A Posteriori Error Analysis of Optimal Control Problems with Control Constraints

Kristina Kohls, Arnd Rösch, Kunibert G. Siebert
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:

\begin{equation}
\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \tag{1.1}
\end{equation}

The module \text{SOLVE} solves the optimal control problem in a finite element space that is defined over a given grid. The module \text{ESTIMATE} computes local error indicators that constitute an estimator for the true error. Based on information of the indicators the module \text{MARK} selects a subset of elements subject to refinement. This is executed by \text{REFINE} returning a suitable refinement of the current grid. In this paper we focus on the module \text{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 \cite{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 \cite{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 \cite{LY01}. Compared to the vast amount of literature about a posteriori error estimators for linear problems the existing results for optimal control problems are
rather limited. We would like to mention [BBM+07, BV09, HHAV08, HHIKOS, HH08, HK10, LY02, LLY07, VW08, ZY08].

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 [Lyo10], 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

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

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 \( \mathbb{Y} \). 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 \( \Gamma \subset \partial \Omega \), which allows us to simultaneously treat distributed controls, i.e., \( \Gamma \subset \mathbb{R}^d \) with positive \( d \)-dimensional Lebesgue measure, and boundary controls, i.e., \( \Gamma \subset \partial \bar{\Omega} \) 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 \( \omega_1 \cap \omega_2 = \emptyset \) and \( v \in \mathbb{Y} \) we have \( \|v\|^2_\mathbb{Y}\omega_1,\omega_2) = \|v\|^2_\mathbb{Y}\omega_1 + \|v\|^2_\mathbb{Y}\omega_2 \), and, similarly for \( \mathbb{U} \).

The convex, closed and non-empty set \( \mathbb{U}^{ad} \subset \mathbb{U} \) 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 \( u \in \mathbb{U} \) with \( \|v\|_\mathbb{U} \leq \gamma \|v\|_\mathbb{Y} \) and \( u \in \mathbb{U} \) implies \( u \in \mathbb{Y}^* \) by

\[ (u, v) := \langle u, v \rangle_{\mathbb{Y}^* \times \mathbb{Y}} = \langle u, v \rangle_{\mathbb{U}} \quad \forall v \in \mathbb{Y}. \]

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

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

\[
\inf_{v \in \mathbb{Y}} \sup_{w \in \mathbb{Y}} \mathcal{B}[v, w] = \inf_{v \in \mathbb{Y}} \sup_{w \in \mathbb{Y}} \mathcal{B}[v, w] = \beta > 0. \tag{2.2}
\]

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

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

For the discretization we additionally suppose that \( \Omega \) and \( \Gamma \) can be meshed; the precise assumptions are stated in \( \S 3.1 \). In \( \S 4 \) and \( \S 5 \) 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 \( \psi \) for the desired state.

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

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

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

for any \( q \in \mathbb{Y}^* \) \( \text{[Neč62, Theorem 5.3]; compare also [NSV09, §2.3]} \) for a more detailed discussion of the inf-sup theory. This particularly implies the existence of solution
operators $S, S^* : \mathbb{Y}^* \to \mathbb{Y}$ 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
\[ \|S\|_{\mathcal{L}(\mathbb{Y}^*, \mathbb{Y})} \leq \beta^{-1} \quad \text{and} \quad \|S^*\|_{\mathcal{L}(\mathbb{Y}^*, \mathbb{Y})} \leq \beta^{-1}. \] (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
\[ \|S\|_{\mathcal{L}(\mathbb{U}, \mathbb{Y})} \leq \beta^{-1} \gamma. \] (2.5)

Utilizing the solution operator $S$, we introduce the reduced cost functional $\mathcal{J}[u] := J[u, S(f + u)]$. The assumptions on $\psi, \alpha > 0$, and the inf-sup condition (2.2) imply that $\mathcal{J}[u]$ is bounded from below and strictly convex. If $U_{ad}$ is not bounded $\mathcal{J}[u]$ is radially unbounded. In combination with the fact that $U_{ad}$ 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 $(\hat{u}, \hat{y}) = (\hat{u}, S(f + \hat{u})) \in U_{ad} \times \mathbb{Y}$.

Introducing the adjoint state $\hat{p} = S^*\psi'(\hat{y}) \in \mathbb{Y}$, the triplet $(\hat{u}, \hat{y}, \hat{p}) \in U_{ad} \times \mathbb{Y} \times \mathbb{Y}$ is characterized as the unique solution of the first order optimality system
\[ \text{state equation:} \quad \mathcal{B}[\hat{y}, v] = (f + \hat{u}, v)_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y}, \quad (2.6a) \]
\[ \text{adjoint equation:} \quad \mathcal{B}[v, \hat{p}] = (\psi'(\hat{y}), v)_{\mathbb{Y}^* \times \mathbb{Y}} \quad \forall v \in \mathbb{Y}, \quad (2.6b) \]
\[ \text{gradient equation:} \quad \hat{u} = \Pi(\hat{p}), \quad (2.6c) \]
where $\Pi : \mathbb{U} \to U_{ad}$ denotes for given $p$ the best approximation of $-\frac{1}{\alpha} p$ in $U_{ad}$, i.e., $\Pi(p)$ is uniquely determined by
\[ \langle \alpha \Pi(p) + p, \Pi(p) - u \rangle_{\mathbb{U}} \leq 0 \quad \forall u \in U_{ad}. \] (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 $a \leq C b$ with a constant that may depend on data constants $\alpha, \beta, \gamma, \|B\|$, 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 $(\hat{u}, \hat{y}, \hat{p}) \in U_{ad} \times \mathbb{Y} \times \mathbb{Y}$ 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 $\Pi : \mathbb{U} \to U_{ad}$ is the nonlinear projection operator of (2.6c), $S, S^* : \mathbb{Y}^* \to \mathbb{Y}$ 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 \( \dot{y} = S(f + \bar{u}) \) we deduce with (2.5)
\[
\|S(f + u) - \dot{y}\|_Y = \|S(u - \bar{u})\|_Y \leq \beta^{-1}\gamma\|u - \bar{u}\|_U.
\]
This implies the first claim by
\[
\|y - \dot{y}\|_Y \leq \|y - S(f + u)\|_Y + \|S(f + u) - \dot{y}\|_Y \leq \|y - S(f + u)\|_Y + \beta^{-1}\gamma\|u - \bar{u}\|_U.
\]
The second bounds follows similarly by interchanging \( \dot{y} \) and \( S(f + u) \).

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

Lemma 2.4. We have
\[
\|p - \bar{p}\|_Y \leq \frac{\gamma L}{\alpha^2} \|u - \bar{u}\|_U + \frac{L}{\beta} \|y - S(f + u)\|_Y + \|p - S^*\psi'(y)\|_Y,
\]
\[
\|p - S^*\psi'(y)\|_Y \leq \frac{L}{\beta} \|y - \dot{y}\|_Y + \|p - \bar{p}\|_Y.
\]

Proof. The adjoint state is \( \bar{p} = S^*(\psi'(\dot{y})) \). Using similar arguments as before we obtain by (2.4) with the Lipschitz constant \( L \) of \( \psi' \)
\[
\|S^*\psi'(y) - \bar{p}\|_Y = \|S^*(\psi'(y) - \psi'(\dot{y}))\|_Y \leq \beta^{-1}\|\psi'(y) - \psi'(\dot{y})\|_Y \leq \beta^{-1}L\|y - \dot{y}\|_Y.
\]
Combining this with Lemma [2.3] gives
\[
\|p - \bar{p}\|_Y \leq \|p - S^*\psi'(y)\|_Y + \|S^*\psi'(y) - \bar{p}\|_Y
\leq \|p - S^*\psi'(y)\|_Y + \beta^{-1}L\|y - S(f + u)\|_Y + \beta^{-1}\gamma\|u - \bar{u}\|_U,
\]
which proves the first bound. The second one is shown similarly by
\[
\|p - S^*\psi'(y)\|_Y \leq \|p - \bar{p}\|_Y + \|S^*(\psi'(\dot{y}) - \psi'(y))\|_Y \leq \|p - \bar{p}\|_Y + \beta^{-1}L\|y - \dot{y}\|_Y. \quad \square
\]

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

Proposition 2.5. We have
\[
\|u - \bar{u}\|_U \leq \left( 1 + \frac{\gamma^2 L}{\alpha^2 \beta^2} \right) \|u - \Pi(p)\|_U + \frac{\gamma L}{\alpha^2 \beta} \|y - S(f + u)\|_Y + \frac{\gamma}{\alpha} \|p - S^*\psi'(y)\|_Y,
\]
\[
\|u - \Pi(p)\|_U \leq \|u - \bar{u}\|_U + \frac{\gamma}{\alpha} \|p - \bar{p}\|_Y.
\]

Proof. We proceed in several steps.

[1] We introduce the auxiliary control \( \bar{u} = \Pi(p) \) and prove first
\[
\|\bar{u} - \bar{u}\|_U \leq \frac{\gamma L}{\alpha^2} \|y - S(f + \bar{u})\|_Y + \frac{\gamma}{\alpha} \|p - S^*\psi'(y)\|_Y. \tag{2.8}
\]
Recalling \( \bar{u} = \Pi(\bar{p}) \) and \( \bar{u} = \Pi(p) \), the definition (2.7) of \( \Pi \) implies
\[
\langle \alpha \bar{u} + p, \bar{u} - \bar{u}\rangle_U \leq 0 \quad \text{and} \quad \langle \alpha\bar{u} + \bar{p}, \bar{u} - \bar{u}\rangle_U \leq 0.
\]
From this we conclude with \( \bar{p} = S^*(\psi'(\dot{y})) \)
\[
\alpha \|\bar{u} - \bar{u}\|_U^2 = \langle \alpha\bar{u} + p, \bar{u} - \bar{u}\rangle_U + \langle \alpha\bar{u} + \bar{p}, \bar{u} - \bar{u}\rangle_U + \langle p - \bar{p}, \bar{u} - \bar{u}\rangle_U
\leq \langle p - \bar{p}, \bar{u} - \bar{u}\rangle_U
\leq \|p - \bar{p}\|_Y \|\bar{u} - \bar{u}\|_U + \|S^*(\psi'(y) - \psi'(\dot{y}))\|_U \|\bar{u} - \bar{u}\|_U.
\]
The first term on the right hand side we bound by
\[
\langle p - S^*\psi'(y), \bar{u} - \bar{u}\rangle_U \leq \|p - S^*\psi'(y)\|_U \|\bar{u} - \bar{u}\|_U \leq \gamma \|p - S^*\psi'(y)\|_Y \|\bar{u} - \bar{u}\|_U.
\]
We start with the Lipschitz constant of $\Pi$ is $K$. KOHLS, A. RÖSCH, AND K. G. SIEBERT, giving the upper bound

$$\gamma \frac{1}{\alpha\beta} \|u - \Pi(p)\|_\mu + \frac{\gamma L}{\alpha} \|u - \Pi(p)\|_\mu + \frac{\gamma}{\alpha} \|p - S^*\psi(y)\|_\mu 
\leq \frac{1}{\alpha\beta} \frac{\gamma^2 L^2}{\alpha^2} \|u - \Pi(p)\|_\mu + \frac{\gamma L}{\alpha \beta} \|y - S(f + u)\|_\mu + \frac{\gamma}{\alpha} \|p - S^*\psi(y)\|_\mu.$$

which implies (2.8).

We have with

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

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

$$\|u - \bar{u}\|_\mu \leq \|u - \bar{u}\|_\mu + \|\Pi(p) - \Pi(p)\|_\mu \leq \|u - \Pi(p)\|_\mu + \frac{\gamma L}{\alpha} \|y - S(f + u)\|_\mu + \frac{\gamma}{\alpha} \|p - S^*\psi(y)\|_\mu.$$

The Lipschitz constant of $\Pi$ is $\alpha^{-1}$, which can be seen from

$$\alpha \|\Pi(u_1) - \Pi(u_2)\|_\mu = \langle \alpha \Pi(u_1) + u_1, \Pi(u_1) - \Pi(u_2)\rangle_\mu + \langle \alpha \Pi(u_2) + u_2, \Pi(u_2) - \Pi(u_1)\rangle_\mu$$

for all $u_1, u_2 \in \mu$ by (2.7). This immediately yields

$$\|u - \Pi(p)\|_\mu \leq \|u - \bar{u}\|_\mu + \|\Pi(\bar{p}) - \Pi(p)\|_\mu \leq \|u - \bar{u}\|_\mu + \alpha^{-1} \gamma \|p - \bar{p}\|_\mu,$$

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

$$\begin{bmatrix} \|u - \bar{u}\|_\mu \\ \|y - \bar{y}\|_\mu \\ \|p - \bar{p}\|_\mu \end{bmatrix} \leq \begin{bmatrix} 1 + \frac{\gamma^2 L^2}{\alpha^2} & \frac{\gamma L}{\alpha \beta} & \frac{\gamma}{\alpha} \\ \frac{\gamma}{\alpha \beta} & 1 + \frac{\gamma^2 L^2}{\alpha^2} & \frac{\gamma L}{\alpha \beta} \\ \frac{\gamma}{\alpha \beta} & \frac{\gamma L}{\alpha \beta} & 1 + \frac{\gamma^2 L^2}{\alpha^2} \end{bmatrix} \begin{bmatrix} \|u - \Pi(p)\|_\mu \\ \|y - S(f + u)\|_\mu \\ \|p - S^*\psi(y)\|_\mu \end{bmatrix}.$$

The respective second estimates immediately yield the lower bound

$$\begin{bmatrix} \|u - \bar{u}\|_\mu \\ \|y - \bar{y}\|_\mu \\ \|p - \bar{p}\|_\mu \end{bmatrix} \leq \begin{bmatrix} 1 & \frac{\gamma}{\alpha \beta} & 0 \\ \frac{\gamma}{\alpha \beta} & 1 & 0 \\ 0 & \frac{\gamma}{\alpha \beta} & 1 \end{bmatrix} \begin{bmatrix} \|u - \bar{u}\|_\mu \\ \|y - \bar{y}\|_\mu \\ \|p - \bar{p}\|_\mu \end{bmatrix}.$$

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 \(\Omega\) that is locally quasi-uniform. Suppose that \(\mathcal{Y}(\mathcal{T}) \subset \mathcal{Y}\) is a conforming finite element space over \(\mathcal{T}\). To also allow for stabilized discretizations we let \(\mathcal{B}_\mathcal{T}: \mathcal{Y}(\mathcal{T}) \times \mathcal{Y}(\mathcal{T}) \to \mathbb{R}\) and \(\langle \cdot, \cdot \rangle_{\mathcal{T}}\) be a discrete continuous bilinear form and a discrete duality pairing on \(\mathcal{Y}(\mathcal{T})^* \times \mathcal{Y}(\mathcal{T})\).

We suppose that \(\mathcal{B}_\mathcal{T}\) satisfies the discrete inf-sup condition

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

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 \(\|\cdot\|_{\mathcal{T}} = \|\cdot\|_{\mathcal{Y}}\). Otherwise, \(\|\cdot\|_{\mathcal{T}}\) is chosen accordingly to the stabilization and it is typically stronger than \(\|\cdot\|_{\mathcal{Y}}\). Any single inf-sup condition implies the other one, thanks to \(\mathcal{Y}(\mathcal{T})\) being finite dimensional. Consequently, the inf-sup theory implies that for any \(q \in \mathcal{Y}(\mathcal{T})^*\) there exist unique solutions to the discrete state and adjoint equation

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

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

In the a posteriori error analysis only the continuous inf-sup constant \(\beta\) 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 \(\beta(\mathcal{T}) \geq \beta > 0\) 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 \(\Pi: \mathbb{U} \to \mathbb{U}^\text{ad}\) introduced in (2.7) is computable for discrete functions \(V \in \mathcal{Y}(\mathcal{T})\). This assumption allows us to pose the following discrete optimal control problem:

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

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 \(\Gamma\). 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

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

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

\[
(\alpha \Pi_\mathcal{G}(P) + P, \Pi_\mathcal{G}(P) - U)_\mathcal{T} \leq 0 \quad \forall U \in \mathbb{U}^\text{ad}(\mathcal{G}) \quad \text{(3.4)}
\]

replaces the continuous projection \(\Pi: \mathbb{U} \to \mathbb{U}^\text{ad}\). The grids \(\mathcal{T}\) and \(\mathcal{G}\) are typically connected in the following way. In case of distributed controls \(\mathcal{G} \subset \mathcal{T}\). In case of boundary controls \(\mathcal{G}\) is the trace of \(\mathcal{T}\) on \(\Gamma\). This allows for an easy transfer of information between the finite element spaces \(\mathcal{Y}(\mathcal{T})\) and \(\mathcal{Y}(\mathcal{G})\) 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 \((\hat{U}, \hat{Y}) \in \mathbb{U}^\text{ad} \times \mathcal{Y}(\mathcal{T})\) to (3.2), and \((\hat{U}, \hat{Y}) \in \mathbb{U}^\text{ad}(\mathcal{G}) \times \mathcal{Y}(\mathcal{T})\) 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
\[
\hat{Y} \in Y(\mathcal{T}) : \quad B_T[\hat{Y}, V] = \langle f + \hat{U}, V \rangle_{\mathcal{T}} \quad \forall V \in Y(\mathcal{T}),
\]
\[
\hat{P} \in Y(\mathcal{T}) : \quad B_T[V, \hat{P}] = \langle \psi'(\hat{Y}), V \rangle_{\mathcal{T}} \quad \forall V \in Y(\mathcal{T})
\]
and
\[
\hat{U} = \Pi(\hat{P}) \in U_{\text{ad}} \quad \text{or} \quad \hat{U} = \Pi_G(\hat{P}) \in U_{\text{ad}}(G)
\]
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 \(U(G)\).

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
\[
\mathcal{E}_y(Y, q; T) = \left( \sum_{T \in \mathcal{T}} \mathcal{E}_y^2(Y, q; T) \right)^{1/2} \quad \text{and} \quad \mathcal{E}_p(P, q; T) = \left( \sum_{T \in \mathcal{T}} \mathcal{E}_p^2(P, q; T) \right)^{1/2}
\]
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 \(T' \subseteq \mathcal{T}\) we set
\[
\text{osc}_y(Y, q; T') = \left( \sum_{T \in T'} \text{osc}_y^2(Y, q; T) \right)^{1/2}
\]
and similarly for osc\(_p\). For \(T \in \mathcal{T}\) we let \(\mathcal{N}_T(T) = \{T' \in \mathcal{T} | T' \cap T \neq \emptyset\}\) be the set of its direct neighbors, and \(\omega_T(T) = \bigcup_{T' \in \mathcal{N}_T(T)} T'\) 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 $E_u$ and $E_p$ provide an upper bound for the true error, i.e., for any $q_1 \in f + U$ and $q_2 \in \text{rg}(\psi')$ and the Galerkin approximations $Y_{q_1}, P_{q_2} \in \mathcal{Y}(T)$ of $(3.1a)$ and $(3.1b)$, respectively, we have
\[
\begin{align*}
\|Y_{q_1} - S q_1\|_\mathcal{Y} &\leq E_u(Y_{q_1}, q_1; T), \\
\|P_{q_2} - S^* q_2\|_\mathcal{Y} &\leq E_p(P_{q_2}, q_2; T).
\end{align*}
\]

(2) **Local Efficiency:** The indicators of $E_u$ and $E_p$ are local lower bounds for the true error up to oscillation, i.e., for any $Y, P \in \mathcal{Y}(T)$ and any $q_1 \in f + U$ and $q_2 \in \text{rg}(\psi')$ we have
\[
\begin{align*}
E_u(Y, q_1; T) &\leq \|Y - S q_1\|_{\mathcal{Y}(\omega_T(T))} + \text{osc}_y(Y, q_1; N_T(T)), \\
E_p(P, q_2; T) &\leq \|P - S^* q_2\|_{\mathcal{Y}(\omega_T(T))} + \text{osc}_p(P, q_2; N_T(T)).
\end{align*}
\]

(3) **Lipschitz Continuity of Indicators:** The indicators of $E_u$ and $E_p$ are Lipschitz continuous with respect to their second arguments, i.e., for $Y, P, q_1, q_2 \in f + U$, and $q_2, q_2 \in \text{rg}(\psi')$ we have
\[
\begin{align*}
|E_u(Y, q_1; T) - E_u(Y, q_1; T)| &\leq \|q_1 - q_1\|_{U(T)}, \\
|E_p(P, q_2; T) - E_p(P, q_2; T)| &\leq \|y - y\|_{\mathcal{Y}(T)},
\end{align*}
\]
where $y, \bar{y} \in \mathcal{Y}$ are such that $\psi'(y) = q_2$ and $\psi'(\bar{y}) = \bar{q}_2$.

We would like to remark that asking for $q_1, q_2 \in \mathcal{Y}^*$ in the definition of $E_u$ and $E_p$ 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 $E_u(U, p; T) = \|U - \Pi(p)\|_{U(T)}$ with the convention $\|U - \Pi(p)\|_{U(\emptyset)} = 0$, and we set
\[
\|U - \Pi(p)\|^2_T = E_u^2(U, p; T) = \sum_{T \in \mathcal{T}} E_u^2(U, p; T).
\]
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 $E_u$ and $E_p$ satisfy Assumption 3.1. Then
\[
E_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; T) := E_u(\hat{U}, \hat{P}; T) + E_p(\hat{Y}, f + \hat{U}; T) + E_p(\hat{P}, \psi'(\hat{Y}); T)
\]
is an estimator for the optimal control problem, which is reliable, i.e.,
\[
\|U - \hat{u}\|_U + \|\hat{Y} - \hat{y}\|_\mathcal{Y} + \|\hat{P} - \hat{p}\|_\mathcal{Y} \leq E_{\text{ocp}}(U, \hat{Y}, \hat{P}; T),
\]
and globally efficient, i.e.,
\[
E_{\text{ocp}}(U, \hat{Y}, \hat{P}; T) \leq \|U - \hat{u}\|_U + \|\hat{Y} - \hat{y}\|_\mathcal{Y} + \|\hat{P} - \hat{p}\|_\mathcal{Y}
+ \text{osc}_y(\hat{Y}, f + \hat{u}, T) + \text{osc}_p(\hat{P}, \psi'(\hat{y}); T).
\]

In case of non-discretized control $f$ we have $E_u(\hat{U}, \hat{P}; T) = 0$ and the estimator $E_{\text{ocp}}$ is locally efficient, i.e., for all $T \in \mathcal{T}$ we have
\[
E_{\text{ocp}}(U, \hat{Y}, \hat{P}; T) \leq \|U - \hat{u}\|_{U(T)} + \|\hat{Y} - \hat{y}\|_{\mathcal{Y}(\omega_T(T))} + \|\hat{P} - \hat{p}\|_{\mathcal{Y}(\omega_T(T))}
+ \text{osc}_y(\hat{Y}, f + \hat{u}, N_T(T)) + \text{osc}_p(\hat{P}, \psi'(\hat{y}); N_T(T)).
\]

**Proof.** 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 $U = \Pi(P)$. This implies $E_u(U, \hat{P}; T) = \|U - \Pi(P\|_U = 0$. 

Looking at the contribution \( \mathcal{E}_u \) we derive from Assumption 3.1 (3) and (2)
\[
\mathcal{E}_u(\bar{y}, f + \bar{u}; T) \leq \mathcal{E}_u(\bar{y}, f + \bar{u}; T) + \|\bar{U} - \bar{u}\|_{U(\Gamma^T)}
\lesssim \|\bar{U} - \bar{u}\|_{U(\Gamma^T)} + \|\bar{Y} - \bar{y}\|_{Y(\Gamma^T)} + \text{osc}_u(\bar{y}, f + \bar{u}; \mathcal{N}_T(T)).
\]
Arguing similarly for \( \mathcal{E}_p \) we derive
\[
\mathcal{E}_p(\tilde{P}, \psi'(\tilde{y}); T) \leq \mathcal{E}_p(\tilde{P}, \psi'(\tilde{y}); T) + \|\tilde{Y} - \tilde{y}\|_{Y(T)}
\lesssim \|\tilde{P} - \tilde{p}\|_{Y(\Gamma^T)} + \|\tilde{Y} - \tilde{y}\|_{Y(T)} + \text{osc}_p(\tilde{P}, \psi'(\tilde{y}); \mathcal{N}_T(T)).
\]
This shows local efficiency in case of non-discretized control.

In the proof of Proposition 3.5 we have shown that \( \Pi \) is globally Lipschitz-continuous with constant \( \alpha^{-1} \). This yields
\[
\mathcal{E}_u(U, P; T) \leq \|\tilde{U} - \tilde{U}(\tilde{p})\|_U + \|\Pi(\tilde{p}) - \Pi(U(\hat{P}))\|_U \leq \|\tilde{U} - \tilde{u}\|_U + \alpha^{-1} \gamma\|\hat{P} - \tilde{p}\|_Y,
\]
utilizing the (global) embedding \( \Psi \hookrightarrow U \).

Local quasi-uniformity of \( \mathcal{T} \) implies
\[
\max_{T \in \mathcal{T}} \#\{T' \in \mathcal{T} \mid |\mathcal{N}_T(T') \cap \mathcal{N}_T(T)| > 0\} \lesssim 1.
\]
Additivity of \( \| \cdot \|_T^2 \) and \( \text{osc}_p^2 \) then in turn yields
\[
\sum_{T \in \mathcal{T}} \|v\|_{\mathcal{Y}(\Gamma^T)}^2 \lesssim \|v\|_{\mathcal{Y}}^2 \quad \text{and} \quad \sum_{T \in \mathcal{T}} \text{osc}_p^2(Y, q; \mathcal{N}_T(T)) \lesssim \text{osc}_p^2(Y, q; T).
\]
This is also true for \( \| \cdot \|_T^2 \) and \( \text{osc}_p^2 \). Consequently, summing up the indicators of step 2 in combination with the bound \( \mathcal{E}_u \) yields global efficiency in any case.

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 \( \Pi \) the term \( \mathcal{E}_u(U, \Pi; T) \) 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 3.1 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 \( \mathcal{E}_{\text{ocp}} \) being locally efficient is \( \mathcal{E}_u \). This term is only locally efficient if \( \Pi \) and the embedding \( \Psi \hookrightarrow U \) have the following local properties:

- The continuous projection operator \( \Pi : U \rightarrow \mathcal{U}_d \) is locally Lipschitz, i. e.,
\[
\|\Pi(u) - \Pi(\bar{u})\|_{U(\Gamma^T)} \lesssim \|u - \bar{u}\|_{U(\Gamma^T)} \quad \forall T \in \mathcal{T}, u, \bar{u} \in U.
\]
- The embedding \( \Psi \hookrightarrow U \) is locally uniform, i. e.,
\[
\|y\|_{U(\Gamma^T)} \lesssim \|y\|_{\mathcal{Y}(T)} \quad \forall T \in \mathcal{T}, y \in \Psi.
\] (3.9)

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

The assumption on \( \Pi \) is for instance fulfilled for box constraints. The embedding \( \Psi \hookrightarrow U \) 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 \( \mathcal{E}_{\text{ocp}} \) 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:

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

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 \( \mathcal{Y} = H^1(\Omega) \) with norm \( \| \cdot \|_Y = \| \cdot \|_{H^1(\Omega)} \),
and the control space is \( \mathcal{U} = L_2(\Gamma) \) with norm \( \| \cdot \|_U = \| \cdot \|_{L_2(\Gamma)} \). The embedding
\( \mathcal{Y} \hookrightarrow \mathcal{U} \hookrightarrow \mathcal{Y}^* \) is naturally given by the trace operator and \( \gamma \) equals its norm.

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

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

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

We use box constraints, i.e., for given \( a, b \in \mathcal{U}(T_0) \) with \( a \leq b \) we let

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

The continuous projection operator \( \Pi: \mathcal{U} \to \mathcal{U}^{ad} \) 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 \( g \in \mathcal{U} \) and desired state \( y_d \in L_2(\Omega) \) we define the objective

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

The Fréchet derivative of \( \psi \) is given by \( \psi'(y) = \langle y - y_d, \cdot \rangle_{L_2(\Omega)} + \langle g, \cdot \rangle_{L_2(\Gamma)} \in \mathcal{Y}^* \).
It 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 \( u \in \mathcal{U} \) solve for \( y = S(f + u) \), i.e.,

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

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

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

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 \( T_0 \) meshes \( \Omega \) exactly and
in such a way that \( T_0 := \{ T \cap \Gamma \mid T \in T_0 \} \) meshes \( \Gamma \) exactly. We concentrate on the case
of discretized control and use piecewise polynomial spaces of degree \( \ell_y \geq 1, \; \ell_u \geq 1 \) such that

\[ \mathcal{Y}(T) := \{ V \in C^0(\Omega) \mid V|_T \in \mathbb{P}_{\ell_y}(T) \forall T \in T \}, \]

\[ \mathcal{U}(G) := \{ V \in L_2(\Gamma) \mid V|_E \in \mathbb{P}_{\ell_u}(E) \forall E \in G \}. \]

The assumption \( a, b \in \mathcal{U}(T_0) \) implies \( \mathcal{U}^{ad}(G) = \mathcal{U}(G) \cap \mathcal{U}^{ad} \neq \emptyset \). We choose to use \( B_T = B \) and \( \langle \cdot, \cdot \rangle_T = \langle \cdot, \cdot \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \) without any stabilization. The bilinear-form \( B \) is coercive on \( \mathcal{Y}(T) \subset H^1(\Omega) \), which yields the discrete inf-sup condition with
\[ \beta_T = \beta = 1. \] Let \((\hat{U}, \hat{Y}, \hat{P}) \in \text{Unad}(\mathcal{G}) \times \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T})\) be the solution to the discrete first order optimality system (3.5) in this setting.

4.2. A Posteriori Error Control. We accept \(\mathcal{E}_i(\hat{U}, \hat{P}; \mathcal{T}) = \|\hat{U} - \Pi(\hat{P})\|_{L_2(\mathcal{T})}\) 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 \(\Sigma\) 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:

\[ \mathcal{E}_y(\hat{Y}, \hat{U} + f; T) = h_T \| - \Delta \hat{Y} + \hat{Y} - f_2\|_{L_2(T)} + h_T^{1/2} \| \hat{Y}\|_{L_2(\partial T \cap \Gamma)} + h_T\| (U + f_1)||_{L_2(\partial T \cap \Gamma)}, \]

\[ \mathcal{E}_p(\hat{P}, \psi'(\hat{Y}); T) = h_T \| - \Delta \hat{P} + \hat{P} - (\hat{Y} - y_0)\|_{L_2(T)} + h_T^{1/2} \| \hat{P}\|_{L_2(\partial T \cap \Gamma)} + h_T\| g\|_{L_2(\partial T \cap \Gamma)}, \]

where \(h_T = |T|^{1/4}\) is the local mesh-size and \([-\] \+] denotes the flux of the normal derivative on inter-element sides and the normal derivative for a boundary side.

The estimators \(\mathcal{E}_y, \mathcal{E}_u\) 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 \(\mathbb{Y}(\mathcal{T})' \supset \mathbb{Y}(\mathcal{T})\). 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

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

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 \(S\) the set of all sides of \(\mathcal{T}\). If \(E \in S\) is a boundary side there exists a unique \(T \in \mathcal{T}\) with \(E \subset \mathcal{T}\) and we set \(\omega_E := T\). Otherwise \(E\) is an inter-element side of \(T, T' \subset \mathcal{T}\) 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 \(\mathbb{Y}(\mathcal{T})' \supset \mathbb{Y}(\mathcal{T})\) that provides for any \(E \in S\) a function \(\Phi_E \in \mathbb{Y}(\mathcal{T})' \setminus \mathbb{Y}(\mathcal{T})\) with

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

We let \(h_e := \text{diam}(E)\). For \(E \subset \partial \Omega\) we define \(\text{osc}_T(y; E) := h_E^{1/2} \|y - \bar{y}\|_{L_2(E)}\) with the mean value \(\bar{y} = |E|^{-1} \int_E gdA\). For an interior side \(E\) we define \(\text{osc}_T(y; E) := 0.\)
The side-oriented indicators on \( E \in S \) for (4.1a) and (4.1b) read:

\[
\mathcal{E}_y^2(\hat{Y}, \hat{U} + f; E) := ||(\text{Res}_y(\hat{Y}, \hat{U} + f), \Phi_E)||^2 \\
+ h_E^2||\hat{Y} - f_2||_{L_2(\omega_E)}^2 + \text{osc}_T(f_1; E)
\]

\[
\mathcal{E}_p^2(\hat{P}, \psi'(\hat{Y}); E) := ||(\text{Res}_p(\hat{P}, \psi'(\hat{Y})), \Phi_E)||^2 \\
+ h_E^2||\hat{P} - (\hat{Y} - y_0)||_{L_2(\omega_E)}^2 + \text{osc}_T(g; E).
\]

One easily obtains element-based indicators on \( T \in T \) 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. This comparison can be found in [KST1] §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 and efficiency and without using the equivalence to the residual estimator can be found in [KST1] §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:

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

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

The state space is \( Y = H^1_0(\Omega; \mathbb{R}^d) \times \{ q \in L_2(\Omega) \mid \langle q, 1 \rangle_{L_2(\Omega)} = 0 \} \). We use the norm \( \| y \|_{Y(\Omega)}^2 = \| \nabla y \|_{L_2(\Omega)}^2 + \| q \|_{L_2(\Omega)}^2 \) for \( y = (y, q) \in Y \). The space \( Y \) is a Hilbert space with respect to \( \| \cdot \|_Y \) thanks to the Friedrichs estimate \( \| y \|_{L_2(\Omega)} \leq C_F \| y \|_{L_2(\Omega)} \) on \( H^1_0(\Omega; \mathbb{R}^d) \). The control space is \( U = L_2(\Omega; \mathbb{R}^d) \) with the natural norm \( \| \cdot \|_{U(\Omega)} = \| \cdot \|_{L_2(\Omega)} \). The embedding \( Y \hookrightarrow U \hookrightarrow Y^* \) is given by the inclusion \( H^1_0(\Omega; \mathbb{R}^d) \subset L_2(\Omega; \mathbb{R}^d) \) and \( \gamma \leq C_F \).

The bilinear form \( B \) for the weak formulation of the PDE is

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

It is continuous with \( ||B|| \leq 3 + C_F \| b \|_{L_\infty(\Omega)} \). Moreover, \( B \) satisfies the inf-sup condition (2.2) thanks to \( \text{div} b = 0 \), 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

\[ \mathbb{U}^\text{ad} = \{ u \in \mathbb{U} \mid \| u \| \leq R \text{ in } \Omega \} \neq \emptyset. \]
The continuous projection operator $\Pi: U \rightarrow U^{\text{ad}}$ is therefore given for almost all $x \in \Omega$ by

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

Aiming at minimizing variations in the velocity $y$ at a low pressure $q$ and recalling $\|y\|_2^2 = \|(y, q)\|_2^2 = \|\nabla y\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)}^2$, we define the objective by

$$J(u, y) = \psi(y) + \frac{\alpha}{2} \|u\|_V^2 := \frac{1}{2} \|y\|_V^2 + \frac{\alpha}{2} \|u\|_V^2.$$ 

The functional $\psi = \frac{1}{2} \| \cdot \|_2'$ is Fréchet differentiable, and $\psi'$ 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 $u \in U$ solve for $y = S(f + u)$, i.e.,

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

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

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

Since $B$ 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 $\Omega$ and suppose that the discrete state space $\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}$ of continuous, piecewise polynomials of fixed degree $(f_y, f_q)$ for velocity and pressure. For the stabilization of the advection derivative $[\nabla.]b$ 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 $B_\mathcal{D}[\cdot, \cdot] = B[\cdot, \cdot] + B_{\text{stab}}[\cdot, \cdot]$ and the discrete duality pairing $\langle \cdot, \cdot \rangle_{\mathcal{T}} = \langle \cdot, \cdot \rangle_{\mathbb{Y}(\mathcal{T}) \times \mathbb{Y}} + \langle \cdot, \cdot \rangle_{\text{stab}}$ 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

$$B_{\text{stab}}[Y, V] := \delta \sum_{T \in \mathcal{T}} h_T^2 \int_T (\nabla V)^T (\nabla y) b + \nabla R \cdot (\Delta Y + [\nabla Y]^T b + \nabla Q) \, dV$$

$$+ \mu \int_\Omega \text{div} V \text{div} Y \, dV.$$ 

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

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

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

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

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

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

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

$$\langle Y, V \rangle_{\text{stab}} = \delta \sum_{T \in \mathcal{T}} h_T^2 \int_T \Delta Y \cdot (\nabla V)^T (\nabla y) b + \nabla R) \, dV + \mu \int_\Omega Q \text{div} V \, dV.$$
In [RST08, Chap. IV, Lemma 3.3] it has been shown that for sufficiently small $\delta > 0$ the discrete bilinear form $B_T$ satisfies a discrete inf-sup condition with respect to the mesh-dependent norm $\|\cdot\|^2_T = \|\cdot\|^2_Y + \|\cdot\|^2_{\text{stab}}$, where

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

Using the inverse estimate $h_T \|\nabla R\|_{L_2(T)} \lesssim \|R\|_{L_2(T)}$, in combination with the bound $\|\div V\|_{L_2(\Omega)} \leq \|\nabla V\|_{L_2(\Omega)}$ we see $\|V\|_{\text{stab}} \leq C(\|b\|_{L_\infty(\Omega)}, \delta, \mu) \|V\|_Y$. The discrete inf-sup constant $\beta(T)$ is bounded from below by $\beta > 0$, which solely depends on $\delta, \mu$ and the shape regularity of $T$. Consequently, both (5.2a) and (5.2b) are uniquely solvable. On top of this, the mesh-dependent norm $\|\cdot\|_T$ 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 $\ell_y \geq 2$ and $\ell_q = \ell_y - 1$ 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 $[\nabla V]b$ but gives at the same time a stable discretization not depending on the chosen polynomial degree for velocity and pressure. In particular, $\ell_y = \ell_q = 1$ is included, which is advantageous in optimal control problems for an easy computation of the projection operator $\Pi$: $\mathbb{Y}(T) \to \mathbb{U}^{\text{rad}}$.

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

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

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

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

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

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

### 5.2 A Posteriori Error Analysis

For non-discretized control we have $\mathcal{E}_u = 0$ by Theorem 3.3. 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 $g = f + u$ and define for $T \in T$ the indicators

$$\mathcal{E}^2_{y,T}(Y, g; T) := h_T^2 \|\nabla Y\|_T - \|\nabla Y\|b + \|\nabla f - g\|^2_{L_2(T)} + h_T \|\nabla Y\|^2_{L_2(\partial T)} + \|\div Y\|^2_{L_2(\Omega)}.$$

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

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

$$\|Y - y\|_T^2 \lesssim \mathcal{E}^2_{y,T}(Y, g; T) = \sum_{T \in T} \mathcal{E}^2_{y,T}(Y, g; T).$$
For any \( \bar{Y} = (\bar{y}, \bar{q}) \in \mathcal{Y}(T) \) we have the local lower bound
\[
\mathcal{E}_Y^2(\bar{Y}, g; T) \lesssim \|\bar{Y} - y\|_{L_2(T)} + \text{osc}^2_y(\bar{Y}, g; N_T(T))
\]
\( \forall T \in T \)
with \( \text{osc}^2(\bar{Y}, g; T) = h_T^2\|g - g_T\|^2_{L_2(T)} + h_T^2\|b - b_T\|^2_{L_2(T)} \), where \( g_T \) and \( b_T \) are suitable finite dimensional approximations of \( g \) and \( b \) on \( T \).

**Proof.** Starting point of the analysis is the residual
\[
\text{Res}_y(\bar{Y}, g) := B[\bar{Y}, \cdot] - \langle g, \cdot \rangle \in \mathcal{Y},
\]
for \( \bar{Y} \in \mathcal{Y}(T) \). Well-posedness of the continuous problem \( 5.1 \) yields for the solution \( y \) the error equivalence
\[
\beta \|\bar{Y} - y\|_Y \leq \|\text{Res}_y(\bar{Y}, g)\|_{\mathcal{Y}} \leq \|B\|\|\bar{Y} - y\|_Y
\]
and it remains to estimate \( \|\text{Res}_y(\bar{Y}, g)\|_{\mathcal{Y}} \).

**Proof.** For the SDFEM solution \( Y \) we have
\[
\langle \text{Res}_y(Y, g), V \rangle_{\mathcal{Y} \times \mathcal{Y}} = -\langle \text{Res}_{\text{stab}}(Y, g), V \rangle_{\mathcal{Y} \times \mathcal{Y}} =: -\langle \text{Res}_{\text{stab}}(Y, g), W \rangle_{\mathcal{Y} \times \mathcal{Y}}
\]
for all \( V \in \mathcal{Y}(T) \), which replaces Galerkin-orthogonality. For given \( v = (v, r) \in \mathcal{Y} \) we let \( V = (I_T v, 0) \in \mathcal{Y}(T) \), where \( I_T \) denotes the Clément or Scott-Zhang interpolation operator into the discrete velocity space \([CR75, SZ90]\). This in turn yields
\[
\|\text{Res}_y(Y, g)\|_{\mathcal{Y}} \leq \sup_{\|v\|_{\mathcal{Y}} = 1} \langle \text{Res}_y(Y, g), v - V \rangle_{\mathcal{Y} \times \mathcal{Y}} + C \sup_{W = (W, 0) \in \mathcal{Y}(T)} \|W\|_{\mathcal{Y}} \langle \text{Res}_{\text{stab}}(Y, g), W \rangle_{\mathcal{Y} \times \mathcal{Y}}
\]
thanks to \( \|V\|_{\mathcal{Y}} \leq C\|v\|_{\mathcal{Y}} \).

**Proof.** For Stokes, i.e., \( b = 0 \), Verfürth has shown
\[
\sup_{\|v\|_{\mathcal{Y}} = 1} \langle \text{Res}_y(Y, g), v - V \rangle_{\mathcal{Y} \times \mathcal{Y}} \lesssim \mathcal{E}_y(Y, g; T)
\]
where the constant hidden in ‘\( \lesssim \)’ only depends on properties of \( I_T [Ver89, Theorem 3.1] \). The generalization of this bound to Oseen, i.e., \( b \neq 0 \), is straightforward.

For the stabilization term we easily deduce for \( W = (W, 0) \) by the definition of \( \mathcal{E}_y \) and \( \|\cdot\|_{\text{stab}} \) the bound
\[
\langle \text{Res}_{\text{stab}}(Y, g), W \rangle = \delta \sum_{T \in T} h_T^2 \int_T [\nabla W] b \cdot [\Delta Y + [\nabla Y] b + \nabla Q - g] \, dV + \mu \int_{\Omega} \text{div} W \, \text{div} Y \, dV
\]
\[
\leq \sqrt{\max\{\delta, \mu\}} \mathcal{E}_y(Y, g; T) (\|W\|_{\text{stab}} + \|\text{div} W\|_{L_2(\Omega)}).
\]
Since \( \|W\|_{\text{stab}} + \|\text{div} W\|_{L_2(\Omega)} \lesssim \|W\|_{\mathcal{Y}} \) we have shown the upper bound.

**Proof.** The lower bound is the straightforward generalization of the lower bound in \[Ver89, Theorem 3.1\] from Stokes to Oseen.

**Adjoint equation.** We define for \( T \in T \) the indicators
\[
\mathcal{E}_p^2(P, \psi'(Y); T) := h_T^2\|\Delta(P - Y) - [\nabla P] b + \nabla S\|^2_{L_2(T)}
\]
\[
+ h_T\|\Delta(P - Y)\|^2_{L_2(\partial T)} + \|\text{div} P - Q\|^2_{L_2(T)},
\]
which constitute an efficient and reliable estimator for the SDFEM discretization of the adjoint equation.
Proposition 5.4. For given $Y \in \mathcal{Y}(\Omega)$ let $p \in \mathcal{Y}$ be the solution to (5.1b) and $P \in \mathcal{Y}(\Omega)$ be the SDFEM approximation given by (5.2b). Then we have the global upper bound
\[
\|P - p\|_Y^2 \lesssim E_y^2(p, \psi'(Y); T) = \sum_{T \in T} E_y^2(p, \psi'(Y); T).
\]
For any $\bar{P} = (\bar{P}, \bar{S}) \in \mathcal{Y}(\Omega)$ we have the local lower bound
\[
E_y^2(\bar{P}, \psi'(Y); T) \lesssim \|\bar{P} - p\|_Y^2 + \text{osc}^2_p(\bar{P}, \psi(Y); N_T(T)) \quad \forall T \in T
\]
with $\text{osc}^2_p(\bar{P}, \psi'(Y); T) = h_T^2\|b - b_T\|_{L_2(T)}^2$.

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

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
\[
E_{\text{ocp}}(\hat{U}, \hat{Y}, \hat{P}; T) = E_y(\hat{Y}, f + \hat{U}; T) + E_p(\hat{P}, \psi'(Y); T)
\]
is a reliable and locally efficient estimator for (2.1). For non-discretized control there is no explicit contribution for the control error $\|\hat{U} - \tilde{u}\|_{L_2(\Omega)}$. However, the total error is influenced by the distance of $\hat{U}$ to $\tilde{u}$.

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

Remark 5.6. In [HYZ09] an a posteriori error estimator for a scalar advection diffusion problem with a simpler objective $\psi$ 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 $\Gamma$ for the boundary control is indicated by bold lines.

6. Numerical Experiments

We conclude the article by two numerical experiments in order 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 $\alpha$ 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.

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):

$$f_1(x) := - \text{Proj}_{-0.5}^{+0.5} \left( -|x|^{\frac{2}{3}} \cos(\frac{2}{3}\phi(x)) \right),$$

$$f_2(x) := 0,$$

$$y_d(x) := -|x|^{\frac{2}{3}} \cos(\frac{2}{3}\phi(x)),$$

$$g(x) := \frac{2}{3}|x|^{-\frac{2}{3}} \left( - \cos(\frac{2}{3}\phi(x)) \cdot x + x^\perp \sin(\frac{2}{3}\phi(x)) \right) \cdot n_\Gamma(x).$$
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 \( n_\Gamma \) is the outer unit normal vector to \( \Omega \).

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

\[
\hat{y}(x) = 0, \quad \hat{p}(x) = |x|^\frac{2}{3} \cos(\hat{\phi}(x)), \quad \hat{u}(x) = \text{Proj}_{0.5}^{1.5} \left( -|x|\frac{2}{3} \cos(\hat{\phi}(x)) \right).
\]

**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 \( E_u(T), E_y(T) \), and \( E_p(T) \) instead of \( E_u(\hat{U}, \hat{P}; T), E_y(\hat{Y}, \hat{U}; T), \) and \( E_p(\hat{P}, \psi'(\hat{Y}); T) \). We then use \( E_{\text{ocp}}(T) = E_u^2(T) + E_y^2(T) + E_p^2(T) \).

We observe that the different contributions \( E_u, E_y, \) and \( E_p \) of \( E_{\text{ocp}} \) 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 \( \{E(T)\}_{T \in \mathcal{T}} \) the maximum strategy \( \text{MS}(\{E(T)\}_{T \in \mathcal{T}}) \) outputs a subset \( \mathcal{M} \subset \mathcal{T} \) of marked elements such that

\[
\mathcal{M} = \{ T \in \mathcal{T} | E(T) \geq \theta E_{\text{max}} \} \quad \text{with} \quad E_{\text{max}} := \max \{ E(T) \mid T \in \mathcal{T} \}.
\]

Two strategies involve all three contributions.

**(AR-1)** We apply the maximum strategy to \( E_{\text{ocp}} \), i.e.,

\[
\mathcal{M} = \text{MS}(\{E_{\text{ocp}}(T)\}_{T \in \mathcal{T}}).
\]
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
$$I_\delta := \{i \in \{u, y, p\} \mid \mathcal{E}_i(T) \geq \delta \max \{\mathcal{E}_u(T), \mathcal{E}_y(T), \mathcal{E}_p(T)\}\}.$$

We then proceed as follows.

(AR-3) For $T \in \mathcal{T}$ set $\mathcal{E}^2(T) := \sum_{i \in I_\delta} \mathcal{E}_{i, T}^2(T)$ and apply the maximum strategy to $\mathcal{E}_i$, 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 $\frac{2}{3}$. 
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 $E_u$, $E_y$, and $E_p$ 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 $\Gamma = \{0\} \times [1, 2] \subseteq \partial \Omega$; 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):
\[
f_1(x) := -\text{Proj}_a^{(2n)} \left( \frac{x}{2n} \left( (2n + 1)(\frac{3}{4}x_2 - 1) - \left( \frac{3}{4}x_2 - 1 \right)^{2n+1} \right) \right)
\]
\[
f_2(x) := e^{-10|x|^2} \left( 41 - 1600|x|^2 \right),
\]
\[
g(x) := 0,
\]
\[
y_d(x) := e^{-10|x|^2} - Z \left( \frac{2n+1}{4}x_2 - 1 + \frac{8n+4}{9} \left( \frac{3}{4}x_2 - 1 \right)^{2n-1} - \left( \frac{3}{4}x_2 - 1 \right)^{2n+1} \right).
\]

The true solution is
\[
\hat{u}(x) = \text{Proj}_a^{(2n)} \left( \frac{x}{2n} \left( (2n + 1)(\frac{3}{4}x_2 - 1) - \left( \frac{3}{4}x_2 - 1 \right)^{2n+1} \right) \right),
\]
\[
\hat{y}(x) = e^{-10|x|^2},
\]
\[
\hat{p}(x) = \frac{2}{2n} \left( (2n + 1)(\frac{3}{4}x_2 - 1) - \left( \frac{3}{4}x_2 - 1 \right)^{2n+1} \right)
\]
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 \(x_2 \in [0, 2, 8]\) 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 \(\Gamma\).

**Active Case:** For \(a = 15\) we have \(-\hat{p} \leq b\) and \(\hat{u} \equiv a\) on \(\Gamma\).

**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 $\#\text{DOFs}^{-1/2}$ 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 $E_u$.

**Remark 6.1.** In all our experiments we have used $E_u(\hat{U}, \hat{P}; T) = \|\hat{U} - \Pi(\hat{P})\|_U$ as indicator for the control error. The exact computation of $E_u$ 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 $E(\hat{U}, \hat{P}; T)$ can be computed exactly. Moreover, Theorem 3.2 states that this term is globally efficient.

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

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 $\Gamma$ in all three cases. This shows, that neither $\|\hat{U} + \frac{1}{\alpha} \hat{P}\|_U$ nor $\|\hat{P} - M_T \hat{P}\|_Y$ can be efficient for our error notion $\|\hat{U} - \hat{u}\|_U + \|\hat{Y} - \hat{y}\|_Y + \|\hat{P} - \hat{p}\|_Y$. Efficiency of the estimator is a prerequisite for an optimal performance of an adaptive method;
compare with [CKNS08, §5]. We therefore conclude that the additional effort for computing $\|\hat{U} - \Pi(\hat{P})\|_U$ exactly pays off.

Efficiency for $\|\hat{U} + \frac{1}{\alpha} \hat{P}\|_U$ or $\|\hat{P} - M_T \hat{P}\|_Y$ 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 $\Gamma$ will persist in all three cases.

**References**


26

K. KOHLS, A. RÖSCH, AND K. G. SIEBERT


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

2012-018 Kimmerle, W.; Konovalov, A.: On the Prime Graph of the Unit Group of Integral Ring Groups of Finite Groups II

2012-017 Stroppel, B.; Stroppel, M.: Desargues, Doily, Dualities, and Exceptional Isomorphisms


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


2012-012 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


2011-026 Müller, S.: Bootstrapping for Bandwidth Selection in Functional Data Regression


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

<table>
<thead>
<tr>
<th>Year</th>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>019</td>
<td>Frank, R.L.; Geisinger, L.</td>
<td>Two-term spectral asymptotics for the Dirichlet Laplacian on a bounded domain</td>
</tr>
<tr>
<td>2011</td>
<td>018</td>
<td>Händel, A.; Schulz, C.; Wirth, J.</td>
<td>Embedded eigenvalues for the elastic strip with cracks</td>
</tr>
<tr>
<td>2011</td>
<td>017</td>
<td>Wirth, J.</td>
<td>Thermo-elasticity for anisotropic media in higher dimensions</td>
</tr>
<tr>
<td>2011</td>
<td>016</td>
<td>Höllig, K.; Hörner, J.</td>
<td>Programming Multigrid Methods with B-Splines</td>
</tr>
<tr>
<td>2011</td>
<td>015</td>
<td>Ferrario, P.</td>
<td>Nonparametric Local Averaging Estimation of the Local Variance Function</td>
</tr>
<tr>
<td>2011</td>
<td>014</td>
<td>Müller, S.; Dippon, J.</td>
<td>k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis</td>
</tr>
<tr>
<td>2011</td>
<td>013</td>
<td>Knarr, N.; Stroppel, M.</td>
<td>Unitals over composition algebras</td>
</tr>
<tr>
<td>2011</td>
<td>012</td>
<td>Knarr, N.; Stroppel, M.</td>
<td>Baer involutions and polarities in Moufang planes of characteristic two</td>
</tr>
<tr>
<td>2011</td>
<td>011</td>
<td>Knarr, N.; Stroppel, M.</td>
<td>Polarities and planar collineations of Moufang planes</td>
</tr>
<tr>
<td>2011</td>
<td>010</td>
<td>Jentsch, T.; Moroianu, A.; Semmelmann, U.</td>
<td>Extrinsic hyperspheres in manifolds with special holonomy</td>
</tr>
<tr>
<td>2011</td>
<td>009</td>
<td>Wirth, J.</td>
<td>Asymptotic Behaviour of Solutions to Hyperbolic Partial Differential Equations</td>
</tr>
<tr>
<td>2011</td>
<td>008</td>
<td>Stroppel, M.</td>
<td>Orthogonal polar spaces and unital</td>
</tr>
<tr>
<td>2011</td>
<td>007</td>
<td>Nagl, M.</td>
<td>Charakterisierung der Symmetrischen Gruppen durch ihre komplexe Gruppenalgebra</td>
</tr>
<tr>
<td>2011</td>
<td>006</td>
<td>Solanes, G.; Teufel, E.</td>
<td>Horo-tightness and total (absolute) curvatures in hyperbolic spaces</td>
</tr>
<tr>
<td>2011</td>
<td>005</td>
<td>Ginoux, N.; Semmelmann, U.</td>
<td>Imaginary Kählerian Killing spinors I</td>
</tr>
<tr>
<td>2011</td>
<td>002</td>
<td>Alexandrov, B.; Semmelmann, U.</td>
<td>Deformations of nearly parallel G_2-structures</td>
</tr>
<tr>
<td>2010</td>
<td>018</td>
<td>Kimmerle, W.; Konovalov, A.</td>
<td>On integral-like units of modular group rings</td>
</tr>
<tr>
<td>2010</td>
<td>017</td>
<td>Gauduchon, P.; Moroianu, A.; Semmelmann, U.</td>
<td>Almost complex structures on quaternion-Kähler manifolds and inner symmetric spaces</td>
</tr>
<tr>
<td>2010</td>
<td>016</td>
<td>Moroianu, A.; Semmelmann, U.</td>
<td>Clifford structures on Riemannian manifolds</td>
</tr>
<tr>
<td>2010</td>
<td>015</td>
<td>Grafarend, E.W.; Kühnel, W.</td>
<td>A minimal atlas for the rotation group SO(3)</td>
</tr>
<tr>
<td>2010</td>
<td>014</td>
<td>Weidl, T.</td>
<td>Semiclassical Spectral Bounds and Beyond</td>
</tr>
<tr>
<td>2010</td>
<td>013</td>
<td>Stroppel, M.</td>
<td>Early explicit examples of non-desarguesian plane geometries</td>
</tr>
<tr>
<td>2010</td>
<td>012</td>
<td>Effenberger, F.</td>
<td>Stacked polytopes and tight triangulations of manifolds</td>
</tr>
<tr>
<td>2010</td>
<td>010</td>
<td>Kohler, M.; Krzyżak, A.; Walk, H.</td>
<td>Estimation of the essential supremum of a regression function</td>
</tr>
</tbody>
</table>
2010-008  Poppitz, S.; Stroppel, M.:  Polarieties of Schellhammer Planes
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
2009-005  Bächle, A, Kimmerle, W.:  Torsion subgroups in integral group rings of finite groups
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
2008-004  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 $PSL(2, q)$
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
2007-003  Lesky, R.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