

## NEW REGULARITY RESULTS AND FINITE ELEMENT ERROR ESTIMATES FOR A CLASS OF PARABOLIC OPTIMAL CONTROL PROBLEMS WITH POINTWISE STATE CONSTRAINTS\*

CONSTANTIN CHRISTOF\*\* AND BORIS VEXLER

**Abstract.** We study first-order necessary optimality conditions and finite element error estimates for a class of distributed parabolic optimal control problems with pointwise state constraints. It is demonstrated that, if the bound in the state constraint and the differential operator in the governing PDE fulfil a certain compatibility assumption, then locally optimal controls satisfy a stationarity system that allows to significantly improve known regularity results for adjoint states and Lagrange multipliers in the parabolic setting. In contrast to classical approaches to first-order necessary optimality conditions for state-constrained problems, the main arguments of our analysis require neither a Slater point, nor uniform control constraints, nor differentiability of the objective function, nor a restriction of the spatial dimension. As an application of the established improved regularity properties, we derive new finite element error estimates for the dG(0) – cG(1)-discretization of a purely state-constrained linear-quadratic optimal control problem governed by the heat equation. The paper concludes with numerical experiments that confirm our theoretical findings.

**Mathematics Subject Classification.** 35K10, 49K20, 49M05, 65N15, 65N30.

Received January 26, 2020. Accepted September 15, 2020.

### 1. INTRODUCTION

The aim of this paper is to study state-constrained, parabolic, distributed optimal control problems of the form

$$\left. \begin{aligned} \text{Minimize } & J(y, u) := \int_0^T (j(\cdot, y) + g(u)) dt + \phi(y(T)) \\ \text{w.r.t. } & u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\ \text{s.t. } & \partial_t y + Ay = u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ & \text{tr}(y) = 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\ & y(0) = y_0 \quad \text{a.e. in } \Omega, \\ \text{and } & y \geq \psi \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T). \end{aligned} \right\} \quad (\text{P})$$

---

\*This research was conducted within the International Research Training Group IGDK 1754, funded by the German Science Foundation (DFG) and the Austrian Science Fund (FWF) under project number 188264188/GRK1754.

*Keywords and phrases:* Optimal control, parabolic partial differential equation, state constraints, first-order necessary optimality condition, regularity result, finite element method, *a priori* error estimate.

Technische Universität München, Chair of Optimal Control, Center for Mathematical Sciences, M17, Boltzmannstraße 3, 85748 Garching, Germany.

\*\* Corresponding author: [christof@ma.tum.de](mailto:christof@ma.tum.de)

Here,  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , is a bounded domain with a sufficiently regular boundary,  $T > 0$  is a given final time,  $j: [0, T] \times L^2(\Omega) \rightarrow [0, \infty)$ ,  $g: L^2(\Omega) \rightarrow (-\infty, \infty]$  and  $\phi: L^2(\Omega) \rightarrow [0, \infty)$  are (not necessarily differentiable) maps with suitable properties,  $A$  is an elliptic second-order partial differential operator, and  $\psi \in H^2(\Omega)$  and  $y_0 \in H_0^1(\Omega)$  are given functions satisfying  $\text{tr}(\psi) \leq 0$  a.e. on  $\partial\Omega$  and  $y_0 \geq \psi$  a.e. in  $\Omega$ . For the precise assumptions on the quantities in (P), see Section 2.

We establish that, if the bound  $\psi$  and the operator  $A$  satisfy the compatibility condition  $A\psi \leq 0$  a.e. in  $\Omega$ , then local minimizers  $\bar{u} \in L^2(0, T; L^2(\Omega))$  of (P) are also solutions of a stationarity system (roughly of weak type, cf. [34] and [56], Sect. 4.1) which involves an adjoint state  $\bar{p} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*)$  and a multiplier  $\bar{\mu} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)) \cap W_0(0, T)^*$ . Here,  $Y_\gamma$  and  $W_0(0, T)$  are defined by  $Y_\gamma := H_0^1(\Omega) \cap H^\gamma(\Omega)$ ,  $\gamma > d/2$ , and  $W_0(0, T) := \{z \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \mid z(0) = 0 \text{ a.e. in } \Omega\}$ , respectively, and  $\mathcal{M}([0, T] \times \text{cl}(\Omega))$  denotes the space of regular Borel measures on  $[0, T] \times \text{cl}(\Omega)$ . We further demonstrate that the adjoint state  $\bar{p}$  is an element of  $L^\infty((0, T) \times \Omega)$  when the spatial dimension  $d$  is smaller than four and  $\phi$  is sufficiently smooth, that optimal controls  $\bar{u}$  of (P) inherit all of the regularity properties of  $\bar{p}$  when  $g$  is equal to  $\|\cdot\|_{L^2(\Omega)}^2$ , and that a full KKT-system (involving all of the aforementioned regularities) can be recovered for (P) when the functions  $j$ ,  $g$  and  $\phi$  are differentiable. Note that these findings significantly improve what was previously known about the  $L^r(0, T; W^{1,q}(\Omega))$ -regularity properties of adjoint states and optimal controls of problems of the type (P). See below for a more detailed discussion of this topic. For the main results of this paper on the properties of local solutions  $\bar{u}$  of (P) and first-order necessary optimality conditions, we refer the reader to Theorems 3.3 and 3.4 and Corollaries 3.5 and 3.6. As an application of our regularity results for  $\bar{u}$ ,  $\bar{p}$  etc., we derive new finite element error estimates for a standard dG(0)-cG(1)-discretization of a purely state-constrained linear-quadratic optimal control problem governed by the heat equation. The order of convergence that we establish in this context -  $\|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))} = \mathcal{O}(k^{1/2} + h)$  up to logarithmic factors - is the same as that proved for optimal control problems with integral constraints on the states  $y$  or  $L^\infty$ -constraints on the controls  $u$  in [32, 41, 42]. For further details on this topic and our precise assumptions on the discretization etc., see Section 4.

Before we begin with our analysis, let us give some background: Recall that, in the classical theory, the derivation of a KKT-system for an optimal control problem with an affine linear control-to-state operator and a convex state constraint of the form  $y \in Y_{ad}$  typically requires the existence of a Slater point, *i.e.*, of an admissible control  $u$  whose state  $y$  is contained in the interior of the set  $Y_{ad}$  of all admissible states (defined w.r.t. a suitable topology). Compare, for instance, with the results in [2, 5, 8, 12, 13, 20, 22, 28, 29, 32, 41, 42, 44, 48, 49] in this context. For the problem (P) with its parabolic PDE and its pointwise-a.e. constraint  $y \geq \psi$ , this causes a lot of difficulties. The most severe are the following:

- For the set of all functions  $y$  with  $y \geq \psi$  a.e. in  $\Omega$  for a.a.  $t \in (0, T)$  to have a non-empty interior, one typically has to work at least with the topology of the space  $L^\infty((0, T) \times \Omega)$ , and this, in turn, is usually only possible if the solution operator  $S: u \mapsto y$  of the governing partial differential equation maps the set of admissible controls into the essentially bounded functions. In the case of the problem (P), the latter requirement can, in general, only be guaranteed in the one-dimensional setting or in the presence of suitable control constraints so that the applicability of Slater-type conditions is often severely limited.
- Since one usually has to work with the space  $L^\infty((0, T) \times \Omega)$  (or  $C([0, T] \times \text{cl}(\Omega))$ , respectively) to be able to satisfy a Slater condition for an inequality of the form  $y \geq \psi$  a.e. in  $\Omega$  for a.a.  $t \in (0, T)$ , the Lagrange multiplier, that is obtained for such a pointwise state constraint by means of the classical theory, is typically only a measure on  $[0, T] \times \text{cl}(\Omega)$ . This implies in particular that the adjoint state  $\bar{p}$  in a KKT-system, that is derived for a problem of the type (P) by exploiting a Slater condition, can, in general, only be shown to possess the comparatively poor regularity  $\bar{p} \in L^r(0, T; W^{1,q}(\Omega))$  for all  $r, q \in [1, 2)$  with  $2/r + d/q > d + 1$ , cf. ([11], Sect. 6) and the results in [8, 20, 22, 32]. Since optimal controls commonly inherit the regularity properties of the associated adjoint states, the suboptimal regularity results for  $\bar{p}$  are often problematic, *e.g.*, in the context of *a priori* finite element error estimates.

For additional comments on this topic, see also the discussions in the introductions of [31, 37, 49, 50].

The strategy that is most commonly used in the literature to circumvent the above difficulties surrounding the Slater condition in the analysis of parabolic distributed optimal control problems of the type (P) is to employ regularization or penalization techniques that either incorporate the constraint  $y \geq \psi$  into the objective function or introduce artificial bounds on the controls  $u$  which ensure a higher regularity of the attainable states  $y$ . Compare, for instance, with the results in [5, 6, 37, 44, 46, 48–50] in this context where Lavrentiev- and Moreau-Yosida-type regularizations are studied, and with [20, 22, 41] where the control space is restricted a priori. For an approach alternative to regularization/restriction techniques, see also [52], where it is proposed to work simultaneously with two different functional analytic settings - one which allows to establish the existence of an optimal control and one which allows to satisfy a Slater condition - to derive KKT-type stationarity systems for problems with states of low regularity.

What (at least to the best of the authors' knowledge) all strategies currently found in the literature have in common is that they do not yield any regularity properties for the optimal controls  $\bar{u}$  or the adjoint states  $\bar{p}$  of a problem of the type (P) that go beyond  $\bar{u}, \bar{p} \in L^2(0, T; L^2(\Omega))$  or the already mentioned  $\bar{p} \in L^r(0, T; W^{1,q}(\Omega))$  for all  $r, q \in [1, 2)$  with  $2/r + d/q > d + 1$ , *cf.*, *e.g.*, the results in [37, 49, 52]. This is very unsatisfactory not only because additional information about the regularity of  $\bar{u}$  and  $\bar{p}$  is crucial for a proper analysis of numerical solution procedures for problems of the type (P), but also because, for optimal control problems governed by elliptic partial differential equations with pointwise state constraints, the adjoint state  $\bar{p}$  can often be shown to possess far more regularity than one would expect for a function that is characterized by a PDE with a measure on the right-hand side (*i.e.*, the adjoint equation). Compare, for instance, with ([12], Thm. 3.1) in this regard, which establishes that the adjoint state is an element of  $H^1(\Omega) \cap L^\infty(\Omega)$  in the elliptic case and thus drastically improves the  $W^{1,q}(\Omega)$ -regularity for all  $1 \leq q < d/(d-1)$  that is obtained from the adjoint equation, and also with the  $W^{1,\infty}(\Omega)$ -regularity result for optimal controls  $\bar{u}$  in ([13], Thm. 4.2). (Note that the proofs in [12, 13] rely heavily on the existence of a Slater point, the classical KKT-theory and the ellipticity of the governing partial differential equation. It thus does not seem to be possible to extend the arguments used in [12, 13] to the parabolic setting.)

In the present paper, we demonstrate that it is indeed possible to derive improved regularity results similar to those in [12, 13] for parabolic optimal control problems of the type (P) when the bound  $\psi$  in the state constraint and the elliptic part  $A$  of the governing PDE satisfy  $A\psi \leq 0$  a.e. in  $\Omega$  (see Thm. 3.3, Thm. 3.4 and Cor. 3.5). We further show that these results can be established without ever working with a Slater condition so that our approach also automatically resolves all of the aforementioned difficulties concerning the existence of Slater points in the parabolic setting. In the situation where the functions  $j$ ,  $g$  and  $\phi$  are differentiable, we are moreover able to recover a full KKT-system for the problem (P) involving all of the classical complementarity and slackness conditions (see Corollary 3.6). This shows that our assumption  $A\psi \leq 0$  a.e. in  $\Omega$  can be interpreted as a non-standard constraint qualification that offers an alternative to the classical Slater condition. (Note that, analogously to the existence of a Slater point, the condition  $A\psi \leq 0$  a.e. in  $\Omega$  can be verified *a priori* and without any knowledge about the optimal controls of (P), and that the assumption  $A\psi \leq 0$  a.e. in  $\Omega$  is trivially satisfied for all constant bounds  $\psi$ .) The main idea of our analysis is to incorporate the state constraint  $y \geq \psi$  in (P) into the control-to-state mapping  $S: u \mapsto y$ . This results in an optimal control problem governed by an evolution variational inequality which can be analyzed very effectively with regularization techniques. We remark that, in the elliptic setting, a similar argumentation has already been used in ([18], Cor. 5.5). For more details on this topic, see also the discussion after the proof of Proposition 3.2. At least to the best of the authors' knowledge, the regularity results and first-order necessary optimality conditions established in this paper are new and have not been explored so far in the literature. The same seems to be the case for the finite element error estimate (4.15) that we derive as an application of our regularity results in Section 4 for a purely state-constrained linear-quadratic optimal control problem governed by the heat equation. We remark that, for problems involving  $L^\infty$ -constraints on the control, estimates analogous to (4.15) have already been obtained in [32].

We conclude this introduction with a short overview of the structure and the content of the paper.

After briefly clarifying the notation that we use in our analysis in Section 1.1, we discuss preliminary results on the existence of solutions to (P) and the properties of parabolic partial differential equations in Section 2. Here, we also summarize our standing assumptions on the quantities in (P), see Assumption 2.1.

In Section 3, we then show that, under the assumption  $A\psi \leq 0$  a.e. in  $\Omega$ , problems of the type (P) can indeed be reformulated as optimal control problems governed by evolution variational inequalities, and that this reformulation indeed allows to derive first-order necessary optimality conditions and improved regularity results for optimal controls and adjoint states (see Thm. 3.3 and the subsequent corollaries). Section 3 also contains three tangible, self-contained examples that illustrate what our results mean in practice.

In Section 4, we demonstrate that the regularity results established in Section 3 allow to derive new *a priori* finite element error estimates for purely state-constrained parabolic optimal control problems. The approach that we employ in this section roughly follows the lines of ([12], Sect. 5) - see Theorem 4.8 for the main result. Section 4 concludes with numerical experiments that confirm our theoretical findings, a corollary on the  $H^s(0, T; L^2(\Omega))$ -regularity of optimal controls, and some remarks on possible generalizations of our analysis.

The appendix of this paper collects some auxiliary results on stability properties of parabolic partial differential equations and truncation operations in Sobolev–Bochner spaces that are needed for the analysis of Sections 2–4. The theorems found here may also be of independent interest.

### 1.1. Remarks on the notation

In what follows, we use the standard symbols  $L^q(\Omega)$ ,  $C^{k,\gamma}(\text{cl}(\Omega))$ ,  $H_0^s(\Omega)$ ,  $H^s(\Omega)$  and  $W^{k,q}(\Omega)$ ,  $1 \leq q \leq \infty$ ,  $k \in \mathbb{N}$ ,  $0 < \gamma \leq 1$ ,  $s > 0$ , for the Lebesgue-, Hölder- and (fractional) Sobolev spaces on a bounded Lipschitz domain  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , respectively. Given a Banach space  $Z$  and a  $T > 0$ , we further denote with  $Z^*$  the topological dual of  $Z$  and with  $L^q(0, T; Z)$ ,  $W^{k,q}(0, T; Z)$ ,  $H^s(0, T; Z)$ ,  $BV([0, T]; Z)$  and  $C([0, T]; Z)$ ,  $1 \leq q \leq \infty$ ,  $k \in \mathbb{N}$ ,  $s > 0$ , the Lebesgue-Bochner spaces, the Sobolev-Bochner spaces, the space of functions of bounded variation, and the space of continuous functions with values in  $Z$ , respectively. Recall that, in the special case  $Z = L^2(\Omega)$ , we have  $L^2(0, T; L^2(\Omega)) \cong L^2((0, T) \times \Omega)$ . This will be exploited frequently in our analysis. For later use, we also introduce the notation

$$W_0(0, T) := \{z \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \mid z(0) = 0 \text{ a.e. in } \Omega\} \quad (1.1)$$

and  $Y_\gamma := H_0^1(\Omega) \cap H^\gamma(\Omega)$  for all  $\gamma > 0$ . As usual, when working with the spaces  $H_0^1(\Omega)$  and  $C(K)$  on an open set  $\Omega \subset \mathbb{R}^d$  and a compact set  $K \subset \mathbb{R}^d$ , we define  $H^{-1}(\Omega)$  to be the dual of  $H_0^1(\Omega)$  with pivot space  $L^2(\Omega)$  and identify  $C(K)^*$  with the space  $\mathcal{M}(K)$  of signed, regular Borel measures on  $K$  *via* the Riesz representation theorem. For details on this topic and the precise definitions of all of the above spaces, see [1, 3, 4, 27, 35, 47]. Norms, scalar products and dual pairings are denoted by  $\|\cdot\|$ ,  $(\cdot, \cdot)$  and  $\langle \cdot, \cdot \rangle$  in this paper, equipped with suitable indices that specify the space we are referring to. For the Euclidean norm, we also use the notation  $|\cdot|$ . The boundary and the topological closure of a set are denoted by  $\partial(\cdot)$  and  $\text{cl}(\cdot)$ , respectively. If we want to emphasize that the closure is taken w.r.t. a certain topology, then we again add a suitable subscript. The arrows  $\rightharpoonup$  and  $\rightarrow$  indicate weak and strong convergence, and the symbols  $\partial$  and  $\partial_c$  are used for the convex and the Clarke subdifferential in the sense of ([24], Def. I-5.1) and ([19], Sect. 2.1). With  $\partial_t$ ,  $\partial_n$ ,  $n = 1, \dots, d$ ,  $\nabla$ ,  $\Delta$ ,  $\text{tr}(\cdot)$ ,  $\mathcal{L}^d$ ,  $\sigma(\cdot)$  and  $\mathcal{O}(\cdot)$ , we denote the weak time derivative in the Sobolev-Bochner sense, the weak spatial partial derivatives, the weak spatial gradient, the weak spatial Laplace operator, the spatial trace operator (with varying domains of definition), the  $d$ -dimensional Lebesgue measure, and the classical Landau symbols, respectively. A prime is used for Gâteaux and Fréchet derivatives in the functional analytic sense. Given a function  $v: (0, T) \times \Omega \rightarrow \mathbb{R}$  and a measurable set  $D \subset (0, T) \times \Omega$ , we further define  $\{v * 0\}$ ,  $* \in \{=, \neq, <, >, \leq, \geq\}$ , to be the set  $\{(t, x) \in (0, T) \times \Omega \mid v(t, x) * 0\}$  and  $\mathbf{1}_D: (0, T) \times \Omega \rightarrow \{0, 1\}$  to be the indicator function of  $D$ . Where appropriate, we consider  $\{v * 0\}$  to be defined up to sets of measure zero and identify  $\mathbf{1}_D$  with an element of  $L^\infty((0, T) \times \Omega)$ . Finally, we set  $a^+ := \max(0, a)$  and  $a^- := \min(0, a)$  for all  $a \in \mathbb{R}$ . Note that, in addition to the above definitions and conventions, further symbols etc. are introduced in this paper wherever necessary. This supplementary notation is defined where it first appears in the text.

## 2. PROBLEM STATEMENT AND PRELIMINARY RESULTS

As already mentioned in the introduction, the main goal of this paper is to study state-constrained, parabolic, distributed optimal control problems of the type

$$\left. \begin{aligned} \text{Minimize } & J(y, u) := \int_0^T (j(\cdot, y) + g(u)) dt + \phi(y(T)) \\ \text{w.r.t. } & u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\ \text{s.t. } & \partial_t y + Ay = u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ & \text{tr}(y) = 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\ & y(0) = y_0 \quad \text{a.e. in } \Omega, \\ \text{and } & y \geq \psi \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T). \end{aligned} \right\} \quad (\text{P})$$

Our standing assumptions on the quantities in (P) are as follows:

**Assumption 2.1** (Standing assumptions for the study of problem (P)).

- $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , is a bounded domain, which is convex or possesses a  $C^{1,1}$ -boundary, and  $T > 0$  is a given final time.
- $j: [0, T] \times L^2(\Omega) \rightarrow [0, \infty)$  is a function with the following properties:
  - (i)  $j(\cdot, z)$  is Lebesgue measurable for all  $z \in L^2(\Omega)$ ,
  - (ii)  $j(\cdot, 0)$  is essentially bounded,
  - (iii) for every  $r > 0$  there exists a constant  $C_r > 0$  independent of  $t$  with

$$|j(t, z_1) - j(t, z_2)| \leq C_r \|z_1 - z_2\|_{L^2(\Omega)}$$

for all  $t \in [0, T]$  and all  $z_1, z_2 \in L^2(\Omega)$  with  $\|z_1\|_{L^2(\Omega)}, \|z_2\|_{L^2(\Omega)} \leq r$ .

- $g: L^2(\Omega) \rightarrow (-\infty, \infty]$  is a function with the following properties:
  - (i)  $g$  is convex, proper and lower semicontinuous,
  - (ii) there exist constants  $C_1 > 0$ ,  $C_2 \in \mathbb{R}$  with

$$g(z) \geq C_1 \|z\|_{L^2(\Omega)}^2 + C_2 \quad \forall z \in L^2(\Omega),$$

- (iii) for all  $z_1, z_2 \in L^2(\Omega)$  which satisfy  $z_1^+ = z_2^+$  and  $z_2^- \leq z_1^-$  a.e. in  $\Omega$ , it holds  $g(z_1) \leq g(z_2)$ . (Recall that  $z^- := \min(0, z) \leq 0$  so that  $z_2^- \leq z_1^-$  a.e. in  $\Omega$  is equivalent to  $|z_1^-| \leq |z_2^-|$  a.e. in  $\Omega$ .)
- $\phi: L^2(\Omega) \rightarrow [0, \infty)$  is a function which is Lipschitz on bounded sets.
- $A$  is a second-order partial differential operator of the form

$$Av = - \sum_{m,n=1}^d \partial_m (a_{mn} \partial_n v) + a_0 v \quad (2.1)$$

with coefficients  $a_{mn} \in C^{0,1}(\text{cl}(\Omega))$ ,  $a_0 \in L^\infty(\Omega)$  such that

$$a_0 \geq 0, \quad a_{mn} = a_{nm} \quad \text{and} \quad \sum_{m,n=1}^d a_{mn} \xi_m \xi_n \geq \alpha |\xi|^2 \quad (2.2)$$

holds a.e. in  $\Omega$  for all  $m, n = 1, \dots, d$  and all  $\xi \in \mathbb{R}^d$  with some  $\alpha > 0$ . In what follows, we interpret  $A$  as a mapping  $A: H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$  or  $A: H^2(\Omega) \rightarrow L^2(\Omega)$  depending on the situation.

- $\psi \in H^2(\Omega)$  is a given function with  $\text{tr}(\psi) \leq 0$  a.e. on  $\partial\Omega$ .
- $y_0 \in H_0^1(\Omega)$  is a given function with  $y_0 \geq \psi$  a.e. in  $\Omega$ .

We remark that we could also consider other boundary conditions here. We restrict the analysis to homogeneous Dirichlet conditions for the sake of simplicity. The assumptions, that we have made above, are quite natural for the study of optimal control problems of the type (P). Only the third condition on  $g$  may appear a bit unusual at first glance. It will become clear in the proof of Proposition 3.2 why we restrict our analysis to functions  $g$  with this property. To demonstrate that the setting introduced in Assumption 2.1 is quite general, we note the following:

**Example 2.2** (Possible choices for  $j$  and  $g$ ). All of the following functions fit into the framework of Assumption 2.1 (as one may easily check):

- The  $L^2$ -tracking term

$$j: [0, T] \times L^2(\Omega) \rightarrow [0, \infty), \quad j(t, z) := \frac{1}{2} \|z - y_D(t)\|_{L^2(\Omega)}^2,$$

with some arbitrary but fixed measurable and bounded function  $y_D: [0, T] \rightarrow L^2(\Omega)$ .

- The  $L^1$ -tracking term

$$j: [0, T] \times L^2(\Omega) \rightarrow [0, \infty), \quad j(t, z) := \|z - y_D(t)\|_{L^1(\Omega)},$$

with some arbitrary but fixed measurable and bounded function  $y_D: [0, T] \rightarrow L^1(\Omega)$ .

- The  $L^2$ -Tikhonov regularization term with pointwise-a.e. control constraints

$$g: L^2(\Omega) \rightarrow [0, \infty], \quad g(z) := \begin{cases} \frac{\nu}{2} \|z\|_{L^2(\Omega)}^2 & \text{if } u_a \leq z \leq u_b \text{ a.e. in } \Omega, \\ \infty & \text{else} \end{cases},$$

with some arbitrary but fixed measurable functions  $u_a, u_b: \Omega \rightarrow [-\infty, \infty]$  with  $u_a \leq 0 \leq u_b$  a.e. in  $\Omega$  and some  $\nu > 0$ . (The choice  $u_a \equiv -\infty$ ,  $u_b \equiv \infty$ , which corresponds to the case without control constraints, is allowed here.)

- The  $L^1$ -Tikhonov regularization term with pointwise-a.e. control constraints

$$g: L^2(\Omega) \rightarrow [0, \infty], \quad g(z) := \begin{cases} \nu \|z\|_{L^1(\Omega)} & \text{if } u_a \leq z \leq u_b \text{ a.e. in } \Omega, \\ \infty & \text{else} \end{cases},$$

with some arbitrary but fixed  $u_a, u_b \in L^2(\Omega)$  satisfying  $u_a \leq 0 \leq u_b$  a.e. in  $\Omega$  and some  $\nu \geq 0$ . (The “bang-bang” case  $\nu = 0$  is allowed here.)

For some examples of problems satisfying the conditions in Assumption 2.1, see also Section 3. To prove that (P) is well-posed, we recall:

**Proposition 2.3** (Properties of the PDE in (P)). *For every  $u \in L^2(0, T; L^2(\Omega))$ , there exists a unique solution  $y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))$  of the PDE*

$$\begin{aligned} \partial_t y + Ay &= u && \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ \text{tr}(y) &= 0 && \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\ y(0) &= y_0 && \text{a.e. in } \Omega. \end{aligned} \tag{2.3}$$

This solution also possesses  $C([0, T]; H_0^1(\Omega))$ -regularity, and there exists a constant  $C > 0$  depending only on  $A$ ,  $\Omega$ , and  $T$  such that the solution map  $S: u \mapsto y$  associated with (2.3) satisfies

$$\|S(u_1) - S(u_2)\|_{L^2(0, T; H_0^1(\Omega))} + \|S(u_1) - S(u_2)\|_{C([0, T]; L^2(\Omega))} \leq C \|u_1 - u_2\|_{L^1(0, T; L^2(\Omega))} \quad (2.4)$$

$$\forall u_1, u_2 \in L^2(0, T; L^2(\Omega)).$$

Further, the operator  $S$  is completely continuous as a function from  $L^2(0, T; L^2(\Omega))$  to  $C([0, T]; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega))$ , i.e., the convergence  $u_n \rightharpoonup u$  in  $L^2(0, T; L^2(\Omega))$  implies  $S(u_n) \rightarrow S(u)$  in  $L^2(0, T; H_0^1(\Omega))$  and  $C([0, T]; L^2(\Omega))$ .

*Proof.* The unique solvability of the PDE (2.3), the  $C([0, T]; H_0^1(\Omega))$ -regularity of the solution  $y$  and the Lipschitz estimate (2.4) follow from ([4], Thm. 4.3), ([15], Thm. 2.3) and the regularity results for the operator  $A$  in ([33], Thm. 3.2.1.2), ([30], Thm. 9.15). To establish the complete continuity of  $S$ , one can use (for example) the Lipschitz estimate (2.4), the theorem of Aubin-Lions (see [53], Thm. 10.12), and a simple bootstrapping argument. We leave the details to the reader as the arguments are fairly standard.  $\square$

As a direct consequence of the last result, we obtain:

**Proposition 2.4** (Solvability of (P)). *Suppose that there exists a function  $u \in L^2(0, T; L^2(\Omega))$  satisfying*

$$S(u) \geq \psi \text{ a.e. in } \Omega \text{ for a.a. } t \in (0, T) \quad \text{and} \quad \int_0^T g(u) dt < \infty. \quad (2.5)$$

*Then the problem (P) admits at least one global solution  $\bar{u} \in L^2(0, T; L^2(\Omega))$ .*

*Proof.* From the properties of  $g$  and ([14], Lem. 2.5.2), we obtain that the map  $G: L^2(0, T; L^2(\Omega)) \rightarrow (-\infty, \infty]$ ,  $z \mapsto \int_0^T g(z) dt$ , is well-defined, convex, lower semicontinuous, proper and coercive in the sense that there exist constants  $C_1 > 0$  and  $C_2 \in \mathbb{R}$  with  $G(z) \geq C_1 \|z\|_{L^2(0, T; L^2(\Omega))}^2 + C_2$  for all  $z \in L^2(0, T; L^2(\Omega))$ . Proposition 2.3 and our standing assumptions on the functions  $j$  and  $\phi$  further imply that the map  $H: L^2(0, T; L^2(\Omega)) \rightarrow [0, \infty)$ ,  $z \mapsto \int_0^T j(\cdot, S(z)) dt + \phi(S(z)(T))$ , is well-defined and completely continuous. The claim now follows straightforwardly from the direct method of calculus of variations.  $\square$

Note that (2.5) just expresses that there should be at least one control  $u \in L^2(0, T; L^2(\Omega))$  that is admissible for (P) and yields a finite value of the reduced objective function  $J(S(u), u)$ . If such a point does not exist, then the objective function is identical  $+\infty$  on the admissible set of (P) and discussing this problem is clearly not very sensible. In practice, the existence of a control with the property (2.5) is often evident from the start or can easily be established, e.g., with a maximum principle. Compare, for instance, with Examples 3.9–3.11 in this context. We would like to point out that, here and in what follows, we only refer to those points as local/global solutions of (P) that satisfy a local/global optimality condition and yield a finite value of the objective function  $J$ , i.e., we work with the following definition:

**Definition 2.5** (Optima of (P)). A point  $\bar{u} \in L^2(0, T; L^2(\Omega))$  is called a local solution of (P) if it is feasible in the sense of (2.5) and if there exists an  $r > 0$  such that  $J(S(\bar{u}), \bar{u}) \leq J(S(u), u)$  holds for all  $u \in L^2(0, T; L^2(\Omega))$  which satisfy (2.5) and  $\|u - \bar{u}\|_{L^2(0, T; L^2(\Omega))} < r$ . If  $r$  can be chosen as infinity, then we call  $\bar{u}$  a global solution.

### 3. FIRST-ORDER OPTIMALITY CONDITIONS AND REGULARITY RESULTS

Having discussed the solvability of the problem (P), we now turn our attention to first-order necessary optimality conditions and questions of regularity. As already explained in the introduction, the approach, that we use in the following to study the properties of optimal controls  $\bar{u}$  of (P), is based on the idea to completely avoid working with a Slater condition and to exploit an alternative constraint qualification instead that does not suffer from the problems discussed in Section 1. (The term ‘‘constraint qualification’’ is actually a bit of

a stretch here since the proofs of Proposition 3.2 and Theorem 3.3 below also rely on the assumptions on the structure of the objective  $J$  of (P) made in Section 2. We use this expression for lack of better terminology.) The condition, that is at the heart of our analysis, reads as follows:

**Definition 3.1** (Condition (X)). The problem (P) is said to satisfy (X) if

$$A\psi \leq 0 \text{ a.e. in } \Omega.$$

Note that condition (X) can be checked *a priori* and that (X) is trivially satisfied for all constant bounds  $\psi \equiv c$  with  $c \in (-\infty, 0]$ . In particular, (X) also covers the case  $\psi \equiv 0$  in which a Slater point cannot exist for (P) in the continuous functions due to the zero boundary conditions in the governing PDE. The main reason for the usefulness of condition (X) is the following observation:

**Proposition 3.2** (Reformulation of (P) under condition (X)). *Suppose that  $\bar{u} \in L^2(0, T; L^2(\Omega))$  is a local (respectively, global) solution of the problem (P) with state  $\bar{y} := S(\bar{u}) \in L^2(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))$ . Then it holds*

$$\bar{u} = \begin{cases} \partial_t \bar{y} + A\bar{y} & \mathcal{L}^{d+1}\text{-a.e. in } \{\bar{y} > \psi\} \\ A\psi & \mathcal{L}^{d+1}\text{-a.e. in } \{\bar{y} = \psi\} \end{cases}, \quad (3.1)$$

and, if (X) is satisfied, then  $\bar{u}$  is also a local (respectively, global) solution of the problem

$$\left. \begin{aligned} \text{Minimize } & J(y, u) := \int_0^T (j(\cdot, y) + g(u)) dt + \phi(y(T)) \\ \text{w.r.t. } & u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\ \text{s.t. } & y(t) \in K \text{ for a.a. } t \in (0, T), \quad y(0) = y_0 \text{ a.e. in } \Omega, \\ & \int_0^T \langle \partial_t y + Ay - u, v - y \rangle_{H_0^1(\Omega)} dt \geq 0 \\ & \forall v \in L^2(0, T; H_0^1(\Omega)), \quad v(t) \in K \text{ for a.a. } t \in (0, T), \end{aligned} \right\} \quad (Q)$$

where  $K$  is the set defined by  $K := \{z \in H_0^1(\Omega) \mid z \geq \psi \text{ a.e. in } \Omega\}$ .

*Proof.* We proceed in several steps: First, we note that (2.3), Lemma A.1, and the regularity results for the state  $\bar{y} := S(\bar{u})$  in Proposition 2.3 yield that  $\bar{u} = \partial_t \bar{y} + A\bar{y}$  holds  $\mathcal{L}^{d+1}$ -a.e. in  $(0, T) \times \Omega$  and that  $\partial_t(\bar{y} - \psi) + A(\bar{y} - \psi) = 0$  holds  $\mathcal{L}^{d+1}$ -a.e. in  $\{\bar{y} = \psi\}$ . (Here and in what follows, we identify elements of  $L^2(\Omega)$  with elements of  $L^2(0, T; L^2(\Omega))$  that are constant in time everywhere where it is appropriate.) If we combine the last two identities, then we immediately arrive at (3.1). This proves the first claim of the proposition.

Next, let us consider the inner variational inequality in (Q), *i.e.*, the problem

$$\begin{aligned} & y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\ & y(t) \in K \text{ for a.a. } t \in (0, T), \quad y(0) = y_0 \text{ a.e. in } \Omega, \\ & \int_0^T \langle \partial_t y + Ay - u, v - y \rangle_{H_0^1(\Omega)} dt \geq 0 \\ & \forall v \in L^2(0, T; H_0^1(\Omega)), \quad v(t) \in K \text{ for a.a. } t \in (0, T). \end{aligned} \quad (3.2)$$

From ([4], Thm. 4.3, Cor. 4.4) and ([15], Thm 2.3), we obtain that this evolution variational inequality possesses a well-defined solution map  $R: L^2(0, T; L^2(\Omega)) \rightarrow L^2(0, T; H^2(\Omega)) \cap C([0, T]; H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$ ,  $u \mapsto y$ , that is globally Lipschitz as a function from  $L^2(0, T; L^2(\Omega))$  to  $L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega))$ . Suppose



now that we are given some  $u \in L^2(0, T; L^2(\Omega))$  with associated  $y := R(u)$ , and let  $\tilde{u} \in L^2(0, T; L^2(\Omega)) \cong L^2((0, T) \times \Omega)$  be defined by

$$\tilde{u} := \begin{cases} u & \mathcal{L}^{d+1}\text{-a.e. in } \{y > \psi\} \\ A\psi & \mathcal{L}^{d+1}\text{-a.e. in } \{y = \psi\} \end{cases}. \quad (3.3)$$

We claim that, if condition **(X)** holds, then this modified control satisfies

$$R(u) = R(\tilde{u}) = S(\tilde{u}) = y \quad \text{and} \quad \int_0^T g(\tilde{u})dt \leq \int_0^T g(u)dt, \quad (3.4)$$

where the identity in (3.4) is to be understood as an equality between elements of the space  $L^2(0, T; H^2(\Omega)) \cap C([0, T]; H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$ . Indeed, if **(X)** is satisfied, then it follows straightforwardly from (3.2), the regularity of  $y$ , Lemma A.1 and the same arguments as in the first part of the proof that  $\tilde{u} = \partial_t y + Ay$  holds  $\mathcal{L}^{d+1}$ -a.e. in  $(0, T) \times \Omega$ . This yields  $R(\tilde{u}) = S(\tilde{u}) = R(u) = y$  (as one may easily check by plugging this formula into (3.2)). Further, we obtain from (3.2) and again the same arguments as in the first part of the proof that  $\tilde{u} - u = A\psi - u = \partial_t y + Ay - u \geq 0$  holds  $\mathcal{L}^{d+1}$ -a.e. in  $\{y = \psi\}$  and, as a consequence, that  $0 \geq A\psi = \tilde{u} \geq u$  holds  $\mathcal{L}^{d+1}$ -a.e. in  $\{y = \psi\}$ . Here, the inequality  $A\psi \leq 0$  follows from condition **(X)**. The above implies that the positive and negative parts of  $u$  and  $\tilde{u}$  satisfy  $u^+ = \tilde{u}^+$  and  $u^- \leq \tilde{u}^-$   $\mathcal{L}^{d+1}$ -a.e. in  $(0, T) \times \Omega$ , and, by our assumptions on  $g$ , that  $\int_0^T g(\tilde{u})dt \leq \int_0^T g(u)dt$  holds. This establishes (3.4) as desired.

Let us assume now that **(X)** holds, that  $\bar{u} \in L^2(0, T; L^2(\Omega))$  is a global solution of **(P)**, that  $u \in L^2(0, T; L^2(\Omega))$  is arbitrary but fixed and that  $\tilde{u}$  and  $R$  are defined as in (3.3) and (3.4). Then the optimality of  $\bar{u}$  in **(P)**, the fact that elements of the space  $H^1(0, T; L^2(\Omega))$  that are identical necessarily also have the same traces w.r.t.  $t$ , and our previous results yield

$$\begin{aligned} & \int_0^T \left( j(\cdot, R(u)) + g(u) \right) dt + \phi(R(u)(T)) \\ & \geq \int_0^T \left( j(\cdot, R(u)) + g(\tilde{u}) \right) dt + \phi(R(u)(T)) = \int_0^T \left( j(\cdot, S(\tilde{u})) + g(\tilde{u}) \right) dt + \phi(S(\tilde{u})(T)) \\ & \geq \int_0^T \left( j(\cdot, S(\bar{u})) + g(\bar{u}) \right) dt + \phi(S(\bar{u})(T)) = \int_0^T \left( j(\cdot, R(\bar{u})) + g(\bar{u}) \right) dt + \phi(R(\bar{u})(T)), \end{aligned}$$

where, in the last line, we have used the trivial identity  $S(\bar{u}) = R(\bar{u})$ , cf. (2.3) and (3.2). This proves that, if **(X)** holds, then every global solution  $\bar{u}$  of **(P)** is also a global solution of **(Q)**.

It remains to prove the claim for local solutions. To this end, we argue by contradiction: Suppose that **(X)** holds and that  $\bar{u} \in L^2(0, T; L^2(\Omega))$  is locally optimal for **(P)** with state  $\bar{y} = S(\bar{u}) = R(\bar{u})$  but not a local solution of **(Q)**. Then there exists a sequence  $\{u_n\} \subset L^2(0, T; L^2(\Omega))$  with  $u_n \rightarrow \bar{u}$  in  $L^2(0, T; L^2(\Omega))$  such that

$$\int_0^T \left( j(\cdot, R(u_n)) + g(u_n) \right) dt + \phi(R(u_n)(T)) < \int_0^T \left( j(\cdot, R(\bar{u})) + g(\bar{u}) \right) dt + \phi(R(\bar{u})(T)) \quad (3.5)$$

holds for all  $n$  where  $R$  again denotes the solution map of (3.2). Define  $y_n := R(u_n)$  and (as in (3.3))

$$\tilde{u}_n := \begin{cases} u_n & \mathcal{L}^{d+1}\text{-a.e. in } \{y_n > \psi\} \\ A\psi & \mathcal{L}^{d+1}\text{-a.e. in } \{y_n = \psi\} \end{cases}.$$

Then we obtain from (3.4) that these modified controls satisfy  $S(\tilde{u}_n) = R(\tilde{u}_n) = R(u_n) = y_n$  and

$$\int_0^T \left( j(\cdot, R(\tilde{u}_n)) + g(\tilde{u}_n) \right) dt + \phi(R(\tilde{u}_n)(T)) \leq \int_0^T \left( j(\cdot, R(u_n)) + g(u_n) \right) dt + \phi(R(u_n)(T)). \quad (3.6)$$

We claim that the sequence  $\{\tilde{u}_n\}$  converges to  $\bar{u}$  in  $L^2(0, T; L^2(\Omega))$ . To see this, we note that (3.1) implies

$$\bar{u} = \begin{cases} \bar{u} & \mathcal{L}^{d+1}\text{-a.e. in } \{\bar{y} > \psi\} \\ A\psi & \mathcal{L}^{d+1}\text{-a.e. in } \{\bar{y} = \psi\} \end{cases}.$$

In tandem with the definition of  $\tilde{u}_n$  and the identification  $L^2(0, T; L^2(\Omega)) \cong L^2((0, T) \times \Omega)$ , the above yields

$$\begin{aligned} \|\tilde{u}_n - \bar{u}\|_{L^2(0, T; L^2(\Omega))} &\leq \|\mathbb{1}_{\{y_n > \psi\}}(u_n - \bar{u})\|_{L^2((0, T) \times \Omega)} + \|\mathbb{1}_{\{y_n = \psi\}}(A\psi - \bar{u})\|_{L^2((0, T) \times \Omega)} \\ &\leq \|u_n - \bar{u}\|_{L^2((0, T) \times \Omega)} + \|\mathbb{1}_{\{y_n = \psi < \bar{y}\}}(A\psi - \bar{u})\|_{L^2((0, T) \times \Omega)}. \end{aligned} \quad (3.7)$$

Since  $u_n \rightarrow \bar{u}$  in  $L^2(0, T; L^2(\Omega))$  and since  $R$  is continuous from  $L^2(0, T; L^2(\Omega))$  to  $L^2(0, T; L^2(\Omega))$ , we know that  $y_n \rightarrow \bar{y}$  holds in  $L^2(0, T; L^2(\Omega))$  and thus (at least after passing over to a subsequence) that  $y_n \rightarrow \bar{y}$  pointwise a.e. in  $(0, T) \times \Omega$ . This convergence implies in particular that  $\mathbb{1}_{\{y_n = \psi < \bar{y}\}} \rightarrow 0$  pointwise a.e. and, by the dominated convergence theorem, that

$$\|\mathbb{1}_{\{y_n = \psi < \bar{y}\}}(A\psi - \bar{u})\|_{L^2((0, T) \times \Omega)} \rightarrow 0.$$

From (3.7), it now follows that  $\tilde{u}_n$  indeed converges to  $\bar{u}$  in  $L^2(0, T; L^2(\Omega))$  for  $n \rightarrow \infty$ . This convergence, in combination with (3.5), (3.6) and the assumed local optimality of  $\bar{u}$  for (P), yields a contradiction. A local solution  $\bar{u}$  of (P) is thus always locally optimal for (Q) if (X) is satisfied and the proof is complete.  $\square$

We remark that, under additional assumptions on the function  $g$ , it is also possible to show that local/global solutions of (Q) are local/global solutions of (P). For details on this topic, see ([18], Sect. 5) where an analogous equivalence is established in the elliptic case and subsequently exploited for the derivation of necessary and sufficient second-order optimality conditions.

The reader might ask at this point why it is useful to make the observation in Proposition 3.2. The control-to-state map of (Q) is, after all, not affine linear (or even differentiable) anymore so that (Q) seems to be significantly more complicated than the problem (P) that we started with. However, the problem (Q) has one fundamental advantage: Since the constraint  $y \geq \psi$  is incorporated into its control-to-state operator, (Q) can be treated far more effectively with regularization techniques than the original problem (P). To see this, note that the governing variational inequality in (Q) can also be written as

$$\begin{aligned} \partial_t y + Ay + \beta(y - \psi) &\ni u && \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ \text{tr}(y) &= 0 && \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\ y(0) &= y_0 && \text{a.e. in } \Omega, \end{aligned} \quad (3.8)$$

where  $\beta$  denotes the maximal monotone graph in  $\mathbb{R} \times \mathbb{R}$  defined by

$$\beta(r) := \begin{cases} \emptyset & \text{if } r < 0 \\ (-\infty, 0] & \text{if } r = 0 \\ 0 & \text{if } r > 0 \end{cases}.$$

Suppose now that we replace  $\beta$  by a suitable, monotonously increasing regularization  $\beta_\varepsilon: \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $\beta_\varepsilon(0) = 0$  and (in an appropriate sense)  $\beta_\varepsilon \rightarrow \beta$  for  $\varepsilon \searrow 0$ , and consider the approximate problem

$$\begin{aligned} \partial_t y + Ay + \beta_\varepsilon(y - \psi) &= u && \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ \text{tr}(y) &= 0 && \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\ y(0) &= y_0 && \text{a.e. in } \Omega. \end{aligned} \tag{3.9}$$

Then it is straightforward to derive first-order necessary optimality conditions for (Q) by replacing the variational inequality (3.8) in (Q) with (3.9) and by subsequently passing to the limit with the regularization parameter  $\varepsilon$ . What is beneficial in the above regularization approach is that, in (3.9), the  $\beta_\varepsilon$ -term constitutes a monotone part of the governing PDE. Because of this, it is possible to exploit the properties of the differential operators  $\partial_t$  and  $A$  to obtain comparatively strong,  $\varepsilon$ -independent bounds for the primal and dual quantities in the necessary optimality conditions associated with the regularized version of (Q). (We will see below that we obtain precisely  $L^2(0, T; H_0^1(\Omega))$ - and  $L^\infty(0, T; L^2(\Omega))$ -regularity here.) A similar effect is not present, when we directly mollify the problem (P), where the inequality  $y \geq \psi$  is a constraint on the upper level and thus, in a sense, separated from the governing partial differential equation. Compare, *e.g.*, with the regularization approaches in [5, 37, 48–50] in this context, where at most  $L^2(0, T; L^2(\Omega))$ -bounds are obtained.

By making the above informal argumentation rigorous, we arrive at:

**Theorem 3.3** (First-order necessary optimality condition for (P)). *Suppose that (X) holds, let  $\gamma$  be an arbitrary but fixed real number with  $\gamma > d/2$ , and let  $Y_\gamma$  and  $W_0(0, T)$  be defined as in Section 1.1. Then, for every local solution  $\bar{u} \in L^2(0, T; L^2(\Omega))$  of the optimal control problem (P) with associated state  $\bar{y} := S(\bar{u})$ , there exist an adjoint state*

$$\bar{p} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*) \tag{3.10}$$

and multipliers

$$\bar{\mu} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)) \cap W_0(0, T)^*, \quad \bar{\eta}_j, \bar{\eta}_g \in L^\infty(0, T; L^2(\Omega)), \quad \bar{\eta}_\phi \in L^2(\Omega),$$

such that

$$\begin{aligned} \bar{\eta}_j(t) &\in \partial_c j(t, \bar{y}(t)) && \text{for a.a. } t \in (0, T), \\ \bar{\eta}_g(t) &\in \partial g(\bar{u}(t)) && \text{for a.a. } t \in (0, T), \\ \bar{p}(t) + \bar{\eta}_g(t) &= 0 && \text{for a.a. } t \in (0, T), \\ \bar{\eta}_\phi &\in \partial_c \phi(\bar{y}(T)), \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} -\partial_t \bar{p} + A\bar{p} &= \bar{\eta}_j - \bar{\mu} && \text{in } (0, T) \times \Omega, \\ \text{tr}(\bar{p}) &= 0 && \text{on } (0, T) \times \partial\Omega, \\ \bar{p}(T) &= \bar{\eta}_\phi && \text{in } \Omega \end{aligned} \tag{3.12}$$

holds. Here,  $\partial_c j(t, z)$ ,  $\partial_c \phi(z)$ , and  $\partial g(z)$  denote Clarke's generalized differential and the convex subdifferential of the maps  $L^2(\Omega) \ni z \mapsto j(t, z) \in \mathbb{R}$ ,  $L^2(\Omega) \ni z \mapsto \phi(z) \in \mathbb{R}$  and  $L^2(\Omega) \ni z \mapsto g(z) \in (-\infty, \infty]$ , respectively, and the PDE for  $\bar{p}$  is understood as a formal identity in  $W_0(0, T)^*$ , *i.e.*, in the sense that, for all  $z \in W_0(0, T)$ ,

we have

$$\int_0^T \left( \langle \partial_t z + Az, \bar{p} \rangle_{H_0^1(\Omega)} - (\bar{\eta}_j, z)_{L^2(\Omega)} \right) dt + \langle \bar{\mu}, z \rangle_{W_0(0,T)} - (z(T), \bar{\eta}_\phi)_{L^2(\Omega)} = 0. \quad (3.13)$$

*Proof.* All assertions of the theorem follow straightforwardly from Proposition 3.2 and the necessary optimality conditions for problems of the type (Q) derived by regularization in ([4], Chap. 5). Compare in particular with ([4], Prop. 5.2, Lem. 5.3, 5.4, Eq. (5.28)) in this context.  $\square$

In certain situations, we can obtain even more information about the adjoint state  $\bar{p}$  in (3.10) as the following result shows:

**Theorem 3.4** (Additional information about the adjoint state). *Suppose that condition (X) is satisfied and consider the situation in Theorem 3.3.*

- (i) *If  $d \in \{1, 2, 3\}$  holds and if  $\phi$  is a  $C^1$ -function such that  $\phi': L^2(\Omega) \rightarrow L^2(\Omega)$  maps bounded subsets of  $L^2(\Omega)$  into bounded subsets of  $L^\infty(\Omega)$ , then the stationarity system in Theorem 3.3 holds true with the additional regularity*

$$\bar{p} \in L^\infty((0, T) \times \Omega). \quad (3.14)$$

- (ii) *If (in addition to the conditions in Assumption 2.1), we know that the function  $g$  is real-valued and Gâteaux differentiable with derivative  $g': L^2(\Omega) \rightarrow L^2(\Omega)$  and that the objective  $J$  of (P) is Gâteaux differentiable and locally Lipschitz continuous as a map from  $L^q(0, T; L^2(\Omega)) \times L^2(0, T; L^2(\Omega))$  to  $\mathbb{R}$  for some  $1 \leq q < \infty$  with partial derivative  $\partial_u J(y, u) = g'(u) \in L^2(0, T; L^2(\Omega))$  for all  $u \in L^2(0, T; L^2(\Omega))$ , then the adjoint state  $\bar{p}$  in (3.10) is unique and also satisfies*

$$\bar{p} \in \text{cl}_{L^2(0,T;H_0^1(\Omega))}(\mathbb{R}^+(\mathcal{K} - \bar{y})),$$

where  $\bar{y} := S(\bar{u})$  denotes the state associated with  $\bar{u}$  and where

$$\mathcal{K} := \{z \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \mid z \geq \psi \text{ a.e. in } (0, T) \times \Omega\}.$$

In particular, in this situation, we have

$$\bar{p}(t) \geq 0 \text{ } H_0^1\text{-q.e. in } \{\bar{y}(t) = \psi\} \text{ for a.a. } t \in (0, T),$$

where “q.e.” is short for quasi-everywhere in the sense of ([9], Sect. 6.4.3).

*Proof.* To obtain the  $L^\infty((0, T) \times \Omega)$ -regularity of  $\bar{p}$  in (3.14), we note that, in the situation of (i), the proof of the stationarity system in Theorem 3.3 in ([4], Chap. 5) implies that there exist functions  $0 \leq \chi_n \in L^\infty((0, T) \times \Omega)$  and bounded sequences  $\{\eta_n\} \subset L^\infty(0, T; L^2(\Omega))$ ,  $\{\vartheta_n\} \subset L^2(\Omega)$  such that the (necessarily unique) weak solutions  $p_n \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$  of the partial differential equations

$$\begin{aligned} -\partial_t p_n + A p_n + \chi_n p_n &= \eta_n && \text{in } (0, T) \times \Omega, \\ \text{tr}(p_n) &= 0 && \text{on } (0, T) \times \partial\Omega, \\ p_n(T) &= \phi'(\vartheta_n) && \text{in } \Omega, \end{aligned} \quad (3.15)$$

converge weakly in  $L^2(0, T; H_0^1(\Omega))$  and weakly- $\star$  in  $L^\infty(0, T; L^2(\Omega))$  to  $\bar{p}$  for  $n \rightarrow \infty$ . Compare in particular with ([4], Eqs. (5.28), (5.34), Lem. 5.3, Prop. 5.2) in this context and note that, due to our assumptions on  $\phi$ , the analysis of [4] holds without an additional regularization of this function. Applying Lemma A.3 to

the PDE (3.15) (or the PDE for  $\tilde{p}_n(t) := p_n(T - t)$ , to be more precise) yields that there exists a constant  $C = C(\Omega, T, d, \alpha) > 0$  with

$$\|p_n\|_{L^\infty((0,T)\times\Omega)} \leq \|\phi'(\vartheta_n)\|_{L^\infty(\Omega)} + C\|\eta_n\|_{L^\infty(0,T;L^2(\Omega))}$$

for all  $n$ . The claim in (i) now follows immediately from the fact that the last estimate, the properties of  $\{\eta_n\}$  and  $\{\vartheta_n\}$ , and our assumptions on  $\phi$  imply that there exists a constant  $\tilde{C} > 0$  independent of  $n$  satisfying  $\|p_n\|_{L^\infty((0,T)\times\Omega)} \leq \tilde{C}$ , the weak convergence of the sequence  $\{p_n\}$  in  $L^2(0, T; H_0^1(\Omega))$  to  $\bar{p}$ , and the fact that the set of all elements of the space  $L^2(0, T; H_0^1(\Omega))$  whose  $L^\infty((0, T) \times \Omega)$ -norm is smaller than  $\tilde{C}$  is convex and closed and thus weakly closed by the lemma of Mazur.

To establish (ii), it suffices to note that the assumptions on  $g$  in (ii) and Theorem 3.3 imply  $\bar{p} = -\bar{\eta}_g = -g'(\bar{u}) = -\partial_u J(\bar{y}, \bar{u})$  and to invoke Proposition 3.2, ([15], Thm. 5.5i) and classical results on tangent cones of sets with unilateral bounds in  $H_0^1(\Omega)$ , see ([45], Lem. 3.2). This completes the proof.  $\square$

We remark that the assumptions on  $\phi$  in Theorem 3.4(i) can be relaxed if it is possible to ensure along different lines (e.g., by exploiting suitable control constraints) that the quantities  $\nabla\phi_0^\varepsilon(y_\varepsilon(T))$  in ([4], Eq. (5.28)) remain bounded in  $L^\infty(\Omega)$  as  $\varepsilon$  tends to zero. We omit a detailed discussion of this topic to avoid redoing the analysis of [4] completely. As usual, the optimal control  $\bar{u}$  in Theorem 3.3 directly inherits the regularity properties of the adjoint state  $\bar{p}$  when the function  $g$  is sufficiently well-behaved. In particular, for an  $L^2$ -Tikhonov regularization term, we have:

**Corollary 3.5** (Regularity properties of optimal controls). *Suppose that (X) holds and that there exist a  $\nu > 0$  and measurable functions  $u_a, u_b: \Omega \rightarrow [-\infty, \infty]$  with  $u_a \leq 0 \leq u_b$  a.e. in  $\Omega$  such that the function  $g$  in (P) has the form*

$$g: L^2(\Omega) \rightarrow [0, \infty], \quad g(z) := \begin{cases} \frac{\nu}{2}\|z\|_{L^2(\Omega)}^2 & \text{if } u_a \leq z \leq u_b \text{ a.e. in } \Omega \\ \infty & \text{else} \end{cases}. \quad (3.16)$$

Then the following is true:

(i) If  $u_a \equiv -\infty$  and  $u_b \equiv \infty$  holds, then every local solution  $\bar{u}$  of (P) satisfies

$$\bar{u} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*)$$

with  $Y_\gamma$ ,  $\gamma > d/2$ , defined as in Section 1.1 and Theorem 3.3.

(ii) If  $u_a, u_b \in H^1(\Omega) \cup \{\pm\infty\}$  holds, then every local solution  $\bar{u}$  of (P) satisfies

$$\bar{u} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)).$$

(iii) If, in the above cases (i) and (ii), the function  $\phi$  and the spatial dimension  $d$  satisfy the assumptions of Theorem 3.4(i), then every local solution  $\bar{u}$  of (P) additionally satisfies  $\bar{u} \in L^\infty((0, T) \times \Omega)$ .

*Proof.* The assertions of Corollary 3.5 follow straightforwardly from Theorem 3.3, Theorem 3.4, Stampacchia's lemma, see ([3], Thm. 5.8.2), and the fact that the inclusion  $-\bar{p}(t) \in \partial g(\bar{u}(t))$  for a.a.  $t \in (0, T)$  can be rewritten as

$$\bar{u}(t) = \max \left( u_a, \min \left( u_b, -\frac{1}{\nu}\bar{p}(t) \right) \right) \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T)$$

for functions  $g$  of the form (3.16).  $\square$

Note that the last three results indeed provide significantly more regularity for the adjoint state  $\bar{p}$ , the multiplier  $\bar{\mu}$  and the optimal control  $\bar{u}$  than approaches that are based on a “direct” regularization of the problem (P) or a classical Slater condition, cf. [8, 11, 20, 22, 32, 37, 49, 52]. Theorem 3.3 and 3.4 and Corollary 3.5 are thus similar in nature to the results derived for elliptic problems in ([12], Thm. 3.1) and ([13], Thms. 4.2, 4.3). We remark that in particular the  $L^\infty((0, T) \times \Omega)$ -regularity of  $\bar{p}$  and  $\bar{u}$  in Theorem 3.4(i) and Corollary 3.5(iii) turns out to be an important property in practical applications. Compare, e.g., with ([49], Rem. 2.9) in this context, where it is explained that the condition  $\bar{u} \in L^\infty((0, T) \times \Omega)$  allows to establish the strong convergence of a certain regularization scheme, and with the FE-error analysis in Section 4, where the pointwise-a.e. boundedness of  $\bar{p}$  is a crucial ingredient.

Regarding the regularity of the multiplier  $\bar{\mu}$  in (3.12), we would like to clarify at this point that the inclusion  $\bar{\mu} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)) \cap W_0(0, T)^*$  in Theorem 3.3 should be understood in the following sense:  $\bar{\mu}$  is an element of the dual space  $W_0(0, T)^*$  such that there exists a measure  $\bar{\lambda} \in \mathcal{M}([0, T] \times \text{cl}(\Omega))$  with

$$\langle \bar{\mu}, z \rangle_{W_0(0, T)} = \langle \bar{\lambda}, z \rangle_{C([0, T] \times \text{cl}(\Omega))} \quad \forall z \in W_0(0, T) \cap C([0, T] \times \text{cl}(\Omega)). \quad (3.17)$$

It is important to realize here that, for a given  $\bar{\mu}$ , there are always multiple measures  $\bar{\lambda}$  which satisfy the condition (3.17) (due to the homogeneous Dirichlet boundary conditions in  $W_0(0, T)$ ). To overcome this issue, one can impose, e.g., additional conditions on  $\bar{\lambda}$ , see Corollary 3.6 below.

A major drawback of the stationarity system established in Theorem 3.3 is that it does not contain any complementarity conditions for the multiplier  $\bar{\mu}$  and the function  $\bar{y} - \psi$ . If the regularization term  $g$  has the form (3.16), then the easiest way to recover such conditions is to use the information about the regularity of the optimal controls  $\bar{u}$  of (P) in Corollary 3.5 to restrict the control space of (P) and to subsequently apply the classical Slater-based KKT-theory. (Note that the solution map  $S: u \mapsto y$  of the PDE (2.3) maps the space  $L^\infty(0, T; L^q(\Omega))$  into the continuous functions for all  $q > d/2$  and all sufficiently regular  $y_0$  so that it is indeed possible to satisfy a classical Slater condition in the restricted setting if the spatial dimension is small enough, see ([23], Thm. 3.1).) If the functions  $j$ ,  $g$  and  $\phi$  are sufficiently smooth, then it is also possible to establish the complementarity  $0 \leq \bar{\mu} \perp \bar{y} - \psi \geq 0$  by a direct computation as the following result shows. This approach has the advantage that it also works in those cases where, e.g.,  $y_0$  and  $\psi$  are identically zero and where the Dirichlet boundary conditions in (2.3) make the existence of a Slater point in  $C([0, T] \times \text{cl}(\Omega))$  impossible regardless of the continuity properties of the states  $y$ .

**Corollary 3.6** (Uniqueness of dual quantities and complementarity conditions). *Suppose that (X) holds, let  $\bar{u} \in L^2(0, T; L^2(\Omega))$  be a local solution of problem (P), and define  $\bar{y} := S(\bar{u})$ . Assume (in addition to the conditions in Assumption 2.1) that  $g$  is real-valued and bounded on bounded sets, that the functions  $j(t, \cdot)$ ,  $g(\cdot)$  and  $\phi(\cdot)$  are Gâteaux differentiable for all  $t \in [0, T]$  with derivatives  $j'(t, \cdot)$ ,  $g'(\cdot)$  and  $\phi'(\cdot)$ , that the derivative  $g'$  is coercive (in the sense of (3.21)), and that  $\partial_c j(t, z) = \{j'(t, z)\}$  and  $\partial_c \phi(z) = \{\phi'(z)\}$  holds for all  $t \in [0, T]$  and all  $z \in L^2(\Omega)$ . Then  $\bar{u}$  is an element of  $L^\infty(0, T; L^2(\Omega))$ , the functions  $\bar{p}$ ,  $\bar{\eta}_j$ ,  $\bar{\eta}_g$  and  $\bar{\eta}_\phi$  in the necessary optimality condition of Theorem 3.3 are unique,  $\bar{\eta}_j$ ,  $\bar{\eta}_g$  and  $\bar{\eta}_\phi$  satisfy*

$$\bar{\eta}_j(t) = j'(t, \bar{y}(t)), \quad \bar{\eta}_g(t) = g'(\bar{u}(t)) \text{ for a.a. } t \in (0, T) \quad \text{and} \quad \bar{\eta}_\phi = \phi'(\bar{y}(T)), \quad (3.18)$$

*the multiplier  $\bar{\mu}$  in Theorem 3.3 is unique and non-negative as an element of the dual space  $W_0(0, T)^*$ , and there exists a unique, non-negative  $\bar{\lambda} \in \mathcal{M}([0, T] \times \text{cl}(\Omega))$  with*

$$\begin{aligned} \bar{\lambda}|_{[0, T] \times \partial\Omega \cup \{0\} \times \Omega} &= 0, \\ \langle \bar{\mu}, z \rangle_{W_0(0, T)} &= \langle \bar{\lambda}, z \rangle_{C([0, T] \times \text{cl}(\Omega))} \quad \forall z \in W_0(0, T) \cap C([0, T] \times \text{cl}(\Omega)). \end{aligned} \quad (3.19)$$

*Moreover, in this situation, we have the following:*

(i) If  $d \in \{1, 2, 3\}$  and  $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$  holds, then  $\bar{y} - \psi$  is an element of  $C([0, T] \times \text{cl}(\Omega))$  and the function  $\bar{y} - \psi$  and the measure  $\bar{\lambda}$  defined by (3.19) are complementary to each other in the sense that

$$\begin{aligned} \bar{y} - \psi &\geq 0 \text{ in } [0, T] \times \text{cl}(\Omega), & \bar{\lambda} &\geq 0 \text{ in } \mathcal{M}([0, T] \times \text{cl}(\Omega)), \\ \langle \bar{\lambda}, \bar{y} - \psi \rangle_{C([0, T] \times \text{cl}(\Omega))} &= 0. \end{aligned}$$

(ii) If the dimension  $d$  is arbitrary and  $y_0 = \psi \in H_0^1(\Omega) \cap H^2(\Omega)$  holds, then  $\bar{y} - \psi$  is an element of the space  $W_0(0, T)$  defined in (1.1) and the function  $\bar{y} - \psi$  and the multiplier  $\bar{\mu}$  are complementary to each other in the sense that

$$\begin{aligned} \bar{y} - \psi &\geq 0 \text{ a.e. in } (0, T) \times \Omega, & \bar{\mu} &\geq 0 \text{ in } W_0(0, T)^*, \\ \langle \bar{\mu}, \bar{y} - \psi \rangle_{W_0(0, T)} &= 0. \end{aligned} \tag{3.20}$$

*Proof.* The identities in (3.18) follow straightforwardly from the assumptions on  $j$ ,  $g$  and  $\phi$  and the results of Theorem 3.3. In particular,  $\bar{\eta}_j$ ,  $\bar{\eta}_g$  and  $\bar{\eta}_\phi$  are uniquely determined, and we obtain from (3.11) and (3.12) that the adjoint state  $\bar{p}$  is unique and that the multiplier  $\bar{\mu}$  is unique as an element of  $W_0(0, T)^*$ . To prove that  $\bar{u}$  is an element of  $L^\infty(0, T; L^2(\Omega))$ , we note that (3.10) and (3.11) imply that there exists a constant  $C > 0$  with

$$\frac{(g'(\bar{u}(t)), \bar{u}(t))_{L^2(\Omega)}}{\|\bar{u}(t)\|_{L^2(\Omega)}} = \frac{(-\bar{p}(t), \bar{u}(t))_{L^2(\Omega)}}{\|\bar{u}(t)\|_{L^2(\Omega)}} \leq C \quad \text{for a.a. } t \in (0, T).$$

The  $L^\infty(0, T; L^2(\Omega))$ -regularity of  $\bar{u}$  now follows easily by contradiction from the coercivity of  $g'$ , *i.e.*, the convergence

$$\lim_{\|z\|_{L^2(\Omega)} \rightarrow \infty} \frac{(g'(z), z)_{L^2(\Omega)}}{\|z\|_{L^2(\Omega)}} = \infty. \tag{3.21}$$

Suppose now that an arbitrary but fixed  $z \in C_c^\infty((0, T) \times \Omega)$  is given such that  $\bar{y} + sz \geq \psi$  holds a.e. in  $\Omega$  for a.a.  $t \in (0, T)$  for all sufficiently small  $s > 0$ . Then it follows from the  $L^\infty(0, T; L^2(\Omega))$ -regularity of  $\bar{u}$ , the boundedness of  $g$  on bounded subsets of  $L^2(\Omega)$ , the properties of  $j$  and  $\bar{p}$ , Proposition 2.3, the regularity of  $z$ , and the local optimality of  $\bar{u}$  that  $g'(\bar{u}) \in L^\infty(0, T; L^2(\Omega))$  and

$$g(\bar{u}), g(\bar{u} + s(\partial_t z + Az)), j(\cdot, \bar{y} + sz), j(\cdot, \bar{y}) \in L^\infty(0, T)$$

holds for all  $s > 0$  and that

$$\int_0^T \left( \frac{j(\cdot, \bar{y} + sz) - j(\cdot, \bar{y})}{s} + \frac{g(\bar{u} + s(\partial_t z + Az)) - g(\bar{u})}{s} \right) dt + \frac{\phi((\bar{y} + sz)(T)) - \phi(\bar{y}(T))}{s} \geq 0 \tag{3.22}$$

holds for all sufficiently small  $s > 0$ . Note that the convexity of  $g$  and the Lipschitz continuity property of  $j$  in Assumption 2.1 imply

$$\begin{aligned} 0 &\leq \frac{g(\bar{u} + s(\partial_t z + Az)) - g(\bar{u})}{s} - (g'(\bar{u}), \partial_t z + Az)_{L^2(\Omega)} \\ &\leq g(\bar{u} + \partial_t z + Az) - g(\bar{u}) - (g'(\bar{u}), \partial_t z + Az)_{L^2(\Omega)} \in L^\infty(0, T) \end{aligned}$$

and

$$\left| \frac{j(\cdot, \bar{y} + sz) - j(\cdot, \bar{y})}{s} \right| \leq C \|z\|_{L^\infty(0, T; L^2(\Omega))} \quad \text{a.e. in } (0, T)$$

for all  $s \in (0, 1]$  with some constant  $C > 0$  independent of  $s$ . This allows us to use the dominated convergence theorem to pass to the limit  $s \searrow 0$  in (3.22) and to obtain

$$\int_0^T \left( (j'(\cdot, \bar{y}), z)_{L^2(\Omega)} + (g'(\bar{u}), \partial_t z + Az)_{L^2(\Omega)} \right) dt + (\phi'(\bar{y}(T)), z(T))_{L^2(\Omega)} \geq 0.$$

Due to (3.12) and the properties of  $\bar{p}$  and  $z$ , the last inequality can also be written as

$$\langle \bar{\mu}, z \rangle_{W_0(0, T)} \geq 0. \quad (3.23)$$

Since  $z \in C_c^\infty((0, T] \times \Omega)$  was an arbitrary function with  $\bar{y} + sz \geq \psi$  for some  $s > 0$ , we may now deduce that

$$\langle \bar{\mu}, z \rangle_{W_0(0, T)} \geq 0 \quad \forall 0 \leq z \in C_c^\infty((0, T] \times \Omega)$$

holds and, by mollification, that  $\bar{\mu}$  is a non-negative element of  $W_0(0, T)^*$ . Consider now two arbitrary but fixed  $\lambda_1, \lambda_2 \in \mathcal{M}([0, T] \times \text{cl}(\Omega))$  that satisfy

$$\langle \bar{\mu}, z \rangle_{W_0(0, T)} = \langle \lambda_1, z \rangle_{C([0, T] \times \text{cl}(\Omega))} = \langle \lambda_2, z \rangle_{C([0, T] \times \text{cl}(\Omega))}$$

for all  $z \in W_0(0, T) \cap C([0, T] \times \text{cl}(\Omega))$ , cf. (3.17). Then it follows straightforwardly from the non-negativity of  $\bar{\mu}$  as an element of  $W_0(0, T)^*$  and mollification that the restrictions of  $\lambda_1$  and  $\lambda_2$  to the set  $(0, T] \times \Omega$  are identical, non-negative Borel measures. By defining  $\bar{\lambda} := \lambda_1|_{(0, T] \times \Omega} \in \mathcal{M}([0, T] \times \text{cl}(\Omega))$ , we now obtain a representation of  $\bar{\mu}$  as a measure which has all of the desired properties and is trivially unique. This proves the first part of the corollary. It remains to establish (i) and (ii). To this end, let us first assume that  $d \in \{1, 2, 3\}$  and  $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$  holds. Then, it follows from ([23], Thm. 3.1) and the regularity  $\bar{u} \in L^\infty(0, T; L^2(\Omega))$  that the function  $\bar{y} - \psi$  is continuous and that the contact set  $\mathcal{C} := \{(t, x) \in [0, T] \times \text{cl}(\Omega) \mid (\bar{y} - \psi)(t, x) = 0\}$  is well-defined in the classical sense. The latter allows us to use test functions  $z \in C_c^\infty((0, T] \times \Omega)$  with support away from  $\mathcal{C}$  in (3.23) to deduce that the support of the measure  $\bar{\lambda}$  defined by (3.19) is contained in  $\mathcal{C}$ . The complementarity in (i) now follows immediately. If, on the other hand, we know that  $y_0 = \psi \in H_0^1(\Omega) \cap H^2(\Omega)$  holds, then  $\bar{y} - \psi$  is clearly an element of  $W_0(0, T)$  and we obtain from the properties of  $\bar{y}$ ,  $\bar{u}$ , and  $\psi$  that  $\partial_t(\bar{y} - \psi) + A(\bar{y} - \psi) = \bar{u} - A\psi \in L^\infty(0, T; L^2(\Omega))$ . Using this regularity and exactly the same arguments as for the functions  $z \in C_c^\infty((0, T] \times \Omega)$  in the first part of the proof as well as the trivial estimates  $\bar{y} + (\psi - \bar{y}) \geq \psi$  and  $\bar{y} + (\bar{y} - \psi) \geq \bar{y}$ , it follows that  $\pm \langle \bar{\mu}, \bar{y} - \psi \rangle_{W_0(0, T)} \geq 0$ . This proves the third line in (3.20). Since the inequalities in (3.20) are trivial due to the feasibility of  $\bar{u}$  and the non-negativity of  $\bar{\mu}$  in  $W_0(0, T)^*$ , this yields the assertion in (ii) and completes the proof.  $\square$

Note that it makes sense that we need additional regularity assumptions on  $\bar{y} - \psi$  in Corollary 3.6 to obtain complementarity conditions. Without these assumptions, it is not even clear if a dual pairing between  $\bar{y} - \psi$  and  $\bar{\mu}$  can be defined.

As usual, in the presence of convexity, the necessary optimality conditions in Corollary 3.6 are also sufficient. To be more precise, we have the following:

**Corollary 3.7** (Necessary and sufficient condition in low dimensions). *Suppose that  $d \in \{1, 2, 3\}$  and  $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$  holds, and that (X) is satisfied. Assume (in addition to the conditions in Assumption 2.1) that  $g$  is real-valued and bounded on bounded sets, that the functions  $j(t, \cdot)$ ,  $g(\cdot)$  and  $\phi(\cdot)$  are Gâteaux differentiable*



for all  $t \in [0, T]$  with derivatives  $j'(t, \cdot)$ ,  $g'(\cdot)$  and  $\phi'(\cdot)$ , that the maps  $j(t, \cdot)$  and  $\phi(\cdot)$  are convex for all  $t \in [0, T]$ , and that the derivative  $g'$  is coercive. Then a control  $\bar{u}$  with associated state  $\bar{y} := S(\bar{u})$  is a local/global solution of (P) if and only if there exist an adjoint state  $\bar{p}$  and a measure  $\bar{\lambda}$  such that the following holds true:

$$\begin{aligned}
& \bar{u} \in L^\infty(0, T; L^2(\Omega)), \quad \bar{\lambda} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)), \\
& \bar{y} \in C([0, T] \times \text{cl}(\Omega)) \cap L^q(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap W^{1,q}(0, T; L^2(\Omega)) \quad \forall 1 < q < \infty, \\
& \bar{p} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*), \\
& \bar{p}(t) + g'(\bar{u}(t)) = 0 \text{ for a.a. } t \in (0, T), \\
& -\partial_t \bar{p} + A\bar{p} = j'(\cdot, \bar{y}) - \bar{\lambda} \text{ in } (0, T) \times \Omega, \\
& \text{tr}(\bar{p}) = 0 \text{ on } (0, T) \times \partial\Omega, \quad \bar{p}(T) = \phi'(\bar{y}(T)) \text{ in } \Omega, \\
& \bar{y} - \psi \geq 0 \text{ in } [0, T] \times \text{cl}(\Omega), \quad \bar{\lambda} \geq 0 \text{ in } \mathcal{M}([0, T] \times \text{cl}(\Omega)), \\
& \langle \bar{\lambda}, \bar{y} - \psi \rangle_{C([0, T] \times \text{cl}(\Omega))} = 0.
\end{aligned} \tag{3.24}$$

Here,  $Y_\gamma$ ,  $\gamma > d/2$ , is again defined as in Section 1.1 and Theorem 3.3.

*Proof.* Recall that, for locally Lipschitz continuous, Gâteaux differentiable and convex functions, the Clarke and the convex subdifferential are always identical and only contain the Gâteaux derivative, see ([19], Prop. 2.2.7) and ([24], Prop. I-5.3). This shows that the conditions on  $j$ ,  $g$  and  $\phi$  in Corollary 3.6 are satisfied in the situation of Corollary 3.7. From ([25], Thm. 5.3) and again ([33], Thm. 3.2.1.2), ([30], Thm. 9.15) and ([23], Thm. 3.1), we obtain further that our assumptions  $d \in \{1, 2, 3\}$  and  $y_0, \psi \in H^2(\Omega)$  imply  $S(u) \in L^q(0, T; H^2(\Omega)) \cap W^{1,q}(0, T; L^2(\Omega))$  and  $S(u), S(u) - \psi \in C([0, T] \times \text{cl}(\Omega))$  for all  $u \in L^\infty(0, T; L^2(\Omega))$  and all  $1 < q < \infty$ . By invoking Theorem 3.3 and Corollary 3.6, it now follows immediately that every local solution  $\bar{u}$  of (P) satisfies (3.24). It remains to prove that every  $\bar{u}$  which satisfies (3.24) is a global optimum of (P). To this end, let us assume that we are given a control  $\bar{u}$  with state  $\bar{y}$  such that (3.24) holds with some  $\bar{p}$  and  $\bar{\lambda}$ . Then we know that  $\bar{u}$  is admissible for (P) and it follows from Proposition 2.4 that (P) possesses a global solution  $\tilde{u}$ . According to the first part of the proof, this solution  $\tilde{u}$  has to satisfy  $\tilde{u} \in L^\infty(0, T; L^2(\Omega))$  and  $\tilde{y} := S(\tilde{u}) \in C([0, T] \times \text{cl}(\Omega)) \cap L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))$ . By exploiting the convexity of the functions  $j(t, \cdot)$ ,  $g(\cdot)$  and  $\phi(\cdot)$  and the properties of  $\tilde{y}$ ,  $\bar{y}$ ,  $\bar{p}$  and  $\bar{\lambda}$ , we may now deduce that

$$\begin{aligned}
0 & \leq \langle \bar{\lambda}, \tilde{y} - \psi \rangle_{C([0, T] \times \text{cl}(\Omega))} \\
& = \langle \bar{\lambda}, \tilde{y} - \bar{y} \rangle_{C([0, T] \times \text{cl}(\Omega))} \\
& = \int_0^T \left( \langle -\partial_t(\tilde{y} - \bar{y}) - A(\tilde{y} - \bar{y}), \bar{p} \rangle_{H_0^1(\Omega)} + (j'(\cdot, \bar{y}), \tilde{y} - \bar{y})_{L^2(\Omega)} \right) dt + ((\tilde{y} - \bar{y})(T), \phi'(\bar{y}(T)))_{L^2(\Omega)} \\
& = \int_0^T \left( (g'(\tilde{u}), \tilde{u} - \bar{u})_{L^2(\Omega)} + (j'(\cdot, \bar{y}), \tilde{y} - \bar{y})_{L^2(\Omega)} \right) dt + (\phi'(\bar{y}(T)), (\tilde{y} - \bar{y})(T))_{L^2(\Omega)} \\
& \leq J(\tilde{y}, \tilde{u}) - J(\bar{y}, \bar{u}).
\end{aligned}$$

Thus,  $J(\bar{y}, \bar{u}) \leq J(\tilde{y}, \tilde{u})$  and  $\bar{u}$  is indeed a global solution of (P). This establishes the reverse implication and completes the proof.  $\square$

**Corollary 3.8** (Necessary and sufficient condition in arbitrary dimensions). *Suppose that  $\psi \equiv y_0$  and  $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$  holds, and that (X) is satisfied. Assume that  $j$ ,  $g$  and  $\phi$  satisfy the conditions in corollary 3.7. Then a control  $\bar{u}$  with associated state  $\bar{y} := S(\bar{u})$  is a local/global solution of (P) if and only if there exist an*

adjoint state  $\bar{p}$  and a multiplier  $\bar{\mu}$  such that the following holds true:

$$\begin{aligned}
\bar{u} &\in L^\infty(0, T; L^2(\Omega)), & \bar{\mu} &\in \mathcal{M}([0, T] \times \text{cl}(\Omega)) \cap W_0(0, T)^*, \\
\bar{y} &\in L^q(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap W^{1,q}(0, T; L^2(\Omega)) \quad \forall 1 < q < \infty, \\
\bar{p} &\in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*), \\
&\bar{p}(t) + g'(\bar{u}(t)) = 0 \text{ for a.a. } t \in (0, T), \\
&-\partial_t \bar{p} + A\bar{p} = j'(\cdot, \bar{y}) - \bar{\mu} \text{ in } (0, T) \times \Omega, \\
\text{tr}(\bar{p}) &= 0 \text{ on } (0, T) \times \partial\Omega, & \bar{p}(T) &= \phi'(\bar{y}(T)) \text{ in } \Omega, \\
\bar{y} - \psi &\geq 0 \text{ a.e. in } (0, T) \times \Omega, & \bar{\mu} &\geq 0 \text{ in } W_0(0, T)^*, \\
&\langle \bar{\mu}, \bar{y} - \psi \rangle_{W_0(0, T)} = 0.
\end{aligned} \tag{3.25}$$

Here,  $Y_\gamma$ ,  $\gamma > d/2$ , and  $W_0(0, T)$  are defined as in Section 1.1 and Theorem 3.3.

*Proof.* From the assumption  $\psi \equiv y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$  and Proposition 2.3, it follows that  $S(u) - \psi \in W_0(0, T)$  holds for all  $u \in L^2(0, T; L^2(\Omega))$ . By invoking Theorem 3.3, Corollary 3.6 and again ([25], Thm. 5.3), we now obtain that every local solution  $\bar{u}$  of (P) satisfies (3.25). To prove the reverse implication, we can use exactly the same arguments as in the proof of Corollary 3.7. Indeed, due to the convexity of the involved functions, we may compute that, for every  $\bar{u}$  which satisfies (3.25) with  $\bar{y} := S(\bar{u})$  and some  $\bar{p}$ ,  $\bar{\mu}$ , we have

$$\begin{aligned}
0 &\leq \langle \bar{\mu}, y - \psi \rangle_{W_0(0, T)} = \langle \bar{\mu}, y - \bar{y} \rangle_{W_0(0, T)} \\
&= \int_0^T \left( (g'(\bar{u}), u - \bar{u})_{L^2(\Omega)} + (j'(\cdot, \bar{y}), y - \bar{y})_{L^2(\Omega)} \right) dt + (\phi'(\bar{y}(T)), (y - \bar{y})(T))_{L^2(\Omega)} \\
&\leq J(y, u) - J(\bar{y}, \bar{u})
\end{aligned}$$

for all admissible  $u \in L^2(0, T; L^2(\Omega))$  with state  $y := S(u)$ . This proves the claim.  $\square$

We conclude this section with three tangible, self-contained examples that illustrate what the previous results mean in practice. We begin with a simple tracking-type optimal control problem governed by the heat equation in dimension  $d \in \{1, 2, 3\}$ :

**Example 3.9** (Heat equation without control constraints in low dimensions). Suppose that a bounded, convex (or  $C^{1,1}$ -) domain  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{1, 2, 3\}$ , a  $\nu > 0$ , a  $T > 0$ , a subharmonic  $\psi \in H^2(\Omega)$  with  $\text{tr}(\psi) \leq 0$  on  $\partial\Omega$ , a  $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$  with  $y_0 \geq \psi$  a.e. in  $\Omega$  and a  $y_D \in L^\infty(0, T; L^2(\Omega))$  are given, and consider the problem

$$\begin{aligned}
\text{Minimize} & \quad \frac{1}{2} \int_0^T \|y - y_D\|_{L^2(\Omega)}^2 dt + \frac{\nu}{2} \int_0^T \|u\|_{L^2(\Omega)}^2 dt \\
\text{w.r.t.} & \quad u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\
\text{s.t.} & \quad \partial_t y - \Delta y = u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
& \quad \text{tr}(y) = 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\
& \quad y(0) = y_0 \quad \text{a.e. in } \Omega, \\
\text{and} & \quad y \geq \psi \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T).
\end{aligned} \tag{3.26}$$

Then the stationary control  $u = -\Delta y_0$  is admissible for (3.26), and it follows from the strict convexity of (3.26), Proposition 2.4, Theorem 3.4 and Corollaries 3.7 and 3.5 that (3.26) admits one and only one solution  $\bar{u}$  with

state  $\bar{y} := S(\bar{u})$  which is uniquely characterized by the KKT-system

$$\begin{aligned}
& \bar{u}, \bar{p} \in L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*), \\
& \bar{y} \in C([0, T] \times \text{cl}(\Omega)) \cap L^q(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap W^{1,q}(0, T; L^2(\Omega)) \quad \forall 1 < q < \infty, \\
& \bar{\lambda} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)), \\
& \bar{p}(t) + \nu \bar{u}(t) = 0 \text{ for a.a. } t \in (0, T), \\
& -\partial_t \bar{p} - \Delta \bar{p} = \bar{y} - y_D - \bar{\lambda} \text{ in } (0, T) \times \Omega, \\
& \text{tr}(\bar{p}) = 0 \text{ on } (0, T) \times \partial\Omega, \quad \bar{p}(T) = 0 \text{ in } \Omega, \\
& \bar{p}(t) \geq 0 \text{ } H_0^1\text{-q.e. in } \{\bar{y}(t) = \psi\} \text{ for a.a. } t \in (0, T), \\
& \bar{y} - \psi \geq 0 \text{ in } [0, T] \times \text{cl}(\Omega), \quad \bar{\lambda} \geq 0 \text{ in } \mathcal{M}([0, T] \times \text{cl}(\Omega)), \\
& \langle \bar{\lambda}, \bar{y} - \psi \rangle_{C([0, T] \times \text{cl}(\Omega))} = 0.
\end{aligned} \tag{3.27}$$

Here,  $Y_\gamma$ ,  $\gamma > d/2$ , is again defined as in Section 1.1 and Theorem 3.3, and “q.e.” is short for quasi-everywhere, see ([9], Sect. 6.4.3).

Note that Example 3.9 also covers, *e.g.*, the degenerate case  $y_0 \equiv \psi \equiv 0$ , where Slater points cannot exist in  $C([0, T] \times \text{cl}(\Omega))$  due to the Dirichlet boundary conditions in (3.26). For arbitrary  $d \geq 1$ , we obtain:

**Example 3.10** (Heat equation without control constraints in all dimensions). Suppose that a bounded, convex (or  $C^{1,1}$ -) domain  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , a  $T > 0$ , a subharmonic  $\psi \in H_0^1(\Omega) \cap H^2(\Omega)$ , a  $y_T \in L^2(\Omega)$ , a  $\nu > 0$ , and a  $y_D \in L^\infty(0, T; L^2(\Omega))$  are given, and consider the problem

$$\begin{aligned}
& \text{Minimize} \quad \frac{1}{2} \int_0^T \|y - y_D\|_{L^2(\Omega)}^2 dt + \frac{\nu}{2} \int_0^T \|u\|_{L^2(\Omega)}^2 dt + \frac{1}{2} \|y(T) - y_T\|_{L^2(\Omega)}^2 \\
& \text{w.r.t.} \quad u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\
& \text{s.t.} \quad \partial_t y - \Delta y = u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
& \quad \text{tr}(y) = 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\
& \quad y(0) = \psi \quad \text{a.e. in } \Omega, \\
& \text{and} \quad y \geq \psi \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T).
\end{aligned} \tag{3.28}$$

Then the stationary control  $u = -\Delta \psi$  is admissible for (3.28), and it follows from the strict convexity of (3.28), Proposition 2.4, Corollary 3.5 and Corollary 3.8 that (3.28) admits one and only one solution  $\bar{u}$  with state  $\bar{y} := S(\bar{u})$  which is uniquely characterized by the KKT-system

$$\begin{aligned}
& \bar{u}, \bar{p} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*), \\
& \bar{y} \in L^q(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap W^{1,q}(0, T; L^2(\Omega)) \quad \forall 1 < q < \infty, \\
& \bar{\mu} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)) \cap W_0(0, T)^*, \\
& \bar{p}(t) + \nu \bar{u}(t) = 0 \text{ for a.a. } t \in (0, T), \\
& -\partial_t \bar{p} - \Delta \bar{p} = \bar{y} - y_D - \bar{\mu} \text{ in } (0, T) \times \Omega, \\
& \text{tr}(\bar{p}) = 0 \text{ on } (0, T) \times \partial\Omega, \quad \bar{p}(T) = \bar{y}(T) - y_T \text{ in } \Omega, \\
& \bar{y} - \psi \geq 0 \text{ a.e. in } (0, T) \times \Omega, \quad \bar{\mu} \geq 0 \text{ in } W_0(0, T)^*, \\
& \langle \bar{\mu}, \bar{y} - \psi \rangle_{W_0(0, T)} = 0.
\end{aligned}$$

Here,  $Y_\gamma$ ,  $\gamma > d/2$ , and  $W_0(0, T)$  are defined as in Section 1.1 and Theorem 3.3.

As a final example, we consider an optimal control problem that not only contains a pointwise state constraint but also a non-smooth term and a control constraint:

**Example 3.11** (An  $L^1$ -tracking-type optimal control problem with a lower control constraint). Suppose that a bounded, convex (or  $C^{1,1}$ -) domain  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{1, 2, 3\}$ , a  $T > 0$ , a subharmonic  $\psi \in H^2(\Omega)$  with  $\text{tr}(\psi) \leq 0$  on  $\partial\Omega$ , a  $u_a \in H^1(\Omega)$  with  $u_a \leq 0$  a.e. in  $\Omega$ , a  $\nu > 0$ , a  $y_T \in L^1(\Omega)$ , and a  $y_D \in L^\infty(0, T; L^1(\Omega))$  are given, and consider the optimal control problem

$$\begin{aligned}
& \text{Minimize} && \int_0^T \|y - y_D\|_{L^1(\Omega)} dt + \frac{\nu}{2} \int_0^T \|u\|_{L^2(\Omega)}^2 dt + \int_\Omega \sqrt{1 + (y(T) - y_T)^2} dx \\
& \text{w.r.t.} && u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\
& \text{s.t.} && \partial_t y - \Delta y = u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
& && \text{tr}(y) = 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\
& && y(0) = 0 \quad \text{a.e. in } \Omega, \\
& \text{and} && y \geq \psi \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
& \text{and} && u \geq u_a \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T).
\end{aligned} \tag{3.29}$$

Then the control  $u = 0$  is admissible for (3.29) (since the maximum principle for subharmonic functions implies  $\psi \leq 0$  a.e. in  $\Omega$ ), and it follows from the strict convexity of (3.29), Proposition 2.4, Theorems 3.3 and 3.4, Corollary 3.5, ([19], Prop. 2.2.7), ([23], Thm. 3.1), ([25], Thm. 5.3), ([33], Thm. 3.2.1.2), and ([30], Thm. 9.15) that (3.29) is uniquely solvable, that the unique solution  $\bar{u}$  of (3.29) and its state  $\bar{y} := S(\bar{u})$  satisfy

$$\begin{aligned}
& \bar{u} \in L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega)), \\
& \bar{y} \in C([0, T] \times \text{cl}(\Omega)) \cap L^q(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap W^{1,q}(0, T; L^2(\Omega)) \quad \forall 1 < q < \infty,
\end{aligned}$$

and that there exist

$$\begin{aligned}
& \bar{p} \in L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*), \\
& \bar{\mu} \in \mathcal{M}([0, T] \times \text{cl}(\Omega)) \cap W_0(0, T)^*, \quad \bar{\eta} \in L^\infty(0, T; L^2(\Omega)),
\end{aligned}$$

such that  $\bar{u}$ ,  $\bar{y}$ ,  $\bar{p}$ ,  $\bar{\mu}$  and  $\bar{\eta}$  satisfy the system

$$\begin{aligned}
& \bar{u}(t) = \max \left( u_a, -\frac{1}{\nu} \bar{p}(t) \right) \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
& \bar{\eta}(t) \in \partial \| \cdot \|_{L^1(\Omega)}(\bar{y}(t) - y_D(t)) \quad \text{for a.a. } t \in (0, T), \\
& -\partial_t \bar{p} - \Delta \bar{p} = \bar{\eta} - \bar{\mu} \quad \text{in } (0, T) \times \Omega, \\
& \text{tr}(\bar{p}) = 0 \quad \text{on } (0, T) \times \partial\Omega, \quad \bar{p}(T) = \frac{\bar{y}(T) - y_T}{\sqrt{1 + (\bar{y}(T) - y_T)^2}} \quad \text{in } \Omega.
\end{aligned}$$

Here,  $Y_\gamma$ ,  $\gamma > d/2$ , and  $W_0(0, T)$  are defined as in Section 1.1 and Theorem 3.3.

#### 4. APPLICATION TO FINITE ELEMENT ERROR ESTIMATES

In what follows, we demonstrate that the regularity results established in Section 3 allow to derive new *a priori* error estimates for the finite element approximation of parabolic, distributed optimal control problems with pointwise state constraints. To avoid obscuring the basic ideas of our analysis with technicalities, we do

not aim for the highest possible generality in this section but focus on the model problem in Example 3.9, *i.e.*, the  $L^2$ -tracking-type optimal control problem for the heat equation

$$\begin{aligned}
& \text{Minimize} && \frac{1}{2} \int_0^T \|y - y_D\|_{L^2(\Omega)}^2 dt + \frac{\nu}{2} \int_0^T \|u\|_{L^2(\Omega)}^2 dt \\
& \text{w.r.t.} && u \in L^2(0, T; L^2(\Omega)), \quad y \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\
& \text{s.t.} && \partial_t y - \Delta y = u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
& && \text{tr}(y) = 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\
& && y(0) = y_0 \quad \text{a.e. in } \Omega, \\
& \text{and} && y \geq \psi \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T).
\end{aligned} \tag{4.1}$$

For comments on possible generalizations of the subsequent results, see Remark 4.10. Our standing assumptions on the quantities in (4.1) are as follows:

**Assumption 4.1** (Standing assumptions for the study of problem (4.1)).

- $\Omega \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$ , is a bounded, convex, polygonal/polyhedral domain.
- $\nu > 0$ ,  $T > 0$ ,  $y_D \in L^\infty(0, T; L^2(\Omega))$  and  $y_0 \in H_0^1(\Omega) \cap W^{2, \infty}(\Omega)$  are given.
- $\psi \in W^{2, \infty}(\Omega)$  is a subharmonic function and the continuous representatives of  $\psi$  and  $y_0$  satisfy  $\psi < 0$  and  $\psi < y_0$  everywhere in  $\text{cl}(\Omega)$ .

Note that a classical Slater condition in  $L^\infty((0, T) \times \Omega)$  cannot be satisfied in the above situation since the solution map  $S: u \mapsto y$  of the governing PDE only maps into the space  $L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)) \not\subset L^\infty((0, T) \times \Omega)$ , *cf.* the discussion in [49, 50]. To calculate an approximation of the (necessarily unique) solution  $\bar{u}$  of (4.1), we use a standard dG(0)-cG(1)-scheme analogous to that in [7, 22, 32, 41, 42] (*i.e.*, FE-functions that are piecewise constant in time and piecewise affine and continuous in space). We collect our standing assumptions on the involved meshes etc. in:

**Assumption 4.2** (Assumptions and notation for the discretization of (4.1)).

- We are given a family of partitions  $0 =: t_0 < t_1 < \dots < t_M := T$  of the interval  $(0, T]$  into cells  $I_m := (t_{m-1}, t_m]$  of length  $k_m := t_m - t_{m-1}$ ,  $m = 1, \dots, M$ , with maximum width  $k := \max_m k_m \leq \min(1/2, T/4)$  such that there exist constants  $c_1, c_2, c_3 > 0$  with

$$c_1^{-1} \leq \frac{k_m}{k_{m+1}} \leq c_1 \quad \forall m = 1, \dots, M-1 \quad \text{and} \quad k_m \geq c_2 k^{c_3} \quad \forall m = 1, \dots, M.$$

- We are given a quasi-uniform family of triangulations  $\mathcal{T}_h$  of the domain  $\Omega$  with mesh size  $h < 1/2$ . (See, *e.g.*, [7, 10] for the precise definitions of these terms.)
- The temporal and spatial widths  $k$  and  $h$  satisfy a compatibility condition of the form  $k \leq Ch^\sigma$  with some constants  $C, \sigma > 0$ .
- $V_h$  is the finite element space defined by

$$V_h := \{v \in C(\text{cl}(\Omega)) \mid v|_\tau \text{ is affine for all cells } \tau \in \mathcal{T}_h \text{ and } v|_{\partial\Omega} = 0\}.$$

- $V_{kh}$  is the space-time finite element space defined by

$$V_{kh} := \{v \in L^2(0, T; V_h) \mid v^m := v|_{I_m} = \text{const} \in V_h \text{ for all } m = 1, \dots, M\}.$$

- $x_l$ ,  $l = 1, \dots, \dim(V_h)$ , are the interior nodes of the triangulation  $\mathcal{T}_h$ .

- $B: V_{kh} \times V_{kh} \rightarrow \mathbb{R}$  is the bilinear form defined by

$$B(v_{kh}, w_{kh}) := \sum_{m=2}^M (v_{kh}^m - v_{kh}^{m-1}, w_{kh}^m)_{L^2(\Omega)} + \sum_{m=1}^M k_m (\nabla v_{kh}^m, \nabla w_{kh}^m)_{L^2(\Omega)^d} + (v_{kh}^1, w_{kh}^1)_{L^2(\Omega)}.$$

Note that we do not require additional  $\partial_t$ -terms in  $B$  here since we consider this bilinear form only on the space  $V_{kh}$ , cf. [43]. As a discrete counterpart of the PDE

$$\begin{aligned} \partial_t y - \Delta y &= u \quad \text{a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ \text{tr}(y) &= 0 \quad \text{a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \\ y(0) &= y_0 \quad \text{a.e. in } \Omega \end{aligned} \tag{4.2}$$

in (4.1), we consider the following variational problem:

$$y_{kh} \in V_{kh}, \quad B(y_{kh}, v_{kh}) = (u, v_{kh})_{L^2(0, T; L^2(\Omega))} + (y_0, v_{kh}^1)_{L^2(\Omega)} \quad \forall v_{kh} \in V_{kh}. \tag{4.3}$$

We summarize the (for us relevant) approximation properties of the discretization scheme (4.3) in:

**Lemma 4.3** (FE-error estimates for the discretization of the heat equation (4.2)). *The problem (4.3) admits a unique solution  $S_{kh}(u) := y_{kh} \in V_{kh}$  for every  $u \in L^2(0, T; L^2(\Omega))$  and all  $k$  and  $h$ . Further, the solution maps  $S: u \mapsto y$  and  $S_{kh}: u \mapsto y_{kh}$  of the heat equation (4.2) and its discretization (4.3) satisfy the following error estimates:*

- (i) *There exists a constant  $C > 0$  independent of  $k$ ,  $h$ ,  $y_0$  and  $u$  with*

$$\|S(u) - S_{kh}(u)\|_{L^2(0, T; L^2(\Omega))} \leq C (k + h^2) (\|u\|_{L^2(0, T; L^2(\Omega))} + \|y_0\|_{H^1(\Omega)}) \quad \forall u \in L^2(0, T; L^2(\Omega)).$$

- (ii) *There exist constants  $C, \rho > 0$  independent of  $k$ ,  $h$ ,  $y_0$  and  $u$  with*

$$\|S(u) - S_{kh}(u)\|_{L^\infty((0, T) \times \Omega)} \leq C |\ln(k)| |\ln(h)| (k^\rho + h^\rho) (\|u\|_{L^\infty((0, T) \times \Omega)} + \|y_0\|_{W^{2, \infty}(\Omega)}) \quad \forall u \in L^\infty((0, T) \times \Omega).$$

- (iii) *For every open set  $D \subset \mathbb{R}^d$  that is compactly contained in  $\Omega$ , there exists a constant  $C > 0$  independent of  $k$ ,  $h$ ,  $y_0$  and  $u$  with*

$$\|S(u) - S_{kh}(u)\|_{L^\infty((0, T) \times D)} \leq C |\ln(k)|^2 |\ln(h)|^5 (k + h^2) (\|u\|_{L^\infty((0, T) \times \Omega)} + \|y_0\|_{W^{2, \infty}(\Omega)}) \quad \forall u \in L^\infty((0, T) \times \Omega).$$

*Proof.* The well-definedness of the solution operator  $S_{kh}: u \mapsto y_{kh}$  follows from classical results for elliptic problems after rewriting (4.3) as a time stepping scheme, cf. the proof of ([43], Thm. 3.2), and the error estimates in (i) and (iii) can be found, e.g., in ([7], Lems. 5.33, A.36). Compare also with the regularity results for  $S$  in ([23], Thm. 2.9b)) in this context and with the FE-analysis in [26]. It remains to prove (ii). To this end, we note that ([23], Thm. 3.1) and ([39], Thm. 1) yield that there exist constants  $C, \rho > 0$  independent of  $k$ ,  $h$ ,  $y_0$  and  $u$  such that the solutions of (4.2) and (4.3) with right-hand side  $u \in L^\infty((0, T) \times \Omega)$  satisfy

$$\|S(u)\|_{C^{0, \rho}([0, T] \times \text{cl}(\Omega))} \leq C (\|u\|_{L^\infty((0, T) \times \Omega)} + \|y_0\|_{W^{2, \infty}(\Omega)})$$

and

$$\|S(u) - S_{kh}(u)\|_{L^\infty((0,T)\times\Omega)} \leq C |\ln(k)| |\ln(h)| \inf_{v_{kh} \in V_{kh}} \|S(u) - v_{kh}\|_{L^\infty((0,T)\times\Omega)}.$$

By plugging in a suitable interpolant of the function  $S(u)$  on the right-hand side of the last inequality and by exploiting the  $C^{0,\rho}$ -regularity of  $S(u)$  on  $(0, T) \times \Omega$ , we arrive at the estimate in (ii) as desired. This completes the proof.  $\square$

Using (4.3), we can discretize (4.1) as follows:

$$\begin{aligned} & \text{Minimize} && \frac{1}{2} \int_0^T \|y_{kh} - y_D\|_{L^2(\Omega)}^2 dt + \frac{\nu}{2} \int_0^T \|u\|_{L^2(\Omega)}^2 dt \\ & \text{w.r.t.} && u \in L^2(0, T; L^2(\Omega)), \quad y_{kh} \in V_{kh}, \\ & \text{s.t.} && B(y_{kh}, v_{kh}) = (u, v_{kh})_{L^2(0,T;L^2(\Omega))} + (y_0, v_{kh}^1)_{L^2(\Omega)} \quad \forall v_{kh} \in V_{kh}, \\ & \text{and} && y_{kh}^m(x_l) \geq \psi(x_l) \quad \forall m = 1, \dots, M \quad \forall l = 1, \dots, \dim(V_h). \end{aligned} \quad (4.4)$$

From standard subdifferential calculus rules, we obtain (analogously to [32]):

**Proposition 4.4** (Solvability and optimality conditions for (4.4)). *The problem (4.4) admits a unique solution  $\bar{u}_{kh} \in L^2(0, T; L^2(\Omega))$  for all  $k$  and all  $h$ . This solution is an element of  $V_{kh}$  and there exist multipliers  $\bar{\lambda}^{ml}$ ,  $m = 1, \dots, M$ ,  $l = 1, \dots, \dim(V_h)$ , and an adjoint state  $\bar{p}_{kh}$  such that  $\bar{u}_{kh}$  and its state  $\bar{y}_{kh} := S_{kh}(\bar{u}_{kh})$  satisfy*

$$\begin{aligned} & \bar{u}_{kh}, \bar{p}_{kh}, \bar{y}_{kh} \in V_{kh}, \quad \bar{\lambda}^{ml} \in \mathbb{R}, \quad \bar{p}_{kh} + \nu \bar{u}_{kh} = 0, \\ & B(v_{kh}, \bar{p}_{kh}) = \int_0^T (\bar{y}_{kh} - y_D, v_{kh})_{L^2(\Omega)} dt - \sum_{m=1}^M \sum_{l=1}^{\dim(V_h)} \bar{\lambda}^{ml} v_{kh}^m(x_l) \quad \forall v_{kh} \in V_{kh}, \\ & (\bar{y}_{kh}^m - \psi)(x_l) \geq 0, \quad \bar{\lambda}^{ml} \geq 0, \quad \bar{\lambda}^{ml} (\bar{y}_{kh}^m - \psi)(x_l) = 0 \quad \forall m, l. \end{aligned} \quad (4.5)$$

*Proof.* Since the reduced objective function of (4.4) is strictly convex and coercive, the existence and uniqueness of a solution  $\bar{u}_{kh} \in L^2(0, T; L^2(\Omega))$  are obvious. It remains to show that  $\bar{u}_{kh} \in V_{kh}$  and (4.5) hold. To this end, we note that the control

$$u := \begin{cases} -\frac{1}{k_1} y_0 & \text{a.e. in } I_1 \\ 0 & \text{else} \end{cases}$$

trivially satisfies  $S_{kh}(u) = 0 > \psi$  everywhere in  $\text{cl}(\Omega)$ . The discrete problem (4.4) thus admits a Slater point, and we may use the sum and the chain rule for the convex subdifferential, see ([24], Props. I5.6, I5.7), to deduce that there exists an element  $\bar{\mu}_{kh} \in V_{kh}^*$  of the negative normal cone of the set

$$\{v_{kh} \in V_{kh} \mid v_{kh}^m(x_l) \geq \psi(x_l) \text{ for all } m = 1, \dots, M \text{ and all } l = 1, \dots, \dim(V_h)\} \quad (4.6)$$

at  $\bar{y}_{kh} := S_{kh}(\bar{u}_{kh})$  such that

$$\int_0^T (\bar{y}_{kh} - y_D, S_{kh}^0(z))_{L^2(\Omega)} dt + \nu (\bar{u}_{kh}, z)_{L^2(0,T;L^2(\Omega))} - \langle \bar{\mu}_{kh}, S_{kh}^0(z) \rangle_{V_{kh}} = 0 \quad (4.7)$$

holds for all  $z \in L^2(0, T; L^2(\Omega))$ , where  $S_{kh}^0: L^2(0, T; L^2(\Omega)) \rightarrow V_{kh}$  denotes the solution map of the discretized PDE (4.3) with initial value zero. If we express this multiplier  $\bar{\mu}_{kh}$  in the basis of  $V_{kh}^*$  spanned by the Dirac

functionals  $\delta^{ml}(v_{kh}) := v_{kh}^m(x_l)$ ,  $m = 1, \dots, M$ ,  $l = 1, \dots, \dim(V_h)$ , and take into account the properties of the set in (4.6), then it follows straightforwardly that  $\bar{\mu}_{kh} = \sum_{m,l} \bar{\lambda}^{ml} \delta^{ml}$  has to hold with coefficients  $\bar{\lambda}^{ml}$  that satisfy the conditions in the last line of (4.5). Suppose now that  $\bar{p}_{kh}$  is the unique element of  $V_{kh}$  that satisfies the variational identity in the second line of (4.5) for these  $\bar{\lambda}^{ml}$ . Then, the definitions of  $S_{kh}^0$  and  $\bar{\lambda}^{ml}$  and (4.7) yield

$$0 = B(S_{kh}^0(z), \bar{p}_{kh}) + (\nu \bar{u}_{kh}, z)_{L^2(0,T;L^2(\Omega))} = (z, \bar{p}_{kh} + \nu \bar{u}_{kh})_{L^2(0,T;L^2(\Omega))}$$

for all  $z \in L^2(0,T;L^2(\Omega))$ . This shows that  $\bar{p}_{kh} + \nu \bar{u}_{kh} = 0$  has to hold, that  $\bar{u}_{kh}$  is indeed an element of  $V_{kh}$ , and that  $\bar{u}_{kh}$ ,  $\bar{y}_{kh}$ ,  $\bar{p}_{kh}$  and  $\bar{\lambda}^{ml}$  satisfy (4.5) as claimed.  $\square$

Note that the control is discretized variationally in the above, cf. [22, 32, 36], and that the structure of (4.5) is completely analogous to that of the KKT-system (3.27) satisfied by  $\bar{u}$ , see Example 3.9. To relate the solutions  $\bar{u}$  and  $\bar{u}_{kh}$  of the problems (4.1) and (4.4) to each other, we need two auxiliary results. The first one collects some properties of the multipliers in (3.27) and (4.5):

**Lemma 4.5.** *The measure  $\bar{\lambda} \in \mathcal{M}([0, T] \times \text{cl}(\Omega))$  and the multipliers  $\bar{\lambda}^{ml} \in \mathbb{R}$  in the optimality systems (3.27) and (4.5) satisfied by the solutions  $\bar{u}$  and  $\bar{u}_{kh}$  of (4.1) and (4.4), respectively, are unique. Further, there exist an open set  $D \subset \mathbb{R}^d$ , which is compactly contained in  $\Omega$ , and constants  $C, \varepsilon > 0$  such that the support of  $\bar{\lambda}$  is a subset of  $[\varepsilon, T] \times D$  and such that, for all sufficiently small  $k$  and  $h$ , it holds*

$$\|\bar{u}_{kh}\|_{L^2(0,T;L^2(\Omega))} + \|\bar{y}_{kh}\|_{L^2(0,T;L^2(\Omega))} + \sum_{m,l} \bar{\lambda}^{ml} \leq C. \quad (4.8)$$

*Proof.* The uniqueness of the multipliers  $\bar{\lambda}^{ml}$  follows immediately from (4.5), and the uniqueness of the measure  $\bar{\lambda}$  is a straightforward consequence of our assumptions  $\psi < 0$  and  $\psi < y_0$ , the continuity and zero boundary conditions of the state  $\bar{y} = S(\bar{u})$ , the conditions on  $\bar{\lambda}$  in (3.27), and Corollary 3.6. In the proof of Corollary 3.6, we have further already established that the support of  $\bar{\lambda}$  is contained in the contact set  $\{(t, x) \in [0, T] \times \text{cl}(\Omega) \mid (\bar{y} - \psi)(t, x) = 0\}$ , and this set is trivially contained in a cylinder of the form  $[\varepsilon, T] \times D$  with some open set  $D$  that is compactly contained in  $\Omega$  and some  $\varepsilon > 0$  due to the continuity and zero boundary conditions of  $\bar{y}$  and again the properties  $\psi < 0$  and  $y_0 > \psi$ . It remains to prove (4.8) for all sufficiently small  $k$  and  $h$ . This boundedness property, however, follows immediately from the fact that the control  $u := -\Delta y_0 \in L^\infty((0, T) \times \Omega)$  satisfies  $C([0, T] \times \text{cl}(\Omega)) \ni S(u) = y_0 \geq \psi + \varepsilon$  in  $\Omega$  for a.a.  $t \in (0, T)$  with some  $\varepsilon > 0$ , the  $L^\infty$ -error estimate in Lemma 4.3(ii), and exactly the same calculation as in the proof of ([32], Lem. 2).  $\square$

Our second auxiliary result is concerned with the regularity and approximability properties of the function that we obtain when we solve the adjoint equation in (3.27) with right-hand side  $-\bar{\lambda}$ . Before we state it, we introduce some notation:

**Definition 4.6** (Ritz projection, discrete Laplacian and nodal interpolation). Henceforth, we denote with:

(i)  $\Delta_h: V_h \rightarrow V_h$  the discrete Laplacian defined by

$$(-\Delta_h v_h, z_h)_{L^2(\Omega)} = (\nabla v_h, \nabla z_h)_{L^2(\Omega)^d} \quad \forall z_h \in V_h, \quad \forall v_h \in V_h.$$

(ii)  $R_h: H_0^1(\Omega) \rightarrow V_h$  the Ritz projection onto  $V_h$  defined by

$$R_h(v) \in V_h, \quad \int_{\Omega} \nabla(v - R_h(v)) \cdot \nabla z_h \, dx = 0 \quad \forall z_h \in V_h \quad \forall v \in H_0^1(\Omega).$$



(iii)  $\mathcal{I}_{kh}: V_{kh} \rightarrow C([0, T] \times \text{cl}(\Omega))$  the interpolation operator defined by

$$\begin{aligned} \mathcal{I}_{kh}(v_{kh})(t) &:= \frac{t - t_0}{k_1} v_{kh}^1 \quad \forall t \in I_1, \\ \mathcal{I}_{kh}(v_{kh})(t) &:= \frac{t - t_{m-1}}{k_m} v_{kh}^m + \frac{t_m - t}{k_m} v_{kh}^{m-1} \quad \forall t \in I_m \quad \forall m = 2, \dots, M. \end{aligned}$$

(iv)  $\mathcal{L}_h: C(\text{cl}(\Omega)) \rightarrow V_h$  the Lagrange interpolation operator associated with  $V_h$ .

We can now prove:

**Lemma 4.7.** *Let  $\bar{\lambda}$  denote the measure in the optimality system (3.27), and let  $\mathcal{I}_{kh}: V_{kh} \rightarrow C([0, T] \times \text{cl}(\Omega))$  be defined as in Definition 4.6(iii). Then there exists a unique, weak solution  $\zeta \in L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega))$  of the PDE*

$$\begin{aligned} -\partial_t \zeta - \Delta \zeta &= -\bar{\lambda} \text{ in } (0, T) \times \Omega, \\ \text{tr}(\zeta) &= 0 \text{ on } (0, T) \times \partial\Omega, \quad \zeta(T) = 0 \text{ in } \Omega, \end{aligned} \tag{4.9}$$

and there is a constant  $C > 0$  independent of  $k$  and  $h$  such that  $\zeta$  and the function  $\zeta_{kh} \in V_{kh}$  defined by

$$B(v_{kh}, \zeta_{kh}) = \langle -\bar{\lambda}, \mathcal{I}_{kh}(v_{kh}) \rangle_{C([0, T] \times \text{cl}(\Omega))} \quad \forall v_{kh} \in V_{kh} \tag{4.10}$$

satisfy

$$\|\zeta - \zeta_{kh}\|_{L^2(0, T; L^2(\Omega))}^2 \leq C |\ln(k)|^3 |\ln(h)|^6 (k + h^2). \tag{4.11}$$

*Proof.* From classical results on the existence, uniqueness, and regularity of solutions of parabolic PDEs, see Proposition 2.3 and ([43], Prop. 2.6), and from the  $L^\infty(0, T; L^2(\Omega))$ -regularity of the function  $\bar{y} - y_D$ , we obtain that there exists a unique  $\vartheta \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega)) \cap L^\infty((0, T) \times \Omega)$  satisfying

$$\begin{aligned} -\partial_t \vartheta - \Delta \vartheta &= \bar{y} - y_D \text{ a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ \text{tr}(\vartheta) &= 0 \text{ a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \quad \vartheta(T) = 0 \text{ a.e. in } \Omega. \end{aligned}$$

Since the adjoint state  $\bar{p}$  in (3.27) is an element of the space  $L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega))$  and a solution of the adjoint equation in (3.27), from the linearity of the involved PDEs, it now follows immediately that the function  $\zeta := \bar{p} - \vartheta$  is an element of  $L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega))$  that satisfies (4.9) (in the weak sense analogous to (3.13)). Since the PDE (4.9) can have at most one solution (as one may easily check by contradiction), this proves the first assertion of Lemma 4.7. To obtain the unique solvability of the discrete problem (4.10), it suffices to note that this variational equality can be rewritten as a time-stepping scheme that proceeds backwards in time and to invoke classical results on the existence and uniqueness of solutions of elliptic problems, cf. the proof of ([43], Thm. 3.2). It remains to establish the estimate (4.11). To this end, we first prove that there exists a constant  $C > 0$  independent of  $k$  and  $h$  such that the functions  $\zeta_{kh}$  defined by (4.10) satisfy

$$\|\zeta_{kh}\|_{L^\infty((0, T) \times \Omega)} \leq C |\ln(k)| |\ln(h)|. \tag{4.12}$$

So let us consider an arbitrary but fixed  $x_0 \in \Omega$  and some  $m \in \{1, \dots, M\}$ , and let  $\tilde{\delta}_{x_0}$  be a mollification of the Dirac measure  $\delta_{x_0}$  at  $x_0$  as constructed in ([51], Appendix A5), i.e., a function  $\tilde{\delta}_{x_0} \in L^1(\Omega)$  supported in

one cell of the mesh  $\mathcal{T}_h$  satisfying  $v_h(x_0) = (\tilde{\delta}_{x_0}, v_h)_{L^2(\Omega)}$  for all  $v_h \in V_h$  and  $\|\tilde{\delta}_{x_0}\|_{L^1(\Omega)} \leq C$  for some constant  $C > 0$  independent of  $k$  and  $h$ . Then there exists a unique solution  $\varrho_{kh} \in V_{kh}$  of the problem

$$B(\varrho_{kh}, v_{kh}) = \int_{I_m} \int_{\Omega} \frac{1}{k_m} \tilde{\delta}_{x_0} v_{kh}^m dx dt = v_{kh}^m(x_0) \quad \forall v_{kh} \in V_{kh},$$

and we obtain from ([40], Thm. 11) that we can find a constant  $C > 0$  independent of  $k$  and  $h$  with

$$\sum_{m=1}^M k_m \|\Delta_h \varrho_{kh}^m\|_{L^1(\Omega)} + \|\varrho_{kh}^1\|_{L^1(\Omega)} + \sum_{m=2}^M \|\varrho_{kh}^m - \varrho_{kh}^{m-1}\|_{L^1(\Omega)} \leq C |\ln(k)|, \quad (4.13)$$

where  $\Delta_h: V_h \rightarrow V_h$  is the discrete Laplacian from Definition 4.6(i). By exploiting the properties of  $\zeta$ ,  $\zeta_{kh}$  and  $\varrho_{kh}$ , (4.13), the definition of the operator  $\mathcal{I}_{kh}$ , the stability estimate  $\|\mathcal{I}_{kh}(v_{kh})\|_{L^1(0,T;L^1(\Omega))} \leq C \|v_{kh}\|_{L^1(0,T;L^1(\Omega))}$  for all  $v_{kh} \in V_{kh}$  (which is easy to check), and the fact that the Ritz projection  $R_h$  satisfies

$$\|R_h(v)\|_{L^\infty(\Omega)} \leq C |\ln(h)| \|v\|_{L^\infty(\Omega)} \quad \forall v \in H_0^1(\Omega) \cap L^\infty(\Omega)$$

for some  $C > 0$  independent of  $k$  and  $h$ , see ([39], Lem. 5), we may now compute

$$\begin{aligned} \zeta_{kh}^m(x_0) &= B(\varrho_{kh}, \zeta_{kh}) = \langle -\bar{\lambda}, \mathcal{I}_{kh}(\varrho_{kh}) \rangle_{C([0,T] \times \text{cl}(\Omega))} \\ &= \int_0^T \left( (\partial_t \mathcal{I}_{kh}(\varrho_{kh}), \zeta)_{L^2(\Omega)} + (\nabla \mathcal{I}_{kh}(\varrho_{kh}), \nabla \zeta)_{L^2(\Omega)^d} \right) dt \\ &= \int_0^T \left( (\partial_t \mathcal{I}_{kh}(\varrho_{kh}), \zeta)_{L^2(\Omega)} + (\nabla \mathcal{I}_{kh}(\varrho_{kh}), \nabla R_h(\zeta))_{L^2(\Omega)^d} \right) dt \\ &\leq C |\ln(h)| \left( \|\partial_t \mathcal{I}_{kh}(\varrho_{kh})\|_{L^1(0,T;L^1(\Omega))} + \|\Delta_h \mathcal{I}_{kh}(\varrho_{kh})\|_{L^1(0,T;L^1(\Omega))} \right) \\ &= C |\ln(h)| \left( \|\varrho_{kh}^1\|_{L^1(\Omega)} + \sum_{m=2}^M \|\varrho_{kh}^m - \varrho_{kh}^{m-1}\|_{L^1(\Omega)} + \|\mathcal{I}_{kh}(\Delta_h \varrho_{kh})\|_{L^1(0,T;L^1(\Omega))} \right) \\ &\leq C |\ln(k)| |\ln(h)|, \end{aligned}$$

where the constant  $C$  may change from step to step but is always independent of  $k$  and  $h$ . This establishes (4.12) as desired. To prove the estimate (4.11), we can now use a standard duality argument: Let  $w \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))$  be the (necessarily unique) solution of the PDE

$$\begin{aligned} \partial_t w - \Delta w &= \text{sgn}(\zeta - \zeta_{kh}) \text{ a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\ \text{tr}(w) &= 0 \text{ a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \quad w(0) = 0 \text{ a.e. in } \Omega, \end{aligned}$$

and let  $w_{kh}$  be the discrete approximation of  $w$  defined by

$$w_{kh} \in V_{kh}, \quad B(w_{kh}, v_{kh}) = (\text{sgn}(\zeta - \zeta_{kh}), v_{kh})_{L^2(0,T;L^2(\Omega))} \quad \forall v_{kh} \in V_{kh}.$$

Then the  $L^\infty((0, T) \times \Omega)$ -regularity of the right-hand side  $\text{sgn}(\zeta - \zeta_{kh})$  implies that the function  $w$  is also an element of  $C([0, T] \times \text{cl}(\Omega))$ , see ([43], Prop. 2.7), and we may use the definition of  $\mathcal{I}_{kh}$  and the properties of  $w$ ,

$\zeta$ ,  $w_{kh}$ , and  $\zeta_{kh}$  to obtain

$$\begin{aligned}
\|\zeta - \zeta_{kh}\|_{L^1((0,T)\times\Omega)} &= (\operatorname{sgn}(\zeta - \zeta_{kh}), \zeta - \zeta_{kh})_{L^2(0,T;L^2(\Omega))} \\
&= \int_0^T \left( (\partial_t w, \zeta)_{L^2(\Omega)} + (\nabla w, \nabla \zeta)_{L^2(\Omega)^d} \right) dt - B(w_{kh}, \zeta_{kh}) \\
&= \langle \bar{\lambda}, \mathcal{I}_{kh}(w_{kh}) - w \rangle_{C([0,T]\times\operatorname{cl}(\Omega))} \\
&\leq \|\bar{\lambda}\|_{\mathcal{M}([0,T]\times\operatorname{cl}(\Omega))} \left( \|w_{kh} - w\|_{L^\infty((\varepsilon,T)\times D)} + \|\mathcal{I}_{kh}(w_{kh}) - w_{kh}\|_{L^\infty((\varepsilon,T)\times D)} \right) \\
&\leq \|\bar{\lambda}\|_{\mathcal{M}([0,T]\times\operatorname{cl}(\Omega))} \|w_{kh} - w\|_{L^\infty((\varepsilon,T)\times D)} \\
&\quad + k \|\bar{\lambda}\|_{\mathcal{M}([0,T]\times\operatorname{cl}(\Omega))} \left( \frac{\|w_{kh}^1\|_{L^\infty(\Omega)}}{k_1} + \max_{m=2,\dots,M} \frac{\|w_{kh}^m - w_{kh}^{m-1}\|_{L^\infty(\Omega)}}{k_m} \right).
\end{aligned} \tag{4.14}$$

Here,  $\varepsilon > 0$  and  $D \subset \Omega$  are chosen as in Lemma 4.5, *i.e.*, such that  $\operatorname{supp}(\bar{\lambda}) \subset [\varepsilon, T] \times D$  and  $\operatorname{cl}(D) \subset \Omega$ . Since Lemma 4.3(iii) implies that  $\|w_{kh} - w\|_{L^\infty((\varepsilon,T)\times D)} \leq C |\ln(k)|^2 |\ln(h)|^5 (k + h^2)$  holds and since the results in ([40], Cor. 2, Thm. 11) allow to bound the second term on the right-hand side of (4.14) by  $C |\ln(k)| k$ , the last estimate yields that there exists a constant  $C > 0$  independent of  $k$  and  $h$  with  $\|\zeta - \zeta_{kh}\|_{L^1((0,T)\times\Omega)} \leq C |\ln(k)|^2 |\ln(h)|^5 (k + h^2)$ . Since  $\zeta$  is an element of  $L^\infty((0, T) \times \Omega)$  and since the sequence  $\{\zeta_{kh}\}$  satisfies (4.12), we may now conclude that the  $L^2$ -error between  $\zeta$  and  $\zeta_{kh}$  satisfies

$$\|\zeta - \zeta_{kh}\|_{L^2(0,T;L^2(\Omega))}^2 \leq \|\zeta - \zeta_{kh}\|_{L^\infty((0,T)\times\Omega)} \|\zeta - \zeta_{kh}\|_{L^1((0,T)\times\Omega)} \leq C |\ln(k)|^3 |\ln(h)|^6 (k + h^2).$$

This completes the proof.  $\square$

We are now in the position to prove the main result of this section:

**Theorem 4.8** (FE-error estimate for (4.1)). *Let  $\bar{u}$  and  $\bar{u}_{kh}$  be the solutions of the problems (4.1) and (4.4), respectively. Then there exists a constant  $C > 0$  which is independent of  $k$  and  $h$  (but may depend on the problem data  $\Omega$ ,  $\psi$ , etc.) such that*

$$\|\bar{u} - \bar{u}_{kh}\|_{L^2(0,T;L^2(\Omega))} \leq C |\ln(k)|^{3/2} |\ln(h)|^3 \left( k^{1/2} + h \right) \tag{4.15}$$

holds for all sufficiently small  $k$  and  $h$ .

*Proof.* To establish (4.15), we proceed along roughly the same lines as in the proof of ([12], Thm. 5.5). Let us first introduce some notation: Henceforth, with  $\zeta$  and  $\zeta_{kh}$ , we again denote the functions in (4.9) and (4.10), respectively, and with  $\bar{\lambda} \in \mathcal{M}([0, T] \times \operatorname{cl}(\Omega))$  and  $\bar{\lambda}^{ml} \in \mathbb{R}$  the multipliers in the systems (3.27) and (4.5). Further, we define  $\vartheta$ ,  $\chi$ ,  $\theta$ ,  $\theta_{kh}$  and  $\xi_{kh}$  to be the (necessarily unique) solutions of the auxiliary problems

$$\begin{aligned}
-\partial_t \vartheta - \Delta \vartheta &= \bar{y} - y_D \text{ a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
\operatorname{tr}(\vartheta) &= 0 \text{ a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \quad \vartheta(T) = 0 \text{ a.e. in } \Omega, \\
-\partial_t \chi - \Delta \chi &= S(\bar{u}_{kh}) - y_D \text{ a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
\operatorname{tr}(\chi) &= 0 \text{ a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \quad \chi(T) = 0 \text{ a.e. in } \Omega, \\
-\partial_t \theta - \Delta \theta &= \bar{y}_{kh} - y_D \text{ a.e. in } \Omega \text{ for a.a. } t \in (0, T), \\
\operatorname{tr}(\theta) &= 0 \text{ a.e. on } \partial\Omega \text{ for a.a. } t \in (0, T), \quad \theta(T) = 0 \text{ a.e. in } \Omega, \\
\theta_{kh} \in V_{kh}, \quad B(v_{kh}, \theta_{kh}) &= (\bar{y}_{kh} - y_D, v_{kh})_{L^2(0,T;L^2(\Omega))} \quad \forall v_{kh} \in V_{kh},
\end{aligned}$$

and

$$\xi_{kh} \in V_{kh}, \quad B(v_{kh}, \xi_{kh}) = - \sum_{m=1}^M \sum_{l=1}^{\dim(V_h)} \bar{\chi}^{ml} v_{kh}^m(x_l) \quad \forall v_{kh} \in V_{kh}.$$

Here,  $S$  again denotes the solution operator of the PDE (4.2). Note that the properties of the functions  $\vartheta$ ,  $\zeta$ ,  $\theta_{kh}$  and  $\xi_{kh}$  and the adjoint states  $\bar{p}$  and  $\bar{p}_{kh}$  in (3.27) and (4.5) imply that  $-\nu\bar{u} = \bar{p} = \vartheta + \zeta$  and  $-\nu\bar{u}_{kh} = \bar{p}_{kh} = \theta_{kh} + \xi_{kh}$  holds, and that Proposition 2.3 and a simple reparametrization argument yield  $\vartheta, \chi, \theta \in L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))$ . From the PDEs for  $\vartheta$  and  $\chi$ , we further obtain

$$\begin{aligned} (\vartheta - \chi, \bar{u} - \bar{u}_{kh})_{L^2(0, T; L^2(\Omega))} &= (\vartheta - \chi, \partial_t(S(\bar{u}) - S(\bar{u}_{kh})) - \Delta(S(\bar{u}) - S(\bar{u}_{kh})))_{L^2(0, T; L^2(\Omega))} \\ &= (-\partial_t(\vartheta - \chi) - \Delta(\vartheta - \chi), S(\bar{u}) - S(\bar{u}_{kh}))_{L^2(0, T; L^2(\Omega))} \\ &= \|S(\bar{u}) - S(\bar{u}_{kh})\|_{L^2(0, T; L^2(\Omega))}^2 \geq 0. \end{aligned}$$

The above estimate allows us to compute

$$\begin{aligned} \nu \|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))}^2 &\leq (\nu\bar{u} + \vartheta - \nu\bar{u}_{kh} - \theta_{kh} + \theta_{kh} - \chi, \bar{u} - \bar{u}_{kh})_{L^2(0, T; L^2(\Omega))} \\ &\leq (-\zeta + \xi_{kh}, \bar{u} - \bar{u}_{kh})_{L^2(0, T; L^2(\Omega))} \\ &\quad + \|\chi - \theta_{kh}\|_{L^2(0, T; L^2(\Omega))} \|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))} \\ &\leq (\xi_{kh} - \zeta_{kh}, \bar{u} - \bar{u}_{kh})_{L^2(0, T; L^2(\Omega))} \\ &\quad + \|\zeta_{kh} - \zeta\|_{L^2(0, T; L^2(\Omega))} \|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))} \\ &\quad + \|\theta - \theta_{kh}\|_{L^2(0, T; L^2(\Omega))} \|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))} \\ &\quad + \|\chi - \theta\|_{L^2(0, T; L^2(\Omega))} \|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))}, \end{aligned}$$

which, in combination with Young's inequality, implies that there exist constants  $C_1, C_2 > 0$  depending only on the Tikhonov parameter  $\nu$  with

$$\begin{aligned} \|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))}^2 &\leq C_1 (\xi_{kh} - \zeta_{kh}, \bar{u} - \bar{u}_{kh})_{L^2(0, T; L^2(\Omega))} \\ &\quad + C_2 \left( \|\zeta_{kh} - \zeta\|_{L^2(0, T; L^2(\Omega))}^2 + \|\theta - \theta_{kh}\|_{L^2(0, T; L^2(\Omega))}^2 + \|\chi - \theta\|_{L^2(0, T; L^2(\Omega))}^2 \right). \end{aligned} \quad (4.16)$$

Note that, from Lemma 4.7, we immediately obtain that

$$\|\zeta - \zeta_{kh}\|_{L^2(0, T; L^2(\Omega))}^2 \leq C |\ln(k)|^3 |\ln(h)|^6 (k + h^2), \quad (4.17)$$

and that Lemma 4.3(i), a simple reparametrization in time and Lemma 4.5 yield

$$\|\theta - \theta_{kh}\|_{L^2(0, T; L^2(\Omega))} \leq C(k + h^2) \|\bar{y}_{kh} - y_D\|_{L^2(0, T; L^2(\Omega))} \leq C(k + h^2) \quad (4.18)$$

for all sufficiently small  $k$  and  $h$ , where  $C$  is again a generic constant independent of  $k$  and  $h$ . From the Lipschitz continuity properties of the map  $S$  in Proposition 2.3, again a reparametrization, Lemma 4.5 and Lemma 4.3(i), we may further deduce that

$$\begin{aligned} \|\chi - \theta\|_{L^2(0, T; L^2(\Omega))} &\leq C \|S(\bar{u}_{kh}) - S_{kh}(\bar{u}_{kh})\|_{L^2(0, T; L^2(\Omega))} \\ &\leq C (k + h^2) (\|\bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))} + \|y_0\|_{H^1(\Omega)}) \leq C (k + h^2). \end{aligned}$$

It remains to estimate the first term on the right-hand side of (4.16). To this end, we note that the properties of  $\xi_{kh}$ ,  $\zeta_{kh}$ ,  $\bar{y}_{kh}$ ,  $\bar{y}$ ,  $\bar{\lambda}$  and  $\bar{\lambda}^{ml}$  yield

$$\begin{aligned}
& (\xi_{kh} - \zeta_{kh}, \bar{u} - \bar{u}_{kh})_{L^2(0,T;L^2(\Omega))} \\
&= B(S_{kh}(\bar{u}) - \bar{y}_{kh}, \xi_{kh} - \zeta_{kh}) \\
&= \langle \bar{\lambda}, \mathcal{I}_{kh}(S_{kh}(\bar{u}) - \bar{y}_{kh}) \rangle_{C([0,T] \times \text{cl}(\Omega))} + \sum_{m,l} \bar{\lambda}^{ml} (\bar{y}_{kh}^m - S_{kh}(\bar{u})^m)(x_l) \\
&= \langle \bar{\lambda}, \mathcal{I}_{kh}(S_{kh}(\bar{u}) - \mathfrak{L}_h(\psi)) \rangle_{C([0,T] \times \text{cl}(\Omega))} + \langle \bar{\lambda}, \mathcal{I}_{kh}(\mathfrak{L}_h(\psi) - \bar{y}_{kh}) \rangle_{C([0,T] \times \text{cl}(\Omega))} \\
&\quad + \sum_{m,l} \bar{\lambda}^{ml} (\psi - S_{kh}(\bar{u})^m)(x_l) \\
&\leq \langle \bar{\lambda}, \mathcal{I}_{kh}(S_{kh}(\bar{u}) - \mathfrak{L}_h(\psi)) \rangle_{C([0,T] \times \text{cl}(\Omega))} + \sum_{m,l} \bar{\lambda}^{ml} \max(0, (\psi - S_{kh}(\bar{u})^m)(x_l)) \\
&= \langle \bar{\lambda}, \mathcal{I}_{kh}(S_{kh}(\bar{u}) - \bar{y}) \rangle_{C([0,T] \times \text{cl}(\Omega))} + \langle \bar{\lambda}, \psi - \mathcal{I}_{kh}(\mathfrak{L}_h(\psi)) \rangle_{C([0,T] \times \text{cl}(\Omega))} \\
&\quad + \sum_{m,l} \bar{\lambda}^{ml} \max(0, (\psi - S_{kh}(\bar{u})^m)(x_l)),
\end{aligned} \tag{4.19}$$

where  $S_{kh}$  denotes the solution map of (4.3), where  $\mathcal{I}_{kh}$  and  $\mathfrak{L}_h$  are the interpolation operators defined in Definition 4.6, and where  $\mathfrak{L}_h(\psi)$  is identified with an element of  $V_{kh}$ . For the first term on the right-hand side of (4.19), we obtain, analogously to the estimate for the quantity  $\langle \bar{\lambda}, \mathcal{I}_{kh}(w_{kh}) - w \rangle_{C([0,T] \times \text{cl}(\Omega))}$  in the proof of Lemma 4.7, from Lemma 4.5, Lemma 4.3(iii) and the results in [40] that

$$\langle \bar{\lambda}, \mathcal{I}_{kh}(S_{kh}(\bar{u}) - \bar{y}) \rangle_{C([0,T] \times \text{cl}(\Omega))} \leq C |\ln(k)|^2 |\ln(h)|^5 (k + h^2).$$

Further, for the second term on the right-hand side of (4.19), it holds

$$\langle \bar{\lambda}, \psi - \mathcal{I}_{kh}(\mathfrak{L}_h(\psi)) \rangle_{C([0,T] \times \text{cl}(\Omega))} \leq Ch^2$$

for all sufficiently small  $k$  and  $h$  due to the definition of  $\mathcal{I}_{kh}$ , the properties of the support of  $\bar{\lambda}$  in Lemma 4.5 and the  $W^{2,\infty}$ -regularity of  $\psi$ . To finally control the third term on the right-hand side of (4.19), we note that Lemma 4.3(ii), our assumption  $\psi < 0$  on  $\partial\Omega$ , and the continuity and zero boundary conditions of  $\bar{y}$  imply that there exist a constant  $\varepsilon > 0$  and an open set  $D \subset \mathbb{R}^d$ , which is compactly contained in  $\Omega$ , such that, for all sufficiently small  $k$  and  $h$ , we have  $S_{kh}(\bar{u}) \geq \psi + \varepsilon$  a.e. in  $(0, T) \times \Omega \setminus D$  and, as a consequence,

$$\max(0, (\psi - S_{kh}(\bar{u})^m)(x_l)) = 0 \text{ for all } x_l \text{ with } x_l \in \Omega \setminus D \text{ and all } m.$$

Using the above, Lemma 4.3(iii) and again Lemma 4.5, we obtain

$$\begin{aligned}
\sum_{m,l} \bar{\lambda}^{ml} \max(0, (\psi - S_{kh}(\bar{u})^m)(x_l)) &\leq C \|\max(0, \psi - S_{kh}(\bar{u}))\|_{L^\infty((0,T) \times D)} \\
&\leq C \|\max(0, \bar{y} - S_{kh}(\bar{u}))\|_{L^\infty((0,T) \times D)} \\
&\leq C |\ln(k)|^2 |\ln(h)|^5 (k + h^2)
\end{aligned}$$

for all sufficiently small  $k$  and  $h$ , where  $C$  is again a generic constant. The desired estimate (4.15) now follows immediately from (4.16), (4.17) and (4.18).  $\square$

Quite surprisingly, the error estimate (4.15), that we have established in the last theorem by exploiting the improved regularity results from Section 3, can itself be used to derive further information about the regularity properties of  $\bar{u}$ . Indeed, by proceeding along the lines of ([42], Sect. 7.1), we obtain:

**Corollary 4.9** ( $H^s(0, T; L^2(\Omega))$ -regularity of optimal controls). *The optimal control  $\bar{u}$  of (4.1) is also an element of  $H^s(0, T; L^2(\Omega))$  for all  $0 < s < 1/2$ .*

*Proof.* Suppose that  $s \in (0, 1/2)$  is arbitrary but fixed, consider the family of equidistant partitions of the time interval  $(0, T)$  associated with the widths  $k_l := 2^{-l}T$ ,  $l \in \mathbb{N}$ , and choose a quasi-uniform sequence of triangulations of  $\Omega$  such that there exists a constant  $c > 0$  with  $c^{-1}k_l^{1/2} \leq h_l \leq ck_l^{1/2}$  for all  $l \in \mathbb{N}$ . Then it follows from (4.15) that we can find constants  $C > 0$  and  $\varepsilon \in (0, 1/2 - s)$  independent of  $l$  such that the solutions  $\bar{u}$  and  $\bar{u}_l := \bar{u}_{k_l h_l}$  of (4.1) and (4.4) satisfy  $\|\bar{u} - \bar{u}_l\|_{L^2(0, T; L^2(\Omega))} \leq Ck_l^{s+\varepsilon}$  for all sufficiently large  $l \in \mathbb{N}$ . Using this inequality and the same inverse estimate as in the proof of ([42], Thm. 7.1), we obtain that, for all  $l_2 \geq l_1 \geq L$ ,  $L \in \mathbb{N}$  sufficiently large, we have

$$\begin{aligned} \|\bar{u}_{l_1} - \bar{u}_{l_2}\|_{H^s(0, T; L^2(\Omega))} &\leq \sum_{l=l_1}^{l_2-1} \|\bar{u}_l - \bar{u}_{l+1}\|_{H^s(0, T; L^2(\Omega))} \leq C \sum_{l=l_1}^{l_2-1} k_{l+1}^{-s} \|\bar{u}_l - \bar{u}_{l+1}\|_{L^2(0, T; L^2(\Omega))} \\ &\leq C \sum_{l=l_1}^{l_2-1} k_{l+1}^{-s} (\|\bar{u}_l - \bar{u}\|_{L^2(0, T; L^2(\Omega))} + \|\bar{u} - \bar{u}_{l+1}\|_{L^2(0, T; L^2(\Omega))}) \\ &\leq C \sum_{l=l_1}^{l_2-1} k_{l+1}^{-s} (k_l^{s+\varepsilon} + k_{l+1}^{s+\varepsilon}) \leq 2^{s+1}CT^\varepsilon \sum_{l=L}^{\infty} 2^{-\varepsilon l} = o(1), \end{aligned}$$

where the symbol  $o(1)$  refers to the limit  $L \rightarrow \infty$ . This shows that  $\{\bar{u}_l\}$  is a Cauchy sequence in  $H^s(0, T; L^2(\Omega))$  and, since we already know that  $\bar{u}_l$  converges to  $\bar{u}$  in  $L^2(0, T; L^2(\Omega))$ , that  $\bar{u} \in H^s(0, T; L^2(\Omega))$  as claimed.  $\square$

In summary (see (3.27) and Cor. 4.9), we have now proved that the optimal control  $\bar{u}$  of (4.1) satisfies

$$\bar{u} \in L^\infty((0, T) \times \Omega) \cap L^2(0, T; H_0^1(\Omega)) \cap BV([0, T]; Y_\gamma^*) \cap H^s(0, T; L^2(\Omega)) \quad (4.20)$$

for all  $s \in (0, 1/2)$  and all  $\gamma > d/2$ . Note that this inclusion recovers in particular all of the regularity properties that have been established in [42] for solutions of parabolic problems with  $L^\infty$ -bounds on the control and state constraints of integral type.

#### 4.1. Numerical experiments

To validate the error estimate (4.15), we consider a simple numerical example with a known analytic solution: Define

$$\rho(s) := \begin{cases} \sin(\pi/4) + \pi \cos(\pi/4) (s - 1/4) & \text{if } s \in (0, 1/4) \\ \sin(\pi s) & \text{if } s \in [1/4, 3/4] \\ \sin(3\pi/4) + \pi \cos(3\pi/4) (s - 3/4) & \text{if } s \in (3/4, 1) \end{cases}$$

and

$$\varphi(s) := \begin{cases} 128s^3 - 48s^2 & \text{if } s \in (0, 1/4) \\ -1 & \text{if } s \in [1/4, 3/4] \\ 128(1-s)^3 - 48(1-s)^2 & \text{if } s \in (3/4, 1) \end{cases},$$

TABLE 1. EOCs in  $L^2(0, T; L^2(\Omega))$  for sequences of meshes that resolve the boundary of the active set  $(1/4, 3/4)^3$  perfectly in space and time. In (A), the spatial and the temporal widths are coupled, and in (B) and (C) one of the widths is fixed.

(A) $k_l = 0.5h_l^2$			(B) $k_l \equiv 1/1024$		(C) $h_l \equiv 1/64$	
$h_l$	$(\text{EOC}^h)_l$	$(\text{EOC}^k)_l$	$h_l$	$(\text{EOC}^h)_l$	$k_l$	$(\text{EOC}^k)_l$
$\frac{1}{4}$	-	-	$\frac{1}{4}$	-	$\frac{1}{16}$	-
$\frac{1}{8}$	0.5541	0.2770	$\frac{1}{8}$	0.3876	$\frac{1}{24}$	0.9101
$\frac{1}{12}$	1.8847	0.9423	$\frac{1}{12}$	1.8794	$\frac{1}{32}$	0.9384
$\frac{1}{16}$	1.9656	0.9828	$\frac{1}{16}$	1.9559	$\frac{1}{40}$	0.9520
$\frac{1}{20}$	1.9884	0.9942	$\frac{1}{20}$	1.9688	$\frac{1}{48}$	0.9596

and set  $\Omega := (0, 1)^2$ ,  $T := 1$ ,  $y_0 := 0$ ,  $\nu := 10^{-5}$  and

$$\begin{aligned}
\psi(x_1, x_2) &:= -\rho(x_1)\rho(x_2), \\
\bar{y}(t, x_1, x_2) &:= \varphi(t) \sin(\pi x_1) \sin(\pi x_2), \\
\bar{u}(t, x_1, x_2) &:= \varphi'(t) \sin(\pi x_1) \sin(\pi x_2) + 2\pi^2 \varphi(t) \sin(\pi x_1) \sin(\pi x_2), \\
y_D(t, x_1, x_2) &:= \varphi(t) \sin(\pi x_1) \sin(\pi x_2) - \nu \varphi''(t) \sin(\pi x_1) \sin(\pi x_2) \\
&\quad + 4\pi^4 \nu \varphi(t) \sin(\pi x_1) \sin(\pi x_2) - \mathbb{1}_{(1/4, 3/4)^3}(t, x_1, x_2)
\end{aligned} \tag{4.21}$$

for a.a.  $t \in (0, 1)$  and  $(x_1, x_2) \in \Omega$ , where  $\varphi''$  is understood in the weak sense. Then it is easy to check that  $\Omega$ ,  $T$ ,  $\nu$ ,  $y_D$ ,  $y_0$  and  $\psi$  satisfy the conditions in Assumption 4.1 and that  $\bar{u}$  and  $\bar{y}$  solve the KKT-system (3.27) with the adjoint state  $\bar{p} := -\nu \bar{u}$  and the multiplier  $\bar{\lambda} := \mathbb{1}_{(1/4, 3/4)^3}$ . This shows that the function  $\bar{u}$  is the (necessarily unique) global minimizer of the problem (4.1) when  $\psi$  etc. are chosen as in (4.21).

The experimental orders of convergence

$$\begin{aligned}
(\text{EOC}^h)_l &:= \frac{\log \|\bar{u} - \bar{u}_{k_l h_l}\|_{L^2(0, T; L^2(\Omega))} - \log \|\bar{u} - \bar{u}_{k_{l-1} h_{l-1}}\|_{L^2(0, T; L^2(\Omega))}}{\log h_l - \log h_{l-1}}, \\
(\text{EOC}^k)_l &:= \frac{\log \|\bar{u} - \bar{u}_{k_l h_l}\|_{L^2(0, T; L^2(\Omega))} - \log \|\bar{u} - \bar{u}_{k_{l-1} h_{l-1}}\|_{L^2(0, T; L^2(\Omega))}}{\log k_l - \log k_{l-1}}
\end{aligned} \tag{4.22}$$

that are obtained in the above situation when  $\bar{u}$  is approximated by means of the discretization scheme (4.4) on meshes with temporal and spatial widths  $k_l$ ,  $h_l$ ,  $l \in \mathbb{N}$ , can be seen in Tables 1 and 2. Here, we have considered standard Friedrichs-Keller triangulations of the spatial domain  $\Omega = (0, 1)^2$  and equidistant partitions of the time interval  $(0, 1)$ . The discrete problems (4.4) have been solved by reformulating the necessary and sufficient optimality condition (4.5) as a system of equations by means of the complementarity function  $\Phi(a, b) := \min(a, b)$ , and by subsequently applying a semismooth Newton method with tolerance  $10^{-10}$ . For an alternative to this solution approach, we refer to [31], where an ADMM-scheme for state-constrained parabolic optimal control problems is developed. We would like to point out that the calculation of the  $y_D$ -integral in (4.5) was carried out with a subdivided three-point Gauss-rule in space and time in our numerical experiments (*i.e.*, overall twelve nodes per triangle and 6 nodes per time interval). This choice was made to ensure that the discontinuous desired state  $y_D$  is integrated properly for all spatial/temporal meshes with  $h^{-1} = 2l$ ,  $k^{-1} = 2m$ ,  $l, m \in \mathbb{N}$ , even in those cases where these meshes do not resolve the boundary of the active set  $(1/4, 3/4)^3$  accurately. The same integration procedure has been used for the calculation of the  $L^2((0, T) \times \Omega)$ -errors in (4.22).

TABLE 2. EOCs in  $L^2(0, T; L^2(\Omega))$  for sequences of meshes that do not resolve the boundary of the active set  $(1/4, 3/4)^3$  in space and time. In (A), the spatial and the temporal widths are coupled, and in (B) and (C) one of the widths is fixed.

(A) $k_l = 0.5h_l^2$			(B) $k_l \equiv 1/1026$		(C) $h_l \equiv 1/66$	
$h_l$	$(\text{EOC}^h)_l$	$(\text{EOC}^k)_l$	$h_l$	$(\text{EOC}^h)_l$	$k_l$	$(\text{EOC}^k)_l$
$\frac{1}{6}$	–	–	$\frac{1}{6}$	–	$\frac{1}{10}$	–
$\frac{1}{10}$	1.1231	0.5615	$\frac{1}{10}$	1.0960	$\frac{1}{18}$	0.8993
$\frac{1}{14}$	1.2088	0.6044	$\frac{1}{14}$	1.1961	$\frac{1}{26}$	0.8865
$\frac{1}{18}$	1.0842	0.5421	$\frac{1}{18}$	1.0745	$\frac{1}{34}$	0.8645
$\frac{1}{22}$	1.0316	0.5158	$\frac{1}{22}$	1.0244	$\frac{1}{42}$	0.8308

As the results in Tables 1 and 2 show, the orders of convergence observed in our numerical experiments depend strongly on how well the employed temporal and spatial meshes fit to the active set  $(1/4, 3/4)^3$  of the exact solution in (4.21). If we choose sequences of meshes that resolve the boundary  $\partial(1/4, 3/4)^3$  precisely, then it holds  $\|\bar{u} - \bar{u}_{k_l h_l}\|_{L^2(0, T; L^2(\Omega))} \approx \mathcal{O}(k_l + h_l^2)$  and the error in the optimal control decays twice as fast as predicted by (4.15); see Table 1. If, however, we consider sequences with  $h_l^{-1} = 2 + 4m_l$  and  $k_l^{-1} = 2 + 4n_l$  for some  $m_l, n_l \in \mathbb{N}$  and, by doing so, ensure that the boundary  $\partial(1/4, 3/4)^3$  always cuts through the middle of the mesh cells, then the rate of convergence is significantly reduced and we end up with a behavior which is roughly of the form  $\|\bar{u} - \bar{u}_{k_l h_l}\|_{L^2(0, T; L^2(\Omega))} \approx \mathcal{O}(k_l^{0.85} + h_l)$ ; see Table 2. Note that the order of convergence observed in  $h$  here is precisely that derived in Theorem 4.8 while the rate in  $k$  is slightly higher than in our *a priori* error estimate. The reason for this faster convergence w.r.t.  $k$  is most likely that the optimal control and the adjoint state in our test case enjoy far more temporal regularity than predicted by (4.20). (It seems to be difficult to construct a situation where the conditions in Assumption 4.1 are satisfied and the control  $\bar{u}$  possesses, *e.g.*, only  $BV$ - or  $H^{1/2}$ -regularity in time.) Similar observations have also been made, for instance, in the numerical experiments conducted in ([32], Sect. 5) and ([42], Sect. 8). We would like to point out that the dependence of the convergence behavior on the resolution of the active set in Tables 1 and 2 offers a possible explanation for the observation that the order of the error  $\|\bar{u} - \bar{u}_{kh}\|_{L^2(0, T; L^2(\Omega))}$  often exceeds  $\mathcal{O}(k^{1/2} + h)$  in numerical experiments for which the analytical solution is not known. If a problem of the type (4.1) with random data is considered, then it has to be expected that the used meshes resolve the boundary of the (in this case unknown) active set relatively well in some parts of the domain and not at all in others and that, overall, a mixture of the two behaviors in Tables 1 and 2 is present. Compare again with the experiments in ([32], Sect. 5) in this context and also with Table 3 which shows the orders of convergence that are obtained when the solution of a problem of the type (4.1) with  $\Omega := (0, 1)^2$ ,  $T := 1$ ,  $y_0 := 0$ ,  $\nu := 10^{-5}$ ,  $\psi := -0.5$ , and  $y_D(t, x_1, x_2) := -\sin(\pi x_1) \sin(\pi x_2)t$  is approximated with the same discretization as in our first example and the solution on a fine mesh is used as a substitute for the optimal control  $\bar{u}$  in (4.22). We remark that the effects that we observe here also closely resemble those documented in [16].

We conclude this paper with some remarks on the results of the last sections and possible extensions of our analysis:

**Remark 4.10.** – It is important to realize that – although we know that the optimal control  $\bar{u}$  is in  $L^\infty((0, T) \times \Omega)$  in the situation of (4.1) – we do not have any *a priori* bound for the  $L^\infty((0, T) \times \Omega)$ -norm of the discrete controls  $\bar{u}_{kh}$ . If such a bound was available, one could introduce artificial control constraints and proceed along the lines of [32] to derive an estimate analogous to (4.15) with improved logarithmic factors. However, at least to the best of the authors' knowledge, there is currently no way to exclude that the  $L^\infty((0, T) \times \Omega)$ -norm of  $\bar{u}_{kh}$  blows up, *e.g.*, logarithmically in  $k$  and  $h$ .



TABLE 3. EOCs in  $L^2(0, T; L^2(\Omega))$  obtained for the problem (4.1) with  $\Omega := (0, 1)^2$ ,  $T := 1$ ,  $y_0 := 0$ ,  $\nu := 10^{-5}$ ,  $\psi := -0.5$ , and  $y_D(t, x_1, x_2) := -\sin(\pi x_1) \sin(\pi x_2)t$ . The discretization was the same as in our first numerical experiment and a solution on a fine mesh was used as a substitute for the analytical solution  $\bar{u}$  in (4.22). In (A), the spatial and the temporal widths are coupled, and in (B) and (C) one of the widths is fixed.

(A) $k_l = 9h_l^2$			(A) $k_l \equiv 1/64$		(A) $h_l \equiv 1/48$	
$h_l$	$(\text{EOC}^h)_l$	$(\text{EOC}^k)_l$	$h_l$	$(\text{EOC}^h)_l$	$k_l$	$(\text{EOC}^k)_l$
$\frac{1}{3}$	–	–	$\frac{1}{3}$	–	$\frac{1}{4}$	–
$\frac{1}{6}$	0.8038	0.4019	$\frac{1}{6}$	0.7280	$\frac{1}{8}$	0.9165
$\frac{1}{12}$	1.3331	0.6665	$\frac{1}{12}$	1.2880	$\frac{1}{16}$	0.9795
$\frac{1}{24}$	1.5396	0.7698	$\frac{1}{24}$	1.5152	$\frac{1}{32}$	1.0819

- Note that introducing the interpolant  $\mathcal{I}_{kh}$  in Lemma 4.7 is essential for the proof of (4.15) as this operator ensures that the discrete auxiliary problem (4.10) with right-hand side  $\bar{\lambda}$  is sensible. Without  $\mathcal{I}_{kh}$ , we would end up with a dual pairing between elements of  $V_{kh}$  and  $\mathcal{M}([0, T] \times \text{cl}(\Omega))$  at this point which is not well-defined due to the discontinuity in time of the functions in  $V_{kh}$ . A similar effect cannot be observed in the elliptic case, see ([12], Sect. 3).
- Using the same techniques as in the proof of Theorem 4.8, it is also possible to derive finite element error estimates for problems involving additional box constraints on the control and/or general second-order partial differential operators of the type (2.1). (Note that, in the presence of control constraints, some care has to be taken to ensure that the adjoint state  $\bar{p}$  is still an element of  $L^\infty((0, T) \times \Omega)$  and that additional assumptions may be necessary to ensure that an operator  $A$  of the form (2.1) satisfies a discrete maximal parabolic regularity estimate analogous to that in ([40], Thm. 11).) We omit a detailed discussion of this topic to keep this paper concise.
- We expect that the regularity results established in Section 3 can also be extended to optimal control problems governed by semilinear parabolic partial differential equations as studied, *e.g.*, in [48]. However, such a generalization requires carefully redoing the analysis in [4] which, in its current form, does not allow for additional non-linearities. We leave this topic for future research.

## APPENDIX A. TRUNCATION OPERATIONS IN SOBOLEV–BOCHNER SPACES

In what follows, we collect and prove several auxiliary results on pointwise properties of Sobolev-Bochner functions that are needed for our analysis. For convenience, we state the lemmas in this section in a general format, *i.e.*, we do not tacitly assume here that the quantities  $\Omega$  etc. satisfy the conditions in Assumption 2.1.

**Lemma A.1** (Stampacchia’s lemma for Sobolev-Bochner functions). *Suppose that a  $T > 0$ , a domain  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , and a  $v \in L^2(0, T; H^1(\Omega))$  are given. Define  $v^+ := \max(0, v)$ , where  $\max(0, \cdot)$  acts pointwise a.e. in  $(0, T) \times \Omega$ . Then the function  $v^+$  is an element of  $L^2(0, T; H^1(\Omega))$ , the gradient  $\nabla(v^+) \in L^2(0, T; L^2(\Omega, \mathbb{R}^d)) \cong L^2((0, T) \times \Omega; \mathbb{R}^d)$  of  $v^+$  satisfies*

$$\nabla(v^+) = \begin{cases} \nabla v & \mathcal{L}^{d+1}\text{-a.e. in } \{v > 0\} \\ 0 & \mathcal{L}^{d+1}\text{-a.e. in } \{v \leq 0\} \end{cases}, \quad (\text{A.1})$$

and we have

$$\nabla v = 0 \quad \mathcal{L}^{d+1}\text{-a.e. in } \{v = 0\}. \quad (\text{A.2})$$

If, further, the function  $v$  additionally possesses  $H^1(0, T; L^2(\Omega))$ -regularity, then  $v^+$  is also an element of  $H^1(0, T; L^2(\Omega))$  and it holds

$$\partial_t(v^+) = \begin{cases} \partial_t v & \mathcal{L}^{d+1}\text{-a.e. in } \{v > 0\} \\ 0 & \mathcal{L}^{d+1}\text{-a.e. in } \{v \leq 0\} \end{cases}, \quad (\text{A.3})$$

as well as

$$\partial_t v = 0 \quad \mathcal{L}^{d+1}\text{-a.e. in } \{v = 0\}, \quad (\text{A.4})$$

and if  $v$  is even in  $L^2(0, T; H^2(\Omega))$ , then the Hessian  $\nabla^2 v \in L^2(0, T; L^2(\Omega, \mathbb{R}^{d \times d})) \cong L^2((0, T) \times \Omega; \mathbb{R}^{d \times d})$  of  $v$  satisfies

$$\nabla^2 v = 0 \quad \mathcal{L}^{d+1}\text{-a.e. in } \{v = 0\}. \quad (\text{A.5})$$

*Proof.* The  $L^2(0, T; H^1(\Omega))$ -regularity of  $v^+$  follows straightforwardly from the results in ([3], Sect. 5.8) and ([35], Sect. 3), the formula (A.1) can be established completely analogously to ([3], Thm. 5.8.2), and to obtain (A.2), it suffices to note that (A.1) and the linearity of the operator  $\nabla$  yield

$$0 = \nabla v - \nabla(v^+ + v^-) = \nabla v - \nabla(v^+) - \nabla(v^-) = \begin{cases} \nabla v & \mathcal{L}^{d+1}\text{-a.e. in } \{v = 0\} \\ 0 & \mathcal{L}^{d+1}\text{-a.e. in } \{v \neq 0\} \end{cases},$$

where  $v^-$  is again short for  $\min(0, v) = -\max(0, -v)$ . This proves the first part of the lemma. Let us assume now that  $v$  is also an element of  $H^1(0, T; L^2(\Omega))$ . Then the  $H^1(0, T; L^2(\Omega))$ -regularity of  $v^+$  and the formula (A.3) follow from ([55], Cor. 2.3, Eq. (2)), and the derivation of (A.4) is completely along the lines of that of (A.2). It remains to establish (A.5). To this end, we first note that, for every  $v \in L^2(0, T; H^2(\Omega))$ , we have (due to (A.2) and since  $\partial_n v \in L^2(0, T; H^1(\Omega))$  holds for all spatial partial derivatives  $\partial_n v$ ,  $n = 1, \dots, d$ )

$$(\partial_n v) \mathbb{1}_{\{v=0\}} = 0 \in L^2((0, T) \times \Omega) \quad \forall n = 1, \dots, d$$

and

$$(\partial_m \partial_n v) \mathbb{1}_{\{\partial_n v=0\}} = 0 \in L^2((0, T) \times \Omega) \quad \forall m, n = 1, \dots, d.$$

The above implies in particular that

$$\mathbb{1}_{\{v=0\}} = \mathbb{1}_{\{\partial_n v=0\}} \mathbb{1}_{\{v=0\}} \in L^2((0, T) \times \Omega) \quad \forall n = 1, \dots, d$$

and, as a consequence, that

$$(\partial_m \partial_n v) \mathbb{1}_{\{v=0\}} = (\partial_m \partial_n v) \mathbb{1}_{\{\partial_n v=0\}} \mathbb{1}_{\{v=0\}} = 0 \in L^2((0, T) \times \Omega) \quad \forall m, n = 1, \dots, d.$$

This establishes (A.5) and completes the proof.  $\square$

We are now in the position to state a criterion for pointwise-a.e. boundedness in the space-time cylinder  $(0, T) \times \Omega$  that extends the results of ([38], Chap. II, Appendix B) to the parabolic setting:

**Lemma A.2** (Pointwise bounds via growth estimates). *Suppose that a  $T > 0$ , a bounded Lipschitz domain  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , and a  $v \in L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$  are given such that the truncations*

$$v_l := \min(0, v + l) + \max(0, v - l) = \begin{cases} v - l & \text{a.e. in } \{v \geq l\} \\ 0 & \text{a.e. in } \{|v| < l\} \\ v + l & \text{a.e. in } \{v \leq -l\} \end{cases}, \quad l \geq 0, \quad (\text{A.6})$$

satisfy

$$\|v_l\|_{L^\infty(0, T; L^2(\Omega))}^2 + \|v_l\|_{L^2(0, T; H^1(\Omega))}^2 \leq \int_0^T \int_\Omega |f| |v_l| dx dt \quad \forall l \geq L \quad (\text{A.7})$$

for some  $f \in L^\infty(0, T; L^q(\Omega))$ ,  $q > \max(d/2, 1)$ , and  $L \geq 0$ . Then there exists a constant  $C > 0$  depending only on  $\Omega$ ,  $T$ ,  $q$  and  $d$  such that

$$\|v\|_{L^\infty((0, T) \times \Omega)} \leq L + C \|f\|_{L^\infty(0, T; L^q(\Omega))}. \quad (\text{A.8})$$

*Proof.* To establish (A.8), we proceed along the lines of ([38], Lem. II.B2) (cf. also ([21], Lem. 3.6)). Let us first assume that  $d > 2$  and define

$$\lambda_l: (0, T) \rightarrow \mathbb{R}, \quad \lambda_l(t) := \mathcal{L}^d(\{x \in \Omega : |v(t, x)| \geq l\}), \quad l \geq 0.$$

Then the function  $\lambda_l$  is well-defined for each representative of  $v \in L^2(0, T; L^2(\Omega))$  and all  $l \geq 0$ , and the upper semicontinuity of the map

$$L^2(\Omega) \ni z \mapsto \mathcal{L}^d(\{x \in \Omega : |z(x)| \geq l\}) \in \mathbb{R}$$

for all  $l \geq 0$  and the Pettis measurability theorem, see ([35], Cor. 3.1.2), imply that  $\lambda_l$  is measurable for all  $l \geq 0$  and can be identified with an element of  $L^\infty(0, T)$ . From the Sobolev embedding  $H^1(\Omega) \hookrightarrow L^{2d/(d-2)}(\Omega)$ , the definition (A.6) and Lemma A.1, we now obtain that there exists a constant  $C > 0$  depending only on  $\Omega$  and  $d$  such that, for all  $0 \leq l \leq m$ , we have

$$\begin{aligned} & \|v_l\|_{L^\infty(0, T; L^2(\Omega))}^2 + \|v_l\|_{L^2(0, T; H^1(\Omega))}^2 \\ & \geq \operatorname{ess\,sup}_{t \in (0, T)} \int_{\{|v(t, \cdot)| \geq l\}} (|v(t, x)| - l)^2 dx + C \int_0^T \left( \int_{\{|v(t, \cdot)| \geq l\}} (|v(t, x)| - l)^{\frac{2d}{d-2}} dx \right)^{\frac{d-2}{d}} dt \\ & \geq (m - l)^2 \left( \|\lambda_m\|_{L^\infty(0, T)} + C \int_0^T \lambda_m(t)^{\frac{d-2}{d}} dt \right). \end{aligned}$$

On the other hand, the  $L^\infty(0, T; L^q(\Omega))$ -regularity of the function  $f$ , the condition  $q > \max(d/2, 1)$ , again the embedding  $H^1(\Omega) \hookrightarrow L^{2d/(d-2)}(\Omega)$ , the inequalities of Hölder and Young, and the elementary estimates

$$\frac{q(d+2)}{2d} > \frac{d+2}{4} \geq 1 \quad \text{and} \quad s := 1 + \frac{4}{d} - \frac{2}{q} > 1$$

allow us to compute that

$$\begin{aligned}
\int_0^T \int_{\Omega} |f| |v_l| dx dt &\leq C \int_0^T \left( \int_{\{|v(t,\cdot)| \geq l\}} |f(t,x)|^{\frac{2d}{d+2}} dx \right)^{\frac{d+2}{2d}} \|v_l(t)\|_{H^1(\Omega)} dt \\
&\leq C \int_0^T \left( \left( \int_{\{|v(t,\cdot)| \geq l\}} |f(t,x)|^q dx \right)^{\frac{2d}{q(d+2)}} \lambda_l(t)^{1 - \frac{2d}{q(d+2)}} \right)^{\frac{d+2}{2d}} \|v_l(t)\|_{H^1(\Omega)} dt \\
&\leq C \int_0^T \|f(t)\|_{L^q(\Omega)} \lambda_l(t)^{\frac{d+2}{2d} - \frac{1}{q}} \|v_l(t)\|_{H^1(\Omega)} dt \\
&\leq C \|f\|_{L^\infty(0,T;L^q(\Omega))}^2 \int_0^T \lambda_l(t)^{\frac{d+2}{d} - \frac{2}{q}} dt + \frac{1}{2} \|v_l\|_{L^2(0,T;H^1(\Omega))}^2 \\
&\leq C \|f\|_{L^\infty(0,T;L^q(\Omega))}^2 \left( \left( \|\lambda_l\|_{L^\infty(0,T)}^{\frac{4}{d} - \frac{2}{q}} \right)^{1/s} \left( \int_0^T \lambda_l(t)^{\frac{d-2}{d}} dt \right)^{1/s} \right)^s + \frac{1}{2} \|v_l\|_{L^2(0,T;H^1(\Omega))}^2
\end{aligned}$$

holds for all  $l \geq 0$ , where  $C > 0$  is a generic constant which depends only on  $d$ ,  $q$  and  $\Omega$  and which may change from step to step. Applying Young's inequality once again on the right-hand side of the last estimate yields

$$\begin{aligned}
\int_0^T \int_{\Omega} |f| |v_l| dx dt &\leq C \|f\|_{L^\infty(0,T;L^q(\Omega))}^2 \left( \left( \|\lambda_l\|_{L^\infty(0,T)}^{\frac{4}{d} - \frac{2}{q}} \right)^{1/(s-1)} + \int_0^T \lambda_l(t)^{\frac{d-2}{d}} dt \right)^s + \frac{1}{2} \|v_l\|_{L^2(0,T;H^1(\Omega))}^2 \\
&= C \|f\|_{L^\infty(0,T;L^q(\Omega))}^2 \left( \|\lambda_l\|_{L^\infty(0,T)} + \int_0^T \lambda_l(t)^{\frac{d-2}{d}} dt \right)^s + \frac{1}{2} \|v_l\|_{L^2(0,T;H^1(\Omega))}^2.
\end{aligned}$$

From (A.7), we may now deduce that there exists a constant  $C = C(q, d, \Omega) > 0$  and an exponent  $s > 1$  such that the (trivially non-negative and non-increasing) function

$$\Lambda: [0, \infty) \rightarrow [0, \infty), \quad \Lambda(l) := \|\lambda_l\|_{L^\infty(0,T)} + \int_0^T \lambda_l(t)^{\frac{d-2}{d}} dt,$$

satisfies

$$\Lambda(m) \leq C(m-l)^{-2} \|f\|_{L^\infty(0,T;L^q(\Omega))}^2 \Lambda(l)^s \quad \forall L \leq l \leq m.$$

The above implies, in combination with ([38], Lem. IIB.1), that there exists a (different) constant  $C = C(q, d, \Omega, T) > 0$  with  $\Lambda(L + C\|f\|_{L^\infty(0,T;L^q(\Omega))}) = 0$ . The estimate (A.8) now follows immediately from the definitions of  $\Lambda$  and  $\lambda_l$ . This proves the claim in the case  $d > 2$ . For  $d \in \{1, 2\}$ , the proof of (A.8) is completely along the same lines and only requires minor modifications due to the degenerate Sobolev conjugate. We leave the details to the reader.  $\square$

As an immediate consequence of Lemma A.2, we obtain the following result that allows to establish the  $L^\infty$ -regularity of the adjoint state  $\bar{p}$  in Theorem 3.4.

**Lemma A.3** (Pointwise estimates for parabolic PDEs). *Suppose that  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , is a bounded Lipschitz domain and that a  $T > 0$  is given. Assume further that  $a_{mn}, a_0 \in L^\infty(\Omega)$ ,  $m, n = 1, \dots, d$ , are functions which satisfy the conditions in (2.2) a.e. in  $\Omega$  for all  $m, n = 1, \dots, d$  and all  $\xi \in \mathbb{R}^d$  with some  $\alpha > 0$ , that  $\chi$  is a*

non-negative element of  $L^\infty((0, T) \times \Omega)$ , and that  $A: H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$  is defined as in (2.1). Then the PDE

$$\begin{aligned} \partial_t v + Av + \chi v &= f && \text{in } (0, T) \times \Omega, \\ \text{tr}(v) &= 0 && \text{on } (0, T) \times \partial\Omega, \\ v(0) &= g && \text{in } \Omega, \end{aligned} \tag{A.9}$$

admits a unique (weak) solution  $v = v(f, g) \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$  for all  $f \in L^2(0, T; H^{-1}(\Omega))$  and all  $g \in L^2(\Omega)$ , and, for every  $q > \max(d/2, 1)$ , there exists a constant  $C > 0$ , which depends only on  $\Omega$ ,  $T$ ,  $d$ ,  $q$  and  $\alpha$ , such that

$$\|v(f, g)\|_{L^\infty((0, T) \times \Omega)} \leq \|g\|_{L^\infty(\Omega)} + C\|f\|_{L^\infty(0, T; L^q(\Omega))} \tag{A.10}$$

holds for all  $f \in L^\infty(0, T; L^q(\Omega))$  and all  $g \in L^\infty(\Omega)$ .

*Proof.* The existence of a unique solution  $v \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$  of the problem (A.9) for all  $f \in L^2(0, T; H^{-1}(\Omega))$  and all  $g \in L^2(\Omega)$  follows straightforwardly from ([4], Thm. 1.14). It remains to prove (A.10). To this end, let us suppose that arbitrary but fixed  $f \in L^\infty(0, T; L^q(\Omega))$  and  $g \in L^\infty(\Omega)$  with  $q > \max(d/2, 1)$  are given, denote with  $v = v(f, g)$  the solution of (A.9) associated with  $f$  and  $g$ , and define  $v_l$ ,  $l \geq 0$ , as in (A.6). Then the embedding  $L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \hookrightarrow C([0, T]; L^2(\Omega))$  and Lemma A.1 yield that  $v_l \in L^2(0, T; H_0^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$  holds for all  $l \geq 0$ , and we may use (A.9), the formula (A.1), and exactly the same regularization argument as in the proof of ([55], Lem. 3.3) to compute that

$$\begin{aligned} \int_0^s \int_\Omega f v_l dx dt &= \int_0^s \left( \langle \partial_t v, v_l \rangle_{H_0^1(\Omega)} + \int_\Omega \sum_{m, n=1}^d a_{mn} (\partial_m v) (\partial_n v_l) + (a_0 + \chi) v v_l dx \right) dt \\ &\geq \int_0^s \left( \langle \partial_t v, v_l \rangle_{H_0^1(\Omega)} + \int_\Omega \sum_{m, n=1}^d a_{mn} (\partial_m v_l) (\partial_n v_l) dx \right) dt \\ &\geq \frac{1}{2} \|v_l(s)\|_{L^2(\Omega)}^2 - \frac{1}{2} \|v_l(0)\|_{L^2(\Omega)}^2 + \alpha \int_0^s \int_\Omega |\nabla v_l|^2 dx dt \quad \forall s \in (0, T). \end{aligned}$$

The above implies, in tandem with Friedrichs' inequality, the fact that  $v_l(0) = \min(0, g + l) + \max(0, g - l) = 0$  holds for all  $l \geq \|g\|_{L^\infty(\Omega)}$  and trivial estimates, that there exists a constant  $C = C(\Omega, \alpha) > 0$  with

$$\|v_l\|_{L^\infty(0, T; L^2(\Omega))}^2 + \|v_l\|_{L^2(0, T; H^1(\Omega))}^2 \leq \int_0^T \int_\Omega C|f||v_l| dx dt \quad \forall l \geq \|g\|_{L^\infty(\Omega)}.$$

The claim now follows immediately from Lemma A.2. This completes the proof.  $\square$

Note that the crucial point of Lemma A.3 is that the constant  $C$  in (A.10) is independent of  $\chi$ . This is what makes it possible to obtain the inclusion (3.14) in Theorem 3.4 by passing to the limit in (3.15).

We remark that results similar to that in Lemma A.3 can also be found, *e.g.*, in ([54], Chap. 5), albeit under different regularity assumptions on the right-hand side and for different boundary conditions. Further, we would like to point out that, completely analogously to the proof of Lemma A.3, one can also establish  $L^\infty$ -estimates for other types of parabolic PDEs and even certain evolution variational inequalities. Compare, *e.g.*, with the results for elliptic problems in ([17], Sect. 2), ([21], Lem. 3.6) and ([38], Sect. IIB) in this context. We do not state these generalizations here since they are not needed for the analysis of the problem (P).

## REFERENCES

- [1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, Oxford & New York (2000).
- [2] M.S. Aronna, J.F. Bonnans and A. Kröner, State-constrained control-affine parabolic problems I: first and second order necessary optimality conditions. Preprint [arXiv:1906.00237](https://arxiv.org/abs/1906.00237) (2019).
- [3] H. Attouch, G. Buttazzo and G. Michaille, *Variational Analysis in Sobolev and BV Spaces*. SIAM, Philadelphia (2006).
- [4] V. Barbu, *Optimal Control of Variational Inequalities*. Research Notes in Mathematics. Pitman (1984).
- [5] M. Bergounioux, Optimal control of parabolic problems with state constraints: a penalization method for optimality conditions. *Appl. Math. Optim.* **29** (1994) 285–307.
- [6] M. Bergounioux and K. Kunisch, Primal-dual strategy for state-constrained optimal control problems. *Comput. Optim. Appl.* **22** (2002) 193–224.
- [7] L. Bonifacius, *Numerical Analysis of Parabolic Time-optimal Control Problems*. Ph.D. thesis, Technische Universität München (2018).
- [8] J. Bonnans and P. Jaisson, Optimal control of a parabolic equation with time-dependent state constraints. *SIAM J. Control Optim.* **48** (2010) 4550–4571.
- [9] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research. Springer, New York (2000).
- [10] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd edn. Springer (2008).
- [11] E. Casas, Pontryagin’s principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.* **35** (1997) 1297–1327.
- [12] E. Casas, M. Mateos and B. Vexler, New regularity results and improved error estimates for optimal control problems with state constraints. *ESAIM: COCV* **20** (2014) 803–822.
- [13] E. Casas and F. Tröltzsch, Recent advances in the analysis of pointwise state-constrained elliptic optimal control problems. *ESAIM Control Optim. Calc. Var.* **16** (2010) 581–600.
- [14] C. Christof, *Sensitivity Analysis of Elliptic Variational Inequalities of the First and the Second Kind*. Ph.D. thesis, Technische Universität Dortmund (2018).
- [15] C. Christof, Sensitivity analysis and optimal control of obstacle-type evolution variational inequalities. *SIAM J. Control Optim.* **57** (2019) 192–218.
- [16] C. Christof and C. Meyer, A note on a priori  $L^p$ -error estimates for the obstacle problem. *Numer. Math.* **139** (2018) 27–45.
- [17] C. Christof and C. Meyer, Sensitivity analysis for a class of  $H_0^1$ -elliptic variational inequalities of the second kind. *Set-Valued Var. Anal.* **27** (2018) 469–502.
- [18] C. Christof and G. Wachsmuth, On second-order optimality conditions for optimal control problems governed by the obstacle problem. To appear in: *Optimization* (2020) 1–41. <https://doi.org/10.1080/02331934.2020.1778686>.
- [19] F.H. Clarke, *Optimization and Nonsmooth Analysis*. SIAM’s Classics in Applied Mathematics. SIAM, Philadelphia, PA (1990).
- [20] J.C. De Los Reyes, P. Merino, J. Rehberg and F. Tröltzsch, Optimality conditions for state-constrained PDE control problems with time-dependent controls. *Control Cybernet.* **37** (2008) 5–38.
- [21] J.C. De Los Reyes and C. Meyer, Strong stationarity conditions for a class of optimization problems governed by variational inequalities of the second kind. *J. Optim. Theory Appl.* **168** (2016) 375–409.
- [22] K. Deckelnick and M. Hinze, Variational discretization of parabolic control problems in the presence of pointwise state constraints. *J. Comput. Math.* **29** (2011) 1–15.
- [23] K. Disser, A.F.M. ter Elst and J. Rehberg, Hölder estimates for parabolic operators on domains with rough boundary. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **XVII** (2017) 65–79.
- [24] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*. North-Holland (1976).
- [25] J. Elschner, J. Rehberg and G. Schmidt, Optimal regularity for elliptic transmission problems including  $C^1$ -interfaces. *Interfaces Free Bound.* **9** (2007) 233–252.
- [26] K. Eriksson, C. Johnson and V. Thomée, Time discretization of parabolic problems by the discontinuous Galerkin method. *ESAIM Math. Model. Numer. Anal.* **19** (1985) 611–643.
- [27] L.C. Evans, *Partial Differential Equations*, 2nd edn. AMS, Providence, RI (2010).
- [28] L.A. Fernández, *State Constrained Optimal Control for Some Quasilinear Parabolic Equations*, edited by K.-H. Hoffmann, G. Leugering, F. Tröltzsch and S. Caesar. *Optimal Control of Partial Differential Equations*. Birkhäuser, Basel, (1999) 145–156.
- [29] L.A. Fernández, Integral state-constrained optimal control problems for some quasilinear parabolic equations. *J. Nonlinear Anal. Optim.* **39** (2000) 977–996.
- [30] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, reprint of the 1998 edn. Springer (2001).
- [31] R. Glowinski, Y. Song and X. Yuan, An ADMM numerical approach to linear parabolic state constrained optimal control problems. *Numer. Math.* **144** (2020) 931–966.
- [32] W. Gong and M. Hinze, Error estimates for parabolic optimal control problems with control and state constraints. *Comput. Optim. Appl.* **56** (2013) 131–151.
- [33] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*. Pitman (1985).
- [34] F. Harder and G. Wachsmuth, Comparison of optimality systems for the optimal control of the obstacle problem. *GAMM-Mitt.* **40** (2018) 312–338.

- [35] J. Heinonen, P. Koselka, N. Shanmugalingam and J.T. Tyson, Sobolev Spaces on Metric Measure Spaces. Vol. 27 of *New Mathematical Monographs*. Cambridge University Press (2015).
- [36] M. Hinze. A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.* **30** (2005) 45–61.
- [37] K. Ito and K. Kunisch, Semi-smooth Newton methods for state-constrained optimal control problems. *Systems Control Lett.* **50** (2003) 221–228.
- [38] D. Kinderlehrer and G. Stampacchia, An Introduction to Variational Inequalities and Their Applications. Vol. 31 of *Classics in Applied Mathematics*. SIAM (2000).
- [39] D. Leykekhman and B. Vexler, Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. *SIAM J. Numer. Anal.* **54** (2016) 1365–1384.
- [40] D. Leykekhman and B. Vexler, Discrete maximal parabolic regularity for Galerkin finite element methods. *Numer. Math.* **135** (2017) 923–952.
- [41] F. Ludovici, I. Neitzel and W. Wollner, A priori error estimates for state-constrained semilinear parabolic optimal control problems. *J. Optim. Theory Appl.* **178** (2018) 317–348.
- [42] D. Meidner, R. Rannacher and B. Vexler, A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time. *SIAM J. Control Optim.* **49** (2011) 1961–1997.
- [43] D. Meidner and B. Vexler, Optimal error estimates for fully discrete Galerkin approximations of semilinear parabolic equations. *ESAIM: M2AN* **52** (2018) 2307–2325.
- [44] C. Meyer, A. Rösch and F. Tröltzsch, Optimal control of PDEs with regularized pointwise state constraints. *Comput. Optim. Appl.* **33** (2006) 209–228.
- [45] F. Mignot, Contrôle dans les inéquations variationelles elliptiques. *J. Funct. Anal.* **22** (1976) 130–185.
- [46] B.S. Mordukhovich and K. Zhang, Optimal control of state-constrained parabolic systems with nonregular boundary controllers, in *Proceedings of the 36th IEEE Conference on Decision and Control*, Vol. 1 (1997) 527–528.
- [47] J.J. Moreau, P.D. Panagiotopoulos and G. Strang, *Topics in Nonsmooth Mechanics*. Birkhäuser, Basel (1988).
- [48] I. Neitzel and F. Tröltzsch, On convergence of regularization methods for nonlinear parabolic optimal control problems with control and state constraints. *Control Cybernet.* **37** (2008) 1013–1043.
- [49] I. Neitzel and F. Tröltzsch, On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints. *ESAIM: COCV* **15** (2009) 426–453.
- [50] I. Neitzel and F. Tröltzsch, Numerical analysis of state-constrained optimal control problems for PDEs, edited by G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich and S. Ulbrich *Constrained Optimization and Optimal Control for Partial Differential Equations*. Springer, Basel (2012) 467–482.
- [51] A.H. Schatz and L.B. Wahlbin, Interior maximum-norm estimates for finite element methods, part II. *Math. Comput.* **64** (1995) 907–928.
- [52] A. Schiela, State constrained optimal control problems with states of low regularity. *SIAM J. Control Optim.* **48** (2009) 2407–2432.
- [53] B. Schweizer, *Partielle Differentialgleichungen*. Springer, Berlin/Heidelberg (2013).
- [54] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen*. Vieweg und Teubner, Wiesbaden, 2nd edn. (2009).
- [55] D. Wachsmuth, The regularity of the positive part of functions in  $L^2(I; H^1(\Omega)) \cap H^1(I; H^1(\Omega)^*)$  with applications to parabolic equations. *Comment. Math. Univ. Carolin.* **57** (2016) 327–332.
- [56] G. Wachsmuth, Towards M-stationarity for optimal control of the obstacle problem with control constraints. *SIAM J. Control Optim.* **54** (2016) 964–986.