www.oeaw.ac.at

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

A-posteriori error estimates for optimal control problems with state and control constraints

A. Rösch, D. Wachsmuth

RICAM-Report 2010-08

A-POSTERIORI ERROR ESTIMATES FOR OPTIMAL CONTROL PROBLEMS WITH STATE AND CONTROL CONSTRAINTS

ARND RÖSCH[1], DANIEL WACHSMUTH[2],

Abstract. We discuss the full discretization of an elliptic optimal control problem with pointwise control and state constraints. We provide the first reliable a-posteriori error estimator that contains only computable quantities for this class of problems. Moreover, we show, that the error estimator converges to zero if one has convergence of the discrete solutions to the solution of the original problem. The theory is illustrated by numerical tests.

1. Introduction. In this paper we consider the optimal control problem of minimizing the cost functional J given by

$$J(y,u) = \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \tag{1.1}$$

subject to

$$\begin{aligned} -\Delta y + y &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{1.2}$$

and

$$u_a \le u \le u_b \quad \text{a.e. in } \Omega, \tag{1.3}$$

$$y_a \le y \le y_b \text{ in } \Omega. \tag{1.4}$$

Let us define the set of admissible controls by

$$U_{ad} := \{u \in L^2(\Omega) : u_a \le u \le u_b \text{ a.e. in } \Omega\}.$$

In order to guarantee existence and regularity of solutions, we assume for the whole paper:

Assumption 1. *$\Omega \subset \mathbb{R}^n$, $n \in \{2,3\}$, is a convex polygonal domain; $u_a < u_b$ are constants; $y_a, y_b \in C(\bar\Omega)$, $y_d \in L^2(\Omega)$, $\alpha \ge 0$.*

Note that our problem and also the discretized counterparts are strictly convex. Therefore, one has only to show the existence of a feasible point to get existence and uniqueness of an optimal control $\bar u$ and state $\bar y$. Under a Slater type assumption we will state such a result in Theorem 2.3. Using the same assumption, one obtains existence and uniqueness of solutions $(\bar y_h, \bar u_h)$ for sufficiently fine discretizations.

The distance of the solution $(\bar y_h, \bar u_h)$ to $(\bar y, \bar u)$ can be estimated a-priori as $\|\bar u_h - \bar u\|_{L^2(\Omega)} + \|\bar y_h - \bar y\|_{H^1(\Omega)} \le Ch^{2-\frac{n}{2}-\varepsilon}$, $\varepsilon \ge 0$, see e.g. [8, 17]. In order to generate adaptively refined meshes, computable and localized a-posteriori error estimators are inevitable. For work on dual-weighted residual error indicators for state constrained problems we refer to [2, 10, 21]. The main drawback of this method is that the resulting

[1]Universität Duisburg-Essen, Fachbereich Mathematik, Forsthausweg 2, D-47057 Duisburg, Germany, arnd.roesch@uni-due.de

[2]Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria, daniel.wachsmuth@oeaw.ac.at

error estimators are not computable since they depend on the unknown solution of the continuous problem. The only work in context of residual error estimates is due to [14]. However, this error estimator contains also an uncomputable part that involves the Lagrange multipliers of the continuous problem. All cited papers contain different heuristics to replace uncomputable quantities by computable ones. Our main goal is to derive *computable* error bounds.

The main difficulty to obtain a computable error bound is that the errors in the dual quantities $\|\bar{p}_h - \bar{p}\|$ and $\|\bar{\mu}_{i,h} - \bar{\mu}_i\|$ cannot be majorized by $\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}$. Hence, we completely avoid the use of the Lagrange multipliers and of the first-order necessary optimality system of Theorem 2.3. In our main result Theorem 3.9 we describe detailed all the different error contributions.

Moreover, we prove that the error estimate converges to zero if the discrete quantities converge in a certain sense, see Section 4. Such a result does not seem to be available even for pure control constrained problems, see the comments at the end of that section.

Let us shortly describe the structure of our paper. Section 2 contains basic properties of the optimal control problems and its discrete counterparts. A-posteriori error estimates are derived in Section 3. The behavior of the error bound when the discrete solutions converge is studied in Section 4. Numerical experiments are presented in Section 5.

2. Optimality system and discretization.

2.1. The undiscretized problem. First, let us define the notion of weak solutions of the state equation (1.2). A function $y \in H_0^1(\Omega)$ is called weak solution of (1.2) if it satisfies the the weak formulation

$$a(y, v) = (u, v)_{L^2(\Omega)} \qquad \forall v \in H_0^1(\Omega), \tag{2.1}$$

where the bilinear form a is defined as

$$a(y, v) = (\nabla y, \nabla v)_{L^2(\Omega)^n} + (y, v)_{L^2(\Omega)}.$$

Let us define the operators $A : V := H^2(\Omega) \cap H_0^1(\Omega) \to L^2(\Omega)$ by $A = -\Delta + I$ and its dual $A^* : L^2(\Omega) \to V^*$ by

$$(Aw, v) = \langle w, A^* v\rangle \quad \forall w \in V,\, v \in L^2(\Omega).$$

LEMMA 2.1. *For each control $u \in L^2(\Omega)$ the state equation* (1.2) *admits a unique weak solution $y \in H^2(\Omega)$, and the mapping $u \mapsto y(u)$ is continuous from $L^2(\Omega)$ to $H^2(\Omega)$, i.e. $\|y\|_{H^2(\Omega)} \le C_0 \|u\|_{L^2(\Omega)}$.*
For the proof we refer to Grisvard [9].

Throughout the article we will assume the existence of a Slater point:

ASSUMPTION 2. *There exists $\hat{u} \in U_{ad}$ and $\tau \in \mathbb{R}$, $\tau > 0$, such that the associated state $\hat{y}$ satisfies $y_a + \tau \le \hat{y} \le y_b - \tau$.*

Please note, that this assumption implies that the state constraints cannot be active on Γ, i.e. it holds $y_a < -\tau$ and $\tau < y_b$ on Γ since $\hat{y} = 0$ on Γ due to the Dirichlet boundary conditions.

Additionally, Assumption 2 implies that the feasible set of the control problem is non-empty. Due to convexity, we get immediately the existence and uniqueness of solutions.

LEMMA 2.2. *Under Assumption 2, the optimal control problem* (1.1)–(1.4) *admits a unique solution* $(\bar{y}, \bar{u})$.

The solution of the optimal control problem can be characterized by means of first-order necessary optimality conditions. Due to Assumption 2, one can prove existence of Lagrange multipliers, see e.g. [5, 20]. In the following, let us denote by $M(\Omega) = C(\bar{\Omega})^*$ the space of regular Borel measures.

THEOREM 2.3. *Let* $(\bar{y}, \bar{u})$ *be a solution of the problem* (1.1)–(1.3). *Then there are* $\bar{\mu}_a, \bar{\mu}_b \in M(\Omega)$ *and* $\bar{p} \in W^{1,s}(\Omega)$, $s < \frac{n}{n-1}$, *such that the following system is satisfied, which consists of adjoint equation*

$$\begin{aligned} -\Delta \bar{p} &= \bar{y} - y_d - \bar{\mu}_a + \bar{\mu}_b && \text{in } \Omega, \\ \bar{p} &= 0 && \text{on } \Gamma = \partial\Omega, \end{aligned} \tag{2.2}$$

complementarity conditions for $\bar{\mu}_a$ *and* $\bar{\mu}_b$

$$\bar{\mu}_i \geq 0, \ \langle \bar{\mu}_i, \bar{y} - y_i \rangle = 0 \quad i \in \{a, b\}, \tag{2.3}$$

and the variational inequality

$$(\alpha \bar{u} + \bar{p}, \ u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}. \tag{2.4}$$

For the proof we refer to Casas [5].

2.2. Discretization. Let us fix the assumptions on the discretization of problem (1.1)-(1.4) by finite elements. First let us specify the notation of regular meshes. Each mesh $\mathcal{T}$ consists of closed cells T (for example triangles, tetrahedra, etc.) such that $\bar{\Omega} = \bigcup_{T \in \mathcal{T}} T$ holds. We assume that the mesh is regular in the following sense: for different cells $T_i, T_j \in \mathcal{T}$, $i \neq j$, the intersection $T_i \cap T_j$ is either empty or a node, an edge, or a face of both cells, i.e. hanging nodes are not allowed. Let us denote the size of each cell by $h_T = \operatorname{diam} T$ and define $h(\mathcal{T}) = \max_{T \in \mathcal{T}} h_T$. For each $T \in \mathcal{T}$, we define R_T to be the diameter of the largest ball contained in T.

We will work with a family of regular meshes $\mathcal{F} = \{\mathcal{T}_h\}_{h>0}$, where the meshes are indexed by their mesh size, i.e. $h(\mathcal{T}_h) = h$. We assume in addition that there exists a positive constant R such that

$$\frac{h_T}{R_T} \leq R$$

holds for all cells $T \in \mathcal{T}_h$ and all $h > 0$. With each mesh $\mathcal{T}_h \in \mathcal{F}$, we associate the finite-dimensional space $V_h \subset H_0^1(\Omega)$ that consists of polynomial finite element functions of degree $l \geq 1$.

Furthermore, let us denote by $U_h \subset L^2(\Omega)$ the corresponding control discretization. Here, we have the following three possibilities in mind: discretization of controls by piecewise constant or linear finite element functions, or the choice $U_h = L^2(\Omega)$, which corresponds to the so-called variational discretization introduced by Hinze [12].

Let us now introduce the discretized version of the optimal control problem (1.1)–(1.4). This problem is given as: Find $y_h \in V_h$ and $u_h \in U_h$ that minimize

$$\min J(y_h, u_h) = \frac{1}{2} \|y_h - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|u_h\|^2_{L^2(\Omega)} \tag{2.5}$$

subject to the discretized state equation

$$a(y_h, v_h) = (u_h, v_h)_{L^2(\Omega)} \qquad \forall v_h \in V_h, \tag{2.6}$$

the control constraints

$$u_a \le u_h \le u_b \quad \text{a.e. in } \Omega, \tag{2.7}$$

and the discrete state constraints

$$y_a(x_i) \le y_h(x_i) \le y_b(x_i) \text{ for } i = 1, .., K. \tag{2.8}$$

Here, we denoted by x_i, $i = 1 \dots K$, the nodes of the mesh $\mathcal{T}$. Before proving existence of solutions of this problem, we first state some auxiliary results

LEMMA 2.4. *Let y^h and y_h be the solution of (2.1) and (2.6) for the control u_h. Let $C_M > 1$ be given such that $\max_{T \in \mathcal{T}_h} h_T \le C_M \min_{T \in \mathcal{T}_h} h_T$ is satisfied. Then the following L^∞-error estimate holds*

$$\|y_h - y^h\|_{L^\infty(\Omega)} \le ch\|u\|_{L^2(\Omega)}.$$

For a proof, we refer to Braess [3].

Let us denote by Π_h the $L^2(\Omega)$-projection onto U_h for piecewise constant functions. For piecewise linear functions one can use a quasi-interpolation, see [18]. In the case of the variational discretization we take $\Pi_h = I$. In all three cases, this operator has the following known approximation property.

LEMMA 2.5. *There is a constant c_I independent of h such that*

$$\|u - \Pi_h u\|_{L^2(\Omega)} \le c_I h \|\nabla u\|_{L^2(\Omega)}$$

is fulfilled for all $u \in H^1(\Omega)$.

Regarding existence of solutions of the discrete optimal control problem, we have the following result.

LEMMA 2.6. *Let $C_M > 1$ be given and $\max_{T \in \mathcal{T}_h} h_T \le C_M \min_{T \in \mathcal{T}_h} h_T$. Then, there exists a mesh size $h_0 > 0$ such that for all $h \le h_0$ a feasible point of the discretized problem exists.*

Proof. We set $\hat{u}_h = \Pi_h \hat{u}$, where $\hat{u}$ is the Slater point from Assumption 2. The function $\hat{u}_h$ satisfies the control constraints. It remains to check the state constraints. Let us denote the solutions of the discrete and continuous state equation associated to $\hat{u}_h$ by $\hat{y}_h$ and $\hat{y}^h$, respectively. We find

$$\begin{aligned} \hat{y}_h(x_i) - y_a(x_i) &\ge \tau - |\hat{y}(x_i) - \hat{y}^h(x_i)| - |\hat{y}^h(x_i) - \hat{y}_h(x_i)| \\ &\ge \tau - \|\hat{y} - \hat{y}^h\|_{L^\infty(\Omega)} - \|\hat{y}^h - \hat{y}_h\|_{L^\infty(\Omega)} \end{aligned}$$

Here, $\|\hat{y} - \hat{y}^h\|_{L^\infty(\Omega)}$ becomes small for small h because of Lemmas 2.5 and 2.1. The term $\|\hat{y}^h - \hat{y}_h\|_{L^\infty(\Omega)}$ reflects the pointwise discretization error, which tends to zero for $h \to 0$ due to Lemma 2.4. Consequently, for sufficiently small h the discrete lower state constraints are fulfilled. Analogously, one shows that the upper state constraint of the discrete problem is satisfied for h small enough. Hence, the point $(\hat{y}_h, \hat{u}_h)$ is admissible for the discrete problem. ∎

As for the continuous problem, we get existence and uniqueness of solutions of the discrete problem.

LEMMA 2.7. *Let Assumption 2 be satisfied and* $\max_{T\in\mathcal{T}_h} h_T \le C_M \min_{T\in\mathcal{T}_h} h_T$. *Let the mesh-size satisfy* $h < h_0$. *Then the optimal control problem* (2.5)–(2.8) *admits a uniquely determined solution* $(\bar{y}_h, \bar{u}_h)$ *for all meshes with mesh-size* $h < h_0$.

Proof. Due to Lemma 2.6 the feasible set of the discrete problem is non-empty for $h < h_0$. By standard arguments one concludes the existence of a unique solution of this problem. □

Let us remark that the existence of a feasible point for the discrete problem is not guaranteed for arbitrary meshes with $h < h_0$. This is due to the fact that a-priori L^∞-error estimates are derived using inverse inequalities. Then the relation between minimal and maximal element size must be limited by some constant C_M. Another possibility are meshes for which a discrete maximum principle holds, see [7]. Usually, sequences of adaptive refined grids do not satisfy these requirements. However, the existence of a feasible point can be verified a-posteriori. If no feasible point exists, then one can refine the mesh in a suitable way to get a feasible point for the new mesh. We will use the following strategy. Our a-posteriori error estimator will contain a term that measures the $L^\infty(\Omega)$-norm of the violation of the state constraint, and thus should ensure the solvability of the discrete systems for the adaptively generated meshes.

ASSUMPTION 3. *In the sequel, we assume that the discrete optimal control problem admits a unique solution.*

Analogous to the continuous problem, one finds that the solution of the discrete problem can be characterized by a first-order optimality system.

THEOREM 2.8. *Let* $(\bar{y}_h, \bar{u}_h)$ *be a solution of the problem* (2.5)–(2.8). *Then there are* $\bar{\mu}_a, \bar{\mu}_b \in M(\Omega)$ *and* $\bar{p}_h \in V_h$, *such that the following system is satisfied, which consists of discrete adjoint equation*

$$a(v_h, \bar{p}_h) = (\bar{y}_h - y_d, v_h)_{L^2(\Omega)} + \langle -\bar{\mu}_{a,h} + \bar{\mu}_{b,h}, v_h\rangle \quad \forall v_h \in V_h, \tag{2.9}$$

complementarity condition

$$\bar{\mu}_{i,h} = \sum_{j=1}^{K} \bar{\mu}^j_{i,h}\delta(x_j),\ \bar{\mu}^j_{i,h} \ge 0,\ \bar{\mu}^j_{i,h}(\bar{y}_h(x_j) - y_i(x)) = 0,\ i \in \{a, b\}, j \in \{1, .., K\}, \tag{2.10}$$

and variational inequality

$$(\alpha\bar{u}_h + \bar{p}_h,\ u_h - \bar{u}_h) \ge 0 \quad \forall u_h \in U_h \cap U_{ad}.$$

Since the state constraints for the discrete problem were prescribed in the mesh nodes, the Lagrange multipliers $\bar{\mu}_{a,h}$ and $\bar{\mu}_{b,h}$ are positive linear combinations of Dirac measures concentrated in the mesh nodes. Due to this representation, the complementarity condition in (2.10) can be written as

$$\langle \bar{\mu}_{i,h},\ \bar{y}_h - y_i\rangle = 0 \quad i \in \{a, b\}.$$

3. A-posteriori error estimates. In this section, we will derive an upper bound for the error $\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}$. In order to avoid the difficulties mentioned in the introduction, we will not work with the first-order optimality system given by Theorem 2.3. We will work with the optimality of $\bar{u}$ instead, i.e. $J(\bar{y}, \bar{u}) \le J(y, u)$ for all (y, u) satisfying the constraints (1.2)–(1.4).

Let us suppose that we have computed the solution $(\bar{y}_h, \bar{u}_h, \bar{p}_h, \bar{\mu}_{a,h}, \bar{\mu}_{b,h})$ of the discrete problem. At first, we have to find a pair $(\tilde{y}, \tilde{u}) \approx (\bar{y}_h, \bar{u}_h)$ that is feasible for the continuous problem, which would give $J(\bar{y}, \bar{u}) \le J(\tilde{y}, \tilde{u})$.

3.1. Construction of feasible control. Let us first define an auxiliary state y^h as the weak solution of the state equation with right-hand side $\bar{u}_h$, i.e.

$$a(y^h, v) = (\bar{u}_h, v) \quad \forall v \in H_0^1(\Omega). \tag{3.1}$$

LEMMA 3.1. *Let $\hat{u}$ denote the Slater point as given by Assumption 2. Let y^h denote the solution of* (3.1). *Let us define the violation of the state constraints by*

$$e_{\mathrm{sc}} := \max(\|(\bar{y}_h - y_a)^-\|_{L^\infty(\Omega)},\ \|(y_b - \bar{y}_h)^-\|_{L^\infty(\Omega)}). \tag{3.2}$$

Then the state $\tilde{y} := (1-\sigma)y^h + \sigma\hat{y}$ associated to the control $\tilde{u} := (1-\sigma)\bar{u}_h + \sigma\hat{u}$ is admissible for the state constraints (1.4) *if σ is chosen as*

$$\sigma = \frac{\|\bar{y}_h - y^h\|_{L^\infty(\Omega)} + e_{\mathrm{sc}}}{\tau + \|\bar{y}_h - y^h\|_{L^\infty(\Omega)} + e_{\mathrm{sc}}}.$$

Proof. We find

$$\begin{aligned}
(1-\sigma)y^h + \sigma\hat{y} &\ge (1-\sigma)(\bar{y}_h - \|\bar{y}_h - y^h\|_{L^\infty(\Omega)}) + \sigma(y_a + \tau)\\
&\ge y_a + \tau + (1-\sigma)(\bar{y}_h - \|\bar{y}_h - y^h\|_{L^\infty(\Omega)} - y_a - \tau)\\
&\ge y_a + \tau + (1-\sigma)(-\|(\bar{y}_h - y_a)^-\|_{L^\infty(\Omega)} - \|\bar{y}_h - y^h\|_{L^\infty(\Omega)} - \tau)\\
&\ge y_a + \tau - (1-\sigma)(\tau + \|\bar{y}_h - y^h\|_{L^\infty(\Omega)} + e_{\mathrm{sc}}),
\end{aligned}$$

which implies $(1-\sigma)y^h + \sigma\hat{y} \ge y_a$ for $1 - \sigma \le \frac{\tau}{\tau + \|\bar{y}_h - y^h\|_{L^\infty(\Omega)} + e_{\mathrm{sc}}}$. An analogous discussion for the upper state constraint yields the claim. $\square$

With the notation of the previous lemma, we have

$$\tilde{u} - \bar{u}_h = \sigma(\hat{u} - \bar{u}_h), \tag{3.3}$$

which allows to estimate the difference $\|\tilde{u} - \bar{u}_h\|_{L^2(\Omega)}$ provided upper bounds for σ are available. The difference in the states $\tilde{y} - \bar{y}_h$ can be written as

$$\begin{aligned}
\tilde{y} - \bar{y}_h = \tilde{y} - y^h + y^h - \bar{y}_h &= \sigma(\hat{y} - y^h) + y^h - \bar{y}_h\\
&= \sigma(\hat{y} - \bar{y}_h) + (1-\sigma)(y^h - \bar{y}_h).
\end{aligned} \tag{3.4}$$

REMARK 3.2. *Please note, that we did not use feasibility of $\bar{y}_h$ for the discrete optimization problem. Hence, Lemma 3.1 is valid without this assumption, which means it is also applicable if the discrete optimization problem is solved for instance by penalty methods.*

Remark 3.3. *If the discrete space V_h is the space of piecewise linear polynomial functions, then we can replace the constraint violation $e_{\rm sc}$ by the interpolation error of the state constraint bounds. Let I_h denote the Lagrange (or nodal) interpolation operator. Then the discrete state constraints* (2.8) *imply $I_h y_a \le \bar y_h \le I_h y_b$ on $\bar\Omega$. And we can estimate in the previous proof*

$$\begin{aligned}(1-\sigma)y^h + \sigma\hat y &\ge (1-\sigma)(\bar y_h - \|\bar y_h - y^h\|_{L^\infty(\Omega)}) + \sigma(y_a+\tau)\\ &\ge y_a + \tau + (1-\sigma)(I_h y_a - \|\bar y_h - y^h\|_{L^\infty(\Omega)} - y_a - \tau),\end{aligned}$$

which shows that we can replace $e_{\rm sc}$ by

$$\tilde e_{\rm sc} := \max(\|(I_h y_a - y_a)^-\|_{L^\infty(\Omega)},\ \|(y_b - I_h y_b)^-\|_{L^\infty(\Omega)}).$$

If the functions y_a, y_b can be exactly represented by piecewise linear functions then $\tilde e_{\rm sc} = 0$. This is in particular the case if both state constraints are constant functions.

3.2. Estimate of error in the control variational inequality. As one ingredient of the final error estimator we will develop an error estimator for the error in the variational inequality (2.4). We will comment on the relation to existing work at the end of Section 4.

At first, let us define the following subsets of Ω

$$\begin{aligned}\Omega_{0,h} := \Big\{x\in\Omega :\ & \bar u_h(x), -\frac{1}{\alpha}\bar p_h(x) \in (u_a,u_b)\\ &\text{or } \bar u_h(x) = u_a, \alpha\bar u_h(x) + \bar p_h(x) < 0\\ &\text{or } \bar u_h(x) = u_b, \alpha\bar u_h(x) + \bar p_h(x) > 0\Big\}\end{aligned} \tag{3.5}$$

and

$$\begin{aligned}\Omega_{a,h} &:= \left\{x : \bar u_h(x)\in(u_a,u_b), -\frac{1}{\alpha}\bar p_h(x) \le u_a\right\}\\ \Omega_{b,h} &:= \left\{x : \bar u_h(x)\in(u_a,u_b), -\frac{1}{\alpha}\bar p_h(x) \ge u_b\right\}.\end{aligned} \tag{3.6}$$

The set $\Omega_{0,h}$ contains the points, where $\bar u_h$ and $-\frac{1}{\alpha}\bar p_h$ are strict between the bounds, and where $\bar u_h$ is at the bound but $\alpha\bar u_h + \bar p_h$ has the wrong sign. The sets $\Omega_{a,h}$ and $\Omega_{b,h}$ contain points, where $\bar u_h$ is strictly between the bounds, but $-\frac{1}{\alpha}\bar p_h$ is not feasible with respect to these bounds. In addition, we have the following properties

$$\begin{aligned}\alpha\bar u_h + \bar p_h &\ge \alpha(\bar u_h - u_a) > 0 \text{ on } \Omega_{a,h},\\ \alpha\bar u_h + \bar p_h &\le \alpha(\bar u_h - u_b) < 0 \text{ on } \Omega_{b,h}.\end{aligned} \tag{3.7}$$

Lemma 3.4. *Let $\bar u_h \in U_{ad}$ and $\bar p_h \in L^2(\Omega)$ be given. Let the sets $\Omega_{0,h}$, $\Omega_{a,h}$, and $\Omega_{b,h}$ be defined according to* (3.5) *and* (3.6). *Then for each $u\in U_{ad}$ it holds*

$$\begin{aligned}(\alpha\bar u_h + \bar p_h, u - \bar u_h) \ge\ & (\chi_{\Omega_{0,h}}(\alpha\bar u_h + \bar p_h), u - \bar u_h)\\ &+ (\chi_{\Omega_{a,h}}(\alpha\bar u_h + \bar p_h), u_a - \bar u_h) + (\chi_{\Omega_{b,h}}(\alpha\bar u_h + \bar p_h), u_b - \bar u_h)\end{aligned}$$

Proof. The proof follows directly from the definition of the sets, and the properties of $\bar{u}_h$ and $\alpha\bar{u}_h + \bar{p}_h$ on these sets, confer also (3.7). ∎

In the derivation of the a-posteriori error estimate in Section 3.4, we will use the following implication of the previous lemma.

LEMMA 3.5. *Let $\bar{u}_h \in U_{ad}$ and $\bar{p}_h \in L^2(\Omega)$ be given. Let the sets $\Omega_{0,h}$, $\Omega_{a,h}$, and $\Omega_{b,h}$ be defined according to* (3.5) *and* (3.6)*. Then for each $u \in U_{ad}$ it holds*

$$(\alpha\bar{u}_h + \bar{p}_h, u - \bar{u}_h) \geq -\frac{\alpha}{4}\|u - \bar{u}_h\|^2_{L^2(\Omega)} - \|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega)},$$

where $\eta_{\mathrm{vi}} = \eta_{\mathrm{vi}}(\bar{u}_h, \bar{p}_h) \in L^2(\Omega)$ is given by

$$\eta^2_{\mathrm{vi}} = \frac{1}{\alpha}\chi_{\Omega_{0,h}}(\alpha\bar{u}_h+\bar{p}_h)^2+\chi_{\Omega_{a,h}}(\alpha\bar{u}_h+\bar{p}_h)(\bar{u}_h-u_a)+\chi_{\Omega_{b,h}}(\alpha\bar{u}_h+\bar{p}_h)(\bar{u}_h-u_b). \quad (3.8)$$

Proof. The claim follows directly from the definition of η_{vi} and the inequality

$$(\chi_{\Omega_{0,h}}(\alpha\bar{u}_h + \bar{p}_h), u - \bar{u}_h) \geq -\frac{\alpha}{4}\|u - \bar{u}_h\|^2_{L^2(\Omega)} - \frac{1}{\alpha}\|\alpha\bar{u}_h + \bar{p}_h\|^2_{L^2(\Omega_{0,h})}.$$

∎

The function η_{vi} will serve as a localizable error estimator for the error in the variational inequality.

3.3. Estimate of cost functional values with respect to the admissible point. Due to the feasibility of $(\tilde{y}, \tilde{u})$ the inequality $0 \leq J(\tilde{y}, \tilde{u}) - J(\bar{y}, \bar{u})$ holds. Now, we will derive an upper bound for $J(\tilde{y}, \tilde{u}) - J(y, u)$ for arbitrary feasible pairs (y, u) in terms of the distance $\|u - \tilde{u}\|_{L^2(\Omega)}$ and of residuals of the optimality system.

LEMMA 3.6. *Let $(\tilde{y}, \tilde{u})$ be given by Lemma 3.1. Then it holds for all (y, u) satisfying* (1.2)–(1.4)

$$J(\tilde{y}, \tilde{u}) - J(y, u) \leq -\frac{\alpha}{4}\|u - \tilde{u}\|^2_{L^2(\Omega)} + \eta_a\|u - \tilde{u}\|_{L^2(\Omega)} + \eta_b,$$

where η_a, η_b are real numbers depending on $(\bar{y}_h, \bar{u}_h, \bar{p}_h, \bar{\mu}_{a,h}, \bar{\mu}_{b,h})$, $(\tilde{y}, \tilde{u})$, and the data of the problem but not on (y, u), see below (3.13).

Proof. Let us introduce the abbreviation $r := \|u - \tilde{u}\|_{L^2(\Omega)}$.

Since $\tilde{y}$ and y are solutions to the elliptic equations for right-hand sides $\tilde{u}$ and u respectively, we have by Lemma 2.1 the regularity $\tilde{y}, y \in H^2(\Omega) \cap H^1_0(\Omega)$. Since we have $\bar{p}_h \in V_h \subset L^2(\Omega)$ the dual product $\langle A\tilde{y} - Ay, \bar{p}_h\rangle$ is well-defined.

Let us write the differences of the cost functional as

$$\begin{aligned} J(\tilde{y}, \tilde{u}) - J(y, u) &= \frac{1}{2}\|\tilde{y} - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|\tilde{u}\|^2_{L^2(\Omega)} - \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} - \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ &\quad + \langle A\tilde{y} - \tilde{u} - Ay + u, \bar{p}_h\rangle \\ &= -\frac{1}{2}\|\tilde{y} - y\|^2_{L^2(\Omega)} - \frac{\alpha}{2}\|\tilde{u} - u\|^2_{L^2(\Omega)} + (\alpha\tilde{u} + \bar{p}_h, \tilde{u} - u) \\ &\quad + \langle -A^*\bar{p}_h + \bar{y}_h - y_d, \tilde{y} - y\rangle + (\tilde{y} - \bar{y}_h, \tilde{y} - y) \\ &\leq -\frac{\alpha}{2}r^2 + \frac{1}{2}\|\tilde{y} - \bar{y}_h\|^2_{L^2(\Omega)} + (\alpha\tilde{u} + \bar{p}_h, \tilde{u} - u) \\ &\quad + \langle -A^*\bar{p}_h + \bar{y}_h - y_d, \tilde{y} - y\rangle. \end{aligned} \quad (3.9)$$

Now we will estimate the third and fourth addend on the right-hand side. By Lemma 3.5 there exists $\eta_{\mathrm{vi}} \in L^2(\Omega)$ such that

$$(\alpha\bar{u}_h + \bar{p}_h,\, u - \bar{u}_h) \geq -\frac{\alpha}{4}\|u - \bar{u}_h\|^2_{L^2(\Omega)} - \|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega)} \quad \forall u \in L^2(\Omega) :\, u_a \leq u \leq u_b.$$

We then obtain

$$\begin{aligned}(\alpha\tilde{u} + \bar{p}_h, \tilde{u} - u) &= \alpha(\tilde{u} - \bar{u}_h, \tilde{u} - u) + (\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) + (\alpha\bar{u}_h + \bar{p}_h, \bar{u}_h - u)\\ &\leq \alpha(\tilde{u} - \bar{u}_h, \tilde{u} - u) + (\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) + \frac{\alpha}{4}\|u - \bar{u}_h\|^2_{L^2(\Omega)} + \|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega)}\\ &\leq \frac{\alpha}{4}r^2 + \frac{3}{2}\alpha\|\tilde{u} - \bar{u}_h\|_{L^2(\Omega)}\, r + \frac{\alpha}{4}\|\tilde{u} - \bar{u}_h\|^2_{L^2(\Omega)}\\ &\qquad + (\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) + \|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega)}.\end{aligned} \tag{3.10}$$

Due to the complementarity condition (2.10) on $\bar{\mu}_{a,h}$ and the feasibility $y_a \leq y$ it holds

$$\begin{aligned}\langle\bar{\mu}_{a,h},\, \tilde{y} - y\rangle &= \langle\bar{\mu}_{a,h},\, \tilde{y} - \bar{y}_h + \bar{y}_h - y_a + y_a - y\rangle\\ &= \langle\bar{\mu}_{a,h},\, \tilde{y} - \bar{y}_h\rangle + \langle\bar{\mu}_{a,h},\, y_a - y\rangle\\ &\leq \langle\bar{\mu}_{a,h},\, \tilde{y} - \bar{y}_h\rangle.\end{aligned}$$

For analogous reasons we find

$$-\langle\bar{\mu}_{b,h},\, \tilde{y} - y\rangle \leq -\langle\bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h\rangle.$$

Thus, the fourth addend in the estimate above can be estimated as

$$\begin{aligned}&\langle -A^*\bar{p}_h + \bar{y}_h - y_d, \tilde{y} - y\rangle\\ &\quad= \langle -A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}, \tilde{y} - y\rangle + \langle\bar{\mu}_{a,h},\, \tilde{y} - y\rangle - \langle\bar{\mu}_{b,h},\, \tilde{y} - y\rangle\\ &\quad\leq \langle -A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}, \tilde{y} - y\rangle + \langle\bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h\rangle\\ &\quad\leq C_0\|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)} \cdot r + \langle\bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h\rangle,\end{aligned} \tag{3.11}$$

where we applied the $H^2(\Omega)$-regularity result Lemma 2.1 in the last step.

Combining (3.9)–(3.11), we can estimate the difference of the values of the cost functional as

$$\begin{aligned}J(\tilde{y}, \tilde{u}) - J(y, u) &\leq -\frac{\alpha}{4}r^2\\ &\quad + \left(\frac{3}{2}\alpha\|\tilde{u} - \bar{u}_h\|_{L^2(\Omega)} + C_0\|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)}\right) r\\ &\quad + \frac{1}{2}\|\tilde{y} - \bar{y}_h\|^2_{L^2(\Omega)} + \frac{\alpha}{4}\|\tilde{u} - \bar{u}_h\|^2_{L^2(\Omega)} + (\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) + \|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega)}\\ &\quad + \langle\bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h\rangle,\end{aligned} \tag{3.12}$$

which yields the claim with

$$\begin{aligned}\eta_a &:= \frac{3}{2}\alpha\|\tilde{u} - \bar{u}_h\|_{L^2(\Omega)} + C_0\|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)}\\ \eta_b &:= \frac{1}{2}\|\tilde{y} - \bar{y}_h\|^2_{L^2(\Omega)} + \frac{\alpha}{4}\|\tilde{u} - \bar{u}_h\|^2_{L^2(\Omega)} + (\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) + \|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega)}\\ &\qquad + \langle\bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h\rangle.\end{aligned} \tag{3.13}$$

□

REMARK 3.7. *If one uses interior point methods to solve the discretized problem, then the discrete complementarity condition* (2.10)*, in particular* $\langle \bar\mu_{i,h},\, \bar y_h - y_i\rangle = 0$*,* $i \in \{a, b\}$*, is not satisfied in general. It turns out that the estimate of the previous lemma holds true if* η_a *is replaced by* $\tilde\eta_a$ *given by*

$$\tilde\eta_a := \eta_a + \langle \bar\mu_{a,h}, \bar y_h - y_a\rangle - \langle \bar\mu_{b,h}, \bar y_h - y_b\rangle,$$

which takes the violation of the discrete complementarity condition into account.

3.4. Upper bound for the error in the control and state. Using the solution of the continuous problem as test functions in Lemma 3.6, we get directly an estimate of the error in the controls.

LEMMA 3.8. *With the notations of the previous Lemma 3.6, it holds*

$$\|\bar u_h - \bar u\|^2_{L^2(\Omega)} \le 2\|\tilde u - \bar u_h\|^2_{L^2(\Omega)} + \frac{8}{\alpha^2}\eta_a^2 + \frac{8}{\alpha}\eta_b.$$

Proof. By Lemma 3.6, optimality of $(\bar y, \bar u)$, and feasibility of $(\tilde y, \tilde u)$, we obtain

$$0 \le J(\tilde y, \tilde u) - J(y, u) \le -\frac{\alpha}{2}\|\bar u - \tilde u\|^2_{L^2(\Omega)} + \eta_a\|\bar u - \tilde u\|_{L^2(\Omega)} + \eta_b,$$

which gives directly

$$\|\bar u - \tilde u\|^2_{L^2(\Omega)} \le \frac{4}{\alpha^2}\eta_a^2 + \frac{4}{\alpha}\eta_b.$$

The claim follows with $\|\bar u_h - \bar u\|^2_{L^2(\Omega)} \le 2(\|\tilde u - \bar u_h\|^2_{L^2(\Omega)} + \|\bar u - \tilde u\|^2_{L^2(\Omega)})$. □

Here, the quantity η_a^2 can be bounded from above, cf. (3.13), by

$$\eta_a^2 \le \frac{9}{2}\alpha^2\|\tilde u - \bar u_h\|^2_{L^2(\Omega)} + 2C_0^2\|-A^*\bar p_h + \bar y_h - y_d - \bar\mu_{a,h} + \bar\mu_{b,h}\|^2_{H^{-2}(\Omega)}. \tag{3.14}$$

Now we have everything at hand to derive the upper bound for the discretization error of control and state. Let us emphasize that all quantities on the right-hand side of the error estimate (3.15) are computable. In particular, the right-hand side *does not* contain any component of the solution $(\bar y, \bar u, \bar p, \bar\mu_a, \bar\mu_b)$ of the continuous optimality system.

THEOREM 3.9. *Let* $(\bar y, \bar u)$ *be the solution of the continuous optimal control problem* (1.1)–(1.4)*. Let* $(\bar y_h, \bar u_h, \bar p_h, \bar\mu_{a,h}, \bar\mu_{b,h})$ *be the solution of the discrete problem satisfying* (2.7)–(2.8), (2.10).

Then there is a constant $c > 0$ *that depends only on* α, Ω, A, τ *such that*

$$\begin{aligned}\|\bar u_h - \bar u\|^2_{L^2(\Omega)} + \|\bar y_h - \bar y\|^2_{L^\infty(\Omega)} \le c\Big(& r^{(1)}_{\mathrm{state},L^\infty} + r^{(2)}_{\mathrm{state},L^\infty} + r^2_{\mathrm{state},L^2} \\ & + r^2_{\mathrm{adjoint},L^2} + r^2_{\mathrm{control},L^2} + e_{\mathrm{state}}\Big)\end{aligned} \tag{3.15}$$

where $r^{(1)}_{\mathrm{state},L^\infty}$ *is the scaled* L^∞*-error of the states given by*

$$\begin{aligned} r^{(1)}_{\mathrm{state},L^\infty} = \Big(&\max\big((\alpha\bar u_h + \bar p_h, \hat u - \bar u_h) + \langle \bar\mu_{a,h} - \bar\mu_{b,h},\, \hat y - \bar y_h\rangle,\ 0\big) \\ &+ \|\bar\mu_{a,h} - \bar\mu_{b,h}\|_{M(\Omega)}\Big)\|y^h - \bar y_h\|_{L^\infty(\Omega)},\end{aligned}$$

$r^{(2)}_{\text{state},L^\infty}$ *is the squared and scaled* L^∞*-error of the states given by*

$$r^{(2)}_{\text{state},L^\infty} = \left(\|\hat{u} - \bar{u}_h\|^2_{L^2(\Omega)} + \|\hat{y} - \bar{y}_h\|^2_{L^2(\Omega)} \right) \|y^h - \bar{y}_h\|^2_{L^\infty(\Omega)},$$

r_{state,L^2} *is the* L^2*-error of the states given by*

$$r_{\text{state},L^2} = \|y^h - \bar{y}_h\|_{L^2(\Omega)},$$

r_{adjoint,L^2} *is the* $H^{-2}(\Omega)$*-residual in the adjoint equation*

$$r_{\text{adjoint},L^2} = \|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)},$$

r_{control,L^2} *is the* $L^2(\Omega)$*-residual in the variational inequality defined by*

$$r_{\text{control},L^2} = \|\eta_{\text{vi}}\|_{L^2(\Omega)}, \text{ where } \eta_{\text{vi}} = \eta_{\text{vi}}(\bar{u}_h, \bar{p}_h) \text{ is given by Lemma 3.5,}$$

and e_{state} *is the weighted violation of the state constraints given by*

$$\begin{aligned} e_{\text{state}} = \max\big((\alpha\bar{u}_h + \bar{p}_h, \hat{u} - \bar{u}_h) + \langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \hat{y} - \bar{y}_h \rangle,\ 0 \big) e_{\text{sc}} \\ + \left(\|\hat{u} - \bar{u}_h\|^2_{L^2(\Omega)} + \|\hat{y} - \bar{y}_h\|^2_{L^2(\Omega)} \right) e_{\text{sc}}^2, \end{aligned}$$

where e_{sc} *is defined in Lemma 3.1, eq.* (3.2).

Proof. Let us first combine the results of Lemma 3.6 with the estimate of η_a^2 in (3.14) and the definition of η_b in Lemma 3.8, eq. (3.13), to obtain

$$\begin{aligned} \|\bar{u}_h - \bar{u}\|^2_{L^2(\Omega)} \le 40 \|\tilde{u} - \bar{u}_h\|^2_{L^2(\Omega)} \\ + \frac{16C_0^2}{\alpha^2} \|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|^2_{H^{-2}(\Omega)} \\ + \frac{8}{\alpha}\Big(\frac{1}{2}\|\tilde{y} - \bar{y}_h\|^2_{L^2(\Omega)} + (\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) + \|\eta_{\text{vi}}\|^2_{L^2(\Omega)} \\ + \langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h \rangle \Big). \end{aligned}$$

At first, we have by Lemma 3.1, cf. (3.3),

$$(\alpha\bar{u}_h + \bar{p}_h, \tilde{u} - \bar{u}_h) = (\alpha\bar{u}_h + \bar{p}_h, \hat{u} - \bar{u}_h)\sigma$$

and

$$\|\tilde{u} - \bar{u}_h\|^2_{L^2(\Omega)} = \sigma^2 \|\hat{u} - \bar{u}_h\|^2_{L^2(\Omega)}.$$

With the help of (3.4) we get

$$\begin{aligned} \langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \tilde{y} - \bar{y}_h \rangle &= \sigma \langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \hat{y} - \bar{y}_h \rangle + (1-\sigma)\langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, y^h - \bar{y}_h \rangle \\ &\le \sigma \langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \hat{y} - \bar{y}_h \rangle + \|\bar{\mu}_{a,h} - \bar{\mu}_{b,h}\|_{M(\Omega)} \|y^h - \bar{y}_h\|_{L^\infty(\Omega)}. \end{aligned}$$

Similarly, we can estimate using $\sigma \ge 0$

$$\begin{aligned} \|\tilde{y} - \bar{y}_h\|^2_{L^2(\Omega)} &= \|\sigma(\hat{y} - \bar{y}_h) + (1-\sigma)(y^h - \bar{y}_h)\|^2_{L^2(\Omega)} \\ &\le 2\left(\sigma^2 \|\hat{y} - \bar{y}_h\|^2_{L^2(\Omega)} + \|y^h - \bar{y}_h\|^2_{L^2(\Omega)} \right). \end{aligned}$$

Hence there is a constant $c > 0$ depending only on α, Ω, A such that

$$
\begin{aligned}
\|\bar{u}_h - \bar{u}\|^2_{L^2(\Omega)} \le c\Big\{ & \|\hat{u} - \bar{u}_h\|^2_{L^2(\Omega)}\sigma^2 \\
& + \|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|^2_{H^{-2}(\Omega)} \\
& + \|\hat{y} - \bar{y}_h\|^2_{L^2(\Omega)}\sigma^2 + (\alpha\bar{u}_h + \bar{p}_h, \hat{u} - \bar{u}_h)\sigma + \|\eta\|^2_{L^2(\Omega)} \\
& + \langle \bar{\mu}_{a,h} - \bar{\mu}_{b,h},\, \hat{y} - \bar{y}_h\rangle\sigma\Big).
\end{aligned}
\tag{3.16}
$$

The value of σ can be bounded according to Lemma 3.1 by

$$
\sigma \le \tau^{-1}(\|y^h - \bar{y}_h\|_{L^\infty(\Omega)} + e_{\mathrm{sc}}),
$$

where e_{sc} is the state constraint violation defined in (3.2). Using this estimate in (3.16), the claim follows. ☐

In Section 4, we will prove convergence to zero of the upper bound (3.15) if the discrete quantities converge in some sense to solutions of the optimality system in Theorem 2.3.

3.5. Localized a-posteriori error estimates. In the previous sections, we developed error bounds for the discretization error. These bounds contain terms that are still not fully accessible. In particular, it needs to be specified, how the L^2- and L^∞-discretization errors of the states as well as the residual of the adjoint equation can be calculated.

For the L^2-error of the states, we have the following result, which is a standard estimate, see e.g. [4]. Recall that $\bar{y}_h$ is the solution of the discretized equation (2.6) with right-hand side $\bar{u}_h$, while y^h is the solution of the elliptic equation (3.1) with the same right-hand side $\bar{u}_h$.

Lemma 3.10. *There is a constant $c > 0$ depending on Ω, the polynomial degree l, and the shape regularity of the triangulation such that*

$$
\|y^h - \bar{y}_h\|_{L^2(\Omega)} \le c\,\eta^2_{state,L^2}
$$

with $\eta^2_{state,L^2} = \sum_{T\in\mathcal{T}} \eta^2_{T,state,L^2}$ and

$$
\eta^2_{T,state,L^2} = \left(h_T^4 \|\Delta\bar{y}_h + \bar{u}_h\|^2_{L^2(T)} + h_T^3 \left\| \left[\frac{\partial \bar{y}_h}{\partial n}\right] \right\|^2_{L^2(\partial T\setminus\Gamma)} \right).
$$

Here, $\left[\frac{\partial \bar{y}_h}{\partial n}\right]$ denotes the jump of the normal derivative across interior edges.

To estimate the L^∞-error we use the following reliable and efficient error estimator from [19].

Lemma 3.11. *There is a constant $c > 0$ depending on Ω, the polynomial degree l, and the shape regularity of the triangulation such that*

$$
\|y^h - \bar{y}_h\|_{L^\infty(\Omega)} \le c\,\eta_{state,L^\infty}
$$

with $\eta_{state,L^\infty} = \max_{T\in\mathcal{T}} \eta_{T,state,L^\infty}$ and

$$
\eta_{T,state,L^\infty} = |\log h_{min}|^2 \left(h_T^2 \|\Delta\bar{y}_h + \bar{u}_h\|_{L^\infty(T)} + h_T \left\| \left[\frac{\partial \bar{y}_h}{\partial n}\right] \right\|_{L^\infty(\partial T\setminus\Gamma)} \right).
$$

It remains to describe the estimation of the H^{-2}-residual of the adjoint equation. For a test function $\phi \in V = H_0^1(\Omega) \cap H^2(\Omega)$ we will employ the Lagrange interpolation $I_h\phi$, which has the property $\phi(x_i) = (I_h\phi)(x_i)$ for all nodes x_i. Due to the assumptions on V_h we get $I_h\phi \in V_h$. Let $\bar{p}_h$ solve the discrete adjoint equation (2.9). We have

$$\begin{aligned}
&\|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)} \\
&\qquad = \sup_{\phi \in V,\, \|\phi\|_{H^2(\Omega)}=1} \langle -A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}, \phi \rangle \\
&\qquad = \sup_{\phi \in V,\, \|\phi\|_{H^2(\Omega)}=1} \langle -A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}, \phi - I_h\phi \rangle.
\end{aligned}$$

Since $\bar{\mu}_{a,h}$ and $\bar{\mu}_{b,h}$ are a linear combination of Dirac measures concentrated in the mesh nodes, it holds

$$\langle -\bar{\mu}_{a,h} + \bar{\mu}_{b,h}, \phi - I_h\phi \rangle = 0,$$

which implies

$$\|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)} = \sup_{\phi \in V,\, \|\phi\|_{H^2(\Omega)}=1} \langle -A^*\bar{p}_h + \bar{y}_h - y_d, \phi - I_h\phi \rangle.$$

Following [1] we obtain

Lemma 3.12. *There is a constant $c > 0$ depending on Ω, the polynomial degree l, and the shape regularity of the triangulation such that*

$$\|-A^*\bar{p}_h + \bar{y}_h - y_d - \bar{\mu}_{a,h} + \bar{\mu}_{b,h}\|_{H^{-2}(\Omega)} \le c\,\eta^2_{adjoint,L^2}$$

with $\eta^2_{adjoint,L^2} = \sum_{T \in \mathcal{T}} \eta^2_{T,adjoint,L^2}$ and

$$\eta^2_{T,adjoint,L^2} = \left(h_T^4 \|\Delta\bar{p}_h + \bar{y}_h - y_d\|^2_{L^2(T)} + h_T^3 \left\| \left[\frac{\partial \bar{p}_h}{\partial n} \right] \right\|^2_{L^2(\partial T \setminus \Gamma)} \right).$$

Although this result of [1] is formulated only for $n = 2$, $l = 1$, and a single Dirac measure, the proofs carry over one-to-one to the case considered here: $n \in \{2, 3\}$, general FE-space V_h with $l \ge 1$, right-hand side consists of linear combination of Dirac measures concentrated in the nodes.

In these Lemmata, we cited only the reliability estimates (i.e. upper error bounds). For all three estimators also lower error bounds (efficiency estimates) are available.

The localization of the estimator of the error in the variational inequality is an obvious choice. Let us define

$$\eta^2_{T,control,L^2} := \eta_{\mathrm{vi}}|_T,$$

where $\eta_{\mathrm{vi}} = \eta(\bar{u}_h, \bar{p}_h)$ is the function constructed in Lemma 3.5.

Combining the estimates of this section with the result of Theorem 3.9, we get our main result, which is the localized a-posteriori error estimate.

Theorem 3.13. *Let $(\bar{y}, \bar{u})$ be the solution of the continuous optimal control problem* (1.1)–(1.4). *Let $(\bar{y}_h, \bar{u}_h, \bar{p}_h, \bar{\mu}_{a,h}, \bar{\mu}_{b,h})$ satisfy* (2.7)–(2.8), (2.10).

Then there is a constant $c > 0$ depending on α, Ω, τ, and the shape regularity of the triangulation, and a weight $\omega_{\infty,h} > 0$ such that

$$\|\bar u_h - \bar u\|^2_{L^2(\Omega)} + \|\bar y_h - \bar y\|^2_{L^\infty(\Omega)} \le c\left(\eta^2_{L^2} + \omega_{\infty,h}\,\eta_{state,L^\infty} + e_{\mathrm{state}}\right)$$

with $\eta^2_{L^2} = \sum_{T\in\mathcal{T}} \eta^2_{T,L^2}$ and

$$\eta^2_{T,L^2} = \eta^2_{T,state,L^2} + \eta^2_{T,control,L^2} + \eta^2_{T,adjoint,L^2}.$$

The weight $\omega_{\infty,h}$ depends on the discrete quantities $(\bar y_h, \bar u_h, \bar p_h, \bar\mu_{a,h}, \bar\mu_{b,h})$. The mapping from $(\bar y_h, \bar u_h, \bar p_h, \bar\mu_{a,h}, \bar\mu_{b,h})$ to $\omega_{\infty,h}$ is bounded from $C(\bar\Omega) \times L^2(\Omega) \times L^2(\Omega) \times M(\Omega) \times M(\Omega)$ to $\mathbb{R}$.

Proof. The result follows directly from Theorem 3.9 and Lemmata 3.10, 3.11, and 3.12. The quantity $\omega_{\infty,h}$ is given by

$$\begin{aligned}\omega_{\infty,h} = \max\big((\alpha\bar u_h + \bar p_h, \hat u - \bar u_h) + \langle \bar\mu_{a,h} - \bar\mu_{b,h},\, \hat y - \bar y_h\rangle,\ 0\big) + \|\bar\mu_{a,h} - \bar\mu_{b,h}\|_{M(\Omega)}\\ + \Big(\|\hat u - \bar u_h\|^2_{L^2(\Omega)} + \|\hat y - \bar y_h\|^2_{L^2(\Omega)}\Big)\|y^h - \bar y_h\|_{L^\infty(\Omega)},\end{aligned}$$

cf. the definition of $r^{(1)}_{\mathrm{state},L^\infty}$ and $r^{(2)}_{\mathrm{state},L^\infty}$ in Theorem 3.9. The mapping $\bar u_h \mapsto y^h$ is bounded from $L^2(\Omega)$ to $L^\infty(\Omega)$ by Lemma 2.1, which proves the claimed boundedness of $\omega_{\infty,h}$. ☐

3.6. Marking strategy. It remains to describe, how to mark elements for refinement. Here, we follow the common strategy to mark elements that have relatively large local error indicators. In our case, see Theorem 3.13, the error indicator contains two terms with different accumulation properties: $\eta^2_{L^2} = \sum_{T\in\mathcal{T}} \eta^2_{T,L^2}$ and $\eta_{state,L^\infty} = \max_{T\in\mathcal{T}}(\eta_{T,state,L^\infty} + e_{T,\mathrm{state}})$.

As marking strategy we employ the one used in [19]. Let us define

$$\begin{aligned}\eta_2^2 &:= \eta^2_{L^2}\\ \eta_\infty &:= \omega_{\infty,h}\,\eta_{state,L^\infty} + e_{\mathrm{state}}\\ \eta &:= \max(\eta_2,\ \eta_\infty).\end{aligned}$$

We choose an error indicator η_i if it is relatively large compared to the total error, that is if $\eta_i \ge \theta_1\eta$, $i \in \{2, \infty\}$. For a chosen error indicator η_i, we mark elements by the maximum strategy, that is, elements $\hat T$ with $\eta_{\hat T,i} \ge \theta_2 \max_{T\in\mathcal{T}_h} \eta_{T,i}$ are marked for refinement. Here, the parameters θ_1, θ_2 are taken from $(0, 1)$. In our computations we used $\theta_1 = 0.2$, $\theta_2 = 0.8$.

4. Convergence of error bound. In this section we will prove the convergence to zero of the error bound of Theorem 3.9 if the solution of the discrete system converges in the following sense.

ASSUMPTION 4. *Let a sequence of meshes $\mathcal{T}_k$ with associated solutions of the discrete problem $\{(\bar y_{h_k}, \bar u_{h_k}, \bar p_{h_k}, \bar\mu_{a,h_k}, \bar\mu_{b,h_k})\}$ be given. Let us assume that we have the following properties of this sequence:*

(i) The sequence $(\bar y_{h_k}, \bar u_{h_k})$ converges strongly to $(\bar y, \bar u)$ in $C(\bar\Omega) \times L^2(\Omega)$.
(ii) The sequence $(\bar p_{h_k}, \bar\mu_{a,h_k}, \bar\mu_{b,h_k})$ is bounded in $W^{1,q}(\Omega) \times M(\Omega) \times M(\Omega)$ with $\frac{2n}{2+n} < q < \frac{n}{n-1}$.

(iii) For each subsequence $(\bar p_{h_{k'}}, \bar\mu_{a,h_{k'}}, \bar\mu_{b,h_{k'}})$ with $\bar p_{h_{k'}} \rightharpoonup \bar p$ in $W^{1,q}(\Omega)$, $q > \frac{2n}{2+n}$, and $(\bar\mu_{a,h_{k'}}, \bar\mu_{b,h_{k'}}) \rightharpoonup^* (\bar\mu_a, \bar\mu_b)$ in $M(\Omega)\times M(\Omega)$ the limit element $(\bar p, \bar\mu_a, \bar\mu_b)$ is a Lagrange multiplier to $(\bar y, \bar u)$, i.e. the system (2.2)–(2.4) is satisfied.

Please note, that the requirements of this assumption are satisfied for a uniform refinement of the mesh, see e.g. [13, Chapter 3]. However, convergence of the mesh size $h \to 0$ is not explicitly required.

Let us recall the error representation of Theorem 3.9, which reads

$$\begin{aligned}\|\bar u_h - \bar u\|^2_{L^2(\Omega)} + \|\bar y_h - \bar y\|^2_{L^\infty(\Omega)} \le c\Big(& r^{(1)}_{\text{state},L^\infty} + r^{(2)}_{\text{state},L^\infty} + r^2_{\text{state},L^2} \\ &+ r^2_{\text{adjoint},L^2} + r^2_{\text{control},L^2} + e_{\text{state}}\Big).\end{aligned}$$

Lemma 4.1. *Let Assumption 4 be satisfied. Then it holds*

$$r^{(1)}_{\text{state},L^\infty} + r^{(2)}_{\text{state},L^\infty} + r^2_{\text{state},L^2} + r^2_{\text{adjoint},L^2} + e_{\text{state}} \to 0$$

for $k \to \infty$.

Proof. Let us define y^{h_k} to be the solution of (3.1) to the control $\bar u_{h_k}$. Then we have

$$\begin{aligned}\|y^{h_k} - \bar y_{h_k}\|_{L^\infty(\Omega)} &\le \|y^{h_k} - \bar y\|_{L^\infty(\Omega)} + \|\bar y - \bar y_{h_k}\|_{L^\infty(\Omega)} \\ &\le C\|\bar u_{h_k} - \bar u\|_{L^2(\Omega)} + \|\bar y - \bar y_{h_k}\|_{L^\infty(\Omega)}.\end{aligned}$$

which proves that $r^{(1)}_{\text{state},L^\infty}$, $r^{(2)}_{\text{state},L^\infty}$, and similarly r_{state,L^2} converge to zero under the Assumption 4.

Because of $y_{h_k} \to \bar y$ in $C(\bar\Omega)$ and the feasibility of $\bar y$ with respect to the control constraints, the state constraint violation e_{sc} and consequently e_{state} tend to zero.

Let weak (weak*) converging subsequences $(\bar p_{h_{k'}}, \bar\mu_{a,h_{k'}}, \bar\mu_{b,h_{k'}})$ be given. Then by compact embeddings and after extracting another subsequence $\bar p_{h_{k''}} \to \bar p$ in $L^2(\Omega)$ and $(\bar\mu_{a,h_{k'}}, \bar\mu_{b,h_{k'}}) \to (\bar\mu_a, \bar\mu_b)$ in $H^{-2}(\Omega)$. Hence $r_{\text{adjoint},L^2} \to 0$ for this subsequence. Since the subsequence (k') was chosen arbitrary, it follows $r_{\text{adjoint},L^2} \to 0$ for $k \to \infty$. ∎

The discussion of the estimator r_{control,L^2} is more involved, and thus we state and prove it separately. For convenience, let us recall its definition. In Theorem 3.9 we set $r_{\text{control},L^2} := \|\eta_{\text{vi}}\|_{L^2(\Omega)}$, where η_{vi} was given by Lemma 3.5 and defined by

$$\eta^2_{\text{vi}} = \frac{1}{\alpha}\chi_{\Omega_{0,h}}(\alpha\bar u_h + \bar p_h)^2 + \chi_{\Omega_{a,h}}(\alpha\bar u_h + \bar p_h)(\bar u_h - u_a) + \chi_{\Omega_{b,h}}(\alpha\bar u_h + \bar p_h)(\bar u_h - u_b),$$

where the sets $\Omega_{0,h}, \Omega_{a,h}, \Omega_{b,h}$ were defined as

$$\begin{aligned}\Omega_{0,h} = \Big\{x \in \Omega :\ & \bar u_h(x), -\frac{1}{\alpha}\bar p_h(x) \in (u_a, u_b) \\ & \text{or } \bar u_h(x) = u_a, \alpha\bar u_h(x) + \bar p_h(x) < 0 \\ & \text{or } \bar u_h(x) = u_b, \alpha\bar u_h(x) + \bar p_h(x) > 0\Big\}\end{aligned}$$

and

$$\begin{aligned}\Omega_{a,h} &= \left\{x : \bar u_h(x) \in (u_a, u_b), -\frac{1}{\alpha}\bar p_h(x) \le u_a\right\} \\ \Omega_{b,h} &= \left\{x : \bar u_h(x) \in (u_a, u_b), -\frac{1}{\alpha}\bar p_h(x) \ge u_b\right\}.\end{aligned}$$

Lemma 4.2. *Let Assumption 4 be satisfied. Then it holds*

$$r_{\mathrm{control},L^2} \to 0.$$

Proof. Since the controls u_{h_k} are feasible for the discrete problem, they are uniformly bounded in $L^\infty(\Omega)$. Hence, the sequence u_{h_k} converges to $\bar u$ in $L^s(\Omega)$ for all $2 \le s < \infty$. Moreover, as argued above, each weakly converging subsequence $\bar p_{h_{k'}}$ has a strongly converging subsequence $\bar p_{h_{k''}} \to \bar p$ in $L^2(\Omega)$. By Assumption 4, we have that the variational inequality (2.4) is satisfied.

In the course of the proof, we will bound $\|\eta_{\mathrm{vi}}\|_{L^2(\Omega)}$ in terms of $\|\bar p_{h_{k''}} - \bar p\|_{L^2(\Omega)}$ and $\|\bar u_{h_{k''}} - \bar u\|_{L^s(\Omega)}$, $s > 2$. To simplify the notation, we will drop the index k''.

At first, let us note that

$$\eta_{\mathrm{vi}}(x)^2 \le \frac{1}{\alpha}|\alpha\bar u_h(x) + \bar p_h(x)|^2 \tag{4.1}$$

follows directly from the definition of η_{vi}, $\Omega_{a,h}$, and $\Omega_{b,h}$, confer also (3.7).

Now, we will discuss upper bounds of η_{vi} on different subsets of Ω.

Case (1): $\Omega_1 = \{x \in \Omega :\ \alpha\bar u(x) + \bar p(x) = 0\}$.

Using (4.1) it follows directly

$$\begin{aligned}\|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega_1)} &\le \alpha^{-1}\|\alpha\bar u_h + \bar p_h - (\alpha\bar u + \bar p)\|^2_{L^2(\Omega_1)}\\ &\le \alpha^{-1}(\alpha\|\bar u_h - \bar u\|_{L^2(\Omega)} + \|\bar p_h - \bar p\|_{L^2(\Omega)})^2.\end{aligned} \tag{4.2}$$

Case (2$_a$): $\Omega_{2,a} = \{x \in \Omega :\ \bar u_h(x) = \bar u(x) = u_a\}$. Here, we have

$$\begin{aligned}\|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega_{2,a})} &= \alpha^{-1}\|(\alpha\bar u_h + \bar p_h)^-\|^2_{L^2(\Omega_{2,a})}\\ &= \alpha^{-1}\left\|(\alpha\bar u_h + \bar p_h)^- - (\alpha\bar u + \bar p)^-\right\|^2_{L^2(\Omega_{2,a})}\\ &\le \alpha^{-1}(\alpha\|\bar u_h - \bar u\|_{L^2(\Omega)} + \|\bar p_h - \bar p\|_{L^2(\Omega)})^2.\end{aligned}$$

Case (2$_b$): $\Omega_{2,b} = \{x \in \Omega :\ \bar u_h(x) = \bar u(x) = u_b\}$. Analogous to *Case (2$_a$)*.

Case (3$_a$): $\Omega_{3,a} = \{x \in \Omega :\ \bar u(x) = u_a\} \cap \Omega_{a,h}$. Here, we have by definition of η_{vi}

$$\eta_{\mathrm{vi}}(x)^2 = (\alpha\bar u_h + \bar p_h)(\bar u_h - u_a) = (\alpha\bar u_h + \bar p_h)(\bar u_h - \bar u).$$

Hence, it holds $\|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega_{3,a})} \le \|\alpha\bar u_h + \bar p_h\|_{L^2(\Omega)}\|\bar u_h - \bar u\|_{L^2(\Omega)}$.

Case (3$_b$): $\Omega_{3,b} = \{x \in \Omega :\ \bar u(x) = u_b\} \cap \Omega_{b,h}$. Analogous to *Case (3$_a$)*.

Case (4$_a$): $\Omega_{4,a} = \{x \in \Omega :\ \bar u_h(x), -\frac{1}{\alpha}\bar p_h(x) \in (u_a, u_b),\ \bar u(x) = u_a,\ -\frac{1}{\alpha}\bar p(x) < u_a\}$. The inequality $-\frac{1}{\alpha}\bar p(x) < u_a = \bar u(x) < -\frac{1}{\alpha}\bar p_h(x)$ implies

$$|\alpha\bar u_h + \bar p_h| \le \alpha|\bar u_h - \bar u| + \alpha|\bar u - (-\alpha^{-1}\bar p_h)| \le \alpha|\bar u_h - \bar u| + |-\bar p - (-\bar p_h)|$$

which proves with (4.1)

$$\|\eta_{\mathrm{vi}}\|^2_{L^2(\Omega_{4,a})} \le \frac{1}{\alpha}\|\alpha\bar u(x) + \bar p(x)\|^2_{L^2(\Omega_{4,a})} \le \|\bar p_h - \bar p\|^2_{L^2(\Omega_{4,a})}.$$

Case (4$_b$): $\Omega_{4,b} = \{x \in \Omega:\ \bar u_h(x), -\frac{1}{\alpha}\bar p_h(x) \in (u_a, u_b),\ \bar u(x) = u_b,\ -\frac{1}{\alpha}\bar p(x) < u_b\}$. Analogous to *Case (4$_a$)*.

Case (5$_a$): $\Omega_{5,a} = \{x \in \Omega:\ \bar u(x) = u_a\} \cap \Omega_{b,h}$. Here, we have $-\frac{1}{\alpha}\bar p(x) \le u_a < u_b < -\frac{1}{\alpha}\bar p_h(x)$. By Chebyshev's inequality, we obtain

$$\begin{aligned}|\Omega_{5,a}| &\le \left|\left\{x:\ \alpha^{-1}|\bar p_h(x) - \bar p(x)| \ge |u_b - u_a|\right\}\right|\\ &\le \int_{\Omega\setminus\Omega_2} \frac{|\bar p_h(x) - \bar p(x)|^2}{\alpha^2|u_b - u_a|^2}\\ &\le \frac{1}{(\alpha|u_b - u_a|)^2}\|\bar p_h - \bar p\|^2_{L^2(\Omega)}.\end{aligned}$$

Together with the definition η_{vi} on $\Omega_{b,h}$, we find

$$\|\eta_{\text{vi}}\|^2_{L^2(\Omega_{5,a})} \le |\Omega_{5,a}|^{1/2}\|\alpha\bar u_h + \bar p_h\|_{L^2(\Omega)}\|\bar u_h - u_b\|_{L^\infty(\Omega)} \le c\|\bar p_h - \bar p\|_{L^2(\Omega)}.$$

Case (5$_b$): $\Omega_{5,b} = \{x \in \Omega:\ \bar u(x) = u_b\} \cap \Omega_{a,h}$. Analogous to *Case (5$_a$)*.

Case (6): $\Omega_6 = \{x \in \Omega:\bar u(x), \bar u_h(x) \in \{u_a, u_b\},\ \bar u(x) \ne \bar u_h(x)\}$. That is, here the control constraints are active at $\bar u$ and $\bar u_h$ but both are not equal. Similarly as in *Case (5$_a$)* we estimate

$$\begin{aligned}|\Omega_6| &\le \left|\left\{x:\ |\bar u(x) - \bar u_h(x)| \ge |u_b - u_a|\right\}\right|\\ &\le \int_{\Omega\setminus\Omega_2} \frac{|\bar u_h(x) - \bar u(x)|^s}{|u_b - u_a|^s}\\ &\le \frac{1}{|u_b - u_a|^s}\|\bar u_h - \bar u\|^s_{L^s(\Omega)},\end{aligned}$$

which yields due to $\Omega_6 \subset \Omega_{0,h}$

$$\|\eta_{\text{vi}}\|^2_{L^2(\Omega_6)} \le \frac{1}{\alpha}|\Omega_6|^{1-2/s}\|\alpha\bar u_h + \bar p_h\|^2_{L^s(\Omega)} \le C\|\bar u_h - \bar u\|^{s-2}_{L^s(\Omega)}.$$

Let us argue that the splitting introduce by the cases above covers Ω. Due to first order optimality conditions, Ω can be divided in sets, where $\alpha\bar u + \bar p = 0$ and $\bar u \in \{u_a, u_b\}$. The first possibility is covered by *Case (1)*. The case that both $\bar u$ and $\bar u_h$ are at the bounds is contained in *Case (2)* and *Case (6)*. Now it remains to cover the set, where $\bar u \in \{u_a, u_b\}$ and $\bar u_h \in (u_a, u_b)$. The subset, where $-\frac{1}{\alpha}\bar p_h$ is not in (u_a, u_b), is discussed in *Case (3)* and *Case (5)*. And *Case (4)* covers the subset, where $-\frac{1}{\alpha}\bar p_h$ is in (u_a, u_b).

Summing up all the estimates, we find the convergence $\eta_{\text{vi}} \to 0$ in $L^2(\Omega)$ for the subsequence (k'') chosen above. This implies that for every subsequence of $(\bar u_{h_k}, \bar p_{h_k})$ we can choose a subsequence such that the corresponding quantity η_{vi} converges to zero, which finishes the proof. ☐

As a consequence of these results we obtain the main result of this section.

THEOREM 4.3. *Let Assumption 4 be satisfied. Then the error bound given by Theorem 3.9 converges to zero for $k \to \infty$, e.g.*

$$\left(r^{(1)}_{\text{state},L^\infty} + r^{(2)}_{\text{state},L^\infty} + r^2_{\text{state},L^2} + r^2_{\text{adjoint},L^2} + r^2_{\text{control},L^2} + e_{\text{state}}\right) \to 0.$$

Let us now comment on different available error estimators for the error in the variational inequality connected to the control constraints. To the best of our knowledge, all estimators in the literature contain in their upper bounds quantities that do not tend to zero as $(\bar{u}_h, \bar{p}_h)$ converges, which makes it difficult to prove a convergence result similar to Lemma 4.2.

Let us first compare our findings to the error estimator as considered by Krumbiegel and Rösch [15]. Their error estimator η_{KR} coincides with η_{vi} as developed here except for the definition on the sets $\Omega_{a,h}, \Omega_{b,h}$, i.e.

$$\eta_{\text{KR}} = \chi_{\Omega_{0,h}} \eta_{\text{vi}} + \frac{1}{\alpha} \chi_{\Omega_{a,h} \cup \Omega_{b,h}} (\alpha u_h + p_h)^2.$$

which implies $\|\eta_{\text{vi}}\|^2_{L^2(\Omega)} \leq \|\eta_{\text{KR}}\|^2_{L^2(\Omega)}$ by (3.7). The error estimator η_{KR} does not allow a convergence proof comparable to Lemma 4.2. To this end consider the following example: let the constant functions $u_a = 0$, $u_b = 1$, $\alpha = 1$, $\bar{u} = 0$, $\bar{p} = \bar{p}_h = 1$, and $\bar{u}_h = h$ be given. That means, the values are chosen in such a way that $\Omega_{a,h} = \Omega$ and correspond to the *Case (3$_a$)* in the proof of Lemma 4.2. Our error estimator gives $\|\eta_{\text{vi}}\|^2_{L^2(\Omega)} = (\alpha \bar{u}_h + \bar{p}_h, \bar{u}_h - u_a) = (\alpha h + 1) \cdot h \cdot |\Omega| \to 0$ as $h \to 0$. The estimator of [15] yields $\|\eta_{\text{KR}}\|^2_{L^2(\Omega)} = \frac{1}{\alpha} \|\alpha h + 1\|^2_{L^2(\Omega)}$, which does not converge to zero as $h \to 0$. Similar situations as in the example will occur if one uses interior point methods to get rid of the control constraints in the discrete optimization problem.

Second, let us comment on the error estimator analyzed by Hintermüller, Hoppe, Iliash, and Kieweg in [11]. There an efficient and reliable error estimator for optimal control problems with a lower control bound is developed. They prove that the estimator is equivalent (up to higher order terms) to the error

$$\|\bar{y} - \bar{y}_h\|_{H^1(\Omega)} + \|\bar{p} - \bar{p}_h\|_{H^1(\Omega)} + \|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} + \|\sigma - \sigma_h\|_{L^2(\Omega)},$$

with $\sigma := \alpha \bar{u} + \bar{p}$ and $\sigma_h := \alpha \bar{u}_h + \Pi_h \bar{p}_h$. However, it is not clear, whether it holds $\|\sigma - \sigma_h\|_{L^2(\Omega)} \to 0$ if (u_h, p_h) converges to $(\bar{u}, \bar{p})$. Indeed, for *Case (3$_a$)* in the proof of Lemma 4.2 one has the following. If (u_h, p_h) are solutions of the discrete optimal control problem it holds $\alpha u_h + \Pi_h p_h = 0$ on this set, since u_h is not at the control bounds. This implies that

$$\|\sigma - \sigma_h\|_{L^2(\Omega)} = \|(\alpha \bar{u} + \bar{p}) - (\alpha u_h + \Pi_h p_h)\|_{L^2(\Omega_{3,a})} = \|\alpha \bar{u} + \bar{p}\|_{L^2(\Omega_{3,a})}.$$

However, since $\bar{u} = u_a$ holds on $\Omega_{3,a}$ by definition, this quantity is in general non-zero. In order to prove $\|\sigma - \sigma_h\|_{L^2(\Omega)} \to 0$, one has to prove in addition that the measure of $\Omega_{3,a}$ tends to zero.

Finally, we comment on the error estimator for control constrained optimal control problems as considered by Li, Liu, and Yan [16]. The error estimator developed there converges to zero under the assumption that $(\bar{y}_h, \bar{u}_h, \bar{p}_h)$ converges to $(\bar{y}, \bar{u}, \bar{p})$, an assumption on the regularity of the active sets, and the assumption $h \to 0$. Clearly our convergence result Lemma 4.2 holds under weaker assumptions on $\bar{u}_h, \bar{p}_h$.

5. Numerical experiment. Let us report on numerical results with adaptive refinement using the error estimator developed in the present article.

The data of the example is taken from [6]. It was originally posed for $\Omega = B_1(0)$. We modified it to work with $\Omega = [-1, 2]^2$.

The data of the problem is given in polar coordinates. For convenience we define $r = \sqrt{x_1^2 + x_2^2}$. Let us set

$$
\begin{aligned}
\bar{u}(r) &= -\frac{1}{2\pi\alpha}\chi_{\{r\leq 1\}}(\log r + r^2 - r^3) \\
\bar{y}(r) &= \frac{1}{2\pi\alpha}\chi_{\{r\leq 1\}}\left(\frac{r^2}{4}(\log r - 2) + \frac{r^3}{4} + \frac{1}{4}\right) \\
\bar{p}(r) &= -\alpha\bar{u}(r) \\
\bar{\mu}_a(r) &= \delta_0(r) \\
f(r) &= \frac{1}{8\pi}\chi_{\{r\leq 1\}}(4 - 9r + 4r^2 - 4r^3) \\
y_d(r) &= \bar{y}(r) + \frac{1}{2\pi}\chi_{\{r\leq 1\}}(4 - 9r) \\
y_a(r) &= \frac{1}{2\pi\alpha}\left(\frac{1}{4} - \frac{r}{2}\right).
\end{aligned}
$$

The problem features one lower state constraint, there are no upper state constraint and no control constraints given.

One can verify that $(\bar{y}, \bar{u}, \bar{p}, \bar{\mu}_a)$ is the solution of the problem: Minimize $J(y, u)$ given by

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

subject to

$$
\begin{aligned}
-\Delta y &= u + f \quad \text{in } \Omega \\
y &= 0 \qquad\quad \text{on } \Gamma,
\end{aligned}
$$

and

$$
y_a \leq y \text{ in } \Omega.
$$

Moreover, one can verify that

$$
\hat{u}(r) = \chi_{\{r\leq 1\}}(-12 + 18r) \qquad \hat{y}(r) = \chi_{\{r\leq 1\}}(1 + 3r^2 - 2r^3)
$$

fulfills Assumption 2 with $\tau = 0.24999$.

Due to the special structure of the problem, the dual quantities are uniquely determined. Please note, that the adjoint state has a pronounced singularity. Since the data of the problem are smooth, there is no chance to perform an a-priori mesh refinement to resolve the singularity.

To avoid superconvergence effects, we ensured that the point $x = (0, 0)$ cannot be a node of the grid for any refined mesh. This is achieved with a coarse grid mesh generated as a uniform triangulation with 8 triangles.

Figure 5.1 shows the adaptively refine mesh after 5 refinement steps. It shows local refinement around the singularity at $(0, 0)$, which appears only in the adjoint equation and thus is not identifiable a-priori. The right-hand plot compares the error $\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}$ for adaptive and uniform mesh refinement. Clearly, the adaptive mesh refinement leads to a better approximation of the solution with respect to the number of unknowns.

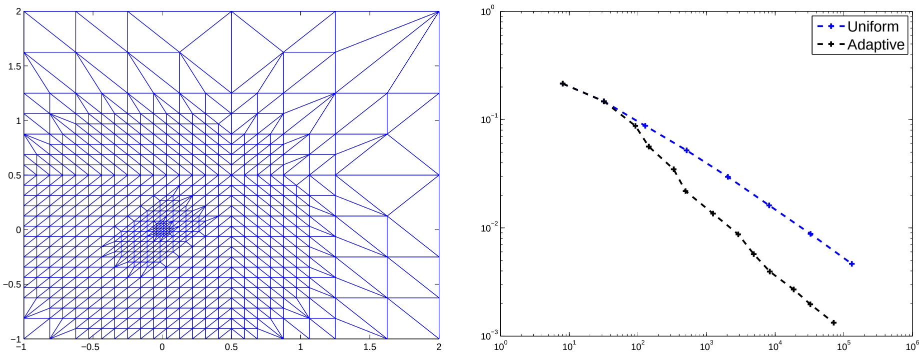

FIG. 5.1. *Adaptively refined mesh; Comparison of* $\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}$ *vs. number of unknowns for uniform vs. adaptive refinement*

6. Conclusion. In this article, we developed a fully computable a-posteriori error estimator for state and control constrained optimal control problems. Moreover, we showed that the estimator tends to zero if the solution of the discretized problems converge to the solution of the undiscretized problem.

REFERENCES

[1] R. Araya, E. Behrens, and R. Rodríguez. A posteriori error estimates for elliptic problems with Dirac delta source terms. *Numer. Math.*, 105(2):193–216, 2006.

[2] O. Benedix and B. Vexler. A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints. *Comput. Optim. Appl.*, 44(1):3–25, 2009.

[3] D. Braess. *Finite Elemente – Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie.* Springer Lehrbuch, Berlin-Heidelberg, 1997.

[4] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer Verlag, New York, 1994.

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

[6] S. Cherednichenko, K. Krumbiegel, and A. Rösch. Error estimates for the Lavrentiev regularization of elliptic optimal control problems. *Inverse Problems*, 24(5):055003, 21, 2008.

[7] P. G. Ciarlet and P.-A. Raviart. Maximum principle and uniform convergence for the finite element method. *Computer Methods in Applied Mechanics and Engeneering*, 2:17–31, 1973.

[8] K. Deckelnick and M. Hinze. Convergence of a finite element approximation to a state-constrained elliptic control problem. *SIAM J. Numer. Anal.*, 45(5):1937–1953, 2007.

[9] P. Grisvard. *Elliptic problems in nonsmooth domains.* Pitman, London, 1985.

[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, R. H. W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM Control Optim. Calc. Var.*, 14(3):540–560, 2008.

[12] M. Hinze. A variational discretization concept in control constrained optimization: the linear-quadratic case. *J. Computational Optimization and Applications*, 30:45–63, 2005.

[13] 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.

[14] R. H. W. Hoppe and M. Kieweg. A posteriori error estimation of finite element approximations of pointwise state constrained distributed control problems. *J. Numer. Math.*, 17(3):219–244, 2009.

[15] K. Krumbiegel and A. Rösch. A new stopping criterion for iterative solvers for control constrained optimal control problems. *Archives of Control Sciences*, 18(1):17–42, 2008.

[16] R. Li, W. Liu, and N. Yan. A posteriori error estimates of recovery type for distributed convex optimal control problems. *J. Sci. Comput.*, 33(2):155–182, 2007.

[17] C. Meyer. Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints. *Control and Cybernetics*, 37(1):51–83, 2008.

[18] C. Meyer, J. C. de los Reyes, and B. Vexler. Finite element error analysis for state-constrained optimal control of the Stokes equations. *Control and Cybernetics*, 37:251–284, 2008.

[19] R. H. Nochetto, A. Schmidt, K. G. Siebert, and A. Veeser. Pointwise a posteriori error estimates for monotone semi-linear equations. *Numer. Math.*, 104(4):515–538, 2006.

[20] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. AMS, Providence, RI, 2010.

[21] 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*, 2008.