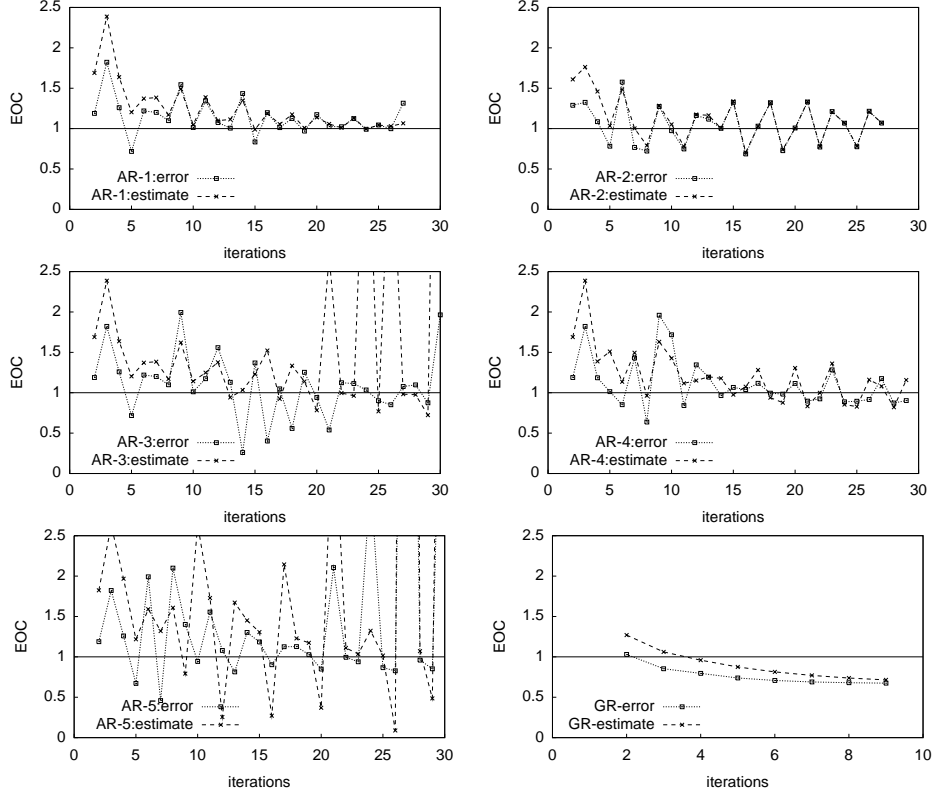


FIGURE 6.2. Experiment 1: Experimental order of convergence for the various strategies. The optimal adaptive order 1 is indicated by the solid line.

Here, $\phi(x) = \arccos(|x|^{-1}x \cdot e_1)$ is the angle between x and e_1 , $x^\perp = [-x_2, x_1]^T$, and n_Γ is the outer unit normal vector to Ω .

Data is chosen such that the true adjoint state has the typical singularity at the reentrant corner of Ω . The true solution is given by

$$\hat{y}(x) = 0, \quad \hat{p}(x) = |x|^{\frac{2}{3}} \cos(\frac{2}{3}\phi(x)), \quad \hat{u}(x) = \text{Proj}_{-0.5}^{+0.5} \left(-|x|^{\frac{2}{3}} \cos(\frac{2}{3}\phi(x)) \right).$$

Marking strategies. We utilize the residual estimator from Example 4.1. For ease of notation we omit the discrete solution $(\hat{U}, \hat{Y}, \hat{P})$ as arguments of the indicators and write $\mathcal{E}_u(T)$, $\mathcal{E}_y(T)$, and $\mathcal{E}_p(T)$ instead of $\mathcal{E}_u(\hat{U}, \hat{P}; T)$, $\mathcal{E}_y(\hat{Y}, \hat{U} + f; T)$, and $\mathcal{E}_p(\hat{P}, \psi'(\hat{Y}); T)$. We then use $\mathcal{E}_{\text{ocp}}^2(T) = \mathcal{E}_u^2(T) + \mathcal{E}_y^2(T) + \mathcal{E}_p^2(T)$.

We observe that the different contributions \mathcal{E}_u , \mathcal{E}_y , and \mathcal{E}_p of \mathcal{E}_{ocp} often differ by orders of magnitudes. To investigate the effects of the different orders we designed and used several marking strategies. Basis of all marking strategies is the *maximum strategy*: Given a parameter $\theta \in [0, 1]$ and a set of indicators $\{\mathcal{E}(T)\}_{T \in \mathcal{T}}$ the maximum strategy $\text{MS}(\{\mathcal{E}(T)\}_{T \in \mathcal{T}})$ outputs a subset $\mathcal{M} \subset \mathcal{T}$ of marked elements such

$$\mathcal{M} = \{T \in \mathcal{T} \mid \mathcal{E}(T) \geq \theta \mathcal{E}_{\max}\} \quad \text{with} \quad \mathcal{E}_{\max} := \max\{\mathcal{E}(T) \mid T \in \mathcal{T}\}.$$

Two strategies involve all three contributions.

(AR-1) We apply the maximum strategy to \mathcal{E}_{ocp} , i. e.,

$$\mathcal{M} = \text{MS}(\{\mathcal{E}_{\text{ocp}}(T)\}_{T \in \mathcal{T}}).$$

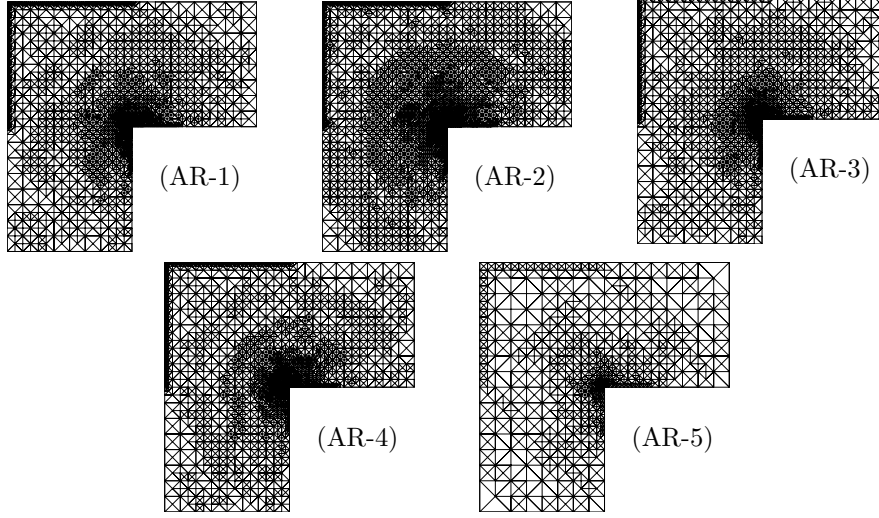


FIGURE 6.3. Experiment 1: Adaptive grids of iteration 12 for the various strategies. The approximation of the singularity of p and the approximation of u is well reflected. It is important to note that the accuracy of the diverse strategies on these grids is quite different.

(AR-2) Separately apply the maximum-strategy to all contributions, i. e.,

$$\mathcal{M} = \text{MS}(\{\mathcal{E}_u(T)\}_{T \in \mathcal{T}}) \cup \text{MS}(\{\mathcal{E}_y(T)\}_{T \in \mathcal{T}}) \cup \text{MS}(\{\mathcal{E}_p(T)\}_{T \in \mathcal{T}}).$$

The other strategies adaptively select particular contributions from $\{\mathcal{E}_u, \mathcal{E}_y, \mathcal{E}_p\}$ depending on their relative size for marking: For given parameter $\delta \in [0, 1]$ determine the index set

$$I_\delta := \{i \in \{u, y, p\} \mid \mathcal{E}_i(\mathcal{T}) \geq \delta \max\{\mathcal{E}_u(\mathcal{T}), \mathcal{E}_y(\mathcal{T}), \mathcal{E}_p(\mathcal{T})\}\}.$$

We then proceed as follows.

(AR-3) For $T \in \mathcal{T}$ set $\mathcal{E}^2(T) := \sum_{i \in I_\delta} \mathcal{E}_i^2(T)$ and apply the maximum strategy to \mathcal{E} , i. e.,

$$\mathcal{M} = \text{MS}(\{\mathcal{E}(T)\}_{T \in \mathcal{T}}).$$

(AR-4) Separately apply the maximum strategy to \mathcal{E}_i , $i \in I_\delta$, i. e.,

$$\mathcal{M} = \bigcup_{i \in I_\delta} \text{MS}(\{\mathcal{E}_i(T)\}_{T \in \mathcal{T}}).$$

(AR-5) Apply the maximum strategy to the largest contribution only, i. e., use strategy (AR-4) with $\delta = 1$. In general, $\#I_1 = 1$.

In all experiments we have used the parameters $\theta = 0.5$ for MS, and $\delta = 0.5$ for (AR-3) and (AR-4). We also compared adaptive refinement to uniform refinement, which we denote by (GR).

Evaluation of the experiments. We first see from Fig. 6.1 that all adaptive strategies show an optimal performance in terms of degrees of freedom (DOFs). This is, the error and estimator decay with rate $\#\text{DOFs}^{-1/2}$, which is the best possible adaptive rate for the H^1 -error with linear finite elements. Only uniform refinement leads to the lower rate $\#\text{DOFs}^{-1/3}$ as predicted by theory. In the log-log plots of Fig. 6.1 error and estimator show the same decay.

We next have a closer look at the experimental order of convergence (EOC). For adaptive refinement we expect an EOC 1 and for uniform refinement an EOC $\frac{2}{3}$.

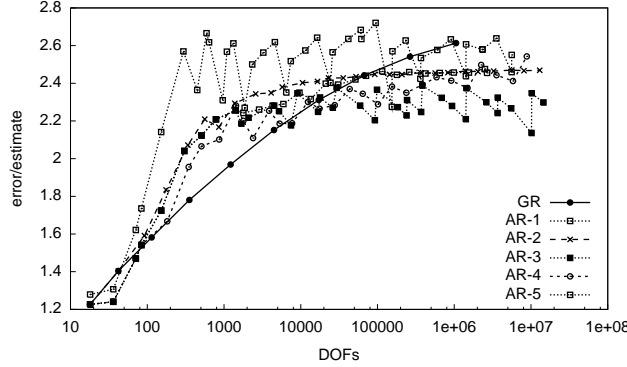


FIGURE 6.4. Experiment 1: ratio error/estimator. The more the strategy adaptively accounts for the different sizes of the contributions \mathcal{E}_u , \mathcal{E}_y , and \mathcal{E}_p the more the EOC oscillates.

This is confirmed by the plots in Fig. 6.2. Yet we have the following surprising observation: The more the strategy adaptively accounts for the different sizes of the contributions \mathcal{E}_u , \mathcal{E}_y , and \mathcal{E}_p the more the EOC oscillates: For (AR-1), (AR-2), and (GR) the EOC behaves nicely, whereas for (AR-3), (AR-4), and (AR-5) we see strong oscillations.

Even more astonishing is the influence of the marking strategy on the ratio of error and estimator, which is depicted in Fig. 6.4. We first observe that the estimator has the tendency to underestimate the true error in early stages. Some minimal resolution seems to be necessary so that the ratio becomes constant. Strategies (AR-1) and (AR-2) cope quickly with this. Uniform refinement (GR) is not able to produce the necessary resolution. The adaptive selection of contributions by (AR-3), (AR-4), and (AR-5) leads to an oscillatory behavior of the ratio.

In Fig. 6.3 we present the adaptive grids of iteration 12 with the different strategies. We see the appropriate refinement accounting on the one hand for the singularity of the adjoint state at the reentrant corner and on the other hand for the approximation of the control in the inactive set

$$I = \{0\} \times [-\sqrt{1/8}, 0] \cup [0, \sqrt{1/8}] \times \{0\} \cup \{-1\} \times [0, 1] \cup [-1, 0] \times \{1\}.$$

The error and estimator in this iteration varies for the different strategies. The true error is approximately 0.0208 for (AR-1), 0.0162 for (AR-2), 0.0212 for (AR-3), 0.0211 for (AR-4) and 0.0361 for (AR-5).

Summarizing the results of the experiments, we find that the adaptive strategies (AR-1) and (AR-2) perform well. On the contrary, strategies (AR-3), (AR-4), and (AR-5) show effects that can turn out to be disadvantageous in applications.

6.2. Experiment 2: Hierarchical Estimator and Active/Inactive Sets.

With this experiment we examine the quality of the hierarchical estimator from Example 4.2 and the sensitivity of the adaptive method with respect to changes of the active and inactive sets. Relying on the experience of §6.1 we only use marking strategy (AR-1).

Data of the problem and true solution. We let $\Omega = [0, 3]^2$ and consider a boundary control supported on $\Gamma = \{0\} \times [1, 2] \subsetneq \partial\Omega$; compare with Fig. 5.1. We use constant box-constraints $a, b = 20$, where we vary in the lower bound a . Remaining data of the problem is chosen as follows (with $Z = 10$, $n = 20$, and x

from the respective domain):

$$\begin{aligned} f_1(x) &:= -\text{Proj}_a^{20} \left(\frac{Z}{2n} \left((2n+1) \left(\frac{2}{3}x_2 - 1 \right) - \left(\frac{2}{3}x_2 - 1 \right)^{2n+1} \right) \right) \\ f_2(x) &:= e^{-10|x|^2} (41 - 1600|x|^2), \\ g(x) &:= 0, \\ y_a(x) &:= e^{-10|x|^2} - Z \left(\frac{2n+1}{2n} \left(\frac{2}{3}x_2 - 1 \right) + \frac{8n+4}{9} \left(\frac{2}{3}x_2 - 1 \right)^{2n-1} - \left(\frac{2}{3}x_2 - 1 \right)^{2n+1} \right). \end{aligned}$$

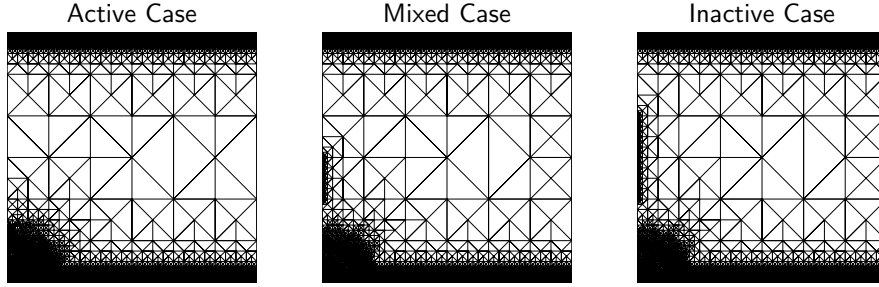


FIGURE 6.5. Experiment 2: Adaptive grids after 10 iterations. The resolution perfectly matches the different features of $(\hat{u}, \hat{y}, \hat{p})$.

The true solution is

$$\begin{aligned} \hat{u}(x) &= \text{Proj}_a^{20} \left(\frac{Z}{2n} \left((2n+1) \left(\frac{2}{3}x_2 - 1 \right) - \left(\frac{2}{3}x_2 - 1 \right)^{2n+1} \right) \right), \\ \hat{y}(x) &= e^{-10|x|^2}, \\ \hat{p}(x) &= \frac{Z}{2n} \left((2n+1) \left(\frac{2}{3}x_2 - 1 \right) - \left(\frac{2}{3}x_2 - 1 \right)^{2n+1} \right) \end{aligned}$$

and has the following particular features. An adaptive approximation of the respective components $\hat{u}, \hat{y}, \hat{p}$ requires local refinement in different regions of the domain. The true state \hat{y} needs a high resolution close to the origin due to shape of a narrow exponential peak. The adjoint state \hat{p} is constant in x_1 . In x_2 it is almost linear for $x_2 \in [0.2, 2.8]$ with a sharp boundary layer close to $x_2 = 0$ and $x_2 = 3$, which calls for refinement. Properties of the control \hat{u} vary with the parameter a , where we consider three cases:

Inactive Case: For $a = -20$ we have $a \leq -\hat{p} \equiv \hat{u} \leq b$ on Γ .

Active Case: For $a = 15$ we have $-\hat{p} \leq b$ and $\hat{u} \equiv a$ on Γ .

Mixed Case: For $a = 0$ we have $a \leq -\hat{p} = \hat{u} \leq b$ for $x_2 \in [1, 1.5]$, and $-\hat{p} \leq a$ and $\hat{u} \equiv a$ for $x_2 \in [1.5, 2]$.

Approximation of \hat{u} only requires refinement in the inactive set I , which is $I = \emptyset$ in the Active Case, $I = \{0\} \times [1, 1.5]$ in the Mixed Case, and $I = \Gamma$ in the Inactive Case.

Evaluation of the experiments. In Fig. 6.5 we have depicted the adaptive grids of the 10th iteration for the three different cases. The resolution for \hat{y} and \hat{p} is as we expected it and is nearly the same in all scenarios. The refinement for \hat{u} is confined to the true inactive set. This indicates that the estimator is not sensitive with respect to changes in the active/inactive sets.

This is also confirmed by Figs. 6.6 and 6.7. In Fig. 6.6 we have depicted the decay of error and estimator vs. #DOFs for adaptive and uniform refinement. Both, error and estimator decay with the optimal rate $\#DOFs^{-1/2}$ and the lines look similar for all three cases. Here, uniform refinement also exhibits the optimal

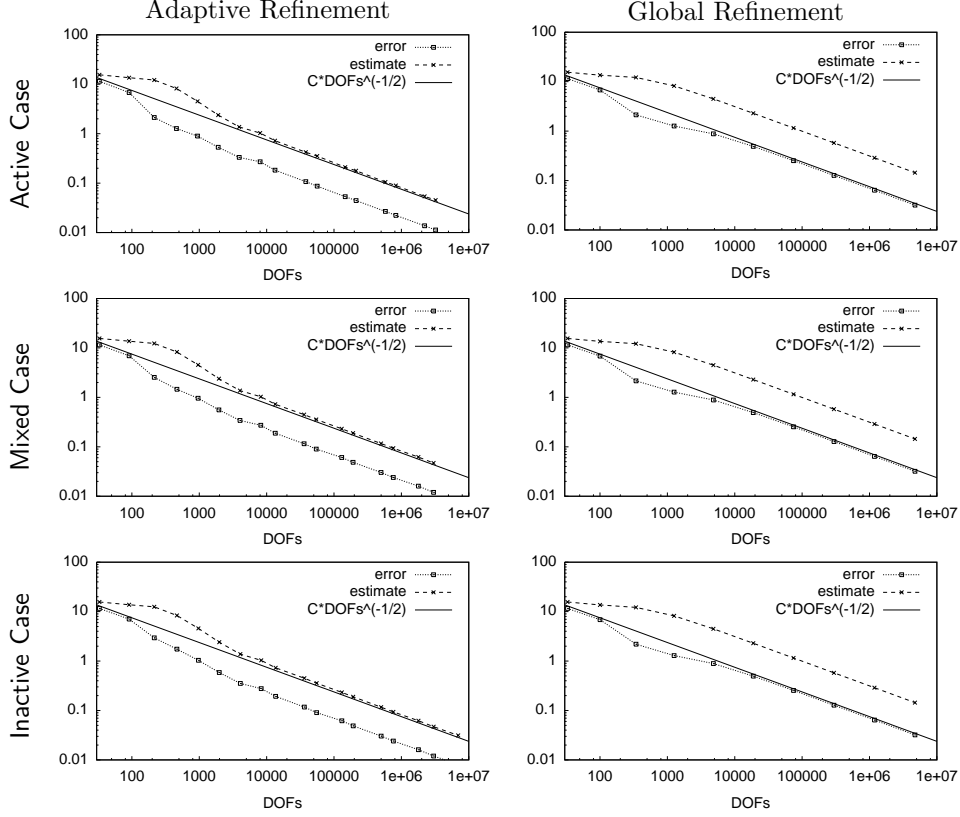


FIGURE 6.6. Experiment 2: Decay of error and estimator vs $\#DOFs$ for the three different cases for adaptive and uniform refinement. The behavior is not sensitive with respect to changes in the active/inactive sets.

decay $\#DOFs^{-1/2}$ since the true solution $(\hat{u}, \hat{y}, \hat{p})$ is regular. Variations of the error/estimator ratios plotted in Fig. 6.7 are small when changing the active/inactive sets. We observe that the estimator underestimates the true error (by an almost constant factor). This is typical for the hierarchical estimator.

We close this discussion with a remark on further estimates of the indicator \mathcal{E}_u .

Remark 6.1. In all our experiments we have used $\mathcal{E}_u(\hat{U}, \hat{P}; \mathcal{T}) = \|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}}$ as indicator for the control error. The exact computation of \mathcal{E}_u is involved since $(\Pi(\hat{P}))|_{T \cap \Gamma}$ is a piecewise affine function over $T \cap \Gamma$ rather than an affine function on $T \cap \Gamma$. None the less, using a suitable decomposition of $T \cap \Gamma$ the indicator $\mathcal{E}(\hat{U}, \hat{P}; T)$ can be computed exactly. Moreover, Theorem 3.2 states that this term is globally efficient.

In order to simplify computations, the term \mathcal{E}_u is often estimated further. Liu and Yan propose in [LY01] to use the bound $\|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}} \leq \|\hat{U} + \frac{1}{\alpha} \hat{P}\|_{\mathbb{U}}$. Denoting by $M_{\mathcal{T}}: \mathbb{U} \rightarrow \mathbb{U}(\mathcal{G})$ the best-approximation in L_2 , Hintermüller et al. suggest in [HHIK08] to use the estimate $\|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}} \lesssim \|\hat{P} - M_{\mathcal{T}} \hat{P}\|_{\mathbb{Y}}$.

Using one of these bounds as indicator for the control error does not lead to a localization of refinement to the inactive set. In fact, we always observe a refinement of whole Γ in all three cases. This shows, that neither $\|\hat{U} + \frac{1}{\alpha} \hat{P}\|_{\mathbb{U}}$ nor $\|\hat{P} - M_{\mathcal{T}} \hat{P}\|_{\mathbb{Y}}$ can be efficient for our error notion $\|\hat{U} - \hat{u}\|_{\mathbb{U}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}}$. Efficiency of the estimator is a prerequisite for an optimal performance of an adaptive method;

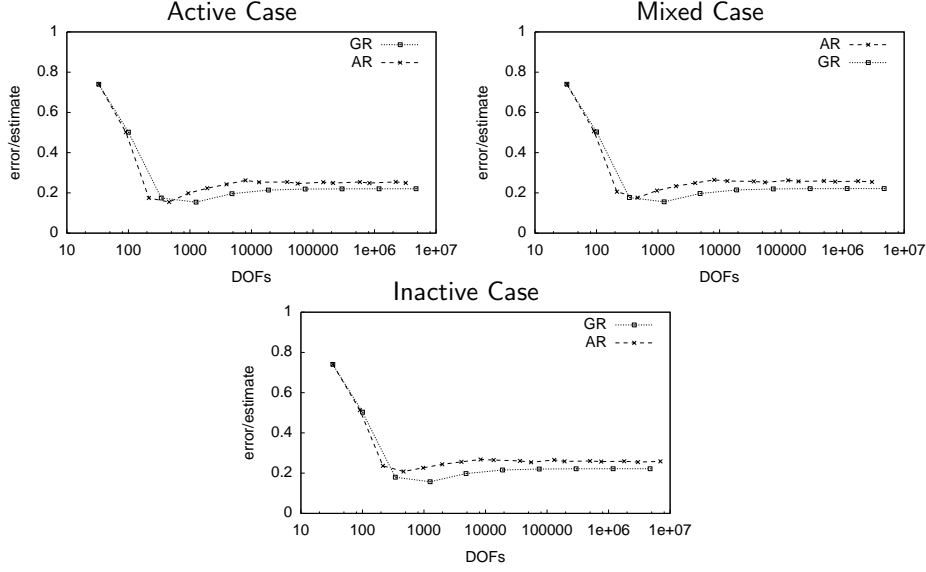


FIGURE 6.7. Experiment 2: Ratio of error/estimator vs #DOFs for the three different cases. The behavior is not sensitive with respect to changes in the active/inactive sets.

compare with [CKNS08, §5]. We therefore conclude that the additional effort for computing $\|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}}$ exactly pays off.

Efficiency for $\|\hat{U} + \frac{1}{\alpha}\hat{P}\|_{\mathbb{U}}$ or $\|\hat{P} - M_{\mathcal{T}}\hat{P}\|_{\mathbb{Y}}$ can be re-established by introducing slack-variables for the control constraints and including these slack-variables in the error notion. Even though the emanating estimator is then efficient with respect to this new error notion, the refinement of whole Γ will persist in all three cases.

REFERENCES

- [AO00] Mark Ainsworth and J. Tinsley Oden, *A posteriori error estimation in finite element analysis*, Wiley-Interscience, New York, 2000.
- [Bab71] Ivo Babuška, *Error-bounds for finite element method*, Numer. Math. **16** (1971), 322–333.
- [BBM⁺07] R. Becker, M. Braack, D. Meidner, R. Rannacher, and B. Vexler, *Adaptive finite element methods for PDE-constrained optimal control problems*, Reactive flows, diffusion and transport, Springer, Berlin, 2007, pp. 177–205.
- [BIK99] Maïtine Bergounioux, Kazufumi Ito, and Karl Kunisch, *Primal-dual strategy for constrained optimal control problems*, SIAM Journal on Control and Optimization **37** (1999), no. 4, 1176–1194.
- [BPS09] Dietrich Braess, Veronika Pillwein, and Joachim Schöberl, *Equilibrated residual error estimates are p-robust*, Comput. Methods Appl. Mech. Engrg. **198** (2009), no. 13-14, 1189–1197.
- [BR78] Ivo Babuška and Werner Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal. **15** (1978), 736–754.
- [BR03] Wolfgang Bangerth and Rolf Rannacher, *Adaptive finite element methods for differential equations*, Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel, 2003.
- [Bra07] Dietrich Braess, *Finite elements*, third ed., Cambridge University Press, Cambridge, 2007, Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.
- [BS01] Ivo Babuška and Theofanis Strouboulis, *The finite element method and its reliability*, Numerical Mathematics and Scientific Computation, The Clarendon Press Oxford University Press, New York, 2001.

- [BV09] Olaf Benedix and Boris Vexler, *A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints*, Comput. Optim. Appl. **44** (2009), no. 1, 3–25.
- [CB02] C. Carstensen and S. Bartels, *Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. part I: low order conforming, nonconforming, and mixed FEM*, Math. Comp. **71** (2002), 945–969.
- [CKNS08] J. Manuel Cascón, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), no. 5, 2524–2550.
- [Clé75] Philippe Clément, *Approximation by finite element functions using local regularization*, R.A.I.R.O. **9** (1975), 77–84.
- [dlRMV08] Juan Carlos de los Reyes, Christian Meyer, and Boris Vexler, *Finite element error analysis for state-constrained optimal control of the Stokes equations*, Control Cybernet. **37** (2008), no. 2, 251–284.
- [Gal94] Giovanni P. Galdi, *An introduction to the mathematical theory of the Navier-Stokes equations. Vol. I*, Springer Tracts in Natural Philosophy, vol. 38, Springer-Verlag, New York, 1994, Linearized steady problems.
- [HH08] Michael Hintermüller and Ronald H.W. Hoppe, *Goal-oriented adaptivity in control-constrained optimal control of partial differential equations*, SIAM Journal on Control and Optimization **47** (2008), no. 4, 1721–1743.
- [HH10] Michael Hintermüller and Ronald H. W. Hoppe, *Goal oriented mesh adaptivity for mixed control-state constrained elliptic optimal control problems*, Applied and numerical partial differential equations, Comput. Methods Appl. Sci., vol. 15, Springer, New York, 2010, pp. 97–111.
- [HHIK08] Michael Hintermüller, Ronald H. W. Hoppe, Yuri Iliash, and Michael Kieweg, *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints*, ESAIM, Control Optim. Calc. Var. **14** (2008), no. 3, 540–560.
- [HIK02] Michael Hintermüller, Kazufumi Ito, and Karl Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim. **13** (2002), no. 3, 865–888 (electronic) (2003).
- [Hin05] Michael Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl. **30** (2005), no. 1, 45–61.
- [HK10] Ronald. H. W. Hoppe and Michael Kieweg, *Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems*, Comput. Optim. Appl. **46** (2010), no. 3, 511–533.
- [HV09] Michael Hinze and Morten Vierling, *Variational discretization and semi-smooth Newton methods; implementation, convergence and globalization in PDE constrained optimization with control constraints*, Hamburger Beiträge zur Angewandten Mathematik No. 2009-13, 2009.
- [HYZ09] Michael Hinze, Ningning Yan, and Zhaojie Zhou, *Variational discretization for optimal control governed by convection dominated diffusion equations*, J. Comput. Math. **27** (2009), no. 2-3, 237–253.
- [Koh13] Kristina Kohls, *An adaptive finite element method for control-constrained optimal control problems*, Ph.D. thesis, Fachbereich Mathematik, Universität Stuttgart, 2013.
- [KRS12] Kristina Kohls, Arnd Rösch, and Kunibert G. Siebert, *A posteriori error estimators for control constrained optimal control problems*, Constrained Optimization and Optimal Control for Partial Differential Equations (Leugering et al., ed.), International Series of Numerical Mathematics 160, Springer, 2012, pp. 431–443.
- [KS11] Christian Kreuzer and Kunibert G. Siebert, *Decay rates of adaptive finite elements with Dörfler Marking*, Numer. Math. **117** (2011), no. 4, 679–716.
- [KR02] Karl Kunisch and Arnd Rösch, *Primal-dual active set strategy for a general class of constrained optimal control problems*, SIAM Journal Optimization **13** (2002), no. 2, 321–334.
- [Lio71] Jacques L. Lions, *Optimal control of systems governed by partial differential equations*, Die Grundlehren der mathematischen Wissenschaften, vol. 170, Springer-Verlag, 1971.
- [LLY07] Ruo Li, Wenbin Liu, and Ningning Yan, *A posteriori error estimates of recovery type for distributed convex optimal control problems*, Journal of Scientific Computing **33** (2007), no. 2, 155–182.
- [LY01] W. Liu and N. Yan, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math **15** (2001), no. 1-4, 285–309.

- [LY02] W. Liu and N. Yan, *A posteriori error estimates for control problems governed by Stokes equations*, SIAM J. Numer. Anal. **40** (2002), no. 5, 1850–1869.
- [MNS03] Pedro Morin, Ricardo H. Nochetto, and Kunibert G. Siebert, *Local problems on stars: A posteriori error estimators, convergence, and performance*, Math. Comp. **72** (2003), 1067–1097.
- [MR11] Mariano Mateos and Arnd Rösch, *On saturation effects in the Neumann boundary control of elliptic optimal control problems*, Comput. Optim. Appl. **49** (2011), no. 2, 359–378.
- [MSV08] Pedro Morin, Kunibert G. Siebert, and Andreas Veese, *A basic convergence result for conforming adaptive finite elements*, Math. Models Methods Appl. **18** (2008), 707–737.
- [Neč62] J. Nečas, *Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle*, Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser. **16** (1962), 305–326 (French).
- [NSV09] Ricardo H. Nochetto, Kunibert G. Siebert, and Andreas Veese, *Theory of adaptive finite element methods: An introduction*, Multiscale, Nonlinear and Adaptive Approximation (Ronald A. DeVore and Angela Kunoth, eds.), Springer, 2009, pp. 409–542.
- [RST08] Hans-Görg Roos, Martin Stynes, and Lutz Tobiska, *Robust numerical methods for singularly perturbed differential equations*, second ed., Springer Series in Computational Mathematics, vol. 24, Springer-Verlag, Berlin, 2008, Convection-diffusion-reaction and flow problems.
- [Sie11] Kunibert G. Siebert, *A convergence proof for adaptive finite elements without lower bound*, IMA J. of Numer. Anal. **31** (2011), no. 3, 947–970.
- [SS05] Alfred Schmidt and Kunibert G. Siebert, *Design of adaptive finite element software. The finite element toolbox ALBERTA*, Lecture Notes in Computational Science and Engineering 42. Springer, Berlin, 2005.
- [SSH⁺] Alfred Schmidt, Kunibert G. Siebert, Claus-Justus Heine, Daniel Köster, and Oliver Kriessl, *ALBERTA: An adaptive hierarchical finite element toolbox*, <http://www.alberta-fem.de/>, Version 1.2 and 2.0.
- [SV07] Kunibert G. Siebert and Andreas Veese, *A unilaterally constrained quadratic minimization with adaptive finite elements*, SIAM J. Optim. **18** (2007), no. 1, 260–289.
- [SZ90] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Mathematics of Computation **54** (1990), no. 190, 483–493.
- [Trö10] Fredi Tröltzsch, *Optimal control of partial differential equations. theory, methods and applications*, Graduate Studies Series in Mathematics, vol. 112, AMS, Providence, 2010.
- [Vee02] Andreas Veese, *Convergent adaptive finite elements for the nonlinear Laplacian*, Numer. Math. **92** (2002), no. 4, 743–770.
- [Ver89] Rüdiger Verfürth, *A posteriori error estimators for the Stokes equations*, Numer. Math. **55** (1989), 309–325.
- [Ver96] Rüdiger Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Adv. Numer. Math., John Wiley, Chichester, UK, 1996.
- [VV09] Andreas Veese and Rüdiger Verfürth, *Explicit upper bounds for dual norms of residuals*, to appear in SIAM J. Numer. Anal., 2009.
- [VW08] Boris Vexler and Winnifred Wollner, *Adaptive finite elements for elliptic optimization problems with control constraints*, SIAM J. Control Optim. **47** (2008), no. 1, 509–534.
- [YZ08] Ningning Yan and Zhaojie 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** (2008), no. 3, 297–320.
- [ZZ87] O. C. Zienkiewicz and J. Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis*, Int. J. Numer. Methods Eng. **24** (1987), 337–357.

KRISTINA KOHLS, INSTITUT FÜR ANGEWANDTE ANALYSIS UND NUMERISCHE SIMULATION, FACHBEREICH MATHEMATIK, UNIVERSITÄT STUTTGART, PFAFFENWALDRING 57, D-70569 STUTTGART, GERMANY

URL: www.ians.uni-stuttgart.de/nmh/

E-mail address: kristina.kohls@ians.uni-stuttgart.de

ARND RÖSCH, FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT DUISBURG-ESSEN, FORSTHAUSWEG 2, D-47057 DUISBURG, GERMANY

URL: www.uni-due.de/mathematik/agroesch/

E-mail address: arnd.roesch@uni-due.de

KUNIBERT G. SIEBERT, INSTITUT FÜR ANGEWANDTE ANALYSIS UND NUMERISCHE SIMULATION, FACHBEREICH MATHEMATIK, UNIVERSITÄT STUTTGART, PFAFFENWALDRING 57, D-70569 STUTTGART, GERMANY

URL: www.ians.uni-stuttgart.de/nmh/

E-mail address: kg.siebert@ians.uni-stuttgart.de

Erschienenene Preprints ab Nummer 2007/2007-001

Komplette Liste: <http://www.mathematik.uni-stuttgart.de/preprints>

- 2013-001 *Kohls, K.; Rösch, A.; Siebert, K.:* A Posteriori Error Analysis of Optimal Control Problems with Control Constraints
- 2012-018 *Kimmerle, W.; Konovalov, A.:* On the Prime Graph of the Unit Group of Integral Group Rings of Finite Groups II
- 2012-017 *Stroppel, B.; Stroppel, M.:* Desargues, Doily, Dualities, and Exceptional Isomorphisms
- 2012-016 *Moroianu, A.; Pilca, M.; Semmelmann, U.:* Homogeneous almost quaternion-Hermitian manifolds
- 2012-015 *Steinke, G.F.; Stroppel, M.J.:* Simple groups acting two-transitively on the set of generators of a finite elation Laguerre plane
- 2012-014 *Steinke, G.F.; Stroppel, M.J.:* Finite elation Laguerre planes admitting a two-transitive group on their set of generators
- 2012-013 *Diaz Ramos, J.C.; Dominguez Vázquez, M.; Kollross, A.:* Polar actions on complex hyperbolic spaces
- 2012-012 *Moroianu, A.; Semmelmann, U.:* Weakly complex homogeneous spaces
- 2012-011 *Moroianu, A.; Semmelmann, U.:* Invariant four-forms and symmetric pairs
- 2012-010 *Hamilton, M.J.D.:* The closure of the symplectic cone of elliptic surfaces
- 2012-009 *Hamilton, M.J.D.:* Iterated fibre sums of algebraic Lefschetz fibrations
- 2012-008 *Hamilton, M.J.D.:* The minimal genus problem for elliptic surfaces
- 2012-007 *Ferrario, P.:* Partitioning estimation of local variance based on nearest neighbors under censoring
- 2012-006 *Stroppel, M.:* Buttons, Holes and Loops of String: Lacing the Doily
- 2012-005 *Hantsch, F.:* Existence of Minimizers in Restricted Hartree-Fock Theory
- 2012-004 *Grundhöfer, T.; Stroppel, M.; Van Maldeghem, H.:* Unitals admitting all translations
- 2012-003 *Hamilton, M.J.D.:* Representing homology classes by symplectic surfaces
- 2012-002 *Hamilton, M.J.D.:* On certain exotic 4-manifolds of Akhmedov and Park
- 2012-001 *Jentsch, T.:* Parallel submanifolds of the real 2-Grassmannian
- 2011-028 *Spreer, J.:* Combinatorial 3-manifolds with cyclic automorphism group
- 2011-027 *Griesemer, M.; Hantsch, F.; Wellig, D.:* On the Magnetic Pekar Functional and the Existence of Bipolarons
- 2011-026 *Müller, S.:* Bootstrapping for Bandwidth Selection in Functional Data Regression
- 2011-025 *Felber, T.; Jones, D.; Kohler, M.; Walk, H.:* Weakly universally consistent static forecasting of stationary and ergodic time series via local averaging and least squares estimates
- 2011-024 *Jones, D.; Kohler, M.; Walk, H.:* Weakly universally consistent forecasting of stationary and ergodic time series
- 2011-023 *Györfi, L.; Walk, H.:* Strongly consistent nonparametric tests of conditional independence
- 2011-022 *Ferrario, P.G.; Walk, H.:* Nonparametric partitioning estimation of residual and local variance based on first and second nearest neighbors
- 2011-021 *Eberts, M.; Steinwart, I.:* Optimal regression rates for SVMs using Gaussian kernels
- 2011-020 *Frank, R.L.; Geisinger, L.:* Refined Semiclassical Asymptotics for Fractional Powers of the Laplace Operator

- 2011-019 *Frank, R.L.; Geisinger, L.:* Two-term spectral asymptotics for the Dirichlet Laplacian on a bounded domain
- 2011-018 *Hänel, A.; Schulz, C.; Wirth, J.:* Embedded eigenvalues for the elastic strip with cracks
- 2011-017 *Wirth, J.:* Thermo-elasticity for anisotropic media in higher dimensions
- 2011-016 *Höllig, K.; Hörner, J.:* Programming Multigrid Methods with B-Splines
- 2011-015 *Ferrario, P.:* Nonparametric Local Averaging Estimation of the Local Variance Function
- 2011-014 *Müller, S.; Dippon, J.:* k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis
- 2011-013 *Knarr, N.; Stroppel, M.:* Unitals over composition algebras
- 2011-012 *Knarr, N.; Stroppel, M.:* Baer involutions and polarities in Moufang planes of characteristic two
- 2011-011 *Knarr, N.; Stroppel, M.:* Polarities and planar collineations of Moufang planes
- 2011-010 *Jentsch, T.; Moroianu, A.; Semmelmann, U.:* Extrinsic hyperspheres in manifolds with special holonomy
- 2011-009 *Wirth, J.:* Asymptotic Behaviour of Solutions to Hyperbolic Partial Differential Equations
- 2011-008 *Stroppel, M.:* Orthogonal polar spaces and unitals
- 2011-007 *Nagl, M.:* Charakterisierung der Symmetrischen Gruppen durch ihre komplexe Gruppenalgebra
- 2011-006 *Solanes, G.; Teufel, E.:* Horo-tightness and total (absolute) curvatures in hyperbolic spaces
- 2011-005 *Ginoux, N.; Semmelmann, U.:* Imaginary Kählerian Killing spinors I
- 2011-004 *Scherer, C.W.; Köse, I.E.:* Control Synthesis using Dynamic D -Scales: Part II — Gain-Scheduled Control
- 2011-003 *Scherer, C.W.; Köse, I.E.:* Control Synthesis using Dynamic D -Scales: Part I — Robust Control
- 2011-002 *Alexandrov, B.; Semmelmann, U.:* Deformations of nearly parallel G_2 -structures
- 2011-001 *Geisinger, L.; Weidl, T.:* Sharp spectral estimates in domains of infinite volume
- 2010-018 *Kimmerle, W.; Konovalov, A.:* On integral-like units of modular group rings
- 2010-017 *Gauduchon, P.; Moroianu, A.; Semmelmann, U.:* Almost complex structures on quaternion-Kähler manifolds and inner symmetric spaces
- 2010-016 *Moroianu, A.; Semmelmann, U.:* Clifford structures on Riemannian manifolds
- 2010-015 *Grafarend, E.W.; Kühnel, W.:* A minimal atlas for the rotation group $SO(3)$
- 2010-014 *Weidl, T.:* Semiclassical Spectral Bounds and Beyond
- 2010-013 *Stroppel, M.:* Early explicit examples of non-desarguesian plane geometries
- 2010-012 *Effenberger, F.:* Stacked polytopes and tight triangulations of manifolds
- 2010-011 *Györfi, L.; Walk, H.:* Empirical portfolio selection strategies with proportional transaction costs
- 2010-010 *Kohler, M.; Krzyżak, A.; Walk, H.:* Estimation of the essential supremum of a regression function
- 2010-009 *Geisinger, L.; Laptev, A.; Weidl, T.:* Geometrical Versions of improved Berezin-Li-Yau Inequalities

- 2010-008 *Poppitz, S.; Stroppel, M.:* Polarities of Schellhammer Planes
- 2010-007 *Grundhöfer, T.; Krinn, B.; Stroppel, M.:* Non-existence of isomorphisms between certain unitals
- 2010-006 *Höllig, K.; Hörner, J.; Hoffacker, A.:* Finite Element Analysis with B-Splines: Weighted and Isogeometric Methods
- 2010-005 *Kaltenbacher, B.; Walk, H.:* On convergence of local averaging regression function estimates for the regularization of inverse problems
- 2010-004 *Kühnel, W.; Solanes, G.:* Tight surfaces with boundary
- 2010-003 *Kohler, M.; Walk, H.:* On optimal exercising of American options in discrete time for stationary and ergodic data
- 2010-002 *Gulde, M.; Stroppel, M.:* Stabilizers of Subspaces under Similitudes of the Klein Quadric, and Automorphisms of Heisenberg Algebras
- 2010-001 *Leitner, F.:* Examples of almost Einstein structures on products and in cohomogeneity one
- 2009-008 *Griesemer, M.; Zenk, H.:* On the atomic photoeffect in non-relativistic QED
- 2009-007 *Griesemer, M.; Moeller, J.S.:* Bounds on the minimal energy of translation invariant n-polaron systems
- 2009-006 *Demirel, S.; Harrell II, E.M.:* On semiclassical and universal inequalities for eigenvalues of quantum graphs
- 2009-005 *Bächle, A.; Kimmerle, W.:* Torsion subgroups in integral group rings of finite groups
- 2009-004 *Geisinger, L.; Weidl, T.:* Universal bounds for traces of the Dirichlet Laplace operator
- 2009-003 *Walk, H.:* Strong laws of large numbers and nonparametric estimation
- 2009-002 *Leitner, F.:* The collapsing sphere product of Poincaré-Einstein spaces
- 2009-001 *Brehm, U.; Kühnel, W.:* Lattice triangulations of E^3 and of the 3-torus
- 2008-006 *Kohler, M.; Krzyżak, A.; Walk, H.:* Upper bounds for Bermudan options on Markovian data using nonparametric regression and a reduced number of nested Monte Carlo steps
- 2008-005 *Kaltenbacher, B.; Schöpfer, F.; Schuster, T.:* Iterative methods for nonlinear ill-posed problems in Banach spaces: convergence and applications to parameter identification problems
- 2008-004 *Leitner, F.:* Conformally closed Poincaré-Einstein metrics with intersecting scale singularities
- 2008-003 *Effenberger, F.; Kühnel, W.:* Hamiltonian submanifolds of regular polytope
- 2008-002 *Hertweck, M.; Höfert, C.R.; Kimmerle, W.:* Finite groups of units and their composition factors in the integral group rings of the groups $PSL(2, q)$
- 2008-001 *Kovarik, H.; Vugalter, S.; Weidl, T.:* Two dimensional Berezin-Li-Yau inequalities with a correction term
- 2007-006 *Weidl, T.:* Improved Berezin-Li-Yau inequalities with a remainder term
- 2007-005 *Frank, R.L.; Loss, M.; Weidl, T.:* Polya's conjecture in the presence of a constant magnetic field
- 2007-004 *Ekholm, T.; Frank, R.L.; Kovarik, H.:* Eigenvalue estimates for Schrödinger operators on metric trees
- 2007-003 *Lesky, P.H.; Racke, R.:* Elastic and electro-magnetic waves in infinite waveguides
- 2007-002 *Teufel, E.:* Spherical transforms and Radon transforms in Moebius geometry
- 2007-001 *Meister, A.:* Deconvolution from Fourier-oscillating error densities under decay and smoothness restrictions