

An a posteriori error analysis for distributed elliptic optimal control problems with pointwise state constraints

Dissertation

zur Erlangung des akademischen Titels eines
Doktors der Naturwissenschaften
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Augsburg

vorgelegt von Michael Kieweg
geboren am 27.01.1982 in Schwabmünchen

Betreuer: Prof. Dr. Ronald H. W. Hoppe

1. Gutachter: Prof. Dr. Ronald H. W. Hoppe
2. Gutachter: Prof. Dr. Kunibert G. Siebert
3. Gutachter: Prof. Dr. Martin Brokate

Mündliche Prüfung: 05. Dezember 2007

DANKSAGUNGEN

Diese Arbeit wäre wohl nicht ohne eine breite Unterstützung und Hilfe von zahlreichen Personen zustande gekommen.

Als erstes gilt mein ganz besonderer Dank dem Betreuer dieser Dissertation Prof. Dr. Ronald H. W. Hoppe von der Universität Augsburg für die intensive Betreuung meines wissenschaftlichen Schaffens und für die zahlreichen informativen und ergebnisreichen Diskussionen über die Thematik dieser Doktorarbeit. Besonders schätze ich auch, dass Prof. Hoppe mir die Möglichkeit gab, Teile dieser Arbeit in Houston/USA zu erstellen.

Ferner danke ich Prof. Dr. Kunibert G. Siebert von der Universität Augsburg und Prof. Dr. Martin Brokate von der TU München dafür, sich als weitere Gutachter zur Verfügung zu stellen.

Mein Dank geht auch an Dr. Yuri Iliash von der Universität Augsburg für die Erlaubnis seine FEM-Bibliotheken zu verwenden.

Diese Dissertation entstand im Rahmen des Elitestudiengangs TopMath. Ich danke allen Gründern und Organisatoren dieses Studiengangs, allen voran Dr. Ralf Franken und Dr. Christian Kredler von der TU München, sowie Prof. Dr. Karl Heinz Borgwardt von der Universität Augsburg, für ihren Einsatz und ihre Hilfsbereitschaft.

Außerdem gilt mein Dank dem Elitenetzwerk Bayern für die finanzielle Unterstützung in Form eines Promotionsstipendiums nach dem bayerischen Eliteförderungsgesetz. Insbesondere danke ich meiner Familie für ihre unermüdliche Unterstützung in jeglicher Hinsicht während meiner gesamten Ausbildungszeit. Besonderer Dank gilt dabei meinem Vater Dr. Werner Kieweg für seine tatkräftige Mithilfe bei der Verfassung dieser Arbeit in englischer Sprache.

A B S T R A C T

This thesis is concerned with the development, analysis, and implementation of an adaptive finite element method for distributed elliptic optimal control problems with pointwise unilateral constraints on the state. In particular, two residual-type a posteriori error estimators will be derived. The first one takes advantage of the modified adjoint state, which is defined as some kind of regularization of the adjoint state. Furthermore, this error estimator will, after minor modification, be transferred to the Lavrentiev regularization of the pure state constrained case. Up to a consistency error and data oscillation, reliability and efficiency results concerning the approximation of the state, the control, and the modified adjoint state can be provided for these error estimators. With two numerical examples, the performance of the adaptive algorithm will be investigated. A benefit compared to an uniform refinement strategy will be noticeable.

The second developed a posteriori error estimator results from a measure extension of the discrete measure appearing in the right-hand side of the adjoint state equation to an element in $L^2(\Omega)$. This error estimator provides, again up to a consistency error and data oscillation, reliability and efficiency for the approximation error in the control, in the state, and in a semi-continuous auxiliary adjoint state. Another numerical example will show that this error estimator might be advantageous.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 11 |
| 2 | The distributed optimal control problem | 17 |
| 2.1 | Model problem - existence and uniqueness | 17 |
| 2.2 | An illustrative example | 20 |
| 2.3 | Optimality conditions | 21 |
| 2.4 | Discretization and discrete optimality conditions | 23 |
| 3 | A posteriori error analysis | 29 |
| 3.1 | The residual-type a posteriori error estimator | 29 |
| 3.2 | Reliability of the error estimator | 32 |
| 3.3 | Efficiency of the error estimator | 36 |
| 4 | Lavrentiev Regularization | 43 |
| 4.1 | Mixed control-state constraints | 44 |
| 4.2 | Optimality conditions of the regularized problem | 45 |
| 4.3 | Convergence to the unregularized problem | 48 |
| 4.4 | Discretization and modified Lagrange multipliers | 52 |
| 4.5 | A posteriori error estimator | 55 |
| 4.6 | Reliability of the error estimator | 57 |
| 4.7 | Efficiency of the error estimator | 60 |
| 5 | Numerical results | 63 |
| 5.1 | The adaptive algorithm | 63 |
| 5.2 | Implementation Issues | 68 |
| 5.3 | Example 1 - Smooth Lagrange Multiplier | 68 |
| 5.4 | Example 2 - Nonsmooth Lagrange multiplier | 80 |
| 6 | The obstacle problem approach | 91 |
| 6.1 | Measure extension error estimator | 92 |
| 6.2 | Reliability | 95 |
| 6.3 | Efficiency | 101 |
| 6.4 | Discussion of the consistency error | 105 |
| 6.5 | Numerical experiment | 110 |

4

CONTENTS

7 Conclusion and Outlook

117

List of symbols

The following list consists of symbols and expressions as they are used in this thesis. An explanation and/or a reference of its definition is quoted for each symbol.

Greek letters

| symbol | explanation and/or reference |
|--|--|
| α | regularization parameter of the objective functional J |
| Γ_D | Dirichlet part of the boundary $\partial\Omega$ |
| δ_n | Dirac-measure in the point n |
| Δ | Laplace operator |
| ε | Lavrentiev regularization parameter |
| $\{\varepsilon_n\}_{n \in \mathbb{N}}$ | arbitrary positive null sequence |
| ζ_E | extension of $\zeta_E \in P_k(E)$ to ω_E |
| η | error estimator, (3.1.3) |
| $\eta_{\hat{p},T}$ | element residual of η , (3.1.2b) |
| $\eta_{y,T}$ | element residual of η , (3.1.2a) |
| $\eta_{\hat{p},E}$ | edge residual of η , (3.1.2d) |
| $\eta_{y,E}$ | edge residual of η , (3.1.2c) |
| $\tilde{\eta}$ | measure extension error estimator, (6.1.9) |
| $\tilde{\eta}_{p,T}$ | element residual of $\tilde{\eta}$, (6.1.8b) |
| $\tilde{\eta}_{y,T}$ | element residual of $\tilde{\eta}$, (6.1.8a) |
| $\tilde{\eta}_{p,E}$ | edge residual of $\tilde{\eta}$, (6.1.8d) |
| $\tilde{\eta}_{y,E}$ | edge residual of $\tilde{\eta}$, (6.1.8c) |
| η_ε | regularized error estimator, (4.5.1) |
| $\eta_{\hat{p},\varepsilon,T}$ | element residual of η_ε , (4.5.3b) |
| $\eta_{y,\varepsilon,T}$ | element residual of η_ε , (4.5.3a) |
| $\eta_{\hat{p},\varepsilon,E}$ | edge residual of η_ε , (4.5.4b) |
| $\eta_{y,\varepsilon,E}$ | edge residual of η_ε , (4.5.4a) |
| $\theta_1, \theta_2, \theta_3, \theta_4$ | constants of the bulk criterion, (5.1.3)-(5.1.6) |
| ϑ_E | edge bubble function |
| ϑ_T | element bubble function |
| λ_1^E, λ_2^E | barycentric coordinates in E |
| $\lambda_1^T, \lambda_2^T, \lambda_3^T$ | barycentric coordinates in T |
| $\mu_h(\cdot), \mu_T(\cdot)$ | lower order data oscillation, (3.1.4), (4.5.5), (6.1.10) |
| ν_E | exterior unit normal vector of the edge E |
| $\sigma, \bar{\sigma}$ | optimal adjoint control, (2.3.1) |
| $\sigma_\varepsilon, \bar{\sigma}_\varepsilon$ | regularized adjoint control, (4.2.16) |
| $\sigma_{\varepsilon,h}$ | discrete regularized adjoint control, (4.4.4) |
| $\sigma_h, \bar{\sigma}_h$ | discrete optimal adjoint control, (2.4.5) |

| | |
|--------------------------------|---|
| $\tilde{\sigma}_h$ | measure extension, (6.1.3) |
| $\hat{\sigma}$ | modified adjoint control, (2.3.1) |
| $\hat{\sigma}_{\varepsilon,h}$ | discrete regularized modified adjoint control, (4.4.16) |
| $\hat{\sigma}_h$ | discrete modified adjoint control, (2.4.10) |
| ψ | unilateral constraint of the state variable, (2.1.3c) |
| ψ_h | approximation of ψ in S_h |
| ω_E | $T_1 \cup T_2$, where $T_1 \cap T_2 = E$ |
| $\tilde{\omega}_E$ | (6.2.12) |
| $\tilde{\omega}_T$ | (6.2.12) |
| Ω | open, polygonal domain in \mathbb{R}^2 |

Latin letters

| symbol | explanation and/or reference |
|------------------------------|--|
| \mathcal{A} | active set, (2.3.13a) |
| \mathcal{A}^ε | regularized active set, (4.4.1a) |
| \mathcal{A}_h | discrete active set, (6.4.1a) |
| $\mathcal{A}_h^\varepsilon$ | discrete regularized active set, (4.4.8) |
| B | linear operator as defined in (6.2.15) |
| c | parameter of the partial differential equation (2.1.1) |
| \tilde{c} | (3.1.12) |
| $C(\overline{\Omega})$ | Banach space of continuous functions on Ω |
| $C_+(\overline{\Omega})$ | positive coin of $C(\overline{\Omega})$ |
| $C_0^\infty(\Omega)$ | infinitely often differentiable functions with compact support in Ω |
| D | Frechet derivative |
| $e_c^{(1)}, e_c^{(2)}$ | consistency error, (4.5.8) |
| e_h^n | nodal basis function in the vertice n , (6.1.2) |
| $\{e_i\}_{i \in \mathbb{N}}$ | basis of eigenvectors of G |
| $e_c(u, u_h)$ | consistency error, (3.1.11) |
| $e_c(u_h)$ | consistency error, (6.1.16) |
| E | edge of the triangulation |
| \mathcal{E}_h | set of all edges of the triangulation |
| $\mathcal{E}_h(D)$ | set of edges in $D \subset \overline{\Omega}$ |
| $\mathcal{E}_h(\Omega)$ | set of all interior edges |
| \mathcal{F}_h | discrete free boundary, (5.1.2) |
| $\mathcal{F}_h^\varepsilon$ | discrete regularized free boundary, (4.4.11) |
| G | solution operator of the partial differential equation, (2.1.1) |
| G_h | discrete solution operator of the partial differential equation, (2.4.2) |
| h_E | length of an edge E |
| h_T | diameter of an element T |
| $H^k(\Omega)$ | $W^{k,2}(\Omega)$ |

| | |
|--------------------------------------|---|
| $H_0^k(\Omega)$ | $W_0^{k,2}(\Omega)$ |
| I | indicator function, (2.3.2) |
| I_h | Clement interpolation operator, cf. [Ver96] |
| \mathcal{I} | continuous inactive set, (2.3.13b) |
| \mathcal{I}^ε | regularized inactive set, (4.4.1b) |
| \mathcal{I}_h | discrete inactive set, (6.4.1b) |
| $\mathcal{I}_h^\varepsilon$ | discrete regularized inactive set, (4.4.9) |
| J | objective functional, (2.1.3a) |
| J_h | discrete objective functional, (2.4.3a) |
| $J_{red}, J(\cdot)$ | reduced objective functional, (2.1.6) |
| \tilde{J} | objective functional as defined in (4.2.2) |
| K | set of admissible states, (2.1.3c) |
| K_h | set of admissible discrete states, (2.4.3c) |
| K_2 | (6.2.17) |
| \hat{K} | (4.1.1c) |
| L | Lagrange functional of Problem (P_ε) , (4.2.15) |
| L_h | Lagrange interpolation operator, (6.1.1) |
| $L^p(\Omega)$ | p-integrable functions |
| \tilde{L} | Lagrange functional of Problem (\tilde{P}_ε) , (4.2.11) |
| $meas_1$ | one dimensional Lebesgue measure |
| M_h | L^2 -projection on the space W_h , (3.1.7) |
| \mathcal{M}^E | selected edges for refinement, (5.1.3) |
| $\mathcal{M}^{\eta,T}$ | selected triangles due to element residuals, (5.1.4) |
| $\mathcal{M}^{fb,T}$ | selected triangles due to the resolution of the discrete free boundary |
| $\mathcal{M}^{\mu,T}$ | selected triangles due to lower order data oscillation, (5.1.5) |
| $\mathcal{M}^{osc,T}$ | selected triangles due to higher order data oscillation, (5.1.6) |
| \mathcal{M}^T | selected triangles for refinement |
| \mathcal{M}_h | (2.4.4) |
| $\mathcal{M}(\Omega)$ | space of Radon measures |
| $\mathcal{M}_+(\Omega)$ | positive coin of $\mathcal{M}(\Omega)$ |
| \mathcal{N}_h | set of all vertices of the triangulation |
| N_{dof} | number of degrees of freedom |
| $\mathcal{N}_h(D)$ | set of vertices of the triangulation in $D \subset \bar{\Omega}$ |
| $osc_h(\cdot), osc_T(\cdot)$ | higher order data oscillation, (3.1.5), (6.1.11) |
| p, \bar{p} | optimal adjoint state, (2.3.1) |
| $p_\varepsilon, \bar{p}_\varepsilon$ | regularized optimal adjoint state, (4.2.20) |
| $p_{\varepsilon,h}$ | discrete regularized optimal adjoint state, (4.4.4) |
| p_h, \bar{p}_h | discrete optimal adjoint state, (2.4.5c) |
| $p(u_h, \tilde{\sigma}_h)$ | auxiliary adjoint state, (6.1.15) |
| \hat{p} | modified adjoint state, (2.3.8) |
| \hat{p}_ε | regularized modified adjoint state, (4.4.13) |
| $\hat{p}_{\varepsilon,h}$ | discrete regularized modified adjoint state, (4.4.15) |

| | |
|--------------------------------------|---|
| \hat{p}_h | discrete modified adjoint state, (2.4.9) |
| $\hat{p}(y_{\varepsilon,h})$ | auxiliary regularized modified adjoint state, (4.5.6b) |
| $\hat{p}(y_h)$ | auxiliary modified adjoint state, (3.1.9) |
| (P_ε) | regularized problem, (4.1.1) |
| (\tilde{P}_ε) | auxiliary problem, (4.2.3) |
| $P_k(D)$ | polynomial of degree k on D |
| $\partial F(a)$ | subdifferential of the function F at the point a |
| ∂K | boundary of the domain F |
| r | regularity of the state |
| R | operator as defined in (4.2.1) |
| \mathcal{R}_h | (6.2.16) |
| s | regularity of the adjoint state |
| S_h | linear finite element space, (2.4.1a) |
| T | triangle of the triangulation $\mathcal{T}_h(\Omega)$ |
| \mathcal{T}_h | $\mathcal{T}_h(\Omega)$ |
| $\mathcal{T}_h(\Omega)$ | shape regular triangulation |
| u, \bar{u} | optimal control of (2.1.3) |
| u^d | control shift, (2.1.3) |
| $u_\varepsilon, \bar{u}_\varepsilon$ | regularized optimal control, (4.1.1) |
| $u_{\varepsilon,h}$ | discrete regularized control, (4.4.4) |
| u_h, \bar{u}_h | discrete optimal control of (2.4.5) |
| u_h^d | approximation of u^d in S_h |
| $\{u_n\}_{n \in \mathbb{N}}$ | sequence of optimal solutions corresponding to $\{(P_{\varepsilon_n})\}_{n \in \mathbb{N}}$ |
| \tilde{u}_n | (4.3.1) |
| U_{ad} | set of admissible controls, (2.1.6) |
| \bar{v}_ε | optimal solution of (\tilde{P}_ε) |
| V_h | (2.4.1b) |
| V_q | (2.1.2) |
| $W^{k,p}(\Omega)$ | Sobolev space |
| $W_0^{k,p}(\Omega)$ | Sobolev space with vanishing boundary condition |
| W_h | space of piecewise constants function, (2.4.1c) |
| y, \bar{y} | optimal state of (2.1.3) |
| y^d | desired state of the optimal control problem, (2.1.3) |
| $y_\varepsilon, \bar{y}_\varepsilon$ | regularized optimal state of (4.1.1) |
| $y_{\varepsilon,h}$ | discrete regularized state, (4.4.4) |
| y_h, \bar{y}_h | discrete optimal state of (2.4.5) |
| y_h^d | $M_h y^d$ |
| $y(\cdot)$ | $G \cdot$ |
| $y_h(u)$ | discrete auxiliary state, (3.1.10) |
| $y_h(u_\varepsilon)$ | discrete auxiliary regularized state, (4.5.7) |
| $y(u_{\varepsilon,h})$ | auxiliary regularized state, (4.5.6a) |
| $y(u_h)$ | auxiliary state, (3.1.8) |

Other notations

| symbol | explanation and/or reference |
|----------------------------------|---|
| \cdot^* | dual space |
| $[\nabla \cdot]$ | jump across an interior edge |
| $\ \cdot\ _{k,p,\Omega}$ | Sobolev norm of the space $W^{k,p}(\Omega)$ |
| $\ \cdot\ _{k,\Omega}$ | $\ \cdot\ _{k,2,\Omega}$ |
| $ \cdot _{1,p,\Omega}$ | $\ \nabla \cdot\ _{0,p,\Omega}$ |
| $\ \cdot\ _{-1,\Omega}$ | dual norm in V^* |
| $(\cdot, \cdot)_{0,\Omega}$ | inner product of $L^2(\Omega)$ |
| $\langle \cdot, \cdot \rangle$ | dual pairing |
| $\langle \cdot, \cdot \rangle_h$ | (6.1.1) |
| \preceq | $\leq C$, with C only depending on the shape regularity of the triangulation |

Chapter 1

Introduction

A lot of models in physics and engineering lead to partial differential equations. However, most of them cannot be solved in an analytic way. With the development of computer science and the increase of computational power, discretization methods for the solution of partial differential equations have become more and more important during the last decades. The best known discretization schemes are the finite difference method, the finite volume method, and the finite element method. Among these methods, the latter one is the most common and its theory for solving partial differential equations is well advanced in development. Of particular interest are adaptive refinement strategies. If one wants to obtain the solution in an efficient way, i.e., with as less degrees of freedom and as less computational time as possible, one uses an adaptive finite element method based on an a posteriori error estimator. These error estimators should be reliable and efficient, i.e., they should not over- or underestimate the real error. With the help of adaptive finite element methods, those regions can be found where the solution is hard to approximate and a higher density of nodal points is necessary. A pioneering work has been done in [BR78]. Especially for elliptic partial differential equations, the analysis of adaptive finite element methods is well examined and even convergence results have been provided (cf. [Doe96], [MNS00], [MN07]). The same has been done for mixed and nonconforming finite element methods in [CH04] and [CH05], respectively. For obstacle problems, also some kind of maturity has been reached. A reliable and efficient a posteriori error estimator was, for example, proposed in [Vee01], and a convergence proof was provided in [BCH].

Optimal control problems governed by a partial differential equation also have a variety of applications in physics and engineering and therefore became an important part of research during the last decade. Thus, a lot of theoretical results are available for elliptic, parabolic, and hyperbolic partial differential equations. A great overview on this subject can be found in the pioneering book [Lio71] and in [Tro05]. The goal of an optimal control problem is that a state, which we will denote by y , should approximate a given desired state y^d as well as possible. The state can be adjusted by a control u . The objective functional, which should be minimized,

consists of one part which measures the approximation of the state to the desired state and an additional term which penalizes high costs of the control and which makes the problem convex and thus easier to handle. The side condition is the partial differential equation. In practice, there are often additional constraints on the control and/or the state, like an upper and/or a lower bound which should be fulfilled. Whether there are such additional constraints or not, makes a big difference for the theory and the numerics of the optimal control problem. If there are no additional constraints on the state or on the control, the optimality conditions consist of a system of partial differential equations. If there are constraints on the control and/or the state, variational inequalities appear in addition. In order to solve the optimality system of an optimal control problem, one also may discretize by means of the powerful finite element method. Then, the finite dimensional optimality system can be solved efficiently with a primal-dual active set strategy or an interior point method. An overview and a comparison of these methods can be found in [BHHK00], where the control constrained case and the state constrained case are investigated.

Even though there are a lot of applications of optimal control problems in practice, a posteriori error analysis is not as far developed as for partial differential equations. For optimal control problems with elliptic partial differential equations, there exist results for the control constrained case. In [HHIK07], [LLMT02], [LY01], and [LY03] reliable and efficient residual-type a posteriori error estimators have been derived, and, under some further assumptions, a convergence proof of the associated finite element method was given in [GHIK07]. For the unconstrained case (cf. [BR03], [BKR00]) and for the control constrained case (cf. [HH07]), the so-called goal-oriented weighted dual approach also has been investigated. Here, the adaptive method takes a specific quantity of interest into account, which is naturally the decrease of the objective functional in the case of optimal control problems.

On the other side, much less results are available if state constraints appear instead of control constraints. This is probably due to the fact that pointwise state constraints are much more difficult to handle since the involved Lagrange multipliers have less regularity (cf. [Cas86]). Unfortunately, the Lagrange multiplier associated with the pointwise state constraints is in general only a measure which causes additional problems for the analysis and the numerics. Some interesting properties of this measure have been studied in [BK02-2]. There, it is shown that under some additional assumptions, the singular part is concentrated on the boundary of the active set. A lot of the characteristics of a state constrained optimal control problem are thus different from those of a control constrained optimal control problem. Some features are somehow similar to those of an obstacle problem. However, they are far from being equal. One difference, for example, is that the optimal state of an optimal control problem is still two times weakly differentiable on compact subsets of the interior of the domain, while this is not necessarily the case for the solution of an obstacle problem if the obstacle is not smooth.

So far, a priori results have been derived (cf. [DH06]) for state constrained optimal control problems, i.e., convergence of the discrete control and the discrete state have been proved if the local mesh size tends globally to 0, and, with the help of a goal-oriented weighted dual approach, an adaptive algorithm was derived in [GH07]. According to the best of my knowledge, this thesis is the first approach to derive a residual-type a posteriori error estimator for state constrained optimal control problems. Such estimates are important in order to solve these kind of problems efficiently. It is worthwhile to mention that the a priori results known so far only provide convergence of the state and the control in the corresponding norms. Neither convergence of the Lagrange multiplier which eliminates the partial differential equation in a higher norm than the L^2 -norm nor convergence of the Lagrange multiplier associated with the pointwise state constraints in a measure norm have been provided. It is even assumed that one may only get weak*-convergence for the Lagrange multiplier associated with pointwise state constraints. The difficulty of the state constrained case for the numerics is that the algorithms which solve the optimality system, like, for example, the interior point methods or active set strategies, are only well-defined for the discrete case and not in the continuous setting. Consequently, the number of steps these iterative algorithms need to find the exact solution increases with the number of degrees of freedom (cf. [BK02], [BHHK00]).

An idea to overcome the difficulties arising from the nonsmoothness of the Lagrange multipliers was proposed in [MPT06] and [MRT06], where the pointwise state constraints were regularized by means of mixed control-state constraints. The advantage of this method is that now the Lagrange multipliers are smooth and have the same properties as for the pure control constrained case. The inventors of this idea are also able to prove that for a sufficiently small regularization parameter, the regularized optimal solution can be arbitrary close to the unregularized exact solution. An additional advantage is that the existing methods for solving the optimality conditions of the regularized problem are now also defined for the continuous problem. In [MPT06] an interior point method and an primal-dual active set strategy are investigated, and from the numerics of this paper, one may assume that the regularization is comparable with the pure state constrained case. Even though this regularization provides similar features as the control constrained case, it is my belief that the a posteriori error analysis known from the control constrained case can not be applied to the mixed control-state constraints since the involved functions blow up in the smooth norms, as the regularization parameter goes to 0. This is due to the fact that the smooth regularized Lagrange multipliers tend to be nonsmooth as the regularization parameter tends to 0. This seems rather undesirable because only a sufficiently small regularization parameter provides a good regularized solution, which is close to the unperturbed one. Thus, one is also in search of an useful a posteriori error estimator for this regularization method, which, according to my knowledge, has not been provided so far.

This thesis is concerned with the development, analysis, and implementation of an

adaptive finite element method based on a residual-type a posteriori error estimator for state constrained elliptic optimal control problems. In particular, we will follow two approaches which will provide two different a posteriori error estimators. The first one results from the viewpoint of control constraints, while the second one takes advantage of some similar properties as they appear for obstacle problems. It will also be shown that the analysis of the first error estimator can be transferred with minor modifications to the mixed control-state constrained case in such a way that the involved constants do not depend on the regularization parameter. In particular, this thesis is organized as follows:

In chapter two, a distributed elliptic optimal control problem for a two-dimensional, second order elliptic partial differential equation with a convex objective functional and unilateral constraints on the state variable will be formulated. This will serve as a model problem throughout the whole thesis. Under the assumption of a Slater Condition, optimality conditions will be derived, which imply complementarity conditions of a Karush-Kuhn-Tucker type. The assumptions on the domain will be a bit weaker than in [Cas86], for example. This will also allow domains with reentrant corners, which, concerning adaptive finite element methods, are of particular interest since the solution of partial differential equations on such domains often have singularities. The optimal control problem will be discretized by linear shape-regular conforming finite elements. The Lagrange multiplier associated with the pointwise state constraints will be approximated by a finite sum of Dirac-measures which are associated with the nodal points of the triangulation. This will lead to a discrete analogon of the continuous complementarity conditions.

In chapter three, the residual type a posteriori error estimator will be stated. The main idea of the analysis of this error estimator is the introduction of a modified adjoint state, which is smoother than the Lagrange multiplier associated with the partial differential equation. This allows to apply similar arguments as known from the control constrained case. As it will be shown, the error estimator provides reliability and efficiency for the approximation error in the state, in the control, and in the modified adjoint state up to data oscillation and a consistency error. The consistency error vanishes if the discrete control is exact.

Chapter four is concerned with mixed control-state constraints, the so-called Lavrentiev regularization. First, we introduce this idea, and then, the existence of smooth Lagrange multipliers will be shown. Furthermore, the convergence proof of the regularized problem to the unperturbed problem as the regularization parameter tends to 0 will be given. In order to get suitable discrete equations, we will not formulate a discrete optimal control problem, but discretize the continuous optimality system. However, for this case, the Lagrange multiplier associated with the control-state constraints will not be approximated by a sum of Dirac-measures, but also in the finite element space. Then, we will see that the error estimator known from the pure state constrained case can also be applied to the regularized problem in such a way that the involved constants do not blow up as the regularization parameter tends to 0. Reliability and efficiency will again be provided for the approximation error in

the state, in the control, and in the modified adjoint state. The use of the modified adjoint state will ensure that the constants do not depend on the regularization parameter. The consistency error associated with the reliability will consist of two parts of which the first one is similar to the one of the pure state constrained case, while the support of the second one is concentrated on the discrete free boundary and thus should decrease if one refines along the discrete free boundary in every iteration of the corresponding adaptive algorithm.

In the following chapter, a detailed documentation of two numerical examples will be given. Both methods, the one with pure state constraints and the mixed method, will be applied and also a comparison will be given. While the first example is a new one and involves smooth Lagrange multipliers, the second one is taken from [MPT06] and has the interesting feature that the Lagrange multipliers are indeed of less regularity. For both examples, one may notice an advantage of the adaptive finite element method compared to an uniform refinement strategy. For small regularization parameters, the Lavrentiev regularized solutions will be equal to the unperturbed ones. However, at least for these examples, no computational advantage of the regularization will be noticeable.

The idea to look at the state constrained optimal control problem from the viewpoint of obstacle problems will be investigated in chapter six. This will result in an error estimator for which the element residuals corresponding to the adjoint state equation involve an extension of the discrete adjoint control to an element in $L^2(\Omega)$. It will be shown that this error estimator provides, up to a consistency error and data oscillation, reliability and efficiency for the approximation of the state, of the control, and of an auxiliary adjoint state, which changes in every iteration. Finally, a further numerical example will be stated for which the error estimator with measure extension provides a benefit compared to the error estimator of chapter three.

Chapter 2

The distributed optimal control problem with state constraints

In this chapter, we state the distributed elliptic optimal control problem with pointwise state constraints. It will be shown that it is well posed in the sense that it has a unique optimal solution. Furthermore, a possible practical application will be presented briefly. We derive optimality conditions for our model problem, which assume less regularity on the domain than in [Cas86], for example. Thus, the advantage is that our analysis includes more general domains. The optimality conditions will also imply complementarity conditions for the Lagrange multiplier associated with the state constraints. Finally, we will present the applied approximation scheme. The discretization is chosen in such a way that we get a discrete analogon to the continuous complementarity conditions, which will be useful in the error estimates of the forthcoming chapter.

2.1 The model problem and existence and uniqueness of an optimal solution

Before we define our model problem, let us first introduce notation as it will be used in the whole thesis. In the following, we assume that Ω is an open, polygonal domain in \mathbb{R}^2 with boundary $\partial\Omega$. Standard notation from Lebesgue and Sobolev space theory will be used. In particular, we refer to $W^{k,p}(\Omega)$ with $k \in \mathbb{N}_0$ and $1 < p < \infty$ as the Sobolev spaces with norms $\|\cdot\|_{k,p,\Omega}$. For $k = 0$, we will shortly write $L^p(\Omega)$, as usual for Lebesgue spaces. If $p = 2$, we shortly write $H^k(\Omega)$ instead of $W^{k,2}(\Omega)$ and we use $(\cdot, \cdot)_{0,\Omega}$ as the inner product in the Hilbert space $L^2(\Omega)$ and $\|\cdot\|_{k,\Omega}$ instead of $\|\cdot\|_{k,2,\Omega}$. For $k = 1$, we refer to $|\cdot|_{1,p,\Omega} := \|\nabla \cdot\|_{0,p,\Omega}$ as the associated seminorm on $W^{1,p}(\Omega)$, which actually is a norm on $W_0^{1,p}(\Omega)$. Here, $W_0^{1,p}(\Omega)$ denotes the closure of $C_0^\infty(\Omega)$ w.r.t. the $\|\cdot\|_{1,p,\Omega}$ -norm, and $C_0^\infty(\Omega)$ stands for the infinitely often differentiable functions with compact support in Ω . If we consider functions in $W^{k,p}(D)$ or the restriction of functions in $W^{k,p}(\Omega)$ on D , with $D \subset \Omega$,

we write $\|\cdot\|_{k,p,D}$ and $|\cdot|_{k,p,D}$, as well as $(\cdot, \cdot)_{0,D}$ if $k = 0$ and $p = 2$. The Banach space of continuous functions on $\overline{\Omega}$ will be denoted by $C(\overline{\Omega})$, and its dual $C(\overline{\Omega})^*$ is isomorphic to the space of Radon measures $\mathcal{M}(\Omega)$ (compare [Alt02], p. 170). $\langle \cdot, \cdot \rangle$ stands for the associated dual pairing, and we refer to $C_+(\overline{\Omega})$ and $\mathcal{M}_+(\Omega)$ as the positive cones of $C(\overline{\Omega})$ and $\mathcal{M}(\Omega)$, respectively. In particular, $\mu \in \mathcal{M}_+(\Omega)$, if and only if $\langle \mu, v \rangle \geq 0$ for all $v \in C_+(\Omega)$.

With the above declared notation at hand, we are now able to formulate our model problem. To make things easier and to be able to concentrate on the difficulties arising from the appearance of the state constraints, we have chosen a simple linear partial differential equation with Dirichlet boundary conditions on part of the boundary of the domain to describe the dependency of the state on the control. In particular, the state y should be determined as follows:

For a given $v \in L^2(\Omega)$, we search $y \in H^1(\Omega)$ as the unique weak solution to the following elliptic partial differential equation:

$$-\Delta y + cy = v \text{ in } \Omega \quad (2.1.1a)$$

$$y|_{\Gamma_D} = 0 \text{ ,} \quad (2.1.1b)$$

where $\Gamma_D \subset \partial\Omega$. In order to ensure that (2.1.1) is uniquely solvable, we assume that $c > 0$ or $meas_1(\Gamma_D) > 0$, where $meas_1$ denotes the usual onedimensional Lebesgue measure. Let us assume the set $\Omega \subset \mathbb{R}^2$ fulfills some extra regularity such that the unique weak solution $y \in H^1(\Omega)$ of (2.1.1) is for all $v \in L^2(\Omega)$ also in $W^{1,r}(\Omega)$ for a $r > 2$. According to an embedding theorem (compare [Alt02], p. 317), this ensures that y is also a continuous function, i.e., $y \in C(\overline{\Omega})$. This additional assumption gives us the opportunity to demand pointwise constraints on the state. Notice that this also allows some nonconvex domains, for example such with reentrant corners (cf. [Gri85]). This is an important advantage since the solutions of partial differential equations on such domains often have singularities, and then, the adaptive finite element method provides a real benefit compared to an uniform mesh. However, domains with a crack are excluded. If one considers a higher dimension d , one has to require $r > \frac{d}{2}$ to get a continuous function as a solution of (2.1.1). Higher dimensions, however, should not be part of this thesis.

For a more practical notation, let us define for all $1 < q < \infty$ the set $V_q \subset W^{1,q}(\Omega)$ by

$$V_q := \{v \in W^{1,q}(\Omega) \mid v|_{\Gamma_D} = 0\} \quad (2.1.2)$$

and set $V := V_2$.

Let us now consider the following optimal control problem for a linear second order

elliptic boundary value problem with pointwise state constraints:

$$\text{minimize } J(y, u) := \frac{1}{2} \|y - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u - u^d\|_{0,\Omega}^2 \quad (2.1.3a)$$

$$\text{over } (y, u) \in V \times L^2(\Omega) ,$$

$$\text{subject to } -\Delta y + cy = u , \quad (2.1.3b)$$

$$Iy \in K := \{v \in C(\overline{\Omega}) \mid v(x) \leq \psi(x) \ \forall x \in \Omega\} . \quad (2.1.3c)$$

Here, I is the embedding operator from V to $C(\overline{\Omega})$, which is justified by the above assumptions. Moreover, it is assumed that:

$$y^d \in L^2(\Omega), \ u^d \in L^2(\Omega), \ \psi \in C(\overline{\Omega}), \ \psi|_{\Gamma_D} > 0, \ \alpha \in \mathbb{R}_{>0} . \quad (2.1.4)$$

J is called objective functional and ψ is called upper bound. Furthermore, y^d and u^d are referred to as the desired state and control shift, respectively. The first term of the objective functional measures the approximation of the state to the desired state, while the second term measures the approximation of the control to the control shift and can be interpreted as a cost term. The parameter $\alpha > 0$, which is typically small, regulates the ratio of the cost term to the approximation term and may be interpreted as a Tikhonov regularization. A vanishing α makes things more difficult since the optimal control problem loses its coercivity and should not be part of this thesis.

Because the partial differential equation is uniquely solvable and fulfills extra regularity, one may define the so called control-to-state mapping $G : L^2(\Omega) \rightarrow C(\overline{\Omega})$ which maps each function $v \in L^2(\Omega)$ to the solution of (2.1.1) with right-hand side v . It is a well-known fact that G is linear and continuous. In order to formulate suitable optimality conditions, we furthermore assume that a Slater Condition is fulfilled, i.e.:

$$\exists v_0 \in L^2(\Omega) \text{ such that } Gv_0 \in \text{int}(K) , \quad (2.1.5)$$

where $\text{int}(K)$ refers to the interior of K according to the max-norm in $C(\overline{\Omega})$. Notice that this can only be fulfilled if we assume that $\psi|_{\Gamma_D} > 0$ in (2.1.4). The Slater Condition will turn out to be useful in order to gain suitable optimality conditions. Now we can formulate an obvious equivalent version of (2.1.3), where the state y has been eliminated, as follows

$$\begin{aligned} \min_{u \in U_{ad}} J_{red}(u) &:= \frac{1}{2} \|y(u) - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u - u^d\|_{0,\Omega}^2 \\ U_{ad} &= \{v \in L^2(\Omega) \mid (Gv)(x) \leq \psi(x) \ \forall x \in \Omega\} . \end{aligned} \quad (2.1.6)$$

Here $y(u) = Gu$.

Next, one has to ask the existence and uniqueness of a solution of (2.1.3). The answer will be given by the following theorem:

Theorem 2.1.1 (Existence and uniqueness of an optimal solution). *If the assumptions of (2.1.4) hold true, the optimal control problem (2.1.3) has a unique optimal solution $(\bar{y}, \bar{u}) \in H^1(\Omega) \times L^2(\Omega)$.*

Proof. Obviously, the reduced objective functional J_{red} is bounded from below by 0. This implies the existence of a minimizing sequence $\{u_n\}_{n \in \mathbb{N}} \subset U_{ad}$ for (2.1.6), which fulfills

$$\lim_{n \rightarrow \infty} J_{red}(u_n) = \inf_{u \in U_{ad}} J_{red}(u) \quad (2.1.7)$$

With the help of Cauchy's and Young's inequality, it is easy to check that the following inequality has to hold:

$$J_{red}(u) \geq \frac{\alpha}{2} \|u - u^d\|_{0,\Omega}^2 \geq \frac{\alpha}{4} \|u\|_{0,\Omega}^2 - \frac{\alpha}{2} \|u^d\|_{0,\Omega}^2. \quad (2.1.8)$$

Since $\alpha > 0$, one may derive that the minimizing sequence $\{u_n\}_{n \in \mathbb{N}}$ has to be bounded. Because $L^2(\Omega)$ is a Hilbert space and thus particularly reflexive, a bounded sequence in $L^2(\Omega)$ has a weakly convergent subsequence, and we may conclude w.l.o.g.

$$u_n \rightharpoonup u^* \quad \text{as } n \rightarrow \infty, \quad (2.1.9)$$

where $u^* \in L^2(\Omega)$. Because of the fact that U_{ad} is closed and convex, it is also weakly closed, which implies $u^* \in U_{ad}$. The last step is now to show that u^* is indeed the optimal solution. The reduced objective functional J_{red} is convex and lower semi-continuous and thus weakly lower semi-continuous (compare [ET99]). Therefore, we may conclude

$$J_{red}(u^*) \leq \liminf_{n \in \mathbb{N}} J_{red}(u_n) \leq \inf_{u \in U_{ad}} J_{red}(u). \quad (2.1.10)$$

This implies that u^* is the optimal solution. The strict convexity of J_{red} implies the uniqueness of the optimal control and the existence and uniqueness of an optimal state y follows from the fact that the underlying partial differential equation (2.1.3b) is uniquely solvable. \square

It should be mentioned that for some cases, one may also show the existence and uniqueness of an optimal solution of the optimal control problem if the involved partial differential equation is not uniquely solvable. Even though this might be interesting from a mathematical point of view, we will, for practical reasons, not deal with this question in this thesis.

2.2 An illustrative example

Since I heard once the prejudice that every talk of a mathematician who deals with numerical analysis involves the heat equation, I also want to go this way and briefly present a possible application of the above stated problem (2.1.3). Assume that Ω is a body in which a specific stationary heat distribution y^d should be achieved as close

as possible. y describes this heat distribution and can be controlled by microwaves, whose intensity u may spatially differ over the body and can be regularized on the whole domain. The dependence of the heat distribution y on the microwaves can be described by means of a partial differential equation. Assume that the body is completely isolated, and therefore, one has to include homogeneous Neumann boundary conditions, i.e., we choose $\Gamma_D = \emptyset$ in (2.1.1b). Furthermore, the body Ω might consist of a specific material which gets damaged if the temperature is too high. This can be avoided if we include a constant upper bound ψ . Even though our main goal is to achieve y^d , the use of microwaves costs money and uses energy, and, of course, we do not want to spend too much on both of them. Thus, the objective functional consists of the term $\|y - y^d\|_{0,\Omega}^2$, which measures how close the temperature distribution is to the desired one, as well as of the term $\|u\|_{0,\Omega}^2$, which measures the costs ($u^d \equiv 0$). With the parameter α , we can adjust what is more important for us, the approximation of the desired temperature or the costs.

In most practical applications, the partial differential equation is, of course, more complicated and probably nonlinear in most cases. However, we use the Laplacian in our model since we want to concentrate on the difficulties arising from the state constraints and we do not want to deal with further severities of the nonlinearity. The derived results should then, with minor modifications, be adjustable to more complicated elliptic partial differential equations.

There are probably more interesting examples of applications for optimal control problems than the above stated one. However, they would be too complicated to be described even shortly, and thus, the interested reader is referred to the literature. For example, applications in crystal growths and in aerospace are given in [Mey06] and [Tro05], respectively.

2.3 Optimality conditions

In this section, we will derive optimality conditions for the optimal control problem with pointwise state constraints. The proof requires slightly less assumptions on the domain than the standard results, as given in [Cas86], for example.

Theorem 2.3.1 (Continuous optimality conditions). *If the assumptions (2.1.4) and the Slater Condition (2.1.5) are fulfilled, the unique optimal solution $(\bar{y}, \bar{u}) \in V \times L^2(\Omega)$ of Problem (2.1.3) is characterized by the existence of $\bar{p} \in V_s$ and $\bar{\sigma} \in \mathcal{M}(\Omega) \cong C(\bar{\Omega})^*$ such that there holds*

$$(\nabla \bar{y}, \nabla v)_{0,\Omega} + c(\bar{y}, v)_{0,\Omega} = (\bar{u}, v)_{0,\Omega} \quad \forall v \in V, \quad I\bar{y} \in K, \quad (2.3.1a)$$

$$(\nabla \bar{p}, \nabla v)_{0,\Omega} + c(\bar{p}, v)_{0,\Omega} = \langle \bar{\sigma}, v \rangle - (y^d - \bar{y}, v)_{0,\Omega} \quad \forall v \in V_r, \quad (2.3.1b)$$

$$\bar{p} + \alpha(\bar{u} - u^d) = 0, \quad (2.3.1c)$$

$$\langle \bar{\sigma}, y - \bar{y} \rangle \leq 0 \quad \forall y \in K, \quad (2.3.1d)$$

where s denotes the conjugate to r , i.e., $\frac{1}{s} + \frac{1}{r} = 1$.

We will refer to \bar{p} as the (optimal) adjoint state and to $\bar{\sigma}$ as the (optimal) adjoint control. One may interpret \bar{p} and $\bar{\sigma}$ as the Lagrange multipliers which eliminate the partial differential equation and the upper bound on the state, respectively. Notice that (2.3.1) is similar to the optimality system in the case of control constraints (compare [HHIK07]). There, however, the adjoint functions have higher regularity, i.e., $\bar{p} \in V$ and $\bar{\sigma} \in L^2(\Omega)$. This minor regularity is the main difficulty of state constrained optimal control problems.

Proof. The proof follows the lines of [Cas86]. However, there, the assumption $y \in H^2(\Omega)$ has to hold. Thus, we will repeat the proof here.

Let us first define $I_K : C(\bar{\Omega}) \rightarrow \mathbb{R} \cup \{\infty\}$ as the indicator function of the set K , i.e.,

$$I_K(v) := \begin{cases} 0, & \text{if } v \in K \\ +\infty, & \text{otherwise} \end{cases} . \quad (2.3.2)$$

Then, the original problem (2.1.3) can be formulated equivalently as the following unconstrained optimality problem:

$$\inf_{v \in L^2(\Omega)} \hat{J}(v) := J_{red}(v) + (I_K \circ G)(v) , \quad (2.3.3)$$

for which we search the optimal solution $\bar{u} \in L^2(\Omega)$, which fulfills

$$\hat{J}(\bar{u}) = \inf_{v \in L^2(\Omega)} \hat{J}(v) , \quad (2.3.4)$$

what is equivalent to the statement

$$0 \in \partial \hat{J}(\bar{u}) , \quad (2.3.5)$$

where $\partial \hat{J}(\bar{u})$ denotes the subdifferential of the convex function \hat{J} at \bar{u} . Since the Slater Condition (2.1.5) holds, one may derive by applying methods from ([Roc70]) that (2.3.5) is equivalent to:

$$0 \in J'_{red}(\bar{u}) + G^* \circ \partial I_K(G\bar{u}) , \quad (2.3.6)$$

which is equivalent to the existence of a $\bar{\sigma} \in \partial I_K(G\bar{u})$ such that

$$(y(\bar{u}) - y^d, y(v))_{0,\Omega} + \alpha(\bar{u} - u^d, v)_{0,\Omega} + (G^*\bar{\sigma}, v)_{0,\Omega} = 0 \quad \forall v \in L^2(\Omega) . \quad (2.3.7)$$

Let us now define $\hat{p} \in V$ as the unique solution of

$$(\nabla \hat{p}, \nabla w)_{0,\Omega} + c(\hat{p}, w)_{0,\Omega} = (y(\bar{u}) - y^d, w)_{0,\Omega} \quad \forall w \in V . \quad (2.3.8)$$

With this definition at hand, one may conclude by considering the definition of $y(\cdot) = G \cdot$ that

$$\begin{aligned} (y(\bar{u}) - y^d, y(v))_{0,\Omega} &= (\nabla \hat{p}, \nabla y(v))_{0,\Omega} + c(\hat{p}, y(v))_{0,\Omega} \\ &= (\hat{p}, v)_{0,\Omega} . \end{aligned} \quad (2.3.9)$$

Next, we set

$$\hat{\sigma} := G^* \bar{\sigma} . \quad (2.3.10)$$

From [Cas86], Theorem 4, it follows that $\hat{\sigma} \in V_q$ for any $1 < q < 2$ and especially that $\hat{\sigma} \in V_s$ with s being the conjugate to r . Therefore, (2.3.10) implies that there holds

$$(\nabla \hat{\sigma}, \nabla v)_{0,\Omega} + c(\hat{\sigma}, v)_{0,\Omega} = \langle \bar{\sigma}, v \rangle \quad \forall v \in V_r . \quad (2.3.11)$$

Now, we may define $\bar{p} := \hat{p} + \hat{\sigma} \in V_s$. Thus, (2.3.7), combined with (2.3.9) and (2.3.10), results in (2.3.1c). Furthermore, one may conclude for an arbitrary $w \in V_r$

$$\begin{aligned} (\nabla \bar{p}, \nabla w)_{0,\Omega} &= (\nabla \hat{p}, \nabla w)_{0,\Omega} + c(\hat{p}, w)_{0,\Omega} + (\nabla \hat{\sigma}, \nabla w)_{0,\Omega} + c(\hat{\sigma}, w)_{0,\Omega} \\ &= (y(\bar{u}) - y^d, w)_{0,\Omega} + \langle \bar{\sigma}, w \rangle, \end{aligned} \quad (2.3.12)$$

which results in (2.3.1b). Notice that $\bar{\sigma} \in \partial I_K(G\bar{u})$ is equivalent to (2.3.1d). \square

We will call $\hat{p} \in V \subset H^1(\Omega)$ the modified adjoint state because it has a higher regularity than the adjoint state \bar{p} . Analogously, we call $\hat{\sigma} \in W^{1,s}(\Omega)$ the modified adjoint control. Especially the modified adjoint state will play an important role in the forthcoming a posteriori error analysis.

For further purposes, we fix the continuous active set $\mathcal{A} \subset \Omega$ and the continuous inactive set $\mathcal{I} \subset \Omega$ according to

$$\mathcal{A} = \{x \in \Omega \mid \bar{y}(x) = \psi(x)\} , \quad (2.3.13a)$$

$$\mathcal{I} = \{x \in \Omega \mid \bar{y}(x) < \psi(x)\} . \quad (2.3.13b)$$

These sets are well-defined since \bar{y} and ψ are both continuous functions. Now, we can formulate the following corollary:

Corollary 2.3.2 (Continuous complementarity conditions). *From (2.3.1d), one may derive the following complementarity conditions (cf. [Cas86]):*

$$\bar{\sigma} \in \mathcal{M}_+(\Omega) \quad (2.3.14a)$$

$$\langle \bar{\sigma}, g \rangle = 0 \text{ for every } g \in C(\bar{\Omega}) \text{ with } g|_{\mathcal{A}} = 0 \quad (2.3.14b)$$

$$\langle \bar{\sigma}, \psi - \bar{y} \rangle = 0 \quad (2.3.14c)$$

These complementarity conditions can be interpreted as an infinitely dimensional version of the Karush-Kuhn-Tucker conditions.

2.4 Discretization and discrete optimality conditions

In order to solve the optimal control problem, we use a linear finite element approximation. Let us assume that $\mathcal{T}_h := \mathcal{T}_h(\Omega)$ is a shape-regular, simplicial, and

conforming triangulation of Ω , which aligns with Γ_D on $\partial\Omega$. We only use one grid for all the involved functions. There may be some situations for which different meshes for the different appearing functions make sense. This might for example be the case, if the state and the adjoint state have a substantial change in different areas of the domain. However, the existence of different meshes for the state, control, adjoint state, and adjoint control also requires projections between the different meshes, what probably often causes more additional effort than what is gained. In addition to that, it is more difficult to implement and to analyze. Thus, we decided to use only one grid for all the involved functions.

In all what follows, the sign ' \preceq ' can be replaced by ' $\leq C$ ' where C represents a constant which depends only on the shape regularity of the triangulation and is independent of the local mesh size. We refer to $\mathcal{N}_h(D)$ and $\mathcal{E}_h(D)$, $D \subseteq \overline{\Omega}$, as the sets of vertices and edges of $\mathcal{T}_h(\Omega)$ in $D \subseteq \overline{\Omega}$, respectively, and let $\mathcal{N}_h := \mathcal{N}_h(\overline{\Omega})$ and $\mathcal{E}_h := \mathcal{E}_h(\overline{\Omega})$. Then, $\mathcal{E}_h(\Omega)$ contains all the interior edges. We denote by h_T the diameter of an element $T \in \mathcal{T}_h(\Omega)$ and by h_E the length of an edge $E \in \mathcal{E}_h(D)$. Furthermore, let

$$S_h := \{v_h \in C(\overline{\Omega}) \mid w_{h|_T} \in P_1(T), T \in \mathcal{T}_h(\Omega)\} , \quad (2.4.1a)$$

$$V_h := \{v_h \in S_h \mid v_{h|_{\Gamma_D}} = 0\} , \quad (2.4.1b)$$

$$W_h := \{v_h \in L^2(\Omega) \mid v_{h|_T} \in P_0(T), T \in \mathcal{T}_h(\Omega)\} . \quad (2.4.1c)$$

S_h is the standard linear finite element space and V_h is its subspace which takes the boundary conditions into account. The space W_h consists of piecewise constant functions. Let furthermore $G_h : L^2(\Omega) \rightarrow V_h$ be the discrete analogon to G , i.e., for a given function $v \in L^2(\Omega)$, let $G_h(v)$ be the unique solution of the discrete partial differential equation, i.e.,

$$(\nabla G_h(v), \nabla v_h)_{0,\Omega} + c(G_h(v), v_h)_{0,\Omega} = (v, v_h)_{0,\Omega} \quad \forall v_h \in V_h . \quad (2.4.2)$$

Let u_h^d be some approximation of u^d in V_h . For continuous u^d , one can use, for example, the interpolant in the nodal points of the triangulation.

With these definitions at hand, we are now able to formulate the discrete optimal control problem as follows:

$$\text{minimize } J_h(y_h, u_h) := \frac{1}{2} \|y_h - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u_h - u_h^d\|_{0,\Omega}^2 \quad (2.4.3a)$$

$$\text{over } (y_h, u_h) \in V_h \times S_h , \quad (2.4.3b)$$

$$\text{subject to } y_h = G_h(u_h) , \quad (2.4.3c)$$

$$y_h \in K_h := \{v_h \in V_h \mid v_h(n) \leq \psi(n) \quad \forall n \in \mathcal{N}_h(\overline{\Omega})\} .$$

The existence and uniqueness of an optimal solution of the discrete problem (2.4.3) follows from similar arguments used in the proof of Theorem 2.1.1. Since we want to

gain a discrete analogon to the complementarity conditions (2.3.14), we will choose the discrete Lagrange multiplier $\sigma_h \in V_h^*$ associated with the state constraints in the space

$$\mathcal{M}_h := \{ \mu_h \in \mathcal{M}(\Omega) \mid \mu_h = \sum_{n \in \mathcal{N}_h(\bar{\Omega} \setminus \Gamma_D)} \mu_h^n \delta_n, \mu_h^n \in \mathbb{R} \}, \quad (2.4.4)$$

where δ_z denotes the Dirac-measure in the point z . In the sum of (2.4.4), one can exclude the nodes lying on the Dirichlet part of the boundary since, due to the Slater Condition (2.1.5), there has to hold $\psi|_{\Gamma_D} > 0$, and thus, the upper bound cannot be sharp in these points. This choice also seems to be natural if one considers that in general the adjoint control is a measure.

Now, we can formulate the discrete optimality conditions as follows:

Theorem 2.4.1 (Discrete optimality conditions). *The discrete optimal solution $(\bar{y}_h, \bar{u}_h) \in V_h \times S_h$ of (2.4.3) is characterized by the existence of $\bar{p}_h \in V_h$ and $\bar{\sigma}_h \in \mathcal{M}_h$ such that:*

$$(\nabla \bar{y}_h, \nabla v_h)_{0,\Omega} + c(\bar{y}_h, v_h)_{0,\Omega} = (\bar{u}_h, v_h)_{0,\Omega} \quad \forall v_h \in V_h, \bar{y}_h \in K_h \quad (2.4.5a)$$

$$(\nabla \bar{p}_h, \nabla v_h)_{0,\Omega} + c(\bar{p}_h, v_h)_{0,\Omega} = \langle \bar{\sigma}_h, v_h \rangle - (y^d - \bar{y}_h, v_h)_{0,\Omega} \quad \forall v_h \in V_h \quad (2.4.5b)$$

$$\bar{p}_h + \alpha(\bar{u}_h - u_h^d) = 0 \quad (2.4.5c)$$

$$\langle \bar{\sigma}_h, v_h - \bar{y}_h \rangle \leq 0 \quad \forall v_h \in K_h. \quad (2.4.5d)$$

Corresponding to the continuous case, we call \bar{p}_h the discrete (optimal) adjoint state and $\bar{\sigma}_h$ the discrete (optimal) adjoint control. The optimality conditions (2.4.5) may be solved, for example, with the help of an interior-point method or a primal-dual active set strategy (cf. [BHHK00]).

Proof. The claim can be proved by standard methods of nonlinear optimization (cf. [Lue84]). An interesting alternative proof was proposed in [DH06], which we will quote here for completeness.

The main idea is to introduce the following semidiscrete optimization problem:

$$\text{minimize } J_{red,h}(u) := \frac{1}{2} \|G_h u - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u - u_h^d\|_{0,\Omega}^2 \quad (2.4.6a)$$

$$\text{over } u \in L^2(\Omega),$$

$$\text{subject to } G_h u \in K_h := \{v_h \in V_h \mid v_h(n) \leq \psi(n) \quad \forall n \in \mathcal{N}_h(\bar{\Omega})\}. \quad (2.4.6b)$$

(2.4.6) is an infinitely dimensional optimization problem with only a finite number of constraints. The uniqueness and existence of a solution of (2.4.7) follows from similar arguments as used in the proof of Theorem 2.1.1. Thus, one may apply Theorem 5.2 from [Cas93] or similar arguments as used in the proof of Theorem 2.4.1. This results in the following optimality system:

The unique optimal solution $(\tilde{u}, \tilde{y}_h) \in L^2(\Omega) \times V_h$ ($\tilde{y}_h = G_h \tilde{u}$) of problem (2.4.6) is characterized by the existence of a $\tilde{p}_h \in V_h$ and a $\tilde{\sigma}_h \in \mathcal{M}_h$ such that:

$$(\nabla \tilde{y}_h, \nabla v_h)_{0,\Omega} + c(\tilde{y}_h, v_h)_{0,\Omega} = (\tilde{u}, v_h)_{0,\Omega} \quad \forall v_h \in V_h, y_h \in K_h \quad (2.4.7a)$$

$$(\nabla \tilde{p}_h, \nabla v_h)_{0,\Omega} + c(\tilde{p}_h, v_h)_{0,\Omega} = \langle \tilde{\sigma}_h, v_h \rangle - (y^d - \tilde{y}_h, v_h)_{0,\Omega} \quad \forall v_h \in V_h \quad (2.4.7b)$$

$$\tilde{p}_h + \alpha(\tilde{u} - u_h^d) = 0 \quad (2.4.7c)$$

$$\langle \tilde{\sigma}_h, v_h - \tilde{y}_h \rangle \leq 0 \quad \forall v_h \in K_h. \quad (2.4.7d)$$

From the equality (2.4.7c), one may now deduce that the optimal solution \tilde{u} is actually also an element of the space S_h . Thus, the optimal solution of (2.4.6) also has to be the optimal solution of (2.4.5), and (2.4.7) also has to be the optimality system if one minimizes only over the space S_h . \square

Since we will only deal with the optimal functions in the following, we drop the line above the optimal solutions.

Corollary 2.4.2 (Discrete complementarity conditions). *From (2.4.5d), one may easily derive the following discrete complementarity conditions, which hold for the representation $\sigma_h = \sum_{n \in \mathcal{N}_h} \sigma_h^n \delta_n$:*

$$\sigma_h^n \geq 0 \quad \forall n \in \mathcal{N}_h, \quad (2.4.8a)$$

$$\sigma_h^n = 0 \quad \text{if } y_h(n) < \psi(n), \quad (2.4.8b)$$

$$\langle \sigma_h, \psi - y_h \rangle = 0. \quad (2.4.8c)$$

Let us furthermore introduce discrete analogies to the modified functions \hat{p} and $\hat{\sigma}$ as defined in (2.3.8) and (2.3.11), respectively, i.e., let the discrete modified adjoint state $\hat{p}_h \in V_h$ be the unique solution of

$$(\nabla \hat{p}_h, \nabla v_h)_{0,\Omega} + c(\hat{p}_h, v_h)_{0,\Omega} = (y_h - y^d, v_h)_{0,\Omega} \quad \forall v_h \in V_h \quad (2.4.9)$$

and fix the discrete modified adjoint control $\hat{\sigma}_h \in V_h$ according to

$$\hat{\sigma}_h := p_h - \hat{p}_h. \quad (2.4.10)$$

Then, there holds:

$$(\nabla \hat{\sigma}_h, \nabla v_h)_{0,\Omega} + c(\hat{\sigma}_h, v_h)_{0,\Omega} = \langle \sigma_h, v_h \rangle \quad \forall v_h \in V_h. \quad (2.4.11)$$

After having solved the discrete problem, the natural question to be asked is how good the discrete solution is, i.e., how close it is to the exact continuous solution. A possible answer to this question for Neumann boundary conditions gives the following a priori result.

Theorem 2.4.3 (Convergence of the finite element approximation). *Assume that $\Gamma_D = \emptyset$ and that $(y, u) \in V \times L^2(\Omega)$ and $(y_h, u_h) \in V_h \times S_h$ are the optimal solutions of (2.1.3) and (2.4.1), respectively. Then, for every $\epsilon > 0$, there exists a constant $C = C(\epsilon)$ such that*

$$\|u - u_h\|_{0,\Omega} + \|y - y_h\|_{1,\Omega} \leq Ch^{1-\epsilon} . \quad (2.4.12)$$

Furthermore, assume that σ and σ_h denote the continuous and discrete adjoint control as they appear in the optimality systems (2.3.1) and (2.4.5), respectively. Then, there holds:

$$\sigma_h \rightharpoonup^* \sigma \text{ as } h \rightarrow 0 . \quad (2.4.13)$$

Proof. The proof can be found in [DH06]. □

In the following chapter, we want to give an answer to the question of the exactness of the discrete solution, which takes the discrete solution into account. This will result in an a posteriori error estimator, which gives rise to an adaptive refinement strategy.

Chapter 3

A posteriori error analysis

The goal of this chapter is to derive and analyze a residual-type a posteriori error estimator. It would be advantageous to gain an a posteriori error estimator which provides an upper and a lower bound for the approximation error in the state and in the control, as well as in the adjoint state and in the adjoint control, like it has been done for control constrained optimal control problems (compare [HHIK07]). However, due to the fewer regularity of the Lagrange multipliers, this seems to be very challenging, and it is assumed that even for a uniform refinement strategy, one only gets weak*-convergence for the adjoint control (compare Theorem 2.4.3). Furthermore, the existing a priori analysis does, according to my knowledge, not provide a convergence of the approximation error in the adjoint state in a higher norm than the $L^2(\Omega)$ -norm so far (cf. [DH06]). Thus, it seems to be a challenging task to derive such estimates with an a posteriori error analysis. To overcome these difficulties, the main idea of this chapter is to derive an a posteriori error estimator which is reliable and efficient for the approximation error in the state, in the control, and in the modified adjoint state. Due to its higher regularity, the latter one seems to be easier to handle compared to the original adjoint state. These estimations will be provided up to a consistency error, which vanishes if one already has reached the exact continuous solution.

3.1 The residual-type a posteriori error estimator

The residual-type a posteriori error estimator consists of easily computable element and edge residuals with respect to the finite element approximations $y_h \in V_h$ and $\hat{p}_h \in V_h$ of the state $y \in V$ and the modified adjoint state $\hat{p} \in V$ (or, more precisely, of the auxiliary state $y(u_h)$ and the auxiliary modified adjoint state $p(y_h)$, as they will be defined in (3.1.8) and (3.1.9), respectively), as well as of data oscillations. In particular, let us define

$$\eta_y := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \eta_{y,T}^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \eta_{y,E}^2 \right)^{\frac{1}{2}}, \quad (3.1.1a)$$

$$\eta_{\hat{p}} := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \eta_{\hat{p},T}^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \eta_{\hat{p},E}^2 \right)^{\frac{1}{2}}. \quad (3.1.1b)$$

Here, the element residuals $\eta_{y,T}$ and $\eta_{\hat{p},T}$ are weighted elementwise L^2 -residuals with respect to the strong form of the state equation (2.3.1a) and the modified adjoint state equation (2.3.8), respectively, and the edge residuals $\eta_{y,E}$ and $\eta_{\hat{p},E}$ are weighted L^2 -norms of the jumps $\boldsymbol{\nu}_E \cdot [\nabla y_h]$ and $\boldsymbol{\nu}_E \cdot [\nabla \hat{p}_h]$ of the normal derivatives across the interior edges. In particular, they read as follows:

$$\eta_{y,T} := h_T \|u_h - cy_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \quad (3.1.2a)$$

$$\eta_{\hat{p},T} := h_T \|y_h - y^d - c\hat{p}_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \quad (3.1.2b)$$

$$\eta_{y,E} := h_E^{\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \quad (3.1.2c)$$

$$\eta_{\hat{p},E} := h_E^{\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla \hat{p}_h]\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \quad (3.1.2d)$$

where $E = T_1 \cap T_2$ with $T_1, T_2 \in \mathcal{T}_h(\Omega)$, and $\boldsymbol{\nu}_E$ is the exterior unit normal vector on E directed towards T_2 , whereas $[\nabla y_h]$ and $[\nabla \hat{p}_h]$ denote the jumps of ∇y_h and $\nabla \hat{p}_h$ across E , i.e., for example, $[\nabla y_h] = \nabla y_h|_{T_1} - \nabla y_h|_{T_2}$.

The total residual-type error estimator η for the finite element approximation of the distributed control problem (2.1.3a)-(2.1.3c) is then given by

$$\eta := \left(\eta_y^2 + \eta_{\hat{p}}^2 \right)^{\frac{1}{2}}. \quad (3.1.3)$$

Moreover, we define the lower order data oscillation according to

$$\mu_h(u^d) := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \mu_T^2(u^d) \right)^{1/2}, \quad (3.1.4)$$

$$\mu_T(u^d) := \|u^d - u_h^d\|_{0,T},$$

as well as the higher order data oscillation

$$\text{osc}_h(y^d) := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \text{osc}_T^2(y^d) \right)^{1/2}, \quad (3.1.5)$$

$$\text{osc}_T(y^d) := h_T \|y^d - y_h^d\|_{0,T},$$

where

$$y_h^d := M_h y^d \quad (3.1.6)$$

and $M_h : L^2(\Omega) \rightarrow W_h$ refers to the L^2 -projection, i.e., it is the operator defined by

$$(M_h v)|_T := |T|^{-1} \int_T v(x) dx, \quad T \in \mathcal{T}_h, \quad v \in L^2(\Omega). \quad (3.1.7)$$

Remark 3.1.1. *Compared to the element residuals $\eta_{y,T}, \eta_{\hat{p},T}$ and the edge residuals $\eta_{y,E}, \eta_{\hat{p},E}$, the data oscillation $\text{osc}_h(y^d)$ is of the same order for nonsmooth y^d and of higher order for smooth y^d , i.e., $y^d \in H^1(\Omega)$. The lower order data oscillation $\mu_h(u^d)$ is only of the same order for $u^d \in H^1(\Omega)$.*

One may wonder why the data oscillations in y^d and in u^d are of different order. The data oscillation in y^d is equal to the one known from partial differential equations because y^d is part of the right-hand side of the modified adjoint state equation, while the data oscillation in u^d has another origin, as will be seen in the proof of the reliability of the error estimator in the forthcoming section.

As for the control constrained case, we define the auxiliary state $y(u_h) \in V$ as the continuous solution with the discrete right-hand side of the discrete optimal state, i.e.

$$(\nabla y(u_h), \nabla v)_{0,\Omega} + c(y(u_h), v)_{0,\Omega} = (u_h, v)_{0,\Omega} \quad \forall v \in V, \quad (3.1.8)$$

and the auxiliary modified adjoint state $\hat{p}(y_h) \in V$ as the continuous solution with the discrete right-hand side of the discrete modified adjoint state

$$(\nabla \hat{p}(y_h), \nabla v)_{0,\Omega} + c(\hat{p}(y_h), v)_{0,\Omega} = (y_h - y^d, v)_{0,\Omega} \quad \forall v \in V. \quad (3.1.9)$$

In addition to that, we introduce the discrete auxiliary state $y_h(u) \in V_h$ as the discrete solution with the continuous right-hand side of the state equation:

$$(\nabla y_h(u), \nabla v_h)_{0,\Omega} + c(y_h(u), v_h)_{0,\Omega} = (u, v_h)_{0,\Omega} \quad \forall v_h \in V_h. \quad (3.1.10)$$

Neither $y(u_h)$ nor $y_h(u)$ have to satisfy the continuous or the discrete state constraints, i.e., it may happen that $y(u_h) \notin K$ and $y_h(u) \notin K_h$. Therefore, we define the consistency error $e_c(u, u_h) \in \mathbb{R}$ according to

$$e_c(u, u_h) := \max(\langle \sigma, y(u_h) - \psi \rangle + \langle \sigma_h, y_h(u) - \psi \rangle, 0). \quad (3.1.11)$$

This consistency error term will appear in the forthcoming reliability results. Notice that $e_c(u, u_h) = 0$ if $u = u_h$ since this implies $y(u_h) = y$ and $y_h(u) = y_h$, and hence, the terms $\langle \sigma_h, y_h(u) - \psi \rangle$ and $\langle \sigma, y(u_h) - \psi \rangle$ vanish due to (2.4.8c) and (2.3.14c), respectively.

Remark 3.1.2. *Since the support of σ is concentrated on the continuous active zone \mathcal{A} and since σ is a positive measure, i.e., $\sigma \in C(\bar{\Omega})_+^*$, the term $\langle \sigma, y(u_h) - \psi \rangle$ penalizes if $y(u_h)$ violates the state constraints in the continuous active zone. Thus, the value of this term should be close to 0 if in the active region $y(u_h)$ is close to y , which fulfills the upper bound. It is even negative if $y(u_h) < \psi$ holds in the active zone. The same remark applies to the term $\langle \sigma_h, y_h(u) - \psi \rangle$ in a discrete sense.*

For the following estimations, let us notice that for the case $c > 0$, the existence of a constant \tilde{c} such that

$$\|v\|_{1,\Omega}^2 \leq \tilde{c}((\nabla v, \nabla v)_{0,\Omega} + c(v, v)_{0,\Omega}) \quad \forall v \in V \quad (3.1.12)$$

holds is obvious. Then, \tilde{c} depends on c . For the case $c = 0$, we have $meas_1(\Gamma_D) > 0$, due to our assumption that the underlying partial differential equation is well-posed. Then, the existence of a constant \tilde{c} such that (3.1.12) holds follows directly from the Poincaré-Friedrich-Inequality (cf. [Alt02], p. 155). For this case, \tilde{c} depends on the domain if Ω is not convex.

3.2 Reliability of the error estimator

The aim of this section is to prove the following theorem, which the reliability of the error estimator is the subject of. The reliability is given for the approximation error of the state in the $\|\cdot\|_{1,\Omega}$ -norm, of the control in the $\|\cdot\|_{0,\Omega}$ -norm, and of the modified adjoint state in the $\|\cdot\|_{1,\Omega}$ -norm up to the consistency error (3.1.11). Due to the fundamental relationships (2.3.1c) and (2.4.5c), this also leads to an upper bound for the L^2 -norm of the discretization error in the adjoint state. However, the theorem will not give any results about the approximation of the adjoint control, for which only weak*-convergence is assumed to hold true.

Theorem 3.2.1 (Reliability). *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let \hat{p} and \hat{p}_h be as defined in (2.3.8) and in (2.4.9), respectively. Furthermore, assume that η is the residual-type error estimator as given by (3.1.3), that $\mu_h(u^d)$ is the lower order data oscillation as given by (3.1.4), and that $e_c(u, u_h)$ denotes the consistency error as given by (3.1.11). Then, there exists a constant C , depending only on α , \tilde{c} , and the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that:*

$$\begin{aligned} & \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{1,\Omega}^2 + \|\hat{p} - \hat{p}_h\|_{1,\Omega}^2 + \|p - p_h\|_{0,\Omega}^2 \\ & \leq C(\eta^2 + \mu_h^2(u^d) + e_c(u, u_h)) \end{aligned} \quad (3.2.1)$$

The idea of the proof of Theorem 3.2.1 is similar to the control constrained case (compare [HHIK07]). First, we will show that the approximation of the state, of the control, and of the modified adjoint state can be bounded from above by the lower order data oscillation, the consistency error, the approximation of the discrete state to the auxiliary state, and the approximation of the discrete modified adjoint state to the auxiliary modified adjoint state. The two latter terms can then be handled with standard methods of a posteriori error analysis since the discrete state y_h and the discrete modified adjoint state \hat{p}_h are just the finite element approximations of the auxiliary state $y(u_h)$ and the auxiliary modified adjoint state $\hat{p}(y_h)$, respectively. As a first step into this direction, we will prove the following auxiliary result:

Lemma 3.2.2. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let \hat{p} and \hat{p}_h be as defined in (2.3.8) and in (2.4.9), respectively. Furthermore, let $y(u_h)$ and $\hat{p}(y_h)$ be the auxiliary state and the auxiliary modified adjoint state as given by (3.1.8) and (3.1.9), respectively. Then, there exists a constant C , depending only on \tilde{c} , such that:*

$$\begin{aligned} & \|y - y_h\|_{1,\Omega}^2 + \|\hat{p} - \hat{p}_h\|_{1,\Omega}^2 \\ & \leq C(\|y(u_h) - y_h\|_{1,\Omega}^2 + \|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2) \end{aligned} \quad (3.2.2)$$

Proof. With the help of the triangle inequality one may find

$$\|y - y_h\|_{1,\Omega}^2 \leq 2\|y - y(u_h)\|_{1,\Omega}^2 + 2\|y(u_h) - y_h\|_{1,\Omega}^2 \quad (3.2.3)$$

Thus, it remains to estimate the first term of (3.2.3). This can be done by applying (3.1.12), (2.3.1a), (3.1.8), and Cauchy's inequality as follows:

$$\begin{aligned} & \|y - y(u_h)\|_{1,\Omega}^2 \\ & \leq \tilde{c}((\nabla(y - y(u_h)), \nabla(y - y(u_h)))_{0,\Omega} + c(y - y(u_h), y - y(u_h))_{0,\Omega}) \\ & = \tilde{c}(u - u_h, y - y(u_h))_{0,\Omega} \\ & \leq \tilde{c}\|u - u_h\|_{0,\Omega}\|y - y(u_h)\|_{0,\Omega} \\ & \leq \tilde{c}\|u - u_h\|_{0,\Omega}\|y - y(u_h)\|_{1,\Omega}. \end{aligned} \quad (3.2.4)$$

Therefore, we may gain

$$\|y - y(u_h)\|_{1,\Omega} \leq \tilde{c}\|u - u_h\|_{0,\Omega}. \quad (3.2.5)$$

Thus, one may conclude from (3.2.3) and (3.2.5) that

$$\|y - y_h\|_{1,\Omega}^2 \leq 2\tilde{c}^2\|u - u_h\|_{0,\Omega}^2 + 2\|y_h - y(u_h)\|_{1,\Omega}^2 \quad (3.2.6)$$

For the second term of the left-hand side of (3.2.2), one proceeds in pretty much the same way. Again, an application of the triangle inequality results in

$$\|\hat{p} - \hat{p}_h\|_{1,\Omega}^2 \leq 2\|\hat{p} - \hat{p}(y_h)\|_{1,\Omega}^2 + 2\|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}^2 \quad (3.2.7)$$

For the first term of (3.2.7), we may obtain from (3.1.12), (2.3.8), (3.1.9), and Cauchy's and Young's inequality

$$\begin{aligned} & \|\hat{p} - \hat{p}(y_h)\|_{1,\Omega}^2 \\ & \leq \tilde{c}((\nabla(\hat{p} - \hat{p}(y_h)), \nabla(\hat{p} - \hat{p}(y_h)))_{0,\Omega} + c(\hat{p} - \hat{p}(y_h), \hat{p} - \hat{p}(y_h))_{0,\Omega}) \\ & = \tilde{c}(y - y^d - y_h + y^d, \hat{p} - \hat{p}(y_h))_{0,\Omega} \\ & \leq \tilde{c}\|y - y_h\|_{0,\Omega}\|\hat{p} - \hat{p}(y_h)\|_{0,\Omega} \\ & \leq \tilde{c}\|y - y_h\|_{0,\Omega}\|\hat{p} - \hat{p}(y_h)\|_{1,\Omega} \\ & \leq \frac{\tilde{c}^2}{2}\|y - y_h\|_{0,\Omega}^2 + \frac{1}{2}\|\hat{p} - \hat{p}(y_h)\|_{1,\Omega}^2 \end{aligned} \quad (3.2.8)$$

Thus, it follows from (3.2.7) and (3.2.8)

$$\|\hat{p} - \hat{p}_h\|_{1,\Omega}^2 \leq 2\tilde{c}^2 \|y - y_h\|_{0,\Omega}^2 + 2\|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}^2, \quad (3.2.9)$$

where the first term of the right-hand side can be estimated with the help of (3.2.6). Then, the combination of (3.2.6) and (3.2.9) gives the claimed result. \square

Next, we have to have a closer look at the approximation error of the control, which will be done in the following lemma.

Lemma 3.2.3. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, let \hat{p} and \hat{p}_h be as defined in (2.3.8) and in (2.4.9), respectively, and let η and $\mu_h(u^d)$ be the residual-type error estimator and the lower order data oscillation as given by (3.1.3) and (3.1.4), respectively. Furthermore, let $y(u_h)$ and $\hat{p}(y_h)$ be the auxiliary state and the auxiliary modified adjoint state as given by (3.1.8) and (3.1.9), respectively, and let $e_c(u, u_h)$ be the consistency error as given by (3.1.11). Then, there exists a constant C , depending only on \tilde{c} and α , such that:*

$$\|u - u_h\|_{0,\Omega}^2 \leq C(\|y(u_h) - y_h\|_{1,\Omega}^2 + \|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}^2 + \mu_h^2(u^d) + e_c(u, u_h)) \quad (3.2.10)$$

Proof. Since we can write $p = \hat{p} + \hat{\sigma}$ and $p_h = \hat{p}_h + \hat{\sigma}_h$, the fundamental relationships (2.3.1c) and (2.4.5c) lead to:

$$\begin{aligned} \|u - u_h\|_{0,\Omega}^2 &= (u - u_h, u^d - u_h^d)_{0,\Omega} + \frac{1}{\alpha}(u - u_h, \hat{\sigma}_h - \hat{\sigma})_{0,\Omega} \\ &\quad + \frac{1}{\alpha}(u - u_h, \hat{p}_h - \hat{p})_{0,\Omega} \end{aligned} \quad (3.2.11)$$

Let us first have a look at the second term of the right-hand side of (3.2.11). An application of the continuous equations (2.3.1a) and (3.1.8), the discrete equations (2.4.5a) and (3.1.10), and the modified control equations (2.3.11) and (2.4.11), as well as the complementarity conditions (2.3.14c) and (2.4.8c), gives us

$$\begin{aligned} &(u - u_h, \hat{\sigma}_h - \hat{\sigma})_{0,\Omega} \quad (3.2.12) \\ &= (\nabla(y_h(u) - y_h), \nabla\hat{\sigma}_h)_{0,\Omega} + c(y_h(u) - y_h, \hat{\sigma}_h)_{0,\Omega} \\ &\quad - (\nabla(y - y(u_h)), \nabla\hat{\sigma})_{0,\Omega} - c(y - y(u_h), \hat{\sigma})_{0,\Omega} \\ &= \langle \sigma_h, y_h(u) - y_h \rangle + \langle \sigma, y(u_h) - y \rangle \\ &= \langle \sigma_h, y_h(u) - \psi \rangle + \underbrace{\langle \sigma_h, \psi - y_h \rangle}_{=0} \\ &\quad + \langle \sigma, y(u_h) - \psi \rangle + \underbrace{\langle \sigma, \psi - y \rangle}_{=0} \\ &\leq e_c(u, u_h), \end{aligned}$$

where (2.3.11) was applicable because there holds $y, y(u_h) \in V_r$ due to the assumption on the domain Ω . The last term of the right-hand side (3.2.11) can be split up as follows:

$$\begin{aligned} & \frac{1}{\alpha}(u - u_h, \hat{p}_h - \hat{p})_{0,\Omega} \\ &= \frac{1}{\alpha}(u - u_h, \hat{p}_h - \hat{p}(y_h))_{0,\Omega} + \frac{1}{\alpha}(u - u_h, \hat{p}(y_h) - \hat{p})_{0,\Omega} . \end{aligned} \quad (3.2.13)$$

While we will apply Cauchy's and Young's inequality to the first term of the right-hand side of (3.2.13), which results in

$$\frac{1}{\alpha}(u - u_h, \hat{p}_h - \hat{p}(y_h))_{0,\Omega} \leq \frac{1}{8}\|u - u_h\|_{0,\Omega}^2 + \frac{2}{\alpha^2}\|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}^2 . \quad (3.2.14)$$

We estimate the second term with the help of (2.3.1a), (3.1.8), (2.3.8), (3.1.9), (3.2.5), and Cauchy's and Young's inequality as follows:

$$\begin{aligned} & \frac{1}{\alpha}(u - u_h, \hat{p}(y_h) - \hat{p})_{0,\Omega} \\ &= \frac{1}{\alpha}((\nabla(y - y(u_h)), \nabla(\hat{p}(y_h) - \hat{p}))_{0,\Omega} + c(y - y(u_h), \hat{p}(y_h) - \hat{p})_{0,\Omega}) \\ &= \frac{1}{\alpha}((y_h - y^d, y - y(u_h))_{0,\Omega} - (y - y^d, y - y(u_h))_{0,\Omega}) \\ &= \frac{1}{\alpha}(y_h - y, y - y(u_h))_{0,\Omega} \\ &= \frac{1}{\alpha}((y_h - y(u_h), y - y(u_h))_{0,\Omega} - \|y - y(u_h)\|_{0,\Omega}^2) \\ &\leq \frac{1}{\alpha}\|y(u_h) - y_h\|_{0,\Omega} \|y - y(u_h)\|_{0,\Omega} \\ &\leq \frac{1}{\alpha}\|y(u_h) - y_h\|_{1,\Omega} \|y - y(u_h)\|_{1,\Omega} \\ &\leq \frac{\tilde{c}}{\alpha}\|y(u_h) - y_h\|_{1,\Omega} \|u - u_h\|_{0,\Omega} \\ &\leq \frac{2\tilde{c}^2}{\alpha^2}\|y(u_h) - y_h\|_{1,\Omega}^2 + \frac{1}{8}\|u - u_h\|_{0,\Omega}^2 \end{aligned} \quad (3.2.15)$$

With a further application of Cauchy's and Young's inequality to the first term of the right-hand side of (3.2.11), we obtain

$$(u - u_h, u^d - u_h^d)_{0,\Omega} \leq \frac{1}{4}\|u - u_h\|_{0,\Omega}^2 + \|u^d - u_h^d\|_{0,\Omega}^2 . \quad (3.2.16)$$

Consequently, one concludes by combining (3.2.11)-(3.2.16) the existence of a constant C , depending on \tilde{c} and α , such that

$$\begin{aligned} \|u - u_h\|_{0,\Omega}^2 &\leq C(\|u^d - u_h^d\|_{0,\Omega}^2 + \|y(u_h) - y_h\|_{1,\Omega}^2 \\ &\quad + \|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}^2 + e_c(u, u_h)) , \end{aligned} \quad (3.2.17)$$

which is the result claimed. \square

In order to prove Theorem 3.2.1, it remains to estimate the two terms $\|y(u_h) - y_h\|_{1,\Omega}$ and $\|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega}$, which will be done in the following lemma.

Lemma 3.2.4. *Let $(y_h, p_h, u_h, \sigma_h)$ be the solution of (2.4.5) and let $y(u_h)$ be the corresponding auxiliary state as defined in (3.1.8). Furthermore, let \hat{p}_h be the discrete modified adjoint state as defined in (2.4.9), let $\hat{p}(y_h)$ be the corresponding auxiliary modified adjoint state as defined in (3.1.9), and let η_y and η_p be as defined in (3.1.1a) and (3.1.1b), respectively. Then, there exists a constant C , depending only on \tilde{c} and the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that*

$$\|y(u_h) - y_h\|_{1,\Omega} \leq C\eta_y, \quad (3.2.18a)$$

$$\|\hat{p}(y_h) - \hat{p}_h\|_{1,\Omega} \leq C\eta_{\hat{p}}. \quad (3.2.18b)$$

Proof. Since the right-hand sides of $y(u_h)$ and y_h and of $\hat{p}(y_h)$ and \hat{p}_h are the same, Galerkin orthogonality holds, and thus, the claim is standard in a posteriori error analysis (compare for example [Ver96]). \square

Proof of Theorem 3.2.1.

If we combine Lemma 3.2.2, Lemma 3.2.3, and Lemma 3.2.4, we may easily obtain

$$\|y - y_h\|_{1,\Omega}^2 + \|\hat{p} - \hat{p}_h\|_{1,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \leq \eta^2 + \mu_h^2(u^d) + e_c(u, u_h). \quad (3.2.19)$$

From (2.3.1c) and (2.4.5c) one easily derives

$$\|p - p_h\|_{0,\Omega}^2 \leq 2\alpha(\|u - u_h\|_{0,\Omega}^2 + \|u^d - u_h^d\|_{0,\Omega}^2). \quad (3.2.20)$$

The combination of (3.2.19) and (3.2.20) immediately gives the claim of Theorem 3.2.1. \square

3.3 Efficiency of the error estimator

This section is devoted to the proof that, up to data oscillation, the error estimator η also provides a lower bound for the approximation error in the state, in the control, and in the modified adjoint state. Similarly to the case of partial differential equations, one takes advantage of the so-called bubble functions (cf. [Ver96]). Let us first give some notations.

We denote by λ_i^T , $1 \leq i \leq 3$, the barycentric coordinates of $T \in \mathcal{T}_h(\Omega)$ and refer to $\vartheta_T := 27 \prod_{i=1}^3 \lambda_i^T$ as the associated element bubble function. Likewise, we define the edge bubble function ϑ_E on $\omega_E := T_1 \cup T_2$, $E = T_1 \cap T_2$, $T_\nu \in \mathcal{T}_h(\Omega)$, $1 \leq \nu \leq 2$, according to $\vartheta_{E|T_\nu} := 4 \prod_{i=1}^2 \lambda_i^{T_\nu}$, where $\lambda_1^{T_\nu}$ and $\lambda_2^{T_\nu}$ denote the barycentric coordinates of T_ν corresponding to the two nodes which are part of the edge E . Recall

from [Ver96] that there exist the constants $c_i, 1 \leq i \leq 5$, depending only on the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that for $\zeta_T \in P_k(T), k \in \mathbb{N}_0$, and $\zeta_E \in P_k(E), k \in \mathbb{N}_0$, there holds

$$\|\zeta_T\|_{0,T}^2 \leq c_1 (\zeta_T, \zeta_T \vartheta_T)_{0,T} \quad , \quad T \in \mathcal{T}_h(\Omega) \quad , \quad (3.3.1a)$$

$$\|\zeta_T \vartheta_T\|_{0,T} \leq c_2 \|\zeta_T\|_{0,T} \quad , \quad T \in \mathcal{T}_h(\Omega) \quad , \quad (3.3.1b)$$

$$|\zeta_T \vartheta_T|_{1,T} \leq c_3 h_T^{-1} \|\zeta_T\|_{0,T} \quad , \quad T \in \mathcal{T}_h(\Omega) \quad , \quad (3.3.1c)$$

$$\|\zeta_E\|_{0,E}^2 \leq c_4 (\zeta_E, \zeta_E \vartheta_E)_{0,E} \quad , \quad E \in \mathcal{E}_h(\Omega) \quad , \quad (3.3.1d)$$

$$\|\zeta_E \vartheta_E\|_{0,E} \leq c_5 \|\zeta_E\|_{0,E} \quad , \quad E \in \mathcal{E}_h(\Omega) \quad . \quad (3.3.1e)$$

For $E \in \mathcal{E}_h(T)$ and $\zeta_E \in P_k(E), k \in \mathbb{N}_0$, we further refer to $\tilde{\zeta}_E$ as the extension of ζ_E to ω_E in the sense that for fixed $E'_\nu \in \mathcal{E}_h(T_\nu) \setminus \{E\}$, for $x \in T_\nu$ we have $\tilde{\zeta}_E(x) := \zeta_E(x_E)$ where $x_E \in E$ is such that $x - x_E$ is parallel to E'_ν . Then, $\tilde{\zeta}_E \vartheta_E$ vanishes on the edges contained in the set $\mathcal{E}_h(\omega_E) \setminus E$. Again referring to [Ver96], there exist positive constants c_6 and c_7 which only depend on the shape regularity of $T \in \mathcal{T}_h(\Omega)$ and which fulfill

$$\|\tilde{\zeta}_E \vartheta_E\|_{0,\omega_E} \leq c_6 h_E^{1/2} \|\zeta_E\|_{0,E} \quad , \quad (3.3.2a)$$

$$|\tilde{\zeta}_E \vartheta_E|_{1,\omega_E} \leq c_7 h_E^{-1/2} \|\zeta_E\|_{0,E} \quad . \quad (3.3.2b)$$

With these tools at hand, one is now in a position to state and prove the following lemmas.

Lemma 3.3.1. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $\eta_{y,T}, T \in \mathcal{T}_h(\Omega)$, be given by (3.1.2a). Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for each $T \in \mathcal{T}_h(\Omega)$, there holds*

$$\eta_{y,T}^2 \leq C \left(\|y - y_h\|_{1,T}^2 + h_T^2 \|u - u_h\|_{0,T}^2 \right) \quad . \quad (3.3.3)$$

Proof. Setting $z_h := (u_h - cy_h)|_T \vartheta_T$, applying (3.3.1a) and Green's formula, observing that $\Delta y_h|_T = 0$, and the fact that z_h is an admissible test function in (2.3.1a) results in

$$\begin{aligned} & \eta_{y,T}^2 & (3.3.4) \\ &= h_T^2 \|u_h - cy_h\|_{0,T}^2 \\ &\leq c_1 h_T^2 (u_h - cy_h, z_h)_{0,T} \\ &= c_1 h_T^2 (u, z_h)_{0,T} + c_1 h_T^2 (\Delta y_h - cy_h, z_h)_{0,T} + c_1 h_T^2 (u_h - u, z_h)_{0,T} \\ &= c_1 h_T^2 (\nabla y, \nabla z_h)_{0,T} + c_1 c h_T^2 (y - y_h, z_h)_{0,T} - c_1 h_T^2 (\nabla y_h, \nabla z_h)_{0,T} \\ &\quad + c_1 h_T^2 (\nabla y_h \cdot \boldsymbol{\nu}_{\partial T}, \underbrace{z_h}_{=0})_{0,\partial T} + c_1 h_T^2 (u_h - u, z_h)_{0,T} \\ &= c_1 h_T^2 (\nabla(y - y_h), \nabla z_h)_{0,T} + c c_1 h_T^2 (y - y_h, z_h)_{0,T} + c_1 h_T^2 (u_h - u, z_h)_{0,T} \quad . \end{aligned}$$

Now, one can apply Cauchy's and Young's inequality and (3.3.1) to (3.3.4), which leads to

$$\begin{aligned}
& \eta_{y,T}^2 && (3.3.5) \\
\leq & c_1 h_T^2 |y - y_h|_{1,T} \underbrace{|z_h|_{1,T}}_{\leq c_3 h_T^{-1} \|u_h - cy_h\|_{0,T}} + c_1 c h_T^2 \|y - y_h\|_{0,T} \underbrace{\|z_h\|_{0,T}}_{\leq c_2 \|u_h - cy_h\|_{0,T}} \\
& + c_1 h_T^2 \|u - u_h\|_{0,T} \|z_h\|_{0,T} \\
\leq & \frac{3}{4} h_T^2 \|u_h - cy_h\|_{0,T}^2 + (c_1 c_3)^2 |y - y_h|_{1,T}^2 \\
& + (c_1 c_2 c)^2 h_T^2 \|y - y_h\|_{0,T}^2 + (c_1 c_2)^2 h_T^2 \|u - u_h\|_{0,T}^2,
\end{aligned}$$

from which we may conclude the result claimed. \square

Notice that the approximation error in the control appearing in the right-hand side of (3.3.3) has the same origin as the data oscillation known from the a posteriori error analysis for partial differential equations. Therefore, the local meshsize appears in front of this term in (3.3.3).

Lemma 3.3.2. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $\eta_{\hat{p},T}$ and $\text{osc}_T(y^d)$, $T \in \mathcal{T}_h(\Omega)$, be given by (3.1.2b) and (3.1.5), respectively. In addition to that, assume that \hat{p} and \hat{p}_h are given by (2.3.8) and (2.4.9), respectively. Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for each $T \in \mathcal{T}_h(\Omega)$, there holds*

$$\eta_{\hat{p},T}^2 \leq C \left(\|\hat{p} - \hat{p}_h\|_{1,T}^2 + h_T^2 \|y - y_h\|_{1,T}^2 + \text{osc}_T^2(y^d) \right). \quad (3.3.6)$$

Proof. An application of the triangle inequality yields

$$\begin{aligned}
\eta_{\hat{p},T}^2 &= h_T^2 \|y^d - y_h + c\hat{p}_h\|_{0,T}^2 && (3.3.7) \\
&\leq 2h_T^2 \|y^d - y_h^d\|_{0,T}^2 + 2h_T^2 \|y_h^d - y_h + c\hat{p}_h\|_{0,T}^2
\end{aligned}$$

By setting $z_h := (y_h^d - y_h + c\hat{p}_h)|_T \vartheta_T$, applying (3.3.1a) and Green's formula, and observing that $\Delta \hat{p}_h|_T = 0$ and that z_h is an admissible test function in (2.3.8), one may conclude

$$\begin{aligned}
& h_T^2 \|y_h^d - y_h + c\hat{p}_h\|_{0,T} && (3.3.8) \\
\leq & c_1 h_T^2 (y_h^d - y_h + c\hat{p}_h, z_h)_{0,T} \\
= & c_1 h_T^2 (y^d - y - \Delta \hat{p}_h, z_h)_{0,T} + c_1 c h_T^2 (\hat{p}_h, z_h)_{0,T} \\
& + c_1 h_T^2 (y_h^d - y^d, z_h)_{0,T} + c_1 h_T^2 (y - y_h, z_h)_{0,T} \\
= & c_1 h_T^2 (\nabla(\hat{p}_h - \hat{p}), \nabla z_h)_{0,T} - c_1 h_T^2 (\nabla \hat{p}_h \cdot \boldsymbol{\nu}_{\partial T}, \underbrace{z_h}_{=0})_{0,\partial T} \\
& + c c_1 h_T^2 (\hat{p}_h - \hat{p}, z_h)_{0,T} + c_1 h_T^2 (y_h^d - y^d, z_h)_{0,T} \\
& + c_1 h_T^2 (y - y_h, z_h)_{0,T}
\end{aligned}$$

Once again, one applies Cauchy's and Young's inequality, as well as the inequalities of (3.3.1) in order to gain

$$\begin{aligned}
& h_T^2 \|y_h^d - y_h + c\hat{p}_h\|_{0,T} \tag{3.3.9} \\
\leq & c_1 h_T^2 |\hat{p}_h - \hat{p}|_{1,T} \underbrace{\|z_h\|_{1,T}}_{\leq c_3 h_T^{-1} \|y_h^d - y_h + c\hat{p}_h\|_{0,T}} \\
& + c c_1 h_T^2 \|\hat{p}_h - \hat{p}\|_{0,T} \underbrace{\|z_h\|_{0,T}}_{\leq c_2 \|y_h^d - y_h + c\hat{p}_h\|_{0,T}} \\
& + c_1 h_T^2 \|y_h^d - y^d\|_{0,T} \|z_h\|_{0,T} + c_1 h_T^2 \|y - y_h\|_{0,T} \|z_h\|_{0,T} \\
\leq & \frac{1}{2} h_T^2 \|y_h^d - y_h + c\hat{p}_h\|_{0,T}^2 + 2(c_1 c_3)^2 |\hat{p}_h - \hat{p}|_{1,T}^2 \\
& + 2(cc_1 c_2)^2 h_T^2 \|\hat{p}_h - \hat{p}\|_{0,T}^2 + 2(c_1 c_2)^2 h_T^2 \|y^d - y_h^d\|_{0,T}^2 \\
& + 2(c_1 c_2)^2 h_T^2 \|y - y_h\|_{1,T}^2 .
\end{aligned}$$

The combination of (3.3.7) and (3.3.9) readily gives the result claimed. \square

Lemma 3.3.3. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $\eta_{y,T}$, $T \in \mathcal{T}_h(\Omega)$, and $\eta_{y,E}$, $E \in \mathcal{E}_h(\Omega)$, be given by (3.1.2a) and (3.1.2c), respectively. Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for every interior edge $E \in \mathcal{E}_h(\Omega)$, there holds*

$$\eta_{y,E}^2 \leq C \left(\|y - y_h\|_{1,\omega_E}^2 + h_E^2 \|u - u_h\|_{0,\omega_E}^2 + \sum_{\nu=1}^2 \eta_{y,T_\nu}^2 \right), \tag{3.3.10}$$

where T_1 and T_2 are the two triangles which have the edge E in common and $\omega_E := T_1 \cup T_2$.

Proof. Set $\zeta_E := (\boldsymbol{\nu}_E \cdot [\nabla y_h])|_E$ and $z_h := \tilde{\zeta}_E \vartheta_E$, where $\tilde{\zeta}_E$ denotes the extension of ζ_E to ω_E as described in the beginning of this section. Then, using (3.3.1d), applying Green's formula, and observing that $\Delta y_h|_T = 0$, that $z_h|_{\partial\omega_E} = 0$, and that z_h is an admissible test function in (2.3.1a), you can find

$$\begin{aligned}
\eta_{y,E}^2 &= h_E \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}^2 \tag{3.3.11} \\
&\leq c_4 h_E (\boldsymbol{\nu}_E \cdot [\nabla y_h], \zeta_E \vartheta_E)_{0,E} \\
&= c_4 h_E \sum_{\nu=1}^2 \{ (\boldsymbol{\nu}_{\partial T_\nu} \cdot \nabla y_h, z_h)_{0,\partial T_\nu} - (\Delta y_h, z_h)_{0,T_\nu} \} \\
&= c_4 h_E \left((\nabla(y_h - y), \nabla z_h)_{0,\omega_E} - c(y, z_h)_{0,\omega_E} \right. \\
&\quad \left. + (u - u_h, z_h)_{0,\omega_E} + (u_h, z_h)_{0,\omega_E} \right) \\
&= c_4 h_E \left((\nabla(y_h - y), \nabla z_h)_{0,\omega_E} + c(y_h - y, z_h)_{0,\omega_E} \right. \\
&\quad \left. + (u - u_h, z_h)_{0,\omega_E} + (u_h - cy_h, z_h)_{0,\omega_E} \right) .
\end{aligned}$$

Applying Cauchy's inequality to (3.3.11) and taking advantage of (3.3.2a) and (3.3.2b), one arrives at

$$\begin{aligned}
\eta_{y,E}^2 &\leq c_4 h_E |y - y_h|_{1,\omega_E} \underbrace{\|z_h\|_{1,\omega_E}}_{\leq c_7 h_E^{-\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}} & (3.3.12) \\
&\quad + c c_4 h_E \|y - y_h\|_{0,\omega_E} \underbrace{\|z_h\|_{0,\omega_E}}_{\leq c_6 h_E^{\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}} \\
&\quad + c_4 h_E \|u - u_h\|_{0,\omega_E} \|z_h\|_{0,\omega_E} \\
&\quad + c_4 h_E \|u_h - cy_h\|_{0,\omega_E} \|z_h\|_{0,\omega_E}
\end{aligned}$$

An application of Young's inequality and the fact that, due to the shape regularity of the triangulation, h_{T_ν} and h_E are of comparable size, i.e., $h_{T_\nu} \preceq h_E \preceq h_{T_\nu}$, allows to conclude the existence of a constant C , being independent of the local mesh size, such that

$$\eta_{y,E}^2 \leq \frac{1}{2} h_E \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}^2 + 2(c_4 c_7)^2 |y - y_h|_{1,\omega_E}^2 \quad (3.3.13)$$

$$+ 2(c_4 c_6)^2 h_E^2 \|y - y_h\|_{0,\omega_E}^2 + 2(c_4 c_6)^2 h_E^2 \|u - u_h\|_{0,\omega_E}^2 \quad (3.3.14)$$

$$+ C \sum_{\nu=1}^2 h_{T_\nu}^2 \|u_h - cy_h\|_{0,T_\nu}^2,$$

which gives the assertion (3.3.10). \square

Lemma 3.3.4. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, let \hat{p} and \hat{p}_h be as defined in (2.3.8) and (2.4.9), respectively, and let $\eta_{\hat{p},T}$, $T \in \mathcal{T}_h(\Omega)$, and $\eta_{\hat{p},E}$, $E \in \mathcal{E}_h(\Omega)$, be given by (3.1.2b) and (3.1.2d), respectively. Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for every interior edge $E \in \mathcal{E}_h(\Omega)$, there holds*

$$\eta_{\hat{p},E}^2 \leq C \left(\|\hat{p} - \hat{p}_h\|_{1,\omega_E}^2 + h_E^2 \|y - y_h\|_{1,\omega_E}^2 + \sum_{\nu=1}^2 \eta_{\hat{p},T_\nu}^2 \right), \quad (3.3.15)$$

where T_1 and T_2 are the two triangles which share the edge E with each other and $\omega_E := T_1 \cup T_2$.

Proof. The claim follows by the same arguments as used in the proof of Lemma 3.3.3. \square

The combination of Lemma 3.3.1, Lemma 3.3.2, Lemma 3.3.3, and Lemma 3.3.4 tells us that local efficiency holds, i.e., the local parts of the error estimator can be bounded from above by the local approximation errors, up to data oscillation. As a summary, one may state the following global result:

Theorem 3.3.5 (Efficiency). *Assume that (y, u, p, σ) and $(y_h, u_h, p_h, \sigma_h)$ are the optimal solutions of (2.3.1) and (2.4.5), respectively. In addition to that, let \hat{p} and \hat{p}_h be the continuous and discrete modified adjoint state as defined in (2.3.8) and (2.4.9), respectively, and let η and $\text{osc}_h(y^d)$ be the error estimator and data oscillation as given by (3.1.3) and (3.1.5), respectively. Then, there holds*

$$\eta \preceq \|y - y_h\|_{1,\Omega} + \|\hat{p} - \hat{p}_h\|_{1,\Omega} + \|u - u_h\|_{1,\Omega} + \text{osc}_h(y^d), \quad (3.3.16)$$

$$\eta \preceq \|y - y_h\|_{1,\Omega} + \|\hat{p} - \hat{p}_h\|_{1,\Omega} + \|u - u_h\|_{1,\Omega} + \|p - p_h\|_{0,\Omega} + \text{osc}_h(y^d). \quad (3.3.17)$$

Proof. The combination of the Lemma 3.3.1, Lemma 3.3.2, Lemma 3.3.3, and Lemma 3.3.4 and considering that summing up over all edges in (3.3.10) and (3.3.15) only gives a finite overlap, allows to conclude that (3.3.16) holds. Obviously, (3.3.17) follows directly from (3.3.16). \square

Remark 3.3.6. *Now the combination of Theorem 3.2.1 and Theorem 3.3.5 tells us that, up to data oscillation and consistency error, the real error, i.e., the approximation error in the state, in the control, and in the modified adjoint state, is neither under- nor overestimated by the error estimator. Furthermore, Lemma 3.3.1, Lemma 3.3.2, Lemma 3.3.3, and Lemma 3.3.4 tell us that even locally the error is not overestimated. This justifies that in the adaptive algorithm derived below, only triangles and edges with relatively large element and edge residuals will be refined.*

Chapter 4

Lavrentiev regularization of the state constraints

In [MPT06], [Tro05-2], and [MRT06] an interesting concept to get smoother Lagrange multipliers was proposed. The main idea is to replace the pure state constraints of the form $y \leq \psi$ by a regularization of the type $y_\varepsilon + \varepsilon u_\varepsilon \leq \psi$ for a small regularization parameter ε . This method is called Lavrentiev regularization. The main advantage is that now the adjoint state and the adjoint control are as regular as in the pure control constrained case, i.e., $\sigma_\varepsilon \in L^2(\Omega)$ and $p_\varepsilon \in H^1(\Omega)$. Thus, numerical methods known from the pure control constrained case, like active set strategies or interior-point methods, for example, can be applied after minor modifications and are now also well-defined in the continuous setting. An additional profit is that we are not restricted to continuous upper bounds, as for pure state constraints. It is also possible to apply an a posteriori error analysis like it has been derived for the control constrained case, as in [HHIK07], for example. However, this does not seem to be very meaningful since for that analysis, the constants blow up as the regularization parameter ε tends to 0. This is simply due to the fact that the unregularized Lagrange multipliers p and σ can, in general, only be chosen in the spaces $W^{1,s}(\Omega)$ with $1 < s < 2$ and $\mathcal{M}(\Omega)$, respectively. Even though the regularized adjoint state and adjoint control are smooth, they tend to those unsmooth unregularized functions at the limit and are thus hard to approximate in the smooth norms if the regularization parameter is small.

In this chapter, it will be shown that if one looks at the Lavrentiev regularization from the point of view of pure state constraints, one may derive an a posteriori error analysis for which the involved constants do not blow up as the regularization parameter tends to 0. The applied ideas are similar to those used in the previous chapter. The drawback compared to pure control constraints will again be that we only get results for the approximation error in the state, in the control, and in the modified adjoint state.

4.1 Mixed control-state constraints

We use the same model problem as in (2.1.3), despite that we now replace the pure state constraints by mixed control-state constraints. In particular, we define Problem (P_ε) as follows:

$$\text{minimize } J(y_\varepsilon, u_\varepsilon) := \frac{1}{2} \|y_\varepsilon - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u_\varepsilon - u^d\|_{0,\Omega}^2 \quad (4.1.1a)$$

$$\text{over } (y_\varepsilon, u_\varepsilon) \in V \times L^2(\Omega) ,$$

$$\text{subject to } -\Delta y_\varepsilon + cy_\varepsilon = u_\varepsilon , \quad (4.1.1b)$$

$$y_\varepsilon + \varepsilon u_\varepsilon \in \hat{K} := \{v \in L^2(\Omega) \mid v(x) \leq \psi(x) \text{ f.a.a. } x \in \Omega\} , \quad (4.1.1c)$$

where the data fulfill the same assumptions as declared in (2.1.4). However, in (4.1.1), it is enough to demand $\psi \in L^\infty(\Omega) \subset L^2(\Omega)$ since, due to the fact that $u_\varepsilon \in L^2(\Omega)$, we only can require the upper bound to be fulfilled almost everywhere anyway. In this chapter, we will show that Problem (P_ε) has an unique optimal solution, and we will derive optimality conditions, which involve the existence of smooth Lagrange multipliers. The proof that the optimal regularized state y_ε and the optimal regularized control u_ε converge to the unregularized optimal solutions will also be provided in this chapter. In order to prove existence and uniqueness, we proceed in the same way as for the pure state constrained case. Again, we refer to $G : L^2(\Omega) \rightarrow V$ as the control-to-state mapping which assigns to a given control $u_\varepsilon \in L^2(\Omega)$ the unique solution $y_\varepsilon = y(u_\varepsilon)$ of (4.1.1b). Then, we may formulate the reduced objective functional according to

$$J_{red}(u) := \frac{1}{2} \|Gu - y^d\|_{0,\Omega}^2 + \frac{1}{2} \|u - u^d\|_{0,\Omega}^2 \quad (4.1.2)$$

For simplicity, we will use $J(\cdot) := J_{red}(\cdot)$ in this chapter. This should not cause any confusions. Problem (4.1.1) can be equivalently formulated as

$$\inf_{Gu_\varepsilon + \varepsilon u_\varepsilon \in \hat{K}} J(u_\varepsilon) . \quad (4.1.3)$$

Now, we may easily prove the following theorem.

Theorem 4.1.1 (Existence and uniqueness). *Problem (P_ε) , as defined in (4.1.1), has an unique optimal solution $(\bar{y}_\varepsilon, \bar{u}_\varepsilon) \in V \times L^2(\Omega)$.*

Proof. The proof follows the same arguments as used in the proof of Theorem 2.1.1. \square

The following two sections are a repetition of previously established results for Lavrentiev regularized state constraints. Readers with knowledge of mixed control-state constraints or readers who are mainly interested in the a posteriori error analysis may skip these parts.

4.2 Optimality conditions of the regularized problem

In this section, we will derive the optimality conditions for the regularized problem (P_ε) . We will see that the involved Lagrange multipliers have the same smoothness as in the pure control constrained case. The arguments follow the lines of [Tro05-2].

Therefore, let us first define

$$R := (\varepsilon Id + G)^{-1} : L^2(\Omega) \rightarrow L^2(\Omega) , \quad (4.2.1)$$

where we now interpret G as the solution operator of the elliptic partial differential equation (2.1.1) with image in $L^2(\Omega)$. G is compact, and thus, the theory of Fredholm operators combined with the fact that G is positive definite tells us that G does not have any negative eigenvalues. As a consequence, (4.2.1) is well-defined. Then, we may fix

$$\tilde{J}(v) := J(Rv) \quad (4.2.2)$$

and define the auxiliary problem (\tilde{P}_ε) according to

$$\inf_{v_\varepsilon \in \tilde{K}} \tilde{J}(v_\varepsilon) . \quad (4.2.3)$$

Obviously, \tilde{P}_ε has an unique optimal solution, which we will denote by \bar{v}_ε . This is inferred from the same arguments used in the proof of Theorem 4.1.1. Now, it is easy to see that if u_ε is admissible for (P_ε) , i.e., if there holds $(\varepsilon Id + G)u_\varepsilon \leq \psi$, then $R^{-1}u_\varepsilon$ is admissible for the problem (\tilde{P}_ε) , i.e., there holds $R^{-1}u_\varepsilon \leq \psi$. Similarly, there holds that if v_ε is admissible for (\tilde{P}_ε) , then Rv_ε is admissible for Problem (P_ε) . With these simple observations at hand, one may now prove the following lemma.

Lemma 4.2.1. *The optimal solution \bar{v}_ε of Problem (\tilde{P}_ε) and the optimal solution \bar{u}_ε of Problem (P_ε) are related according to $\bar{u}_\varepsilon = R\bar{v}_\varepsilon$.*

Proof. Assume by contradiction that the claim is false, i.e., $\bar{v}_\varepsilon \neq R^{-1}\bar{u}_\varepsilon$. Since $R^{-1}\bar{u}_\varepsilon$ is admissible for (\tilde{P}_ε) and \tilde{J} is strictly convex, we get

$$\tilde{J}(\bar{v}_\varepsilon) < \tilde{J}(R^{-1}\bar{u}_\varepsilon) , \quad (4.2.4)$$

which is equivalent to

$$J(R\bar{v}_\varepsilon) < J(\bar{u}_\varepsilon) . \quad (4.2.5)$$

This is a contradiction because $R\bar{v}_\varepsilon$ is admissible for (P_ε) . Since both problems have an unique optimal solution, the claim follows. \square

As a consequence of Lemma 4.2.1, the following relationship between the Gâteaux derivatives of J and \tilde{J} is valid:

Lemma 4.2.2. *Assume that \bar{v}_ε is the optimal solution of (\tilde{P}_ε) and that \bar{u}_ε is the optimal solution of (P_ε) . Furthermore, let R and \tilde{J} be as defined in (4.2.1) and (4.2.2), respectively. Then, the following relationship holds:*

$$\tilde{J}'(\bar{v}_\varepsilon) = R^* J'(\bar{u}_\varepsilon) \quad (4.2.6)$$

Proof. On the one hand, one may find by means of a simple computation that for every $u \in L^2(\Omega)$, there holds

$$J'(\bar{u}_\varepsilon)(u) = (G(\bar{u}_\varepsilon) - y^d, G(u))_{0,\Omega} + \alpha(\bar{u}_\varepsilon - u^d, u)_{0,\Omega} . \quad (4.2.7)$$

After identifying $(L^2(\Omega))^*$ with $L^2(\Omega)$, what is justified by means of the Riesz representation theorem, one arrives at:

$$J'(\bar{u}_\varepsilon) = G^*(G(\bar{u}_\varepsilon) - y^d) - \alpha(\bar{u}_\varepsilon - u^d) . \quad (4.2.8)$$

On the other hand, one may easily find

$$\tilde{J}'(\bar{v}_\varepsilon)(v) = (GR\bar{v}_\varepsilon - y^d, GRv)_{0,\Omega} + \alpha(R\bar{v}_\varepsilon - u^d, Rv)_{0,\Omega} . \quad (4.2.9)$$

Then, by applying again the Riesz representation theorem to (4.2.9) and combining it with (4.2.8), one may conclude

$$\tilde{J}'(\bar{v}_\varepsilon) = R^* G^*(GR\bar{v}_\varepsilon - y^d) + \alpha R^*(R\bar{v}_\varepsilon - u^d) = R^* J'(\bar{u}_\varepsilon) , \quad (4.2.10)$$

where we also used the observation $\bar{u}_\varepsilon = R\bar{v}_\varepsilon$. \square

Before we state the existence of a smooth adjoint control $\sigma_\varepsilon \in L^2(\Omega)$, we fix the Lagrange functional corresponding to the problem (\tilde{P}_ε) as follows:

$$\tilde{L}(v, \sigma) := \tilde{J}(v) + (\sigma, v - \psi)_{0,\Omega} \quad (4.2.11)$$

With this definition at hand, we may state the following lemma.

Lemma 4.2.3. *The Problem (\tilde{P}_ε) has a smooth Lagrange multiplier, i.e., there exists a $\bar{\sigma}_\varepsilon \in L^2(\Omega)$ such that for the optimal solution \bar{v}_ε , the following system is fulfilled:*

$$\bar{\sigma}_\varepsilon \geq 0 \quad \text{a.e. in } \Omega , \quad (4.2.12a)$$

$$(\bar{\sigma}_\varepsilon, \bar{v}_\varepsilon - \psi)_{0,\Omega} = 0 , \quad (4.2.12b)$$

$$\frac{\partial}{\partial v} \tilde{L}(\bar{v}_\varepsilon, \bar{\sigma}_\varepsilon) = 0 . \quad (4.2.12c)$$

Proof. Because \bar{v}_ε is the optimal solution and \hat{K} is a convex set, the inequality

$$\tilde{J}'(\bar{v}_\varepsilon)(v - \bar{v}_\varepsilon) \geq 0 \quad \forall v \in \hat{K} \quad (4.2.13)$$

has to hold. Notice that $\tilde{J}'(\bar{v}_\varepsilon) \in (L^2(\Omega))^*$. Considering the Riesz representation lemma, one may interpret $\tilde{J}'(\bar{v}_\varepsilon) \in L^2(\Omega)$. Thus, one may derive from (4.2.13) the

following complementarity conditions, which have to be interpreted in the a.e. sense:

$$\tilde{J}'(\bar{v}_\varepsilon) = 0, \quad \text{if } \bar{v}_\varepsilon < \psi, \quad (4.2.14a)$$

$$\tilde{J}'(\bar{v}_\varepsilon) \leq 0, \quad \text{if } \bar{v}_\varepsilon = \psi. \quad (4.2.14b)$$

Now, we simply *define* $\bar{\sigma}_\varepsilon := -\tilde{J}'(\bar{v}_\varepsilon)$. Then, (4.2.12c) holds by definition and (4.2.12a) and (4.2.12b) are a simple consequence of (4.2.14). \square

The next step is to show that the Lagrange multiplier $\bar{\sigma}_\varepsilon$ of Problem (\tilde{P}_ε) is also a Lagrange multiplier for the problem (P_ε) . Therefore, we first define the Lagrange functional L corresponding to the problem (P_ε) as follows:

$$L(u, \sigma) := J(u) + (\sigma, \varepsilon u + Gu - \psi)_{0,\Omega}. \quad (4.2.15)$$

Then, there holds:

Lemma 4.2.4. *The Lagrange multiplier $\bar{\sigma}_\varepsilon$ from Lemma 4.2.3 is also an Lagrange multiplier for the problem (P_ε) , i.e., for its optimal solution \bar{u}_ε , there holds*

$$\bar{\sigma}_\varepsilon \geq 0 \quad \text{a.e. in } \Omega, \quad (4.2.16a)$$

$$(\bar{\sigma}_\varepsilon, G\bar{u}_\varepsilon + \varepsilon\bar{u}_\varepsilon - \psi)_{0,\Omega} = 0, \quad (4.2.16b)$$

$$\frac{\partial}{\partial u} L(\bar{u}_\varepsilon, \bar{\sigma}_\varepsilon) = 0 \quad (4.2.16c)$$

Proof. Since we have $\bar{v}_\varepsilon = \varepsilon\bar{u}_\varepsilon + G\bar{u}_\varepsilon$, one may obtain

$$(\bar{\sigma}_\varepsilon, \varepsilon\bar{u}_\varepsilon + G\bar{u}_\varepsilon - \psi)_{0,\Omega} = (\bar{\sigma}_\varepsilon, \bar{v}_\varepsilon - \psi)_{0,\Omega} = 0, \quad (4.2.17)$$

which proves (4.2.16b). With the help of the Riesz representation theorem, a simple computation results in

$$\frac{\partial}{\partial u} L(\bar{u}_\varepsilon, \bar{\sigma}_\varepsilon) = J'(\bar{u}_\varepsilon) + (R^*)^{-1}\bar{\sigma}_\varepsilon. \quad (4.2.18)$$

The right-hand side of (4.2.18) vanishes since

$$R^* J'(\bar{u}_\varepsilon) + \bar{\sigma}_\varepsilon = 0 \quad (4.2.19)$$

due to (4.2.12c) and Lemma 4.2.2, what finally shows (4.2.16c). \square

With Lemma 4.2.4 at hand, one may now state the optimality conditions for Problem (P_ε) :

Theorem 4.2.5 (Regularized optimality conditions). *The unique optimal solution $(\bar{y}_\varepsilon, \bar{u}_\varepsilon) \in V \times L^2(\Omega)$ of (4.1.1) is characterized by the existence of an optimal adjoint state $\bar{p}_\varepsilon \in V$ and an optimal adjoint control $\bar{\sigma}_\varepsilon \in L^2(\Omega)$ such that the following conditions are fulfilled:*

$$(\nabla \bar{y}_\varepsilon, \nabla v)_{0,\Omega} + c(\bar{y}_\varepsilon, v)_{0,\Omega} = (\bar{u}_\varepsilon, v)_{0,\Omega} \quad \forall v \in V \quad (4.2.20a)$$

$$(\nabla \bar{p}_\varepsilon, \nabla v)_{0,\Omega} + c(\bar{p}_\varepsilon, v)_{0,\Omega} = (\bar{y}_\varepsilon + \bar{\sigma}_\varepsilon - y^d, v)_{0,\Omega} \quad \forall v \in V \quad (4.2.20b)$$

$$\bar{p}_\varepsilon + \alpha(\bar{u}_\varepsilon - u^d) + \varepsilon \bar{\sigma}_\varepsilon = 0 \quad \text{a.e. in } \Omega \quad (4.2.20c)$$

$$\bar{\sigma}_\varepsilon \geq 0, \quad \varepsilon \bar{u}_\varepsilon + \bar{y}_\varepsilon \leq \psi, \quad (\bar{\sigma}_\varepsilon, \bar{y}_\varepsilon + \varepsilon \bar{u}_\varepsilon - \psi)_{0,\Omega} = 0 \quad (4.2.20d)$$

Proof. We can use (4.2.20b) as a definition of \bar{p}_ε . Then, considering Lemma 4.2.4, only (4.2.20c) remains to be shown. (4.2.16c) implies that for every $v \in L^2(\Omega)$, there holds

$$(\bar{y}_\varepsilon - y^d, Gv)_{0,\Omega} + \alpha(\bar{u}_\varepsilon - u^d, v)_{0,\Omega} + \varepsilon(\bar{\sigma}_\varepsilon, v)_{0,\Omega} + (\bar{\sigma}_\varepsilon, Gv)_{0,\Omega} = 0. \quad (4.2.21)$$

Since $Gv \in V$, one may apply (4.2.20b) and obtains for the combination of the first and the last term of the right-hand side of (4.2.21):

$$(\bar{y}_\varepsilon - y^d + \bar{\sigma}_\varepsilon, Gv)_{0,\Omega} = (\nabla \bar{p}_\varepsilon, \nabla Gv)_{0,\Omega} + c(\bar{p}_\varepsilon, Gv)_{0,\Omega}. \quad (4.2.22)$$

By applying the definition of G to the right-hand side of (4.2.22), one may conclude from (4.2.22) that

$$(\bar{y}_\varepsilon - y^d + \bar{\sigma}_\varepsilon, Gv)_{0,\Omega} = (v, \bar{p}_\varepsilon)_{0,\Omega}. \quad (4.2.23)$$

If one now inserts (4.2.23) into (4.2.21), the claim (4.2.20c) follows as an equality in $L^2(\Omega)$. \square

In the remaining part of this chapter, we will only deal with the optimal solutions of the regularized problem (P_ε) . For clarity, we thus drop the line above the optimal functions in all what follows.

4.3 Convergence to the unregularized problem

Obviously, the regularization only makes sense if the regularized optimal solutions converge to the unregularized ones as $\varepsilon \rightarrow 0$. This section is devoted to the proof of this important property. The proof follows the lines of [MPT06]. However, the problem is a bit different, and thus, we will repeat the proof here. Notice that we may also state the unregularized optimal control problem (2.1.3) with the upper bound ψ being taken from the space $L^\infty(\Omega)$. The property $\psi \in C(\bar{\Omega})$ was only important to derive optimality conditions.

In order to prove the convergence to the unperturbed problem, let $\{\varepsilon_n\}_{n \in \mathbb{N}}$ be a positive null sequence, and let $\{u_{\varepsilon_n}\}_{n \in \mathbb{N}}$ be the corresponding optimal solutions of the sequence $\{(P_{\varepsilon_n})\}_{n \in \mathbb{N}}$. For a better notation, we shortly write $u_n = u_{\varepsilon_n}$ for all

$n \in \mathbb{N}$. Now, the goal is to show that the sequence $\{u_n\}_{n \in \mathbb{N}}$ converges to the optimal solution u of the unregularized problem (2.1.3).

As a first step in this direction, we set

$$\tilde{u}_n := (\varepsilon_n Id + G)^{-1} y, \quad n \in \mathbb{N}, \quad (4.3.1)$$

where still $y = Gu$. With these definitions at hand, we may now formulate the following lemma:

Lemma 4.3.1. *Let u be the optimal solution of the unregularized problem (2.1.3) and let $\{\tilde{u}_n\}_{n \in \mathbb{N}}$ be as defined in (4.3.1). Then, one may claim*

$$\lim_{n \rightarrow \infty} \|u - \tilde{u}_n\|_{0,\Omega} = 0. \quad (4.3.2)$$

Proof. It is easy to see that there holds

$$\begin{aligned} u - \tilde{u}_n &= (\varepsilon_n Id + G)^{-1} (\varepsilon_n Id + G)u - (\varepsilon_n Id + G)^{-1} (Gu) \\ &= (\varepsilon_n Id + G)^{-1} \varepsilon_n u \\ &= \varepsilon_n (\varepsilon_n Id + G)^{-1} u. \end{aligned} \quad (4.3.3)$$

Because the differential operator of the partial differential equation (2.1.1) is symmetric, G is self-adjoint. In addition to that, G , with image $L^2(\Omega)$, is compact and positive definite, and thus, G has a 0-dimensional kernel and there exists an orthonormal basis $\{e_i\}_{i \in \mathbb{N}}$ of eigenvectors of G with corresponding positive eigenvalues $\{\lambda_i\}_{i \in \mathbb{N}}$ (compare [Alt02]). Since G is furthermore continuous, (4.3.3) may be written equivalently as follows:

$$\begin{aligned} u - \tilde{u}_n &= \varepsilon_n (\varepsilon_n Id + G)^{-1} \sum_{i=1}^{\infty} (u, e_i)_{0,\Omega} e_i \\ &= \varepsilon_n \sum_{i=1}^{\infty} (u, e_i)_{0,\Omega} (\varepsilon_n Id + G)^{-1} e_i \\ &= \varepsilon_n \sum_{i=1}^{\infty} (u, e_i)_{0,\Omega} \frac{1}{\varepsilon_n + \lambda_i} e_i \\ &= \sum_{i=1}^{\infty} \frac{\varepsilon_n}{\varepsilon_n + \lambda_i} (u, e_i)_{0,\Omega} e_i \end{aligned} \quad (4.3.4)$$

Taking into account that $\lambda_i > 0 \forall i \in \mathbb{N}$, one gets with the help of the Parseval

equality

$$\begin{aligned}
\|u - \tilde{u}_n\|_{0,\Omega}^2 &= \left\| \sum_{i=1}^{\infty} \frac{\varepsilon_n}{\varepsilon_n + \lambda_i} (u, e_i)_{0,\Omega} e_i \right\|_{0,\Omega}^2 & (4.3.5) \\
&= \sum_{i=1}^{\infty} \underbrace{\left(\frac{\varepsilon_n}{\varepsilon_n + \lambda_i} \right)^2}_{\leq 1} (u, e_i)_{0,\Omega}^2 \\
&\leq \sum_{i=1}^{\infty} (u, e_i)_{0,\Omega}^2 \\
&= \|u\|_{0,\Omega}^2 .
\end{aligned}$$

Thus, the real valued function series $\sum_{i=1}^{\infty} \phi_i(\varepsilon)$, with

$$\phi_i(\varepsilon) := \left(\frac{\varepsilon}{\varepsilon + \lambda_i} \right)^2 (u, e_i)_{0,\Omega}^2 \leq (u, e_i)_{0,\Omega}^2 , \quad (4.3.6)$$

converges uniformly, and consequently, we are allowed to exchange summation and pass to the limit. Therefore, we may conclude:

$$\lim_{n \rightarrow \infty} \|u - \tilde{u}_n\|_{0,\Omega}^2 = \sum_{i=1}^{\infty} \underbrace{\lim_{n \rightarrow \infty} \left(\frac{\varepsilon_n}{\varepsilon_n + \lambda_i} \right)^2 (u, e_i)_{0,\Omega}^2}_{= 0 \ \forall i \in \mathbb{N}} = 0 \quad (4.3.7)$$

□

Lemma 4.3.2. *The sequence $\{\|u_n\|_{0,\Omega}\}_{n \in \mathbb{N}}$ of the L^2 -norms of the optimal solutions corresponding to the sequence $\{(P_{\varepsilon_n})\}_{n \in \mathbb{N}}$ is bounded in $L^2(\Omega)$.*

Proof. Because $y = Gu$ is feasible for the unregularized problem, we may obtain

$$G\tilde{u}_n + \varepsilon_n \tilde{u}_n - \psi = (\varepsilon_n Id + G)\tilde{u}_n - \psi = y - \psi \leq 0$$

Thus, \tilde{u}_n is feasible for (P_{ε_n}) for every $n \in \mathbb{N}$. This implies

$$J(u_n) \leq J(\tilde{u}_n) \quad \forall n \in \mathbb{N} . \quad (4.3.8)$$

Considering Lemma 4.3.1 and the fact that J is continuous, we get

$$J(\tilde{u}_n) \longrightarrow J(u) \quad \text{for } n \rightarrow \infty . \quad (4.3.9)$$

Combined with (4.3.8), this implies the existence of a $n_0 \in \mathbb{N}$ such that

$$J(u_n) \leq J(\tilde{u}_n) \leq J(u) + 1 \quad \forall n \geq n_0 . \quad (4.3.10)$$

From (4.3.10), one may conclude with the help of a basic computation

$$\|u_n\|_{0,\Omega}^2 \leq \frac{4}{\alpha}(J(u) + 1) + 2\|u^d\|_{0,\Omega}^2 \quad \forall n \geq n_0, \quad (4.3.11)$$

which readily gives the assertion. \square

Because $L^2(\Omega)$ is a Hilbert space and thus particularly reflexive, we can select a subsequence of $\{u_n\}_{n \in \mathbb{N}}$ which is weakly convergent. W.l.o.g., we can assume that the whole sequence converges, i.e., that there exists an $\tilde{u} \in L^2(\Omega)$:

$$u_n \rightharpoonup \tilde{u} \quad \text{for } n \rightarrow \infty. \quad (4.3.12)$$

Now, it remains to be shown that \tilde{u} is equal to the optimal control u of the unregularized problem and that the convergence is indeed strong. As a first step in this direction, we have to verify the following result for \tilde{u} :

Lemma 4.3.3. *The weak limit \tilde{u} of the sequence $\{u_n\}_{n \in \mathbb{N}}$ is feasible for the unregularized problem (2.1.3).*

Proof. Since G is compact, $\{\varepsilon_n\}_{n \in \mathbb{N}}$ is a null sequence, and $\{\|u_n\|_{0,\Omega}\}_{n \in \mathbb{N}}$ is bounded, one gets

$$\underbrace{Gu_n}_{\rightarrow G\tilde{u}} + \underbrace{\varepsilon_n u_n}_{\rightarrow 0} - \psi \leq 0 \quad \text{a.e. in } \Omega, \quad (4.3.13)$$

which implies

$$G\tilde{u} \leq \psi \quad \text{a.e. in } \Omega. \quad (4.3.14)$$

Due to our assumptions on the domain Ω , $G\tilde{u}$ lives in the space $C(\overline{\Omega})$. Therefore, (4.3.14) has to hold for all $x \in \Omega$ if we demand that $\psi \in C(\overline{\Omega})$. \square

Now, we are in the position to prove the following theorem, which is the main result of this section.

Theorem 4.3.4 (Convergence to the unregularized problem). *The regularized optimal control u_ε converges to the unregularized control u as the regularization parameter ε tends to 0, i.e., for the sequence $\{u_n\}_{n \in \mathbb{N}}$, there holds:*

$$\lim_{n \rightarrow \infty} \|u - u_n\|_{0,\Omega} = 0. \quad (4.3.15)$$

Proof. Since J is convex and lower semi-continuous, it is also weakly lower semi-continuous. Thus, taking into account that \tilde{u}_n is admissible for (P_{ε_n}) for all $n \in \mathbb{N}$ and applying (4.3.8), one may conclude

$$J(u) = \lim_{n \rightarrow \infty} J(\tilde{u}_n) \geq \limsup_{n \rightarrow \infty} J(u_n) \geq \liminf_{n \rightarrow \infty} J(u_n) \geq J(\tilde{u}) \geq J(u), \quad (4.3.16)$$

where we also have used Lemma 4.3.3 for the last inequality. Thus, all inequality signs in (4.3.16) can be replaced by equality signs. Because J is strictly convex for

$\alpha > 0$, there has to hold $\tilde{u} = u$. Therefore, $u_n \rightharpoonup u$ as $n \rightarrow \infty$, and it remains to show that the convergence is in fact strong. In order to do that, we consider the definition of J and thus get

$$\|u_n - u^d\|_{0,\Omega}^2 = \frac{2}{\alpha} (J(u_n) - \frac{1}{2} \|y_n - y^d\|_{0,\Omega}^2). \quad (4.3.17)$$

Now, we may conclude from (4.3.16) that $J(u_n) \rightarrow J(u)$ for $n \rightarrow \infty$. Since furthermore, the operator G is compact if one interpretes it with image $L^2(\Omega)$, we may conclude

$$\lim_{n \rightarrow \infty} \|u_n - u^d\|_{0,\Omega}^2 = \frac{2}{\alpha} (J(u) - \frac{1}{2} \|y - y^d\|_{0,\Omega}^2) = \|u - u^d\|_{0,\Omega}^2. \quad (4.3.18)$$

Then, using the fundamental relationship

$$\|a - c\|^2 = \|a - b\|^2 - \|c - b\|^2 - 2(c - b, a - c) \quad (4.3.19)$$

with $a = u_n$, $b = u^d$, and $c = u$, one may conclude

$$\|u_n - u\|_{0,\Omega}^2 = (\|u_n - u^d\|_{0,\Omega}^2 - \|u - u^d\|_{0,\Omega}^2) - 2(u - u^d, u_n - u)_{0,\Omega}. \quad (4.3.20)$$

If one takes the limit in (4.3.20), the term in brackets tends to 0, due to (4.3.18), while the last term converges to 0 because of the weak convergence of u_n to u . Thus, the claim has been proved. \square

4.4 Discretization and modified Lagrange multipliers

Assume that the triangulation and the discrete spaces are defined as described for the pure state constrained case in the previous chapter. For the discretization of the control u , the state y , and the adjoint state p , we use the same discrete spaces as for the pure state constrained case. Since now, the adjoint control σ lives in the space $L^2(\Omega)$ and is part of the fundamental relationship (4.2.20c), we also want to approximate it in the space S_h and not in the discrete measure space M_h , as we did for the pure state constrained case. However, the discretization of the control in the space of piecewise linear functions causes some difficulties. This issue is discussed in [CM], for example. The problem is that projection arguments like they are valid in the continuous case or if one discretizes with piecewise constant functions can not be applied anymore. Instead, the variational inequality of the associated discrete optimal control problem leads to a L^2 -projection, which is hard to deal with from a theoretical and numerical point of view.

In order to overcome this problem, we go the same way as it has been done in [MPT06] and do not formulate a discrete optimal control problem, but just discretize the optimality system (4.2.20) instead. This may be interpreted as an approach which first optimizes and then discretizes. However, this will result in a

difference within the complementarity conditions and, as a consequence, in an additional consistency error term. The support of this quantity is concentrated on the discrete free boundary as it will be defined in (4.4.11). Therefore, this term should not cause any problems if one refines along the discrete free boundary a priori.

In all what follows, the regularization parameter ε can be an arbitrary positive number but is assumed to be fixed.

In order to get an useful discretization of the optimality system, we have to reformulate it equivalently. Therefore, we define the continuous regularized active set $\mathcal{A}^\varepsilon \subset \Omega$ and the continuous regularized inactive set $\mathcal{I}^\varepsilon \subset \Omega$ according to

$$\mathcal{A}^\varepsilon := \{x \in \Omega \mid \varepsilon u_\varepsilon(x) + y_\varepsilon(x) + \sigma_\varepsilon(x) > \psi(x)\} , \quad (4.4.1a)$$

$$\mathcal{I}^\varepsilon := \Omega \setminus \mathcal{A}^\varepsilon . \quad (4.4.1b)$$

These sets are well-defined up to measure 0. One can easily verify (or cf. [MPT06]) that the optimality system (4.2.20) is equivalent to

$$(\nabla y_\varepsilon, \nabla v)_{0,\Omega} + c(y_\varepsilon, v)_{0,\Omega} = (u_\varepsilon, v)_{0,\Omega} \quad \forall v \in V \quad (4.4.2a)$$

$$(\nabla p_\varepsilon, \nabla v)_{0,\Omega} + c(p_\varepsilon, v)_{0,\Omega} = (y_\varepsilon + \sigma_\varepsilon - y^d, v)_{0,\Omega} \quad \forall v \in V \quad (4.4.2b)$$

$$p_\varepsilon + \alpha(u_\varepsilon - u^d) + \varepsilon \sigma_\varepsilon = 0 \quad \text{a.e. in } \Omega \quad (4.4.2c)$$

$$\sigma_\varepsilon = 0 \quad \text{a.e. in } \mathcal{I}^\varepsilon , \quad \varepsilon u_\varepsilon + y_\varepsilon = \psi \quad \text{a.e. in } \mathcal{A}^\varepsilon \quad (4.4.2d)$$

Notice that the upper bound on the state and the positiveness of the adjoint control, which do not explicitly appear in (4.4.2), are involved in the definitions of \mathcal{A}^ε and \mathcal{I}^ε .

The state equation (4.4.2a) and the adjoint state equation (4.4.2b) are now discretized by linear finite elements, as done in the previous chapter. Furthermore, we replace the equality (4.4.2c) by its discrete counterpart

$$p_{\varepsilon,h} + \alpha(u_{\varepsilon,h} - u_h^d) + \varepsilon \sigma_{\varepsilon,h} = 0 , \quad (4.4.3)$$

where u_h^d is again a suitable approximation of u^d in the space S_h . Actually, it is sufficient to demand (4.4.3) for all nodes of the triangulation. Now, we also use a suitable discretization ψ_h of the upper bound ψ in S_h . If ψ is continuous, which is not necessary in the regularized case, one may use the interpolant. For the approximation of (4.4.2d), assume that $\{n_1, \dots, n_N\} = \mathcal{N}_h(\overline{\Omega})$ are the numbered nodes of the actual triangulation. Then, the discrete solution consists of $(y_{\varepsilon,h}, u_{\varepsilon,h}, p_{\varepsilon,h}, \sigma_{\varepsilon,h}) \in (V_h, S_h, V_h, S_h)$ and two subsets of indexes, $\mathcal{A}_{h,idx}^\varepsilon, \mathcal{I}_{h,idx}^\varepsilon$, such that the following system is fulfilled:

$$(\nabla y_{\varepsilon,h}, \nabla v_h)_{0,\Omega} + c(y_{\varepsilon,h}, v_h)_{0,\Omega} = (u_{\varepsilon,h}, v_h)_{0,\Omega} , \quad v_h \in V_h , \quad (4.4.4a)$$

$$(\nabla p_{\varepsilon,h}, \nabla v_h)_{0,\Omega} + c(p_{\varepsilon,h}, v_h)_{0,\Omega} = (y_h - y^d + \sigma_{\varepsilon,h}, v_h)_{0,\Omega} , \quad v_h \in V_h , \quad (4.4.4b)$$

$$p_{\varepsilon,h} + \alpha(u_{\varepsilon,h} - u_h^d) + \varepsilon \sigma_{\varepsilon,h} = 0 , \quad (4.4.4c)$$

$$\sigma_{\varepsilon,h}(n_i) = 0 \quad \text{if } i \in \mathcal{I}_{h,idx}^\varepsilon , \quad \varepsilon u_{\varepsilon,h}(n_i) + y_{\varepsilon,h}(n_i) = \psi_h(n_i) \quad \text{if } i \in \mathcal{A}_{h,idx}^\varepsilon , \quad (4.4.4d)$$

where the active and inactive index sets are given by

$$\mathcal{A}_{h,idx}^\varepsilon := \{i \in \{1, \dots, N\} \mid \varepsilon u_{\varepsilon,h}(n_i) + y_{\varepsilon,h}(n_i) + \sigma_{\varepsilon,h}(n_i) > \psi_h(n_i)\} , \quad (4.4.5a)$$

$$\mathcal{I}_{h,idx}^\varepsilon := \{1, \dots, N\} \setminus \mathcal{A}_{h,idx}^\varepsilon . \quad (4.4.5b)$$

Such a discrete solution can, for example, be achieved by a primal-dual active set strategy (cf. [MPT06]).

One may easily derive from (4.4.4d) and from the definitions (4.4.5) the following complementarity conditions:

Corollary 4.4.1 (Regularized complementarity conditions). *The regularized solution $(y_{\varepsilon,h}, u_{\varepsilon,h}, p_{\varepsilon,h}, \sigma_{\varepsilon,h}) \in (V_h, S_h, V_h, S_h)$ of (4.4.4) fulfills*

$$\begin{aligned} \varepsilon u_{\varepsilon,h} + y_{\varepsilon,h} - \psi_h &\leq 0 , \quad \sigma_{\varepsilon,h} \geq 0 , \\ \sigma_{\varepsilon,h}(n) * (\varepsilon u_{\varepsilon,h}(n) + y_{\varepsilon,h}(n) - \psi_h(n)) &= 0 \quad \forall n \in \mathcal{N}_h(\Omega) . \end{aligned} \quad (4.4.6)$$

Unfortunately, these complementarity conditions are not similar to those of the unconstrained case (compare (2.4.8)) since in general there holds

$$(\sigma_{\varepsilon,h}, \varepsilon u_{\varepsilon,h} + y_{\varepsilon,h} - \psi_h)_{0,\Omega} \neq 0 . \quad (4.4.7)$$

However, we can derive something similar for a subset of the triangles. Therefore, we define the discrete regularized active set $\mathcal{A}_h^\varepsilon \subset \mathcal{T}_h(\Omega)$ according to

$$\mathcal{A}_h^\varepsilon := \{T \in \mathcal{T}_h(\Omega) : \varepsilon u_{\varepsilon,h}(n) + y_{\varepsilon,h}(n) = \psi_h(n) , n \in \mathcal{T}_h(T)\} \quad (4.4.8)$$

and the discrete regularized inactive set $\mathcal{I}_h^\varepsilon \subset \mathcal{T}_h(\Omega)$ according to

$$\mathcal{I}_h^\varepsilon := \{T \in \mathcal{T}_h(\Omega) : \varepsilon u_{\varepsilon,h}(n) + y_{\varepsilon,h}(n) < \psi_h(n) , n \in \mathcal{T}_h(T)\} . \quad (4.4.9)$$

Notice that from (4.4.6), one may infer that

$$\sigma_{\varepsilon,h|_T} := 0 , T \in \mathcal{I}_h^\varepsilon . \quad (4.4.10)$$

The remaining triangles are referred to as the discrete free boundary. In particular, we set

$$\mathcal{F}_h^\varepsilon := \mathcal{T}_h(\Omega) \setminus \{\mathcal{A}_h^\varepsilon \cup \mathcal{I}_h^\varepsilon\} . \quad (4.4.11)$$

With these definitions at hand and taking (4.4.10) into account, it is easy to see that there holds

$$(\sigma_{\varepsilon,h}, \varepsilon u_{\varepsilon,h} + y_{\varepsilon,h} - \psi_h)_{0,\Omega} = \sum_{T \in \mathcal{F}_h^\varepsilon} (\sigma_{\varepsilon,h}, \varepsilon u_{\varepsilon,h} + y_{\varepsilon,h} - \psi_h)_{0,T} . \quad (4.4.12)$$

As already mentioned above, it is only useful to derive an a posteriori error estimation for which the involved constants do not depend on the regularization parameter

ε because only a sufficiently small value of ε ensures a good approximation of the regularized solution to the unregularized one. This can also be seen in the numerical examples of the forthcoming chapter. In the following, we show that the error estimator derived in the previous chapter is also suitable for the regularized constraints with the involved constants not depending on ε . In order to achieve this goal, we again have to define a modified adjoint state \hat{p}_ε and a modified adjoint control $\hat{\sigma}_\varepsilon$. Even so p_ε and σ_ε are smooth functions, this seems to be necessary since they tend to be unsmooth functions as the Lavrentiev regularization parameter ε tends to 0. Similar to the pure state constrained case, we therefore define the modified (regularized) adjoint state $\hat{p}_\varepsilon \in V$ as the unique solution of

$$(\nabla \hat{p}_\varepsilon, \nabla v)_{0,\Omega} + c(\hat{p}_\varepsilon, v)_{0,\Omega} = (y_\varepsilon - y^d, v)_{0,\Omega}, \quad v \in V, \quad (4.4.13)$$

and the modified (regularized) adjoint control $\hat{\sigma}_\varepsilon \in V$ as the unique solution of

$$(\nabla \hat{\sigma}_\varepsilon, \nabla v)_{0,\Omega} + c(\hat{\sigma}_\varepsilon, v)_{0,\Omega} = (\sigma_\varepsilon, v)_{0,\Omega}, \quad v \in V. \quad (4.4.14)$$

In addition to that, we define discrete modified Lagrange multipliers. Let the discrete modified adjoint state $\hat{p}_{\varepsilon,h} \in V_h$ be the unique solution of

$$(\nabla \hat{p}_{\varepsilon,h}, \nabla v_h)_{0,\Omega} + c(\hat{p}_{\varepsilon,h}, v_h)_{0,\Omega} = (y_{\varepsilon,h} - y^d, v_h)_{0,\Omega}, \quad v_h \in V_h, \quad (4.4.15)$$

and let the discrete modified adjoint control $\hat{\sigma}_{\varepsilon,h} \in V$ be the unique solution of

$$(\nabla \hat{\sigma}_{\varepsilon,h}, \nabla v_h)_{0,\Omega} + c(\hat{\sigma}_{\varepsilon,h}, v_h)_{0,\Omega} = (\sigma_{\varepsilon,h}, v_h)_{0,\Omega}, \quad v_h \in V_h. \quad (4.4.16)$$

Then, it is easy to see that there holds

$$p_{\varepsilon,h} = \hat{p}_{\varepsilon,h} + \hat{\sigma}_{\varepsilon,h}. \quad (4.4.17)$$

4.5 A posteriori error estimator

The residual-type a posteriori error estimator for the regularized problem is pretty much the same as for the unregularized case. It involves terms corresponding to the state equation and to the modified adjoint state. The difference is that now, the consistency error consists of two parts.

In particular, the total residual-type a posteriori error estimator corresponding to the regularized optimal control problem (4.1.1) is given by

$$\eta_\varepsilon := \eta_{y_\varepsilon} + \eta_{\hat{p}_\varepsilon}, \quad (4.5.1)$$

where η_{y_ε} and $\eta_{\hat{p}_\varepsilon}$ are again the sum of weighted element and edge residuals according to

$$\eta_{y_\varepsilon} := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \eta_{y_\varepsilon, T}^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \eta_{y_\varepsilon, E}^2 \right)^{1/2}, \quad (4.5.2a)$$

$$\eta_{\hat{p}_\varepsilon} := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \eta_{\hat{p}_\varepsilon, T}^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \eta_{\hat{p}_\varepsilon, E}^2 \right)^{1/2}. \quad (4.5.2b)$$

The element residuals $\eta_{y_\varepsilon, T}$ and $\eta_{\hat{p}_\varepsilon, T}$, $T \in \mathcal{T}_h(\Omega)$, are weighted elementwise L^2 -residuals with respect to the strong forms of the state equation (4.1.1b) and the modified adjoint state equation (4.4.13), respectively:

$$\eta_{y_\varepsilon, T} := h_T \|u_{\varepsilon, h} - cy_{\varepsilon, h}\|_{0, T}, \quad T \in \mathcal{T}_h(\Omega), \quad (4.5.3a)$$

$$\eta_{\hat{p}_\varepsilon, T} := h_T \|y_{\varepsilon, h} - y^d - c\hat{p}_{\varepsilon, h}\|_{0, T}, \quad T \in \mathcal{T}_h(\Omega). \quad (4.5.3b)$$

The edge residuals $\eta_{y_\varepsilon, E}$ and $\eta_{\hat{p}_\varepsilon, E}$, $E \in \mathcal{E}_h(\Omega)$, are again weighted L^2 -norms of the jumps $\nu_E \cdot [\nabla y_{\varepsilon, h}]$ and $\nu_E \cdot [\nabla \hat{p}_{\varepsilon, h}]$ of the normal derivatives across the interior edges

$$\eta_{y_\varepsilon, E} := h_E^{1/2} \|\nu_E \cdot [\nabla y_{\varepsilon, h}]\|_{0, E}, \quad E \in \mathcal{E}_h(\Omega), \quad (4.5.4a)$$

$$\eta_{\hat{p}_\varepsilon, E} := h_E^{1/2} \|\nu_E \cdot [\nabla \hat{p}_{\varepsilon, h}]\|_{0, E}, \quad E \in \mathcal{E}_h(\Omega). \quad (4.5.4b)$$

Similarly, we use the lower order data oscillation in u^d , as defined in (3.1.4), and the higher order data oscillation in y^d , as defined in (3.1.5). Since an approximation of ψ appears in the discrete problem (4.4.4) and in the consistency error terms as they will be defined in (4.5.8), we will also include lower order data oscillation in the upper bound in our adaptive algorithm. Thus, we define

$$\mu_h(\psi) := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \mu_T^2(\psi) \right)^{1/2}, \quad (4.5.5)$$

$$\mu_T(\psi) := \|\psi - \psi_h\|_{0, T}, \quad T \in \mathcal{T}_h(\Omega).$$

However, the data oscillation in ψ will not appear explicitly in the forthcoming reliability and efficiency results.

We will show that, up to data oscillations and consistency errors as they will be defined below, the residual-type a posteriori error estimator (4.5.1) provides an upper and a lower bound for the discretization errors in the state, in the control, and in the modified adjoint state.

For the proof, it is again useful to fix an auxiliary state $y(u_{\varepsilon, h}) \in V$ and an auxiliary modified adjoint state $\hat{p}(y_{\varepsilon, h}) \in V$. These functions are again defined as the continuous solutions with the discrete right-hand sides, i.e., as the unique solutions of the following variational equations:

$$(\nabla y(u_{\varepsilon, h}), \nabla v)_{0, \Omega} + c(y(u_{\varepsilon, h}), v)_{0, \Omega} = (u_{\varepsilon, h}, v)_{0, \Omega}, \quad v \in V, \quad (4.5.6a)$$

$$(\nabla \hat{p}(y_{\varepsilon, h}), \nabla v)_{0, \Omega} + c(\hat{p}(y_{\varepsilon, h}), v)_{0, \Omega} = (y_{\varepsilon, h} - y^d, v)_{0, \Omega}, \quad v \in V. \quad (4.5.6b)$$

We further introduce a discrete auxiliary state $y_h(u_\varepsilon) \in V_h$ as the solution of the finite dimensional variational problem

$$(\nabla y_h(u_\varepsilon), \nabla v_h)_{0,\Omega} + c(y_h(u_\varepsilon), v_h)_{0,\Omega} = (u_\varepsilon, v_h)_{0,\Omega}, \quad v_h \in V_h. \quad (4.5.7)$$

Like for the unregularized case, the auxiliary states $y(u_{\varepsilon,h}) \in V$ and $y_h(u_\varepsilon) \in V_h$ do not necessarily satisfy the state constraints, i.e., it may happen that $\varepsilon u_\varepsilon + y_h(u_\varepsilon) \not\leq \psi_h$ or $\varepsilon u_{\varepsilon,h} + y(u_{\varepsilon,h}) \not\leq \psi$. This will again result in a consistency error term which is similar to the one before. In addition to that, we now get an additional consistency error term, which has its origin in the observation (4.4.7). In particular, we introduce the consistency errors

$$e_c^{(1)}(u_\varepsilon, u_{\varepsilon,h}) := \max((\sigma_{\varepsilon,h}, \varepsilon u_\varepsilon + y_h(u_\varepsilon) - \psi_h)_{0,\Omega} + (\sigma_\varepsilon, \varepsilon u_{\varepsilon,h} + y(u_{\varepsilon,h}) - \psi)_{0,\Omega}, 0), \quad (4.5.8a)$$

$$e_c^{(2)}(h) := \sum_{T \in \mathcal{F}_h^\varepsilon} (\sigma_{\varepsilon,h}, \psi_h - (\varepsilon u_{\varepsilon,h} + y_{\varepsilon,h}))_{0,T}. \quad (4.5.8b)$$

Again, we note that, like for the pure state constrained case, we get $y(u_{\varepsilon,h}) = y_\varepsilon$ and $y_h(u_\varepsilon) = y_{\varepsilon,h}$ if $u_\varepsilon = u_{\varepsilon,h}$. Hence, $e_c^{(1)}(u_\varepsilon, u_{\varepsilon,h}) = 0$ because both scalar products in (4.5.8a) have non-positive values. For example, for the first scalar product, there holds $\sigma_{\varepsilon,h} \geq 0$ and $\varepsilon u_\varepsilon + y_h(u_\varepsilon) - \psi_h = \varepsilon u_{\varepsilon,h} + y_{\varepsilon,h} - \psi_h \leq 0$ due to Corollary 4.4.1. (4.4.2d) implies a similar argument for the second scalar product. The term $e_c^{(2)}(h)$ is actually fully a posteriori and thus computable. However, its support is concentrated on the discrete free boundary and therefore, we will not include this quantity in our a posteriori error estimator. Instead, in our adaptive algorithm, we will refine triangles being part of the discrete free boundary in each iteration a priori. Then, $e_c^{(2)}(h)$ seems likely to decrease, as its support gets smaller in each iteration. On the other side, it does not necessarily vanish for the case $u = u_h$.

4.6 Reliability of the error estimator

The goal of this section is to prove that the error estimator as defined in the previous section provides an upper bound for the discretization error in the regularized state, in the regularized control, and in the modified regularized adjoint state. The proof is similar to the one given in the previous chapter for the pure state constrained case. However, there are some differences, and thus, the proof will be given here. Let us first state the main theorem. Its proof will be provided by the following series of lemmas.

Theorem 4.6.1 (Reliability). *Let $(y_\varepsilon, u_\varepsilon, p_\varepsilon, \sigma_\varepsilon)$ be the continuous regularized solution of (4.2.20), let $(y_{\varepsilon,h}, u_{\varepsilon,h}, p_{\varepsilon,h}, \sigma_{\varepsilon,h})$ be the discrete regularized solution of (4.4.4),*

and let \hat{p}_ε and $\hat{p}_{\varepsilon,h}$ be as defined in (4.4.13) and (4.4.15), respectively. Further, let η_ε , $\mu_h(u^d)$, $e_c^{(1)}(u_\varepsilon, u_{\varepsilon,h})$, and $e_c^{(2)}(h)$ be as defined in (4.5.1), (3.1.4), (4.5.8a), and (4.5.8b), respectively. Then, independently of the regularization parameter ε , there exists a constant C , depending on α , \tilde{c} , and the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that the following estimation holds:

$$\begin{aligned} & \|y_\varepsilon - y_{\varepsilon,h}\|_{1,\Omega}^2 + \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2 + \|\hat{p}_\varepsilon - \hat{p}_{\varepsilon,h}\|_{1,\Omega}^2 \\ & \leq C(\eta_\varepsilon^2 + \mu_h^2(u^d) + e_c^{(1)}(u_\varepsilon, u_{\varepsilon,h}) + e_c^{(2)}(h)). \end{aligned} \quad (4.6.1)$$

As a first step to prove Theorem 4.6.1, we state

Lemma 4.6.2. *In addition to the assumptions of Theorem 4.6.1, let $y(u_\varepsilon)$ and $\hat{p}(y_{\varepsilon,h})$ be the auxiliary state and auxiliary modified adjoint state according to (4.5.6). Then, there holds:*

$$\begin{aligned} & \|y_\varepsilon - y_{\varepsilon,h}\|_{1,\Omega}^2 + \|\hat{p}_\varepsilon - \hat{p}_{\varepsilon,h}\|_{1,\Omega}^2 \\ & \leq C(\|y(u_\varepsilon) - y_{\varepsilon,h}\|_{1,\Omega}^2 + \|\hat{p}(y_{\varepsilon,h}) - \hat{p}_{\varepsilon,h}\|_{1,\Omega}^2 + \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2), \end{aligned} \quad (4.6.2)$$

where C depends only on the constant \tilde{c} .

Proof. All involved variables have the same right-hand sides as the unperturbed functions. The unregularized functions appearing in the right-hand sides are just replaced by the regularized ones. Thus, the proof is exactly the same as the proof of Lemma 3.2.2. \square

The next step is to estimate the approximation error in the control, which differs a bit from the pure state constrained case. Thus, the proof will be given explicitly:

Lemma 4.6.3. *Under the same assumptions as in Lemma 4.6.2, there exists a constant C , depending on α and \tilde{c} , such that*

$$\begin{aligned} \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2 & \leq C(\|y(u_\varepsilon) - y_{\varepsilon,h}\|_{1,\Omega}^2 + \|\hat{p}(y_{\varepsilon,h}) - \hat{p}_{\varepsilon,h}\|_{1,\Omega}^2 \\ & \quad + \mu_h^2(u^d) + e_c^{(1)}(u_\varepsilon, u_{\varepsilon,h}) + e_c^{(2)}(h)) \end{aligned} \quad (4.6.3)$$

Proof. Using $p_\varepsilon = \hat{p}_\varepsilon + \hat{\sigma}_\varepsilon$, $p_{\varepsilon,h} = \hat{p}_{\varepsilon,h} + \hat{\sigma}_{\varepsilon,h}$, and the fundamental relationships (4.2.20c) and (4.4.4c), we find

$$\begin{aligned} & \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2 \\ & = \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{p}_{\varepsilon,h} - \hat{p}_\varepsilon)_{0,\Omega} + \frac{\varepsilon}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \sigma_{\varepsilon,h} - \sigma_\varepsilon)_{0,\Omega} \\ & \quad + \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{\sigma}_{\varepsilon,h} - \hat{\sigma}_\varepsilon)_{0,\Omega} + (u_\varepsilon - u_{\varepsilon,h}, u^d - u_h^d)_{0,\Omega}. \end{aligned} \quad (4.6.4)$$

Now, we can split up the first term of the right-hand side of (4.6.4) according to

$$\begin{aligned} & \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{p}_{\varepsilon,h} - \hat{p}_\varepsilon)_{0,\Omega} \\ & = \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{p}_{\varepsilon,h} - \hat{p}(y_{\varepsilon,h}))_{0,\Omega} + \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{p}(y_{\varepsilon,h}) - \hat{p}_\varepsilon)_{0,\Omega}. \end{aligned} \quad (4.6.5)$$

The first term of the right-hand side of (4.6.5) can be estimated directly by an application of Cauchy's and Young's inequality. This yields

$$\frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{p}_{\varepsilon,h} - \hat{p}(y_{\varepsilon,h}))_{0,\Omega} \leq \frac{1}{16}\|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2 + \frac{4}{\alpha^2}\|\hat{p}_{\varepsilon,h} - \hat{p}(y_{\varepsilon,h})\|_{0,\Omega}^2 \quad (4.6.6)$$

In order to estimate the second term of the right-hand side of (4.6.5), one may choose $v = \hat{p}(y_{\varepsilon,h}) - \hat{p}_\varepsilon$ in (4.2.20a) and in (4.5.6a), as well as $v = y_\varepsilon - y(u_{\varepsilon,h})$ in (4.4.13) and (4.5.6b) and thus gains with the help of (3.2.5), which also holds for the regularized functions, Cauchy's and Young's inequality

$$\begin{aligned} & \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{p}(y_{\varepsilon,h}) - \hat{p}_\varepsilon)_{0,\Omega} \quad (4.6.7) \\ &= \frac{1}{\alpha}((\nabla(y_\varepsilon - y(u_{\varepsilon,h})), \nabla(\hat{p}(y_{\varepsilon,h}) - \hat{p}_\varepsilon))_{0,\Omega} + c(y_\varepsilon - y(u_{\varepsilon,h}), \hat{p}(y_{\varepsilon,h}) - \hat{p}_\varepsilon)_{0,\Omega}) \\ &= \frac{1}{\alpha}((y_{\varepsilon,h} - y^d, y_\varepsilon - y(u_{\varepsilon,h}))_{0,\Omega} + (y^d - y_\varepsilon, y_\varepsilon - y(u_{\varepsilon,h}))_{0,\Omega}) \\ &= \frac{1}{\alpha}(y_{\varepsilon,h} - y_\varepsilon, y_\varepsilon - y(u_{\varepsilon,h}))_{0,\Omega} \\ &= \frac{1}{\alpha}(y_{\varepsilon,h} - y(u_{\varepsilon,h}), y_\varepsilon - y(u_{\varepsilon,h}))_{0,\Omega} - \frac{1}{\alpha}\|y_\varepsilon - y(u_{\varepsilon,h})\|_{0,\Omega}^2 \\ &\leq \frac{1}{\alpha}\|y(u_{\varepsilon,h}) - y_{\varepsilon,h}\|_{1,\Omega} \|y_\varepsilon - y(u_{\varepsilon,h})\|_{1,\Omega} \\ &\leq \frac{\tilde{c}}{\alpha}\|y(u_{\varepsilon,h}) - y_{\varepsilon,h}\|_{1,\Omega} \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega} \\ &\leq \frac{1}{16}\|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2 + \frac{4\tilde{c}^2}{\alpha^2}\|y(u_{\varepsilon,h}) - y_{\varepsilon,h}\|_{1,\Omega}^2. \end{aligned}$$

Let us next have a closer look at the third term of the right-hand side of (4.6.4). One may choose $v_h = \hat{\sigma}_{\varepsilon,h}$ in (4.5.7) and in (4.4.4a), $v = \hat{\sigma}_\varepsilon$ in (4.2.20a) and in (4.5.6a), and $v_h = y_h(u_\varepsilon) - y_{\varepsilon,h}$ and $v = y_\varepsilon - y(u_{\varepsilon,h})$ in (4.4.16) and in (4.4.14), respectively. This results in

$$\begin{aligned} & \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{\sigma}_{\varepsilon,h} - \hat{\sigma}_\varepsilon)_{0,\Omega} \quad (4.6.8) \\ &= \frac{1}{\alpha}((\nabla(y_h(u_\varepsilon) - y_{\varepsilon,h}), \nabla\hat{\sigma}_{\varepsilon,h})_{0,\Omega} + c(y_h(u_\varepsilon) - y_{\varepsilon,h}, \hat{\sigma}_{\varepsilon,h})_{0,\Omega} \\ &\quad - (\nabla(y_\varepsilon - y(u_{\varepsilon,h})), \nabla\hat{\sigma}_\varepsilon)_{0,\Omega} - c(y_\varepsilon - y(u_{\varepsilon,h}), \hat{\sigma}_\varepsilon)_{0,\Omega}) \\ &= \frac{1}{\alpha}((\sigma_{\varepsilon,h}, y_h(u_\varepsilon) - y_{\varepsilon,h})_{0,\Omega} + (\sigma_\varepsilon, y(u_{\varepsilon,h}) - y_\varepsilon)_{0,\Omega}). \end{aligned}$$

Now, we combine (4.6.8) with the second term of (4.6.4). Then, one gets with the help of the continuous complementarity conditions (4.2.20d), the discrete comple-

mentarity conditions (4.4.12), and the definitions (4.5.8):

$$\begin{aligned}
& \frac{\varepsilon}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \sigma_{\varepsilon,h} - \sigma_\varepsilon)_{0,\Omega} + \frac{1}{\alpha}(u_\varepsilon - u_{\varepsilon,h}, \hat{\sigma}_{\varepsilon,h} - \hat{\sigma}_\varepsilon)_{0,\Omega} \\
&= \frac{1}{\alpha}((\sigma_{\varepsilon,h}, \varepsilon u_\varepsilon + y_h(u_\varepsilon) - (\varepsilon u_{\varepsilon,h} + y_{\varepsilon,h}))_{0,\Omega} - (\sigma_\varepsilon, \varepsilon u_\varepsilon + y_\varepsilon - (\varepsilon u_{\varepsilon,h} + y(u_{\varepsilon,h})))_{0,\Omega}) \\
&= \frac{1}{\alpha}((\sigma_{\varepsilon,h}, \varepsilon u_\varepsilon + y_h(u_\varepsilon) - \psi_h)_{0,\Omega} + \underbrace{(\sigma_{\varepsilon,h}, \psi_h - (\varepsilon u_{\varepsilon,h} + y_{\varepsilon,h}))_{0,\Omega}}_{= e_c^{(2)}(h)}) \\
&\quad + (\sigma_\varepsilon, \varepsilon u_{\varepsilon,h} + y(u_{\varepsilon,h}) - \psi)_{0,\Omega} + \underbrace{(\sigma_\varepsilon, \psi - \varepsilon u_\varepsilon - y_\varepsilon)_{0,\Omega}}_{= 0} \\
&\leq \frac{1}{\alpha}e_c^{(1)}(u_\varepsilon, u_{\varepsilon,h}) + \frac{1}{\alpha}e_c^{(2)}(h)
\end{aligned} \tag{4.6.9}$$

The last term of (4.6.4) can be estimated easily with the help of Cauchy's and Young's inequality as follows:

$$(u_\varepsilon - u_{\varepsilon,h}, u^d - u_h^d)_{0,\Omega} \leq \frac{1}{8}\|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega}^2 + 2\mu_h^2(u^d) \tag{4.6.10}$$

Now, the combination of (4.6.4), (4.6.5), (4.6.6), (4.6.7), (4.6.9), and (4.6.10) allows to conclude the claimed inequality. \square

It is worthwhile to mention that, due to the discrete complementarity conditions of Corollary 2.4.2, the term corresponding to the term which has been estimated against the consistency error $e_c^{(2)}(h)$ in (4.6.9) vanishes in the pure state constrained case (compare (3.2.12)).

Finally, one has to estimate the approximation of the auxiliary state and of the auxiliary modified adjoint state. That will be done by the following lemma.

Lemma 4.6.4. *Under the same assumptions as in Lemma 4.6.3, there holds:*

$$\|y(u_{\varepsilon,h}) - y_{\varepsilon,h}\|_{1,\Omega} \preceq \eta_{y_\varepsilon} , \tag{4.6.11a}$$

$$\|\hat{p}(y_{\varepsilon,h}) - \hat{p}_{\varepsilon,h}\|_{1,\Omega} \preceq \eta_{\hat{p}_\varepsilon} . \tag{4.6.11b}$$

Proof. Since Galerkin orthogonality holds for both cases, the claim is standard in a posteriori error analysis (cf. [Ver96]). \square

Proof of Theorem 4.6.1. The claim is a direct consequence of Lemma 4.6.2, Lemma 4.6.3, and Lemma 4.6.4. \square

4.7 Efficiency of the error estimator

In the final section of this chapter, we show that the error estimator for the mixed control-state constrained case also provides efficiency. The main result of this section is the following theorem.

Theorem 4.7.1. *Let $(y_\varepsilon, u_\varepsilon, p_\varepsilon, \sigma_\varepsilon)$ and $(y_{\varepsilon,h}, u_{\varepsilon,h}, p_{\varepsilon,h}, \sigma_{\varepsilon,h})$ be the unique solutions of (4.2.20) and (4.4.4), respectively, and let η_ε and $\text{osc}_h(y^d)$ be the error estimator and the data oscillation as given by (4.5.1) and (3.1.5), respectively. Further, let \hat{p}_ε and $\hat{p}_{\varepsilon,h}$ be the modified adjoint states as given by (4.4.13) and (4.4.15), respectively. Then, there holds*

$$\eta_\varepsilon \preceq \|y_\varepsilon - y_{\varepsilon,h}\|_{1,\Omega} + \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\Omega} + \|\hat{p}_\varepsilon - \hat{p}_{\varepsilon,h}\|_{1,\Omega} + \text{osc}_h(y^d). \quad (4.7.1)$$

The proof of Theorem 4.7.1 will again be done by establishing local efficiency of the local parts of the error estimator. Thus, Theorem 4.7.1 is a direct consequence of the next two lemmas. The first one reads as follows:

Lemma 4.7.2. *In addition to the assumptions of Theorem 4.7.1, let $\eta_{y_\varepsilon,T}$ and $\eta_{\hat{p}_\varepsilon,T}$, $T \in \mathcal{T}_h(\Omega)$, be the element residuals as given by (4.5.3) and let $\text{osc}_T(y^d)$, $T \in \mathcal{T}_h(\Omega)$, denote the data oscillation as defined in (3.1.5). Then, there holds for each triangle $T \in \mathcal{T}_h(\Omega)$*

$$\eta_{y_\varepsilon,T}^2 \preceq \|y_\varepsilon - y_{\varepsilon,h}\|_{1,T}^2 + h_T^2 \|u_\varepsilon - u_{\varepsilon,h}\|_{0,T}^2, \quad (4.7.2)$$

$$\eta_{\hat{p}_\varepsilon,T}^2 \preceq \|\hat{p}_\varepsilon - \hat{p}_{\varepsilon,h}\|_{1,T}^2 + h_T^2 \|y_\varepsilon - y_{\varepsilon,h}\|_{0,T}^2 + \text{osc}_T^2(y^d). \quad (4.7.3)$$

Proof. The proof follows the lines of the proofs of Lemma 3.3.1 and of Lemma 3.3.2. \square

Finally, one has to deal with the edge residuals.

Lemma 4.7.3. *In addition to the assumptions of Theorem 4.7.1, let $\eta_{y_\varepsilon,T}$, $\eta_{\hat{p}_\varepsilon,T}$, $T \in \mathcal{T}_h(\Omega)$, and $\eta_{y_\varepsilon,E}$, $\eta_{\hat{p}_\varepsilon,E}$, $E \in \mathcal{E}_h(\Omega)$, be the element and edge residuals as given by (4.5.3) and (4.5.4), respectively. Then, there holds for each interior edge $E \in \mathcal{E}_h(\Omega)$*

$$\eta_{y_\varepsilon,E}^2 \preceq \|y_\varepsilon - y_{\varepsilon,h}\|_{1,\omega_E}^2 + h_E^2 \|u_\varepsilon - u_{\varepsilon,h}\|_{0,\omega_E}^2 + \sum_{\nu=1}^2 \eta_{y_\varepsilon,T_\nu}^2, \quad (4.7.4)$$

$$\eta_{\hat{p}_\varepsilon,E}^2 \preceq \|\hat{p}_\varepsilon - \hat{p}_{\varepsilon,h}\|_{1,\omega_E}^2 + h_E^2 \|y_\varepsilon - y_{\varepsilon,h}\|_{0,\omega_E}^2 + \sum_{\nu=1}^2 \eta_{\hat{p}_\varepsilon,T_\nu}^2, \quad (4.7.5)$$

where T_1 and T_2 are the two triangles which share the edge E and $\omega_E := T_1 \cup T_2$.

Proof. The proof follows the lines of the proof of Lemma 3.3.3 and of Lemma 3.3.4. \square

Remark 4.7.4. *It is worthwhile to mention that we do not get an upper or lower bound for the approximation error of the adjoint state in the L^2 -norm. This is due to the fact that in the fundamental relationships (4.2.20c) and (4.4.4c), the adjoint control appears.*

Chapter 5

Numerical results

In this chapter, we derive adaptive algorithms based on the above developed a posteriori error estimators in order to solve an optimal control problem with pointwise state constraints. The pure state constrained case and the Lavrentiev regularization method will both be approached. We illustrate the performance of the adaptive algorithms by two numerical test examples. For both examples, the exact solution is known, and thus, the approximation errors can be calculated exactly. We will also compare the pure state constrained case with the Lavrentiev regularization method for different values of the regularization parameter.

5.1 The adaptive algorithm

We apply the standard adaptive algorithm, as it is known for partial differential equations, of the form

$$\mathbf{SOLVE} \rightarrow \mathbf{ESTIMATE} \rightarrow \mathbf{MARK} \rightarrow \mathbf{REFINE} \quad (5.1.1)$$

to our optimal control problem with pointwise state constraints.

For the step **SOLVE**, we used a primal-dual active set strategy as proposed in [BK02] in order to solve the discrete optimality system (2.4.5). For further details and a convergence proof, we refer to this paper. For the mixed control-state constrained case, a primal-dual active set strategy was proposed in [MPT06], which was adopted here. While each iteration of the primal-dual active set strategy for the pure state constrained case involves only the state, the adjoint state, and the adjoint control, the primal-dual active set strategy corresponding to the Lavrentiev regularization also has to determine the control. Consequently, in each iteration step the latter one has to solve a linear system with four times the number of degrees of freedom as the number of unknowns, whereas for the unregularized case, the number of unknowns of the linear system is only three times the number of degrees of freedom. This also has to be taken into account if one compares the efficiency of these two methods with each other.

The step **ESTIMATE** consists of the computation of the quantities of the error estimator and the data oscillation. This also requires the evaluation of the discrete modified adjoint state \hat{p}_h , which is not computed by the primal-dual active set strategy. However, the right-hand side $y^d - y_h$ has already been calculated by the algorithm, and thus, the modified adjoint state can be determined as the discrete solution of a partial differential equation on the actual grid. Compared to the computational requirements of the primal-dual active set strategy, this effort seems to be negligible.

In the step **MARK**, we selected a subset of edges and a subset of triangles which both should be refined for the next iteration. This is done by a bulk criterion as it was originally proposed in [Doe96] for partial differential equations. This method has been used in the convergence proofs for conforming adaptive finite element methods for elliptic partial differential equations (compare [Doe96], [MN07], and [MNS00]), for mixed and nonconforming finite element methods (compare [CH04] and [CH05]), as well as for control constrained optimal control problems (compare [GHIK07]). Further details about the realization of the bulk criterion for our case will be given below.

Finally, the step **REFINE** refines the actual triangulation in such a way that at least each marked triangle and each marked edge is refined and a shape-regular and conforming triangulation is maintained.

Let us have a closer look at the step **MARK**. The applied bulk criterion is similar to the one used in [HHIK07]. For completeness, we will repeat it here.

Before we state the bulk criterion, let us define the discrete free boundary $\mathcal{F}_h \subset \mathcal{T}_h(\Omega)$ as follows:

$$\mathcal{F}_h := \{T \in \mathcal{T}_h(\Omega) \mid \exists n_1, n_2 \in \mathcal{N}_h(T), y_h(n_1) = \psi(n_1), y_h(n_2) < \psi(n_2)\} \quad (5.1.2)$$

In the case of mixed control-state constraints, the discrete free boundary is defined correspondingly, as it has been done in (4.4.11).

The idea of the bulk criterion is the following:

For given, universal constants $\Theta_i, 1 \leq i \leq 4$, with $0 < \Theta_i < 1$, the outcome of the algorithm should be a subset of edges $\mathcal{M}^E \subset \mathcal{E}_h(\Omega)$ and subsets of elements $\mathcal{M}^{\eta,T}, \mathcal{M}^{\mu,T}, \mathcal{M}^{osc,T} \subset \mathcal{T}_h(\Omega)$ such that

$$\Theta_1 \sum_{E \in \mathcal{E}_h(\Omega)} (\eta_{y,E}^2 + \eta_{\hat{p},E}^2) \leq \sum_{E \in \mathcal{M}^E} (\eta_{y,E}^2 + \eta_{\hat{p},E}^2), \quad (5.1.3)$$

$$\Theta_2 \sum_{T \in \mathcal{T}_h(\Omega)} (\eta_{y,T}^2 + \eta_{\hat{p},T}^2) \leq \sum_{T \in \mathcal{M}^{\eta,T}} (\eta_{y,T}^2 + \eta_{\hat{p},T}^2), \quad (5.1.4)$$

$$\Theta_3 \sum_{T \in \mathcal{T}_h(\Omega)} \mu_T^2(u^d) \leq \sum_{T \in \mathcal{M}^{\mu,T}} \mu_T^2(u^d), \quad (5.1.5)$$

$$\Theta_4 \sum_{T \in \mathcal{T}_h(\Omega)} osc_T^2(y^d) \leq \sum_{T \in \mathcal{M}^{osc,T}} osc_T^2(y^d). \quad (5.1.6)$$

For the Lavrentiev regularized case, we want to include lower order data oscillation in ψ , and therefore, we demand the inequality

$$\Theta_3 \sum_{T \in \mathcal{T}_h(\Omega)} (\mu_T^2(u^d) + \mu_T^2(\psi)) \leq \sum_{T \in \mathcal{M}^{\mu,T}} (\mu_T^2(u^d) + \mu_T^2(\psi)) \quad (5.1.7)$$

instead of (5.1.5). In practice, it is common to choose $\Theta_i \approx 0.7$, $1 \leq i \leq 4$. Usually, this ensures that one obtains an efficient grid in an efficient way.

Due to the appearance of consistency errors in the reliability proofs of Theorem 3.2.1 and Theorem 4.6.1, it seems to make sense to refine along the discrete free boundary a priori. This should guarantee a good resolution of the continuous free boundary and a decline of the consistency error terms in each refinement step.

Thus, we set

$$\mathcal{M}^T := \mathcal{M}^{\eta,T} \cup \mathcal{M}^{\mu,T} \cup \mathcal{M}^{osc,T} \cup \mathcal{F}_h$$

and refine an element $T \in \mathcal{T}_h(\Omega)$ regularly (i.e., subdividing it into four congruent subtriangles by joining the midpoints of the edges) if

- $T \in \mathcal{M}^T$ or
- at least two edges $E \in \mathcal{E}_h(T)$ belong to \mathcal{M}^E .

If only one edge of a triangle is marked, we refine this edge by bisection of the two triangles containing the edge.

Further refinements by bisection are only performed in order to guarantee that the refined triangulation is geometrically conforming and maintains shape-regularity.

Of course, we want the inequalities (5.1.3)-(5.1.6) to be fulfilled with as few selected edges and selected triangles as possible, i.e., with the edges and the triangles with the largest local error quantities or the largest local data oscillations. This should guarantee that we are at least close to an optimal grid. In order to achieve this goal, the bulk criterion (5.1.3)-(5.1.6) may be realized by the following so-called greedy algorithm, which involves the initialization with the discrete free boundary \mathcal{F}_h :

Algorithm (Bulk Criterion):

Step 1. Initialization

Set

$$\mathcal{M}_0^E := \emptyset, \quad k = 0 \quad \text{and} \quad l = 0.$$

Step 2. Iteration loop:

Step 2a. Check edge residuals

If

$$\Theta_1 \sum_{E \in \mathcal{E}_h(\Omega)} (\eta_{y,E}^2 + \eta_{\hat{p},E}^2) \leq \sum_{E \in \mathcal{M}_k^E} (\eta_{y,E}^2 + \eta_{\hat{p},E}^2),$$

go to Step 2b, else select

$$F \in \mathcal{E}_h(\Omega) \setminus \mathcal{M}_k^E$$

such that

$$\eta_{y,F}^2 + \eta_{\hat{p},F}^2 = \max_{G \in \mathcal{E}_H(\Omega) \setminus \mathcal{M}_k^E} \left(\eta_{y,G}^2 + \eta_{\hat{p},G}^2 \right)$$

and set

$$\mathcal{M}_{k+1}^E := \mathcal{M}_k^E \cup \{F\} \quad , \quad k := k + 1 .$$

Step 2b. Check element residuals

Set

$$\mathcal{M}_0^{\eta,T} := \mathcal{F}_h \cup \{T \in \mathcal{T}_h(\Omega) \mid \text{card} \left(\mathcal{E}_h(T) \cap \mathcal{M}_k^E \right) \geq 2\} .$$

If

$$\Theta_2 \sum_{T \in \mathcal{T}_h(\Omega)} (\eta_{y,T}^2 + \eta_{\hat{p},T}^2) \leq \sum_{T \in \mathcal{M}_l^{\eta,T}} (\eta_{y,T}^2 + \eta_{\hat{p},T}^2) ,$$

go to Step 2c, else select

$$S \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\eta,T}$$

such that

$$\eta_{y,S}^2 + \eta_{\hat{p},S}^2 = \max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\eta,T}} \left(\eta_{y,T}^2 + \eta_{\hat{p},T}^2 \right)$$

and set

$$\mathcal{M}_{l+1}^{\eta,T} := \mathcal{M}_l^{\eta,T} \cup \{S\} \quad , \quad l := l + 1 .$$

Step 2c. Check low order data residuals

Set

$$\mathcal{M}_l^{\mu,T} := \mathcal{M}_l^{\eta,T} .$$

If

$$\Theta_3 \sum_{T \in \mathcal{T}_h(\Omega)} \mu_T^2(u^d) \leq \sum_{T \in \mathcal{M}_l^{\mu,T}} \mu_T^2(u^d) ,$$

go to Step 2d, else select

$$S \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\mu,T}$$

such that

$$\mu_S(u^d) = \max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\mu,T}} \mu_T(u^d)$$

and set

$$\mathcal{M}_{l+1}^{\mu,T} := \mathcal{M}_l^{\mu,T} \cup \{S\} \quad , \quad l := l + 1 .$$

Step 2d. Check higher order data residuals

Set

$$\mathcal{M}_l^{\text{osc},T} := \mathcal{M}_l^{\mu,T} .$$

If

$$\Theta_4 \sum_{T \in \mathcal{T}_h(\Omega)} \text{osc}_T^2(y^d) \leq \sum_{T \in \mathcal{M}_l^{\text{osc},T}} \text{osc}_T^2(y^d) ,$$

go to Step 3, else select

$$S \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\text{osc},T}$$

such that

$$\text{osc}_S(y^d) = \max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\text{osc},T}} \text{osc}_T(y^d)$$

and set

$$\mathcal{M}_{l+1}^{\text{osc},T} := \mathcal{M}_l^{\text{osc},T} \cup \{S\} \quad , \quad l := l + 1 .$$

Step 3. Final output

Output the set of marked edges and elements

$$\mathcal{M}^E := \mathcal{M}_k^E \quad , \quad \mathcal{M}^T := \mathcal{M}_l^{\text{osc},T} .$$

For the case of regularized mixed control-state constraints, we also include the lower order data term of the upper bound ψ . In order to fulfill (5.1.7), we change Step 2c as follows:

Step 2c. Check low order data residuals

Set

$$\mathcal{M}_l^{\mu,T} := \mathcal{M}_l^{\eta,T} .$$

If

$$\Theta_3 \left(\sum_{T \in \mathcal{T}_h(\Omega)} (\mu_T^2(u^d) + \mu_T^2(\psi)) \right) \leq \sum_{T \in \mathcal{M}_l^{\mu,T}} (\mu_T^2(u^d) + \mu_T^2(\psi)) ,$$

go to Step 2d, else select

$$S \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\mu,T}$$

such that

$$\begin{aligned} & \max(\mu_S(u^d), \mu_S(\psi)) \\ = & \max \left(\max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\mu,T}} \mu_T(u^d), \max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_l^{\mu,T}} \mu_T(\psi) \right) \end{aligned} \quad (5.1.8)$$

and set

$$\mathcal{M}_{l+1}^{\mu,T} := \mathcal{M}_l^{\mu,T} \cup \{S\} \quad , \quad l := l + 1 .$$

Remark 5.1.1. *It would also make sense to use the sum of the two terms in (5.1.8), instead of the maximum of these terms. However, the maximum might have advantages if one of the parts of the lower order data oscillations has much larger values than the other.*

5.2 Implementation Issues

The algorithm was implemented in C++ under LINUX. The implementation was built up on the partial differential equations library of Dr. Yuri Iliash from the University of Augsburg. In order to solve the linear systems of the primal-dual active set strategy, the package UMFPACK, which contains a direct solver for sparse matrices, was applied.

In both numerical examples, we do not know the exact analytic solution of the modified adjoint state. Since we know the exact right-hand side, we calculated a discrete solution on a very fine mesh and used this function as the exact modified adjoint state. This should provide an appropriate approximation to the real formula of this function, and we are thus able to estimate the approximation error of the modified adjoint state.

Furthermore, we want to calculate the consistency error for the pure state constrained case in every refinement step. However, it involves the function $y(u_h)$, the continuous solution with the discrete right-hand side, which is also unknown. Therefore, we use an approximation which is determined as follows. After having solved the discrete optimality conditions, we know the right-hand side of $y(u_h)$. Thus, we perform four additional adaptive refinement steps with the usual a posteriori error estimator for partial differential equations in order to get a good approximation of $y(u_h)$. The function $y_h(u)$, which also appears in the consistency error, can be calculated exactly on the actual grid since the optimal control u is known in our examples.

Even though the reliability and efficiency results of the previous chapter deal with the discretization error of the continuous regularized solution, we will calculate the approximation of the discrete regularized functions to the continuous unregularized optimal solutions in the case of the Lavrentiev regularization because this is what really matters. Of course, there is a point in the adaptive algorithm when further refinements do not lead to better approximations since the error made from the regularization starts to blur the results and can not be compensated by a finer mesh. The larger the regularization parameter is chosen, the quicker this point is reached. However, the convergence of the regularized to the unregularized solution tells us that this point can be an arbitrary fine grid if the regularization parameter is sufficiently small. We did not calculate the consistency error for the mixed control-state constrained case since we do not know the optimal regularized solutions.

5.3 Example 1 - Smooth Lagrange Multiplier

In this section, we apply the above derived finite element method to a test problem which features an optimal state that strongly oscillates around the origin and whose coincidence set is a connected subdomain with smooth boundary. We first state the problem, then apply the algorithm for pure state constraints, and after that adopt

the Lavrentiev regularization. Before we denote the data of the problem, let us first define the functions which we want to have as its solution. Therefore, we define two real valued functions as follows:

$$\gamma_1(r) := \begin{cases} 1, & r < 0.25 \\ 0, & r > 0.75 \\ -192(r - 0.25)^5 + 240(r - 0.25)^4 - 80(r - 0.25)^3 + 1, & \text{otherwise} \end{cases},$$

$$\gamma_2(r) := \begin{cases} 1, & r < 0.75 \\ 0, & \text{otherwise} \end{cases}.$$

Now, we fix $y \in H_0^1(\Omega)$, $p \in H_0^1(\Omega) \subset W_0^{1,s}(\Omega)$, $u \in L^2(\Omega)$, and $\sigma \in L^2(\Omega) \subset \mathcal{M}(\Omega)$ in polar coordinates according to

$$y(r) := -r^{\frac{4}{3}}\gamma_1(r) \quad (5.3.1)$$

$$u(r) := -\Delta y(r) \quad (5.3.2)$$

$$p(r) := \gamma_2(r)\left(r^4 - \frac{3}{2}r^3 + \frac{9}{16}r^2\right) \quad (5.3.3)$$

$$\sigma(r) := \begin{cases} 0.1, & r < 0.75 \\ 0.0, & \text{otherwise} \end{cases}. \quad (5.3.4)$$

Notice that for this example, the Lagrange multipliers p and σ have higher regularity than in general. Indeed, there even holds $y, p \in H^2(\Omega)$. This will not be the case in the second test example. In order to get the above defined functions as an optimal solution, we define the data of the example as follows:

$$\begin{aligned} \Omega &:= (-2, 2)^2, \quad y^d := y(r) + \Delta p(r) + \sigma(r), \quad u^d := u(r) + \alpha^{-1} p(r), \\ \psi(r) &:= \begin{cases} y(r), & r < 0.75 \\ (r - 0.75), & \text{otherwise} \end{cases}, \quad \alpha := 0.1, \quad c = 0, \quad \Gamma_D := \partial\Omega. \end{aligned}$$

The active region, i.e., the area where the upper bound is sharp, is the interior of the circle $B_{0.75}(0)$. It is easy to check that the above defined functions are indeed the optimal solution for the given data.

A visualization of the desired state y^d and the control shift u^d can be found in Figure 5.1. The optimal state y and the optimal control u can be found in Figure 5.2, while the adjoint state p and the modified adjoint state \hat{p} are visualized in Figure 5.3.

The initial simplicial triangulation \mathcal{T}_{h_0} was chosen according to a subdivision of Ω by joining the four vertices, resulting in four congruent right-angled triangles and one interior nodal point. Then, two uniform refinement steps were performed in advance. The parameters Θ_i in the bulk criteria have been specified according to $\Theta_i = 0.75, 1 \leq i \leq 4$. Figure 5.4 shows the adaptively generated triangulations after twelve (left) and fourteen (right) refinement steps with 1376 and 3261 nodes, respectively. Here, the red colored area indicates the discrete active zone and the yellow colored area the discrete inactive zone. In particular, a triangle is colored red

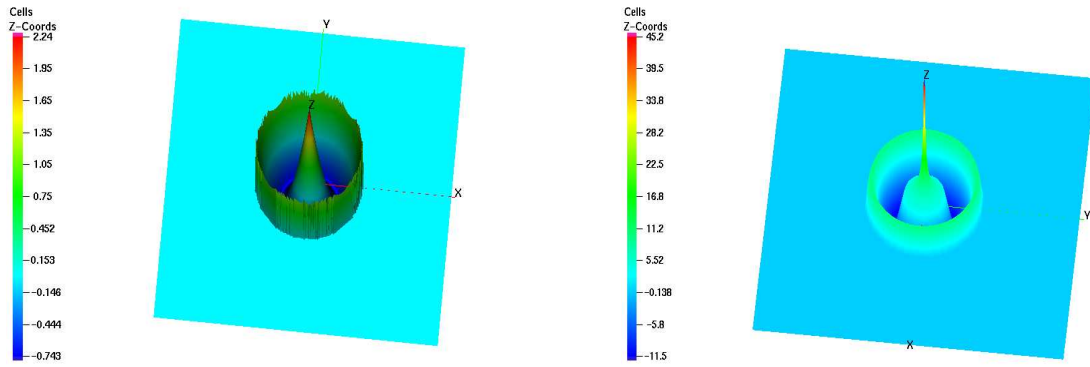


Figure 5.1: Example 1: Visualization of the desired state y^d (left) and the control shift u^d (right).

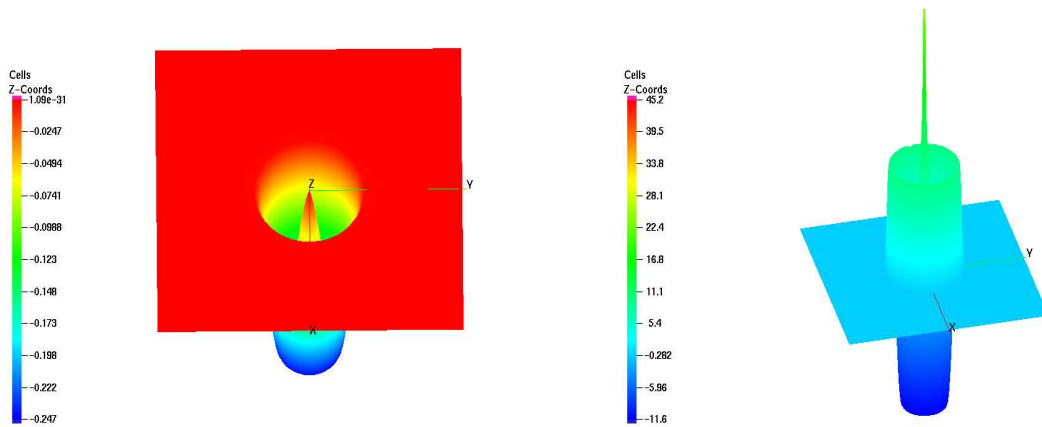


Figure 5.2: Example 1: Visualization of the optimal state y (left) and the optimal control u (right).

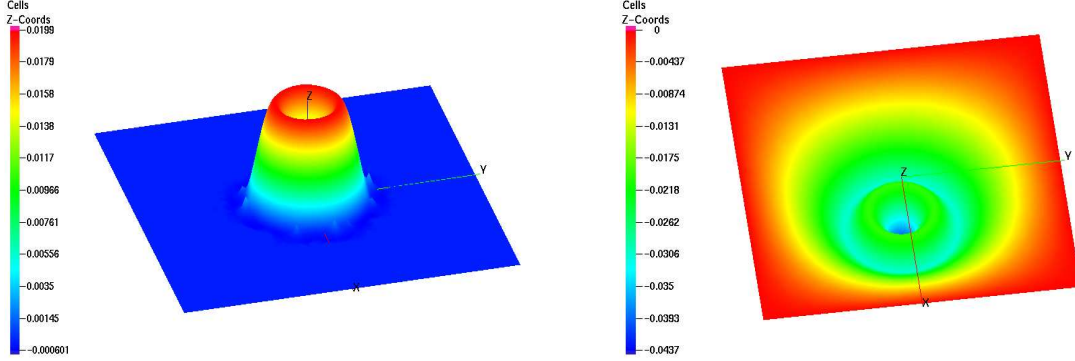


Figure 5.3: Example 1: Visualization of the adjoint state p (left) and the regularized adjoint state \hat{p} (right).

if in all of its nodes the upper bound is sharp. One can also see that the continuous free boundary, $\mathcal{F} = \{(x_1, x_2) \in \Omega \mid (x_1^2 + x_2^2)^{\frac{1}{2}} = 0.75\}$, is well resolved by the adaptive refinement process, which occurs due to the initializing step in the bulk criterion. This also leads to a good reduction of the consistency error $e_c(u, u_h)$, as it can be seen in Table 5.3. It should be emphasized that only one grid for all variables y , u , p , \hat{p} , and σ is used. Hence, the grid reflects regions of substantial change in all these variables.

More detailed information can be found in Table 5.1 to Table 5.4. Here, ℓ stands for the refinement level, and if ℓ is used as an index, it describes the corresponding discrete function at level ℓ . In particular, Table 5.1 displays the error reduction in the total error,

$$\|z - z_\ell\| := \|y - y_\ell\|_{1,\Omega} + \|u - u_\ell\|_{0,\Omega} ,$$

the H^1 -error in the state, the L^2 -errors in the control and in the adjoint state, and the H^1 -error in the modified adjoint state. The actual components η_y and $\eta_{\hat{p}}$ of the residual type a posteriori error estimator, the data oscillations $\mu_h(u^d)$ and $\text{osc}_h(y^d)$, and the consistency error $e_c(u, u_h)$ are given in Table 5.2, whereas Table 5.3 contains the average values of the local element and edge contributions of the error estimator, as well as the average values of the data oscillations. One may notice that the consistency error is only non-vanishing at the very beginning of the algorithm. Finally, Table 5.4 lists the percentages of elements and edges that have been marked for refinement according to the bulk criteria or the initializing step. Here, $M^{fb,T}$, $M^{\eta,T}$, $M^{\mu,T}$, and $M^{\text{osc},T}$ stand for the level ℓ elements marked for refinement due to the resolution of the discrete free boundary, the element residuals, and the lower and higher order data oscillation, respectively, whereas M^E refers to the edges marked for refinement with regard to the edge residuals. As it is the case for the two coarsest grids, the sum of the percentages may exceed 100 % since an element $T \in \mathcal{T}_h(\Omega)$ may satisfy more than one criterion in the bulk criteria. One

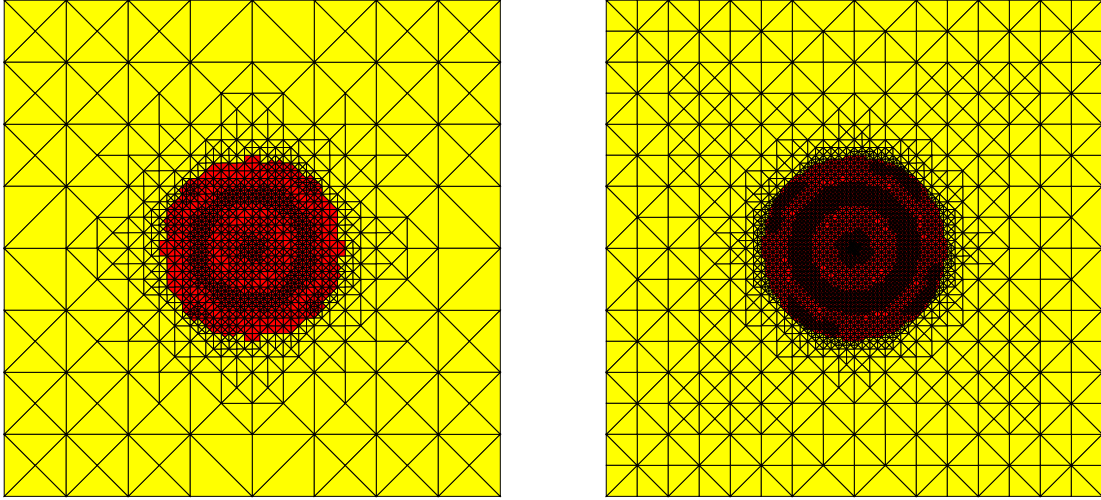


Figure 5.4: Example 1: Adaptively generated grid after 12 (left) and 14 (right) refinement steps with 1376 and 3261 nodes, respectively, $\Theta_i = 0.75$.

can see that at the very beginning, the refinement is dominated by the resolution of the discrete free boundary, whereas at later stages the element residuals prevail. The edge residuals do not play a dominant role since in this example both functions, the state and the modified adjoint state, are smooth, i.e., we have $y, \hat{p} \in H^2(\Omega)$.

Figure 5.5 shows the benefit of the adaptive versus the uniform refinement by displaying the discretization errors as a function of the number of degrees of freedom. The axes are logarithmically scaled. The left picture shows the benefit for the state y and the right one for the control u . Especially for the approximation of the state, one may notice that the line is steeper for the adaptive case compared to the uniform one. This benefit will actually be more noticeable in the second example.

Now, we also apply the Lavrentiev regularization with different regularization parameters 10^{-2} , 10^{-4} , and 10^{-6} . We use the same initial triangulation as before, and again, we apply two uniform refinement steps first.

We noticed that for this example the regularization parameter $\varepsilon = 0.01$ is too large and quickly blurs the results. For smaller parameters, the results are, up to a 16000 degrees of freedom, comparable to the pure state constrained case. This can already be seen if one draws the discrete solutions. Figure 5.6 shows the adaptively generated states y for the regularization parameters $\varepsilon = 10^{-2}$ (left) and $\varepsilon = 10^{-4}$ (right) on grids with comparable number of degrees of freedom.

The adaptively generated grids for the regularization parameter $\varepsilon = 10^{-6}$ after 12 and 14 refinement steps with 1523 and 3985 degrees of freedom, respectively, are shown in Figure 5.7. The discrete active region is a bit fuzzy due to the regularization, and thus, we did not colorize this area in red in the pictures. One can notice that the adaptively generated grids for the Lavrentiev regularization looks similar

Table 5.1: Example 1: Convergence history of the adaptive FEM, Part I: Total discretization error and discretization errors in the state, control, adjoint state, and modified adjoint state.

| ℓ | N | $\ z - z_\ell\ $ | $\ y - y_\ell\ _1$ | $\ u - u_\ell\ _0$ | $\ p - p_\ell\ _0$ | $\ \hat{p} - \hat{p}_\ell\ _1$ |
|--------|------|------------------|--------------------|--------------------|--------------------|--------------------------------|
| 0 | 5 | 2.38e+01 | 2.79e+00 | 2.10e+01 | 7.68e-01 | 1.11e+00 |
| 1 | 13 | 2.53e+01 | 1.73e+00 | 2.36e+01 | 1.09e+00 | 9.82e-01 |
| 2 | 41 | 1.46e+01 | 1.02e+00 | 1.36e+01 | 1.30e-02 | 1.32e-01 |
| 3 | 53 | 1.16e+01 | 9.56e-01 | 1.06e+01 | 1.24e-02 | 1.06e-01 |
| 4 | 69 | 1.02e+01 | 7.44e-01 | 9.42e+00 | 2.86e-02 | 5.91e-02 |
| 5 | 97 | 8.94e+00 | 6.03e-01 | 8.34e+00 | 7.57e-03 | 4.08e-02 |
| 6 | 133 | 6.58e+00 | 5.44e-01 | 6.04e+00 | 5.96e-03 | 3.89e-02 |
| 7 | 201 | 4.24e+00 | 3.67e-01 | 3.87e+00 | 5.58e-03 | 3.00e-02 |
| 8 | 303 | 3.40e+00 | 2.79e-01 | 3.12e+00 | 3.00e-03 | 2.51e-02 |
| 9 | 441 | 3.16e+00 | 2.43e-01 | 2.92e+00 | 6.72e-03 | 2.28e-02 |
| 10 | 669 | 2.66e+00 | 2.08e-01 | 2.45e+00 | 5.86e-03 | 2.06e-02 |
| 11 | 940 | 2.06e+00 | 1.72e-01 | 1.89e+00 | 3.06e-03 | 1.61e-02 |
| 12 | 1376 | 1.59e+00 | 1.29e-01 | 1.46e+00 | 1.29e-03 | 1.23e-02 |
| 13 | 2125 | 1.28e+00 | 1.04e-01 | 1.18e+00 | 1.62e-03 | 1.03e-02 |
| 14 | 3261 | 1.08e+00 | 7.64e-02 | 9.99e-01 | 1.13e-03 | 7.56e-03 |
| 15 | 5409 | 9.29e-01 | 6.00e-02 | 8.69e-01 | 8.29e-04 | 5.49e-03 |
| 16 | 9108 | 8.09e-01 | 4.50e-02 | 7.64e-01 | 7.68e-04 | 3.77e-03 |

Table 5.2: Example 1: Convergence history of the adaptive FEM, Part II: Components of the error estimator, the data oscillations, and the consistency error.

| ℓ | N_{dof} | η_y | $\eta_{\hat{p}}$ | $\mu_\ell(u^d)$ | $\text{osc}_\ell(y^d)$ | $e_c(u, u_\ell)$ |
|--------|------------------|----------|------------------|-----------------|------------------------|------------------|
| 0 | 5 | 3.95e+01 | 6.19e+00 | 1.48e+01 | 3.93e+00 | 8.09e-01 |
| 1 | 13 | 2.18e+01 | 1.88e+00 | 1.37e+01 | 8.60e-01 | 0.00e+00 |
| 2 | 41 | 9.91e+00 | 6.98e-01 | 1.36e+01 | 6.29e-01 | 7.40e-01 |
| 3 | 53 | 5.17e+00 | 5.83e-01 | 1.06e+01 | 5.06e-01 | 3.64e-01 |
| 4 | 69 | 3.68e+00 | 3.80e-01 | 9.42e+00 | 3.40e-01 | 0.00e+00 |
| 5 | 97 | 2.21e+00 | 2.69e-01 | 8.33e+00 | 2.01e-01 | 0.00e+00 |
| 6 | 133 | 1.72e+00 | 1.85e-01 | 6.02e+00 | 1.03e-01 | 0.00e+00 |
| 7 | 201 | 1.49e+00 | 1.48e-01 | 3.87e+00 | 8.04e-02 | 0.00e+00 |
| 8 | 303 | 1.20e+00 | 1.20e-01 | 3.11e+00 | 5.61e-02 | 0.00e+00 |
| 9 | 441 | 9.88e-01 | 1.03e-01 | 2.91e+00 | 4.71e-02 | 0.00e+00 |
| 10 | 669 | 8.25e-01 | 8.57e-02 | 2.44e+00 | 3.71e-02 | 0.00e+00 |
| 11 | 940 | 7.03e-01 | 6.70e-02 | 1.89e+00 | 2.65e-02 | 0.00e+00 |
| 12 | 1376 | 5.61e-01 | 5.07e-02 | 1.46e+00 | 1.68e-02 | 0.00e+00 |
| 13 | 2125 | 4.41e-01 | 3.83e-02 | 1.18e+00 | 1.06e-02 | 0.00e+00 |
| 14 | 3261 | 3.47e-01 | 2.95e-02 | 9.99e-01 | 7.23e-03 | 0.00e+00 |
| 15 | 5409 | 2.63e-01 | 2.28e-02 | 8.69e-01 | 5.26e-03 | 0.00e+00 |
| 16 | 9108 | 2.01e-01 | 1.76e-02 | 7.64e-01 | 4.04e-03 | 0.00e+00 |

Table 5.3: Example 1: Convergence history of the adaptive FEM, Part III: Average values of the local estimators.

| ℓ | N_{dof} | $\eta_{y,T}$ | $\eta_{\hat{p},T}$ | $\eta_{y,E}$ | $\eta_{\hat{p},E}$ | $\mu_{\ell}(u^d)$ | $\text{osc}_{\ell}(y^d)$ |
|--------|------------------|--------------|--------------------|--------------|--------------------|-------------------|--------------------------|
| 0 | 5 | 1.87e+01 | 1.69e+00 | 6.41e+00 | 2.59e+00 | 7.40e+00 | 1.96e+00 |
| 1 | 13 | 3.57e+00 | 1.64e-01 | 7.48e-01 | 2.62e-01 | 1.71e+00 | 1.07e-01 |
| 2 | 41 | 6.01e-01 | 3.33e-02 | 3.50e-02 | 3.77e-03 | 7.57e-01 | 2.65e-02 |
| 3 | 53 | 3.35e-01 | 2.95e-02 | 3.85e-02 | 5.20e-03 | 6.47e-01 | 2.30e-02 |
| 4 | 69 | 2.15e-01 | 2.05e-02 | 2.55e-02 | 3.82e-03 | 5.24e-01 | 1.65e-02 |
| 5 | 97 | 1.02e-01 | 1.32e-02 | 1.42e-02 | 2.14e-03 | 3.71e-01 | 8.27e-03 |
| 6 | 133 | 7.55e-02 | 8.19e-03 | 8.97e-03 | 1.36e-03 | 2.60e-01 | 3.98e-03 |
| 7 | 201 | 5.64e-02 | 5.42e-03 | 4.53e-03 | 7.78e-04 | 1.21e-01 | 2.31e-03 |
| 8 | 303 | 3.78e-02 | 3.23e-03 | 2.38e-03 | 4.16e-04 | 6.49e-02 | 1.11e-03 |
| 9 | 441 | 2.61e-02 | 2.19e-03 | 1.38e-03 | 2.63e-04 | 4.34e-02 | 6.96e-04 |
| 10 | 669 | 1.75e-02 | 1.49e-03 | 7.71e-04 | 1.66e-04 | 2.51e-02 | 4.15e-04 |
| 11 | 940 | 1.26e-02 | 1.07e-03 | 4.47e-04 | 9.77e-05 | 1.52e-02 | 2.51e-04 |
| 12 | 1376 | 8.62e-03 | 7.11e-04 | 2.27e-04 | 5.48e-05 | 8.55e-03 | 1.35e-04 |
| 13 | 2125 | 5.58e-03 | 4.52e-04 | 1.20e-04 | 3.05e-05 | 4.49e-03 | 6.78e-05 |
| 14 | 3261 | 3.64e-03 | 2.93e-04 | 5.76e-05 | 1.62e-05 | 2.42e-03 | 3.48e-05 |
| 15 | 5409 | 2.19e-03 | 1.75e-04 | 2.73e-05 | 7.26e-06 | 1.20e-03 | 1.69e-05 |
| 16 | 9108 | 1.30e-03 | 1.04e-04 | 1.22e-05 | 3.58e-06 | 5.89e-04 | 7.86e-06 |

Table 5.4: Example 1: Convergence history of the adaptive FEM, Part IV: Percentages in the bulk criteria.

| ℓ | N_{dof} | $\mathcal{M}^{fb,T}$ | \mathcal{M}^E | $\mathcal{M}^{\eta,T}$ | $\mathcal{M}^{\mu,T}$ | $\mathcal{M}^{\text{osc},T}$ |
|--------|------------------|----------------------|-----------------|------------------------|-----------------------|------------------------------|
| 0 | 5 | 0.0 | 100.0 | 75.0 | 75.0 | 75.0 |
| 1 | 13 | 25.0 | 25.0 | 25.0 | 25.0 | 18.8 |
| 2 | 41 | 0.0 | 15.9 | 12.5 | 9.4 | 6.2 |
| 3 | 53 | 18.2 | 12.1 | 27.3 | 18.2 | 13.6 |
| 4 | 69 | 13.3 | 7.6 | 27.5 | 22.5 | 15.0 |
| 5 | 97 | 27.3 | 7.8 | 23.3 | 18.2 | 17.0 |
| 6 | 133 | 19.4 | 15.9 | 24.6 | 24.2 | 17.7 |
| 7 | 201 | 43.8 | 15.0 | 30.5 | 15.9 | 9.4 |
| 8 | 303 | 17.0 | 11.6 | 30.4 | 7.5 | 5.6 |
| 9 | 441 | 12.0 | 9.4 | 33.1 | 5.2 | 4.7 |
| 10 | 669 | 8.8 | 6.0 | 30.6 | 3.0 | 3.6 |
| 11 | 940 | 7.6 | 5.5 | 30.6 | 2.2 | 3.7 |
| 12 | 1376 | 6.9 | 11.1 | 36.2 | 1.3 | 3.2 |
| 13 | 2125 | 5.6 | 9.3 | 35.6 | 0.2 | 2.6 |
| 14 | 3261 | 4.5 | 12.7 | 41.5 | 0.1 | 2.2 |
| 15 | 5409 | 3.5 | 16.4 | 44.7 | 0.1 | 1.5 |
| 16 | 9108 | 2.7 | 14.7 | 44.1 | 0.0 | 0.7 |

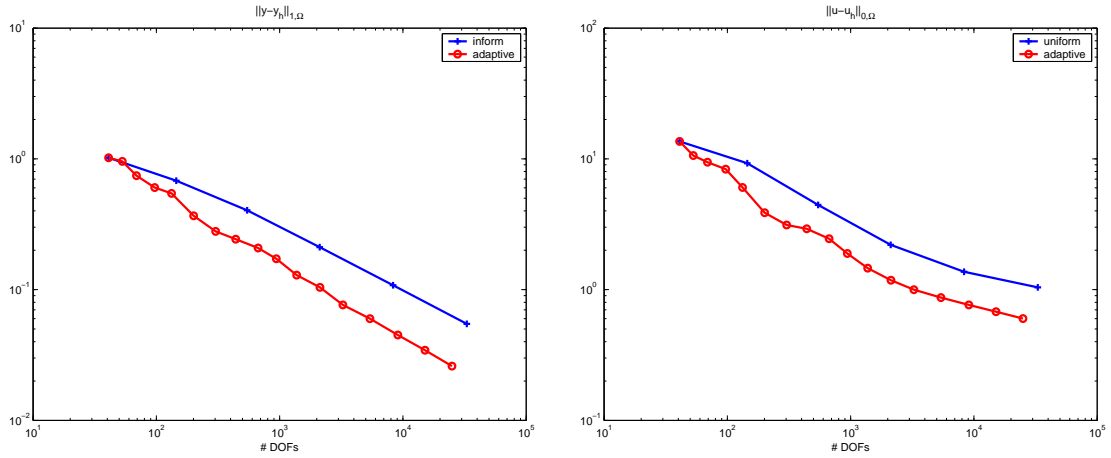


Figure 5.5: Example 1: Adaptive versus uniform refinement for the state (left) and the control (right).

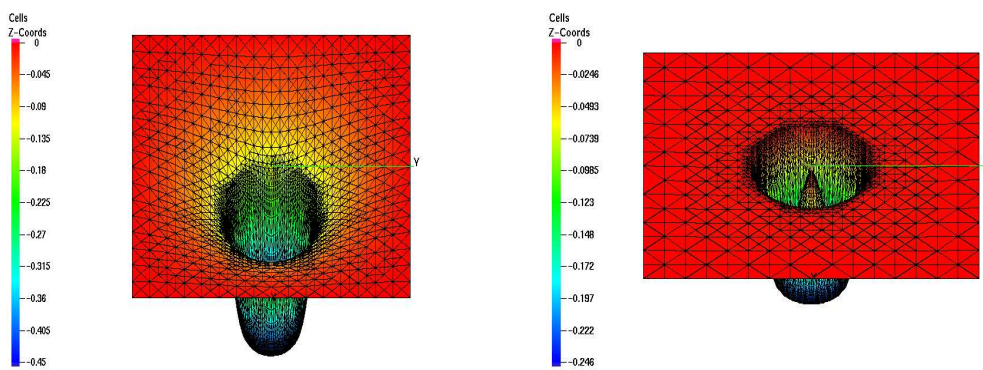


Figure 5.6: Example 1: Adaptively generated state y for the regularization parameter $\varepsilon = 10^{-2}$ (left) and $\varepsilon = 10^{-4}$ (right).

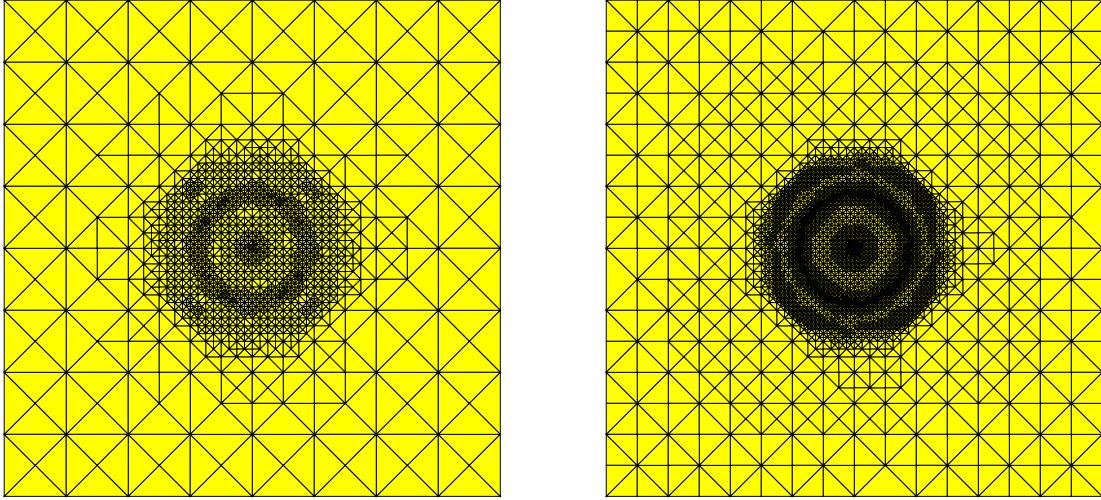


Figure 5.7: Example 1: Adaptively generated grid after 12 (left) and 14 (right) refinement steps with 1523 and 3985 nodes, respectively, with regularization parameter $\varepsilon = 10^{-6}$, $\Theta_i = 0.75$, $1 \leq i \leq 4$.

to the unregularized grids (compare Figure 5.4).

Table 5.5 to Table 5.8 provide the same information as it was given for the unconstrained case. We mention again that in Table 5.5, we are comparing the discrete regularized solutions with the continuous unregularized functions since this is the goal which should be achieved. We see that even though a regularization was applied, the discretization errors in the control and the state decrease in every refinement step. However, the regularization starts to blur the approximation of the adjoint state and the modified adjoint state quickly.

In Figure 5.8, one may see that also for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, the adaptive finite element method provides a benefit compared to an uniform refinement strategy for the approximation error in the state and in the control.

Next, we compare the convergence depending on the regularization parameter ε . For the value $\varepsilon = 10^{-2}$, the error resulting from the regularization quickly dominates the discretization error, and consequently, the perturbed problem does not lead to a good approximation of the real solution. However, for the values 10^{-4} and 10^{-6} , the results are, at least for the approximation error in the state and in the control, up to 16000 degrees of freedom, comparable to the unregularized case. This can be derived from Figure 5.9, which shows the approximation error in the state and the control for different regularization parameters and the unconstrained case.

Finally, we want to compare the number of iterations the applied primal-dual active set strategies need to find the exact solution. For the pure state constrained case, the applied primal-dual active set strategy is not well-defined for the continuous problem. Therefore, the number of iterations the algorithm needs until the solution is reached, increases with the number of degrees of freedom. One of the advantages of

Table 5.5: Example 1: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part I: Total discretization error and discretization errors in the state, control, adjoint state, and modified adjoint state.

| ℓ | N_{dof} | $\ z - z_\ell\ $ | $\ y - y_\ell\ _1$ | $\ u - u_\ell\ _0$ | $\ p - p_\ell\ _0$ | $\ \hat{p} - \hat{p}_\ell\ _1$ |
|--------|------------------|------------------|--------------------|--------------------|--------------------|--------------------------------|
| 0 | 5 | 1.49e+01 | 6.33e-01 | 1.43e+01 | 9.49e-02 | 2.54e-01 |
| 1 | 13 | 1.47e+01 | 1.02e+00 | 1.37e+01 | 3.01e-02 | 2.01e-01 |
| 2 | 41 | 1.19e+01 | 1.68e+00 | 1.02e+01 | 8.59e-01 | 1.99e-01 |
| 3 | 62 | 1.13e+01 | 1.20e+00 | 1.01e+01 | 2.58e-01 | 9.51e-02 |
| 4 | 76 | 1.08e+01 | 1.32e+00 | 9.45e+00 | 7.53e-02 | 1.11e-01 |
| 5 | 106 | 7.77e+00 | 9.18e-01 | 8.46e+00 | 6.79e-02 | 7.57e-02 |
| 6 | 151 | 6.58e+00 | 6.46e-01 | 5.93e+00 | 3.50e-02 | 4.21e-02 |
| 7 | 217 | 4.81e+00 | 4.75e-01 | 4.33e+00 | 2.88e-02 | 3.75e-02 |
| 8 | 310 | 3.61e+00 | 3.51e-01 | 3.26e+00 | 1.81e-02 | 3.97e-02 |
| 9 | 437 | 3.07e+00 | 2.89e-01 | 2.78e+00 | 8.89e-03 | 3.43e-02 |
| 10 | 655 | 2.24e+00 | 2.14e-01 | 2.03e+00 | 7.38e-03 | 3.35e-02 |
| 11 | 949 | 1.73e+00 | 1.58e-01 | 1.57e+00 | 9.75e-03 | 3.65e-02 |
| 12 | 1523 | 1.40e+00 | 1.21e-01 | 1.28e+00 | 1.18e-02 | 3.43e-02 |
| 13 | 2459 | 1.15e+00 | 9.33e-02 | 1.06e+00 | 1.28e-02 | 3.37e-02 |
| 14 | 3985 | 9.89e-01 | 7.08e-02 | 9.18e-01 | 1.46e-02 | 3.31e-02 |
| 15 | 6247 | 8.61e-01 | 5.80e-02 | 8.03e-01 | 1.56e-02 | 3.33e-02 |
| 16 | 10525 | 7.58e-01 | 4.48e-02 | 7.13e-01 | 1.61e-02 | 3.35e-02 |

Table 5.6: Example 1: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part II: Components of the error estimator and the data oscillations.

| ℓ | N_{dof} | η_{y_ε} | $\eta_{\hat{p}_\varepsilon}$ | $\mu_\ell(u^d)$ | $\mu_\ell(\psi)$ | $osc_\ell(y^d)$ |
|--------|------------------|------------------------|------------------------------|-----------------|------------------|-----------------|
| 0 | 5 | 4.11e+00 | 3.20e+00 | 1.48e+01 | 2.50e+00 | 3.93e+00 |
| 1 | 13 | 7.48e-01 | 1.23e+00 | 1.37e+01 | 9.10e-01 | 8.60e-01 |
| 2 | 41 | 4.51e+00 | 8.47e-01 | 1.36e+01 | 2.64e-01 | 6.29e-01 |
| 3 | 62 | 4.74e+00 | 5.99e-01 | 1.06e+01 | 1.68e-01 | 5.06e-01 |
| 4 | 76 | 3.85e+00 | 4.09e-01 | 9.59e+00 | 1.03e-01 | 3.40e-01 |
| 5 | 106 | 2.39e+00 | 2.96e-01 | 8.54e+00 | 8.11e-02 | 2.16e-01 |
| 6 | 151 | 1.71e+00 | 2.02e-01 | 5.95e+00 | 7.22e-02 | 1.19e-01 |
| 7 | 217 | 1.44e+00 | 1.51e-01 | 4.36e+00 | 6.54e-02 | 7.95e-02 |
| 8 | 310 | 1.21e+00 | 1.18e-01 | 3.28e+00 | 6.39e-02 | 5.64e-02 |
| 9 | 437 | 1.01e+00 | 9.66e-02 | 2.79e+00 | 6.34e-02 | 4.37e-02 |
| 10 | 655 | 8.28e-01 | 7.45e-02 | 2.04e+00 | 5.80e-02 | 3.11e-02 |
| 11 | 949 | 6.68e-01 | 5.70e-02 | 1.57e+00 | 3.48e-02 | 2.13e-02 |
| 12 | 1523 | 5.12e-01 | 4.46e-02 | 1.27e+00 | 2.27e-02 | 1.52e-02 |
| 13 | 2459 | 3.95e-01 | 3.39e-02 | 1.05e+00 | 1.74e-02 | 1.01e-02 |
| 14 | 3985 | 3.06e-01 | 2.66e-02 | 9.05e-01 | 1.11e-02 | 7.82e-03 |
| 15 | 6247 | 2.41e-01 | 2.10e-02 | 7.88e-01 | 6.26e-03 | 6.11e-03 |
| 16 | 10525 | 1.83e-01 | 1.68e-02 | 6.95e-01 | 4.19e-03 | 5.00e-03 |

Table 5.7: Example 1: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part III: Average values of the local estimators.

| ℓ | N_{dof} | $\eta_{y_\varepsilon, T}$ | $\eta_{\hat{p}_\varepsilon, T}$ | $\eta_{y_\varepsilon, E}$ | $\eta_{\hat{p}_\varepsilon, E}$ | $\mu_\ell(u^d)$ | $\mu_\ell(\psi)$ | $\text{osc}_\ell(y^d)$ |
|--------|------------------|---------------------------|---------------------------------|---------------------------|---------------------------------|-----------------|------------------|------------------------|
| 0 | 5 | 1.34e+00 | 1.55e+00 | 1.55e+00 | 3.88e-01 | 7.40e+00 | 1.25e+00 | 1.96e+00 |
| 1 | 13 | 1.36e-01 | 1.81e-01 | 7.42e-02 | 9.71e-02 | 1.71e+00 | 2.03e-01 | 1.07e-01 |
| 2 | 41 | 3.99e-01 | 6.95e-02 | 1.30e-01 | 1.79e-02 | 7.57e-01 | 2.64e-02 | 2.65e-02 |
| 3 | 62 | 2.75e-01 | 2.68e-02 | 5.63e-02 | 4.44e-03 | 5.37e-01 | 1.23e-02 | 1.91e-02 |
| 4 | 76 | 1.99e-01 | 2.10e-02 | 4.86e-02 | 4.88e-03 | 4.80e-01 | 7.58e-03 | 1.48e-02 |
| 5 | 106 | 1.07e-01 | 1.31e-02 | 1.95e-02 | 2.52e-03 | 3.51e-01 | 4.85e-03 | 7.92e-03 |
| 6 | 151 | 6.73e-02 | 7.52e-03 | 8.03e-03 | 1.12e-03 | 2.15e-01 | 3.07e-03 | 3.81e-03 |
| 7 | 217 | 5.12e-02 | 4.76e-03 | 4.55e-03 | 6.39e-04 | 1.18e-01 | 1.83e-03 | 2.07e-03 |
| 8 | 310 | 3.72e-02 | 3.09e-03 | 2.47e-03 | 3.51e-04 | 6.98e-02 | 1.18e-03 | 1.17e-03 |
| 9 | 437 | 2.68e-02 | 2.19e-03 | 1.48e-03 | 2.32e-04 | 4.46e-02 | 8.02e-04 | 7.43e-04 |
| 10 | 655 | 1.81e-02 | 1.43e-03 | 7.65e-04 | 1.38e-04 | 2.40e-02 | 4.88e-04 | 4.13e-04 |
| 11 | 949 | 1.27e-02 | 9.57e-04 | 3.99e-04 | 7.75e-05 | 1.40e-02 | 2.76e-04 | 2.26e-04 |
| 12 | 1523 | 7.88e-03 | 5.88e-04 | 1.89e-04 | 3.86e-05 | 6.92e-03 | 1.42e-04 | 1.09e-04 |
| 13 | 2459 | 4.86e-03 | 3.59e-04 | 9.21e-05 | 2.02e-05 | 3.53e-03 | 7.43e-05 | 5.40e-05 |
| 14 | 3985 | 3.01e-03 | 2.20e-04 | 4.29e-05 | 1.03e-05 | 1.86e-03 | 3.67e-05 | 2.73e-05 |
| 15 | 6247 | 1.92e-03 | 1.39e-04 | 2.23e-05 | 4.98e-06 | 9.66e-04 | 1.82e-05 | 1.41e-05 |
| 16 | 10525 | 1.14e-03 | 8.22e-05 | 9.86e-06 | 2.38e-06 | 4.90e-04 | 8.63e-06 | 7.06e-06 |

Table 5.8: Example 1: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part IV: Percentages in the bulk criteria.

| ℓ | N_{dof} | $\mathcal{M}^{fb, T}$ | \mathcal{M}^E | $\mathcal{M}^{\eta, T}$ | $\mathcal{M}^{\mu, T}$ | $\mathcal{M}^{\text{osc}, T}$ |
|--------|------------------|-----------------------|-----------------|-------------------------|------------------------|-------------------------------|
| 0 | 5 | 0.0 | 100.0 | 75.0 | 75.0 | 75.0 |
| 1 | 13 | 25.0 | 35.0 | 25.0 | 25.0 | 18.8 |
| 2 | 41 | 25.0 | 14.8 | 29.7 | 9.4 | 6.2 |
| 3 | 62 | 32.0 | 10.6 | 18.9 | 15.1 | 11.3 |
| 4 | 76 | 0.0 | 20.7 | 23.1 | 20.9 | 13.4 |
| 5 | 106 | 0.0 | 14.5 | 20.1 | 18.0 | 14.4 |
| 6 | 151 | 5.5 | 13.6 | 21.5 | 18.7 | 13.4 |
| 7 | 217 | 7.4 | 13.1 | 26.9 | 11.5 | 9.1 |
| 8 | 310 | 3.7 | 13.1 | 30.2 | 9.0 | 9.0 |
| 9 | 437 | 2.5 | 15.4 | 33.1 | 6.2 | 6.7 |
| 10 | 655 | 2.2 | 10.6 | 31.0 | 4.3 | 5.4 |
| 11 | 949 | 2.8 | 15.0 | 36.2 | 3.1 | 3.9 |
| 12 | 1523 | 2.8 | 15.6 | 40.2 | 0.6 | 2.2 |
| 13 | 2459 | 2.9 | 14.4 | 42.2 | 0.2 | 1.8 |
| 14 | 3985 | 2.8 | 15.1 | 41.2 | 0.1 | 1.3 |
| 15 | 6247 | 2.9 | 16.9 | 45.8 | 0.1 | 0.9 |
| 16 | 10525 | 3.0 | 16.3 | 47.9 | 0.0 | 0.6 |

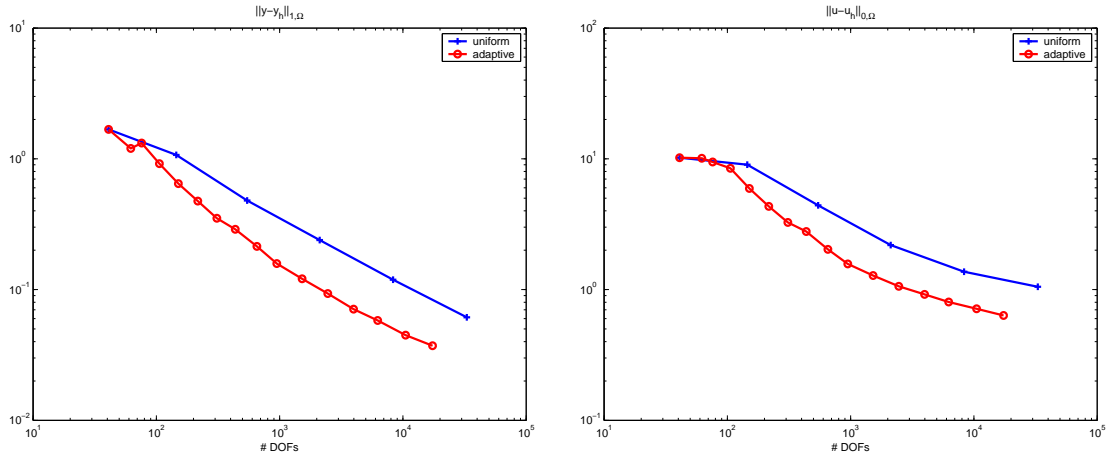


Figure 5.8: Example 1: Adaptive versus uniform refinement for the state (left) and the control (right) with regularization parameter $\varepsilon = 10^{-6}$.

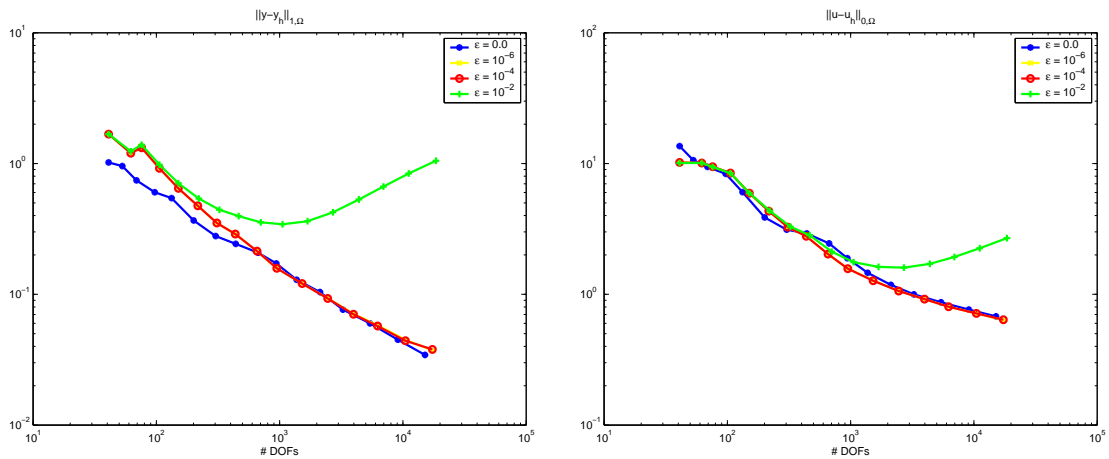


Figure 5.9: Example 1: Comparison of the approximation error in the state (left) and the control (right) for different regularization parameters.

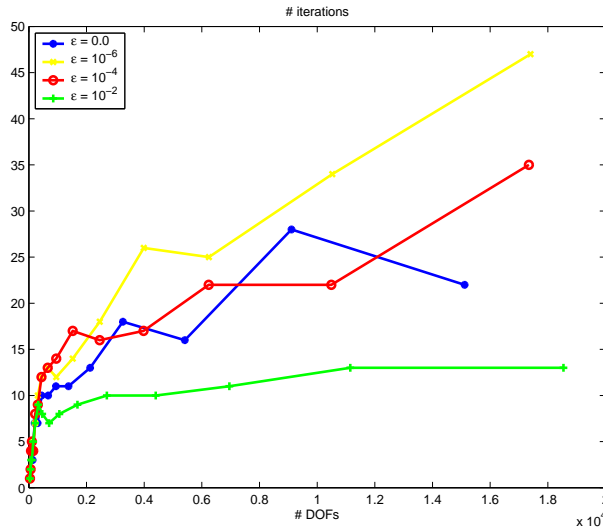


Figure 5.10: Example 1: Comparison of the number of iterations the active set strategy needs to find the exact solution for different regularization parameters and for the unregularized case.

the Lavrentiev regularization is that the primal-dual active set strategy (and interior-point methods) is also well-defined in the continuous setting. Consequently, one may hope that the number of iterations the algorithm needs is smaller than for the pure state constrained case. This, however, does not seem to be the case. Figure 5.10 gives a comparison of the iteration numbers the active set strategy needs to determine the exact solution depending on the number of degrees of freedom for different regularization parameters and the unregularized case. One may see that only for the regularization parameter $\varepsilon = 10^{-2}$, one gets fewer iteration numbers compared to the pure state constrained case. This value, however, is too big and does not lead to an useful solution. For smaller regularization parameters, the number of iterations is equal or even slightly larger than for the unregularized case. This behavior is probably due to the fact that the whole system converges to the unregularized one, as the regularization parameter tends to 0 and is thus for the continuous case also undefined at the limit. Taking into account that for the Lavrentiev regularization, the linear systems which have to be solved in each iteration step are larger, the Lavrentiev regularization does not provide a benefit in efficiency compared to the unregularized case for this example.

5.4 Example 2 - Nonsmooth Lagrange multiplier

The second example is taken from [MPT06]. The interesting feature of this example is that the involved Lagrange multipliers are not smooth. Here, $p \in W^{1,s}(\Omega)$ holds only for $1 < s < 2$ and the optimal adjoint control σ is not a L^2 -function but indeed

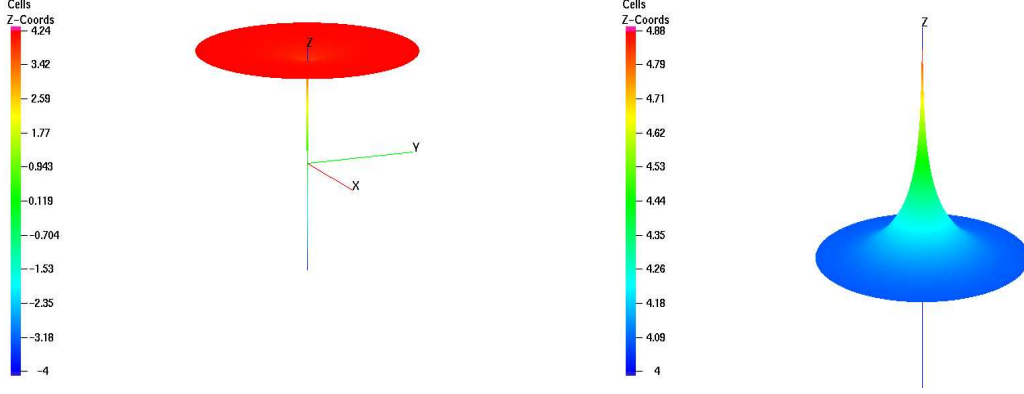


Figure 5.11: Example 2: Visualization of the desired state y^d (left) and the control shift u^d (right).

only a measure. In particular, the data of the problem in polar coordinates reads as follows:

$$\begin{aligned} \Omega &:= B(0,1) \quad , \quad \Gamma_D = \emptyset \quad , \quad y^d(r) := 4 + \frac{1}{\pi} - \frac{1}{4\pi}r^2 + \frac{1}{2\pi}\ln(r) \quad , \\ u^d(r) &:= 4 + \frac{1}{4\pi}r^2 - \frac{1}{2\pi}\ln(r) \quad , \quad \alpha := 1.0 \quad , \quad c := 1.0 \quad , \quad \psi(r) := r + 4 \quad . \end{aligned}$$

For this data, it can easily be shown that the optimal solution is given according to

$$\begin{aligned} y(r) &\equiv 4 \quad , \quad p(r) = \frac{1}{4\pi}r^2 - \frac{1}{2\pi}\ln(r) \quad , \\ u(r) &\equiv 4 \quad , \quad \sigma = \delta_0 \quad , \end{aligned}$$

where δ_0 denotes the Dirac-measure concentrated at the origin. While the control and the state are smooth, the adjoint state p has a singularity at the origin, and thus, it only lives in the spaces $W^{1,s}(\Omega)$, $1 < s < 2$. Notice that the state is only active at the origin. Therefore, the active zone is degenerated to the single point $(0,0)^T$.

Figures 5.11, 5.12, and 5.13 show a visualization of the desired state y^d and the control shift u^d , as well as the discrete state y_ℓ , the discrete control u_ℓ , the discrete adjoint state p_ℓ , and the discrete modified adjoint state \bar{p}_ℓ with respect to a simplicial triangulation consisting of 6735 nodal points. Even though the optimal state and the optimal control could be approximated exactly on any grid, especially for the control, there is still some error occurring around the origin. In Figure 5.13, one may also notice the smoothness of modified adjoint state \hat{p} compared to the adjoint state p .

The initial simplicial triangulation \mathcal{T}_{h_0} has been chosen by means of the five nodal points $(0,0)$, $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$, resulting in five congruent triangles. During

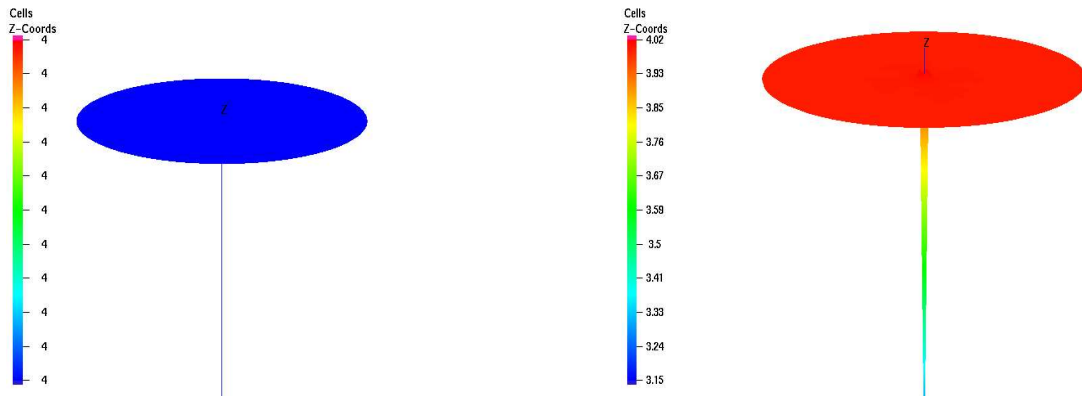


Figure 5.12: Example 2: Visualization of the discrete state y_l (left) and the discrete control u_l (right) on an adaptive generated mesh with 6735 nodes.

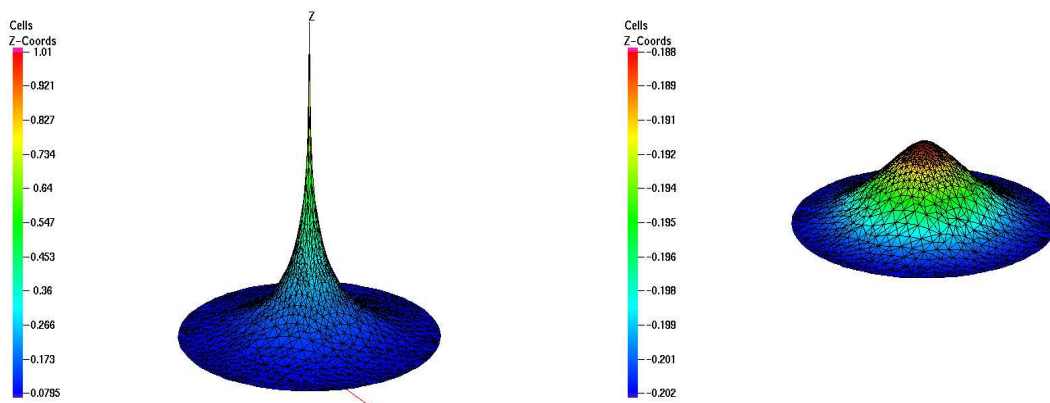


Figure 5.13: Example 2: Visualization of the discrete adjoint state p_l (left) and the discrete modified adjoint state \bar{p}_l (right) on an adaptive generated mesh with 6735 nodes.

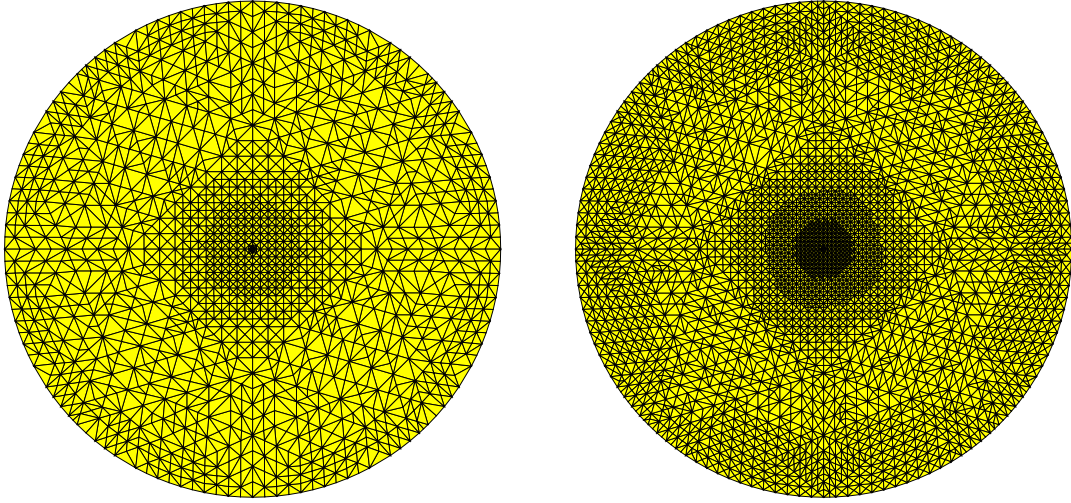


Figure 5.14: Example 2: Adaptively generated grid after 12 (left) and 14 (right) refinement steps with 1621 and 2491 nodal points, respectively, $\Theta_i = 0.7$.

the refinement process, each new point on a boundary edge has been projected onto the boundary $\partial B(0,1)$. Again, it was refined uniformly in the first two iteration steps.

One can notice that during the refinement process the optimal state y is even on coarse grids almost exactly approximated, while for the discrete controls, some error occurs around the origin (compare Figure 5.12). Thus, the adaptive refinement process should especially refine at this part of the domain.

In the adaptive algorithm, we fixed $\Theta_i = 0.7$, $1 \leq i \leq 4$. Figure 5.14 shows the adaptive generated grids after 12 and 14 refinement steps with 1621 and 2491 nodal points, respectively. Notice that indeed a high density of nodes occurs around the origin. For this example, it does not make a difference whether one refines along the discrete free boundary or not because all discrete solutions are only sharp at the origin and, due to the large local terms of the data oscillation in y^d and u^d around the origin, this area would be refined anyway.

Table 5.9 to Table 5.12 contain the same data as in Example 1, documenting the history of the adaptive refinement process. Figure 5.15 displays again the approximation errors in the control and in the state for the adaptive and for the uniform refinement. For both approximations, one can see that a great benefit is gained by the application of the adaptive strategy. This is remarkable because for partial differential equations such a noticeable benefit normally occurs only for nonconvex domains since these may provide solutions with singularities. However, due to the appearance of measures in the right-hand side of the adjoint state, the optimal solution of a state constrained optimal control problem may even contain singularities on convex domains. This is the case in this example, and it explains the great benefit noticed.

Table 5.9: Example 2: Convergence history of the adaptive FEM, Part I: Total discretization error and discretization errors in the state, control, adjoint state, and modified adjoint state.

| ℓ | N_{dof} | $\ z - z_\ell\ $ | $\ y - y_\ell\ _1$ | $\ u - u_\ell\ _0$ | $\ p - p_\ell\ _0$ | $\ \hat{p} - \hat{p}_\ell\ _1$ |
|--------|------------------|------------------|--------------------|--------------------|--------------------|--------------------------------|
| 0 | 5 | 1.55e-01 | 1.20e-02 | 1.43e-01 | 6.46e-02 | 3.81e-02 |
| 1 | 13 | 1.13e-01 | 8.51e-03 | 1.04e-01 | 3.73e-02 | 1.74e-02 |
| 2 | 41 | 7.39e-02 | 4.43e-03 | 6.95e-02 | 1.86e-02 | 9.01e-03 |
| 3 | 57 | 7.55e-02 | 1.59e-03 | 7.39e-02 | 1.07e-02 | 7.91e-03 |
| 4 | 73 | 5.96e-02 | 2.30e-03 | 5.73e-02 | 1.00e-02 | 7.36e-03 |
| 5 | 89 | 4.68e-02 | 2.57e-03 | 4.42e-02 | 9.83e-03 | 7.10e-03 |
| 6 | 121 | 3.60e-02 | 1.79e-03 | 3.42e-02 | 7.41e-03 | 6.11e-03 |
| 7 | 159 | 2.74e-02 | 1.30e-03 | 2.61e-02 | 5.24e-03 | 4.75e-03 |
| 8 | 243 | 2.10e-02 | 1.07e-03 | 1.99e-02 | 4.13e-03 | 4.02e-03 |
| 9 | 389 | 1.58e-02 | 6.53e-04 | 1.51e-02 | 2.82e-03 | 3.17e-03 |
| 10 | 604 | 1.18e-02 | 4.02e-04 | 1.14e-02 | 1.95e-03 | 2.43e-03 |
| 11 | 965 | 8.81e-03 | 2.71e-04 | 8.54e-03 | 1.36e-03 | 1.87e-03 |
| 12 | 1621 | 6.55e-03 | 1.60e-04 | 6.39e-03 | 9.26e-04 | 1.52e-03 |
| 13 | 2491 | 4.87e-03 | 1.02e-04 | 4.77e-03 | 6.49e-04 | 1.17e-03 |
| 14 | 3991 | 3.62e-03 | 6.81e-05 | 3.55e-03 | 4.55e-04 | 8.79e-04 |
| 15 | 6735 | 2.67e-03 | 3.97e-05 | 2.63e-03 | 3.14e-04 | 6.79e-04 |

Table 5.10: Example 2: Convergence history of the adaptive FEM, Part II: Components of the error estimator, the data oscillations, and the consistency error.

| ℓ | N_{dof} | η_y | $\eta_{\hat{p}}$ | $\mu_\ell(u^d)$ | $osc_\ell(y^d)$ | $e_c(u, u_\ell)$ |
|--------|------------------|----------|------------------|-----------------|-----------------|------------------|
| 0 | 5 | 1.91e-01 | 1.38e-01 | 1.73e-01 | 1.36e-01 | 0.00e+00 |
| 1 | 13 | 7.32e-02 | 7.62e-02 | 1.29e-01 | 4.36e-02 | 0.00e+00 |
| 2 | 41 | 2.45e-02 | 3.83e-02 | 8.14e-02 | 1.26e-02 | 0.00e+00 |
| 3 | 57 | 1.85e-02 | 2.97e-02 | 7.60e-02 | 9.79e-03 | 0.00e+00 |
| 4 | 73 | 1.02e-02 | 2.54e-02 | 5.95e-02 | 7.78e-03 | 0.00e+00 |
| 5 | 89 | 5.64e-03 | 2.37e-02 | 4.63e-02 | 6.75e-03 | 0.00e+00 |
| 6 | 121 | 3.11e-03 | 1.97e-02 | 3.56e-02 | 4.96e-03 | 0.00e+00 |
| 7 | 159 | 1.67e-03 | 1.66e-02 | 2.71e-02 | 3.12e-03 | 0.00e+00 |
| 8 | 243 | 9.10e-04 | 1.32e-02 | 2.06e-02 | 1.87e-03 | 0.00e+00 |
| 9 | 389 | 4.89e-04 | 1.05e-02 | 1.55e-02 | 1.34e-03 | 0.00e+00 |
| 10 | 604 | 2.59e-04 | 8.07e-03 | 1.17e-02 | 8.27e-04 | 0.00e+00 |
| 11 | 965 | 1.37e-04 | 6.12e-03 | 8.75e-03 | 4.93e-04 | 0.00e+00 |
| 12 | 1621 | 7.22e-05 | 4.75e-03 | 6.54e-03 | 3.16e-04 | 0.00e+00 |
| 13 | 2491 | 3.83e-05 | 3.75e-03 | 4.87e-03 | 2.16e-04 | 0.00e+00 |
| 14 | 3991 | 2.01e-05 | 2.89e-03 | 3.62e-03 | 1.41e-04 | 0.00e+00 |
| 15 | 6735 | 1.05e-05 | 2.23e-03 | 2.68e-03 | 8.65e-05 | 0.00e+00 |

Table 5.11: Example 2: Convergence history of the adaptive FEM, Part III: Average values of the local estimators.

| ℓ | N_{dof} | $\eta_{y,T}$ | $\eta_{\bar{p},T}$ | $\eta_{y,E}$ | $\eta_{\bar{p},E}$ | $\mu_\ell(u^d)$ | $\text{osc}_\ell(y^d)$ |
|--------|------------------|--------------|--------------------|--------------|--------------------|-----------------|------------------------|
| 0 | 5 | 9.95e-02 | 6.73e-02 | 8.01e-03 | 1.43e-02 | 8.67e-02 | 6.80e-02 |
| 1 | 13 | 1.03e-02 | 1.58e-02 | 1.01e-03 | 3.21e-03 | 1.91e-02 | 8.67e-03 |
| 2 | 41 | 9.39e-04 | 3.61e-03 | 8.46e-05 | 6.85e-04 | 3.39e-03 | 1.08e-03 |
| 3 | 57 | 6.68e-04 | 2.60e-03 | 1.77e-05 | 4.44e-04 | 2.82e-03 | 7.29e-04 |
| 4 | 73 | 3.20e-04 | 1.94e-03 | 1.99e-05 | 3.14e-04 | 1.75e-03 | 4.88e-04 |
| 5 | 89 | 1.79e-04 | 1.63e-03 | 1.66e-05 | 2.46e-04 | 1.23e-03 | 3.86e-04 |
| 6 | 121 | 8.10e-05 | 1.19e-03 | 8.36e-06 | 1.74e-04 | 7.40e-04 | 2.49e-04 |
| 7 | 159 | 3.43e-05 | 8.53e-04 | 3.15e-06 | 1.03e-04 | 4.24e-04 | 1.45e-04 |
| 8 | 243 | 1.44e-05 | 5.36e-04 | 1.33e-06 | 5.34e-05 | 2.16e-04 | 7.00e-05 |
| 9 | 389 | 5.73e-06 | 3.40e-04 | 4.24e-07 | 2.83e-05 | 1.07e-04 | 3.73e-05 |
| 10 | 604 | 2.20e-06 | 2.15e-04 | 1.37e-07 | 1.44e-05 | 5.28e-05 | 1.85e-05 |
| 11 | 965 | 7.51e-07 | 1.30e-04 | 4.10e-08 | 6.73e-06 | 2.50e-05 | 8.58e-06 |
| 12 | 1621 | 2.80e-07 | 7.80e-05 | 1.24e-08 | 3.36e-06 | 1.15e-05 | 4.07e-06 |
| 13 | 2491 | 1.22e-07 | 5.04e-05 | 4.23e-09 | 1.73e-06 | 5.77e-06 | 2.12e-06 |
| 14 | 3991 | 4.32e-08 | 3.09e-05 | 1.31e-09 | 8.29e-07 | 2.76e-06 | 1.04e-06 |
| 15 | 6735 | 1.61e-08 | 1.83e-05 | 3.70e-10 | 3.93e-07 | 1.27e-06 | 4.82e-07 |

Table 5.12: Example 2: Convergence history of the adaptive FEM, Part IV: Percentages in the bulk criteria.

| ℓ | N_{dof} | $\mathcal{M}^{fb,T}$ | \mathcal{M}^E | $\mathcal{M}^{\eta,T}$ | $\mathcal{M}^{\mu,T}$ | $\mathcal{M}^{\text{osc},T}$ |
|--------|------------------|----------------------|-----------------|------------------------|-----------------------|------------------------------|
| 0 | 5 | 100.0 | 75.0 | 75.0 | 75.0 | 75.0 |
| 1 | 13 | 25.0 | 30.0 | 25.0 | 31.2 | 25.0 |
| 2 | 41 | 6.3 | 27.3 | 7.8 | 6.2 | 12.5 |
| 3 | 57 | 8.7 | 16.4 | 17.4 | 7.6 | 18.5 |
| 4 | 73 | 6.5 | 10.8 | 25.8 | 5.6 | 17.7 |
| 5 | 89 | 5.3 | 8.3 | 30.9 | 4.6 | 15.8 |
| 6 | 121 | 3.9 | 7.5 | 30.0 | 3.4 | 16.4 |
| 7 | 159 | 2.8 | 21.3 | 33.2 | 2.5 | 21.9 |
| 8 | 243 | 1.8 | 11.3 | 31.7 | 1.6 | 21.4 |
| 9 | 389 | 1.2 | 8.3 | 32.6 | 0.8 | 17.5 |
| 10 | 604 | 0.7 | 15.2 | 37.9 | 0.5 | 17.1 |
| 11 | 965 | 0.4 | 10.2 | 40.8 | 0.3 | 17.7 |
| 12 | 1621 | 0.3 | 8.1 | 39.8 | 0.2 | 15.9 |
| 13 | 2491 | 0.2 | 7.9 | 43.5 | 0.1 | 11.4 |
| 14 | 3991 | 0.1 | 8.7 | 47.4 | 0.1 | 9.8 |
| 15 | 6735 | 0.0 | 5.2 | 45.0 | 0.0 | 9.7 |

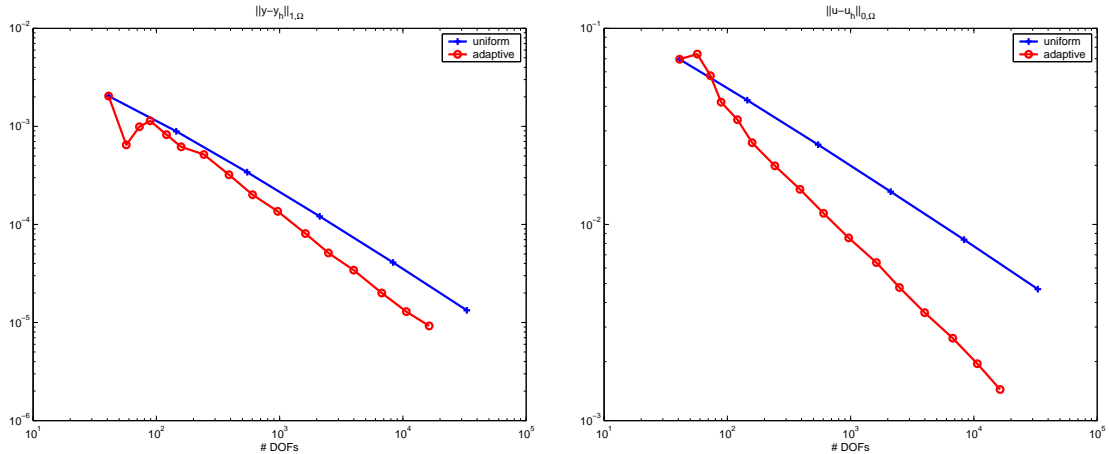


Figure 5.15: Example 2: Adaptive versus uniform refinement for the state (left) and the control (right).

Like in the previous example, we want to compare the results of the pure state constrained case with the Lavrentiev regularization. Again, we used the values 10^{-2} , 10^{-4} , and 10^{-6} as regularization parameters. Similarly, for sufficiently small ε , the adaptively generated grids look equal to the unconstrained ones, as in Figure 5.14. Table 5.13 to Table 5.16 consist of the same information as for the pure state constrained case for a regularization parameter $\varepsilon = 10^{-6}$. Now, in contrast to the previous example, also the discretization error in the adjoint state and in the modified adjoint state increases at each iteration step until the termination of the algorithm.

Again, we compare the adaptive refinement to an uniform refinement strategy for the regularization parameter $\varepsilon = 10^{-6}$. As one can see in Figure 5.16, the adaptive algorithm also provides a noticeable benefit for the approximation error in the state and the control in the corresponding norms.

Like in the previous example, we have a look at the behavior for different regularization parameters. Figure 5.17 shows the approximation error in the state and the control for the regularization parameters 10^{-2} , 10^{-4} , and 10^{-6} and for the unregularized case. Again, it can be seen that the value $\varepsilon = 10^{-2}$ is too big and does not lead to useful results. The behavior for the regularization parameters $\varepsilon = 10^{-4}$ is interesting. While it is comparable to the unregularized case up to a number of degrees of freedom of 16000 for the approximation error in the control, this is not the case for the approximation error in the state. Here, the regularization starts at around 1200 degrees of freedom to blur the results. Again, the regularization parameter $\varepsilon = 10^{-6}$ is comparable to the unregularized case up to 16000 degrees of freedom and is even slightly better for the approximation error in the state.

Finally, we have again a look at the number of iterations the primal-dual active set strategy needs to find the exact solution. This is visualized in Figure 5.18. Like in the previous example, the Lavrentiev regularization does not provide a computational

Table 5.13: Example 2: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part I: Total discretization error and discretization errors in the state, control, adjoint state, and modified adjoint state.

| ℓ | N_{dof} | $\ z - z_\ell\ $ | $\ y - y_\ell\ _1$ | $\ u - u_\ell\ _0$ | $\ p - p_\ell\ _0$ | $\ \hat{p} - \hat{p}_\ell\ _1$ |
|--------|------------------|------------------|--------------------|--------------------|--------------------|--------------------------------|
| 0 | 5 | 6.28e-02 | 9.62e-03 | 5.32e-02 | 1.48e-01 | 3.54e-02 |
| 1 | 13 | 1.04e-01 | 1.75e-02 | 8.62e-02 | 6.73e-02 | 2.24e-02 |
| 2 | 41 | 7.51e-02 | 1.01e-02 | 6.50e-02 | 2.93e-02 | 1.21e-02 |
| 3 | 57 | 8.17e-02 | 1.15e-02 | 7.02e-02 | 2.26e-02 | 1.20e-02 |
| 4 | 73 | 6.25e-02 | 6.89e-03 | 5.56e-02 | 1.54e-02 | 9.36e-03 |
| 5 | 89 | 4.74e-02 | 3.91e-03 | 4.35e-02 | 1.19e-02 | 7.73e-03 |
| 6 | 121 | 3.63e-02 | 2.35e-03 | 3.39e-02 | 8.30e-03 | 6.02e-03 |
| 7 | 159 | 2.73e-02 | 1.33e-03 | 2.60e-02 | 5.70e-03 | 4.68e-03 |
| 8 | 243 | 2.05e-02 | 6.96e-04 | 1.98e-02 | 4.35e-03 | 3.91e-03 |
| 9 | 389 | 1.55e-02 | 4.04e-04 | 1.51e-02 | 2.92e-03 | 3.04e-03 |
| 10 | 603 | 1.16e-02 | 2.25e-04 | 1.14e-02 | 2.00e-03 | 2.32e-03 |
| 11 | 964 | 8.67e-03 | 1.25e-04 | 8.54e-03 | 1.38e-03 | 1.82e-03 |
| 12 | 1618 | 6.46e-03 | 6.98e-05 | 6.39e-03 | 9.37e-04 | 1.46e-03 |
| 13 | 2485 | 4.81e-03 | 4.07e-05 | 4.77e-03 | 6.56e-04 | 1.14e-03 |
| 14 | 3989 | 3.57e-03 | 2.58e-05 | 3.55e-03 | 4.57e-04 | 8.54e-04 |
| 15 | 6721 | 2.65e-03 | 1.50e-05 | 2.63e-03 | 3.16e-04 | 6.61e-04 |
| 16 | 10656 | 1.96e-03 | 1.10e-05 | 1.95e-03 | 2.21e-04 | 4.76e-04 |

Table 5.14: Example 2: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part II: Components of the error estimator and the data oscillations.

| ℓ | N_{dof} | η_{y_ε} | $\eta_{\hat{p}_\varepsilon}$ | $\mu_\ell(u^d)$ | $\mu_\ell(\psi)$ | $osc_\ell(y^d)$ |
|--------|------------------|------------------------|------------------------------|-----------------|------------------|-----------------|
| 0 | 5 | 7.51e-02 | 1.34e-01 | 1.73e-01 | 2.11e-01 | 1.36e-01 |
| 1 | 13 | 6.15e-02 | 7.38e-02 | 1.29e-01 | 1.11e-01 | 4.36e-02 |
| 2 | 41 | 2.29e-02 | 3.76e-02 | 8.14e-02 | 3.25e-02 | 1.26e-02 |
| 3 | 57 | 1.77e-02 | 2.92e-02 | 7.60e-02 | 2.42e-02 | 9.79e-03 |
| 4 | 73 | 1.00e-02 | 2.52e-02 | 5.95e-02 | 2.13e-02 | 7.78e-03 |
| 5 | 89 | 5.65e-03 | 2.35e-02 | 4.63e-02 | 1.79e-02 | 6.75e-03 |
| 6 | 121 | 3.11e-03 | 2.01e-02 | 3.56e-02 | 1.23e-02 | 4.96e-03 |
| 7 | 159 | 1.68e-03 | 1.66e-02 | 2.71e-02 | 8.29e-03 | 3.12e-03 |
| 8 | 243 | 9.15e-04 | 1.32e-02 | 2.06e-02 | 5.27e-03 | 1.87e-03 |
| 9 | 389 | 4.89e-04 | 1.05e-02 | 1.55e-02 | 3.62e-03 | 1.34e-03 |
| 10 | 603 | 2.59e-04 | 8.12e-03 | 1.17e-02 | 2.25e-03 | 8.28e-04 |
| 11 | 964 | 1.37e-04 | 6.13e-03 | 8.75e-03 | 1.41e-03 | 4.95e-04 |
| 12 | 1618 | 7.23e-05 | 4.76e-03 | 6.54e-03 | 9.86e-04 | 3.17e-04 |
| 13 | 2485 | 3.84e-05 | 3.76e-03 | 4.87e-03 | 6.68e-04 | 2.19e-04 |
| 14 | 3989 | 2.01e-05 | 2.89e-03 | 3.62e-03 | 3.95e-04 | 1.42e-04 |
| 15 | 6721 | 1.05e-05 | 2.24e-03 | 2.68e-03 | 2.54e-04 | 8.68e-05 |
| 16 | 10656 | 5.53e-06 | 1.78e-03 | 1.98e-03 | 1.86e-04 | 5.89e-05 |

Table 5.15: Example 2: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part III: Average values of the local estimators.

| ℓ | N_{dof} | $\eta_{y_\varepsilon, T}$ | $\eta_{\hat{p}_\varepsilon, T}$ | $\eta_{y_\varepsilon, E}$ | $\eta_{\hat{p}_\varepsilon, E}$ | $\mu_\ell(u^d)$ | $\mu_\ell(\psi)$ | $osc_\ell(y^d)$ |
|--------|------------------|---------------------------|---------------------------------|---------------------------|---------------------------------|-----------------|------------------|-----------------|
| 0 | 5 | 3.65e-02 | 6.56e-02 | 8.61e-03 | 1.38e-02 | 8.67e-02 | 1.05e-01 | 6.80e-02 |
| 1 | 13 | 1.14e-02 | 1.52e-02 | 2.68e-03 | 3.12e-03 | 1.91e-02 | 2.66e-02 | 8.67e-03 |
| 2 | 41 | 1.18e-03 | 3.54e-03 | 2.53e-04 | 6.74e-04 | 3.39e-03 | 3.84e-03 | 1.08e-03 |
| 3 | 57 | 8.93e-04 | 2.56e-03 | 1.72e-04 | 4.37e-04 | 2.82e-03 | 2.16e-03 | 7.29e-04 |
| 4 | 73 | 3.87e-04 | 1.93e-03 | 6.32e-05 | 3.12e-04 | 1.75e-03 | 1.47e-03 | 4.88e-04 |
| 5 | 89 | 2.00e-04 | 1.62e-03 | 2.44e-05 | 2.45e-04 | 1.23e-03 | 1.11e-03 | 3.86e-04 |
| 6 | 121 | 8.87e-05 | 1.18e-03 | 8.56e-06 | 1.74e-04 | 7.40e-04 | 6.81e-04 | 2.49e-04 |
| 7 | 159 | 3.74e-05 | 8.52e-04 | 3.04e-06 | 1.03e-04 | 4.24e-04 | 4.29e-04 | 1.45e-04 |
| 8 | 243 | 1.61e-05 | 5.35e-04 | 9.89e-07 | 5.34e-05 | 2.16e-04 | 2.11e-04 | 7.00e-05 |
| 9 | 389 | 6.09e-06 | 3.40e-04 | 3.43e-07 | 2.82e-05 | 1.07e-04 | 1.08e-04 | 3.73e-05 |
| 10 | 603 | 2.34e-06 | 2.15e-04 | 1.03e-07 | 1.45e-05 | 5.29e-05 | 5.53e-05 | 1.85e-05 |
| 11 | 964 | 8.54e-07 | 1.30e-04 | 3.27e-08 | 6.74e-06 | 2.50e-05 | 2.60e-05 | 8.61e-06 |
| 12 | 1618 | 3.01e-07 | 7.81e-05 | 9.52e-09 | 3.37e-06 | 1.16e-05 | 1.29e-05 | 4.08e-06 |
| 13 | 2485 | 1.34e-07 | 5.05e-05 | 3.51e-09 | 1.75e-06 | 5.79e-06 | 6.78e-06 | 2.14e-06 |
| 14 | 3989 | 4.94e-08 | 3.09e-05 | 1.05e-09 | 8.30e-07 | 2.76e-06 | 3.19e-06 | 1.04e-06 |
| 15 | 6721 | 1.75e-08 | 1.83e-05 | 3.02e-10 | 3.95e-07 | 1.27e-06 | 1.50e-06 | 4.84e-07 |
| 16 | 10656 | 7.35e-09 | 1.16e-05 | 1.03e-10 | 1.98e-07 | 6.09e-07 | 7.89e-07 | 2.46e-07 |

Table 5.16: Example 2: Convergence history of the adaptive FEM for the Lavrentiev regularization with $\varepsilon = 10^{-6}$, Part IV: Percentages in the bulk criteria.

| ℓ | N_{dof} | $\mathcal{M}^{fb, T}$ | \mathcal{M}^E | $\mathcal{M}^{\eta, T}$ | $\mathcal{M}^{\mu, T}$ | $\mathcal{M}^{osc, T}$ |
|--------|------------------|-----------------------|-----------------|-------------------------|------------------------|------------------------|
| 0 | 5 | 100.0 | 75.0 | 75.0 | 75.0 | 75.0 |
| 1 | 13 | 25.0 | 30.0 | 25.0 | 31.2 | 25.0 |
| 2 | 41 | 6.3 | 28.4 | 9.4 | 6.2 | 12.5 |
| 3 | 57 | 8.7 | 19.5 | 19.6 | 7.6 | 18.5 |
| 4 | 73 | 6.5 | 11.9 | 25.8 | 5.6 | 17.7 |
| 5 | 89 | 5.3 | 8.3 | 30.9 | 4.6 | 15.8 |
| 6 | 121 | 3.9 | 7.5 | 30.0 | 3.4 | 16.4 |
| 7 | 159 | 2.8 | 21.3 | 33.2 | 2.5 | 21.9 |
| 8 | 243 | 1.8 | 11.3 | 31.7 | 1.6 | 21.4 |
| 9 | 389 | 1.1 | 8.3 | 32.6 | 0.8 | 17.5 |
| 10 | 603 | 0.7 | 15.3 | 37.9 | 0.5 | 17.2 |
| 11 | 964 | 0.4 | 10.1 | 40.9 | 0.3 | 17.6 |
| 12 | 1618 | 0.3 | 8.1 | 39.8 | 0.2 | 15.8 |
| 13 | 2485 | 0.3 | 7.4 | 43.6 | 0.1 | 10.9 |
| 14 | 3989 | 0.1 | 8.8 | 47.5 | 0.1 | 9.7 |
| 15 | 6721 | 0.1 | 5.1 | 45.1 | 0.0 | 9.6 |
| 16 | 10656 | 0.1 | 4.6 | 44.8 | 0.0 | 6.9 |

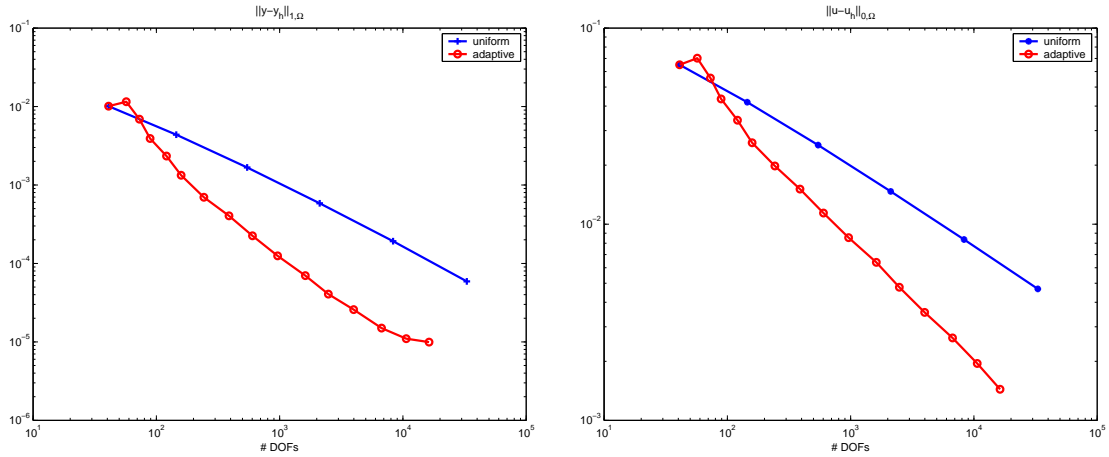


Figure 5.16: Example 2: Adaptive versus uniform refinement for the state (left) and the control (right) with regularization parameter $\varepsilon = 10^{-6}$.

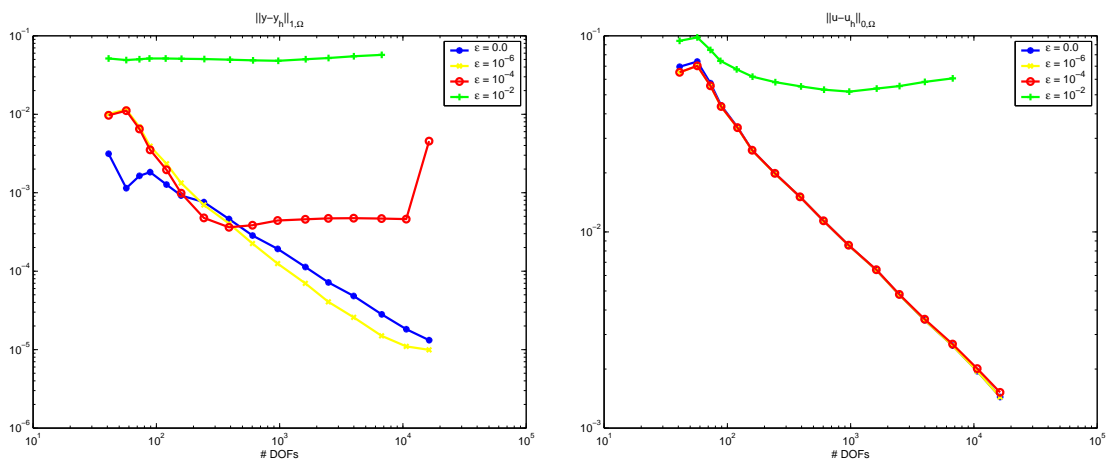


Figure 5.17: Example 2: Comparison of the approximation error in the state (left) and the control (right) for different regularization parameters.

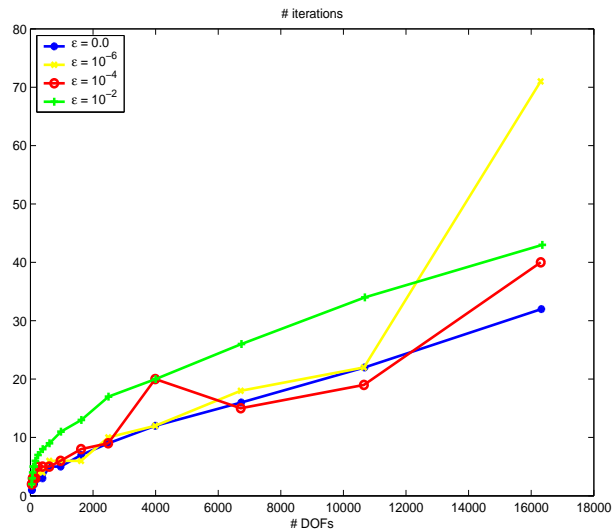


Figure 5.18: Example 2: Comparison of the number of iterations the active set strategy needs to find the exact solution for different regularization parameters and the unregularized case.

benefit compared to the unregularized case. In this example, even the iteration numbers for the regularization parameter $\varepsilon = 10^{-2}$ are worse. Remarkable is the behavior for $\varepsilon = 10^{-6}$ in the last step of the adaptive algorithm, where the number of iterations blows up to 71.

Chapter 6

An approach from the viewpoint of obstacle problems

In order to obtain an a posteriori error analysis, this thesis so far approached the state constrained optimal control problem with similar ideas as known from the control constrained case. This was made possible by the introduction of the modified adjoint state. Consequently, the a posteriori error estimates also involve the approximation error of this auxiliary function. On the other hand, the a posteriori error estimates known from the control constrained case also involve the approximation of the adjoint control and the adjoint state. Since these functions are also part of the optimality system, big values of the corresponding approximation errors probably also have a bad influence on the exactness of the discrete state and the discrete control, which are the quantities of real interest. Furthermore, one may assume that if the approximation error of the adjoint control is large, the discrete active zone and the continuous active zone are still apart from each other. This, especially, should be true if the supports of the continuous and the discrete adjoint control differ spatially. Conversely, the meaning of the discretization error of the modified adjoint state is not obvious, if there is one at all. Even though the adaptive algorithm worked well for our numerical test examples, there is even the danger that, at least for certain examples, the adaptive finite element method strongly refines in a region where the modified adjoint state is hard to approximate, but where a high density of nodes would not be necessary in order to gain a reasonable solution of the original optimal control problem.

As already mentioned in the introduction, the nature of state constrained optimal control problems is somehow similar to the one of obstacle problems. Thus, the last chapter of this thesis is devoted to an approach which takes advantage of this similarity and provides a different a posteriori error estimator. Consequently, this error estimator is similar to the one known from obstacle problems (cf. [Vee01]) and also implies a measure extension. Results concerning the reliability and the efficiency will be provided of which the first one also holds up to a consistency error. In addition to that, this chapter contains a numerical example for which this error

estimator provides a benefit compared to the uniform refinement and to the error estimator derived in the previous chapters.

Throughout the entire chapter, we deal with the same state constrained elliptic optimal control problem as given by (2.1.3) and also use the same discretization by linear finite elements as proposed in Section 2.4. In addition, we use the same notations as before.

6.1 Measure extension error estimator

In order to formulate the second residual-type a posteriori error estimator, we first have to expand the discrete Lagrange multiplier $\sigma_h \in \mathcal{M}_h$, which appears in the right-hand side of the adjoint state equation, to an element in $L^2(\Omega)$. Let $L_h : C(\Omega) \rightarrow S_h$ denote the Lagrange interpolation operator and set for two arbitrary functions $v_h, w_h \in S_h$

$$\langle v_h, w_h \rangle_h := \int_{\Omega} L_h(v_h w_h) = \sum_{n \in \mathcal{N}_h} v_h(n) w_h(n) \int_{\Omega} e_h^n, \quad (6.1.1)$$

where $e_h^n \in S_h$ denotes the nodal basis functions in the node n , i.e.,

$$e_h^n(\tilde{n}) := \begin{cases} 1, & \tilde{n} = n \\ 0, & \text{otherwise} \end{cases}, \quad \tilde{n} \in \mathcal{N}_h. \quad (6.1.2)$$

Now, we define the extension $\tilde{\sigma}_h \in S_h$ as follows:

$$\tilde{\sigma}_h(n) := \left(\int_{\Omega} e_h^n \right)^{-1} \sigma_h^n, \quad n \in \mathcal{N}_h, \quad (6.1.3)$$

where the values $\sigma_h^n \in \mathbb{R}$ are taken from the unique representation

$$\sigma_h = \sum_{n \in \mathcal{N}_h} \sigma_h^n \delta_n. \quad (6.1.4)$$

Notice that there holds

$$\langle \tilde{\sigma}_h, v_h \rangle_h = \langle \sigma_h, v_h \rangle \quad \forall v_h \in V_h. \quad (6.1.5)$$

For an easier notation in the forthcoming analysis, we furthermore define $\langle \tilde{\sigma}_h, \cdot \rangle : L^2(\Omega) \rightarrow \mathbb{R}$ according to

$$\langle \tilde{\sigma}_h, v \rangle := (\tilde{\sigma}_h, v)_{0,\Omega} \quad \forall v \in L^2(\Omega). \quad (6.1.6)$$

The main difference of the error estimator of this chapter compared to the error estimators of the previous chapters is that we now want to use error terms which

are closer related to the adjoint state and do not take the modified adjoint state into account. Since the right-hand side of p involves a measure which it is hard to deal with, we use the just stated measure extension. In particular, let us define

$$\tilde{\eta}_y := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \tilde{\eta}_{y,T}^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \tilde{\eta}_{y,E}^2 \right)^{\frac{1}{2}}, \quad (6.1.7a)$$

$$\tilde{\eta}_p := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \tilde{\eta}_{p,T}^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \tilde{\eta}_{p,E}^2 \right)^{\frac{1}{2}}. \quad (6.1.7b)$$

Here, the element residuals $\tilde{\eta}_{y,T}$ and $\tilde{\eta}_{p,T}$ and the edge residuals $\tilde{\eta}_{y,E}$ and $\tilde{\eta}_{p,E}$ are again given by weighted L^2 -residuals. This time, they read as follows:

$$\tilde{\eta}_{y,T} := h_T \|u_h - cy_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \quad (6.1.8a)$$

$$\tilde{\eta}_{p,T} := h_T \|y_h - y^d + \tilde{\sigma}_h - cp_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \quad (6.1.8b)$$

$$\tilde{\eta}_{y,E} := h_E^{\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \quad (6.1.8c)$$

$$\tilde{\eta}_{p,E} := h_E^{\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla p_h]\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \quad (6.1.8d)$$

where again $E = T_1 \cap T_2$ with $T_1, T_2 \in \mathcal{T}_h(\Omega)$, and $\boldsymbol{\nu}_E$ is the exterior unit normal vector on E directed towards T_2 , whereas $[\nabla y_h]$ and $[\nabla p_h]$ denote the jumps of ∇y_h and ∇p_h across E .

Again, the total residual-type error estimator $\tilde{\eta}$ for the finite element approximation of the distributed optimal control problem (2.1.3) is then given by the summation of the y -component and the p -component, i.e.,

$$\tilde{\eta} := \left(\tilde{\eta}_y^2 + \tilde{\eta}_p^2 \right)^{\frac{1}{2}}. \quad (6.1.9)$$

Moreover, the lower order data oscillation in u^d is given by the same formula as before, according to

$$\mu_h(u^d) := \left(\sum_{T \in \mathcal{T}_h(\Omega)} \mu_T(u^d)^2 \right)^{1/2}, \quad (6.1.10)$$

$$\mu_T(u^d) := \|u^d - u_h^d\|_{0,T}.$$

In addition to that, we now have, beside the higher order data oscillation in y^d , also data oscillation in y_h and in p_h . Therefore, we fix

$$osc_h(y^d) := \left(\sum_{T \in \mathcal{T}_h(\Omega)} osc_T(y^d)^2 \right)^{1/2}, \quad (6.1.11a)$$

$$osc_T(y^d) := h_T \|y^d - y_h^d\|_{0,T},$$

$$osc_h(p_h) := \left(\sum_{T \in \mathcal{T}_h(\Omega)} osc_T(p_h)^2 \right)^{1/2}, \quad (6.1.11b)$$

$$\begin{aligned}
osc_T(p_h) &:= h_T \|y_h - M_h y_h\|_{0,T} , \\
osc_h(p_h) &:= \left(\sum_{T \in \mathcal{T}_h(\Omega)} osc_T(p_h)^2 \right)^{1/2} , \\
osc_T(p_h) &:= h_T \|p_h - M_h p_h\|_{0,T} .
\end{aligned} \tag{6.1.11c}$$

Remember that $M_h : L^2(\Omega) \rightarrow W_h$ is the operator defined by

$$(M_h v)|_T := |T|^{-1} \int_T v(x) dx , \quad T \in \mathcal{T}_h , \quad v \in L^2(\Omega) \tag{6.1.12}$$

and

$$y_h^d := M_h y^d . \tag{6.1.13}$$

Remark 6.1.1. *The statements of Remark 3.1.2 concerning the data oscillation in u^d and y^d also apply here. $osc_h(p_h)$ and $osc_h(y_h)$ are always of higher order, since $y_h, p_h \in H^1(\Omega)$. However, the functions change in every iteration.*

As we did before, we define the auxiliary state $y(u_h) \in V$ as the unique solution of

$$(\nabla y(u_h), \nabla v)_{0,\Omega} + c(y(u_h), v)_{0,\Omega} = (u_h, v)_{0,\Omega} \quad \forall v \in V . \tag{6.1.14}$$

The auxiliary adjoint state $p(u_h, \tilde{\sigma}_h) \in V$ involves the measure extension and is defined as the unique solution of

$$(\nabla p(u_h, \tilde{\sigma}_h), \nabla v)_{0,\Omega} + c(p(u_h, \tilde{\sigma}_h), v)_{0,\Omega} = (\tilde{\sigma}_h - y^d + y(u_h), v)_{0,\Omega} \quad \forall v \in V . \tag{6.1.15}$$

Notice that the auxiliary adjoint state is smoother than the adjoint state, since we used the measure extension $\tilde{\sigma}_h$, which is a L^2 -function, for its right-hand side. If one uses the discrete measure $\sigma_h \in \mathcal{M}_h$ in the right-hand side, the result would be a pretty complicated function because it would only be part of the spaces $W^{1,q}(\Omega)$ with $1 < q < 2$, and it would have singularities in every nodal point of the triangulation. Now, $p(u_h, \tilde{\sigma}_h)$ has the same regularity as the modified adjoint state used in the previous chapters, but, due to the appearance of the measure extension in its right-hand side, it should be closer to the adjoint state p .

Since the auxiliary state $y(u_h)$ fulfills neither the continuous nor the discrete complementarity conditions, we define the consistency error as follows:

$$e_c(u_h) := \max \left(\langle \sigma, y(u_h) - \psi \rangle + \langle \tilde{\sigma}_h, \psi - y(u_h) \rangle, 0 \right) \tag{6.1.16}$$

For this consistency error, similar arguments as given in Remark 3.1.2 apply:

Remark 6.1.2. *Since the support of σ is concentrated on the continuous active zone \mathcal{A} and σ is a positive measure, i.e., $\sigma \in \mathcal{M}_+(\Omega)$, the term $\langle \sigma, y(u_h) - \psi \rangle$ penalizes if $y(u_h)$ violates the state constraints in the continuous active zone. Thus, the value of this term should be close to 0 if $y(u_h)$ is close to y , which fulfills the upper bound. It is even negative if $y(u_h) < \psi$ holds in the active zone. Furthermore, there should hold $\langle \tilde{\sigma}_h, \psi - y(u_h) \rangle \approx \langle \sigma_h, \psi - y(u_h) \rangle$ since $\tilde{\sigma}_h$ is defined as an extension of the measure σ_h , and thus, the second term should be close to 0 if $y(u_h) \approx y_h$ due to the complementarity condition (2.4.8c). However, these are just some heuristic arguments.*

The error estimator will be discussed in more detail in a forthcoming chapter.

6.2 Reliability of the measure extension error estimator

As we did for our first error estimator, we also want to establish reliability. In this section, we will show that the above declared error estimator, derived from a measure extension, is an upper bound for the discretization error in the state, in the control, and in the auxiliary adjoint state. Similarly, this will also provide an upper bound for the approximation of the adjoint state in the L^2 -norm. The theoretical disadvantage compared to error estimator η is that the latter of these functions is not completely continuous but somehow semidiscrete. However, it still might provide more information in some cases than the modified adjoint state does. As for the error estimator η , the result will be provided up to data oscillation and consistency error. Let us first state the main result of this chapter.

Theorem 6.2.1 (Reliability of the measure extension error estimator). *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $p(u_h, \tilde{\sigma}_h)$ be the auxiliary adjoint state as defined in (6.1.15). Furthermore, let $\tilde{\eta}$, $\mu_h(u^d)$, $\text{osc}_h(y^d)$, $\text{osc}_h(y_h)$, and $\text{osc}_h(p_h)$ be the residual-type error estimator and the data oscillation as given by (6.1.9), (6.1.10), (6.1.11a), (6.1.11b), and (6.1.11c), respectively, and denote by $e_c(u_h)$ the consistency error (6.1.16). Then, there exists a constant C , depending only on α , \tilde{c} , and the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that:*

$$\begin{aligned} & \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{0,\Omega}^2 + \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega}^2 \quad (6.2.1) \\ & \leq C(\tilde{\eta}^2 + \mu_h^2(u^d) + \text{osc}_h^2(y^d) + \text{osc}_h^2(y_h) + c^2 \text{osc}_h^2(p_h) + e_c(u_h)) \end{aligned}$$

Notice that the function $p(u_h, \tilde{\sigma}_h)$ changes in each iteration and that in Theorem 6.2.1, the data oscillation in p_h only occurs if the parameter c does not vanish, as it is the case if one establishes discrete local efficiency for general elliptic partial differential equations (cf. [MN07]). Local efficiency, on the other side, may be proved

without this term. The data oscillation in y_h appears independently of c because it is part of the right-hand side of the adjoint state equation.

As a first step to show Theorem 6.2.1, we are going to prove the following auxiliary result.

Lemma 6.2.2. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $y(u_h)$ and $p(u_h, \tilde{\sigma}_h)$ be the auxiliary state and auxiliary adjoint state as given by (6.1.14) and (6.1.15), respectively. In addition to that, let $\mu_h(u^d)$ be the data oscillation as given by (6.1.10) and $e_c(u_h)$ be the consistency error according to (6.1.16). Then, there exists a constant C , depending only on α and Ω , such that:*

$$\begin{aligned} & \|y - y_h\|_{1,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \\ & \leq C(\|y(u_h) - y_h\|_{1,\Omega}^2 + \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega}^2 + \mu_h^2(u^d) + e_c(u_h)) \end{aligned} \quad (6.2.2)$$

Proof. With the help of the triangle inequality one may find

$$\|y - y_h\|_{1,\Omega}^2 \leq 2\|y - y(u_h)\|_{1,\Omega}^2 + 2\|y(u_h) - y_h\|_{1,\Omega}^2 \quad (6.2.3)$$

The first term of the right-hand side of (6.2.3) can be estimated by applying (3.1.12), (2.3.1a), and (6.1.14), as well as Cauchy's inequality according to

$$\begin{aligned} & \|y - y(u_h)\|_{1,\Omega}^2 \\ & \leq \tilde{c}[(\nabla(y - y(u_h)), \nabla(y - y(u_h)))_{0,\Omega} + c(y - y(u_h), y - y(u_h))_{0,\Omega}] \\ & = \tilde{c}(u - u_h, y - y(u_h))_{0,\Omega} \\ & \leq \tilde{c}\|u - u_h\|_{0,\Omega}\|y - y(u_h)\|_{0,\Omega} \\ & \leq \tilde{c}\|u - u_h\|_{0,\Omega}\|y - y(u_h)\|_{1,\Omega}. \end{aligned} \quad (6.2.4)$$

From (6.2.3) and (6.2.4), one may conclude

$$\|y - y_h\|_{1,\Omega}^2 \leq 2\tilde{c}^2\|u - u_h\|_{0,\Omega}^2 + 2\|y_h - y(u_h)\|_{1,\Omega}^2. \quad (6.2.5)$$

Now, one may apply the relationships (2.3.1c) and (2.4.5c) to the term $\|u - u_h\|_{0,\Omega}$, which results in

$$\begin{aligned} & \|u - u_h\|_{0,\Omega}^2 \\ & = (u^d - \frac{1}{\alpha}p + \frac{1}{\alpha}p_h - u_h^d, u - u_h)_{0,\Omega} \\ & = (u^d - u_h^d, u - u_h)_{0,\Omega} + \frac{1}{\alpha}(p_h - p(u_h, \tilde{\sigma}_h), u - u_h)_{0,\Omega} \\ & \quad + \frac{1}{\alpha}(p(u_h, \tilde{\sigma}_h) - p, u - u_h)_{0,\Omega}. \end{aligned} \quad (6.2.6)$$

By applying the variational equalities (2.3.1a), (6.1.14), (2.3.1b), and (6.1.15), one obtains for the last term of (6.2.6)

$$\begin{aligned}
& (p(u_h, \tilde{\sigma}_h) - p, u - u_h)_{0,\Omega} & (6.2.7) \\
&= (\nabla(p(u_h, \tilde{\sigma}_h) - p), \nabla(y - y(u_h)))_{0,\Omega} + c(p(u_h, \tilde{\sigma}_h) - p, y - y(u_h))_{0,\Omega} \\
&= (\tilde{\sigma}_h - y^d + y(u_h), y - y(u_h))_{0,\Omega} - \langle \sigma, y - y(u_h) \rangle + (y^d - y, y - y(u_h)) \\
&= \langle \tilde{\sigma}_h - \sigma, y - y(u_h) \rangle - \|y - y(u_h)\|^2 \\
&\leq \langle \tilde{\sigma}_h, y - y(u_h) \rangle + \langle \sigma, y(u_h) - y \rangle .
\end{aligned}$$

For the second term of the right-hand side of (6.2.7), one gets by applying the complementarity condition (2.3.14c)

$$\begin{aligned}
\langle \sigma, y(u_h) - y \rangle &= \langle \sigma, y(u_h) - \psi \rangle + \underbrace{\langle \sigma, \psi - y \rangle}_{=0} & (6.2.8) \\
&\leq \langle \sigma, y(u_h) - \psi \rangle
\end{aligned}$$

One can easily derive that, due to the definition of $\tilde{\sigma}_h$ in (6.1.3) and the positiveness of σ_h (compare (2.4.8a)), there also holds $\tilde{\sigma}_h \geq 0$. Since furthermore the optimal state y has to be feasible for the continuous problem, one may conclude

$$\begin{aligned}
\langle \tilde{\sigma}_h, y - y(u_h) \rangle &= \underbrace{\langle \tilde{\sigma}_h, y - \psi \rangle}_{\geq 0} + \underbrace{\langle \tilde{\sigma}_h, \psi - y(u_h) \rangle}_{\leq 0} & (6.2.9) \\
&\leq \langle \tilde{\sigma}_h, \psi - y(u_h) \rangle .
\end{aligned}$$

Now, the combination of the right-hand sides of (6.2.8) and (6.2.9) results in the consistency error term $e_c(u_h)$. If one now applies Cauchy's and Young's inequality to the first and second term of (6.2.6) and combines the estimates (6.2.6) -(6.2.9), one arrives at

$$\begin{aligned}
& \|u - u_h\|_{0,\Omega}^2 & (6.2.10) \\
&\leq \frac{1}{2} \|u - u_h\|_{0,\Omega}^2 + \|u^d - u_h^d\|_{0,\Omega}^2 + \frac{1}{\alpha^2} \|p_h - p(u_h, \tilde{\sigma}_h)\|_{0,\Omega}^2 + \frac{1}{\alpha} e_c(u_h) ,
\end{aligned}$$

which together with (6.2.5) gives the claimed result. \square

In order to prove Theorem 6.2.1, it remains to estimate the two terms $\|y(u_h) - y_h\|_{1,\Omega}$ and $\|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega}$. That will be done by the two following claims of Lemma 6.2.3 and Lemma 6.2.4.

Before we can prove them, we need some additional tools. Thus, we will refer in the following to $I_h : L^1(\Omega) \rightarrow V_h$ as the Clément interpolation operator, as defined in [Ver96], for example. This operator fulfills the following estimates:

$$\|\varphi - I_h \varphi\|_{k,2,T} \leq c_8 h_T^{l-k} \|\varphi\|_{l,2,\tilde{\omega}_T} , \quad 0 \leq k \leq l \leq 2 , \quad \varphi \in H^l(\tilde{\omega}_T) , \quad (6.2.11a)$$

$$\|\varphi - I_h \varphi\|_{0,2,E} \leq c_9 h_E^{l-\frac{1}{2}} \|\varphi\|_{l,2,\tilde{\omega}_E} , \quad 1 \leq l \leq 2 , \quad \varphi \in H^l(\tilde{\omega}_E) , \quad (6.2.11b)$$

where c_8 and c_9 are constants which only depend on the shape regularity of the triangulation, and $\tilde{\omega}_T$ and $\tilde{\omega}_E$ are defined as follows:

$$\tilde{\omega}_T := \bigcup_{\mathcal{N}(T) \cap \mathcal{N}(T') \neq \emptyset} T' , \quad \tilde{\omega}_E := \bigcup_{\mathcal{N}(E) \cap \mathcal{N}(T') \neq \emptyset} T' . \quad (6.2.12)$$

Lemma 6.2.3. *Let $(y_h, p_h, u_h, \sigma_h)$ be the solution of (2.4.5) and let $y(u_h)$ be the corresponding auxiliary state as defined in (6.1.14). Furthermore, let $\tilde{\eta}_y$ be as defined in (6.1.7a). Then, there exists a constant C , depending only on \tilde{c} and the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that*

$$\|y_h - y(u_h)\|_{1,\Omega} \leq C \tilde{\eta}_y . \quad (6.2.13)$$

Proof. Due to Galerkin orthogonality, the claim is standard in a posteriori error analysis (compare for example [Ver96]) and follows from the same arguments as used in the proof of Lemma 6.2.4. \square

Lemma 6.2.4. *Let $(y_h, p_h, u_h, \sigma_h)$ be the solution of (2.4.5) and let $p(u_h, \tilde{\sigma}_h)$ be the corresponding auxiliary adjoint state as defined in (6.1.15). Furthermore, let $\tilde{\eta}_y$ and $\tilde{\eta}_p$ be as defined in (6.1.7). Then, there exists a constant C , depending only on \tilde{c} and the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$, such that*

$$\|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega} \leq C (\tilde{\eta}_y + \tilde{\eta}_p + \text{osc}_h(y^d) + \text{osc}_h(y_h) + c \text{osc}_h(p_h)) . \quad (6.2.14)$$

Proof. Due to different right-hand sides of the auxiliary adjoint state and the discrete adjoint state, the proof will be presented here. A lot of the arguments used in the following are similar to those known from obstacle problems (cf. [Vee01]).

For a more practical notation, let us first define the linear operator $B : V \rightarrow V^*$ by

$$(Bw)(v) := (\nabla w, \nabla v)_{0,\Omega} + c(w, v)_{0,\Omega} \quad \forall w, v \in V , \quad (6.2.15)$$

the functional $\mathcal{R}_h \in V^*$ by

$$\langle \mathcal{R}_h, v \rangle := (Bp_h)(v) - (\tilde{\sigma}_h - y^d + y_h, v)_{0,\Omega} \quad \forall v \in V , \quad (6.2.16)$$

and

$$K_2 := \{\varphi \in V : \|\varphi\|_{1,\Omega} = 1\} . \quad (6.2.17)$$

With these definitions at hand and applying (3.1.12), (6.1.15), and Cauchy's in-

equality, one may obtain

$$\begin{aligned}
& \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega}^2 & (6.2.18) \\
\leq & \tilde{c}((\nabla(p(u_h, \tilde{\sigma}_h) - p_h), \nabla(p(u_h, \tilde{\sigma}_h) - p_h))_{0,\Omega} \\
& + c(p(u_h, \tilde{\sigma}_h) - p_h, p(u_h, \tilde{\sigma}_h) - p_h)_{0,\Omega}) \\
= & \tilde{c}((\tilde{\sigma}_h - y^d + y(u_h), p(u_h, \tilde{\sigma}_h) - p_h)_{0,\Omega} - (Bp_h)(p(u_h, \tilde{\sigma}_h) - p_h)) \\
= & \tilde{c}((\tilde{\sigma}_h - y^d + y_h, p(u_h, \tilde{\sigma}_h) - p_h)_{0,\Omega} + (y(u_h) - y_h, p(u_h, \tilde{\sigma}_h) - p_h)_{0,\Omega} \\
& - (Bp_h)(p(u_h, \tilde{\sigma}_h) - p_h)) \\
= & \tilde{c}(\langle \mathcal{R}_h, p_h - p(u_h, \tilde{\sigma}_h) \rangle + (y(u_h) - y_h, p(u_h, \tilde{\sigma}_h) - p_h)_{0,\Omega}) \\
\leq & \tilde{c}(\|\mathcal{R}_h\|_{-1,\Omega} \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega} + \|y(u_h) - y_h\|_{1,\Omega} \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega})
\end{aligned}$$

Combined with Lemma 6.2.3, this results in

$$\|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega} \leq \|\mathcal{R}_h\|_{-1,\Omega} + \tilde{\eta}_y . \quad (6.2.19)$$

Thus, it remains to estimate the term $\|\mathcal{R}_H\|_{-1,\Omega}$. Let now $v_h := I_h v \in V_h$ for $v \in V$. Then, we may use (2.4.5b), (6.2.15), (6.1.5), and (6.1.6) in order to gain

$$\begin{aligned}
& \|\mathcal{R}_h\|_{-1,\Omega} & (6.2.20) \\
= & \sup_{v \in K_2} \{(Bp_h)(v) - (\tilde{\sigma}_h - y^d + y_h, v)_{0,\Omega}\} \\
= & \sup_{v \in K_2} \{(Bp_h)(v) - (\tilde{\sigma}_h - y^d + y_h, v)_{0,\Omega} - (Bp_h)(v_h) + \langle \tilde{\sigma}_h, v_h \rangle_h - (y^d - y_h, v_h)\} \\
= & \sup_{v \in K_2} \{(Bp_h)(v - v_h) - (\tilde{\sigma}_h - y^d + y_h, v - v_h)_{0,\Omega} + \langle \tilde{\sigma}_h, v_h \rangle_h - \langle \tilde{\sigma}_h, v_h \rangle\} \\
\leq & \sup_{v \in K_2} \{\langle \tilde{\sigma}_h, v_h \rangle_h - \langle \tilde{\sigma}_h, v_h \rangle\} + \sup_{v \in K_2} \{(\nabla p_h, \nabla(v - v_h))_{0,\Omega} \\
& + c(p_h, v - v_h)_{0,\Omega} - (\tilde{\sigma}_h - y^d + y_h, v - v_h)_{0,\Omega}\} .
\end{aligned}$$

For further estimates, we use Theorem 28.1 from [Cia91], which tells us that there holds

$$\int_T (L_h(\tilde{\sigma}_h v_h) - \tilde{\sigma}_h v_h) dx \leq Ch_T^2 \|D^2[\tilde{\sigma}_h v_h]\|_{0,1,T} , \quad (6.2.21)$$

where D denotes the Fréchet derivative, and C is a constant which is independent of the local mesh size. With the help of (6.1.1) and (6.2.21) and an application of Hölder's inequality for sums and integrals, we may get for the first term of the

right-hand side of (6.2.20)

$$\begin{aligned}
& \sup_{v \in K_2} \{ \langle \tilde{\sigma}_h, v_h \rangle_h - \langle \tilde{\sigma}_h, v_h \rangle \} \tag{6.2.22} \\
&= \sup_{v \in K_2} \left\{ \sum_{T \in \mathcal{T}_h} \int_T (L_h(\tilde{\sigma}_h v_h) - \tilde{\sigma}_h v_h) dx \right\} \\
&\leq C \sup_{v \in K_2} \left\{ \sum_{T \in \mathcal{T}_h} h_T^2 \|D^2[\tilde{\sigma}_h v_h]\|_{0,1,T} \right\} \\
&\leq C \sup_{v \in K_2} \left\{ \sum_{T \in \mathcal{T}_h} h_T^2 \|\nabla \tilde{\sigma}_h\|_{0,T} \|\nabla v_h\|_{0,T} \right\} \\
&\leq C \sup_{v \in K_2} \left\{ \left(\sum_{T \in \mathcal{T}_h} h_T^4 \|\nabla \tilde{\sigma}_h\|_{0,T}^2 \right)^{\frac{1}{2}} \underbrace{\left(\sum_{T \in \mathcal{T}_h} \|\nabla v_h\|_{0,T}^2 \right)^{\frac{1}{2}}}_{\leq 1} \right\} \\
&\leq C \left(\sum_{T \in \mathcal{T}_h} h_T^4 \|\nabla \tilde{\sigma}_h\|_{0,T}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Since $M_h p_h$, y_h^d , and $M_h y_h$ are constant on each triangle T , one may obtain for the right-hand side of (6.2.22) with the help of the inverse argument of Theorem 6.8 of [BR03], which holds on the discrete space:

$$\begin{aligned}
& h_T^4 \|\nabla \tilde{\sigma}_h\|_{0,T}^2 \tag{6.2.23} \\
&= h_T^4 \|\nabla(\tilde{\sigma}_h - y_h^d + M_h y_h - c M_h p_h)\|_{0,T}^2 \\
&\leq C h_T^2 \|\tilde{\sigma}_h - y_h^d + M_h y_h - c M_h p_h\|_{0,T}^2 \\
&\leq C h_T^2 (\|\tilde{\sigma}_h - y^d + y_h - c p_h\|_{0,T}^2 + c^2 \|p_h - M_h p_h\|_{0,T}^2 \\
&\quad + \|y^d - y_h^d\|_{0,T}^2 + \|y_h - M_h y_h\|_{0,T}^2),
\end{aligned}$$

where the right-hand side only contains parts of the error estimator and data oscillations, and C represents a generic constant which is independent of the local mesh size.

For the second term of the right-hand side of (6.2.20), one may now apply Green's formula, consider $\Delta p_h|_T = 0$, and take advantage of the facts that v and v_h vanish on each Dirichlet boundary edge. In addition to that, $\nabla p_h \cdot \nu_E$ has value 0 on all other boundary edges due to the natural Neumann conditions. Therefore, one may conclude:

$$\begin{aligned}
& \sup_{v \in K_2} \{ (\nabla p_h, \nabla(v - v_h))_{0,\Omega} + c(p_h, v - v_h)_{0,\Omega} - (\tilde{\sigma}_h - y^d + y_h, v - v_h)_{0,\Omega} \} \tag{6.2.24} \\
&= \sup_{v \in K_2} \sum_{T \in \mathcal{T}_h(\Omega)} \left\{ \underbrace{(-\Delta p_h, v - v_h)_{0,T}}_{=0} + (\nabla p_h \cdot \nu_{\partial T}, v - v_h)_{0,\partial T} \right. \\
&\quad \left. + (c p_h - y_h + y^d - \tilde{\sigma}_h, v - v_h)_{0,T} \right\} \\
&= \sup_{v \in K_2} \left\{ \sum_{E \in \mathcal{E}_h(\Omega)} ([\nabla p_h] \cdot \nu_E, v - v_h)_{0,E} + \sum_{T \in \mathcal{T}_h(\Omega)} (c p_h - y_h + y^d - \tilde{\sigma}_h, v - v_h)_{0,T} \right\}
\end{aligned}$$

Finally, we apply Cauchy's inequality, (6.2.11) with $k = 0$ and $l = 1$, and Hölder's inequality for sums:

$$\begin{aligned}
& \sup_{v \in K_2} \{ (\nabla p_h, \nabla(v - v_h))_{0,\Omega} + c(p_h, v - v_h)_{0,\Omega} - (\tilde{\sigma}_h - y^d + y_h, v - v_h)_{0,\Omega} \} \quad (6.2.25) \\
& \leq \sup_{v \in K_2} \left\{ \sum_{E \in \mathcal{E}_h(\Omega)} \|[\nabla p_h] \cdot \boldsymbol{\nu}_E\|_{0,E} \|v - v_h\|_{0,E} \right. \\
& \quad \left. + \sum_{T \in \mathcal{T}_h(\Omega)} \|cp_h - y_h + y^d - \tilde{\sigma}_h\|_{0,T} \|v - v_h\|_{0,T} \right\} \\
& \leq \sup_{v \in K_2} \left\{ \sum_{E \in \mathcal{E}_h(\Omega)} c_9 h_E^{\frac{1}{2}} \|[\nabla p_h] \cdot \boldsymbol{\nu}_E\|_{0,E} \|v\|_{1,\tilde{\omega}_E} \right. \\
& \quad \left. + \sum_{T \in \mathcal{T}_h(\Omega)} c_8 h_T \|cp_h - y_h + y^d - \tilde{\sigma}_h\|_{0,T} \|v\|_{1,\tilde{\omega}_T} \right\} \\
& \leq \sup_{v \in K_2} \left\{ c_9 \left(\sum_{E \in \mathcal{E}_h(\Omega)} h_E \|[\nabla p_h] \cdot \boldsymbol{\nu}_E\|_{0,E}^2 \right)^{\frac{1}{2}} \left(\sum_{E \in \mathcal{E}_h(\Omega)} \|v\|_{1,\tilde{\omega}_E}^2 \right)^{\frac{1}{2}} \right. \\
& \quad \left. + c_8 \left(\sum_{T \in \mathcal{T}_h(\Omega)} h_T^2 \|cp_h - y_h + y^d - \tilde{\sigma}_h\|_{0,T}^2 \right)^{\frac{1}{2}} \left(\sum_{T \in \mathcal{T}_h(\Omega)} \|v\|_{1,\tilde{\omega}_T}^2 \right)^{\frac{1}{2}} \right\} \\
& \preceq \left(\sum_{E \in \mathcal{E}_h(\Omega)} h_E \|[\nabla p_h] \cdot \boldsymbol{\nu}_E\|_{0,E}^2 \right)^{\frac{1}{2}} + \left(\sum_{T \in \mathcal{T}_h(\Omega)} h_T^2 \|cp_h - y_h + y^d - \tilde{\sigma}_h\|_{0,T}^2 \right)^{\frac{1}{2}},
\end{aligned}$$

where for the last step, we used the fact that the summation of $\tilde{\omega}_E$ over all edges and of $\tilde{\omega}_T$ over all triangles only gives a finite overlap. Now, the combination of (6.2.19), (6.2.20), (6.2.22), (6.2.23), and (6.2.25) allows to conclude. \square

Proof of Theorem 6.2.1.

The claim is a direct consequence of Lemma 6.2.2, Lemma 6.2.3, Lemma 6.2.4, and the same estimate for the term $\|p - p_h\|_{0,\Omega}$ as used in the proof of Theorem 3.2.1. \square

6.3 Efficiency of the measure extension error estimator

This section is devoted to the proof that also the measure extension error estimator $\tilde{\eta}$ is efficient. The main result is the following:

Theorem 6.3.1 (Efficiency of the measure extension error estimator). *Assume that (y, u, p, σ) and $(y_h, u_h, p_h, \sigma_h)$ are the optimal solutions of (2.3.1) and (2.4.5), respectively. In addition to that, let $p(u_h, \tilde{\sigma}_h)$ be the auxiliary adjoint state, as defined in (6.1.15), and let $\tilde{\eta}$ and $\text{osc}_h(y^d)$ be the error estimator and data oscillation.*

lation as given by (6.1.9) and (6.1.11a), respectively. Then, there holds

$$\tilde{\eta} \preceq \|y - y_h\|_{1,\Omega} + \|u - u_h\|_{1,\Omega} + \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega} + \text{osc}_h(y^d), \quad (6.3.1a)$$

$$\begin{aligned} \tilde{\eta} \preceq & \|y - y_h\|_{1,\Omega} + \|u - u_h\|_{1,\Omega} + \|p - p_h\|_{0,\Omega} \\ & + \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\Omega} + \text{osc}_h(y^d). \end{aligned} \quad (6.3.1b)$$

In this section, we use the same notations for the bubble functions and the corresponding inequalities as previously established (compare (3.3.1) and (3.3.2)). Again, the proof of Theorem 6.3.1 will be given by the following series of lemmas which will provide local efficiency for the parts of the error estimator $\tilde{\eta}$.

Lemma 6.3.2. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $\tilde{\eta}_{y,T}$ be given by (6.1.8a). Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for each $T \in \mathcal{T}_h(\Omega)$, there holds*

$$\tilde{\eta}_{y,T}^2 \leq C \left(\|y - y_h\|_{1,T}^2 + h_T^2 \|u - u_h\|_{0,T}^2 \right). \quad (6.3.2)$$

Proof. Since the error estimator parts $\eta_{y,T}$ and $\tilde{\eta}_{y,T}$ are the same, the proof is equal to the proof of Lemma 3.3.1. \square

Lemma 6.3.3. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, let $p(u_h, \tilde{\sigma}_h)$ be as defined in (6.1.15), and let $\tilde{\eta}_{p,T}$ and $\text{osc}_h(y^d)$ be given by (6.1.8b) and (6.1.11a), respectively. Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for each $T \in \mathcal{T}_h(\Omega)$, there holds*

$$\begin{aligned} \tilde{\eta}_{p,T}^2 \leq & C \left(\|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,T}^2 + h_T^2 \|y - y_h\|_{0,T}^2 \right. \\ & \left. + h_T^2 \|y - y(u_h)\|_{0,T}^2 + \text{osc}_h^2(y^d) \right). \end{aligned} \quad (6.3.3)$$

Proof. Since in contrast to the error estimator terms $\eta_{\tilde{p},T}$, the term $\tilde{\eta}_{p,T}$ also includes the measure extension and gives a result according to the approximation of the auxiliary adjoint state, we will declare the proof here.

By using the triangle inequality, one gains

$$\begin{aligned} \tilde{\eta}_{p,T}^2 &= h_T^2 \|y_h - y^d + \tilde{\sigma}_h - cp_h\|_{0,T}^2 \\ &\leq 2h_T^2 \|y^d - y_h^d\|_{0,T}^2 + 2h_T^2 \|cp_h + y_h^d - y_h - \tilde{\sigma}_h\|_{0,T}^2 \end{aligned} \quad (6.3.4)$$

The second term of the right-hand side of (6.3.4) can now be estimated by setting $z_h := (cp_h + y_h^d - y_h - \tilde{\sigma}_h)|_T \vartheta_T$, using (3.3.1a), (6.1.15), and the fact that $\Delta p_h|_T = 0$,

and an integration by parts as follows:

$$\begin{aligned}
& h_T^2 \|cp_h + y_h^d - y_h - \tilde{\sigma}_h\|_{0,T}^2 \tag{6.3.5} \\
& \leq c_1 h_T^2 (cp_h + y_h^d - y_h - \tilde{\sigma}_h, z_h)_{0,T} \\
& = c_1 h_T^2 (cp_h + y_h^d - y_h - \tilde{\sigma}_h - \Delta p_h, z_h)_{0,T} \\
& = c_1 h_T^2 ((\nabla p_h, \nabla z_h)_{0,T} - (\nabla p_h \cdot \boldsymbol{\nu}_{\partial T}, \underbrace{z_h}_{=0})_{0,\partial T} - (\nabla p(u_h, \tilde{\sigma}_h), \nabla z_h)_{0,T} \\
& \quad + c(p_h - p(u_h, \tilde{\sigma}_h), z_h)_{0,T} - (y^d - y(u_h), z_h)_{0,T} + (y_h^d - y_h, z_h)_{0,T}) \\
& = c_1 h_T^2 ((\nabla(p_h - p(u_h, \tilde{\sigma}_h)), \nabla z_h)_{0,T} + c(p_h - p(u_h, \tilde{\sigma}_h), z_h)_{0,T} \\
& \quad + (y_h^d - y^d, z_h)_{0,T} + (y(u_h) - y_h, z_h)_{0,T})
\end{aligned}$$

Now, we may obtain with the help of Cauchy's and Young's inequality, (3.3.1b), and (3.3.1c)

$$\begin{aligned}
& h_T^2 \|cp_h + y_h^d - y_h - \tilde{\sigma}_h\|_{0,T}^2 \tag{6.3.6} \\
& \leq c_1 h_T^2 (|p_h - p(u_h, \tilde{\sigma}_h)|_{1,T} \underbrace{\|z_h\|_{1,T}}_{\leq c_3 h_T^{-1} \|cp_h + y_h^d - y_h - \tilde{\sigma}_h\|_{0,T}} \\
& \quad + c \|p_h - p(u_h, \tilde{\sigma}_h)\|_{0,T} \underbrace{\|z_h\|_{0,T}}_{\leq c_2 \|cp_h + y_h^d - y_h - \tilde{\sigma}_h\|_{0,T}} \\
& \quad + \|y_h^d - y^d\|_{0,T} \|z_h\|_{0,T} + \|y(u_h) - y_h\|_{0,T} \|z_h\|_{0,T}) \\
& \leq \frac{1}{2} h_T^2 \|cp_h + y_h^d - y_h - \tilde{\sigma}_h\|_{0,T}^2 + 2c_1^2 c_3^2 \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,T}^2 \\
& \quad + 2c_1^2 c_2^2 c^2 h_T^2 \|p(u_h, \tilde{\sigma}_h) - p_h\|_{0,T}^2 + 2h_T^2 c_1^2 c_2^2 \|y_h^d - y^d\|_{0,T}^2 \\
& \quad + 2h_T^2 c_1^2 c_2^2 \|y(u_h) - y_h\|_{0,T}^2
\end{aligned}$$

Now one may insert y into the last term of the right-hand side of (6.3.6), and then, the combination of (6.3.4) and (6.3.6) gives the claimed result. \square

Lemma 6.3.4. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, and let $\tilde{\eta}_{y,T}$ and $\tilde{\eta}_{y,E}$ be given by (6.1.8a) and (6.1.8c), respectively. Then, there exists a positive constant C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for each $E \in \mathcal{E}_h(\Omega)$, there holds*

$$\tilde{\eta}_{y,E}^2 \leq C \left(\|y - y_h\|_{1,\omega_E}^2 + h_E^2 \|u - u_h\|_{0,\omega_E}^2 + \sum_{\nu=1}^2 \tilde{\eta}_{y,T_\nu}^2 \right). \tag{6.3.7}$$

Proof. Again we refer to the proof of Lemma 3.3.3, which is similar. \square

Lemma 6.3.5. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3.1) and (2.4.5), respectively, let $p(u_h, \tilde{\sigma}_h)$ be as defined in (6.1.15), and let $\tilde{\eta}_{p,T}$ and $\tilde{\eta}_{p,E}$ be given by (6.1.8b) and (6.1.8d), respectively. Then, there exists a positive constant*

C , depending only on the shape regularity of $\mathcal{T}_h(\Omega)$, such that for each $E \in \mathcal{E}_h(\Omega)$, there holds

$$\begin{aligned} \tilde{\eta}_{p,E}^2 &\leq C \left(\|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\omega_E}^2 + h_E^2 \|y - y(u_h)\|_{0,\omega_E}^2 \right. \\ &\quad \left. + h_E^2 \|y - y_h\|_{0,\omega_E}^2 + \sum_{\nu=1}^2 \tilde{\eta}_{p,T_\nu}^2 \right). \end{aligned} \quad (6.3.8)$$

Proof. Again, we will state the proof because the claim is different from the one of Lemma 3.3.4. Set $\zeta_E := (\boldsymbol{\nu}_E \cdot [\nabla p_h])|_E$ and $z_h := \tilde{\zeta}_E \vartheta_E$. One can observe that $\Delta p_h|_T = 0$, that $z_h|_{\partial\omega_E} = 0$, and that z_h is an admissible test function in (6.1.15). Thus, applying Green's formula and using (3.3.1d), one may find

$$\begin{aligned} \tilde{\eta}_{p,E}^2 &= h_E \|\boldsymbol{\nu}_E \cdot [\nabla p_h]\|_{0,E}^2 \\ &\leq c_4 h_E (\boldsymbol{\nu}_E \cdot [\nabla p_h], \zeta_E \vartheta_E)_{0,E} \\ &= c_4 h_E \sum_{\nu=1}^2 \{ (\boldsymbol{\nu}_{\partial T_\nu} \cdot \nabla p_h, z_h)_{0,\partial T_\nu} - \underbrace{(\Delta p_h, z_h)_{0,T_\nu}}_{=0} \} \\ &= c_4 h_E \left((\nabla(p_h - p(u_h, \tilde{\sigma}_h)), \nabla z_h)_{0,\omega_E} + c(p_h - p(u_h, \tilde{\sigma}_h), z_h)_{0,\omega_E} \right. \\ &\quad \left. + \langle \tilde{\sigma}_h, z_h \rangle + (y(u_h) - y^d - cp_h, z_h)_{0,\omega_E} \right) \\ &= c_4 h_E \left((\nabla(p_h - p(u_h, \tilde{\sigma}_h)), \nabla z_h)_{0,\omega_E} + c(p_h - p(u_h, \tilde{\sigma}_h), z_h)_{0,\omega_E} \right. \\ &\quad \left. + (y(u_h) - y, z_h) + (y_h - y^d + \tilde{\sigma}_h - cp_h, z_h)_{0,\omega_E} + (y - y_h, z_h)_{0,\omega_E} \right) \end{aligned} \quad (6.3.9)$$

Now, we may apply Cauchy's inequality, as well as (3.3.2a) and (3.3.2b) in order to obtain

$$\begin{aligned} &\tilde{\eta}_{p,E}^2 \\ &\leq c_4 h_E (\|p_h - p(u_h, \tilde{\sigma}_h)\|_{1,\omega_E} \underbrace{\|z_h\|_{1,\omega_E}}_{\leq c_7 h_E^{-\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla p_h]\|_{0,E}} \\ &\quad + \|p_h - p(u_h, \tilde{\sigma}_h)\|_{0,\omega_E} \underbrace{\|z_h\|_{0,\omega_E}}_{\leq c_6 h_E^{\frac{1}{2}} \|\boldsymbol{\nu}_E \cdot [\nabla p_h]\|_{0,E}} + \|y - y(u_h)\|_{0,\omega_E} \|z_h\|_{0,\omega_E} \\ &\quad + \|y_h - y^d + \tilde{\sigma}_h - cp_h\|_{0,\omega_E} \|z_h\|_{0,\omega_E} + \|y_h - y\|_{0,\omega_E} \|z_h\|_{0,\omega_E}) \end{aligned} \quad (6.3.10)$$

An application of Young's inequality results in

$$\begin{aligned} \tilde{\eta}_{p,E}^2 &\leq \frac{5}{8} h_E \|\boldsymbol{\nu}_E \cdot [\nabla p_h]\|_{0,E}^2 + 2c_4^2 c_7^2 \|p(u_h, \tilde{\sigma}_h) - p_h\|_{1,\omega_E}^2 \\ &\quad + 2c_4^2 c_6^2 h_E^2 \|p(u_h, \tilde{\sigma}_h) - p_h\|_{0,\omega_E}^2 + 2c_4^2 c_6^2 h_E^2 \|y - y(u_h)\|_{0,\omega_E}^2 \\ &\quad + 2c_4^2 c_6^2 h_E^2 \|y - y_h\|_{0,\omega_E}^2 + 2c_4^2 c_6^2 \sum_{\nu=1}^2 h_E^2 \|y_h - y^d + \tilde{\sigma}_h - cp_h\|_{0,T_\nu}^2, \end{aligned} \quad (6.3.11)$$

which gives the assertion (6.3.8) since, due to the shape-regularity, there holds $h_E \preceq h_{T_\nu}$ for $\nu = 1, 2$. \square

Proof of Theorem 6.3.1.

We see again that (6.3.1a) follows from Lemma 6.3.2, Lemma 6.3.3, Lemma 6.3.4, and Lemma 6.3.5 if one considers that summing up over all edges only gives a finite overlap. The quantity $\|y - y(u_h)\|_{0,\Omega}$, which results from the corresponding local terms of Lemma 6.3.3 and Lemma 6.3.5, may be handled with the help of inequality (3.2.5). The claim (6.3.1b) is a direct consequence of (6.3.1a). \square

6.4 Discussion of the consistency error

In this section, we want to examine the consistency error as given by (6.1.16). Remark 6.1.2 gives some arguments why it should not cause any problems. There are even some special cases, for which it vanishes. This section will deal with these situations. They involve strong assumptions which concern the feasibility of the auxiliary state and the approximation of the continuous active zone, but they also show the similarities to obstacle problems.

Before we may state the results, we define the discrete active set $\mathcal{A}_h \subset \mathcal{T}_h(\Omega)$ and the discrete inactive set $\mathcal{I}_h \subset \mathcal{T}_h(\Omega)$ according to

$$\mathcal{A}_h := \{T \in \mathcal{T}_h(\Omega) \mid y_h(n) = \psi(n) \forall n \in \mathcal{N}_h(T)\} , \quad (6.4.1a)$$

$$\mathcal{I}_h := \{T \in \mathcal{T}_h(\Omega) \mid y_h(n) < \psi(n) \forall n \in \mathcal{N}_h(T)\} . \quad (6.4.1b)$$

This corresponds to the definition of the discrete free boundary, as it was given in (5.1.2), i.e., there holds

$$\mathcal{F}_h = \mathcal{T}_h(\Omega) \setminus (\mathcal{A}_h \cup \mathcal{I}_h) . \quad (6.4.2)$$

Lemma 6.4.1. *Assume that $u = u_h$ and $\mathcal{A}_h \cup \mathcal{F}_h \subset \mathcal{A}$. Then, the consistency error $e_c(u_h)$ vanishes.*

Proof. Since $u = u_h$, we have $y(u_h) = y$, and thus, one may conclude by applying (2.3.14c) that the first term of the consistency error vanishes. On the other side, on each triangle $T \in \mathcal{T}_h(\Omega)$ there holds that $\tilde{\sigma}_{h|T} = 0$ or $y(u_h) = y = \psi$, due to (2.3.14b) and the assumption $\mathcal{A}_h \cup \mathcal{F}_h \subset \mathcal{A}$. Consequently, the second term of the consistency error also has the value 0. \square

Lemma 6.4.2. *Assume that the upper bound ψ is a linear function, that $\Gamma_D = \partial\Omega$, and that the following conditions hold:*

$$y(u_h) \leq \psi , \quad (6.4.3a)$$

$$\{x \in \Omega \mid y_h(x) = \psi(x)\} \subset \{x \in \Omega \mid y(u_h)(x) = \psi(x)\} . \quad (6.4.3b)$$

Then, (6.2.1) is also valid if the consistency error $e_c(u_h)$ does not occur in the right-hand side of (6.2.1).

The assumptions (6.4.3) demand that $y(u_h)$ is feasible for the continuous problem and active where y_h is active. Before we come to the proof of Lemma 6.4.1, let us first state the following auxiliary result which is taken from [NSV03]:

Lemma 6.4.3. *Let $w_h \in S_h$, $n \in \mathcal{N}_h$, be an interior node and $T \in \mathcal{T}_h$ a triangle with $n \in T$. Furthermore, assume that $w_h(n) = 0$ and either $w_h \leq 0$ or $w_h \geq 0$ on T . Then, there exists a constant C_1 , depending only on the shape regularity of the triangulation, such that:*

$$\|w_h\|_{0,T} \leq C_1 h_T \left(\sum_{E \subset \tilde{\omega}_T} h_E \|\llbracket \nabla w_h \rrbracket \cdot \boldsymbol{\nu}_E\|_{0,E}^2 \right)^{\frac{1}{2}}, \quad (6.4.4)$$

where $\tilde{\omega}_T$ is defined as in (6.2.12).

With the help of Lemma 6.4.3, one can now formulate the

proof of Lemma 6.4.2.

Obviously, it is sufficient to prove that Lemma 6.2.2 holds without the consistency error appearing in the right-hand side. The consistency error does not vanish in general, but can be estimated from above by other parts of the right-hand side of (6.2.1). Therefore, we will prove a new version of the estimate (6.2.2) with a right-hand side which includes parts of the error estimator and data oscillation as they appear in (6.2.1).

If one follows the lines of the proof of Lemma 6.2.2, one may notice that, due to the assumption (6.4.3a) and (2.3.14a), the right-hand side of (6.2.8) is non-positive. Thus, it remains to deal with the first term of the right-hand side of (6.2.7). Instead of inserting ψ , as we did in (6.2.9), we insert y_h . That leads to

$$\langle \tilde{\sigma}_h, y - y(u_h) \rangle = \langle \tilde{\sigma}_h, y - y_h \rangle + \langle \tilde{\sigma}_h, y_h - y(u_h) \rangle \quad (6.4.5)$$

Let us first have a closer look at the second term of the right-hand side of (6.4.5), which we can split according to

$$\begin{aligned} & \langle \tilde{\sigma}_h, y_h - y(u_h) \rangle & (6.4.6) \\ = & \sum_{T \in \mathcal{A}_h} (\tilde{\sigma}_h, \psi - y(u_h))_{0,T} + \sum_{T \in \mathcal{I}_h} (\tilde{\sigma}_h, y_h - y(u_h))_{0,T} + \sum_{T \in \mathcal{F}_h} (\tilde{\sigma}_h, y_h - y(u_h))_{0,T}, \end{aligned}$$

where the first term has, due to assumption (6.4.3b), value 0 and the second sum vanishes due to (2.4.8b). For the remaining part of the right-hand side of (6.4.6),

one may find with the help of Hölder's and Young's inequality

$$\begin{aligned}
& \sum_{T \in \mathcal{F}_h} (\tilde{\sigma}_h, y_h - y(u_h))_{0,T} & (6.4.7) \\
& \leq \sum_{T \in \mathcal{F}_h} \|\tilde{\sigma}_h\|_{0,T} \|y_h - y(u_h)\|_{0,T} \\
& \leq \sum_{T \in \mathcal{F}_h} h_T^2 \|\nabla \tilde{\sigma}_h\|_{0,T} \|\nabla(y_h - y(u_h))\|_{0,T} \\
& \leq \sum_{T \in \mathcal{F}_h} h_T^4 \|\nabla \tilde{\sigma}_h\|_{0,T}^2 + \|\nabla(y_h - y(u_h))\|_{0,\Omega}^2
\end{aligned}$$

Here, the Poincaré-Friedrich inequality has been applied twice, which is possible since for $T \in \mathcal{F}_h$, there exists a node $n_1 \in \mathcal{N}_h(T)$ such that $\tilde{\sigma}_h(n_1) = 0$ and a node $n_2 \in \mathcal{N}_h(T)$ such that $(y(u_h))(n_2) = \psi(n_2) = y_h(n_2)$ (apply assumption (6.4.3b)). The first term of the right-hand side of (6.4.7) can now be estimated as in (6.2.23). This results in

$$\begin{aligned}
\langle \tilde{\sigma}_h, y_h - y(u_h) \rangle & \leq \|y(u_h) - y_h\|_{1,\Omega}^2 + \text{osc}_h^2(y^d) + \text{osc}_h^2(y_h) & (6.4.8) \\
& \quad + \sum_{T \in \mathcal{T}(\Omega)} \tilde{\eta}_{p,T}^2 + c^2 \text{osc}_h^2(p_h)
\end{aligned}$$

The right-hand side of (6.4.8) consists only of parts of the right-hand side of (6.2.2) or parts of the error estimator or data oscillation as they appear in (6.2.1).

Finally, we look closer at the first term of (6.4.5), which is a bit more involved. The facts that $\tilde{\sigma}_h = 0$ on each $T \in \mathcal{I}_h$ and that $\tilde{\sigma}_h \geq 0$ and $y - y_h = y - \psi \leq 0$ on each $T \in \mathcal{A}_h$ imply the estimate

$$\langle \tilde{\sigma}_h, y - y_h \rangle \leq \sum_{T \in \mathcal{F}_h} \int_T \tilde{\sigma}_h (y - y_h) . \quad (6.4.9)$$

In order to estimate the right-hand side of (6.4.9), we examine each addend separately. Therefore, let us now distinguish two cases for $T \in \mathcal{F}_h$:

Case 1: $T \cap \partial\Omega \neq \emptyset$

Remember that I_h denotes the Clément interpolation operator. With Cauchy's inequality, we may find

$$\begin{aligned}
& \int_T \tilde{\sigma}_h (y - y_h) & (6.4.10) \\
& = \int_T \tilde{\sigma}_h [(y - y_h) - I_h(y - y_h)] + \int_T \tilde{\sigma}_h I_h(y - y_h) \\
& \leq \|\tilde{\sigma}_h\|_{0,T} (\|(y - y_h) - I_h(y - y_h)\|_{0,T} + \|I_h(y - y_h)\|_{0,T}) .
\end{aligned}$$

Since for at least on node $n_2 \in \mathcal{N}_h(T)$, there holds $\tilde{\sigma}_h(n_2) = 0$, we can estimate

$$\|\tilde{\sigma}_h\|_{0,T} \leq Ch_T \|\nabla \tilde{\sigma}_h\|_{0,T} \quad (6.4.11)$$

Here and in the remaining part of this proof, C denotes a generic constant which is independent of the local mesh size. Furthermore, there holds $y - y_h \in H_0^1(\Omega)$ due to our assumption $\Gamma_D = \partial\Omega$. Since $T \cap \partial\Omega \neq \emptyset$, we can use the following property of the Clément interpolation operator:

$$\|(y - y_h) - I_h(y - y_h)\|_{0,T} \leq Ch_T \|\nabla(y - y_h)\|_{0,T} \quad (6.4.12)$$

On the other side, by considering $T \cap \partial\Omega \neq \emptyset$, which implies that $y = y_h = 0$ in at least one node of T , one furthermore can estimate

$$\|I_h(y - y_h)\|_{0,T} \leq Ch_T \|\nabla(y - y_h)\|_{0,T} . \quad (6.4.13)$$

Inserting (6.4.11), (6.4.12), and (6.4.13) into (6.4.10) and applying Young's inequality, one may conclude that there exists a constant C such that for each $\epsilon > 0$

$$\int_T \tilde{\sigma}_h(y - y_h) \leq \epsilon \|\nabla(y - y_h)\|_{0,T}^2 + \frac{C}{\epsilon} h_T^4 \|\nabla \tilde{\sigma}_h\|_{0,T}^2 , \quad (6.4.14)$$

where one may handle the last term of (6.4.14) as in (6.2.23).

Case 2: $T \cap \partial\Omega = \emptyset$

Because of the positiveness of $\tilde{\sigma}_h$, the feasibility of y , and Cauchy's inequality, there holds

$$\begin{aligned} & \int_T \tilde{\sigma}_h(y - y_h) \quad (6.4.15) \\ &= \int_T \underbrace{\tilde{\sigma}_h}_{\geq 0} \underbrace{(y - \psi)}_{\leq 0} + \int_T \tilde{\sigma}_h(\psi - y_h) \\ &\leq \int_T \tilde{\sigma}_h(\psi - y_h) \\ &\leq \|\tilde{\sigma}_h\|_{0,T} \|\psi - y_h\|_{0,T} \end{aligned}$$

Notice that, due to the shape-regularity of the triangulation $\mathcal{T}_h(\Omega)$, $h_T \preceq h_E \preceq h_T$ if $E \subset T$.

Now one can apply Lemma 6.4.3 with $w_h = \psi - y_h \geq 0$, which results in

$$\|\psi - y_h\|_{0,T} \leq Ch_T \left(\sum_{E \in \mathcal{E}_h(\omega_T)} \tilde{\eta}_{y,E}^2 \right)^{\frac{1}{2}} , \quad (6.4.16)$$

where we have used the assumption that ψ is a linear function. Since (6.4.11) is also valid for the second case, one gains by inserting (6.4.16) into (6.4.15) the existence of a constant C such that

$$\int_T \tilde{\sigma}_h(y - y_h) \leq C(h_T^4 \|\nabla \tilde{\sigma}_h\|_{0,T}^2 + \sum_{E \in \omega_T} \tilde{\eta}_{y,E}^2), \quad (6.4.17)$$

for all $T \in \mathcal{F}_h$ with $T \cap \partial\Omega = \emptyset$. Here, the first term on the right-hand side can again be handled as in (6.2.23).

Then, the combination of (6.4.5), (6.4.8), (6.4.9), (6.4.14), and (6.4.17) implies for each $\epsilon > 0$ the existence of a constant $C(\epsilon)$ such that

$$\begin{aligned} & \langle \tilde{\sigma}_h, y - y(u_h) \rangle \quad (6.4.18) \\ & \leq \epsilon \|y - y_h\|_{1,\Omega}^2 + C(\epsilon) \left(\sum_{T \in \mathcal{T}_h} \tilde{\eta}_{p,T}^2 + \sum_{E \in \mathcal{E}_h} \tilde{\eta}_{y,E}^2 + \text{osc}_h^2(y^d) + \text{osc}_h^2(y_h) + c^2 \text{osc}_h^2(p_h) \right) \end{aligned}$$

Taking (6.2.5) into account and choosing $\epsilon = \frac{\alpha}{8\tilde{c}^2}$, one arrives at

$$\begin{aligned} & \langle \tilde{\sigma}_h, y - y(u_h) \rangle \quad (6.4.19) \\ & \leq \frac{\alpha}{4} \|u - u_h\|_{0,\Omega}^2 + C \left(\|y(u_h) - y_h\|_{1,\Omega}^2 + \sum_{T \in \mathcal{T}_h} \tilde{\eta}_{p,T}^2 + \sum_{E \in \mathcal{E}_h} \tilde{\eta}_{y,E}^2 \right. \\ & \quad \left. + \text{osc}_h^2(y^d) + \text{osc}_h^2(y_h) + c^2 \text{osc}_h^2(p_h) \right) \end{aligned}$$

Now, the combination of (6.4.19) with (6.2.6), (6.2.7), and the fact that the second term of the right-hand side of (6.2.7) vanishes, results in a new version of the estimate (6.2.10), which reads as follows:

$$\begin{aligned} & \|u - u_h\|_{0,\Omega}^2 \quad (6.4.20) \\ & \leq \frac{3}{4} \|u - u_h\|_{0,\Omega}^2 + C \left(\|u^d - u_h^d\|_{0,\Omega}^2 + \|p_h - p(u_h, \tilde{\sigma}_h)\|_{1,\Omega}^2 + \|y(u_h) - y_h\|_{1,\Omega}^2 \right. \\ & \quad \left. + \sum_{T \in \mathcal{T}_h} \tilde{\eta}_{p,T}^2 + \sum_{E \in \mathcal{E}_h} \tilde{\eta}_{y,E}^2 + \text{osc}_h^2(y^d) + \text{osc}_h^2(y_h) + c^2 \text{osc}_h^2(p_h) \right) \end{aligned}$$

Then, it is easy to see that (6.2.5) and (6.4.20) imply the following new version of Lemma 6.2.2:

$$\begin{aligned} & \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{1,\Omega}^2 \quad (6.4.21) \\ & \leq C \left(\|p_h - p(u_h, \tilde{\sigma}_h)\|_{1,\Omega}^2 + \|y(u_h) - y_h\|_{1,\Omega}^2 + \sum_{T \in \mathcal{T}_h} \tilde{\eta}_{p,T}^2 + \sum_{E \in \mathcal{E}_h} \tilde{\eta}_{y,E}^2 \right. \\ & \quad \left. + \mu_h^2(u^d) + \text{osc}_h^2(y^d) + \text{osc}_h^2(y_h) + c^2 \text{osc}_h^2(p_h) \right), \end{aligned}$$

The observation at the beginning of this proof told us that (6.4.21) is sufficient for the claim of Theorem 6.4.2. \square

Of course, the assumptions of Lemma 6.4.1 and Lemma 6.4.2 are far from being true in general. However, they might hold asymptotically if the auxiliary state $y(u_h)$ tends to fulfill the discrete and continuous complementarity conditions.

6.5 Numerical experiment

In this section, we will finally test the measure extension error estimator $\tilde{\eta}$ numerically. Furthermore, we want to compare it with the error estimator η as derived in the previous chapters. We will not deal with the Lavrentiev regularization method in this section since it does not seem to be a real advantage compared to the unregularized case.

For the measure extension error estimator, we use the same adaptive algorithm as proposed in the previous chapter. We just replace the quantities of the error estimator by the corresponding quantities of the measure extension error estimator $\tilde{\eta}$ and enlarge the inequalities of the bulk criterion in (5.1.6) with the data oscillation in y_h and maybe p_h .

We applied the error estimator $\tilde{\eta}$ to the two numerical test examples of the previous chapter. However, since there is no great difference in the performance compared to the error estimator η , we will not state the explicit results here. Instead, we will present a new example for which we indeed noticed a difference. This example features an optimal state and an optimal control which have a singularity at the origin and an optimal adjoint state which strongly differs from the modified adjoint state. As we will see in the following, this difference of the adjoint functions causes the error estimator η to refine in a region where a high density of nodes is not necessary and does not lead to better approximations of the optimal control or the optimal state. This undesirable refinement does not occur for the error estimator $\tilde{\eta}$.

In particular, the data of Example 3 read as follows:

$$\begin{aligned} \Omega &:= (-2, 2)^2 \setminus (0, 2) \times (-2, 0) \quad , \quad y^d := y(r, \varphi) + \Delta p(r, \varphi) + \sigma(r, \varphi) \quad , \\ u^d &:= u(r, \varphi) + \alpha^{-1} p(r, \varphi) \quad , \quad \alpha := 0.1 \quad , \quad c = 0 \quad , \quad \Gamma_D := \partial\Omega \quad . \end{aligned}$$

The upper bound ψ is chosen in such a way that it has value 0 in a L-shaped region inside the domain (indicated by the blue colored area in the left picture of Figure 6.1) and is larger than 0 outside this region (compare Figure 6.1, left). The optimal solutions $y = y(r, \varphi)$, $u = u(r, \varphi)$, and $p = p(r, \varphi)$ are given in polar coordinates as follows:

$$\begin{aligned} y(r, \varphi) &= -r^{\frac{2}{3}} \gamma_1(r) \sin\left(\frac{2\varphi}{3}\right) \quad , \\ p(r, \varphi) &= \gamma_2(r) \left(r^4 - \frac{3}{2} r^3 + \frac{9}{16} r^2 \right) \sin\left(\frac{2\varphi}{3}\right) \quad , \\ u(r, \varphi) &= \frac{7}{3} r^{-\frac{1}{2}} \gamma_1'(r) \sin\left(\frac{2\varphi}{3}\right) + r^{\frac{2}{3}} \gamma_1''(r) \sin\left(\frac{2\varphi}{3}\right) \quad , \end{aligned}$$

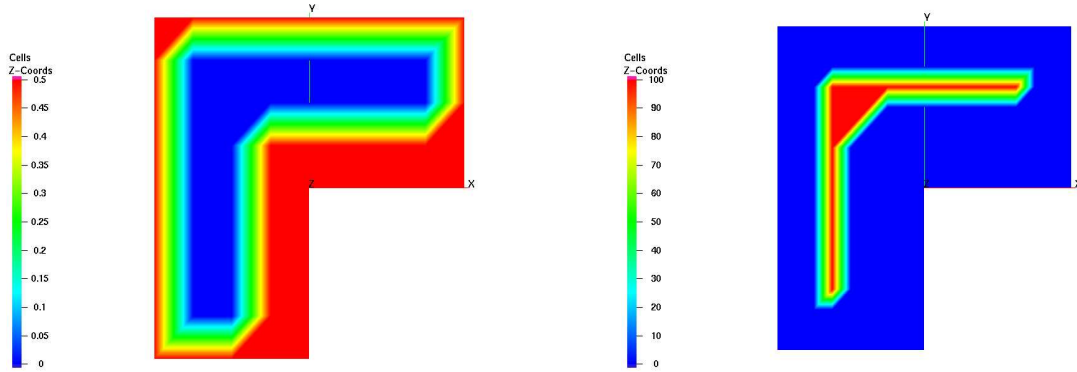


Figure 6.1: Example : Visualization of the upper bound ψ (left) and the Lagrange multiplier σ (right).

where

$$\gamma_1(r) := \begin{cases} 1, & r < 0.25 \\ 0, & r > 0.75 \\ -192(r - 0.25)^5 + 240(r - 0.25)^4 - 80(r - 0.25)^3 + 1, & \text{otherwise} \end{cases}$$

$$\gamma_2(r) := \begin{cases} 1, & r < 0.75 \\ 0, & \text{otherwise} \end{cases}.$$

Notice that the active zone is given by the region where the upper bound ψ takes the value 0. Thus, the optimal adjoint control has to take non-negative values inside the active region and has to vanish outside of it. Consequently, we chose the adjoint control $\sigma \in L^2(\Omega) \subset C(\Omega)^*$ as shown in the right picture of Figure 6.1. The adjoint control and therefore also the desired state have large values inside the active region. We do not declare the explicit formulas for σ and ψ since you have to distinguish a lot of cases to describe them. It is easy to check that the declared functions are indeed the optimal solution of the given problem.

While Figure 6.2 shows the desired state y^d (left) and the control shift u^d (right), Figure 6.3 shows the optimal control (left) and optimal state (right). Figure 6.4 shows the optimal adjoint state p (left) and the modified adjoint state \hat{p} (right), as it is used for $\eta_{\hat{p}}$. It can be seen that these two functions are quite different, especially outside of the circle $B_{0.75}(0)$. This will lead to an overrefinement of the error estimator η_p outside of this circle.

The initial simplicial triangulation was chosen such that the domain Ω was subdivided into right-angled triangles and such that the upper bound ψ can be interpolated exactly on the initial grid. The parameters $\theta_i, 1 \leq i \leq 4$, were all chosen with the value 0.7. Figure 6.5 shows the adaptively generated grids resulting from the error estimator $\tilde{\eta}$ after seven (left) and ten (right) refinement steps with 3766 and 11042 nodes, respectively. There, the red colored zone indicates the discrete active region and the yellow colored area the discrete inactive region. Remember that a triangle is colored red if the upper bound is sharp in all of its nodes. One may notice

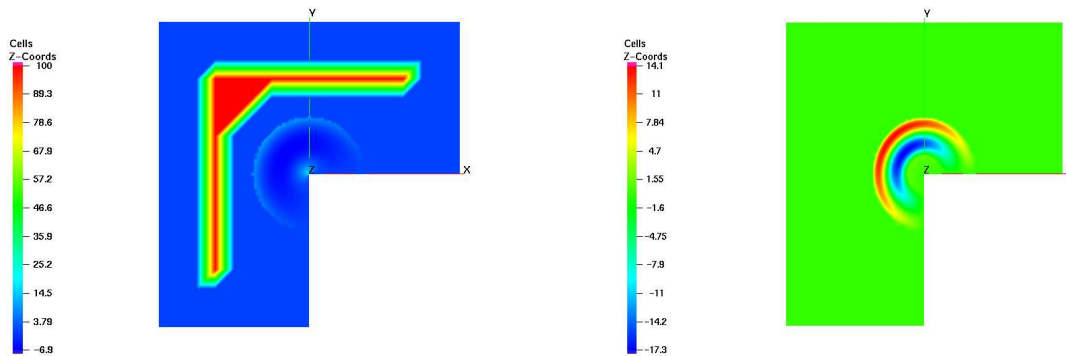


Figure 6.2: Example 3 : Visualization of the desired state y^d (left) and the control shift u^d (right).

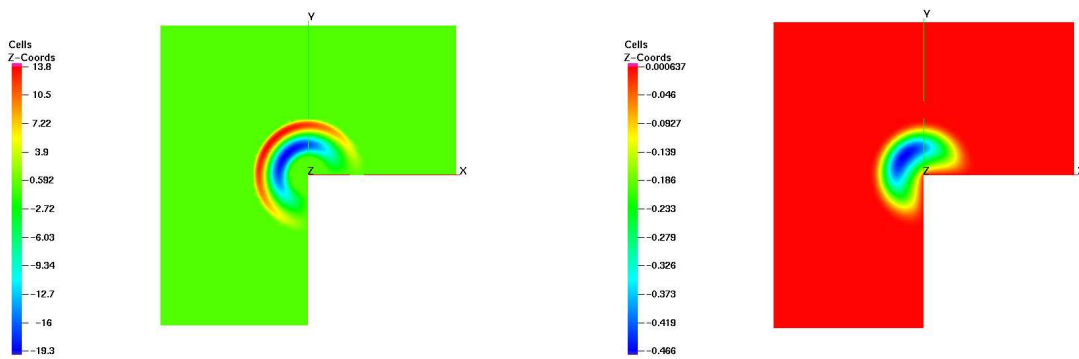


Figure 6.3: Example 3 : Visualization of the optimal control u (left) and the optimal state y (right).

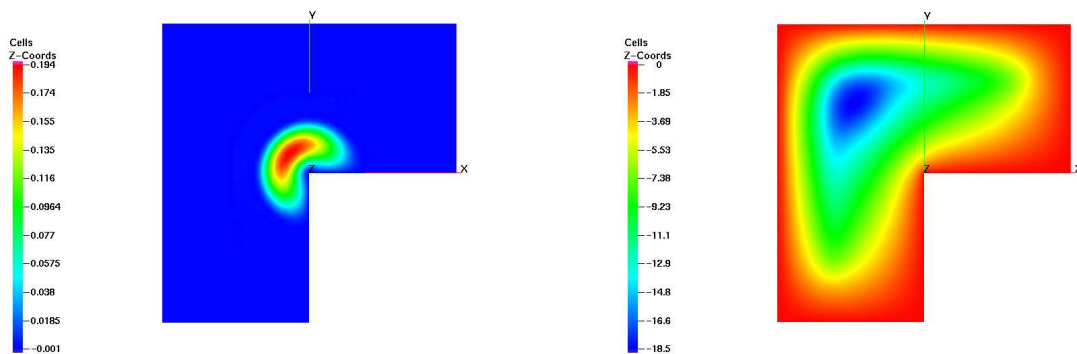


Figure 6.4: Example 3 : Visualization of the optimal adjoint state p (left) and the modified adjoint state \hat{p} (right).

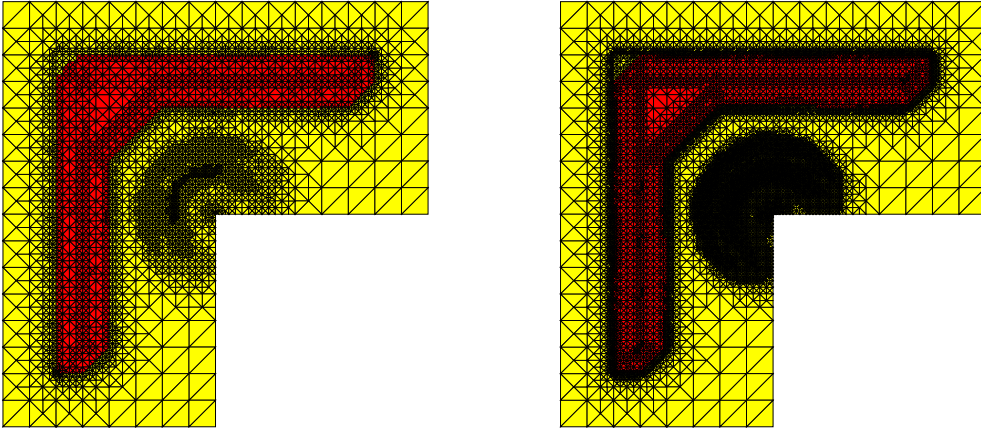


Figure 6.5: Example 3 : Adaptively generated grids with the error estimator $\tilde{\eta}$ after 7 (left) and 10 (right) refinement steps, with 3766 and 11042 nodes, respectively.

that especially the area inside of the circle $B_{0.75}(0)$ is well dissolved. Additionally, the discrete free boundary is well dissolved. However, for this example, one gets even slightly better results if one does not refine along the discrete free boundary since the upper bound ψ and the continuous active zone can be approximated exactly on the initial grid.

Table 6.1 to Table 6.4 contain the same information as in the previous examples. Thus, the same remarks as already mentioned before also apply here. We again estimated the consistency error similarly as described in the previous chapter. One can see that it decreases in every iteration step, and thus, it does not seem to cause any problems.

On the other side, Figure 6.6 shows the adaptively generated grids according to the error estimators η after 5 (left) and 8 (right) refinement steps, with 2736 and 11672 nodes, respectively. One may notice that η_p dissolves the active region almost better than the singularity at the origin. This is due to the fact that the modified adjoint state is especially hard to approximate outside the circle $B_{0.75}(0)$. There, however, a high density of nodes is useless.

Finally, Figure 6.7 shows the discretization error in the control (left) and the state (right) for the error estimator $\tilde{\eta}$, η , and an uniform refinement strategy. It is noticeable that only the error estimator $\tilde{\eta}$ provides a benefit compared to the uniform refinement strategy, whereas the error estimator η is even a bit worse than the uniform refinement strategy. This is due to the fact that η refines strongly in a region in which the refinement does not result in a better approximation of the control or the state. The modified adjoint state \hat{p} , which is used in the error estimator η , is defined by deleting the adjoint control in the right-hand side of p . This example shows that it may happen that the neglect of the adjoint control in the right-hand side of the adjoint state equation may cause a loss of crucial information.

However, one has to admit that this example is really constructive since the desired

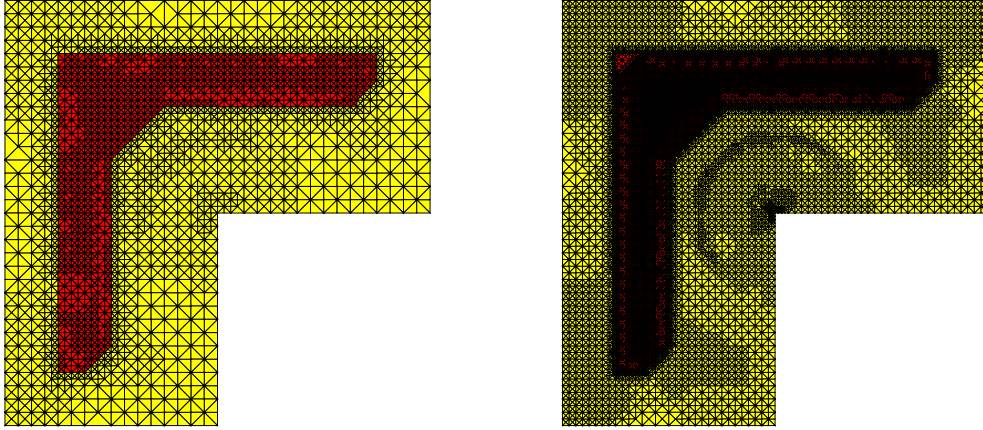


Figure 6.6: Example 3 : Adaptively generated grids for the error estimator η after 5 (left) and 8 (right) refinement steps, with 2736 and 11672 nodes, respectively.

Table 6.1: Example 3: Convergence history of the adaptive FEM, Part I: Total discretization error and discretization errors in the state, control, adjoint state and modified adjoint state.

| ℓ | N | $\ z - z_\ell\ $ | $\ y - y_\ell\ _1$ | $\ u - u_\ell\ _0$ | $\ p - p_\ell\ _0$ |
|--------|-------|------------------|--------------------|--------------------|--------------------|
| 0 | 225 | 6.92e+00 | 6.69e-01 | 6.89e+00 | 2.32e-02 |
| 1 | 358 | 4.74e+00 | 5.40e-01 | 4.71e+00 | 1.50e-02 |
| 2 | 577 | 3.03e+00 | 3.85e-01 | 3.01e+00 | 8.83e-03 |
| 3 | 900 | 1.93e+00 | 2.93e-01 | 1.91e+00 | 5.60e-03 |
| 4 | 1295 | 1.33e+00 | 2.13e-01 | 1.31e+00 | 3.48e-03 |
| 5 | 1881 | 8.97e-01 | 1.69e-01 | 8.81e-01 | 2.67e-03 |
| 6 | 2650 | 5.61e-01 | 1.27e-01 | 5.46e-01 | 2.23e-03 |
| 7 | 3766 | 4.01e-01 | 9.94e-02 | 3.88e-01 | 1.86e-03 |
| 8 | 5426 | 2.86e-01 | 7.81e-02 | 2.75e-01 | 1.45e-03 |
| 9 | 7779 | 2.16e-01 | 5.97e-02 | 2.08e-01 | 1.21e-03 |
| 10 | 11042 | 1.59e-01 | 4.79e-02 | 1.52e-01 | 1.07e-03 |

Table 6.2: Example 3: Convergence history of the adaptive FEM, Part II: Components of the error estimator and the data oscillations and the consistency error.

| ℓ | N_{dof} | $\tilde{\eta}_y$ | $\tilde{\eta}_p$ | $\mu_\ell(u^d)$ | $\text{osc}_\ell(y^d)$ | $e_c(u_h)$ |
|--------|------------------|------------------|------------------|-----------------|------------------------|------------|
| 0 | 225 | 1.91e+00 | 1.07e+01 | 6.75e+00 | 1.26e+00 | 1.72e+00 |
| 1 | 358 | 1.66e+00 | 5.47e+00 | 4.61e+00 | 7.05e-01 | 7.05e-01 |
| 2 | 577 | 1.40e+00 | 3.06e+00 | 2.96e+00 | 4.41e-01 | 2.78e-01 |
| 3 | 900 | 1.13e+00 | 1.77e+00 | 1.90e+00 | 3.00e-01 | 1.16e-01 |
| 4 | 1295 | 8.71e-01 | 1.09e+00 | 1.31e+00 | 2.11e-01 | 5.39e-02 |
| 5 | 1881 | 6.72e-01 | 6.70e-01 | 8.80e-01 | 1.41e-01 | 3.16e-02 |
| 6 | 2650 | 5.09e-01 | 4.90e-01 | 5.45e-01 | 9.56e-02 | 1.32e-02 |
| 7 | 3766 | 3.91e-01 | 3.62e-01 | 3.88e-01 | 6.56e-02 | 6.59e-03 |
| 8 | 5426 | 2.96e-01 | 2.96e-01 | 2.76e-01 | 4.78e-02 | 4.56e-03 |
| 9 | 7779 | 2.28e-01 | 1.83e-01 | 2.08e-01 | 3.63e-02 | 2.18e-03 |
| 10 | 11042 | 1.77e-01 | 1.48e-01 | 1.52e-01 | 2.64e-02 | 1.42e-03 |

Table 6.3: Example 3: Convergence history of the adaptive FEM, Part III: Average values of the local estimators.

| ℓ | N_{dof} | $\tilde{\eta}_{y,T}$ | $\tilde{\eta}_{p,T}$ | $\tilde{\eta}_{y,E}$ | $\tilde{\eta}_{p,E}$ | $\mu_\ell(u^d)$ | $\text{osc}_\ell(y^d)$ |
|--------|------------------|----------------------|----------------------|----------------------|----------------------|-----------------|------------------------|
| 0 | 225 | 2.81e-02 | 3.67e-01 | 6.11e-03 | 3.89e-03 | 1.07e-01 | 2.20e-02 |
| 1 | 358 | 2.00e-02 | 1.47e-01 | 2.91e-03 | 1.78e-03 | 5.69e-02 | 8.44e-03 |
| 2 | 577 | 1.38e-02 | 6.21e-02 | 1.36e-03 | 7.85e-04 | 2.74e-02 | 3.57e-03 |
| 3 | 900 | 9.38e-03 | 2.76e-02 | 6.85e-04 | 3.53e-04 | 1.39e-02 | 1.65e-03 |
| 4 | 1295 | 6.50e-03 | 1.41e-02 | 3.42e-04 | 1.81e-04 | 7.80e-03 | 8.49e-04 |
| 5 | 1881 | 4.43e-03 | 7.33e-03 | 1.88e-04 | 9.47e-05 | 4.28e-03 | 4.36e-04 |
| 6 | 2650 | 3.12e-03 | 4.30e-03 | 9.96e-05 | 5.38e-05 | 2.29e-03 | 2.32e-04 |
| 7 | 3766 | 2.18e-03 | 2.80e-03 | 5.33e-05 | 2.85e-05 | 1.33e-03 | 1.22e-04 |
| 8 | 5426 | 1.51e-03 | 1.87e-03 | 2.90e-05 | 1.64e-05 | 7.27e-04 | 6.52e-05 |
| 9 | 7779 | 1.05e-03 | 9.95e-04 | 1.49e-05 | 8.76e-06 | 4.16e-04 | 3.71e-05 |
| 10 | 11042 | 7.36e-04 | 6.85e-04 | 8.36e-06 | 4.97e-06 | 2.41e-04 | 2.07e-05 |

Table 6.4: Example 3: Convergence history of the adaptive FEM, Part IV: Percentages in the bulk criteria.

| ℓ | N_{dof} | $\mathcal{M}^{fb,T}$ | \mathcal{M}^E | $\mathcal{M}^{\tilde{\eta},T}$ | $\mathcal{M}^{\mu,T}$ | $M^{\text{osc},T}$ |
|--------|------------------|----------------------|-----------------|--------------------------------|-----------------------|--------------------|
| 0 | 225 | 25.5 | 2.9 | 21.1 | 3.6 | 4.9 |
| 1 | 358 | 22.3 | 3.4 | 22.9 | 3.8 | 3.4 |
| 2 | 577 | 18.2 | 2.8 | 19.2 | 3.6 | 2.6 |
| 3 | 900 | 15.9 | 2.4 | 17.4 | 3.1 | 1.8 |
| 4 | 1295 | 13.3 | 3.5 | 17.8 | 2.5 | 1.5 |
| 5 | 1881 | 12.3 | 3.2 | 17.5 | 2.1 | 1.4 |
| 6 | 2650 | 10.9 | 3.9 | 17.3 | 2.0 | 1.2 |
| 7 | 3766 | 9.0 | 4.9 | 20.3 | 1.4 | 1.0 |
| 8 | 5426 | 8.4 | 5.6 | 21.8 | 0.9 | 0.8 |
| 9 | 7779 | 8.1 | 6.6 | 23.2 | 0.6 | 0.7 |
| 10 | 11042 | 7.3 | 7.6 | 27.0 | 0.5 | 0.5 |

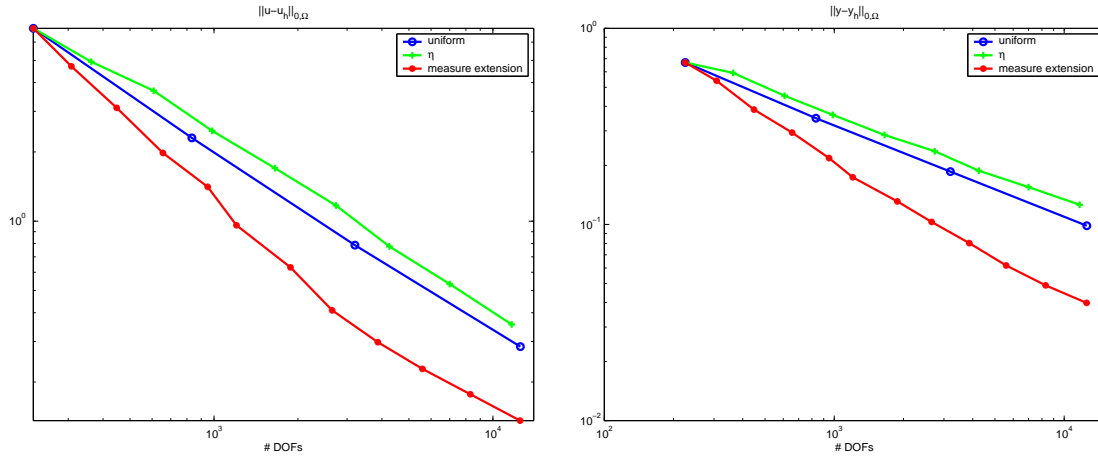


Figure 6.7: Example 3: Comparison of the approximation error in the control (left) and the state (right) for the two adaptive refinement strategies and a uniform refinement.

state y^d has values which are impossible to reach for the state. In particular, y^d has values up to 100 in the active region, while the upper bound ψ enforces the state to stay below the value 0.

Chapter 7

Conclusion and Outlook

In this thesis, the theory and the application of the powerful adaptive finite element method was extended to state constrained elliptic optimal control problems. In particular, two reliable and efficient residual-type a posteriori error estimators η and $\tilde{\eta}$ have been developed, analyzed, and implemented.

The first one of these error estimators was also transferred to the Lavrentiev regularization of the pure state constraints. Two numerical examples show the benefit of the adaptive refinement strategy compared to an uniform refinement. The Lavrentiev regularization provided a discrete solution which was comparable to the unregularized discrete solution for sufficiently small regularization parameters. However, no computational benefit for the mixed control-state constrained case was noticeable. The computational performance was, due to the larger involved linear systems which have to be solved, even a bit worse. In addition to that, more memory is needed for the representation of the involved matrices. Since furthermore the choice of a sufficiently small value for the regularization parameter does not seem to be an easy task, it seems to be better to use the unregularized algorithm. No matter how small a regularization parameter is, one probably always may construct an example for which it is still too large. On the other hand, if one considers a state constrained optimal control problem which contains an upper bound which is not continuous, the Lavrentiev regularization provides useful optimality conditions, whereas the optimality system for the pure state constrained case demands the continuity of the upper bound.

The error estimator $\tilde{\eta}$, resulting from a measure extension, has not such nice reliability and efficiency results since the only semi-continuous auxiliary adjoint state is involved in these estimates. However, we did not find an example for which it works worse than the error estimator η and, as the above stated Example 3 shows, there might be cases for which it even provides a better performance. On the other side, the error estimator η should work well in practice if the adjoint state and the modified adjoint state do not differ too much from each other. As Example 3 shows, a big difference may be noticeable if the desired state involves values which by far can not be reached by the state due to the upper bound. For practical applications,

this might probably be avoidable by a reformulation of the optimal control problem. Thus, both error estimators should work well in practice.

Of course, the gained a posteriori error estimates derived in this thesis are, due to the appearance of consistency errors, not as nice as those known for the control constrained case (cf. [HHIK07]). However, some admittedly heuristic arguments, like, for example, given in Remark 3.1.2, in Remark 6.2.1, or in Chapter 6.4, show that this error terms should not cause any big problems. These arguments are also supported by the numerical examples of this thesis. In my opinion, further a priori estimates for the consistency error does not make sense in an a posteriori error analysis which should give rise to an adaptive refinement strategy. For a pure a posteriori analysis, it is our belief that these consistency error terms can only be controlled if one applies methods known from free boundary value problems. In order to achieve a good approximation of the optimal solution of (2.1.3), this, however, might be an overkill.

Even though in the model problem of this thesis a simple linear partial differential equation was used, the derived results should be adjustable to more complicated, maybe nonlinear elliptic partial differential equations. Therefore, the knowledge about more general elliptic partial differential equations has to be adopted to the state equation and to the modified state equation or to the auxiliary adjoint state equation. This, of course, also should be investigated accurately. With minor modifications of the ideas of this thesis, it should also be possible to handle boundary optimal control problems with state constraints, like it has been done in [HIIS06] for control constrained boundary optimal control problems.

Since optimal control of partial differential equations has a lot of applications in practice, I hope that this thesis might be useful in order to achieve the solution of a lot of problems in science and technique in a more efficient way. Even though the computational power of our computers has increased a lot during the last years, efficiency is still an important goal for computational methods, as it is shown by the tremendous actual literature available for adaptive finite element methods. Especially the computational requirements of three dimensional partial differential equations are often too big. Even though we restricted ourselves to the two dimensional case, the extension to three dimensions should not cause any big difficulties.

Even though the error estimators of this thesis do, up to data oscillation and consistency error terms, neither under- nor overestimate the real error, the convergence of the corresponding adaptive finite element method still has to be proved, as it has been done in [GHIK07] for the control constrained case. However, due to the appearance of consistency error terms, this seems to be a challenging task, if this is possible at all.

Another very interesting field of research, where still a lot of investigations are necessary, is the optimal control of state constrained parabolic optimal control problems. The big challenge of the optimal control of parabolic partial differential equations

is that the optimality system is not evolutionary because the adjoint state equation is a backward partial differential equation. Thus, adaptive finite element methods known from parabolic partial differential equations, as they have been derived in [CJ04], for example, can, at least not directly, be applied to these kind of problems. Consequently, there are still a lot of open problems for parabolic optimal control problems. So far, there are, according to my knowledge, only reliability results available for control constrained parabolic optimal control problems (cf. [LMTY04], [LY03-2]). The goal-oriented weighted dual approach for these kind of problems was investigated in [MV07]. An additional difficulty for state constrained parabolic optimal control problems is the fact that the continuity of the optimal state can only be ensured for the one-dimensional case if there are no additional constraints on the control. Thus, it is even hard to derive useful optimality conditions. Then, the applications of mixed control-state constraints seems to be a good method to overcome these problems, as it has been shown in [NT07]. The investigation of parabolic partial differential equations and the optimal control of parabolic partial differential equations may be one of the challenging task for the next years.

Bibliography

- [Alt02] H.W. Alt (2002). *Lineare Funktionalanalysis*. Springer, Berlin-Heidelberg-New York, 4. Auflage.
- [BR78] I. Babuska and W. Rheinboldt (1978). *Error estimates for adaptive finite element computations*. SIAM J. Numer. Anal. 15, pp. 736–754.
- [BR03] W. Bangerth and R. Rannacher (2003). *Adaptive Finite Element Methods for Differential Equations*. Lectures in Mathematics. ETH-Zürich. Birkhäuser, Basel.
- [BKR00] R. Becker, H. Kapp, and R. Rannacher (2000). *Adaptive finite element methods for optimal control of partial differential equations: Basic concepts*. SIAM J. Control Optim. 39, pp. 113–132.
- [BHHK00] M. Bergounioux, M. Haddou, M. Hintermüller and K. Kunisch (2000). *A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problem*. SIAM Journal on Control and Optimization, 11, pp. 495-521.
- [BK02-2] M. Bergounioux and K. Kunisch (2002). *On the structure of Lagrange multipliers for state-constrained optimal control problem*. Systems and Control Letters, 48, pp. 169-176.
- [BK02] M. Bergounioux and K. Kunisch (2002). *Primal-dual strategy for state-constrained optimal control problem*. Computational Optimization and Applications, 22, pp. 193-224.
- [Bra03] D. Braess. (2003) *Finite Elemente*. Springer, Berlin-Heidelberg-New York, 2003, 3. Auflage.
- [BCH] D. Braess, C. Carstensen, R.H.W. Hoppe. *Convergence analysis of a conforming adaptive finite element method for an obstacle problem*. submitted.
- [CH04] C. Carstensen and R.H.W. Hoppe (2004). *Error reduction and convergence for an adaptive mixed finite element method*. Math. Comp. (in press).

- [CH05] C. Carstensen and R.H.W. Hoppe (2005). *Convergence analysis of an adaptive nonconforming finite element method*. J. Numer. Math. 13, pp. 19-32.
- [Cas93] E. Casas (1993). *Boundary control of semilinear elliptic equations with pointwise state constraints*. SIAM J. Control and Optimization, Vol. 31, No. 4, pp. 993-1006.
- [Cas86] E. Casas (1986). *Control of an elliptic problem with pointwise state constraints*. SIAM J. Control and Optimization, Vol. 24, No. 6, pp. 1309-1318.
- [CM] E. Casas and M. Mateos. *Error estimates for the numerical approximation of Neumann control problems*. to appear.
- [Cia91] P.G. Ciarlet (1991). *Basic error estimates for elliptic problems*. Handbook of numerical analysis, Vol. 2, P.G. Ciarlet and J.-L. Lions, eds. North-Holland, pp. 17-351.
- [CJ04] Z. Chen and F. Jia (2004). *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*. Math. Comp. 73, pp. 1167-1193.
- [DH06] K. Deckelnick and M. Hinze (2006). *Convergence of a finite element approximation to a state constrained elliptic control problem*. Priority Program 1253, German Research Foundation, Preprint-Number SPP1253-08-02.
- [Doe96] W. Dörfler (1996). *A convergent adaptive algorithm for Poisson's equation*. SIAM J. Numer. Anal. 33, pp. 1106-1124.
- [ET99] I. Ekeland and R. Témam (1999) *Convex analysis and variational problems*. SIAM, Philadelphia.
- [GHIK06] A. Gaevskaya, R.H.W. Hoppe, Y. Iliash and M. Kieweg. (2006). *A posteriori error analysis of control constrained distributed and boundary control problems*. Proc. Conf. Advances in Scientific Computing, Moscow, Russia (O. Pironneau et al.; eds.), Russian Academy of Sciences, Moscow.
- [GHIK07] A. Gaevskaya, R.H.W. Hoppe, Y. Iliash and M. Kieweg. (2007) *Convergence analysis of an adaptive finite element method for distributed elliptic optimal control problems with control constraints*. Proc. Conf. Optimal Control for PDEs, Oberwolfach, Germany (G.Leuring et al.; eds.), Springer, Berlin-Heidelberg-New York, pp. 308-326.

- [Gri85] P. Grisvard (1985). *Elliptic problems in nonsmooth domains*. Pitman, Boston.
- [GH07] A. Günther and M. Hinze (2007). *A posteriori error control of a state constrained elliptic control problem*. Preprint.
- [HH07] M. Hintermüller and R.H.W. Hoppe (2007). *Goal-oriented adaptivity in control constrained optimal control of partial differential equations*. Submitted to SIAM J. on Control and Optimization.
- [HHIK07] M. Hintermüller, R.H.W. Hoppe, J. Iliash and M. Kieweg (2007). *An a posteriori error analysis of adaptive finite element methods for distributed elliptic optimal control problems*. ESAIM, Control, Optimization and Calculus of Variations, Vol. 13, No. 4.
- [HHS06] R.H.W. Hoppe, Y. Iliash, C. Iyyunni and N.H. Sweilam. *A posteriori error estimates for adaptive finite element discretizations of boundary control problems*. J. Numer. Math., Vol. 14, No. 1, pp. 57-82.
- [LLMT02] R. Li, W. Liu, H. Ma and T. Tang (2002). *Adaptive finite element approximation for distributed elliptic optimal control problems*. SIAM Journal on Control and Optimization 41, pp. 1321-1349
- [Lio71] J.L. Lions (1971). *Optimal control of systems governed by partial differential equations*. Springer, Berlin-Heidelberg-New York.
- [LMTY04] W. Liu, H. Ma, T. Tang and N. Yan (2004). *A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations*. SIAM J. Numer. Anal., Vol. 42, No. 3, pp. 1032-1061.
- [LY03] W. Liu and N. Yan (2003). *A posteriori error estimates for convex boundary value problems*. Preprint, Institute of Mathematics and Statistics, University of Kent, Canterbury.
- [LY01] W. Liu and N. Yan (2001). *A posteriori error estimates for distributed optimal control problems*. Adv. Comp. Math. 15, pp. 285-309.
- [LY03-2] W. Liu and N. Yan (2003). *A posteriori error estimates for optimal control problems governed by parabolic equations*. Numer. Math. 93, pp. 497-521.
- [Lue84] D.G. Luenberger (1984). *Linear and nonlinear programming (second edition)*. Addison-Wesley Publishing Company, Reading, 1984.
- [MV07] D. Meidner and B. Vexler (2007). *Adaptive space-time finite element methods for parabolic optimization problems*. SIAM J. Control Optim., Vol. 46, No. 1, pp. 116-142.

- [MN07] K. Mekchay and R.H. Nochetto (2007). *Convergence of adaptive finite element methods for general second order linear elliptic pde.* to appear.
- [Mey06] C. Meyer (2006). *Optimal control of semilinear elliptic equations with applications to sublimation crystal growths.* Dissertation.
- [MPT06] C. Meyer, U. Prüfert, F. Tröltzsch. *On two numerical methods for state constrained elliptic optimal control problems.* Tech. Report 5-2005, Inst. of Math., TU Berlin. To appear in Optimization Methods and Software.
- [MRT06] C. Meyer, A. Rösch, F. Tröltzsch (2006). *Optimal control of PDEs with regularized pointwise state constraints.* Computational Optimization and Applications 33, pp. 209-228.
- [MT05] C. Meyer and F. Tröltzsch (2005). *On an elliptic optimal control problem with pointwise mixed control-state constraints.* Recent Advantages in Optimization. Proceedings of the 12th French-German-Spanish Conference on Optimization held in Avignon, September 20-24, 2004, A. Seeger et al., Lecture notes in economics and mathematical systems, Springer.
- [Roc70] R.T. Rockafellar (1970). *Convex Analysis.* Princeton University Press, Princeton, NJ.
- [MNS00] P. Morin, R.H. Nochetto, and K.G. Siebert (2000). *Data oscillation and convergence of adaptive FEM.* SIAM J. Numer. Anal., Vol. 38, No. 2, pp. 466-488.
- [NT07] I. Neitzel and F. Tröltzsch (2007). *On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints.* Deutsche Forschungsgemeinschaft, Priority Program 1253, Preprint-Number SPP1253-24-01.
- [NSV03] R.H. Nochetto, K.G. Siebert, A. Veese (2003). *Pointwise a posteriori error control for elliptic obstacle problems.* Numer. Math. 95, pp. 163-195.
- [Tro05] F. Tröltzsch (2005). *Optimale Steuerung partieller Differentialgleichungen.* Vieweg, Wiesbaden, 2005.
- [Tro05-2] F. Tröltzsch (2005). Regular Lagrange multipliers for control problems with mixed control-state constraints. SIAM J. on Optimization, Vol. 15, Issue 2, pp. 616-634.
- [Vee01] A. Veese (2001). *Efficient and reliable a posteriori error estimators for elliptic obstacle problems* SIAM J. Numer. Anal. 39, pp. 146-167.

- [Ver96] R. Verfürth (1996). *A review of a posteriori estimation and adaptive mesh-refinement techniques*. Wiley-Teubner, New York, Stuttgart.

Index

- Fréchet derivative, 99
- a posteriori, 23, 38, 43
 - error estimation, 54
 - error estimator, 27, 29
- a priori, 26, 29
- active set
 - continuous, 23
 - continuous regularized, 53
 - discrete, 105
 - regularized discrete, 54
- active set strategy, 43
- adaptive
 - algorithm, 41, 63, 110
 - refinement, 27, 68
- adjoint
 - control, 22, 24, 25, 43
 - functions, 22
 - state, 22, 24, 32, 43, 53
- application, 20
- approximation error, 41
- auxiliary
 - adjoint state, 94
 - modified adjoint state, 31, 56
 - state, 31, 56, 57, 94
- barycentric coordinates, 36
- benefit, 72, 76, 80, 83, 86, 90, 113
- bisection, 65
- boundary condition, 24
- bubble function, 36, 102
- bulk criterion, 64, 65, 69
- C++, 68
- Clément interpolation, 97
- coercivity, 19
- compact, 45, 49
- complementarity conditions, 23, 25, 26, 34, 47, 53, 54, 60
 - continuous, 23
 - discrete, 26
- conforming triangulation, 64
- conjugate, 22, 23
- consistency error, 29, 31, 32, 53, 56, 57, 68, 105
- constraints, 18, 25
- continuous
 - conditions, 23
 - optimality conditions, 21
- control, 18, 24, 29
- control constraints, 22
- control shift, 19
- control-to-state mapping, 19, 44
- convergence, 27, 29, 48, 51
 - weak**, 27, 29, 32
 - uniform, 50
 - weak, 20, 51
- convexity, 20
- data oscillation, 29–31, 36, 38, 56, 93, 94
 - higher order, 30, 56, 93
 - lower order, 30, 56, 93
- desired state, 19
- Dirac-measure, 25, 81
- direct solver, 68
- Dirichlet boundary, 18
- discrete
 - adjoint
 - control, 25
 - state, 25
 - auxiliary state, 31, 57
 - complementarity conditions, 26

- free boundary, 54, 64
 - modified
 - adjoint control, 26
 - adjoint state, 26
 - optimality conditions, 25
- discretization, 23, 52, 53
- dual, 18
- edge residual, 30, 31, 56, 93
- efficiency, 29, 36, 41, 60, 101
- eigenvalue, 45, 49
- eigenvector, 49
- element residual, 30, 31, 56, 93
- error estimator, 29, 55, 92
- existence and uniqueness, 20, 25, 44
- finite element
 - approximation, 23, 29
 - space, 24
- Fredholm operator, 45
- free boundary, 64
- Gâteaux derivative, 45
- Galerkin orthogonality, 36, 60, 98
- greedy algorithm, 65
- heat equation, 20
- Hilbert space, 17, 20, 51
- implementation, 24, 68
- inactive set
 - continuous, 23
 - continuous regularized, 53
 - discrete, 105
 - regularized discrete, 54
- indicator function, 22
- interior point method, 43
- interpolant, 24
- Karush-Kuhn-Tucker conditions, 23
- L^2 -Projection, 30
- Lagrange functional, 46, 47
- Lagrange multiplier, 22, 25, 43–47, 55, 69
- Lavrentiev regularization, 43
- Lebesgue
 - measure, 18
 - space, 17
- LINUX, 68
- lower semi-continuous, 20, 51
 - weakly, 20, 51
- measure extension, 92, 94
- mixed control-state constraints, 44, 60
- modified
 - adjoint
 - control, 23, 55
 - control, discrete, 26
 - state, 23, 29, 55
 - state, discrete, 26
 - function, 26
- Neumann boundary, 21
- nonlinearity, 21
- numerical example, 55, 68, 110
- objective functional, 19–21
- obstacle problem, 91
- optimal control problem, 18, 20
 - discrete, 24, 52
- optimal solution, 20, 45
- optimality conditions, 19, 21, 45
 - continuous, 21
 - discrete, 25
 - regularized, 47
- optimality system, 22, 25, 53
- orthonormal basis, 49
- Parseval equality, 50
- partial differential equation, 18, 20–22, 24, 31, 49
- Poincaré-Friedrich-Inequality, 32
- positive definite, 45, 49
- primal-dual active set strategy, 54, 63, 68, 76
- Radon measure, 18
- reflexive, 20, 51
- regular refinement, 65

- regularization, 43, 48
- regularization parameter, 43, 54, 55
- regularized constraints, 55
- regularized optimality conditions, 47
- reliability, 29, 32, 57, 95
- residual
 - edge, 30, 31, 56, 93
 - element, 30, 31, 56, 93
- residual-type error estimator, 29, 55, 56, 92
- Riesz representation, 46, 47

- self-adjoint, 49
- semidiscrete, 25
- shape regularity, 23, 24, 37, 64
- Slater Condition, 19, 21, 22, 25
- Sobolev space, 17
- sparse matrices, 68
- state, 18, 24, 29, 53
- state constraints, 19, 21, 43
- subdifferential, 22

- Tikhonov regularization, 19

- UMFPACK, 68
- uniform convergence, 50
- upper bound, 19, 21, 22, 25, 53

- weak limit, 51
- weak solution, 18

L E B E N S L A U F

Name: Michael Kieweg
Geboren: 27. Januar 1982 in Schwabmünchen
Staatsangehörigkeit: deutsch

Ausbildung:

09/1988 - 08/1992 Sankt-Ulrich-Grundschule Schwabmünchen
09/1992 - 06/2001 Leonhard-Wagner-Gymnasium Schwabmünchen;
Allgemeine Hochschulreife
09/2001 - 06/2002 Zivildienst im Krankenhaus Schwabmünchen
10/2002 Beginn des Mathematikstudiums mit Nebenfach
Informatik an der Universität Augsburg
04/2004 Vordiplom in Mathematik
10/2004 Wechsel in den Studiengang 'TopMath' der Universität
Augsburg; Wissenschaftliche Betreuung innerhalb
dieses Studiengangs durch Prof. Dr. Ronald H. W. Hoppe
09/2005 Bachelor of Science in Mathematik
08/2006 - 05/2007 Auslandsaufenthalt an der University of Houston/USA
09/2007 Master of Science in Mathematik
12/2007 Verteidigung der vorliegenden Dissertation