Hamburger Beiträge
zur Angewandten Mathematik

A goal-oriented adaptive Moreau-Yosida algorithm for control- and state-constrained elliptic control problems

Andreas Günther and Moulay Hicham Tber

Nr. 2009-17
December 2009

A goal-oriented adaptive Moreau-Yosida algorithm for control- and state-constrained elliptic control problems

Andreas Günther * & Moulay Hicham Tber †

December 3, 2009

Abstract

In this work we develop an adaptive algorithm for solving elliptic optimal control problems with simultaneously appearing state and control constraints. The algorithm combines a Moreau-Yosida technique for handling state constraints with a semi-smooth Newton method for solving the optimality systems of the regularized sub-problems. The state and co-state variables are discretized using continuous piecewise linear finite elements while a variational discretization concept is applied for the control. To perform the adaptive mesh refinements cycle we derive local error estimators which extend the goal-oriented error approach to our setting. The performance of the overall adaptive solver is assessed by numerical examples.

Mathematics Subject Classification (2000): 49J20, 49K20, 65K10, 65N50

Keywords: Elliptic optimal control problem, control and state constraints, Moreau-Yosida regularization, semi-smooth Newton method, variational discretization, goal-oriented adaptivity.

1 Introduction

Optimal control problems with state constraints have been the topic of an increasing number of theoretical and numerical studies. The challenging character of these problems roots in the fact that state constraints feature low regular Lagrange multipliers [4, 7]. This low regularity does not allow a pointwise interpretation which complicates not only the analysis of

*Bereich Optimierung und Approximation, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany.

†Institut für Mathematik und wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria.

these problems but also their numerical treatment as well. In addition, in the presence of control constraints, the solution may exhibit subsets of the underlying domain where both control and state are active simultaneously. In this case, the uniqueness of Lagrange multipliers can not be guaranteed [25] which yields to undetermined optimality systems. To overcome these difficulties several techniques in the literature have been proposed. Very popular are relaxation concepts for state constraints such as Lavrentiev, interior point and Moreau-Yosida regularization. The former one is investigated by Tröltzsch in [26] and together with Meyer and Rösch in [20]. Barrier methods in function space ([29]) applied to state constrained optimal control problems are considered by Schiela in [23]. Relaxation by Moreau-Yosida regularization is matter of subject for the fully discrete case in [3, 5] as well as in function space in the work [15] by Hintermüller and Kunisch. Recently in [16] a generalized Moreau-Yosida-based framework also applies for constraints on the gradient of the state. However, as far as we are concerned with adaptive approaches, experiences with this type of problems stay limited. Residual-type a posteriori error estimators for mixed control-state constrained problems are derived in [18]. On the other hand the dual weighted residual method proposed in [1] is applied to derive goal-oriented adaptive meshes for better resolving a certain quantity of interest. This technique is extended for pde-constrained optimization to the presence of control constraints in [11, 27] as well as to state constraints in [2, 10, 12]. Within the framework of goal-oriented adaptive function space algorithms a Lavrentiev regularization approach is considered in [13, 18] while an adaptive interior point method is proposed in the works [24, 30].

In this work we design an adaptive finite element algorithm to solve elliptic optimal control problems with control and pointwise state constraints. Following [16], our algorithm combines a Moreau-Yosida regularization approach with a semi-smooth Newton solver [14]. We apply the variational discretization concept [8, 17] to the state equation of the regularized optimal control problem. Moreover, for a fixed regularization parameter, we develop a goal-oriented a posteriori error estimate to assess the performance of the variational discretization in terms of the objective functional. We therefore derive a regularized extension of the error representation obtained in [10] to the control and state constrained case. In particular no residual associated to the first order optimality condition with respect to the control appears in our approach. We mention here that we are not interested in the error involved by the regularization parameter. Our aim is rather performing a first attempt to understand the behaviour of a goal-oriented based error estimate in connection with a Moreau-Yosida regularization. An overall error reduction which ties the regularization parameter with the current mesh size is subject of an ongoing research work.

The rest of this paper is organized as follows: In the next section we present the optimal control problem under consideration and recall its first

order necessary optimality system. In Section 3 we introduce the regularized version of the problem and state the main convergence theorem. In Section 4 we first apply variational discretization to the regularized sub-problems and propose a semi-smooth Newton solver for the resulting discrete systems. In Section 5 we derive the error representation in terms of the objective functional and we adress the related implementation issues. Finally, numerical examples are reported in Section 6.

2 Optimal control problem

Let Ω be a bounded polygonal and convex domain in $\mathbb{R}^d\,(d = 2, 3)$ with boundary $\partial\Omega$. We consider the general elliptic partial differential operator $\mathcal{A} : H^1(\Omega) \longrightarrow H^1(\Omega)^*$ defined by

$$\mathcal{A}y := \sum_{i,j=1}^{d} \partial_{x_j}(a_{ij} y_{x_i}) + \sum_{i=1}^{d} b_i y_{x_i} + cy$$

along with its formal adjoint operator $\mathcal{A}^*$

$$\mathcal{A}^* y = \sum_{i=1}^{d} \partial_{x_i}\Big(\sum_{j=1}^{d} a_{ij} y_{x_j} + b_i y\Big) + cy.$$

We subsequently assume the coefficients a_{ij}, b_i and c $(i, j = 1, \ldots, d)$ to be sufficiently smooth functions on $\bar{\Omega}$. Moreover we suppose that there exists $c_0 > 0$ such that $\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geq c_0$ for almost all x in Ω and all ξ in $\mathbb{R}^d$. Corresponding to the operator $\mathcal{A}$ we associate the bilinear form $a(\cdot, \cdot) : H^1(\Omega) \times H^1(\Omega) \longrightarrow \mathbb{R}$ with

$$a(y, v) := \int_\Omega \Big(\sum_{i,j=1}^{d} a_{ij} y_{x_i} v_{x_j} + \sum_{i=1}^{d} b_i y_{x_i} v + cyv\Big).$$

Suppose that the form a is coercive on $H^1(\Omega)$, i.e. there exists $c_1 > 0$ such that $a(v, v) \geq c_1 \|v\|^2_{H^1(\Omega)}$ for all v in $H^1(\Omega)$. This follows for instance when

$$\inf \mathrm{ess}_{x\in\Omega}\Big(c - \frac{1}{2}\sum_{i=1}^{d} \partial_{x_i} b_i\Big) > 0 \text{ and } \inf \mathrm{ess}_{x\in\partial\Omega}\Big(\sum_{i=1}^{d} b_i \nu_i\Big) \geq 0$$

holds. Here ν denotes the unit outward normal at $\partial\Omega$.

For given $u \in L^2(\Omega)$ and fixed $f \in L^2(\Omega)$ the homogeneous Neumann boundary value problem

$$\begin{aligned} \mathcal{A}y &= u + f && \text{in } \Omega \\ \partial_{\nu_\mathcal{A}} y := \textstyle\sum_{i,j=1}^{d} a_{ij} y_{x_i} \nu_j &= 0 && \text{on } \partial\Omega \end{aligned} \tag{2.1}$$

has a unique solution $y =: \mathcal{G}(u) \in H^2(\Omega)$. Moreover, there exists a constant C depending on f and the domain Ω such that

$$\|\mathcal{G}(u)\|_{H^2(\Omega)} \le C\|u\|_{L^2(\Omega)} + C.$$

We notice that (2.1) should be interpreted in the variational form

$$a(y, v) = (u + f, v) \qquad \forall v \in H^1(\Omega), \tag{2.2}$$

where $(\,\cdot, \cdot\,)$ stands for the standard inner product in $L^2(\Omega)$ or $L^2(\Omega)^d$ respectively.

Now for given $u_d, y_d \in L^2(\Omega)$, $\alpha > 0$ and u_a, u_b, y_a and $y_b \in \mathbb{R}$ with $u_a < u_b$ and $y_a < y_b$ we focus on the optimal control problem

$$\begin{aligned} J(y,u) &:= \tfrac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \tfrac{\alpha}{2}\|u - u_d\|^2_{L^2(\Omega)} \to \min \\ \text{s.t.} \quad y = \mathcal{G}(u), &\quad u \in U_{ad}, \quad \text{and} \quad y_a \le y \le y_b \quad \text{a.e. in } \Omega, \end{aligned} \tag{2.3}$$

where U_{ad} is the set of admissible controls given by

$$U_{ad} = \left\{ u \in L^2(\Omega) \,:\, u_a \le u \le u_b \quad \text{in } \Omega \right\}.$$

Under the Slater condition

$$\exists u_s \in U_{ad} : \qquad y_a < \mathcal{G}(u_s) < y_b \quad \text{in } \Omega, \tag{2.4}$$

Theorem 2.1 holds true (see for instance [6]). Condition (2.4) holds for instance if there exists some $\beta \in (0,1)$ such that $u_a \le ((1-\beta)y_a + \beta y_b)c(x) - f(x) \le u_b$ a.e. in Ω. Below the space of Radon measures $\mathcal{M}(\bar\Omega)$ is identified to the dual space of $C^0(\bar\Omega)$ such that

$$\langle \mu, y\rangle := \langle \mu, y\rangle_{\mathcal{M}(\bar\Omega), C^0(\bar\Omega)} := \int_{\bar\Omega} y \,\mathrm{d}\mu \quad \forall \mu \in \mathcal{M}(\bar\Omega)\ \forall y \in C^0(\bar\Omega).$$

and

$$\forall \mu \in \mathcal{M}(\bar\Omega) : \quad \mu \ge 0 \Longleftrightarrow \langle \mu, y\rangle \ge 0 \quad \forall y \in C^0(\bar\Omega) \text{ with } y \ge 0.$$

Theorem 2.1. *The optimal control problem* (2.3) *has a unique solution* $(y^*, u^*) \in H^2(\Omega) \times U_{ad}$. *Moreover there exist* $p^* \in W^{1,s}(\Omega)$ *for all* $1 \le s < d/(d-1)$, λ_a^*, $\lambda_b^* \in L^2(\Omega)$ *and* μ_a^*, $\mu_b^* \in \mathcal{M}(\bar\Omega)$ *satisfying the optimality system*

$$\begin{aligned} y^* &= \mathcal{G}(u^*), \\ (p^*, \mathcal{A}v) = (y^* - y_d, v) + \langle \mu_a^* + \mu_b^*, v\rangle \quad &\forall v \in W^{1,1-\frac{1}{s}}(\Omega) \text{ with } \partial_{\nu_{\mathcal{A}}} v|_{\partial\Omega} = 0, \\ \alpha(u^* - u_d) + p^* + \lambda_a^* + \lambda_b^* &= 0, \\ \lambda_a^* \le 0, \quad u^* \ge u_a, \quad (\lambda_a^*, u^* - u_a) &= 0, \\ \lambda_b^* \ge 0, \quad u^* \le u_b, \quad (\lambda_b^*, u^* - u_b) &= 0, \\ \mu_a^* \le 0, \quad y^* \ge y_a, \quad \langle \mu_a^*, y^* - y_a\rangle &= 0, \\ \mu_b^* \ge 0, \quad y^* \le y_b, \quad \langle \mu_b^*, y^* - y_b\rangle &= 0. \end{aligned} \tag{2.5}$$

3 Moreau-Yosida regularization

From the previous theorem we can see that the state constraints in (2.3) features low regularity Lagrange multipliers. The adjoint equation in the optimality system (2.5) is posed in a very weak sense and μ_a^*, μ_b^* are lying only on the measure space $\mathcal{M}(\bar{\Omega})$. This low regularity does not allow a pointwise interpretation of the multipliers which complicates its analysis as well as its numerical treatment. In this section, we overcome this difficulty by applying a Moreau-Yosida regularization technique. In our case this technique penalizes the state constraints $y_a \leq y \leq y_b$ by modifying the objective functional J. The corresponding regularized optimal control problem reads

$$\begin{gathered} J^\gamma(y,u) := J(y,u) + \tfrac{\gamma}{2}\|\min(0, y - y_a)\|^2 + \tfrac{\gamma}{2}\|\max(0, y - y_b)\|^2 \to \min \\ \text{s.t.} \quad y = \mathcal{G}(u) \quad \text{and} \quad u \in U_{ad}, \end{gathered} \tag{3.1}$$

where $\gamma > 0$ denotes a regularization parameter tending to $+\infty$ later on. The max- and min-expressions in the regularized objective functional J^γ arise from regularizing the indicator function corresponding to the set of admissible states.
Notice that (3.1) is a pure control constrained optimal control problem that has a unique solution $(y^\gamma,\, u^\gamma) \in H^2(\Omega) \times U_{ad}$. Furthermore, we can prove the existence of Lagrange multipliers $\left(p^\gamma, \lambda_a^\gamma, \lambda_b^\gamma\right) \in L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)$ using standard theory of mathematical programming in Banach spaces [31] such that

$$\begin{aligned} y^\gamma &= \mathcal{G}(u^\gamma) \\ (p^\gamma, \mathcal{A}v) = (y^\gamma - y_d, v) + \left(\mu_a^\gamma + \mu_b^\gamma, v\right) \quad &\forall v \in H^2(\Omega) \text{ with } \partial_{\nu_\mathcal{A}} v|_{\partial\Omega} = 0, \\ \alpha(u^\gamma - u_d) + p^\gamma + \lambda_a^\gamma + \lambda_b^\gamma &= 0, \\ \lambda_a^\gamma \leq 0, \quad u^\gamma \geq u_a, \quad &(\lambda_a^\gamma, u^\gamma - u_a) = 0, \\ \lambda_b^\gamma \geq 0, \quad u^\gamma \leq u_b, \quad &\left(\lambda_b^\gamma, u^\gamma - u_b\right) = 0 \end{aligned} \tag{3.2}$$

holds, where

$$\mu_a^\gamma = \gamma \min(0, y^\gamma - y_a) \quad \text{and} \quad \mu_b^\gamma = \gamma \max(0, y^\gamma - y_b).$$

The convergence of the solutions of the regularized problems is the purpose of the next result whose proof follows from the discussion in [16].

Theorem 3.1. *Let* $\{(y^\gamma, u^\gamma, p^\gamma, \lambda_a^\gamma, \lambda_b^\gamma)\}_{\gamma>0}$ *be a sequence of solutions of* (3.2). *Then, there exists a subsequence still denoted by* $\{(y^\gamma, u^\gamma, p^\gamma, \lambda_a^\gamma, \lambda_b^\gamma)\}_{\gamma>0}$ *and*

$$(p^\star, \lambda_a^\star, \lambda_b^\star, \mu_a^\star, \mu_b^\star) \in L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathcal{M}(\bar{\Omega}) \times \mathcal{M}(\bar{\Omega})$$

such that, under the Slater condition (2.4),

$$
\begin{aligned}
y^\gamma &\to y^* \ \textit{in}\ C^0(\bar\Omega),\\
y^\gamma &\rightharpoonup y^* \ \textit{in}\ H^2(\Omega)',\\
u^\gamma &\rightharpoonup u^* \ \textit{in}\ L^2(\Omega),\\
p^\gamma &\rightharpoonup p^\star \ \textit{in}\ L^2(\Omega),\\
\lambda_a^\gamma &\rightharpoonup \lambda_a^\star \ \textit{in}\ L^2(\Omega),\\
\lambda_b^\gamma &\rightharpoonup \lambda_b^\star \ \textit{in}\ L^2(\Omega),\\
\mu_a^\gamma &\rightharpoonup \mu_a^\star \ \textit{in}\ \mathcal{M}(\bar\Omega),\\
\mu_b^\gamma &\rightharpoonup \mu_b^\star \ \textit{in}\ \mathcal{M}(\bar\Omega),
\end{aligned}
$$

as $\gamma \to +\infty$*, with* $(y^*, u^*, p^\star, \lambda_a^\star, \lambda_b^\star, \mu_a^\star, \mu_b^\star)$ *being a solution to the optimality system* (2.5).

4 Optimality system

Regarding the previous theorem, to recover the solution of the optimal control problem (2.3) an overall algorithm can be designed by solving (3.1) for a sequence $\gamma \to \infty$. For (3.1) with γ fixed, a locally superlinear semi-smooth Newton method can be applied (see [16]). In order to discretize the corresponding optimality system (3.2), we follow the variational discretization concept introduced in [17], we approximate the space of state variables using finite elements but keeping the infinite dimensional space $U_{ad} \subset L^2(\Omega)$ as set of admissible controls.

Variational discretization

In the sequel and for the computational purposes we consider a shape-regular simplicial triangulation $\mathcal{T}_h$ of Ω. Since Ω is assumed to be a polyhedral, the boundary $\partial\Omega$ is exactly represented by the boundaries of simplices $T \in \mathcal{T}_h$. We refer to $\mathcal{N}_h = \cup_{i=1}^{np}\{x_i\}$ as the set of nodes of $\mathcal{T}_h$. For each element T in $\mathcal{T}_h$, we denote by h_T and $|T|$ the diameter and Lebesgue $\mathbb{R}^d$ measure of T, respectively. The overall mesh size is defined by $h := \max_{T\in\mathcal{T}_h} \operatorname{diam} T$. Further, we associate with $\mathcal{T}_h$ the continuous piecewise linear finite element space

$$V_h = \{v \in C_0(\bar\Omega) \,:\, v|_T \in P_1(T),\ \forall T \in \mathcal{T}_h\},$$

where $P_1(T)$ is the space of first-order polynomials on T. The standard nodal basis of V_h denoted by $\{\phi_i\}_{i=1}^{np}$ satisfies $\phi_i(x_j) = \delta_{ij}$ for all x_j in $\mathcal{N}_h$ and $i, j \in \{1, \ldots, np\}$. Here, δ_{ij} represents the Kronecker symbol. Furthermore for all $v \in C^0(\bar\Omega)$ we denote by $i_h v := \sum_{i=1}^{np} v(x_i)\phi_i$ the Lagrange interpolation of v with x_i denoting the i-th vertex in $\mathcal{T}_h$. In analogy to (2.2) we define for

given $u \in L^2(\Omega)$ the discrete solution operator $\mathcal{G}_h$ by

$$y_h =: \mathcal{G}_h(u) \Longleftrightarrow y_h \in V_h \text{ and } a(y_h, v_h) = (u + f, v_h) \quad \forall v_h \in V_h.$$

We approximate the objective functional J by a sequence of objectives J_h

$$J_h(y_h, u_h) := \frac{1}{2}\|y_h - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u_h - u_{d,h}\|^2_{L^2(\Omega)},$$

where $u_{d,h} \in V_h$ denotes a finite element function corresponding to the given shift control $u_d \in L^2(\Omega)$ with the property

$$\|u_d - u_{d,h}\|_{L^2(\Omega)} \leq Ch.$$

This can be realized by the standard L^2-projection of u_d onto V_h.

We are now in the position to apply variational discretization [17] to problem (3.1). We therefore consider

$$\begin{gathered} J_h^\gamma(y_h, u_h) := J_h(y_h, u_h) + \tfrac{\gamma}{2}\|\min(0, y_h - y_a)\|^2 + \tfrac{\gamma}{2}\|\max(0, y_h - y_b)\|^2 \to \min \\ \text{s.t.} \quad y_h = \mathcal{G}_h(u_h) \quad \text{and} \quad u_h \in U_{ad}. \end{gathered} \tag{4.1}$$

The existence of a solution of (4.1) as well as Lagrange multipliers again follows from standard arguments. The corresponding first order optimality system of (4.1) leads to the variationally discretized counterpart of (3.2)

$$\begin{aligned} y_h^\gamma &= \mathcal{G}_h(u_h^\gamma), \\ a(v_h, p_h^\gamma) &= (v_h, y_h^\gamma - y_d + \mu_{a,h}^\gamma + \mu_{b,h}^\gamma) \quad \forall v_h \in V_h, \\ \alpha(u_h^\gamma - u_{d,h}) + p_h^\gamma + \lambda_{a,h}^\gamma + \lambda_{b,h}^\gamma &= 0, \\ \lambda_{a,h}^\gamma \leq 0, \quad u_h^\gamma \geq u_a, \quad (\lambda_{a,h}^\gamma, u_h^\gamma - u_a) &= 0, \\ \lambda_{b,h}^\gamma \geq 0, \quad u_h^\gamma \leq u_b, \quad (\lambda_{b,h}^\gamma, u_h^\gamma - u_b) &= 0, \end{aligned} \tag{4.2}$$

where $y_h^\gamma, p_h^\gamma \in V_h$ and $u_h^\gamma, \lambda_{a,h}^\gamma, \lambda_{b,h}^\gamma \in L^2(\Omega)$. The multipliers corresponding to the regularized state constraints $\mu_{a,h}^\gamma$ and $\mu_{b,h}^\gamma$ are given by

$$\mu_{a,h}^\gamma = \gamma \min(0, y_h^\gamma - y_a) \quad \text{and} \quad \mu_{b,h}^\gamma = \gamma \max(0, y_h^\gamma - y_b).$$

We mention here that (4.1) is a function space optimization problem and the optimal control u_h^γ is not lying in a finite element space in general. However, regarding (4.2), u_h^γ corresponds to the projection of a finite element quantity onto the admissible set U_{ad}

$$u_h^\gamma = \Pi_{[u_a, u_b]}\left(-\frac{1}{\alpha}p_h^\gamma + u_{d,h}\right),$$

where $\Pi_{[u_a,u_b]}$ is the orthogonal projection onto U_{ad}. This special structure of u_h^γ allows a matricial representation of (4.2). In what follows we extend

the algorithm prescribed in [8] to the regularized problem (4.1). Therefore let us introduce the standard mass- and system matrix as

$$\mathbf{M} = \left[\int_\Omega \phi_i \phi_j\right]_{i,j=1}^{np} \quad \text{and} \quad \mathbf{A} = \left[a(\phi_i, \phi_j)\right]_{i,j=1}^{np}.$$

We represent $L^2(\Omega)$-projections of the given data f, u_d, y_d onto V_h by the vectors

$$\mathbf{f} = \mathbf{M}^{-1}\left[\int_\Omega f\phi_i\right]_{i=1}^{np}, \quad \mathbf{u_d} = \mathbf{M}^{-1}\left[\int_\Omega u_d\phi_i\right]_{i=1}^{np}, \quad \mathbf{y_d} = \mathbf{M}^{-1}\left[\int_\Omega y_d\phi_i\right]_{i=1}^{np}$$

with corresponding finite element functions

$$f_h = \sum_{i=1}^{np} f_i\phi_i, \quad u_{d,h} = \sum_{i=1}^{np} u_{d,h}^i\phi_i, \quad y_{d,h} = \sum_{i=1}^{np} y_{d,h}^i\phi_i.$$

Further we denote

$$\mathbf{u_a} = \left[u_a\right]_{i=1}^{np}, \mathbf{u_b} = \left[u_b\right]_{i=1}^{np}, \mathbf{y_a} = \left[y_a\right]_{i=1}^{np}, \mathbf{y_b} = \left[y_b\right]_{i=1}^{np},$$

and

$$\mathbf{y}^\gamma = \left[y_h^\gamma(x_i)\right]_{i=1}^{np}, \quad \mathbf{p}^\gamma = \left[p_h^\gamma(x_i)\right]_{i=1}^{np}.$$

Of particular importance is the following vector representation of the action of an arbitrary $u_h^\gamma \in L^2(\Omega)$ on V_h basis functions

$$\mathbf{u}^\gamma = \left[\int_\Omega u_h^\gamma\phi_i\right]_{i=1}^{np},$$

which allows avoiding an explicit discretization of the control u. To determine the active sets of the control we use the projection formula between u_h^γ and p_h^γ. Hence, we define the control inactive, lower active and upper active sets, respectively, as

$$i(\mathbf{p}^\gamma) = \{x \in \Omega \,:\, u_a < -\tfrac{1}{\alpha}p_h^\gamma(x) + u_{d,h}(x) < u_b\}, \tag{4.3}$$
$$a(\mathbf{p}^\gamma) = \{x \in \Omega \,:\, -\tfrac{1}{\alpha}p_h^\gamma(x) + u_{d,h}(x) \le u_a\}, \tag{4.4}$$
$$b(\mathbf{p}^\gamma) = \{x \in \Omega \,:\, u_b \le -\tfrac{1}{\alpha}p_h^\gamma(x) + u_{d,h}(x)\}. \tag{4.5}$$

With respect to these sets we additively split the mass matrix $\mathbf{M}$ into

$$\mathbf{M} = \mathbf{M_i^p} + \mathbf{M_a^p} + \mathbf{M_b^p},$$

where

$$\mathbf{M_i^p} = \left[\int_{i(\mathbf{p}^\gamma)} \phi_i\phi_j\right]_{i,j=1}^{np}, \ \mathbf{M_a^p} = \left[\int_{a(\mathbf{p}^\gamma)} \phi_i\phi_j\right]_{i,j=1}^{np}, \ \mathbf{M_b^p} = \left[\int_{b(\mathbf{p}^\gamma)} \phi_i\phi_j\right]_{i,j=1}^{np}.$$

Similarly we define $\mathbf{M_a^y}$ and $\mathbf{M_b^y}$ by

$$\mathbf{M_a^y} = \left[\int_{a(\mathbf{y}^\gamma)} \phi_i\phi_j\right]_{i,j=1}^{np}, \quad \mathbf{M_b^y} = \left[\int_{b(\mathbf{y}^\gamma)} \phi_i\phi_j\right]_{i,j=1}^{np}$$

with

$$a(\mathbf{y}^\gamma) = \{x \in \Omega \,:\, y_h^\gamma(x) \le y_a\}, \quad b(\mathbf{y}^\gamma) = \{x \in \Omega \,:\, y_b \le y_h^\gamma(x)\}. \tag{4.6}$$

Let us emphasize that assembling all appearing mass matrices $\mathbf{M}_\bullet^\bullet$ can be vectorized within the few cases how triangles are active and / or inactive. The number of those cases is insignificantly compared to the total number of elements. The main CPU-time is required for solving the linearized systems we are going to introduce now. The matrix form of (4.2) reads

$$\mathbf{A}\mathbf{y}^\gamma - \mathbf{u}^\gamma - \mathbf{M}\mathbf{f} = \mathbf{0}, \tag{4.7}$$

$$\mathbf{A}^T\mathbf{p}^\gamma - \mathbf{M}(\mathbf{y}^\gamma - \mathbf{y_d}) - \gamma\mathbf{M}_\mathbf{a}^\mathbf{y}(\mathbf{y}^\gamma - \mathbf{y_a}) - \gamma\mathbf{M}_\mathbf{b}^\mathbf{y}(\mathbf{y}^\gamma - \mathbf{y_b}) = \mathbf{0}, \tag{4.8}$$

$$\mathbf{u}^\gamma - \Big(\mathbf{M}_\mathbf{i}^\mathbf{p}\big(-\frac{1}{\alpha}\mathbf{p}^\gamma + \mathbf{u_d}\big) + \mathbf{M}_\mathbf{a}^\mathbf{p}\mathbf{u_a} + \mathbf{M}_\mathbf{b}^\mathbf{p}\mathbf{u_b}\Big) = \mathbf{0}. \tag{4.9}$$

Solution algorithm

We reduce (4.7)-(4.9) to a nonlinear system in $\mathbf{x}^\gamma = [\mathbf{y}^\gamma; \mathbf{p}^\gamma]$

$$G^\gamma(\mathbf{x}^\gamma) := \begin{bmatrix} \mathbf{A}\mathbf{y}^\gamma - \big(\mathbf{M}_\mathbf{i}^\mathbf{p}(-\frac{1}{\alpha}\mathbf{p}^\gamma + \mathbf{u_d}) + \mathbf{M}_\mathbf{a}^\mathbf{p}\mathbf{u_a} + \mathbf{M}_\mathbf{b}^\mathbf{p}\mathbf{u_b}\big) - \mathbf{M}\mathbf{f} \\ \mathbf{A}^T\mathbf{p}^\gamma - \mathbf{M}(\mathbf{y}^\gamma - \mathbf{y_d}) - \gamma\mathbf{M}_\mathbf{a}^\mathbf{y}(\mathbf{y}^\gamma - \mathbf{y_a}) - \gamma\mathbf{M}_\mathbf{b}^\mathbf{y}(\mathbf{y}^\gamma - \mathbf{y_b}) \end{bmatrix} = \mathbf{0}. \tag{4.10}$$

Notice that, due to the presence of max- and min-operations involved in (4.3)-(4.6), G^γ is not Fréchet-differentiable and a classical Newton method can not be applied to solve (4.10). Nevertheless, a generalized Jacobian can be defined for $G^\gamma(\mathbf{x})$ with $\mathbf{x} = [\mathbf{y}; \mathbf{p}] \in \mathbb{R}^{2np}$ by

$$DG^\gamma(\mathbf{x}) := \begin{bmatrix} \mathbf{A} & \frac{1}{\alpha}\mathbf{M}_\mathbf{i}^\mathbf{p} \\ -(\mathbf{M} + \gamma\mathbf{M}_\mathbf{a}^\mathbf{y} + \gamma\mathbf{M}_\mathbf{b}^\mathbf{y}) & \mathbf{A}^T \end{bmatrix}.$$

To solve (4.10) we therefore perform semi-smooth Newton iterations (see for instance [14, 21, 22]): Given $x_0 \in \mathbb{R}^{2np}$, the iteration step reads

$$\mathbf{x}_{n+1} = \mathbf{x}_n - DG^\gamma(\mathbf{x}_n)^{-1}G^\gamma(\mathbf{x}_n) \quad \text{for } n = 0, 1, \dots \tag{4.11}$$

until some stopping criterion is satisfied. With an approximate solution of $G^\gamma(\mathbf{x}^\gamma) = \mathbf{0}$ at hand we recover the $L^2(\Omega)$-function u_h^γ via

$$u_h^\gamma(x) = \Pi_{[u_a,u_b]}\left(-\frac{1}{\alpha}p_h^\gamma(x) + u_{d,h}(x)\right).$$

Proposition 4.1. *The semi-smooth Newton iteration* (4.11) *is well defined. The sequence* $(\mathbf{x}_n)_{n\in\mathbb{N}}$ *generated by* (4.11) *converges to a solution* $\mathbf{x}^\gamma := [\mathbf{y}^\gamma; \mathbf{p}^\gamma]$ *of* (4.10) *provided that* $\|\mathbf{x}^\gamma - \mathbf{x}_0\|$ *is small enough.*

Proof. In order to show this proposition it suffices to prove that DG has got an inverse which is bounded in some neighborhood of $\mathbf{x}^\gamma$.
For an arbitrarily chosen $\mathbf{x} := [\mathbf{y}; \mathbf{p}] \in \mathbb{R}^{2np}$, we know that $\mathbf{C} := \mathbf{M} + \gamma \mathbf{M}_{\mathbf{a}}^{\mathbf{y}} + \gamma \mathbf{M}_{\mathbf{b}}^{\mathbf{y}}$ is symmetric and positive definite, $\mathbf{A}$ is positive definite and $\frac{1}{\alpha}\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}$ is symmetric positive semi-definite. A Schur complement of the matrix block $DG^\gamma(\mathbf{x})$ reads

$$\mathbf{S} := \mathbf{A} + \frac{1}{\alpha}\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}\mathbf{A}^{-T}\mathbf{C},$$

which can be written as

$$\mathbf{S} = \mathbf{A}(\mathbf{I} + \frac{1}{\alpha}\mathbf{A}^{-1}\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}\mathbf{A}^{-T}\mathbf{C}). \tag{4.12}$$

From [19, Thm. 7.6.3] it follows that the product of a real symmetric positive definite matrix and a real symmetric positive semi-definite one is a positive semi-definite matrix (which is not necessarily symmetric). Therefore $\mathbf{A}^{-1}\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}\mathbf{A}^{-T}\mathbf{C}$ is a positive semi-definite matrix and, from (4.12), $\mathbf{S}$ is invertible. Moreover, for a given $\mathbf{r} = [\mathbf{r_1}; \mathbf{r_2}] \in \mathbb{R}^{2np}$, the solution $\mathbf{d} = [\mathbf{d_1}; \mathbf{d_2}] \in \mathbb{R}^{2np}$ to the linear system

$$DG^\gamma(\mathbf{x})\mathbf{d} = \mathbf{r}$$

can be computed using

$$\mathbf{d_1} = \mathbf{S}^{-1}\mathbf{r_1} - \frac{1}{\alpha}\mathbf{S}^{-1}\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}\mathbf{A}^{-T}\mathbf{r_2}, \tag{4.13}$$

$$\mathbf{d_2} = \mathbf{A}^{-T}\mathbf{r_2} + \mathbf{A}^{-T}\mathbf{C}\mathbf{d_1}, \tag{4.14}$$

where

$$\mathbf{S}^{-1} = (\mathbf{I} + \frac{1}{\alpha}\mathbf{A}^{-1}\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}\mathbf{A}^{-T}\mathbf{C})^{-1}\mathbf{A}^{-1}.$$

Notice that (taking for instance the matrix norm induced by $\|\cdot\|_1$)

$$\|\mathbf{S}^{-1}\| \leq \mathrm{Cst}\|\mathbf{A}^{-1}\|, \tag{4.15}$$

$$\max\left(\|\mathbf{M}_{\mathbf{i}}^{\mathbf{p}}\|, \|\mathbf{M}_{\mathbf{a}}^{\mathbf{y}}\|, \|\mathbf{M}_{\mathbf{b}}^{\mathbf{y}}\|\right) \leq \mathrm{Cst}\|\mathbf{M}\|, \tag{4.16}$$

with Cst being a generic positive constant not depending on $\mathbf{x}$. Consequently, from (4.13), (4.14), (4.15) and (4.16) we infer that $\|DG^\gamma(\mathbf{x})^{-1}\|$ is bounded independently of $\mathbf{x}$ which completes the proof of this proposition. □

5 Error representation and estimator

To achieve high accuracies in an optimal fashion, we marry our regularization semi-smooth Newton solver with an adaptive mesh refinement process based on a goal-oriented approach. As quantity of interest we consider the objective functional J which is corresponding to the tracking part in the

objective functional of the regularized optimal control problem (3.1). For a fixed regularization parameter γ we derive hereafter an error representation in J for the solutions of (3.1) and (4.1) respectively. In this section we make the following

Assumption 5.1. $u_d = u_{d,h}$.

As a consequence of this assumption it holds $J = J_h$. Including more general desired controls u_d would lead to additional weighted data oscillation quantities $(u_d - u_{d,h}, \cdot)$ in the error representation (5.1). For residual type a posteriori estimators this was done in [18].
We mention here that the previous assumption is fulfilled by affine linear functions or, more precisely, by a piecewise linear function over the coarsest mesh in refinement processes which is not restrictive from practical point of view. Indeed, in contrast to y_d, u_d is not a desired control but a background control. In many applications, it is corresponding to the result of trial and error experiments performed with a small number of degrees of freedom.

Following [10] we define the following residuals

$$\begin{aligned}\rho^{p^\gamma}(\cdot) &:= J_y(y_h^\gamma, u_h^\gamma)(\cdot) - a(\cdot, p_h^\gamma) + (\mu_h^\gamma, \cdot)\\ \rho^{y^\gamma}(\cdot) &:= -a(y_h^\gamma, \cdot) + (u_h^\gamma + f, \cdot)\end{aligned}$$

with

$$\begin{aligned}\mu^\gamma &:= \gamma \min(0, y^\gamma - y_a) + \gamma \max(0, y^\gamma - y_b),\\ \mu_h^\gamma &:= \gamma \min(0, y_h^\gamma - y_a) + \gamma \max(0, y_h^\gamma - y_b).\end{aligned}$$

As $\gamma \to \infty$, μ^γ and μ_h^γ play the role of the measure Lagrange multipliers corresponding to state constraints in the limit problem (2.3) (compare with [10, Thm. 4.1, Rem. 4.1]). Moreover we abbreviate

$$\lambda^\gamma := \lambda_a^\gamma + \lambda_b^\gamma \quad \text{and} \quad \lambda_h^\gamma := \lambda_{a,h}^\gamma + \lambda_{b,h}^\gamma.$$

Theorem 5.2. *Let (u^γ, y^γ) and (u_h^γ, y_h^γ) be the solutions of the optimal control problems (3.1) and (4.1) with corresponding adjoint states p^γ, p_h^γ and multipliers associated to the control and state constraints $\lambda^\gamma, \lambda_h^\gamma, \mu^\gamma, \mu_h^\gamma$. Then*

$$\begin{aligned}&2(J(y^\gamma, u^\gamma) - J_h(y_h^\gamma, u_h^\gamma)) =\\ &\rho^{p^\gamma}(y^\gamma - i_h y^\gamma) + \rho^{y^\gamma}(p^\gamma - i_h p^\gamma) + (\mu^\gamma + \mu_h^\gamma, y_h^\gamma - y^\gamma) + (\lambda^\gamma + \lambda_h^\gamma, u_h^\gamma - u^\gamma).\end{aligned} \tag{5.1}$$

Proof. For ease of exposition, we omit the superscript γ in this proof for the quantities $y^\gamma, u^\gamma, p^\gamma, \lambda^\gamma, \mu^\gamma$ and their discrete counterparts. We have

$$\begin{aligned}
&2(J(y,u) - J_h(y_h,u_h)) \\
=\ &\alpha((u-u_d)+(u_h-u_d),(u-u_d)-(u_h-u_d)) \\
&+((y-y_d)+(y_h-y_d),(y-y_d)-(y_h-y_d)) \\
=\ &\alpha(u_h-u_d,u)+(-\alpha(u-u_d),u_h-u)+(-\alpha(u_h-u_d),u_h) \\
&+(y_h-y_d,y)+a(y,p)-a(y_h,p_h)-a(y_h,p) \\
&+(\mu_h,y)-(\mu,y)+(\mu_h,y_h)+(\mu,y_h) \\
&-(\mu_h,y).
\end{aligned}$$

For the last step, the adjoint equation was used three times and a zero was added. The last four terms can be summed up to $(\mu+\mu_h, y_h - y)$. The term $(y_h - y_d, y) + (\mu_h, y)$ already belongs to the dual residual, while $-a(y_h, p)$ belongs to the primal residual. The remaining both bilinear forms with a are expressed by using the both primal equations. Furthermore $(p_h, u+f) - a(y, p_h) = 0$ is added to the equation. We obtain:

$$\begin{aligned}
&2(J(y,u) - J_h(y_h,u_h)) \\
=\ &-a(y_h,p) \\
&+(y_h-y_d,y)+(\mu_h,y)-a(y,p_h) \\
&+(\mu+\mu_h,y_h-y) \\
&+\alpha(u_h-u_d,u)+(-\alpha(u-u_d),u_h-u)+(-\alpha(u_h-u_d),u_h) \\
&+(-p,-u-f)+(-p_h,u_h+f) \\
&+(p_h,u+f) \\
&+(u_h,p)-(p,u_h) \\
=\ &-a(y_h,p)+(u_h+f,p) \\
&-a(y,p_h)+(y_h-y_d,y)+(\mu_h,y) \\
&+(\mu+\mu_h,y_h-y) \\
&+(\alpha(u_h-u_d)+p_h,u) \\
&+(-\alpha(u-u_d)-p,u_h-u)+(-\alpha(u_h-u_d)-p_h,u_h) \\
=\ &\rho^y(p)+\rho^p(y)+(\mu+\mu_h,y_h-y) \\
&+\underbrace{(\alpha(u_h-u_d)+p_h+\lambda_h,u)}_{=0}-(\lambda_h,u)+(\lambda,u_h-u)+(\lambda_h,u_h) \\
=\ &\rho^y(p)+\rho^p(y)+(\mu+\mu_h,y_h-y)+(\lambda+\lambda_h,u_h-u).
\end{aligned}$$

Let us emphasize that in last intermediate step due to variational discretization the residual for the control vanishes. Because of Galerkin orthogonality of the error in the state and costate equation we could subtract arbitrary functions $i_h p$ and $i_h y \in V_h$ within the residuals ρ^y and ρ^p and end up with the assertion. □

Let us now define the elementwise residuals

$$\begin{aligned} R_{|T}^{y_h^\gamma} &:= u_h^\gamma + f - \mathcal{A} y_h^\gamma, \\ R_{|T}^{p_h^\gamma} &:= y_h^\gamma - y_d - \mathcal{A}^* p_h^\gamma, \\ R_{|T}^{p^\gamma} &:= y^\gamma - y_d - \mathcal{A}^* p^\gamma, \end{aligned}$$

and the edge residuals

$$\begin{aligned} r_{|\Gamma}^{y_h^\gamma} &:= \begin{cases} \frac{1}{2}\nu \cdot [\nabla y_h^\gamma \cdot (a_{ij})], & \Gamma \subset \partial T \setminus \partial\Omega \\ \nu \cdot (\nabla y_h^\gamma \cdot (a_{ij})), & \Gamma \subset \partial\Omega \end{cases}, \\ r_{|\Gamma}^{p_h^\gamma} &:= \begin{cases} \frac{1}{2}\nu \cdot [(a_{ij}) \nabla p_h^\gamma], & \Gamma \subset \partial T \setminus \partial\Omega \\ \nu \cdot ((a_{ij}) \nabla p_h^\gamma + p_h^\gamma b), & \Gamma \subset \partial\Omega \end{cases}, \\ r_{|\Gamma}^{p^\gamma} &:= \begin{cases} \frac{1}{2}\nu \cdot [(a_{ij}) \nabla p^\gamma], & \Gamma \subset \partial T \setminus \partial\Omega \\ \nu \cdot ((a_{ij}) \nabla p^\gamma + p^\gamma b), & \Gamma \subset \partial\Omega \end{cases}. \end{aligned}$$

Here $[\cdot]$ denotes the jump across the inter-element edge Γ. Now by integration by parts we can localize the error representation (5.1) by

$$\begin{aligned} 2(J(y^\gamma, u^\gamma) - J_h(y_h^\gamma, u_h^\gamma)) = \sum_{T \in \mathcal{T}_h} & (y^\gamma - y_h^\gamma, R_{|T}^{p_h^\gamma})_T - (y^\gamma - y_h^\gamma, r_{\partial T}^{p_h^\gamma})_{\partial T} \\ & + (R_{|T}^{y_h^\gamma}, p^\gamma - i_h p^\gamma)_T - (r_{|\partial T}^{y_h^\gamma}, p^\gamma - i_h p^\gamma)_{\partial T} \\ & + (y^\gamma - y_h^\gamma, R_{|T}^{p^\gamma})_T - (y^\gamma - y_h^\gamma, r_{|\partial T}^{p^\gamma})_{\partial T} \\ & + (\lambda^\gamma + \lambda_h^\gamma, u_h^\gamma - u^\gamma)_T. \end{aligned}$$

Since this localized sum still contains unknown quantities, we make use of local higher order approximation ([1, Sec. 5.1]) which has shown to be a successful heuristic technique for a posteriori error estimation. More precisely we take the local higher order quadratic interpolant operator $i_{2h}^{(2)} : V_h \to P_2(T)$ for some $T \in \mathcal{T}_h$ as already introduced in [10] for $d = 2$. The technique for computing $i_{2h}^{(2)} v_h$ for some $v_h \in V_h$ can easily be carried over to three space dimensions. However this is supposed to be numerically expensive. In order to derive a computable estimator we now replace the unknown functions y^γ and p^γ in (5.1) by $i_{2h}^{(2)} y_h^\gamma$ and $i_{2h}^{(2)} p_h^\gamma$. Since $u^\gamma = \Pi_{[u_a, u_b]} \left(-\frac{1}{\alpha} p^\gamma + u_d\right)$ holds, a reasonable locally computable approximation is

$$\tilde{u}^\gamma = \Pi_{[u_a, u_b]} \left(-\frac{1}{\alpha} i_{2h}^{(2)} p_h^\gamma + u_d \right)$$

as already suggested in [27]. Similarly for $\lambda^\gamma = -p^\gamma - \alpha(u^\gamma - u_d)$ we locally compute

$$\tilde{\lambda}^\gamma = -i_{2h}^{(2)} p_h^\gamma - \alpha \left(\tilde{u}^\gamma - u_d \right)$$

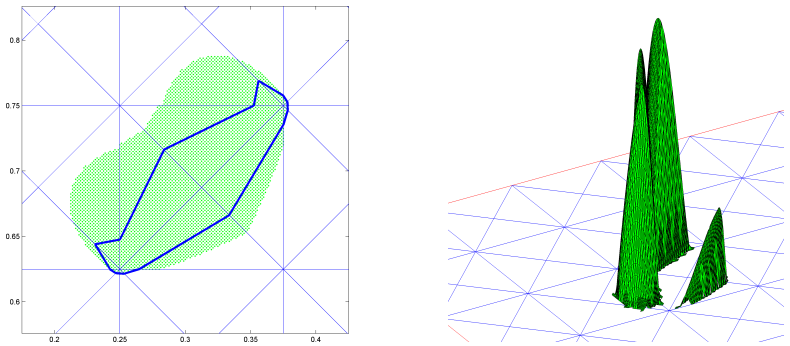


Figure 1: u_a active set for Example 1: blue by u_h^γ, green by $\tilde{u}^\gamma$ (left), integrand $(\tilde{\lambda}^\gamma + \lambda_h^\gamma)(u_h^\gamma - \tilde{u}^\gamma)$ with support on symmetric difference of active sets (right).

instead.

The estimator η^γ now reads

$$\eta^\gamma = \sum_{T \in \mathcal{T}_h} \eta_T^\gamma,$$

where

$$\begin{aligned} 2\eta_T^\gamma =& (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, R_{|T}^{p_h^\gamma})_T - (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, r_{|\partial T}^{p_h^\gamma})_{\partial T} \\ &+ (R_{|T}^{y_h^\gamma}, i_{2h}^{(2)} p_h^\gamma - p_h^\gamma)_T - (r_{|\partial T}^{y_h^\gamma}, i_{2h}^{(2)} p_h^\gamma - p_h^\gamma)_{\partial T} \\ &+ (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, R_{|T}^{i_{2h}^{(2)} p_h^\gamma})_T - (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, r_{|\partial T}^{i_{2h}^{(2)} p_h^\gamma})_{\partial T} \\ &+ (\tilde{\lambda}^\gamma + \lambda_h^\gamma, u_h^\gamma - \tilde{u}^\gamma)_T. \end{aligned}$$

While for the other quantities in η_T^γ quadrature rules of moderate order are suited, one has to take care for the last term

$$(\tilde{\lambda}^\gamma + \lambda_h^\gamma, u_h^\gamma - \tilde{u}^\gamma)_T = \int_T (\tilde{\lambda}^\gamma + \lambda_h^\gamma)(u_h^\gamma - \tilde{u}^\gamma). \tag{5.2}$$

The integrand is continuous but has a support within the symmetric difference of the control active set of the variational discrete solution and the locally improved quantities. Such a situation is depicted in Figure 1. One recognizes that $\tilde{u}^\gamma$ keeps the activity structure as u_h^γ has but smoothes the control active boundary towards the exact control active boundary. The kidney-shaped green area resolves the true control active set from Example 1 already very well even on a coarse mesh (compare also Figure 2 (right)). Finally for computing (5.2) we just provide the integrand and a desired tolerance and apply an adaptive quadrature routine given in [28, Algo. 31] for triangles containing the boundary of the control active set.

In order to study the efficiency of our implemented estimator, we define the effectivity of the estimator as

$$I_{\text{eff}} := \frac{J(y^\gamma, u^\gamma) - J_h(y_h^\gamma, u_h^\gamma)}{\eta^\gamma}.$$

Remark 5.3. *Let us remark that the adjoint variable p admits less regularity at state active sets such that higher order interpolation is not completely satisfying. However this circumstance only leads to local higher weights in the estimator and therefore reasonably suggests to refine at those regions. The efficiency of the estimator is not affected as we are going to see in the numerical experiments. Another thinkable heuristic technique to derive a computable approximation for $p^\gamma - i_h p^\gamma$ is to substitute the best known object p_h^γ for p^γ and compute $p_h^\gamma - p_h^\gamma(x_T)$, where x_T denotes the barycenter of the element T.*

Since the analytic solutions of the numerical examples are not known, we approximate $J(y^\gamma, u^\gamma)$ by $J_h(y_h^\gamma, u_h^\gamma)$ computed on a very fine mesh via the expression

$$\begin{aligned} J_h(y_h^\gamma, u_h^\gamma) = \frac{1}{2}\mathbf{y}^{\gamma T}\mathbf{M}\mathbf{y}^\gamma - \mathbf{y}^{\gamma T}\mathbf{M}\mathbf{y_d} + \frac{1}{2}\int_\Omega y_d^2 + \frac{1}{2\alpha}\mathbf{p}^{\gamma T}\mathbf{M_i^p}\mathbf{p}^\gamma \\ + \frac{\alpha}{2}(\mathbf{u_a} - \mathbf{u_d})^T\mathbf{M_a^p}(\mathbf{u_a} - \mathbf{u_d}) + \frac{\alpha}{2}(\mathbf{u_b} - \mathbf{u_d})^T\mathbf{M_b^p}(\mathbf{u_b} - \mathbf{u_d}). \quad (5.3) \end{aligned}$$

6 Numerical experiments

Based on the previous error estimations and the semi-smooth Newton solvers described earlier, we design an adaptive finite element algorithm to solve (4.1). The algorithm consists in performing cycles of the form

$$\text{Solve} \Longrightarrow \text{Estimate} \Longrightarrow \text{Mark} \Longrightarrow \text{Refine}.$$

In the Mark step, elements are selected according to a bulk-type criterion [9]. We select, for fixed specified $0 < \theta_i < 1$ ($i \in \{1, 2, 3\}$) the set $\mathcal{M} = \cup_{i=1}^3 \mathcal{M}_i \subset \mathcal{T}_h$ such that

$$\begin{aligned} \theta_1 \left| \sum_{T \in \mathcal{T}_h} \tau_{\mathring{T}} \right| &\leq \left| \sum_{T \in \mathcal{M}_1} \tau_{\mathring{T}} \right|, \\ \theta_2 \left| \sum_{T \in \mathcal{T}_h} \tau_{\partial T} \right| &\leq \left| \sum_{T \in \mathcal{M}_2} \tau_{\partial T} \right|, \\ \theta_3 \left| \sum_{T \in \mathcal{T}_h} \tau_{\lambda} \right| &\leq \left| \sum_{T \in \mathcal{M}_3} \tau_{\lambda} \right|, \end{aligned}$$

where the local quantities $\tau_{\mathring{T}}$, $\tau_{\partial T}$ and τ_λ are defined by

$$
\begin{aligned}
2\tau_{\mathring{T}} &:= (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, R_{|T}^{p_h^\gamma})_T + (R_{|T}^{y_h^\gamma}, i_{2h}^{(2)} p_h^\gamma - p_h^\gamma)_T + (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, R_{|T}^{i_{2h}^{(2)} p_h^\gamma})_T, \\
2\tau_{\partial T} &:= (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, r_{|\partial T}^{p_h^\gamma})_{\partial T} + (r_{|\partial T}^{y_h^\gamma}, i_{2h}^{(2)} p_h^\gamma - p_h^\gamma)_{\partial T} + (i_{2h}^{(2)} y_h^\gamma - y_h^\gamma, r_{|\partial T}^{i_{2h}^{(2)} p_h^\gamma})_{\partial T}, \\
2\tau_\lambda &:= (\tilde{\lambda}^\gamma + \lambda_h^\gamma, u_h^\gamma - \tilde{u}^\gamma)_T.
\end{aligned}
$$

Flagging elements in such three separate steps has the advantage of properly handling possible scaling difference between jump, element and multiplier contributions in particular if the regularization parameter γ tends to infinity. Once all the elements to be refined are marked, a new finer mesh is generated using the longest bisection rule implemented within the Matlab pde-toolbox. To assess the performance of the overall adaptive finite element algorithm we compare it with a uniform mesh refinement by monitoring values of the objective functional versus the numbers of degrees of freedom $N_{dof} := np$. Uniform refinement levels and the corresponding number of nodes np, number of triangles nt and grid size h are documented in Table 1.
In the sequel we provide the documentation for two numerical examples. For both examples, the analytic solution is not known, so for obtaining the efficiency index we compute a reference solution on the finest grid in Table 1 and hence an approximation of $J(y^\gamma, u^\gamma)$. The semi-smooth Newton solver converges generally in few iterations provided an appropriate update strategy is used for the regularization coefficient. In our experiments we use a simple continuation method. However more sophisticated techniques might be used (see for instance [15]). We stop the semi-smooth Newton solver as soon as

$$
\|G^\gamma(\mathbf{x}_n^\gamma)\|_2 \leq \epsilon_{\text{rel}} \|G^\gamma(\mathbf{x}_0^\gamma)\|_2 + \epsilon_{\text{abs}}, \quad n = 1, \ldots, n_{max},
$$

for some user-specified maximum number of iterations n_{max} and tolerances ϵ_{rel} and ϵ_{abs}. In our experiments we used $n_{max} = 100$. The absolute and relative tolerances are chosen more and more stringent as $\gamma \to \infty$ such that the final values are

$$
\epsilon_{\text{rel}} = 10^{-12}, \quad \epsilon_{\text{abs}} = 10^{-8}.
$$

Example 1

As a first example we consider problem (2.3) with data

$$
\begin{gathered}
\Omega = (0,\,1)^2, \quad \mathcal{A} = -\Delta + \mathrm{Id}, \quad y_d = \sin(2\pi x_1)\sin(2\pi x_2), \quad f = u_d = 0, \\
u_a = -30, \quad u_b = 30, \quad y_a = -0.55, \quad y_b = 0.55, \quad \alpha = 10^{-4}.
\end{gathered}
$$

Its numerical solution in terms of $-\frac{1}{\alpha} p_h^\gamma$ as well as the optimal state y_h^γ is displayed in Figure 2 for $\gamma = 10^{14}$ on the mesh $l = 14$. The projection

l	np	nt	h
1	81	128	0.17678
2	145	256	0.12500
3	289	512	0.08839
4	545	1024	0.06250
5	1089	2048	0.04419
6	2113	4096	0.03125
7	4225	8192	0.02210
8	8321	16384	0.01563
9	16641	32768	0.01105
10	33025	65536	0.00781
11	66049	131072	0.00552
12	131585	262144	0.00391
13	263169	524288	0.00276
14	525313	1048576	0.00195

Table 1: Mesh parameters for Example 1 (global refinement).

of $-\frac{1}{\alpha}p_h^\gamma$ onto $[u_a, u_b]$ corresponds to the optimal control u_h^γ which represents together with y_h^γ our best approximation to the solution of (3.1). The boundaries of the control active sets are depicted as solid lines, while the state active sets are coded as star and cross markers. The color blue corresponds to the lower bound while the color red highlights the upper bound. Now by using the expression (5.3) we get $J(y^\gamma, u^\gamma) \approx 0.0375586175$. In Table 2 we depict the efficiency coefficient and the convergence history of the quantity of interest. Notice that the values of the efficiency coefficient are close to 1 which illustrate the good performance of our error estimator. A comparison between our adaptive finite element algorithm and a uniform mesh refinement in terms of number of degrees of freedom is reported in Figure 3. The adaptive refinement process performs well even though the benefit in this example is not big since the characteristic features of the optimal solution occupy an important area of the computational domain as illustrated by the adapted grid in Figure 3. Our motivation for including this example is to illustrate the variational discretization effect on the mesh refinement process. If variational discretization for the control would not have been used, one would expect also some refinement at the boundary of the control active set.

Example 2

In this example we set the computational domain to $\Omega = (-1, 1) \times (-1, 1)$ and $\mathcal{A} = -\Delta + \mathrm{Id}$. We take $\alpha = 10^{-3}$ and $u_d = y_d = (-3x_1^4 + 4x_1^3)\chi_{[0,1]}(x_1)$, where χ_A denotes the characteristic function of a set A. Furthermore we

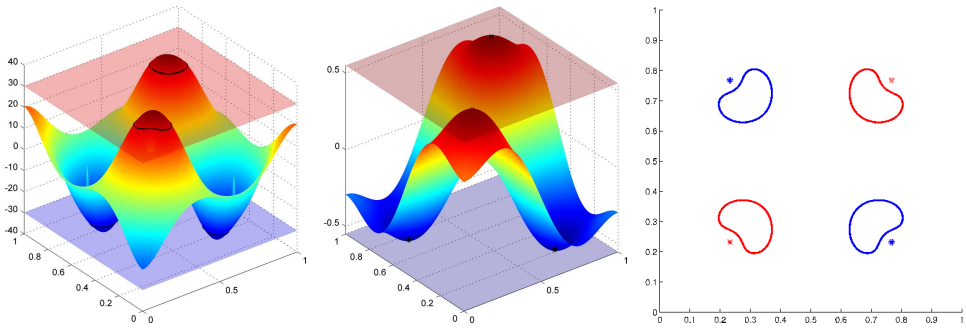


Figure 2: $u_a, -\frac{1}{\alpha}p_h^\gamma, u_b$ (left), $y_a \le y_h^\gamma \le y_b$ (middle) and active sets (right) for Example 1 and $l = 14$.

k	np	$J(y^\gamma, u^\gamma) - J_h(y_h^\gamma, u_h^\gamma)$	I_{eff}
1	81	$4.275 \cdot 10^{-3}$	1.622
2	140	$2.259 \cdot 10^{-3}$	1.543
3	200	$1.380 \cdot 10^{-3}$	1.390
4	301	$7.904 \cdot 10^{-4}$	1.119
5	470	$5.369 \cdot 10^{-4}$	1.176
6	657	$3.643 \cdot 10^{-4}$	1.269
7	948	$2.343 \cdot 10^{-4}$	1.127
8	1405	$1.790 \cdot 10^{-4}$	1.187
9	2075	$1.133 \cdot 10^{-4}$	1.227
10	3123	$7.148 \cdot 10^{-5}$	1.144
11	4469	$5.115 \cdot 10^{-5}$	1.137
12	6775	$3.281 \cdot 10^{-5}$	1.172
13	9799	$2.360 \cdot 10^{-5}$	1.165
14	14305	$1.546 \cdot 10^{-5}$	1.181
15	20977	$1.161 \cdot 10^{-5}$	1.186
16	30445	$7.763 \cdot 10^{-6}$	1.256
17	44958	$5.524 \cdot 10^{-6}$	1.289
18	63389	$3.996 \cdot 10^{-6}$	1.290

Table 2: Adaptive refinement for Example 1 (bulk criterion, $\theta_i = 0.6$).

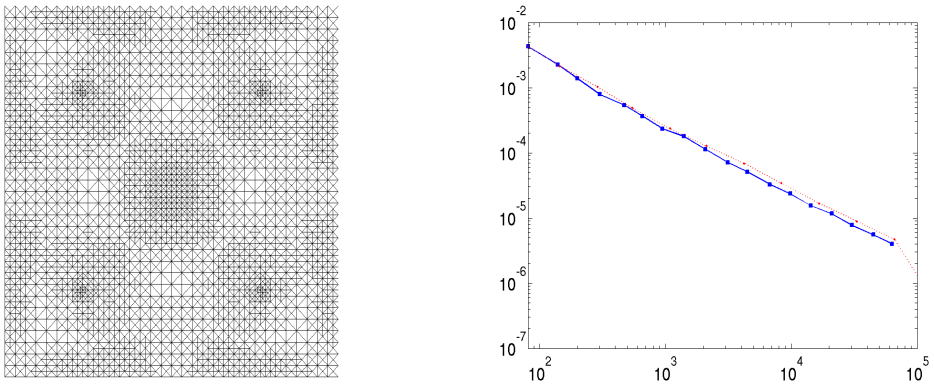


Figure 3: Adaptive mesh for $k = 10$ (left), comparison of error decrement in the quantity of interest (right) for Example 1.

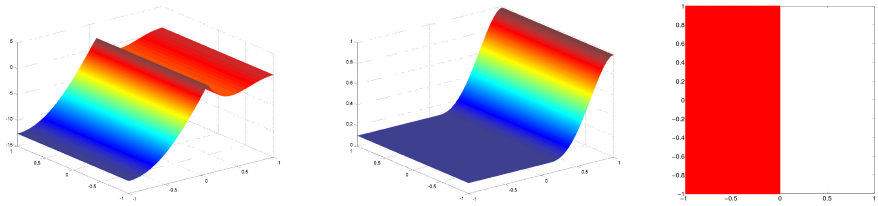
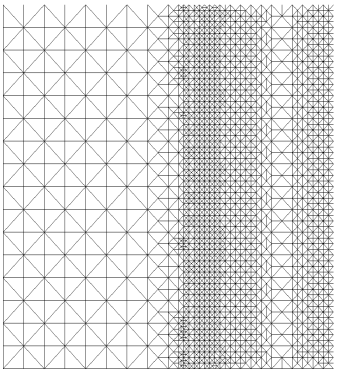


Figure 4: $-\frac{1}{\alpha}p_h^\gamma + u_{d,h}$ (left), y_h^γ (middle) and active sets intersection (right) for Example 2 and $l = 14$.

fix $f = (36x_1^2 - 24x_1)\chi_{[0,1]}(x_1)$ and the bounds $0.1 \leq u \leq 2$, $0.1 \leq y \leq 2$. This data is chosen such that the optimal control and optimal state exhibit active sets whose intersection is not empty (see Figure 4). An approximation $J(y^\gamma, u^\gamma) \approx 0.0130624289$ of the optimal quantity of interest is computed on the mesh level $l = 14$. We notice that the globally refined meshes have the same numbers of nodes and elements as denoted in Table 1 for Example 1 but due to the enlarged domain the doubled mesh parameter h. Figure 4 displays the corresponding state y_h^γ and the finite element quantity $-\frac{1}{\alpha}p_h^\gamma + u_{d,h}$. Throughout our computations we take $\gamma = 10^8$. The history of the efficiency coefficients as well as the convergence of the quantities of interest are reported in Table 3. As for the previous example we notice the high accuracy of our error estimator illustrated by the fact that the efficiency coefficient stays close to 1 during the adaptive procedure. The performance of our adaptive algorithm is illustrated in Figure 5. In the same figure (left) we clearly observe that the characteristic features of the solution are tracked on the adapted grid.

k	np	$J(y^\gamma, u^\gamma) - J_h(y_h^\gamma, u_h^\gamma)$	I_{eff}
1	289	$2.482 \cdot 10^{-4}$	1.261
2	330	$1.805 \cdot 10^{-4}$	1.128
3	411	$1.635 \cdot 10^{-4}$	1.307
4	483	$8.344 \cdot 10^{-5}$	1.674
5	604	$5.544 \cdot 10^{-5}$	1.215
6	758	$4.051 \cdot 10^{-5}$	1.000
7	993	$3.370 \cdot 10^{-5}$	1.155
8	1261	$2.463 \cdot 10^{-5}$	1.198
9	1628	$1.684 \cdot 10^{-5}$	1.202
10	2287	$1.292 \cdot 10^{-5}$	1.140
11	3110	$9.290 \cdot 10^{-6}$	1.155
12	4242	$6.399 \cdot 10^{-6}$	1.167
13	5526	$4.136 \cdot 10^{-6}$	1.168
14	7942	$3.184 \cdot 10^{-6}$	1.109
15	11281	$2.268 \cdot 10^{-6}$	1.121
16	15531	$1.537 \cdot 10^{-6}$	1.144
17	20867	$1.041 \cdot 10^{-6}$	1.148
18	30498	$7.828 \cdot 10^{-7}$	1.095

Table 3: Adaptive refinement for Example 2 (bulk criterion, $\theta_i = 0.5$).



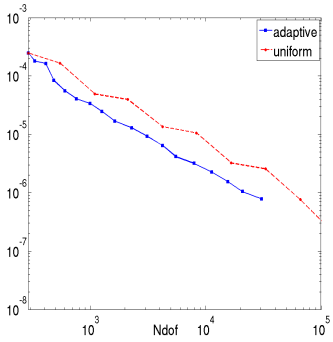


Figure 5: Adaptive mesh for $k = 10$ (left), comparison of error decrement in the quantity of interest (right) for Example 2.

Acknowledgements

The authors thank Michael Hintermüller from Humboldt Universität zu Berlin and Michael Hinze from Universität Hamburg for many helpful discussions. We further greatfully acknowledge support for meetings in Graz and Hamburg by the DFG Priority Program 1253 through grants DFG06-381, DFG06-382 and by the Austrian Ministry of Science and Research and the Austrian Science Fund FWF under START-grant Y305 "Interfaces and Free Boundaries".

References

[1] R. Becker and R. Rannacher: *An optimal control approach to a posteriori error estimation in finite element methods.* Acta Numerica **10**, 1–102 (2001).

[2] O. Benedix and B. Vexler: *A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints.* Comput Optim Appl, DOI 10.1007/s10589-008-9200-y (2008).

[3] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch: *A comparison of a Moreau-Yosida based active strategy and interior point methods for constrained optimal control problems.* SIAM J. Optim. **11**(2), 495–521 (2000).

[4] M. Bergounioux and K. Kunisch: *On the structure of the Lagrange multiplier for state-constrained optimal control problems.* Syst. Control Lett. **48**, 169–176 (2002).

[5] M. Bergounioux and K. Kunisch: *Primal-dual strategy for state-constrained optimal control problems.* Comput Optim Appl **22**(2), 193–224 (2002).

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

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

[8] K. Deckelnick and M. Hinze: *A finite element approximation to elliptic control problems in the presence of control and state constraints.* Hamburger Beiträge zur Angewandten Mathematik, Universtität Hamburg, Preprint No. HBAM2007-01 (2007).

[9] D. Dörfler: *A convergent adaptive algorithm for Poisson's equation.* SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996).

[10] A. Günther and M. Hinze: *A posteriori error control of a state constrained elliptic control problem.* J. Numer. Math. **16**(4), 307–322 (2008).

[11] M. Hintermüller and R.H.W. Hoppe: *Goal-oriented adaptivity in control constrained optimal control of partial differential equations.* SIAM J. Control Optim. **47**(4), 1721–1743 (2008).

[12] M. Hintermüller and R.H.W. Hoppe: *Goal-oriented adaptivity in pointwise state constrained optimal control of partial differential equations.* Institut für Mathematik, Universität Augsburg, Preprint No. 2009-16 (2009).

[13] M. Hintermüller and R.H.W. Hoppe: *Goal oriented mesh adaptivity for mixed control-state constrained elliptic optimal control problems.* Institut für Mathematik, Universität Augsburg, Preprint No. 2008-20 (2008).

[14] M. Hintermüller, K. Ito and K. Kunisch: *The primal-dual active set strategy as a semismooth Newton method.* SIAM J. Optim. **13**(3), 865–888 (2003).

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

[16] 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).

[17] M. Hinze: *A variational discretization concept in control constrained optimization: the linear-quadratic case.* Comput Optim Appl **30**(1), 45–63 (2005).

[18] R.H.W. Hoppe and M. Kieweg: *Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems.* Comput Optim Appl, DOI 10.1007/s10589-008-9195-4 (2008).

[19] R. Horn and C. R. Johnson: Matrix Analysis. Cambridge University Press, New York (1985).

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

[21] R. Mifflin: *Semismooth and semiconvex functions in constrained optimization.* SIAM J. Control Optim. **15**(6), 957–972 (1977).

[22] L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Math. Program. **58**(3), 353–367 (1993).

[23] A. Schiela: *State constrained optimal control problems with states of low regularity.* SIAM J. Control Optim. **48**(4), 2407–2432 (2009).

[24] A. Schiela and A. Günther: *Interior point methods in function space for state constraints - inexact Newton and adaptivity.* DFG Schwerpunktprogramm 1253, Preprint No. SPP1253-08-06 (2009).

[25] A. Shapiro: *On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints.* SIAM J. Optim. **7**(2), 508–518 (1997).

[26] F. Tröltzsch: *Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints.* SIAM J. Optim. **15**(2) 616–634 (2005).

[27] B. Vexler and W. Wollner: *Adaptive finite elements for elliptic optimization problems with control constraints.* SIAM J. Control Optim. **47**(1), 509–534 (2008).

[28] W. Vogt: *Adaptive Verfahren zur numerischen Quadratur und Kubatur.* IfMath TU Ilmenau, Preprint No. M 1/06 (2006).

[29] M. Weiser: *Interior point methods in function space.* SIAM J. Control Optim. **44**(5), 1766–1786 (2005).

[30] W. Wollner: *A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints.* Comput Optim Appl, DOI 10.1007/s10589-008-9209-2 (2008).

[31] J. Zowe and S. Kurcyusz: *Regularity and stability for the mathematical programming problem in Banach spaces.* Appl. Math. Optim. **5**(1), 49–62 (1979).