Deutsche Forschungsgemeinschaft

Priority Programme 1962

An Augmented Lagrange Method for Elliptic State Constrained Optimal Control Problems

Veronika Karl, Daniel Wachsmuth

SPP 1962

Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-008

received on February 16, 2017

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

An augmented Lagrange method for elliptic state constrained optimal control problems *

Veronika Karl[†] Daniel Wachsmuth[‡]

February 16, 2017

Abstract

In the present work we apply an augmented Lagrange method to solve pointwise state constrained elliptic optimal control problems. We prove strong convergence of the primal variables as well as weak convergence of the adjoint states and weak-* convergence of the multipliers associated to the state constraint. In addition, numerical results are presented.

Keywords: optimal control, state constraints, augmented Lagrange method.

AMS subject classification: 49M20, 65K10, 90C30.

1 Introduction

This paper deals with the solution of a convex optimal control problem with an elliptic state equation and pointwise control and state constraints. The problem is given by

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

subject to

$$\begin{aligned} Ay &= u && \text{in } \Omega, \\ \partial_{\nu_A} y &= 0 && \text{on } \partial\Omega, \\ y &\le \psi && \text{in } \Omega, \\ u_a \le u &\le u_b && \text{in } \Omega. \end{aligned}$$

Here A denotes an elliptic operator of second-order. The setting of the optimal control problem will be made precise below in Section 2.1.

Under suitable constraint qualifications, first-order necessary conditions to problem (P) can be proven. In general, the Lagrange multiplier associated to the state constraint $y \le \psi$ is a measure in $C(\bar{\Omega})^*$, see, e.g., [6]. Because of this low regularity of the Lagrange multiplier, the numerical solution of state constrained optimal control problems is challenging. Thus, in recent years different approaches were studied to overcome this problem. These approaches have in common that the state constraint is relaxed in a suitable way. Let us mention Lavrentiev-regularization [14, 24] turning the control problem into problems with mixed control-state constraints. Penalization-based approaches were studied in [10, 13, 17], their combination with a

*This research was supported by the German Research Foundation (DFG) within the priority program "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization" (SPP 1962) under grant number Wa 3626/3-1.

[†]University of Würzburg, Institute of Mathematics, Emil-Fischer-Str. 30, 97074 Würzburg, Germany; veronika.karl@mathematik.uni-wuerzburg.de

[‡]University of Würzburg, Institute of Mathematics, Emil-Fischer-Str. 30, 97074 Würzburg, Germany; daniel.wachsmuth@mathematik.uni-wuerzburg.de

path-following strategy was investigated in [11, 12]. Both types of methods are obtained as special cases of the so-called virtual control regularization approach developed in [20, 21].

Augmented Lagrange methods are well-known in optimization. However, there is only a limited number of publications dedicated to the application of such methods to optimal control problems with state constraints. In [1, 2] the state equation is augmented but the inequality constraints on the state are still present in the augmented Lagrange sub-problem. In [3, 16] the case of finitely many state constraints of the type $\Lambda y \in K$ with Λ having finite-dimensional range is studied. The goal of the present paper is to analyze the classical augmented Lagrange method in the general setting of problems with state constraints: state constraints in $C(\bar{\Omega})$ (not in a – possibly finite-dimensional – Hilbert space) with multipliers in $C(\bar{\Omega})^*$.

Let us mention the recent contribution [19]. There, a modified augmented Lagrange method is investigated, which is in the spirit of recent developments for finite-dimensional optimization problems [5]. The modification allows for a simpler convergence analysis. In contrast to this work, we study an algorithm with the classical Lagrange multiplier update.

After collecting preliminary results in Section 2, we develop the augmented Lagrange method in Section 3. In order to guarantee the boundedness of generated multiplier approximations, we investigate a special multiplier update rule: the classical multiplier update is performed only if a certain measure of feasibility and violation of complementarity shows sufficient decrease, see Section 3.3. The convergence of the method is studied in Section 3.5. The main results of this section are boundedness of iterates (Lemma 3.10) and their convergence (Theorem 3.11) to the solution of the original problem. We demonstrate the performance of the method for selected problems in Section 4.

Notation. Throughout the article we will use the following notation. The inner product in $L^2(\Omega)$ is denoted by $(\cdot,\cdot)$. Duality pairings will be denoted by $\langle\cdot,\cdot\rangle$. The dual of $C(\bar{\Omega})$ is denoted by $\mathcal{M}(\bar{\Omega})$, which is the space of regular Borel measures on $\bar{\Omega}$.

2 Preliminary results

2.1 Setting of the control problem

Let $\Omega \subset \mathbb{R}^N$, $N = \{2, 3\}$ be a bounded domain with $C^{1,1}$-boundary Γ. Let Y denote the space $Y := H^1(\Omega) \cap C(\bar{\Omega})$, and set $U := L^2(\Omega)$. We want to solve the following state-constrained optimal control problem: Minimize

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

over all $(y,u) \in Y \times U$ subject to the elliptic equation

$$\begin{aligned} (Ay)(x) &= u(x) \quad \text{in } \Omega, \\ (\partial_{\nu_A} y)(x) &= 0 \quad \text{on } \Gamma, \end{aligned}$$

and subject to the pointwise state and control constraints

$$\begin{aligned} y(x) &\le \psi(x) \quad \text{in } \Omega, \\ u_a(x) \le u(x) &\le u_b(x) \quad \text{in } \Omega. \end{aligned}$$

In the sequel, we will work with the following set of standing assumptions.

Assumption 1. *1. The given data satisfy $y_d \in L^2(\Omega)$, $\alpha > 0$, $u_a, u_b \in L^2(\Omega)$ with $u_a \le u_b$, $\psi \in C(\bar{\Omega})$.*

2. The differential operator A is given by

$$(Ay)(x) := -\sum_{i,j=1}^{N} \partial_{x_j}(a_{ij}(x)\partial_{x_i} y(x)) + a_0(x)y(x)$$

with $a_{i,j} \in C^{0,1}(\bar\Omega)$ and $a_0 \in L^\infty(\Omega)$. The operator A is assumed to be strongly elliptic, i.e., there is $\delta > 0$ such that

$$\sum_{i,j=1}^N a_{ij}(x)\xi_i\xi_j \geq \delta|\xi|^2 \quad \forall \xi \in \mathbb{R}^N, \text{ a.e. on } \Omega.$$

Further $a_0(x) \geq 0$ a.e. in Ω and $a_0 \not\equiv 0$.

3. *The normal derivative $\partial_{\nu_A} y$ is given by*

$$\partial_{\nu_A} y = \sum_{i,j=1}^N a_{ij}(x)\partial_{x_i} y(x)\nu_j(x),$$

where ν denotes the outward unit normal vector on Γ.

A function $y \in H^1(\Omega)$ is called a weak solution of the state equation if it holds

$$\int_\Omega -\sum_{i,j=1}^N a_{ij}(x)\partial_{x_i} y(x)\partial_{x_j} v(x)\,\mathrm{d}x + \int_\Omega a_0(x)y(x)v(x)\,\mathrm{d}x = \int_\Omega u(x)v(x)\,\mathrm{d}x \quad \forall v \in H^1(\Omega).$$

Due to the assumptions above, for every $u \in L^2(\Omega)$ there exists a uniquely determined weak solution y of the state equation.

Theorem 2.1. *Let Ω be a bounded domain in $\mathbb{R}^N$, $N = 2, 3$ with $C^{1,1}$-boundary. Then, for every $(f, g) \in L^r(\Omega) \times L^s(\Gamma)$ with $r > N/2$, $s > N - 1$, the elliptic partial differential equation*

$$\begin{aligned} Ay &= f \quad \text{in } \Omega,\\ \partial_{\nu_A} y &= g \quad \text{on } \Gamma \end{aligned} \tag{1}$$

admits a unique weak solution $y \in H^1(\Omega) \cap C(\bar\Omega)$, and it holds

$$\|y\|_{H^1(\Omega)} + \|y\|_{C(\bar\Omega)} \leq c\left(\|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Gamma)}\right) \tag{2}$$

with $c > 0$ independent of u.

If in addition (f_n, g_n) are such that $f_n \rightharpoonup f$ in $L^r(\Omega)$ and $g_n \to g$ in $L^s(\Gamma)$ then the corresponding solutions (y_n) of (1) converge strongly in $H^1(\Omega)$ and $C(\bar\Omega)$ to the solution y of (1) to data (f, g).

Proof. The proof can be found in Casas [7, Theorem 3.1]. □

This result shows that the control-to-state mapping $S : u \mapsto y$ is continuous from $L^2(\Omega)$ to $H^1(\Omega) \cap C(\bar\Omega)$. In the following, we will use the feasible sets with respect to the state and control constraints denoted by

$$\begin{aligned} U_{ad} &= \{u \in L^2(\Omega) \mid u_a(x) \leq u(x) \leq u_b(x) \text{ a.e. in } \Omega\},\\ Y_{ad} &= \{y \in C(\bar\Omega) \mid y(x) \leq \psi(x)\ \forall x \in \Omega\}. \end{aligned}$$

The feasible set of the optimal control problem is denoted by

$$F_{ad} = \{(y, u) \in Y \times U \mid (y, u) \in Y_{ad} \times U_{ad},\ y = Su\}.$$

2.2 Existence of solutions

Under the standing assumptions, we can show existence and uniqueness of solutions.

Theorem 2.2. *Let Assumption 1 be satisfied. Assume that the feasible set F_{ad} is non-empty. Then, there exists a uniquely determined solution $(\bar y, \bar u)$ of (P).*

Proof. This can be proven by standard arguments, see, e.g., [15, Theorem 1.43]. □

2.3 Optimality conditions

Existence of Lagrange multipliers to state-constrained optimal control problems is not guaranteed without any regularity assumptions. In the sequel, we will work with the following Slater point condition.

Assumption 2. *We assume that there exists $\hat{u} \in U_{ad}$ and $\sigma > 0$ such that for $\hat{y} = S\hat{u}$ it holds*

$$\hat{y}(x) \leq \psi(x) - \sigma \quad \forall x \in \Omega.$$

Theorem 2.3. *Let $(\bar{y}, \bar{u})$ be a solution of the problem (P). Furthermore, let Assumption 2 be fulfilled. Then, there exists an adjoint state $\bar{p} \in W^{1,s}(\Omega)$, $s < N/(N-1)$, and a Lagrange multiplier $\bar{\mu} \in \mathcal{M}(\bar{\Omega})$ with $\bar{\mu} = \bar{\mu}|_\Omega + \bar{\mu}|_\Gamma$, such that the following optimality system*

$$\begin{aligned} A\bar{y} &= \bar{u} && \text{in } \Omega, \\ \partial_{\nu_A}\bar{y} &= 0 && \text{on } \Gamma, \end{aligned} \tag{3a}$$

$$\begin{aligned} A^*\bar{p} &= \bar{y} - y_d + \bar{\mu}_\Omega && \text{in } \Omega, \\ \partial_{\nu_A^*}\bar{p} &= \bar{\mu}_\Gamma && \text{on } \Gamma, \end{aligned} \tag{3b}$$

$$(\bar{p} + \alpha\bar{u}, u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}, \tag{3c}$$

$$\langle \bar{\mu}, \bar{y} - \psi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})} = 0, \quad \bar{\mu} \geq 0, \tag{3d}$$

is fulfilled. Here, the inequality $\bar{\mu} \geq 0$ means $\langle \bar{\mu}, \varphi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})} \geq 0$ for all $\varphi \in C(\bar{\Omega})$ with $\varphi \geq 0$.

Proof. The proof can be found in [6]. □

Let us state a result about existence and regularity of solutions of the adjoint equation with measures.

Theorem 2.4. *Let $\bar{\mu} \in \mathcal{M}(\bar{\Omega})$ be a regular Borel measure with $\bar{\mu} = \bar{\mu}|_\Omega + \bar{\mu}|_\Gamma$. Then the adjoint state equation*

$$\begin{aligned} A^*\bar{p} &= \bar{y} - y_d + \bar{\mu}_\Omega && \text{in } \Omega, \\ \partial_{\nu_{A^*}}\bar{p} &= \bar{\mu}_\Gamma && \text{on } \Gamma \end{aligned}$$

has a unique very weak solution $\bar{p} \in W^{1,s}(\Omega)$, $s \in (1, N/(N-1))$, and it holds

$$\|\bar{p}\|_{W^{1,s}(\Omega)} \leq c \left(\|\bar{y}\|_{L^2(\Omega)} + \|y_d\|_{L^2(\Omega)} + \|\bar{\mu}\|_{\mathcal{M}(\bar{\Omega})} \right). \tag{4}$$

Proof. This result is due to [7, Theorem 4.3]. □

In general, the adjoint state $\bar{p}$ and the Lagrange multiplier $\bar{\mu}$ from Theorem 2.3 need not to be unique.

3 The augmented Lagrange method

Since the Lagrange multiplier corresponding to the pointwise state constraint is only a measure, its numerical treatment causes difficulties. To overcome these, we replace the inequality constraint $y \leq \psi$ by an augmented penalization term. The precise formulation of this augmentation will be described in the next section. The complete algorithm will be presented in Section 3.4 below.

3.1 The augmented Lagrange optimal control problem

Let $\rho > 0$ be a given penalty parameter, and let $\mu \in L^2(\Omega)$ with $\mu \geq 0$ be a given approximation of the Lagrange multiplier. In the sequel, we will work with the following augmented Lagrange functional P depending on these parameters defined by

$$P(y, \rho, \mu) := \frac{1}{2\rho} \int_\Omega \left((\mu + \rho(y - \psi))_+\right)^2 - \mu^2 \, \mathrm{d}x, \tag{5}$$

where $(z)_+ := \max(0, z)$. For given penalty parameter $\rho > 0$, the mapping $(y, \mu) \to P(y, \rho, \mu)$ is well-defined from $L^2(\Omega) \times L^2(\Omega) \to \mathbb{R}$. Moreover, it is continuously Frechet-differentiable from $L^2(\Omega) \times L^2(\Omega) \to \mathbb{R}$. Clearly, P is not twice Frechet differentiable. In the literature, there exist twice continuously differentiable penalty functions. We refer to Birgin [4] for a numerical comparison of various penalty functions for finite-dimensional optimization problems.

Let now $\rho > 0$ and $\mu \in L^2(\Omega)$ be given. Then in each step of the augmented Lagrange method the following sub-problem has to be solved: Minimize

$$J(y_\rho, u_\rho, \mu, \rho) := \frac{1}{2}||y_\rho - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u_\rho||^2_{L^2(\Omega)} + P(y_\rho, \rho, \mu) \tag{$P_{AL}^{\rho,\mu}$}$$

subject to the state equation and the control constraints

$$y_\rho = S u_\rho, \quad u_\rho \in U_{ad}.$$

In the following, existence of an optimal control and existence of an corresponding adjoint state will be proven. We start with a theorem that states existence of an optimal control.

Theorem 3.1 (Existence of solutions of the augmented Lagrange sub-problem). *For every $\rho > 0$, $\mu \in L^2(\Omega)$ with $\mu \geq 0$ the augmented Lagrange control problem ($P_{AL}^{\rho,\mu}$) admits a unique solution $\bar{u}_\rho \in U_{ad}$ with associated optimal state $\bar{y}_\rho \in Y$.*

Proof. Since U_{ad} is closed, bounded and convex and J is coercive, weakly lower semi-continuous and strictly convex, problem ($P_{AL}^{\rho,\mu}$) has a unique solution $\bar{u}_\rho \in U_{ad}$. For more details see [26] and [8]. □

Since the problem ($P_{AL}^{\rho,\mu}$) has no state constraints, the first-order optimality system is fulfilled without any further regularity assumptions.

Theorem 3.2 (First-order necessary optimality conditions). *Let $(\bar{y}_\rho, \bar{u}_\rho)$ be the solution of ($P_{AL}^{\rho,\mu}$). Then, there exists a unique adjoint state $\bar{p}_\rho \in H^1(\Omega)$ associated with the optimal control $\bar{u}_\rho$, satisfying the following system.*

$$\begin{aligned} A\bar{y}_\rho &= \bar{u}_\rho \quad \text{in } \Omega, \\ \partial_{\nu_A}\bar{y}_\rho &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{6a}$$

$$\begin{aligned} A^*\bar{p}_\rho &= \bar{y}_\rho - y_d + \bar{\mu}_\rho \quad \text{in } \Omega, \\ \partial_{\nu_A^*}\bar{p}_\rho &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{6b}$$

$$(\bar{p}_\rho + \alpha\bar{u}_\rho, u - \bar{u}_\rho) \geq 0 \quad \forall u \in U_{ad}, \tag{6c}$$

$$\bar{\mu}_\rho := (\mu + \rho(\bar{y}_\rho - \psi))_+ . \tag{6d}$$

Proof. Can be found in [15, Corollary 1.3, p. 73]. □

Due to the choice of $\bar{\mu}_\rho$ in (6d), the optimality system (6) of the augmented problem is very similar to optimality system (3) of the original problem (P). In fact, if $(\bar{y}_\rho, \bar{u}_\rho, \bar{p}_\rho, \bar{\mu}_\rho)$ solves (6), $\bar{y}_\rho$ is feasible, and $(\bar{\mu}_\rho, \bar{y}_\rho - \psi) = 0$ holds, then $(\bar{y}_\rho, \bar{u}_\rho, \bar{p}_\rho, \bar{\mu}_\rho)$ is a KKT-point of the original problem.

Another observation is that it is enough to control the L^1-norm of $\bar{\mu}_\rho$ in order to derive bounds on the solution $(\bar{y}_\rho, \bar{u}_\rho, \bar{p}_\rho)$ of (6). Here, we have the following theorem.

Theorem 3.3. *Let $\rho > 0$ and $\mu \in L^2(\Omega)$ be given. Let $s \in (1, N/(N-1))$. Then there is a constant $c > 0$ independent of ρ and μ such that for all solutions $(\bar{y}_\rho, \bar{u}_\rho, \bar{p}_\rho, \bar{\mu}_\rho)$ of (6) it holds*

$$\|\bar{y}_\rho\|_{H^1(\Omega)} + \|\bar{y}_\rho\|_{C(\bar{\Omega})} + \|\bar{u}_\rho\|_{L^2(\Omega)} + \|\bar{p}_\rho\|_{W^{1,s}(\Omega)} \leq c(\|\bar{\mu}_\rho\|_{L^1(\Omega)} + 1).$$

Proof. Let us test the state equation (6a) with $\bar{p}_\rho$ and the adjoint equation (6b) with $\bar{y}_\rho$. This yields

$$(\bar{p}_\rho, \bar{u}_\rho) = (\bar{y}_\rho - y_d, \bar{y}_\rho) + (\bar{\mu}_\rho, \bar{y}_\rho).$$

Using the optimal control $\bar{u}$ of the original problem as test function in (6c), we obtain

$$(\bar{y}_\rho - y_d, \bar{y}_\rho) + (\bar{\mu}_\rho, \bar{y}_\rho) \leq (\alpha \bar{u}_\rho, \bar{u} - \bar{u}_\rho) + (\bar{p}_\rho, \bar{u}).$$

By Young's inequality, we have

$$\|\bar{y}_\rho\|^2_{L^2(\Omega)} + \alpha \|\bar{u}_\rho\|^2_{L^2(\Omega)} \leq 2\|\bar{\mu}_\rho\|_{L^1(\Omega)} \|\bar{y}_\rho\|_{C(\bar{\Omega})} + \alpha \|\bar{u}\|^2_{L^2(\Omega)} + \|y_d\|^2_{L^2(\Omega)} + 2\|\bar{p}_\rho\|_{L^2(\Omega)} \|\bar{u}\|_{L^2(\Omega)}.$$

Let us fix $\bar{s} \in (1, N/(N-1))$ such that $W^{1,\bar{s}}(\Omega)$ is continuously embedded in $L^2(\Omega)$. Then we obtain from Theorems 2.1 and 2.4 that there exists $c > 0$, which is independent of ρ and μ, such that

$$\|\bar{y}_\rho\|^2_{L^2(\Omega)} + \alpha \|\bar{u}_\rho\|^2_{L^2(\Omega)} \leq c \left(\|\bar{\mu}_\rho\|_{L^1(\Omega)} \|\bar{u}_\rho\|_{L^2(\Omega)} + \|\bar{u}\|^2_{L^2(\Omega)} + \|y_d\|^2_{L^2(\Omega)} \right).$$

This implies the bound on the L^2-norms of $\bar{u}_\rho$ and $\bar{y}_\rho$. Using again the regularity results from Theorems 2.1 and 2.4 the claim is proven. □

3.2 The general augmented Lagrange algorithm

In this section we will briefly present a prototypical augmented Lagrange algorithm. In the following, let $(P_{AL})_k$ denote the augmented Lagrange sub-problem $(P^{\rho,\mu}_{AL})$ for given penalty parameter $\rho := \rho_k$ and multiplier $\mu := \mu_k$. We will denote its solution by $(\bar{y}_k, \bar{u}_k)$ with adjoint state $\bar{p}_k$ and updated multiplier $\bar{\mu}_k$, which is given by (6d).

Algorithm 1. *Let $\rho_1 > 0$ and $\mu_1 \in L^2(\Omega)$ be given with $\mu_1 \geq 0$. Choose $\theta > 1$.*

1. *Solve $(P_{AL})_k$, and obtain $(\bar{y}_k, \bar{u}_k, \bar{p}_k, \bar{\mu}_k)$.*
2. *If the step is successful set $\mu_{k+1} := \bar{\mu}_k$, $\rho_{k+1} := \rho_k$.*
3. *Otherwise set $\mu_{k+1} := \mu_k$, increase penalty parameter $\rho_{k+1} := \theta \rho_k$.*
4. *If the stopping criterion is not satisfied set $k := k + 1$ and go to step 1.*

We will describe the decision about successful steps in the next section. The stopping criterion will be based on the same quantities. It ensures that the approximations are close to solutions of the original problem. For convenience, let us restate the system $(P_{AL})_k$ that is solved by $(\bar{y}_k, \bar{u}_k, \bar{p}_k, \bar{\mu}_k)$:

$$\begin{aligned} A\bar{y}_k &= \bar{u}_k \quad \text{in } \Omega, \\ \partial_{\nu_A} \bar{y}_k &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{7a}$$

$$\begin{aligned} A^* \bar{p}_k &= \bar{y}_k - y_d + \bar{\mu}_k \quad \text{in } \Omega, \\ \partial_{\nu_A^*} \bar{p}_k &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{7b}$$

$$\bar{u}_k \in U_{ad}, \quad (\bar{p}_k + \alpha \bar{u}_k, u - \bar{u}_k) \geq 0 \quad \forall u \in U_{ad}, \tag{7c}$$

$$\bar{\mu}_k := (\mu_k + \rho_k(\bar{y}_k - \psi))_+ . \tag{7d}$$

3.3 The multiplier update rule

Let us start this section with a basic estimate, which will be useful in the sequel.

Lemma 3.4. *Let $(\bar y, \bar u, \bar p, \bar\mu)$ be a solution of (3), and let $(\bar y_k, \bar u_k, \bar p_k, \bar\mu_k)$ solve (7). Then it holds*

$$\|\bar y - \bar y_k\|^2_{L^2(\Omega)} + \alpha \|\bar u - \bar u_k\|^2_{L^2(\Omega)} \le (\bar\mu_k, \psi - \bar y_k) + \langle \bar\mu, \bar y_k - \psi\rangle. \tag{8}$$

Proof. Using (3b) and (7b), we obtain

$$\begin{aligned}\|\bar y - \bar y_k\|^2_{L^2(\Omega)} &= (A^*(\bar p - \bar p_k), \bar y - \bar y_k) - (\bar\mu - \bar\mu_k, \bar y - \bar y_k)\\ &= ((\bar p - \bar p_k), \bar u - \bar u_k) - (\bar\mu - \bar\mu_k, \bar y - \bar y_k)\\ &\le -\alpha \|\bar u - \bar u_k\|^2_{L^2(\Omega)} - (\bar\mu - \bar\mu_k, \bar y - \bar y_k),\end{aligned}$$

which implies

$$\|\bar y - \bar y_k\|^2_{L^2(\Omega)} + \alpha \|\bar u - \bar u_k\|^2_{L^2(\Omega)} \le (\bar\mu_k - \bar\mu, \bar y - \bar y_k). \tag{9}$$

The term on the right-hand side of equation (9) can be split into two parts:

$$(\bar\mu_k, \bar y - \bar y_k) = (\bar\mu_k, \bar y - \psi) + (\bar\mu_k, \psi - \bar y_k) \le (\bar\mu_k, \psi - \bar y_k) \tag{10}$$

and

$$-\langle \bar\mu, \bar y - \bar y_k\rangle = -\langle \bar\mu, \bar y - \psi\rangle - \langle \bar\mu, \psi - \bar y_k\rangle = \langle \bar\mu, \bar y_k - \psi\rangle. \tag{11}$$

Here, we used the complementarity relation (3d) as well as $\bar y \le \psi$ and $\bar\mu_k \ge 0$. Putting the inequalities (9), (10), and (11) together, we get

$$\|\bar y - \bar y_k\|^2_{L^2(\Omega)} + \alpha \|\bar u - \bar u_k\|^2_{L^2(\Omega)} \le (\bar\mu_k, \psi - \bar y_k) + \langle \bar\mu, \bar y_k - \psi\rangle.$$

which is the claim. □

Our multiplier update decision is motivated by the following result, which estimates the difference of solutions of the augmented Lagrange sub-problem to the solution of the original problem. The upper bound of the error contains the violation of the state constraint and the mismatch in the complementarity condition.

Lemma 3.5. *Let $(\bar y, \bar u, \bar p, \bar\mu)$ and $(\bar y_k, \bar u_k, \bar p_k, \bar\mu_k)$ be given as in Lemma 3.4. Then it holds*

$$\|\bar y - \bar y_k\|^2_{L^2(\Omega)} + \alpha \|\bar u - \bar u_k\|^2_{L^2(\Omega)} \le \|\bar\mu\|_{\mathcal{M}(\Omega)} \|(\bar y_k - \psi)_+\|_{C(\bar\Omega)} + (\bar\mu_k, \psi - \bar y_k). \tag{12}$$

Proof. The claim follows directly from Lemma 3.4 using the estimate

$$\langle \bar\mu, \bar y_k - \psi\rangle \le \|\bar\mu\|_{\mathcal{M}(\Omega)} \|(\bar y_k - \psi)_+\|_{C(\bar\Omega)}.$$

□

This result shows that the iterates $(\bar y_k, \bar u_k)$ will converge to the solution of the original problem if the quantity

$$\|(\bar y_k - \psi)_+\|_{C(\bar\Omega)} + |(\bar\mu_k, \psi - \bar y_k)|$$

tends to zero for $k \to \infty$. We will say that a step of Algorithm 1 is successful if this quantity decreases sufficiently fast. In fact, we will say that step k was successful if the condition

$$\|(\bar y_k - \psi)_+\|_{C(\bar\Omega)} + |(\bar\mu_k, \psi - \bar y_k)| \le \tau \left(\|(\bar y_n - \psi)_+\|_{C(\bar\Omega)} + |(\bar\mu_n, \psi - \bar y_n)| \right)$$

is satisfied with $\tau \in (0,1)$. Here, we denoted by step n, $n < k$, the previous successful step. Moreover, the quantity above can be used as termination criterion, where the iteration is stopped if this quantity is small enough.

3.4 The augmented Lagrange algorithm in detail

Let us now state the concrete algorithm based on the general algorithm above with the update rule as described in the previous section.

Algorithm 2. *Let $\rho_1 > 0$ and $\mu_1 \in L^2(\Omega)$ be given with $\mu_1 \geq 0$. Choose $\theta > 1$, $\tau \in (0,1)$, $\epsilon \geq 0$, $R_0^+ >> 1$. Set $k = 1$ and $n = 1$.*

1. *Solve $(P_{AL})_k$, and obtain $(\bar{y}_k, \bar{u}_k, \bar{p}_k, \bar{\mu}_k)$.*
2. *Compute $R_k := \|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)|$.*
3. *If $R_k \leq \tau R_{n-1}^+$ then the step k is successful, set $\mu_{k+1} := \bar{\mu}_k$, $\rho_{k+1} := \rho_k$ and define $(y_n^+, u_n^+, p_n^+) := (\bar{y}_k, \bar{u}_k, \bar{p}_k)$, as well as $\mu_n^+ := \mu_{k+1}$ and $R_n^+ := R_k$. Set $n := n + 1$.*
4. *Otherwise the step k is not successful, set $\mu_{k+1} := \mu_k$, increase penalty parameter $\rho_{k+1} := \theta \rho_k$.*
5. *If $R_n^+ \leq \epsilon$ then stop, otherwise set $k := k + 1$ and go to step 1.*

The algorithm is well-defined. Still it needs to be proven that infinitely many steps are successful. Otherwise the algorithm would produce a sequence of iterates which becomes constant after finitely many steps.

In order to do so, we will investigate the solutions of the augmented Lagrange KKT system (6) with fixed multiplier approximation μ and penalization parameter ρ tending to infinity. In this situation, the method reduces to a penalty method with additional shift parameter μ. Such a scheme was already investigated in [11]. However, there a much stronger regularity condition was imposed, which forces to consider the state constraints in $H^2(\Omega)$.

Lemma 3.6. *Let (ρ_k) be a sequence of positive numbers with $\rho_k \to \infty$. Let $\mu \in L^2(\Omega)$ with $\mu \geq 0$ be given. Let $(\bar{y}_k, \bar{u}_k, \bar{p}_k)$ be solutions of*

$$\begin{aligned} A\bar{y}_k &= \bar{u}_k \quad \text{in } \Omega, \\ \partial_{\nu_A} \bar{y}_k &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{13a}$$

$$\begin{aligned} A^*\bar{p}_k &= \bar{y}_k - y_d + \bar{\mu}_k \quad \text{in } \Omega, \\ \partial_{\nu_A^*} \bar{p}_k &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{13b}$$

$$\bar{u}_k \in U_{ad}, \quad (\bar{p}_k + \alpha \bar{u}_k, u - \bar{u}_k) \geq 0 \quad \forall u \in U_{ad}, \tag{13c}$$

$$\bar{\mu}_k := (\mu + \rho_k(\bar{y}_k - \psi))_+ \,. \tag{13d}$$

Then it holds $(\bar{y}_k, \bar{u}_k) \to (\bar{y}, \bar{u})$ in $(H^1(\Omega) \cap C(\bar{\Omega})) \times L^2(\Omega)$ for $k \to \infty$.

Proof. The general idea of the proof follows [11]. Using an observation from the proof of [18, Theorem 3.1], we find

$$\begin{aligned} (\bar{\mu}_k, \bar{y} - \bar{y}_k) &= (\bar{\mu}_k, -\frac{\mu}{\rho_k} - \bar{y}_k + \psi - \psi + \bar{y} + \frac{\mu}{\rho_k}) \\ &= -\frac{1}{\rho_k} \|\bar{\mu}_k\|^2_{L^2(\Omega)} + \frac{1}{\rho_k}(\bar{\mu}_k, \mu) + (\bar{\mu}_k, \bar{y} - \psi) \\ &\leq -\frac{1}{\rho_k} \|\bar{\mu}_k\|^2_{L^2(\Omega)} + \frac{1}{2\rho_k} \left(\|\bar{\mu}_k\|^2_{L^2(\Omega)} + \|\mu\|^2_{L^2(\Omega)} \right) \\ &= -\frac{1}{2\rho_k} \|\bar{\mu}_k\|^2_{L^2(\Omega)} + \frac{1}{2\rho_k} \|\mu\|^2_{L^2(\Omega)} \,. \end{aligned} \tag{14}$$

From inequality (9) in the proof of Lemma 3.4, we get

$$\begin{aligned} \|\bar{y} - \bar{y}_k\|^2_{L^2(\Omega)} + \alpha \|\bar{u} - \bar{u}_k\|^2_{L^2(\Omega)} &\leq (\bar{\mu}_k, \bar{y} - \bar{y}_k) - \langle \bar{\mu}, \bar{y} - \bar{y}_k \rangle \\ &\leq (\bar{\mu}_k, \bar{y} - \bar{y}_k) + c \|\bar{\mu}\|_{\mathcal{M}(\bar{\Omega})} \|\bar{y} - \bar{y}_k\|_{C(\bar{\Omega})} \\ &\leq (\bar{\mu}_k, \bar{y} - \bar{y}_k) + \frac{\alpha}{2} \|\bar{u} - \bar{u}_k\|^2_{L^2(\Omega)} + \frac{c^2}{2\alpha} \|\bar{\mu}\|^2_{\mathcal{M}(\bar{\Omega})} \,, \end{aligned}$$

where we used Young's inequality and the regularity result from Theorem 2.1. With inequality (14) this leads to

$$\|\bar{y}-\bar{y}_k\|^2_{L^2(\Omega)}+\frac{\alpha}{2}\|\bar{u}-\bar{u}_k\|^2_{L^2(\Omega)}+\frac{1}{2\rho_k}\|\bar{\mu}_k\|^2_{L^2(\Omega)}\le\frac{1}{2\rho_k}\|\mu\|^2_{L^2(\Omega)}+\frac{c^2}{2\alpha}\|\bar{\mu}\|^2_{\mathcal{M}(\bar{\Omega})}. \tag{15}$$

Hence, the sequence $(\bar{u}_k)$ is bounded in $L^2(\Omega)$, implying the boundedness of $(\bar{y}_k)$ in $H^1(\Omega)\cap C(\bar{\Omega})$. This allows to extract weakly converging subsequences $\bar{u}_{k'}\rightharpoonup u^*$ in $L^2(\Omega)$ and $\bar{y}_{k'}\rightharpoonup y^*$ in $H^1(\Omega)$. Since the embedding $H^1(\Omega)\cap C(\bar{\Omega})\hookrightarrow L^2(\Omega)$ is compact, the sequence $(\bar{y}_k)$ converges strongly in $L^2(\Omega)$. By Theorem 2.1, the convergence $\bar{y}_{k'}$ to y^* is strong in $C(\bar{\Omega})$. In order to prove $y^*\le\psi$, we use the identity

$$\frac{1}{\rho_k}\|\bar{\mu}_k\|^2_{L^2(\Omega)}=\rho_k\left\|\max\left(0,\frac{\mu}{\rho_k}+\bar{y}_k-\psi\right)\right\|^2_{L^2(\Omega)}, \tag{16}$$

which is bounded because of (15). As $\max\left(0,\frac{\mu}{\rho_{k'}}+\bar{y}_{k'}-\psi\right)$ converges to $\max(0,y^*-\psi)$ in $L^2(\Omega)$ for $k'\to\infty$, we obtain $y^*\le\psi$ by passing to the limit in (16). This shows that y^* is feasible. To argue that $y^*=\bar{y}$ and $u^*=\bar{u}$, we use again inequality (9) to conclude

$$\begin{aligned}\|\bar{y}-\bar{y}_k\|^2_{L^2(\Omega)}+\alpha\|\bar{u}-\bar{u}_k\|^2_{L^2(\Omega)}&=(\bar{\mu}_k,\bar{y}-\bar{y}_k)-\langle\bar{\mu},\bar{y}-\psi\rangle+\langle\bar{\mu},\bar{y}_k-\psi\rangle\\&\le\frac{1}{2\rho_k}\|\mu\|^2_{L^2(\Omega)}+\langle\bar{\mu},\bar{y}_k-\psi\rangle.\end{aligned} \tag{17}$$

Passing to the limit $k'\to\infty$ yields

$$0\le\lim_{k'\to\infty}\|\bar{y}-\bar{y}_{k'}\|^2_{L^2(\Omega)}+\alpha\|\bar{u}-\bar{u}_{k'}\|^2_{L^2(\Omega)}\le\langle\bar{\mu},y^*-\psi\rangle\le 0,$$

and consequently $\bar{u}_{k'}\to\bar{u}$ in $L^2(\Omega)$. Because of Theorem 2.1 we immediately get the strong convergence $\bar{y}_{k'}\to\bar{y}$ in $H^1(\Omega)\cap C(\bar{\Omega})$. As the limit is independent of the taken subsequence, we obtain convergence of the whole sequences (u_k) and (y_k) to $\bar{u}$ and $\bar{y}$, respectively. □

Lemma 3.7. *Under the same assumptions as in Lemma 3.6, it holds*

$$\lim_{k\to\infty}(\bar{\mu}_k,\psi-\bar{y}_k)=0.$$

Proof. First we estimate

$$\begin{aligned}(\bar{\mu}_k,\psi-\bar{y}_k)&=\frac{1}{\rho_k}(\bar{\mu}_k,-\mu+\rho_k(\psi-\bar{y}_k)+\mu)=-\frac{1}{\rho_k}\|\bar{\mu}_k\|^2_{L^2(\Omega)}+\frac{1}{\rho_k}(\bar{\mu}_k,\mu)\\&\le-\frac{1}{\rho_k}\|\bar{\mu}_k\|^2_{L^2(\Omega)}+\frac{1}{2\rho_k}\|\bar{\mu}_k\|^2_{L^2(\Omega)}+\frac{1}{2\rho_k}\|\mu\|^2_{L^2(\Omega)}\\&\le\frac{1}{2\rho_k}\|\mu\|^2_{L^2(\Omega)},\end{aligned}$$

which proves

$$\limsup_{k\to\infty}(\bar{\mu}_k,\psi-\bar{y}_k)\le 0. \tag{18}$$

From Lemma 3.4 we get

$$(\bar{\mu}_k,\psi-\bar{y}_k)\ge\|\bar{y}-\bar{y}_k\|^2_{L^2(\Omega)}+\alpha\|\bar{u}-\bar{u}_k\|^2_{L^2(\Omega)}+\langle\bar{\mu},\psi-\bar{y}_k\rangle,$$

which leads with Lemma 3.6 to

$$\liminf_{k\to\infty}(\bar{\mu}_k,\psi-\bar{y}_k)\ge 0. \tag{19}$$

The inequalities (18) and (19) yield the claim. □

Using these two results, we can show that an infinite number of successful steps are done.

Lemma 3.8. *The augmented Lagrange algorithm makes infinitely many successful steps.*

Proof. We assume the algorithm to do a finite number of successful steps only. Then there is an index m such that all steps k with $k > m$ are not successful. According to Algorithm 2 it holds $\mu_k = \mu_m$ for all $k > m$, $R_k > \tau R_m$ and $\rho_k \to \infty$. However by Lemma 3.6 and Lemma 3.7 we get

$$\lim_{k\to\infty} R_k = \lim_{k\to\infty} \|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| = 0,$$

yielding a contradiction. □

3.5 Convergence of the algorithm

Let us recall that the sequence (y_n^+, u_n^+, p_n^+) denotes the solution of the n-th successful iteration of Algorithm 2 with μ_n^+ being the corresponding approximation of the Lagrange multiplier. Here, we want to show convergence of the algorithm. The most important part is proving L^1-boundedness of the Lagrange multipliers μ_n^+, which is accomplished in Lemma 3.10 below.

Lemma 3.9. *Let y_n^+, μ_n^+ be given as defined in Algorithm 2. Then it holds*

$$|(\mu_n^+, \psi - y_n^+)| \le \tau^{n-1} \left(\left\|(y_1^+ - \psi)_+\right\|_{C(\bar{\Omega})} + \left\|\mu_1^+\right\|_{L^2(\Omega)} \left\|(\psi - y_1^+)_+\right\|_{L^2(\Omega)} \right). \tag{20}$$

Proof. By definition of a successful step in Algorithm 2, we get the result directly by induction and the Cauchy-Schwarz inequality. □

Let us now show the L^1-boundedness of the Lagrange multipliers (μ_n^+).

Lemma 3.10 (Boundedness of the Lagrange multiplier). *Let Assumption 2 be fulfilled. Then Algorithm 2 generates an infinite sequence of bounded iterates, i.e., there is a constant $C > 0$ such that for all n it holds*

$$\left\|y_n^+\right\|_{H^1(\Omega)} + \left\|y_n^+\right\|_{C(\bar{\Omega})} + \left\|u_n^+\right\|_{L^2(\Omega)} + \left\|p_n^+\right\|_{W^{1,s}(\Omega)} + \left\|\mu_n^+\right\|_{L^1(\Omega)} \le C.$$

Proof. Let $(\hat{y}, \hat{u})$ be the Slater point given by Assumption 2, i.e., there exists $\sigma > 0$, such that $\hat{y} + \sigma \le \psi$. Then we can estimate

$$\begin{aligned} \sigma\|\mu_n^+\|_{L^1(\Omega)} &= \int_\Omega \sigma\mu_n^+ \,\mathrm{d}x \le \int_\Omega \mu_n^+(\psi - \hat{y}) \,\mathrm{d}x = \int_\Omega \mu_n^+(\psi - y_n^+ + y_n^+ - \hat{y}) \,\mathrm{d}x \\ &= \underbrace{\int_\Omega \mu_n^+(\psi - y_n^+) \,\mathrm{d}x}_{(I)} + \underbrace{\int_\Omega \mu_n^+(y_n^+ - \hat{y}) \,\mathrm{d}x}_{(II)}. \end{aligned}$$

The first part (I) can be estimated with Lemma 3.9 yielding

$$\begin{aligned} (I) \le |(\mu_n^+, \psi - y_n^+)| &\le \tau^{n-1} \left(\left\|(y_1^+ - \psi)_+\right\|_{C(\bar{\Omega})} + \left\|\mu_1^+\right\|_{L^2(\Omega)} \left\|(\psi - y_1^+)_+\right\|_{L^2(\Omega)} \right) \\ &=: \tau^{n-1} C. \end{aligned} \tag{21}$$

The second part (II) can be estimated using Young's Inequality as follows

$$\begin{aligned} &\int_\Omega \mu_n^+(y_n^+ - \hat{y}) \,\mathrm{d}x = \langle A^* p_n^+ - (y_n^+ - y_d), y_n^+ - \hat{y}\rangle \\ &\quad = \langle p_n^+, A(y_n^+ - \hat{y})\rangle - (y_n^+ - y_d, y_n^+ - \hat{y}) = (p_n^+, u_n^+ - \hat{u}) - (y_n^+ - y_d, y_n^+ - \hat{y}) \\ &\quad \le -(\alpha u_n^+, u_n^+ - \hat{u}) - (y_n^+ - y_d, y_n^+ - \hat{y}) = (\alpha u_n^+, \hat{u} - u_n^+) + (y_n^+ - y_d, \hat{y} - y_n^+) \\ &\quad = \alpha(u_n^+ - \hat{u}, \hat{u} - u_n^+) + \alpha(\hat{u}, \hat{u} - u_n^+) + (y_n^+ - \hat{y}, \hat{y} - y_n^+) + (\hat{y} - y_d, \hat{y} - y_n^+) \\ &\quad \le -\frac{\alpha}{2} \left\|\hat{u} - u_n^+\right\|_{L^2(\Omega)}^2 - \frac{1}{2} \left\|\hat{y} - y_n^+\right\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|\hat{u}\|_{L^2(\Omega)}^2 + \frac{1}{2} \|\hat{y} - y_d\|_{L^2(\Omega)}^2. \end{aligned} \tag{22}$$

Putting (21) and (22) together yields

$$\left\|\mu_n^+\right\|_{L^1(\Omega)} + \frac{\alpha}{2}\left\|\hat{u} - u_n^+\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\left\|\hat{y} - y_n^+\right\|_{L^2(\Omega)}^2 \leq \frac{\tau^{n-1}}{\sigma}C + \frac{\alpha}{2}\left\|\hat{u}\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\left\|\hat{y} - y_d\right\|_{L^2(\Omega)}^2.$$

Since $\tau \in (0,1)$ by assumption, the right-hand side is bounded. Consequently we get boundedness of (u_n^+) in $L^2(\Omega)$ and boundedness of (μ_n^+) in $L^1(\Omega)$ By the regularity result Theorem 2.1, the sequence (y_n^+) is uniformly bounded in $H^1(\Omega)\cap C(\bar{\Omega})$. Boundedness of (p_n^+) follows directly from Theorem 3.3. □

Remark 1. *Let us note that the proof of the previous Lemma 3.10 yields boundedness of (u_n^+) without using boundedness of the admissible set U_{ad}.*

Theorem 3.11 (**Convergence of solutions of the augmented Lagrange algorithm**). *As $n \to \infty$ we have for the sequence (y_n^+, u_n^+) generated by Algorithm 2*

$$(y_n^+, u_n^+) \to (\bar{y}, \bar{u}), \text{ in } (H^1(\Omega)\cap C(\bar{\Omega})) \times L^2(\Omega).$$

Proof. Since the algorithm yields an infinite number of successful steps (Lemma 3.8) we get

$$\lim_{n\to\infty} R_n^+ = \lim_{n\to\infty} \left\|(y_n^+ - \psi)_+\right\|_{C(\bar{\Omega})} + |(\mu_n^+, \psi - y_n^+)| = 0. \tag{23}$$

From Lemma 3.4 we get the following inequality

$$\begin{aligned}\left\|\bar{y} - y_n^+\right\|_{L^2(\Omega)}^2 + \alpha\left\|\bar{u} - u_n^+\right\|_{L^2(\Omega)}^2 &\leq \langle\bar{\mu}, y_n^+ - \psi\rangle + |(\mu_n^+, \psi - y_n^+)| \\ &\leq \|\bar{\mu}\|_{\mathcal{M}(\bar{\Omega})}\left\|(y_n^+ - \psi)_+\right\|_{C(\bar{\Omega})} + |(\mu_n^+, \psi - y_n^+)|.\end{aligned}$$

With (23) from above, we conclude

$$0 \leq \lim_{n\to\infty} \left\|\bar{y} - y_n^+\right\|_{L^2(\Omega)}^2 + \alpha\left\|\bar{u} - u_n^+\right\|_{L^2(\Omega)}^2 \leq 0.$$

yielding $y_n^+ \to \bar{y}$ in $L^2(\Omega)$ and $u_n^+ \to \bar{u}$ in $L^2(\Omega)$. In addition, we get strong convergence of $y_n^+ \to \bar{y}$ in $H^1(\Omega)\cap C(\bar{\Omega})$ by Theorem 2.1. □

The next step in the convergence analysis is to show the convergence of the dual quantities (μ_n^+) and (p_n^+) to multipliers and adjoint states of the original problem (P). Since these sequences are bounded in $L^1(\Omega)$ and $W^{1,s}(\Omega)$, $s \in (1, \frac{N}{N-1})$, we can extract weak-star and weakly converging subsequences. These weak subsequential limits are indeed Lagrange multipliers for the original problem.

Theorem 3.12 (**Subsequential convergence of dual quantities**). *Let subsequences $(p_{n_j}^+, \mu_{n_j}^+)$ of (p_n^+, μ_n^+) be given such that $\mu_{n_j}^+ \rightharpoonup \bar{\mu}$ in $\mathcal{M}(\bar{\Omega})$ and $p_{n_j}^+ \rightharpoonup \bar{p}$ in $W^{1,s}(\Omega)$, $s \in (1, \frac{N}{N-1})$.*

Then $(\bar{y}, \bar{u}, \bar{p}, \bar{\mu})$ satisfies the optimality system (3) of the original problem (P).

Proof. The proof that the limits satisfy the adjoint equation (3b) can be found in [14, Lemma 2.6]. It remains to prove that the weak-* limit of $\mu_{n_j}^+$ is indeed a Lagrange multiplier. First, we prove the positivity property $\langle\bar{\mu}, \varphi\rangle \geq 0, \quad \forall\varphi \in C(\bar{\Omega})$ with $\varphi \geq 0$. By construction of the update of the Lagrange multiplier we get $\mu_n^+ \geq 0$ implying

$$\int_\Omega \mu_n^+ \varphi \,\mathrm{d}x \geq 0, \quad \forall\varphi \in C(\bar{\Omega}) \text{ with } \varphi \geq 0,$$

which in turn yields

$$0 \leq \int_\Omega \mu_{n_j}^+ \varphi \,\mathrm{d}x \to \langle\bar{\mu}, \varphi\rangle, \quad \forall\varphi \in C(\bar{\Omega}) \text{ with } \varphi \geq 0.$$

Next, we show that the complementary slackness condition $\langle\bar{\mu}, \bar{y} - \psi\rangle = 0$ is fulfilled. From Theorem 3.11 we get $y_{n_j} \to \bar{y}$ in $C(\bar{\Omega})$. With Lemma 3.9, we get

$$0 = \lim_{j\to\infty} |(\mu_{n_j}^+, \psi - y_{n_j}^+)| = |(\bar{\mu}, \psi - \bar{y})|,$$

and hence the validity of the complementary condition. The inequality $(\bar{p} + \alpha\bar{u}, u - \bar{u}) \geq 0$ for $u \in U_{ad}$ follows with $u_{n_j}^+ \to \bar{u}$ in $L^2(\Omega)$ and $p_{n_j}^+ \rightharpoonup \bar{p}$ in $L^2(\Omega)$ from (7c). This shows that $(\bar{y}, \bar{u}, \bar{p}, \bar{\mu})$ satisfies (3). □

Since Lagrange multipliers are not uniquely determined in general, we cannot expect weak convergence of the whole sequences (μ_n^+) and (p_n^+). If we assume uniqueness of multipliers then this is possible indeed.

Corollary 3.12.1. *Let $(\bar{y}, \bar{u}, \bar{p}, \bar{\mu})$ satisfy (3). Let us assume that $(\bar{p}, \bar{\mu})$ are uniquely determined Lagrange multipliers. Then it holds*

$$\begin{aligned} p_n^+ &\rightharpoonup \bar{p} \quad \text{in } L^2(\Omega), \\ \mu_n^+ &\rightharpoonup^* \bar{\mu} \ \text{ in } \mathcal{M}(\bar{\Omega}). \end{aligned}$$

A sufficient condition for uniqueness of the adjoint state $\bar{p}$ and the Lagrange multiplier $\bar{\mu}$ is a certain separation condition on the active sets corresponding, see [21, Lemma 1]. There the following result was proven:

Lemma 3.13. *Let $\bar{u}$ be an optimal control of (P) and let Assumption 2 be fulfilled. Moreover, there exits $\delta > 0$ such that it holds for the active sets*

$$\begin{aligned} A_y &= \{x \in \bar{\Omega} \mid \bar{y}(x) = \psi(x)\} \\ A_u &= \{x \in \bar{\Omega} \mid \bar{u}(x) = u_a(x) \ \vee \ \bar{u}(x) = u_b(x)\} \end{aligned}$$

$dist(\bar{A}_y, \bar{A}_u) \geq \delta$, i.e., the active sets are well separated. Then, the corresponding adjoint state $\bar{p}$ and the Lagrange multiplier $\bar{\mu}$ are uniquely determined.

4 Numerical tests

In this section we report on numerical results for the solution of an elliptic pointwise state constrained optimal control problem in two dimensions. All optimal control problems have been solved using the above stated augmented Lagrange algorithm implemented with FEniCS [22] using the DOLFIN [23] Python interface. In the following, (y_h, u_h, p_h, μ_h) denote the calculated solutions after the stopping criterion is reached.

Example 1

We consider an optimal control problem given by

$$\min \ J(y,u) := \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)}$$
$$s.t. \begin{cases} -\Delta y = u & \text{in } \Omega, \\ \quad y = 0 & \text{on } \Gamma, \\ y(x) \leq \psi(x) & \text{in } \Omega, \end{cases}$$

with $\Omega = [0,1] \times [0,1]$. It is well known, that the state equation admits for every $u \in L^p(\Omega)$, $p > N/2$ a unique weak solution $y \in H_0^1(\Omega) \cap C(\bar{\Omega})$, see [9]. Hence, all results can be transferred to this type of state equation. For our example we choose the data as in [11], which is given by $y_d = 10(\sin(2\pi x_1) + x_2)$, $\psi = 0.01$ and $\alpha = 0.1$.

We choose the parameter in the decision concerning successful steps to be $\tau = 0.8$. If a step has not been successful, the penalization parameter is increased by the factor $\theta = 10$. The algorithm was stopped as soon as

$$R_n^+ := \left\|(y_n^+ - \psi)_+\right\|_{C(\bar{\Omega})} + |(\mu_n^+, \psi - y_n^+)| \leq 2 \cdot 10^{-10}$$

was satisfied.

The Figures 1 and 2 show the numerical solution of Example 1. All figures depict results gained for a triangular mesh with 10^4 degrees of freedom (dofs).

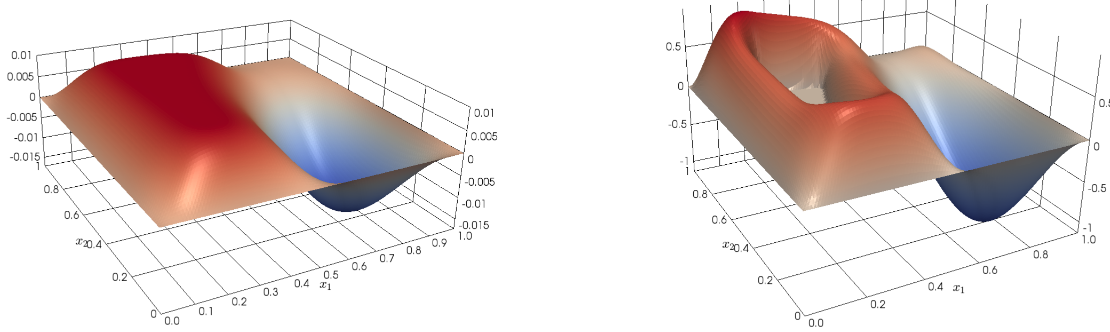

Figure 1: (Example 1) Computed discrete optimal state y_h (left) and optimal control u_h (right)

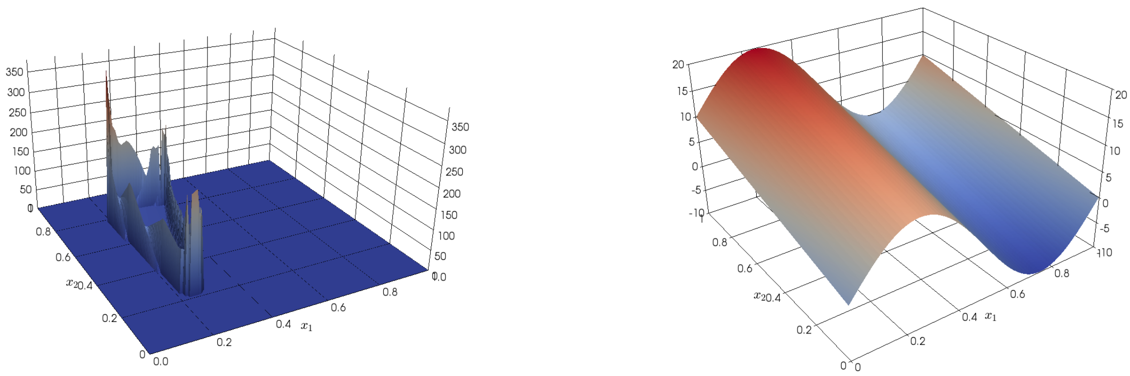

Figure 2: (Example 1) Computed discrete multiplier μ_h (left) and desired state y_d (right)

Figure 3 illustrates the $L^1(\Omega)$-Norm of the approximated Lagrange multiplier μ_k during the iterations. Clearly, this sequence is bounded in $L^1(\Omega)$. In addition, the values of the penalization parameters ρ_k are depicted in logarithmic scale. As can be seen, this sequence is not bounded. If it would have been bounded, then the sequence (μ_k) would be bounded in $L^2(\Omega)$ due to inequality (15).

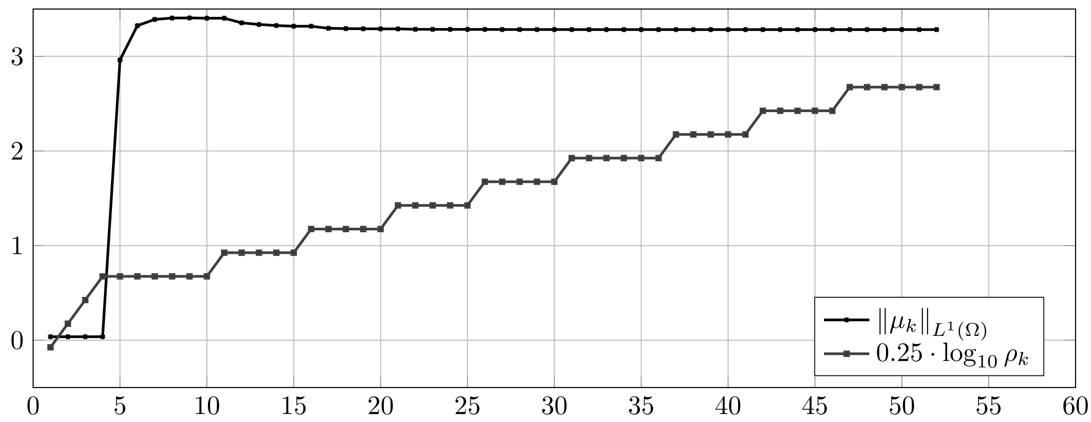

Figure 3: (Example 1) $L^1(\Omega)$-norm of discrete multipliers μ_k, penalty parameters ρ_k vs. iteration number

Example 2

As a second example we choose from [25]. This example is given by

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

$$s.t. \begin{cases} -\Delta y = u + f & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma, \\ y(x) \le \psi(x) & \text{in } \Omega, \end{cases}$$

with $\Omega = (-1,2)^2$. Setting $r = \sqrt{x_1^2 + x_2^2}$ and

$$y_d(r) = \bar{y}(r) - \frac{1}{2\pi}\chi_{r\le 1}(4-9r), \qquad \psi(r) = -\frac{1}{2\pi\alpha}\left(\frac{1}{4} - \frac{r}{2}\right), \qquad f(r) = \frac{1}{8}\chi_{r\le 1}(4 - 9r + 4r^2 - 4r^3).$$

The exact solution of the optimal control problem is given by

$$\bar{y}(r) = -\frac{1}{2\pi\alpha}\chi_{r\le 1}\left(\frac{r^2}{4}(\log r - 2) + \frac{r^3}{4} + \frac{1}{4}\right), \qquad \bar{u}(r) = \frac{1}{2\pi\alpha}\chi_{r\le 1}(\log r + r^2 - r^3),$$

$$\bar{p}(r) = \alpha\bar{u}(r), \qquad \bar{\mu}(r) = \delta_0(r).$$

Figure 4 shows the desired state y_d and the state constraint ψ.

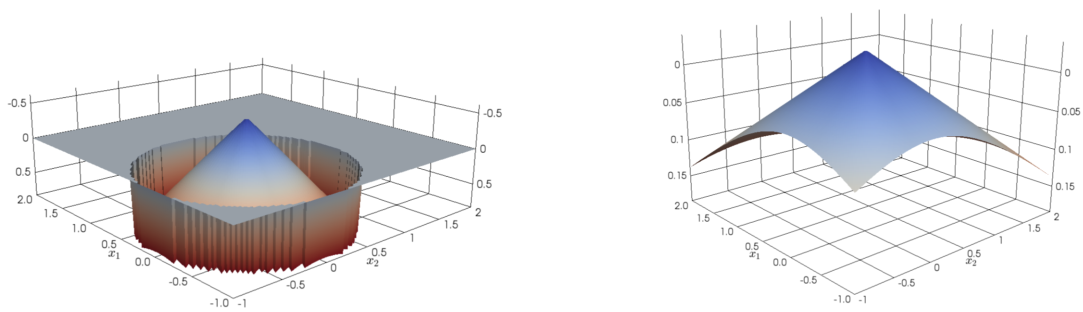

Figure 4: Desired state y_d (left) and state constraint ψ (right)

The Figures 5 and 6 show our numerical results for Example 2 using the same parameters as in Example 1. This solution was computed on a triangular mesh with 10^4 degrees of freedom. The computed Lagrange multiplier behaves like expected, approximating $\delta_0(r)$.

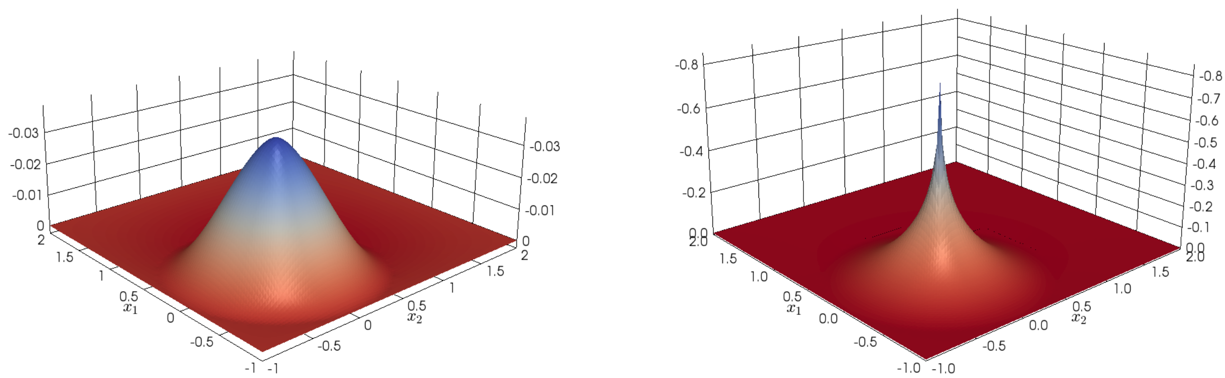

Figure 5: (Example 2) Computed discrete optimal state y_h (left) and optimal control u_h (right)

Since the exact solution of the problem is known, the errors $\|u_h - \bar{u}\|_{L^2(\Omega)}$ and $\|y_h - \bar{y}\|_{L^2(\Omega)}$ can be evaluated. Figure 7 depicts the errors depending on the numbers of degrees of freedom, showing once again convergence of our algorithm.

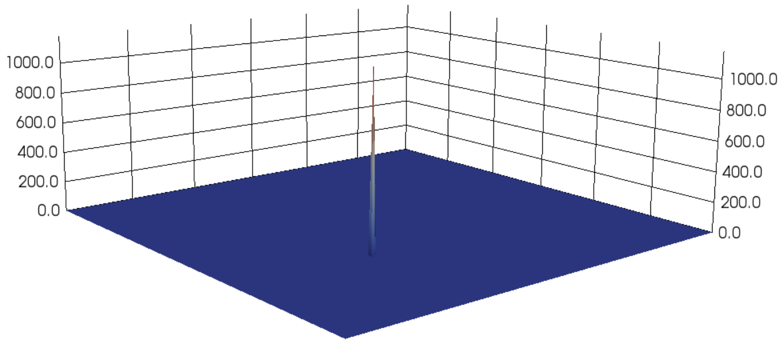

Figure 6: (Example 2) Computed discrete multiplier μ_h

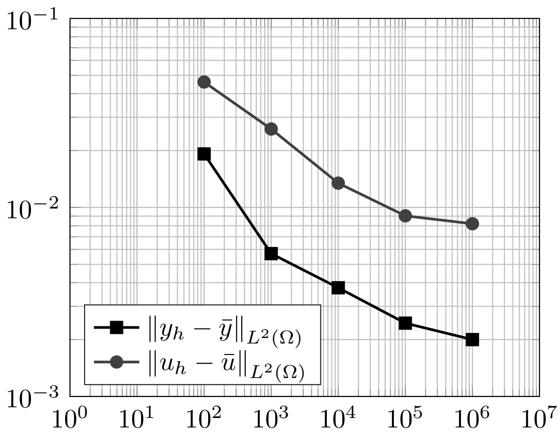

Figure 7: (Example 2) Errors $\|u_h - \bar{u}\|_{L^2(\Omega)}$ and $\|y_h - \bar{y}\|_{L^2(\Omega)}$ vs. degrees of freedom.

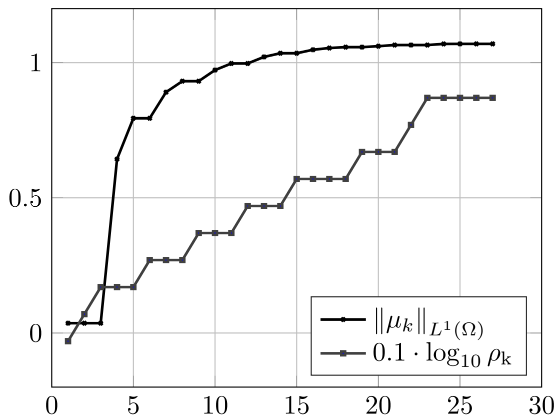

Figure 8: (Example 2) $L^1(\Omega)$-norm of discrete multipliers μ_k, penalty parameters ρ_k vs. iteration number

In Figure 8 the computed L^1-norm and the behaviour of the penalization parameter ρ are shown for a mesh with 10^5 degrees of freedom.

Iteration numbers and penalization parameter

Finally, let us report about the number of iterations and the final penalization parameter for different refinements of the mesh in both examples. Table 1 shows the number of iterations until the stopping criterion is reached for our two examined examples. It also represents the penalization parameter ρ_{max} of the final iteration and the L^1-norm of the approximated Lagrange multiplier.

Degrees of freedom		10^2	10^3	10^4	10^5
Example 1	it	33	41	46	53
	ρ_{max}	$5 \cdot 10^4$	$5 \cdot 10^6$	$5 \cdot 10^8$	$5 \cdot 10^{10}$
	$\|\mu_h\|_{L^1(\Omega)}$	3.284	3.284	3.282	3.282
Example 2	it	33	21	22	27
	ρ_{max}	$5 \cdot 10^3$	$5 \cdot 10^4$	$5 \cdot 10^5$	$5 \cdot 10^8$
	$\|\mu_h\|_{L^1(\Omega)}$	0.911	1.018	1.059	1.069

Table 1: Iteration history for both examples and different discretizations

References

[1] M. Bergounioux. Augmented Lagrangian method for distributed optimal control problems with state constraints. *J. Optim. Theory Appl.*, 78(3):493–521, 1993.

[2] M. Bergounioux. On boundary state constrained control problems. *Numer. Funct. Anal. Optim.*, 14(5-6):515–543, 1993.

[3] M. Bergounioux and K. Kunisch. Augmented Lagrangian techniques for elliptic state constrained optimal control problems. *SIAM J. Control Optim.*, 35(5):1524–1543, 1997.

[4] E. G. Birgin, R. A. Castillo, and J. M. Martínez. Numerical comparison of augmented Lagrangian algorithms for nonconvex problems. *Comput. Optim. Appl.*, 31(1):31–55, 2005.

[5] E. G. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014.

[6] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986.

[7] E. Casas. Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control Optim.*, 31(4):993–1006, 1993.

[8] J. C. De los Reyes. *Numerical PDE-constrained optimization*. Springer, Heidelberg, 2015.

[9] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag, Berlin-New York, 1977. Grundlehren der Mathematischen Wissenschaften, Vol. 224.

[10] M. Hintermüller and M. Hinze. Moreau-Yosida regularization in state constrained elliptic control problems: error estimates and parameter adjustment. *SIAM J. Numer. Anal.*, 47(3):1666–1683, 2009.

[11] M. Hintermüller and K. Kunisch. Feasible and noninterior path-following in constrained minimization with low multiplier regularity. *SIAM Journal on Control and Optimization*, 45(4):1198–1221, 2006.

[12] M. Hintermüller and K. Kunisch. Path-following methods for a class of constrained minimization problems in function space. *SIAM J. Optim.*, 17(1):159–187 (electronic), 2006.

[13] M. Hintermüller and K. Kunisch. PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative. *SIAM J. Optim.*, 20(3):1133–1156, 2009.

[14] M. Hinze and C. Meyer. Variational discretization of Lavrentiev-regularized state constrained elliptic optimal control problems. *Comput. Optim. Appl.*, 46(3):487–510, 2010.

[15] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.

[16] K. Ito and K. Kunisch. Augmented Lagrangian methods for nonsmooth, convex optimization in Hilbert spaces. *Nonlinear Anal.*, 41(5-6, Ser. A: Theory Methods):591–616, 2000.

[17] K. Ito and K. Kunisch. Semi-smooth Newton methods for state-constrained optimal control problems. *Systems Control Lett.*, 50(3):221–228, 2003.

[18] K. Ito and K. Kunisch. Semi-smooth Newton methods for variational inequalities of the first kind. *M2AN Math. Model. Numer. Anal.*, 37(1):41–62, 2003.

[19] C. Kanzow, D. Steck, and D. Wachsmuth. An augmented Lagrangian method for optimization problems in Banach spaces. Preprint SPP1962-003 of priority program "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization" (SPP 1962), 2016.

[20] K. Krumbiegel, I. Neitzel, and A. Rösch. Sufficient optimality conditions for the Moreau-Yosida-type regularization concept applied to semilinear elliptic optimal control problems with pointwise state constraints. *Ann. Acad. Rom. Sci. Ser. Math. Appl.*, 2(2):222–246, 2010.

[21] K. Krumbiegel, I. Neitzel, and A. Rösch. Regularization for semilinear elliptic optimal control problems with pointwise state and control constraints. *Comput. Optim. Appl.*, 52(1):181–207, 2012.

[22] A. Logg, K.-A. Mardal, G. N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.

[23] A. Logg and G. N. Wells. Dolfin: Automated finite element computing. *ACM Transactions on Mathematical Software*, 37(2), 2010.

[24] C. Meyer, A. Rösch, and F. Tröltzsch. Optimal control of PDEs with regularized pointwise state constraints. *Comput. Optim. Appl.*, 33(2-3):209–228, 2006.

[25] A. Rösch and D. Wachsmuth. A-posteriori error estimates for optimal control problems with state and control constraints. *Numer. Math.*, 120(4):733–762, 2012.

[26] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg, Wiesbaden, 2005.