Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Endenicher Allee 19b • 53115 Bonn • Germany
phone +49 228 73-69828 • fax +49 228 73-69847
www.ins.uni-bonn.de

F. Hoppe, I. Neitzel

Optimal control of quasilinear parabolic PDEs with state constraints

INS Preprint No. 2004

December 2020

Optimal control of quasilinear parabolic PDEs with state constraints

Fabian Hoppe and Ira Neitzel

Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Endenicher Allee 19b, 53115 Bonn (Germany)

Abstract. We discuss state-constrained optimal control of a quasilinear parabolic PDE. Existence of optimal controls and first-order necessary optimality conditions are derived for a rather general setting including pointwise in time and space constraints on the state. Second-order sufficient optimality conditions are obtained for averaged-in-time and pointwise in space state-constraints under general regularity assumptions for the equation, and for pointwise in time and space state-constraints when restricting in return to a more regular setting for the state equation.

1. Introduction

This paper is concerned with optimal control of a quasilinear parabolic partial differential equation (PDE) with pointwise control-constraints, and additional constraints on the state variable. In particular, we prove existence of optimal controls and derive first-order necessary optimality conditions under rather general assumptions on the state equation and pointwise in time and space state-constraints. Under additional assumptions we provide second-order sufficient optimality conditions (SSCs). For the rather general assumptions on the state equation we restrict the analysis to averaged-in-time state-constraints. Pointwise in time and space state-constraints are discussed for a more regular state equation and purely timedependent controls.

Optimal control of linear and nonlinear PDEs has been subject to active research for several years, see e.g. [**28**] or the more recent monographs [**22, 40**]. PDE-constrained optimization problems with pointwise constraints on the state are particularly challenging, since one usually has to guarantee continuity of the state function, because a typical Slater-type constraint-qualification can only be fulfilled in this setting. This unfortunately results in rather low regularity of the corresponding Lagrange multipliers, see for instance [**6, 7**].

If nonlinear PDEs are considered, first-order necessary optimality conditions are in general not sufficient due to nonconvexity of the optimal control problem. Analysis of second-order sufficient conditions therefor plays an important role in optimal control. We refer e.g. to the survey [**12**] and references therein for an overview covering different aspects of the topic, as well as the to our knowledge first contribution to SSCs for PDE-constrained optimization, [**19**], and point out only a few more particular aspects. A typical difficulty arising in second-order optimality conditions in PDE-constrained optimization is the so-called two-norm

Acknowledgments: Partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 211504053 - SFB 1060.

discrepancy, see e.g. [**11**, **24**]: Differentiability of the reduced functional and coercivity of its second derivative do often only hold with respect to different, non-equivalent, norms. Since this dictates the function spaces for all further analysis, avoiding the two-norm discrepancy is generally desirable. In any case, a careful regularity analysis of the underlying PDE is necessary, and often leads to restrictions on e.g. the spatial dimension in particular for parabolic problems [**9**] or purely timedependent controls [**15**]. In [**27**] the authors obtain SSCs for semilinear parabolic PDEs in space dimension 2 and 3 and distributed control via a careful regularity analysis utilizing the concept of maximal parabolic regularity. Regarding research concerned with second-order necessary optimality conditions for pure state-constraints we only mention [**25**], as well as both second-order necessary and sufficient conditions with emphasis on a possibly small gap between them in [**26**, **38**] for the different setting of pointwise mixed control-state-constraints. Finally, we mention the recent publications [**14**, **41**, **42**] which also aim at minimizing the gap between necessary and sufficient second-order conditions by utilizing an abstract optimization-theoretic approach.

For results on optimal control problems governed by quasilinear PDEs, we restrict the introduction to the parabolic setting. Early results include the works [**17**, **18**], where existence of optimal controls and first-order necessary optimality conditions are shown for a problem with averaged-in-space and pointwise in time, or finitely many state-constraints of integral-type, respectively, yet under somewhat more restrictive regularity assumptions than in the present paper. Rather recently, well-posedness of the state equation and existence of optimal controls under rather general regularity assumptions on domain and coefficients has been proven in [**34**]. First and second-order optimality conditions have been derived in [**4**]. Based on this, convergence of the SQP method applied to the respective optimization problem has been shown in [**23**]. Optimality conditions for a similar model problem with slightly more regular coefficients and domain, but in return unbounded nonlinearities, have been analyzed in [**8**]. Optimal control of the so-called thermistor problem, a coupled system consisting of a quasilinear parabolic and a nonlinear elliptic equation, is addressed in [**32**, **33**]. We also refer to the references mentioned in the introduction of the mentioned papers for a more detailed overview on the history of this area. The non-trivial existence and regularity theory for solutions of the underlying PDEs poses the main difficulty in the analysis of such problems.

This paper contributes to the theory of nonlinear optimal control problems in several ways. First, we establish existence of optimal controls and first-order necessary in the presence of state-constraints, extending the results from [**4**] and [**8**] without state-constraints. In particular, note that the regularity assumptions from [**4**, **34**] are fairly general, and include certain types of nonsmooth domains, mixed boundary conditions and nonsmooth coefficients. Furthermore, to our best knowledge, second-order sufficient optimality conditions for state-constrained optimal control of a quasilinear parabolic PDE have not been addressed in the existing literature. We present a detailed analysis, restricting our setting to either averaged-in-time state-constraints when keeping the general regularity setting of [**4**], or to the strengthened regularity assumptions of [**8**] and purely timedependent controls when considering pointwise in time and space state-constraints. Extending the the work of Casas and Tröltzsch [**11**] towards the inclusion of state-constraints we pay particular attention on avoiding the two-norm gap.

The structure of the paper is the following: In Section 2 we introduce the model problem and state the corresponding assumptions. Further, we apply well-known techniques to prove existence of optimal controls and derive first-order necessary optimality conditions in this rather general context. Section 3 is devoted to the proof of SSCs to an abstract optimization problem extending the result from [**11**] but avoiding the two-norm discrepancy. The topic of Section 4 is twofold: First, we explain why our abstract result from Section 3 does not apply to the model problem as stated in Section 2. Second, we prove our first main result, i.e. SSCs avoiding the two-norm gap for a modified version of our model problem where the regularity assumptions remain unchanged but the pointwise state-constraints are replaced by averaged-in-time state-constraints. In Section 5 we come back to the consideration of pointwise state-constraints and prove SSCs for this situation, but now assuming a more regular setting for the state equation along the lines of [**8**] and purely timedependent controls. In the final section we shortly comment on how results change in case that constraints on the control are omitted.

Notation. Let us clarify some notation used throughout the paper: Given an interval $I \subset \mathbb{R}$ and a domain $\Omega \subset \mathbb{R}^d$ we abbreviate the corresponding space time cylinder by $Q := I \times \Omega$. Regarding Hölder-, (Bochner-)Lebesgue- and (Bochner-)-Sobolev-spaces we follow [**4**] and apply standard notation. Since the spatial domain Ω does not change during the paper we will omit it, i.e. we write $W^{1,p}$ instead of $W^{1,p}(\Omega)$ for the $W^{1,p}$-regular functions on Ω, but e.g. $C(\overline{Q})$ for the continuous functions on $\overline{Q}$. For real and complex interpolation spaces, respectively, we use the standard notation, see for instance [**3, 39**].

The domain of a densely defined operator $A\colon X \to Y$ between Banach spaces X, Y is denoted by $\mathrm{Dom}_X(A)$, and equipped with the canonical graph norm. For a closed convex subset $K \subset X$ in a Banach space X and some $x \in K$ we denote by $\mathcal{R}(K, x)$ the radial cone, and by $T_C(x)$ the contingent cone, see e.g. [**5**] for the definition.

2. Existence of optimal controls and first-order necessary optimality conditions

In this section we introduce the model problem from [**4**], yet with additional pointwise state-constraints, state our assumptions, and collect some results from [**4**] on the control-to-state map. Finally, following the standard techniques, we derive existence of optimal controls and first-order necessary optimality conditions.

2.1. Model Problem and Assumptions.

The optimal control problem under consideration reads as follows:

$$\text{(OCP)} \qquad \begin{cases} \min\limits_{y,u} J(y,u) := \dfrac{1}{2}\|y - y_d\|^2_{L^2(I\times\Omega)} + \dfrac{\gamma}{2}\|u\|^2_{L^2(\Lambda)}, \\ \\ \text{subject to} \quad u \in U_{ad}, \quad y \in Y_{ad}, \quad \text{and} \quad \text{(SE)}, \end{cases}$$

with the quasilinear parabolic state equation (SE) given by

$$\text{(SE)} \qquad \partial_t y + \mathcal{A}(y)y = Bu \qquad \text{on } Q, \quad y(0) = y_0 \qquad \text{on } \Omega.$$

Here, the quasilinear differential operator $\mathcal{A}$ is defined as

$$\mathcal{A}(y) := -\nabla \cdot \xi(y)\mu\nabla.$$

The control space $L^s(\Lambda)$, the control operator B, as well as the set of admissible controls $U_{ad} \subset L^s(\Lambda)$ will be introduced precisely in the following. The set of admissible states will be clarified in each section, and will be given either by pointwise in space and time inequality-constraints, i.e.

$$Y_{ad} = \left\{ y \in C(\overline{Q}) \colon \quad y_a(t,x) \leq y(t,x) \leq y_b(t,x) \quad \forall (t,x) \in \overline{Q} \right\},$$

or, if we require a weaker type of constraints for our analysis, by pointwise in space and averaged-in-time bounds of type

$$Y_{ad} = \left\{ y \in L^1(I, C(\overline{\Omega})) \colon \quad y_a(x) \leq \int_0^T y(t,x)dt \leq y_b(x) \quad \forall x \in \overline{\Omega} \right\}.$$

The assumptions required for the analysis of the state equation are close to [**4**], but we forego those parts that refer to the improved regularity analysis from [**4**] on Bessel-potential spaces and stick to the setting of [**34**]:

Assumption 2.1. 1. $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is a bounded domain with boundary $\partial\Omega$. $\Gamma_N \subset \partial\Omega$ is relatively open and denotes the Neumann boundary part whereas $\Gamma_D = \partial\Omega \setminus \Gamma_N$ denotes the Dirichlet boundary part. We assume that $\Omega \cup \Gamma_N$ is Gröger regular [**20**] such that every chart map in the definition of Gröger regularity can be chosen volume-preserving. The time interval $I = (0,T)$ with $T > 0$ is fixed.

2. The function $\xi \colon \mathbb{R} \to \mathbb{R}$ is twice differentiable with ξ'' being Lipschitz continuous on bounded subsets of $\mathbb{R}$. Let $\mu \colon \Omega \to \mathbb{R}^{d\times d}$ be measurable and uniformly bounded and coercive in the following sense:

$$0 < \mu_\bullet := \inf_{x\in\Omega} \inf_{z\in\mathbb{R}^d\setminus\{0\}} \frac{z^T\mu(x)z}{z^Tz}, \qquad \mu^\bullet := \sup_{x\in\Omega} \sup_{1\leq i,j\leq d} |\mu_{i,j}(x)| < \infty$$

We assume a coercivity condition $0 < \xi_\bullet \leq \xi \leq \xi^\bullet$ for ξ as well. With this we define as above

$$\langle \mathcal{A}(y)\varphi, \psi \rangle_{L^2(I,W_D^{1,2})} := \int_I \int_\Omega \xi(y)\mu\nabla\varphi\nabla\psi \, dx dt, \qquad \varphi,\psi \in L^2(I, W_D^{1,2}).$$

3. We assume that there is $p \in (d,4)$ such that

$$-\nabla \cdot \mu\nabla + 1 \colon W_D^{1,p} \to W_D^{-1,p}$$

is a topological isomorphism and fix this choice of p.

4. Let $s > 2$ be fixed such that $\frac{1}{s} < \frac{1}{2}\left(1 - \frac{d}{p}\right)$ holds. For a measure space (Λ, ρ) we define the control space $U := L^s(\Lambda)$ and the admissible set

$$U_{ad} = \{u \in L^s(\Lambda) \colon u_a(x) \leq u(x) \leq u_b(x) \quad \text{for a.a. } x \in \Lambda\}$$

with $u_a, u_b \in L^\infty(\Lambda)$, $u_a \leq u_b$ almost everywhere. The control operator

$$B \colon U = L^s(\Lambda) \to L^s(I, W_D^{-1,p})$$

is bounded linear and admits a bounded linear extension

$$B \colon L^2(\Lambda) \to L^2(I, W_D^{-1,p}).$$

Finally, the initial condition $y_0 \in (W_D^{-1,p}, W_D^{1,p})_{1-1/s,s}$ and the desired state $y_d \in L^\infty(I, L^2)$ are fixed.

The constants p and s are fixed from now on. Note that in part (4), we only assume B to be continuous from $L^s(\Lambda)$ to $L^s(I, W_D^{-1,p})$, instead from $L^\infty(\Lambda)$ to $L^s(I, W_D^{-1,p})$ as in [**4**]). This does not destroy applicability of the assumption to the full range of situations described in [**4**, section 2.2]: Distributed control ($U = L^s(I \times \Omega)$), Neumann boundary control ($U = L^s(I \times \Gamma_N)$) in dimension $d = 2$, and purely timedependent control ($U = L^s(I, \mathbb{R}^m)$) with m fixed control functions from $W_D^{-1,p}$ in dimensions $d = 2, 3$. We would like to mention that Theorem 2.1 covers several real-world constellations with certain nonsmooth μ, nonsmooth Ω, or mixed boundary conditions [**16**]. For more details concerning applicability of theorem 2.1 (1) and (3) we refer to [**4**, Remarks 2.1 and 2.3].

Note that in [**4**] some additional assumptions are stated that are only necessary for improved regularity analysis on Bessel-potential spaces, but not for the results regarding first- and second-order analysis therein which can be obtained completely within the $W_D^{-1,p}$-$W_D^{1,p}$-setting described in Theorem 2.1. Since we will only rely on these results, we decided not to include the improved regularity assumptions of [**4**].

2.2. Control-to-state map and reduced functional. For convenience of the reader, we recall some auxiliary results from [**4**] that we will refer to lateron. Due to [**4**, Proposition 3.5] (see also [**34**, Corollary 5.8]) the solution map of the equation

$$\partial_t y + \mathcal{A}(y)y = v, \qquad y(0) = y_0, \tag{1}$$

defined by $y := G(v)$ if and only if (1), is well-defined as map G: $L^s(I, W_D^{-1,p}) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$. By composition with B we obtain the control-to-state map

$$S\colon L^s(\Lambda) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p}), \qquad u \mapsto G(Bu).$$

From [**4**] we recall the following notation for the derivatives of the nonlinear term:

$$\mathcal{A}'(y)v := -\nabla \cdot (\xi'(y) v \mu \nabla y),$$
$$\mathcal{A}''(y)[v_1, v_2] := -\nabla \cdot (\xi'(y)(v_1 \mu \nabla v_2 + v_2 \mu \nabla v_1) + \xi''(y) v_1 v_2 \mu \nabla y)$$

for $v, v_1, v_2 \in W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p})$, $r \in (1, \infty)$ and a measurable function y on Q. The following differentiability properties of G hold true [**4**, Proposition 4.4 and Lemma 4.5]:

Lemma 2.2. *Let Theorem 2.1 be satisfied.*

(1) *The map G: $L^s(I, W_D^{-1,p}) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$ is twice continuously Fréchet differentiable with derivatives $G'(v)h = w$ and $G''(v)[h_1, h_2] = z$ given by the unique solutions of*

$$\partial_t w + \mathcal{A}(y)w + \mathcal{A}'(y)w = h, \qquad w(0) = 0, \tag{2}$$

$$\partial_t z + \mathcal{A}(y)z + \mathcal{A}'(y)z = \mathcal{A}''(y)[G'(v)h_1, G'(u)h_2], \qquad z(0) = 0 \tag{3}$$

for $y = G(v)$, respectively.

(2) *The nonautonomous operator $\mathcal{A}(y) + \mathcal{A}'(y)$ exhibits maximal parabolic regularity on $L^r(I, W_D^{-1,p})$ for $r \in (1, s]$. In particular it holds*

$$G'(v) \in \mathcal{L}(L^r(I, W_D^{-1,p}), W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}))$$

for all $v \in L^s(I, W_D^{-1,p})$, $r \in (1, s]$.

The well-definedness of the control-to-state map allows to introduce the reduced objective functional

$$j\colon L^s(\Lambda) \to \mathbb{R}, \qquad u \mapsto J(S(u), u).$$

From [**4**, Lemma 4.6] we recall the fact that the reduced functional j is twice continuously Fréchet differentiable on $L^s(\Lambda)$ with gradient

$$\nabla j(u) = B^* G'(Bu)^*(y - y_d) + \gamma u \in L^2(\Lambda).$$

By [**4**, Proposition 4.7], the term $G'(Bu)^*(y - y_d)$ can be interpreted as solution of a nonautonomous backward parabolic PDE, the so-called adjoint equation. This, however, requires slightly sharper regularity assumptions on the state equation than ours, cf. [**4**, Assumption 4]. We will come back to this issue lateron when discussing first-order necessary optimality conditions in Sections 2.4 and 4.1.

2.3. Existence of optimal controls. Let us now consider the following setting of the state-constraints, until noted otherwise when discussing SSC:

Assumption 2.3. 1. The set of admissible states $Y_{ad} \subset C(\overline{Q})$ is given by

$$Y_{ad} = \{y \in C(\overline{Q}) : y_a(t,x) \leq y(t,x) \leq y_b(t,x) \quad \forall (t,x) \in \overline{Q}\},$$

with bounds $y_a, y_b \in C(\overline{Q})$ satisfying $y_a(t,x) < y_b(t,x)$ for all $(t,x) \in \overline{Q}$. We allow for unilateral bounds, i.e. $y_a \equiv -\infty$ or $y_b \equiv \infty$.

2. There is a feasible point, i.e. there is $(y,u) \in Y_{ad} \times U_{ad}$ such that y and u fulfill the state equation (SE).

Together with Assumption 2.1 we can prove existence of a minimizer for (OCP) as it has already been done for the case without state-constraints in [**4**, Lemma 4.1] resp. [**34**, Proposition 6.4]. An analogous result for the state-constrained thermistor problem has already been obtained in [**32**].

Theorem 2.4. *Let Theorems 2.1 and 2.3 hold. Then there exists a globally optimal control $\bar{u} \in U_{ad}$ for the optimal control problem* (OCP)*.*

Proof. The proof follows standard arguments in the calculus of variations, cf. [**4, 34**], [**22, 40**]. In particular, note that existence of an infimizing sequence is provided by the existence of a feasible point (Theorem 2.3 (2)). In the proof of [**34**, Proposition 6.4] it is shown that a subsequence of the corresponding sequence of states converges in $C(\overline{Q})$ to the optimal state. Since Y_{ad} is closed in $C(\overline{Q})$ the limit is still in Y_{ad}, i.e. it fulfills the state-constraints. □

For comments on the question how Theorem 2.4 and our subsequent results change in the case without or with only unilateral control-constraints we refer the interested reader to Section 6 at the end of this paper.

2.4. First-order optimality conditions. Having established existence of a global solution to (OCP) in the previous subsection, we now characterize local solutions of (OCP) that fulfill a Slater-type constraint-qualification by first-order necessary optimality conditions. The main difficulty in this section is the following issue that is well-known in state-constrained optimal control of PDEs, cf. e.g. [**6, 7**]: In order to ensure that the Slater-type condition can be fulfilled, we have to choose the underlying spaces in such a way that Y_{ad} has nonempty interior. Since Y_{ad} is defined by pointwise inequality-constraints, this excludes to consider states in $L^q(\bar{Q})$, $1 \leq q < +\infty$; instead we have to consider them in the space $C(\overline{Q})$, which results in the appearance regular Borel measures, i.e. the corresponding dual objects, in the KKT-system. Therefore, the adjoint state exhibits rather poor regularity in general.

In order to apply an abstract result on KKT-conditions for optimization problems posed in Banach spaces [**7**, Theorem 5.2] to our optimal control problem (OCP) we have to formulate an additional assumption first.

Assumption 2.5. Let $\bar{u} \in U_{ad}$ be an $L^2(\Lambda)$-local solution to (OCP) with associated state $\bar{y} = S(\bar{u}) \in Y_{ad}$, i.e. there is $\epsilon > 0$ such that

$$j(u) \geq j(\bar{u}) \qquad \forall u \in \mathbb{B}_\epsilon^{L^2(\Lambda)}(\bar{u}) \cap U_{ad} \quad \text{s.t.} \quad S(u) \in Y_{ad},$$

holds. Further, assume that the following linearized Slater-condition is fulfilled at $\bar{u}$: There is $u_{\mathrm{Sl}} \in U_{ad}$ such that

$$\text{(SL)} \qquad \bar{y} + S'(\bar{u})(u_{\mathrm{Sl}} - \bar{u}) \in \mathring{Y}_{ad},$$
$$\text{i.e. } y_a(t,x) < \bar{y}(t,x) + S'(\bar{u})(u_{\mathrm{Sl}} - \bar{u})(t,x) < y_b(t,x) \quad \forall (t,x) \in \overline{Q}.$$

Note that since the L^s-norm is stronger than the L^2-norm, any $L^2(\Lambda)$-local solution is in particular a $L^s(\Lambda)$-local solution.

Theorem 2.6. *Under Theorems 2.1 and 2.5 and Theorem 2.3 (1) there exists a regular Borel measure $\bar{\mu} \in \mathcal{M}(\overline{Q}) = C(\overline{Q})^*$ on $\overline{Q}$ and the so-called adjoint state $\bar{p} \in L^{r'}(I, W_D^{1,p'})$, $r' \in (1, \frac{2p}{p+d})$, such that the optimality system*

$$\text{(4)} \qquad \begin{aligned} -\partial_t \bar{p} + \mathcal{A}(\bar{y})^* \bar{p} + \mathcal{A}'(\bar{y})^* \bar{p} &= \bar{y} - y_d + \bar{\mu}, \\ \bar{p}(T) &= 0, \end{aligned}$$

$$\text{(5)} \qquad \langle \bar{\mu}, y - \bar{y} \rangle_{\mathcal{M}(\overline{Q}), C(\overline{Q})} \leq 0 \qquad \text{for all } y \in Y_{ad},$$

$$\text{(6)} \qquad \langle B^* \bar{p} + \gamma \bar{u}, u - \bar{u} \rangle_{L^{s'}(\Lambda), L^s(\Lambda)} \geq 0 \qquad \text{for all } u \in U_{ad}$$

is satisfied. The so-called adjoint equation (4) *has to be understood in the weak sense outlined in the proof below.*

Proof. In [**7**, Theorem 5.2] choose $U = L^s(\Lambda)$, $Z = C(\overline{Q})$, $J = j$, $G = S$, $K = U_{ad}$ and $C = Y_{ad}$. Note that the embedding $W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p}) \hookrightarrow C(\overline{Q})$ [**4**, Proposition 3.3] ensures that the control-to-state operator maps $L^s(\Lambda)$ into $C(\overline{Q})$. Utilizing the formula for the gradient of the reduced functional we obtain (6) from [**7**, (5.3)] by introducing $\bar{p} := G'(B\bar{u})^*(\bar{y} - y_d + \bar{\mu})$; equation (4) has to be understood in this sense. Hereby, the notation as backward parabolic PDE is motivated by [**4**, Proposition 4.7]. Regarding the regularity of $\bar{p}$, note that

$$G'(B\bar{u}) \in \mathcal{L}\left(L^r(I, W_D^{-1,p}), W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p})\right)$$

for any $r \in (1, s]$, see Theorem 2.2 (2), together with the embedding $W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}) \hookrightarrow C(\overline{Q})$ for $r \in (\frac{2p}{p-d}, \infty)$ [**4**, Proposition 3.3] implies

$$G'(B\bar{u})^* \in \mathcal{L}\left(\mathcal{M}(\overline{Q}), L^{r'}(I, W_D^{1,p'})\right)$$

for all $r' \in (1, \frac{2p}{p+d})$. The claim follows from $\bar{y} - y_d + \bar{\mu} \in \mathcal{M}(\overline{Q})$. □

Remark 2.7. Condition (5) can be rewritten in a more illustrative way: The Jordan decomposition $\bar{\mu} = \bar{\mu}^+ - \bar{\mu}^-$ into non-negative measures $\bar{\mu}^+, \bar{\mu}^- \geq 0$ satisfies

$$\operatorname{supp} \bar{\mu}^+ \subset \{(t,x) \in \overline{Q}: \quad \bar{y}(t,x) = y_b(t,x)\},$$
$$\operatorname{supp} \bar{\mu}^- \subset \{(t,x) \in \overline{Q}: \quad \bar{y}(t,x) = y_a(t,x)\}.$$

For a proof we refer e.g. to [**10**, Proposition 2.5].

Remark 2.8. Because $\bar{\mu}$ is, in general, only a Borel measure, we cannot improve regularity of the adjoint state $\bar{p}$ along the lines of [**4**, Proposition 4.7] using the improved regularity analysis of the state equation on Bessel-potential spaces. However, we mention that improved regularity for adjoint states in state-constrained optimal control has been obtained under additional assumptions and with different techniques in case of linear and semilinear elliptic [**10**] and parabolic [**13**] PDEs.

We conclude this section by noting that due to $\frac{2p}{p+d} < 2$ and $p' < 2$ Theorem 2.6 shows rather poor temporal and spatial regularity for the adjoint state $\bar{p}$. This is one of the typical difficulties that have to be overcome during the analysis of second-order optimality conditions for (OCP), as we will outline in Sections 4 and 5. In particular, as already stated in the introduction, we will either have to modify the type of state-constraints (Theorem 4.1) or assume a more regular setting for the state equation (Theorem 5.1) in order to achieve our goal.

3. An abstract result on second-order sufficient conditions

In this section we extend the abstract framework of [**11**] towards inclusion of state-constraints, i.e. we give second-order sufficient conditions (SSCs) for an abstract optimization problem similar to the one from [**7**, Theorem 5.2], but now enriched with the presence of two norms as typical for PDE-constrained optimization. However, we prove second-order sufficient conditions that avoid the two-norm gap. The framework is developed having in particular the setting and the arguments from [**15**] in mind.

We start by introducing the abstract optimization problem

$$\text{(P)} \qquad \min j(u) \qquad \text{s.t. } u \in K, \quad g(u) \in C,$$

with the detailed assumptions given below. Note that the suppositions on the functional j and the underlying spaces U_2, U_∞, respectively, will be identical to those from [**11**]. In the present paper we extend this work towards the inclusion of a state-constraint-like constraint of type "$g(u) \in C$". Since the set $K \cap g^{-1}(C)$ will be non-convex in general, this situation is not covered by the results of [**11**]. Further, note that j and g are differentiable with respect to the U_∞-norm, but not necessarily with respect to the weaker U_2-norm. Of course, we have in mind the case $U_2 = L^2(\Sigma, m)$ and $U_\infty = L^p(\Sigma, m)$ with some $p \in (2, \infty]$ for some measure space $(\Sigma, \mathrm{d}m)$. The presence of such two norms goes back to [**24**], and is typical for PDE-constrained optimization, see e.g. the exposition in [**11, 12, 40**].

Before going into the details of the assumption, let us briefly put the result into context: To our best knowledge, prior results on second-order sufficient conditions in the context of state-constraints without two-norm gap required differentiability of j and g with respect to L^2, cf. [**38**, Section 4], [**9**, Theorem 4.3] - an assumption that can be avoided in our result. In particular we are able to state SSCs for the same semilinear parabolic optimal control problem as in [**15**], but without norm gap, see

Theorem 3.3 below. In [**14**] both second-order necessary, and sufficient conditions for certain optimization problems in infinite dimensions are proven. The results rely on the concept of a directional curvature functional for the (possibly non-convex) admissible set. The authors state that it is possible to include cases with two-norm discrepancy (see Remark 4.6(iv)), but the special case of the present paper and [**11**], in which such a discrepancy appears but can be avoided in the formulation of second-order conditions, is not addressed. Further, the explicit computation of the directional curvature term in the presence of pointwise state-constraints is left as topic of further research. We believe that our approach, that is explicitly tailored to situations as e.g. (OCP), [**15**], and [**11**], respectively, on the other hand, is of independent interest.

Assumption 3.1. Let U_2 be a Hilbert space and U_∞ a Banach space such that there is a continuous embedding $U_\infty \hookrightarrow U_2$. With $\|\cdot\|_\infty$, $\|\cdot\|_2$ resp. $\langle\cdot,\cdot\rangle_2$ we denote the corresponding norms resp. the U_2-scalar product. Further, let Z be a Banach space with norm $\|\cdot\|_Z$ and duality pairing $\langle\cdot,\cdot\rangle_{Z,Z^*}$.

1. Let $\emptyset \neq K \subset U_\infty$ be convex and $A \supset K$ be open in U_∞. We fix $\bar{u} \in K$. The functional j: $A \to \mathbb{R}$ is assumed to be twice continuously Fréchet differentiable w.r.t. $\|\cdot\|_\infty$ and to fulfill the following properties:

1a. The derivatives of j taken with respect to the space U_∞ extend to continuous linear resp. bilinear forms on U_2, i.e.

$$j'(u) \in \mathcal{L}(U_2,\mathbb{R}) \quad \text{and} \qquad j''(u) \in \mathcal{L}(U_2 \otimes U_2,\mathbb{R}), \qquad u \in A.$$

1b. Let $(u_k)_k \subset K$, $(v_k)_k \subset U_2$ be arbitrary sequences such that $u_k \to \bar{u}$ strongly with respect to the U_2-norm and $v_k \rightharpoonup v$ weakly in U_2 as $k \to \infty$. Then it holds:

1bi. $j'(\bar{u})v = \lim_{k\to\infty} j'(u_k)v_k$
1bii. $j''(\bar{u})v^2 \le \liminf_{k\to\infty} j''(u_k)v_k^2$
1biii. If $v = 0$, there is some $\gamma > 0$ such that

$$\gamma \liminf_{k\to\infty} \|v_k\|_2^2 \le \liminf_{k\to\infty} j''(u_k)v_k^2.$$

2. Let g: $A \to Z$ be twice continuously Fréchet differentiable w.r.t. $\|\cdot\|_\infty$ such that the following properties hold true:

2a. The derivatives of g taken with respect to U_∞ extend to continuous linear resp. bilinear forms on U_2, i.e.

$$g'(u) \in \mathcal{L}(U_2, Z) \quad \text{and} \quad g''(u) \in \mathcal{L}(U_2 \otimes U_2, Z), \qquad u \in A.$$

2b. Let $(u_k)_k \subset K$, $(v_k)_k \subset U_2$ be arbitrary sequences such that $u_k \to \bar{u}$ strongly with respect to the U_2-norm and $v_k \rightharpoonup v$ weakly in U_2 as $k \to \infty$. Then it holds:

2bi. $g'(u_k)v_k \rightharpoonup g'(\bar{u})v$ weakly in Z
2bii. $g''(u_k)v_k^2 \rightharpoonup g''(\bar{u})v^2$ weakly in Z

The following result is –on the abstract level– the main result of this paper, and extends [**11**, Theorem 2.3] towards the inclusion of a state-constraint-like constraint of type "$g(u) \in C$".

Theorem 3.2. *Let Theorem 3.1 hold. Let $C \subset Z$ be a closed convex set and let $\bar{u} \in K$, $g(\bar{u}) \in C$, and $\bar{\mu} \in Z^*$ fulfill the following properties:*

$$\langle j'(\bar{u}) + g'(\bar{u})^* \bar{\mu}, u - \bar{u} \rangle_2 \geq 0 \qquad \forall u \in K, \tag{7}$$

$$\langle \bar{\mu}, z - g(\bar{u}) \rangle_{Z^*,Z} \leq 0 \qquad \forall z \in C, \tag{8}$$

i.e. the KKT-conditions for the problem (P). *Assume further that it holds*

$$j''(\bar{u})v^2 + \langle \bar{\mu}, g''(\bar{u})v^2 \rangle_{Z^*,Z} > 0 \qquad \forall v \in C_{\bar{u}} \setminus \{0\} \tag{9}$$

with

$$C_{\bar{u}} := \mathrm{cl}_{U_2}\left(\mathcal{R}(K, \bar{u})\right) \cap \{v \in U_2\colon \quad j'(\bar{u})v = 0, \quad \langle g'(\bar{u})^* \bar{\mu}, v \rangle_2 = 0, \quad g'(\bar{u})v \in T_C(g(\bar{u}))\}.$$

Then, there are $\epsilon, \delta > 0$ such that the quadratic growth condition

$$j(u) \geq j(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|_2^2$$

holds for all $u \in K$ that satisfy $\|u - \bar{u}\|_2 \leq \epsilon$ and $g(u) \in C$; in particular, $\bar{u}$ is an U_2-local minimizer for (P).

We follow the the proof of [**11**, Theorem 2.3], and abstract the techniques of several similar results in this context, see e.g. [**9**, **26**, **38**], and, in particular, [**15**].

Proof. Assume the contrary, i.e. there exist $(u_k)_k \subset K$ such that

$$\|u - u_k\|_2 < \frac{1}{k}, \qquad j(u_k) < j(\bar{u}) + \frac{1}{2k}\|u_k - \bar{u}\|_2^2 \qquad \text{and} \quad g(u_k) \in C.$$

Define $\rho_k := \|u_k - \bar{u}\|_2$ and $v_k := \frac{1}{\rho_k}(u_k - \bar{u})$. Since $(v_k)_k \subset U_2$ is bounded by definition and U_2 is a Hilbert space we can w.l.o.g. assume that $v_k \rightharpoonup v$ with some $v \in U_2$. We prove $v \in C_{\bar{u}}$ in four steps (A)-(D):

A. From weak convergence and (7) we derive immediately:

$$\begin{aligned}\langle j'(\bar{u}) + g'(\bar{u})^* \bar{\mu}, v \rangle_2 &= \lim_{k \to \infty} \langle j'(\bar{u}) + g'(\bar{u})^* \bar{\mu}, v_k \rangle_2 \\ &= \lim_{k \to \infty} \frac{1}{\rho_k} \langle j'(\bar{u}) + g'(\bar{u})^* \bar{\mu}, u_k - \bar{u} \rangle_2 \geq 0.\end{aligned}$$

B. To show $\langle g'(\bar{u})\bar{\mu}, v \rangle_2 \leq 0$ observe that

$$\langle \bar{\mu}, g'(u_k^\theta)v_k \rangle_2 = \frac{1}{\rho_k}\langle \bar{\mu}, g'(u_k^\theta)(u_k - \bar{u}) \rangle_2 = \frac{1}{\rho_k}\langle \bar{\mu}, g(u_k) - g(\bar{u}) \rangle_2 \overset{(8)}{\leq} 0$$

with some $u_k^\theta := \theta_k u_k + (1 - \theta_k)\bar{u}$, $(\theta_k)_k \subset [0, 1]$ originating from the mean value theorem. Utilizing Theorem 3.1 (2bi) we obtain:

$$\langle g'(\bar{u})^* \bar{\mu}, v \rangle_2 = \lim_{k \to \infty} \langle \bar{\mu}, g'(u_k^\theta)v_k \rangle_2 \leq 0.$$

Similarly we obtain for arbitrary but fixed $\eta \in Z^*$

$$\langle \eta, \frac{1}{\rho_k}(g(u_k) - g(\bar{u})) \rangle_{Z^*,Z} = \langle \eta, g'(u_k^{\theta,\eta})v_k \rangle \to \langle \eta, g'(\bar{u})v \rangle$$

due to Theorem 3.1 (2bi), i.e. $g'(\bar{u})v \in \text{weak-cl}_Z(\mathcal{R}(C, g(\bar{u}))) = T_C(g(\bar{u}))$, since C is assumed to be closed and convex: From [**5**, Proposition 2.55] we infer that $T_C(g(\bar{u})) = \mathrm{cl}_Z(\mathcal{R}(C, g(\bar{u})))$. The radial cone $\mathcal{R}(C, g(\bar{u}))$ is convex due to convexity of C, and hence its (strong) closure in Z is equal to its weak closure weak-$\mathrm{cl}_Z(\mathcal{R}(C, g(\bar{u})))$, see [**5**, Theorem 2.23 (ii)] for instance.

C. As in the proof of [**11**, Theorem 2.3] we find with help of the mean value theorem that $j'(\bar{u})v \leq 0$ holds. Together with (B) we obtain $\langle j'(\bar{u}) + g'(\bar{u})^*\bar{\mu}, v\rangle_2 \leq 0$ and therefore with (A):

$$\langle j'(\bar{u}) + g'(\bar{u})^*\bar{\mu}, v\rangle_2 = 0$$

D. Now, by (B) we have $j'(u)v = -\langle g'(\bar{u})^*\bar{\mu}, v\rangle_2 \geq 0$, which implies together with $j'(\bar{u})v \leq 0$ that $j'(\bar{u})v = 0$. Finally it follows $\langle g'(\bar{u})^*\bar{\mu}, v\rangle_2 = 0$ by (C).

As in [**11**] one can show that $v \in \mathrm{cl}_{U_2}(\mathcal{R}(K, \bar{u}))$ and hence it follows from (A)-(D) that $v \in C_{\bar{u}}$. Now, using our assumption and Taylor expansion we find

$$\frac{\rho_k^2}{2k} > j(u_k) - j(\bar{u}) = j'(\bar{u})(u_k - \bar{u}) + \frac{1}{2} j''(u_k^\theta)(u_k - \bar{u})^2$$

with some $u_k^\theta = \theta_k u_k + (1-\theta_k)\bar{u}$, $(\theta_k) \subset [0,1]$. Exploiting (7) and (8) it follows

$$\begin{aligned}
\frac{\rho_k^2}{2k} &\overset{(7)}{>} -\langle \bar{\mu}, g'(\bar{u})(u_k - \bar{u})\rangle_{Z^*,Z} + \frac{1}{2} j''(u_k^\theta)(u_k - \bar{u})^2 \\
&= -\langle \bar{\mu}, g(u_k) - g(\bar{u})\rangle_{Z^*,Z} + \langle \bar{\mu}, g(u_k) - g(\bar{u}) - g'(\bar{u})(u_k - \bar{u})\rangle_{Z^*,Z} \\
&\qquad + \frac{1}{2} j''(u_k^\theta)(u_k - \bar{u})^2 \\
&\overset{(8)}{\geq} \langle \bar{\mu}, g(u_k) - g(\bar{u}) - g'(\bar{u})(u_k - \bar{u})\rangle_{Z^*,Z} + \frac{1}{2} j''(u_k^\theta)(u_k - \bar{u})^2 \\
&= \frac{1}{2}\rho_k^2 \left(\langle \bar{\mu}, g''(\tilde{u}_k^\theta)v_k^2\rangle + j''(u_k^\theta)v_k^2\right),
\end{aligned}$$

where we used the mean value theorem for the last equality with some $\tilde{u}_k^\theta = \tilde{\theta}_k u_k + (1-\tilde{\theta}_k)\bar{u} \in K$, $(\tilde{\theta}_k) \subset [0,1]$. From

$$\frac{1}{k} > \langle \bar{\mu}, g''(\tilde{u}_k^\theta)v_k^2\rangle_{Z^*,Z} + j''(u_k^\theta)v_k^2$$

and $u_k^\theta \to \bar{u}$, $\tilde{u}_k^\theta \to \bar{u}$ in U_2, $v_k \rightharpoonup v$ weakly in U_2 we find with Theorem 3.1 (1bii), (2bii):

$$j''(\bar{u})v^2 + \langle \bar{\mu}, g''(\bar{u})v^2\rangle_{Z^*,Z} = 0.$$

Since (9) and $v \in C_{\bar{u}}$ holds, we conclude $v = 0$. Using Theorem 3.1 (1biii) at (♣) and (2bii) at (★) we finally arrive at

$$0 < \gamma = \gamma \liminf_{k\to\infty} \|v_k\|_2^2 \overset{(\clubsuit)}{\leq} \liminf_{k\to} j''(u_k^\theta)v_k^2 \leq \liminf_{k\to\infty} \left(\frac{1}{k} - \langle \bar{\mu}, g''(\tilde{u}_k^\theta)v_k^2\rangle_{Z^*,Z}\right) \overset{(\bigstar)}{=} 0,$$

which is the desired contradiction. □

As a first application of Theorem 3.2 we briefly indicate how it allows to extend a result from the literature:

Example 3.3. The reader may easily verify along the lines of [**11, 15**] that the semilinear parabolic optimal control problem with pointwise constraints on the state from [**15**] fits into the framework of Theorem 3.1. Therefore, Theorem 3.2 allows to reformulate [**15**, Theorem 5] with L^∞- replaced by L^2-neighbourhoods.

Remark 3.4. Let us for a moment replace convergence $u_k \to \bar{u}$ in U_2 in Theorem 3.1 by the stronger convergence $u_k \to \bar{u}$ in V, where $(V, \|\cdot\|_V)$ is some

Banach space such that $V \hookrightarrow U_\infty$ and $K \subset V$. The proof of Theorem 3.2 still shows that a quadratic growth condition of type

$$j(u) \geq j(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|_2^2$$

holds, but now only for those $u \in K$ that fulfill $\|u - \bar{u}\|_V < \epsilon$ and $g(u) \in C$, i.e. there is a so-called two-norm-gap in the quadratic growth condition. Consequently, $\bar{u}$ can only be shown to be a V-local minimizer for (P), which corresponds –on the abstract level– for $V = U_\infty$ to the result of [**15**].

The following example, although of completely artificial nature, illustrates that the assumptions in the formulation of Theorem 3.2 are necessary. For examples concerned with the necessity of the assumptions on j we refer to [**11**]. Therefore we only concentrate on the assumptions on g.

Example 3.5. With $U_\infty = L^\infty([0,1])$, $U_2 = L^2([0,1])$, $Z = C([0,1])$ we consider the problem

$$\text{(EX)} \qquad \min_{u \in L^2([0,1])} j(u) := \int_0^1 u(t)^2 dt \qquad \text{s.t.} \qquad -1 \leq u(t) \leq 1 \quad \text{a.e. on } [0,1], \qquad [g(u)](t) \geq t \quad \text{on } [0,1],$$

with $[g(u)](t) := \int_0^t \left(1 - \cos\left(\frac{\pi}{2}u(s)\right)\right) ds$. First, note that j obviously satisfies Theorem 3.1. Further, observe that g: $L^2([0,1]) \to C([0,1])$ is well-defined. However, since the superposition operator associated to the cosine-function is known to be Fréchet differentiable on $L^\infty([0,1])$, but not on $L^2([0,1])$, we only have at hand twice Fréchet-differentiablity of g as map $L^\infty([0,1]) \to C([0,1])$.

One easily verifies that $\bar{u} \equiv 1$ is feasible for (EX), and also satisfies the first-order necessary conditions (7) and (8) with $\bar{\mu} = -\frac{2}{\pi}\delta_1 \in C([0,1])^*$. Hereby, we denoted by δ_1 the Dirac-measure concentrated at $t = 1$. Moreover, also the coercivity condition (9) is trivially satisfied at $(\bar{u}, \bar{\mu})$, because $C_{\bar{u},\bar{\mu}} = \{0\}$. Furthermore, the second derivative of the functional at $\bar{u}$ is even $L^2([0,1])$-coercive.

Nevertheless, any $u_n \in L^2([0,1])$ defined by $u_n(t) = -1$ for $t \in [0, \frac{1}{n}]$ and $u_n(t) = 1$, else, is also feasible for (EX) and satisfies $j(u_n) = j(\bar{u})$. Together with $u_n \to \bar{u}$ w.r.t. the $L^2([0,1])$-norm, this shown that a quadratic growth condition around $\bar{u}$ cannot hold. The reason for this behavior is the following: Theorem 3.2 cannot be applied, because Theorem 3.1 (2) fails to hold. Choose $v_n := n^{\frac{1}{2}}\mathbf{1}_{(0,\frac{1}{n})}$, then it holds $v_n \rightharpoonup 0$ weakly in $L^2([0,1])$, but for $\hat{u}_n := \frac{1}{2}(u_n + \bar{u})$ we obtain $\hat{u}_n \to \hat{u}$ strongly in $L^2([0,1])$ and

$$\langle \delta_1, g''(\hat{u}_n)v_n^2 \rangle = \frac{\pi^2}{4}\int_0^1 \cos\left(\frac{\pi}{2}\hat{u}_n(t)\right) v_n^2(t) dt = \frac{\pi^2}{4} \nrightarrow 0 = \langle \delta_1, g''(\bar{u})v^2 \rangle$$

which disproves Theorem 3.1 (2bii). However, note that due to continuous Fréchet differentiability of g w.r.t. $L^\infty([0,1])$, the previous remark applies: $\bar{u}$ is an $L^\infty([0,1])$-local, but not an $L^2([0,1])$-local solution of (EX).

We conclude this section by pointing out an open problem in this context. An important property of the SSCs in [**11**, Theorem 2.3] is their minimal gap to corresponding necessary optimality conditions, if the admissible set K is polyhedric: Positivity of $j''(\bar{u})$ on a cone $C_{\bar{u}} \subset U_2$ of certain directions is –together with first-order necessary conditions– a sufficient optimality condition for $\bar{u}$, while non-negativity of $j''(\bar{u})$ on the *same* cone is already necessarily implied by local optimality of $\bar{u}$ [**11**, Theorem 2.2]. Obtaining similar second-order necessary optimality conditions (SNCs) for (P) seems to be a challenging topic and is beyond the scope of the present paper. For recent results concerning no-gap second-order conditions we refer e.g. to [**26, 38**] in case of optimal control of semilinear elliptic PDEs with mixed control-state constraints, or to [**14**] for an abstract optimization-theoretic result with different applications to PDE-constrained optimization.

4. SSCs for averaged-in-time state-constraints

This section contains the first part of our discussion of second-order sufficient conditions for (OCP): We replace the pointwise in space and time state-constraints from by averaged-in-time state-constraints, see Theorem 4.1 below. For the resulting modified model problem we are able to prove second-order sufficient conditions avoiding the two-norm gap while keeping the relatively low regularity requirements on the state equation from Theorem 2.1. This is our first main result for our model problem. Since Theorem 2.1 on the state equation and the control operator still holds, our results apply to the full range of situations described in [**4**, Section 2.2], yet with additional state constraints.

Since part (1) of Theorem 3.1 referring to the unchanged state equation and the objective functional, has already been verified for (OCP) in [**4**, Section 4.3], the remaining work is to check Theorem 3.1 (2). This requires a careful regularity analysis for the first- and second-order derivatives of the control-to-state map. The results of this analysis also highlight the obstructions that prevent us from applying Theorem 3.2 under Theorems 2.1 and 2.3 directly, and therefore motivate the introduction of averaged-in-time state-constraints. In particular, note that in this aspect the analysis of the quasilinear problem (OCP) is quite different from the discussion of the semilinear problem mentioned in Theorem 3.3 due to the more complicated structure of derivatives of the nonlinearity in the differential operator. This fact allows to obtain slightly better regularity results in the case of semilinear equations.

4.1. Averaged-in-time state-constraints. We start by introducing our modified state-constraints and postpone their mathematical motivation to Section 4.2.

Assumption 4.1. 1. The set of admissible states $Y_{ad} \subset L^1(I, C(\overline{\Omega}))$ is given by

$$Y_{ad} = \{y \in L^1(I, C(\overline{\Omega})) : y_a(x) \leq \int_0^T y(t,x)dt \leq y_b(x) \quad \forall x \in \overline{\Omega}\},$$

with bounds $y_a, y_b \in C(\overline{\Omega})$ satisfying $y_a(x) < y_b(x)$ for all $x \in \overline{\Omega}$. We allow for unilateral bounds, i.e. $y_a \equiv -\infty$ or $y_b \equiv \infty$.

2. There is a feasible point, i.e. there is $(y,u) \in Y_{ad} \times U_{ad}$ such that y and u fulfill the state equation (SE).

Intuitively, this means, e.g. in the case of controling temperature, keeping the average temperature over the time interval at each point of an object in a certain desired range. Of course, in order to get closer to the original pointwise in time formulation, it is also possible to consider averaging on a finite number of subintervals of I separately. However, since the latter is only a technicality, we decided to keep the above assumption as simple as possible. Note that considering averaged-type constraints instead of purely pointwise constraints is common in the existing optimal control literature; see e.g. [**17**, **19**, **29**, **31**] for averaged-in-space and pointwise in time state-constraints, or [**30**] for averaged-in-space and pointwise in time bounds on the gradient of the state.

Existence of an optimal control for (OCP) with averaged-in-time state-constraints is proven analogously to Theorem 2.4. We only state the result.

Theorem 4.2. *Let Theorems 2.1 and 4.1 hold. Then there exists a globally optimal control $\bar{u} \in U_{ad}$ for the optimal control problem* (OCP).

In order to state first-order necessary optimality conditions, we first require a suitable constraint-qualification:

Assumption 4.3. Under Theorem 4.1 (1) let $\bar{u} \in U_{ad}$ be an $L^2(\Lambda)$-local solution to (OCP) with associated state $\bar{y} = S(\bar{u}) \in Y_{ad}$ such that that the following linearized Slater-condition is fulfilled at $\bar{u}$: There is $u_{\text{Sl}} \in U_{ad}$ such that

$$\text{(SL)} \qquad \bar{y} + S'(\bar{u})(u_{\text{Sl}} - \bar{u}) \in \mathring{Y}_{ad},$$
$$\text{i.e. } y_a(x) < \int_\Omega [\bar{y}(t,x) + S'(\bar{u})(u_{\text{Sl}} - \bar{u})(t,x)]\, dt < y_b(x) \quad \forall x \in \overline{\Omega}.$$

As in Section 2.4 the derivation of the following result is based on [**7**, Theorem 5.2].

Theorem 4.4. *Under Theorems 2.1 and 4.3 and Theorem 4.1 (1) there exists a regular Borel measure $\bar{\mu} \in \mathcal{M}(\overline{\Omega}) = C(\overline{\Omega})^*$ on $\overline{\Omega}$ and the adjoint state $\bar{p} \in L^{r'}(I, W_D^{1,p'})$, $r' \in (1,\infty)$, such that the optimality system*

$$\text{(10)} \qquad \begin{aligned} -\partial_t \bar{p} + \mathcal{A}(\bar{y})^* \bar{p} + \mathcal{A}'(\bar{y})^* \bar{p} &= \bar{y} - y_d + dt \otimes \bar{\mu}, \\ \bar{p}(T) &= 0, \end{aligned}$$

$$\text{(11)} \qquad \begin{aligned} \operatorname{supp}(\bar{\mu}^+) &\subset \left\{ x \in \overline{\Omega} \colon \int_0^T \bar{y}(t,x) dt = y_b(x) \right\}, \\ \operatorname{supp}(\bar{\mu}^-) &\subset \left\{ x \in \overline{\Omega} \colon \int_0^T \bar{y}(t,x) dt = y_a(x) \right\}, \end{aligned}$$

$$\text{(12)} \qquad \langle B^* \bar{p} + \gamma \bar{u}, u - \bar{u} \rangle_{L^{s'}(\Lambda), L^s(\Lambda)} \geq 0, \qquad \textit{for all } u \in U_{ad},$$

is satisfied. Here, $\bar{\mu} = \bar{\mu}^+ - \bar{\mu}^-$ denotes the Jordan-decomposition of $\bar{\mu}$, cf. Theorem 2.7, and the adjoint equation (10) *has to be understood in the weak sense.*

Because the proof is completely analogous to the proof of Theorem 2.6, we only comment on the differences with respect to the new type of state-constraints: In [**7**, Theorem 5.2] we now choose $Z = C(\overline{I})$ and $G := A \circ S$: , where S is the control-to-state map and $A \in \mathcal{L}(L^1(I, C(\overline{\Omega})), C(\overline{\Omega}))$ is defined by averaging w.r.t. time, i.e. $A\varphi := \left(x \mapsto \int_0^T \varphi(t,x) dt \right)$. A short computation shows that $A^* \bar{\mu} = dt \otimes \bar{\mu}$ and therefore $\nabla j(\bar{u}) + g'(\bar{u})^* \bar{\mu} = B^* \bar{p} + \gamma \bar{u}$. For the regularity of $\bar{p}$, note that

$$G'(B\bar{u}) \in \mathcal{L}\left(L^r(I, W_D^{-1,p}), L^1(I, C(\overline{\Omega})) \right)$$

holds for $r \in (1, \infty)$. Taking adjoints it follows that

$$G'(B\bar{u})^* A^* \in \mathcal{L}\left(\mathcal{M}(\overline{\Omega}), L^{r'}(I, W_D^{1,p'})\right),$$

which shows the claimed regularity for $\bar{p}$.

Remark 4.5. Let, in addition to Theorem 2.1, the enhanced regularity assumptions from [**4**, Assumption 4] hold that enable the improved regularity analysis of the state on Bessel-potential spaces from [**4**, Theorem 3.20]. Consequently, [**4**, Proposition 4.7] shows that $G'(B\bar{u})^*$ is the solution operator of the backward-parabolic PDE

$$-\partial_t z + \mathcal{A}(\bar{y})^* z + \mathcal{A}'(\bar{y})^* z = w, \qquad z(T) = 0.$$

Moreover, for any $r' \in [s', \infty)$ the map $w \mapsto z$ is bounded linear

$$L^{r'}(I, W^{-1,p'}) \to W^{1,r'}(I, W^{-1,p'}) \cap L^{r'}(I, W_D^{1,p'}).$$

Because of $p > d$ we have $dt \otimes \bar{\mu} \in L^\infty(I, \mathcal{M}(\overline{\Omega})) \hookrightarrow L^\infty(I, W_D^{-1,p'})$, and we obtain improved regularity

$$\bar{p} \in W^{1,r'}(I, W^{-1,p'}) \cap L^{r'}(I, W_D^{1,p'}), \qquad r' \in [s', \infty),$$

for the adjoint state from Theorem 4.4. Moreover, the adjoint equation (10) even holds in the distributional sense in the respective space.

4.2. Regularity of the derivatives of the control-to-state map. From Section 2.2 recall the definition of the control-to-state map and its differentiability properties stated in Theorem 2.2. In this subsection we carry out a more detailed analysis with respect to regularity, continuity and extension properties of the derivatives. Moreover, we use these results subsequently to motivate the introduction of the averaged-in-time state-constraints.

Proposition 4.6. *Let Theorem 2.1 hold.*

(1) *Fix $u \in L^s(\Lambda)$.*
  (a) *The first derivative $S'(u)$ of the control-to-state map extends to a continuous linear map from $L^2(\Lambda)$ to $L^q(I, C(\overline{\Omega}))$ for any $q \in \left(1, \frac{2p}{d}\right)$.*
  (b) *The second derivative $S''(u)$ extends to a continuous bilinear map from $L^2(\Lambda) \times L^2(\Lambda)$ to $W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p})$ for any $r \in \left(1, \frac{2p}{p+d}\right)$.*

(2) *Let $(u_k)_k \subset L^s(\Lambda)$ converge to $\bar{u}$ strongly in $L^s(\Lambda)$ and $(v_k)_k \subset L^2(\Lambda)$ converge to some v weakly in $L^2(\Lambda)$. Then it holds:*

$$S'(u_k)v_k \to S'(\bar{u})v_k \qquad \text{strongly in } L^q(I, C(\overline{\Omega})),$$
$$S''(u_k)v_k^2 \rightharpoonup S''(\bar{u})v^2 \qquad \text{weakly in } W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}),$$

for q and r as in part (1).

Proof. 1. From Theorem 2.2 we know that

$$S'(u) \in \mathcal{L}\left((L^2(\Lambda), W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})\right)$$

for any $u \in L^s(\Lambda)$. Therefore, (1a) follows from the embedding

$$W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p}) \hookrightarrow_c L^q(I, C(\overline{\Omega})), \qquad q \in \left(1, \frac{2p}{d}\right), \tag{13}$$

see e.g. [**4**, Proposition 3.3]. For later purpose let us already point out here that embedding (13) is compact. Consequently, $S'(u) \in \mathcal{L}(L^2(\Lambda), L^q(I, C(\overline{\Omega})))$ is compact as well.

Regarding (1b) it suffices due to Theorem 2.2 to show that for $r \in \left(1, \frac{2p}{p+d}\right)$

$$\|\mathcal{A}''(y)[w_1, w_2]\|_{L^r(I, W_D^{-1,p})} \leq c\|w_1\|_{W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})} \cdot \|w_2\|_{W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})}$$

holds for any $w_1, w_2 \in W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})$ and $y = S(u)$. This however follows from the definition of $\mathcal{A}''$ by straightforward application of Hölders inequality and (13).

2. In the proof of [**4**, Proposition 4.9] it has been shown that

$$G'(Bu_k) \to G'(B\bar{u}), \quad \text{in } \mathcal{L}\left(L^r(I, W^{-1,p}), W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p})\right), \tag{14}$$

as long as $r \leq \frac{2p}{p-d}$; see Section 2.4 for the meaning of G. In particular,

$$S'(u_k) \to S'(\bar{u}) \qquad \text{in } \mathcal{L}\left(L^2(\Lambda), W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})\right)$$

is true, from which we conclude the first statement of (2).

Regarding the second derivative, we write

$$\begin{aligned} S''(u_k)v_k^2 - S''(\bar{u})v^2 = {} & \left(G'(Bu_k) - G'(B\bar{u})\right)\mathcal{A}''(y_k)[S'(u_k)v_k]^2 \\ & + G'(B\bar{u})\left(\mathcal{A}''(y_k)[S'(u_k)v_k]^2 - \mathcal{A}''(\bar{y})[S'(\bar{u})v]^2\right), \end{aligned} \tag{15}$$

with $y_k = S(u_k)$ and $\bar{y} = S(\bar{u})$. Note that convergence in (14) is in particular true for $r \in \left(1, \frac{2p}{p+d}\right)$. Hence, it suffices to show that

$$\mathcal{A}''(y_k)[S'(u_k)v_k]^2 \rightharpoonup \mathcal{A}''(\bar{y})[S'(\bar{u})v]^2 \qquad \text{weakly in } L^r(I, W_D^{-1,p}), \tag{16}$$

which can be easily done utilizing Hölders inequality and the results obtained so far. ☐

Let us point out why Theorem 4.6 motivates the introduction of averaged-in-time state-constraints: Assume that we want to apply Theorem 3.2 to (OCP) in case of pointwise in time and space state-constraints (Theorem 2.3). Consequently, we have to verify Theorem 3.1 for $U_\infty = L^s(\Lambda)$, $U_2 = L^2(\Lambda)$, $K = U_{ad}$, $Z = C(\overline{Q})$, $C = Y_{ad}$, j being the reduced functional and $g = S$ being the control-to-state map of (OCP). In particular, we would have to show that $S'(u)$ extends to a bounded linear map $L^2(\Lambda) \to C(\overline{Q})$, and that $S''(u)$ extends to a continuous bilinear map $L^2(\Lambda) \times L^2(\Lambda) \to C(\overline{Q})$, for any fixed $u \in U_{ad}$. The proof of Theorem 4.6, however, shows this already fails to hold for the first derivative: From Theorem 2.2 we know that the extension

$$S'(u)\colon \quad L^2(\Lambda) \to W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p}),$$

is the best possible we can expect. However, there is no embedding $W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p}) \hookrightarrow C(\overline{Q})$. Due to $\frac{2p}{p+d} < 2$, the situation is even worse for $S''(u)$.

Similarly, one can observe that application to averaged-in-space and pointwise in time state-constraints [**31**], i.e.

$$Y_{ad} = \left\{ y\colon \quad y_a(t) \le \int_\Omega y(t,x)\omega(x)dx \le y_b(t) \qquad \forall t \in I \right\},$$

with continuous functions $y_a, y_b \in C(I)$ and a weight function $\omega \in L^\infty$, would require an embedding

$$W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}) \hookrightarrow C(I, L^1)$$

for some $r \in (1, \frac{2p}{p+d})$ in order to verify Theorem 3.1 (2bii). Unfortunately, such an embedding cannot be true. However, the embedding

$$W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}) \hookrightarrow L^1(I, W_D^{1,p}) \hookrightarrow L^1(I, C(\overline{\Omega}))$$

is obvious. Therefore, averaging in time –instead of averaging in space– seems to be a reasonable idea, which resulted in the formulation of Theorem 4.1.

Remark 4.7. The improved regularity analysis of [**4**] and considering the linearized state equation on certain Bessel-potential spaces instead of $W_D^{-1,p}$ does not improve the situation significantly, as can be seen along the lines of the proof of Theorem 4.6.

Moreover, we would like to point out that the appearance of the $\mathcal{A}''$-term in the second derivative of the control-to-state map, and hence in the second derivative of the Lagrangian of (OCP), makes it impossible to repeat the approach of [**27**], cf. in particular [**27**, Proposition 3.8]: The reason for the failure of this technique in our case is the presence of differential operators in $\mathcal{A}''$ that have to be applied to solutions of the linearized state equation. In contrast, for the semilinear equation discussed in [**27**] all terms in the second derivative of the nonlinearity are of order zero, which allows to get along with less regularity for the linearized state equation.

4.3. Second-order sufficient optimality conditions. Finally, using the previously obtained auxiliary results we can formulate and prove second-order sufficient conditions for (OCP). As already pointed out, the proof relies on Theorem 3.2. For convenience, we introduce the "regular part" $\hat{p}$ of the adjoint state $\bar{p}$ defined by the following equation

$$\begin{aligned} -\partial_t \hat{p} + \mathcal{A}(\bar{y})^* \hat{p} + \mathcal{A}'(\bar{y})^* \hat{p} &= \bar{y} - y_d, \\ \hat{p}(T) &= 0, \end{aligned} \tag{17}$$

Note that this allows us to express the first derivative of the reduced functional j as $j'(\bar{u})v = \langle B^*\hat{p} + \gamma\bar{u}, v\rangle_{L^2(\Lambda)}$, cf. Section 2.2.

Theorem 4.8. *Let Theorem 2.1 and Theorem 4.1 (1) hold, and let* $\bar{u} \in U_{ad}$*,* $\bar{y} = S(\bar{u}) \in Y_{ad}$*,* $\bar{\mu} \in \mathcal{M}(\overline{\Omega})$ *fulfill the optimality system* (10)-(12) *from Theorem 4.4. We define the critical cone by*

$$C_{\bar{u},\bar{\mu}} := \left\{ v \in L^2(\Lambda)\colon \text{(18)-(20) } \textit{hold} \right\};$$

$$\int_\Lambda (\gamma\bar{u} + B^*\hat{p})v = 0, \qquad \int_\Omega \int_0^T [S'(\bar{u})v](t,x)dtd\bar{\mu} = 0, \tag{18}$$

$$\begin{aligned} &\int_0^T [S'(\bar{u})v](t,x)dt \geq 0, \qquad \textit{if} \quad \int_0^T \bar{y}(t,x)dt = y_a(x), \\ &\int_0^T [S'(\bar{u})v](t,x)dt \leq 0, \qquad \textit{if} \quad \int_0^T \bar{y}(t,x)dt = y_b(x), \end{aligned} \tag{19}$$

$$v \leq 0, \qquad \textit{on } \{\bar{u} = u_b\}, \quad v \geq 0, \qquad \textit{on } \{\bar{u} = u_a\}, \tag{20}$$

where $\bar{p}$ and $\hat{p}$ are defined by (10) *and* (17)*, respectively. If the positivity condition*

$$\begin{aligned} \gamma\|v\|^2_{L^2(\Lambda)} + \int_Q \big(|S'(\bar{u})v|^2 - \nabla\bar{p}\cdot(2\xi'(\bar{y})[S'(\bar{u})v]\mu\nabla[S'(\bar{u})v] \\ +\xi''(\bar{y})[S'(\bar{u})v]^2\mu\nabla\bar{y})\big)\,dxdt \quad > 0 \qquad \forall v \in C_{\bar{u},\bar{\mu}} \setminus \{0\}, \end{aligned} \tag{21}$$

holds, there are $\epsilon, \delta > 0$ such that the quadratic growth condition

$$j(u) \geq j(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|^2_{L^2(\Lambda)}$$

is satisfied for all $u \in U_{ad}$ such that $\|u - \bar{u}\|_{L^2(\Lambda)} < \epsilon$ and $y_a(x) \leq \int_{t_1}^{t_2} S(u)(t,x)dt \leq y_b(x)$ for all $x \in \overline{\Omega}$. In particular, $\bar{u}$ is an local solution of (OCP) *with respect to the $L^2(\Lambda)$-topology.*

Proof. We apply Theorem 3.2 with $U_\infty = L^s(\Lambda)$, $U_2 = L^2(\Lambda)$, $K = U_{ad}$, $Z = C(\overline{\Omega})$, and $C = \{z \in C(\overline{\Omega})\colon y_a \leq z \leq y_b \text{ on } \overline{\Omega}\}$. The properties for the reduced functional j, $j(u) = J(S(u), u)$, required in Theorem 3.1 have already been checked in [**4**, Theorem 4.14]. Note that the average-in-time map is linear and continuous from both $L^q(I, C(\overline{\Omega}))$ and $W^{1,r}(I, W^{-1,p}) \cap L^r(I, W^{1,p}) \hookrightarrow L^r(I, C(\overline{\Omega}))$ into $C(\overline{\Omega})$ for any $q, r \geq 1$. Hence, extension and continuity properties for the derivatives of $g := A \circ S$ in Theorem 3.1 (2) immediately follow from Theorem 4.6. Hereby, observe that convergence of $(u_k)_k \subset U_{ad}$ to $\bar{u}$ w.r.t. $L^2(\Lambda)$ implies, due to $L^\infty(\Lambda)$-boundedness of U_{ad}, also convergence w.r.t. $L^s(\Lambda)$ by the Riesz-Thorin interpolation theorem. Therefore, application of Theorem 4.6 is possible. □

Although (OCP) with averaged-in-time state-constraints is slightly easier than (OCP) with pointwise in time and space state-constraints from an analytical point of view, Theorem 4.8 still illustrates the full strength of our abstract Theorem 3.2. In order to prove C^2-differentiability of the control-to-state map we require to consider controls in $L^s(\Lambda)$ with $s \gg 1$ as in Theorem 2.1, cf. [**4**], because already existence of solutions to (SE) relies on such an assumption [**34**]. Hence, C^2-differentiability, and even well-definedness, of the reduced functional j is guaranteed on $L^s(\Lambda)$, but not necessarily on $L^2(\Lambda)$. However, we cannot hope for a coercivity or positivity condition like (21) with the increments v coming from $L^s(\Lambda)$. The latter condition can only hold for v coming from $L^2(\Lambda)$, cf. [**12, 24, 40**]. For the same reason, a similar situation holds for $g := A \circ S$. It is immediately clear that g is well-defined and C^2-differentiable on $L^s(\Lambda)$. The question, however, whether g is even well-defined on $L^2(\Lambda)$ is not clear. Although the problem necessarily requires us to refer to two non-equivalent norms, as just explained, a norm gap in the formulation of Theorem 4.8 can be avoided. This is main benefit and novelty of Theorem 3.2.

5. SSCs for pointwise state-constraints

In the previous section we decided to relax the type of state-constraints while keeping the regularity assumptions for the equation unchanged. In the present section we proceed the other way round and strengthen the regularity assumptions and restrict ourselves to purely timedependent controls. In return, we establish SSCs for (OCP) with pointwise in time and space state-constraints as introduced in Section 2, which is our second main result for our precise model problem. We replace Theorem 2.1 by a slightly smoother setting that allows to apply stronger regularity results achieved by Casas and Chrysafinos [**8**]. Based on their analysis we derive a result analogous to Theorem 4.6 in the L^p-$W^{2,p}$-setting that finally allows to apply Theorem 3.2 also in case of pointwise in time and space state-constraints.

5.1. Regularity assumptions for the state equation. For brevity, we do not exploit the results of [**8**] in their full generality, that allows, contrary to [**4**, **34**], e.g. for unbounded nonlinearities and a semilinear term in the state equation. Instead, we state the following regularity assumptions for domain, coefficients, and initial conditions that are those of [**8**] applied to the setting described in Theorem 2.1:

Assumption 5.1. 1. $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ is a bounded domain with $C^{1,1}$-boundary Γ, and homogeneous Dirichlet boundary conditions hold on the entire boundary Γ.

2. Let Theorem 2.1 (2) on μ and ξ hold and assume in addition that μ is Lipschitz-continuous as map $\Omega \to \mathbb{R}^{d\times d}$.

3. Choose $p > d$ and $s > 2$ such that $\frac{1}{s} < \frac{1}{2}\left(1 - \frac{d}{p}\right)$. The set of admissible controls is given by

$$U_{ad} := \{u \in L^{2s}(I, \mathbb{R}^m) : u_a \le u \le u_b \text{ on } I\}$$

with control-bounds $u_a, u_b \in L^\infty(I, \mathbb{R}^m)$, and for fixed control-functions $b_i \in L^p$, $i = 1, ..., m$ we define

$$B\colon L^{2s}(I, \mathbb{R}^m) \to L^{2s}(I, L^p), \quad u \mapsto \sum_{i=1}^{m} u_i b_i.$$

The initial value y_0 for the state equation fulfills

$$y_0 \in (L^p, W^{2,p} \cap W_D^{1,p})_{1-1/s,s} \cap (W_D^{-1,2p}, W_D^{1,2p})_{1-1/(2s),2s} \cap C(\Omega),$$

and the desired state has regularity $y_d \in L^\infty(I, L^2)$.

Unlike in [**8**] we have to restrict ourselves to purely timedependent controls as introduced in [**15**]. The reason is the following, cf. also [**15**, Remark 2]: When switching from controls in $U_\infty = L^{2s}(I, \mathbb{R}^m)$ to controls in $U_2 = L^2(I, \mathbb{R}^m)$, only time integrability decreases, but the spatial regularity of the right-hand-sides of the PDEs is not affected. This turns out to be crucial for obtaining the required regularity for the derivatives of the control-to-state map. From the applied point of view, having only finitely many pre-defined actuators to influence a system might also seem reasonable. However, note that L^p-regularity (unlike $W_D^{-1,p}$-regularity in [**4**, Example 2.5]) of the fixed control-functions now excludes any possibility of boundary control.

Remark 5.2. From [**8**, p. 609] we recall the following observation: $C^{1,1}$-smoothness of Γ, combined with homogeneous Dirichlet boundary conditions and Lipschitz-continuity of μ implies that

$$-\nabla\cdot\mu\nabla+1: \qquad W_D^{1,q}\to W_D^{-1,q}$$

is a topological isomorphism for any $q\in(1,\infty)$. Consequently, Theorem 5.1 is indeed a tightened version of Theorem 2.1.

5.2. Improved regularity of the state. We start our analysis by recalling the following regularity result by Casas and Chrysafinos that will be the cornerstone of our further analysis:

Theorem 5.3 ([**8**], Theorem 2.3)**.** *Let Theorem 5.1 (1) and (2) hold and fix $p,s\in[2,\infty)$ such that $\frac{1}{s}+\frac{d}{p}<2$. Given*

$$v\in L^{2s}(I,L^p),$$
$$y_0\in(L^p,W^{2,p}\cap W_D^{1,p})_{1-1/s,s}\cap(W_D^{-1,2p},W_D^{1,2p})_{1-1/(2s),2s}\cap C(\Omega)$$

there is a unique solution y to equation (1) with regularity

$$y\in W^{1,s}(I,L^p)\cap L^s(I,W^{2,p}\cap W_D^{1,p}).$$

In particular, the control-to-state map S introduced in Section 2.2 is well defined from $L^{2s}(I,\mathbb{R}^m)$ to $W^{1,s}(I,L^p)\cap L^s(I,W^{2,p}\cap W_D^{1,p})$ under Theorem 5.1 .

Note that [**8**, Remark 2], that addresses relaxation of the smoothness of Γ in case of convex polygonal/polyhedral domains, does only apply to our case in dimension $d=2$: To have sufficient spatial regularity in the following corollary we require $p>d$ which excludes the mentioned relaxations for $d=3$.

Corollary 5.4. *Under Theorem 5.1 there are some $\rho,\kappa>0$ such that*

$$W^{1,s}(I,L^p)\cap L^s(I,W^{2,p}\cap W_D^{1,p})\hookrightarrow C^{0,\rho}(I,C^{1,\kappa}).$$

Proof. Choose $\frac{1}{2}\left(1+\frac{d}{p}\right)<\theta<1-\frac{1}{s}$ and set $\rho=1-\frac{1}{s}-\theta>0$. Then it holds

$$W^{1,s}(I,L^p)\cap L^s(I,W^{2,p}\cap W_D^{1,p})\hookrightarrow C^{0,\rho}(I,(L^p,W^{2,p})_{\theta,1})$$

by standard Bochner-Sobolev embedding, see e.g. [**1**]. Further, it is well-known that $(L^p,W^{2,p})_{\theta,1}\hookrightarrow[L^p,W^{2,p}]_\theta$. Since Ω is in particular a domain with Lipschitz boundary, there is a bounded linear extension operator $L^p\to L^p(\mathbb{R}^d)$ that restricts to a bounded extension operator $W^{2,p}\to W^{2,p}(\mathbb{R}^d)$ [**37**]. Thus, a standard argument utilizing the retraction-coretraction theorem ([**39**, Theorem 1.2.4], [**3**, Proposition I.2.3.2]) shows that it suffices to prove $[L^p(\mathbb{R}^d),W^{2,p}(\mathbb{R}^d)]_\theta\hookrightarrow C^{1,\kappa}(\mathbb{R}^d)$. The latter follows from $[L^p(\mathbb{R}^d),W^{2,p}(\mathbb{R}^d)]_\theta=H^{2\theta,p}(\mathbb{R}^d)$ [**39**, Theorem 4.3.2.2] and standard Sobolev embeddings on $\mathbb{R}^d$ with $\kappa=2\theta-\frac{d}{p}-1>0$ [**39**, Theorem 2.8.1]. □

5.3. Improved regularity for derivatives of the control-to-state map. The goal of this subsection is to provide an improved version of Theorem 2.2 under the strengthened regularity Theorem 5.1. Hereby, the improved regularity of the state from Theorem 5.3 is the crucial point, because we can show that the domain of $-\nabla\cdot\xi(y(t))\mu\nabla$ in L^p is independent of $t\in I$ for $y\in C^{0,\rho}(I,C^{1,\kappa})$. Therefore, it is possible to show that $\mathcal{A}(y)$ and $\mathcal{A}(y)+\mathcal{A}'(y)$ exhibit maximal parabolic regularity [**1**, **2**] on L^p-spaces, which finally allows to prove the desired regularity result analogous to Theorems 2.2 and 4.6. The approach is similar to [**4**] with the essential difference that the weaker regularity $y \in W^{1,s}(I, W^{1,p}_D) \cap L^s(I, W^{1,p}_D)$ for the states in [**4**, Section 3.2] suffices to show constant domains and maximal parabolic regularity on $H^{-\zeta,p}_D$ for certain $\zeta \in (0,1)$ close to 1, but not on L^p, cf. the proof of [**4**, Proposition 3.17]. However, an analysis carried out on $H^{-\zeta,p}_D$ will not suffice for the derivation of SSCs for (OCP) in case of pointwise in time and space state-constraints as explained in Theorem 4.7.

Although the following observation is rather trivial in our case, we explicitly state it due to its high importance for the following results.

Lemma 5.5. *Under Theorem 5.1 (1) and (2) let $\eta \in W^{1,\infty}$ with $\eta \geq \eta_\bullet > 0$ on Ω. Then it holds:*

(1) $\mathrm{Dom}_{L^p}(-\nabla \cdot \eta\mu\nabla + 1) \cong \mathrm{Dom}_{L^p}(-\nabla \cdot \mu\nabla + 1) = W^{2,p} \cap W^{1,p}_D$, *i.e.* $-\nabla \cdot \eta\mu\nabla + 1$ *is a topological isomorphism* $W^{2,p} \cap W^{1,p}_D \to L^p$.
(2) *The map* $\eta \mapsto -\nabla \cdot \eta\mu\nabla$ *is bounded linear as map* $W^{1,\infty} \to \mathcal{L}(W^{2,p} \cap W^{1,p}_D, L^p)$.

Similar results have been obtained in [**21**, Lemmas 6.5, 6.7, Corollary 6.8] if (SE) is considered on certain Bessel-potential spaces instead of L^p. The following lemma is a first step towards the analysis of the linearized state equation on L^p.

Lemma 5.6. *Let Theorem 5.1 (1) and (2) hold. If $y \in W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$, the nonautonomous operator $\mathcal{A}(y)$ exhibits maximal parabolic regularity on $L^r(I, L^p)$, $r \in (1,\infty)$. Moreover, the operator norms of*

$$(\partial_t + \mathcal{A}(y), \mathrm{tr}_0)^{-1}\colon \qquad L^r(I, L^p) \times (L^p, W^{2,p} \cap W^{1,p}_D)_{1/r',r} \\ \to W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D)$$

are uniformly bounded for y coming from a bounded set in $W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$.

The proof relies on the same technique as in [**4**, Theorem 3.20]. Nevertheless, the present situation is slightly easier than in [**4**], because the additional regularity assumptions ensure that the domains of $\mathcal{A}(y(t))$ in L^p stay independent of t.

Proof. We apply [**4**, Lemma D.1], see also [**35**, Corollary 14]. First, note that L^p is an UMD space, see [**3**, Section III.4.4] for the definition. Uniform resolvent estimates and uniform $\mathcal{R}$-sectoriality for $A(t) := -\nabla \cdot \xi(y(t))\mu\nabla$ on L^p have already been established, see formula (3.16) and Lemma 3.12 in [**4**], respectively; note that uniformity already holds for y's coming from a bounded set in $C^\alpha(\overline{Q})$, which is a much weaker assumption than in the present case. It remains to check the so-called Acquistapace-Terreni condition on L^p. Note that the latter was done in [**4**] only on the Bessel-potential spaces $H^{-\zeta,p}_D$ for appropriate $\zeta \in (0,1)$, but not on L^p. However, having at hand Theorem 5.5 the same technique as in the proof of [**4**, Proposition 3.18] yields the required condition. □

Now, we consider maximal parabolic regularity for the linearized state equation. This extends Theorem 2.2 (see also [**4**, Proposition 4.4], [**8**, Theorem 3.2]), where maximal parabolic regularity on $W^{-1,p}$ has been dealt with.

Lemma 5.7. *Let Theorem 5.1 (1) and (2) hold, and fix $y \in W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$. For any $r \in (1, s]$ and $f \in L^r(I, L^p)$, the linearized state equation*

$$\partial_t w + \mathcal{A}(y)w + \mathcal{A}'(y)w = f, \qquad w(0) = 0,$$

has a unique solution $w \in W^r(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D)$. The nonautonomous operator $\mathcal{A}(y) + \mathcal{A}'(y)$ has maximal parabolic regularity on $L^r(I, L^p)$ for $r \in (1, s]$.

Proof. The proof works analogous to the proof of [**4**, Proposition 4.4] and relies on application of [**2**, Theorem 7.1] (for $r < s$) and [**36**, Corollary 3.4] (for $r = s$), respectively. Then main steps are to observe that

$$I \to \mathcal{L}(W^{2,p} \cap W^{1,p}_D, L^p), \qquad t \mapsto -\nabla \cdot \xi(y(t))\mu\nabla,$$

is continuous, and that

$$\begin{aligned}\mathcal{A}'(y) = (t \mapsto (\psi \mapsto -\nabla \cdot \xi'(y)\psi\mu\nabla y)) \in L^s\left(I, \mathcal{L}(W^{1,\infty}, L^p)\right) \\ \hookrightarrow L^s\left(I, \mathcal{L}((L^p, W^{2,p})_{\theta,\infty}, L^p)\right)\end{aligned}$$

with some θ such that $1 - 1/s > \theta > \frac{1}{2} + \frac{d}{2p}$. Hereby, we made use of the embedding $(L^p, W^{2,p})_{\theta,\infty} \hookrightarrow W^{1,\infty}$, cf. the proof of Theorem 5.4. □

Let us point out that Theorem 5.7 and Theorem 5.3 do *not* allow immediately to conclude differentiability of the solution map of (1) from $L^r(I, L^p)$ to $W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D)$! Of course, for $\frac{1}{r} < 1 - \frac{d}{2p}$, e.g. $r = 2$, there is an embedding

$$W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D) \hookrightarrow C(\overline{Q}),$$

which can shown with a similar argument as for Theorem 5.4. Hence, the map

$$\begin{aligned}F\colon W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D) \times L^r(I, L^p) \\ \to L^r(I, L^p) \times (L^p, W^{2,p} \cap W^{1,p}_D)_{1/r',r}, \\ (y, v) \mapsto (\partial_t y + \mathcal{A}(y) - v, y(0) - y_0),\end{aligned}$$

is continuously Fréchet differentiable. Further, for $r \in (1, s]$ the partial derivative $\partial_y F(y, v)$ is even continuously invertible as Theorem 5.7 shows. Nevertheless, the fact that prevents us from application of the implicit function theorem is that we first require a *well-defined* solution map $v \mapsto y(v)$ associated to $F(y, v) = 0$, and we do not have such a map at hand: To obtain solutions to (1) in $W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$ we need right-hand-sides $v \in L^{2s}(I, L^p)$ and not in $L^s(I, L^p)$, see Theorem 5.3: For $v \in L^s(I, L^p)$ we do not know whether there exists some $y \in W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$ such that $F(y, v) = 0$. On the other hand, $\partial_y F(y, v)$ cannot be invertible from $W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$ to $L^{2s}(I, L^p)$, because invertibility of $\partial_y F(y, v)$ always holds between $W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D)$ and $L^r(I, L^p)$, $r \in (1, s]$, cf. Theorem 5.7.

Remark 5.8. The reason for requiring double time integrability on the right-hand-side of (1) in Theorem 5.3, is due to the technique applied in the proof of [**8**, Theorem 2.3].

The following lemma is the first step of establishing an analogue to Theorem 4.6 (1) under Theorem 5.1. In particular, it can be seen that the regularity of the $\mathcal{A}''$-term appearing in the second derivative of the control-to-state map can be essentially improved in the present case. Nevertheless, the fact that even in the this highly regular setting $\mathcal{A}''(y)w^2$ will be from $L^r(I, W_D^{-1,p})$, i.e. a distributional object in general, again illustrates the difficulties when dealing with this term.

Lemma 5.9. *Given $y \in W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W_D^{1,p})$ and $w \in W^{1,2}(I, L^p) \cap L^2(I, W^{2,p} \cap W_D^{1,p})$ it holds*

$$\|\mathcal{A}''(y)w^2\|_{L^r(I,W_D^{-1,p})} \leq c(y,r)\|w\|^2_{W^{1,2}(I,L^p)\cap L^2(I,W^{2,p}\cap W_D^{1,p})}$$

for $r \in (1, \infty)$. The constant $c(y, r)$ can be chosen uniformly with respect to y coming from a bounded set in $L^\infty(I, W^{1,p})$.

Proof. This follows from the definition of $\mathcal{A}''$ and Hölders inequality. We have to make use of the embeddings

$$W^{1,2}(I, L^p) \cap L^2(I, W^{2,p} \cap W_D^{1,p}) \hookrightarrow C(\overline{Q}), \tag{22}$$

$$\text{and} \quad W^{1,2}(I, L^p) \cap L^2(I, W^{2,p} \cap W_D^{1,p}) \hookrightarrow L^q(I, W^{1,p}) \tag{23}$$

for every $q \in (1, \infty)$, that can be shown similarly as in Theorem 5.4. □

The following lemma concerning continuity of the control-to-state map in the improved regularity setting is the last auxiliary result before we will be able to verify the assumptions of Theorem 3.2 in the proposition thereafter.

Lemma 5.10. *The solution map of the state equation* (1) *is continuous from $L^{2s}(I, L^p)$ to $W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W_D^{1,p})$.*

Note that this result is not explicitly contained in [**8**]. There, differentiability and, consequently, continuity of the control-to-state map have been addressed in the $W_D^{-1,p}$-setting instead, cf. [**8**, Theorem 3.2]. As outlined after Theorem 5.7, arguing via the implicit function theorem is not possible here. However, to prove continuous dependence it suffices to go through the steps in [**8**] tracking continuous dependence of the quantities under consideration. We omit the details.

As announced, the following proposition is our analogon to Theorem 4.6 for the present section. It provides the main steps in checking Theorem 3.1 for the setting described by Theorem 5.1, and therefore it forms the main part of the proof of our second main result, SSCs for (OCP) in case of pointwise in time and space state-constraints, below.

Proposition 5.11. *Let Theorem 5.1 hold. The control-to-state map is twice continuously Fréchet differentiable as map S: $L^s(I, \mathbb{R}^m) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$. Moreover, the following continuation and continuity properties hold for the respective derivatives:*

(1) *Fix $u \in L^{2s}(I, \mathbb{R}^m)$. The first derivative $S'(u)$ extends to a continuous linear map from $L^2(I, \mathbb{R}^m)$ to $C(\overline{Q})$, and the second derivative $S''(u)$ extends to a continuous bilinear map from $L^2(I, \mathbb{R}^m) \times L^2(I, \mathbb{R}^m)$ to $C(\overline{Q})$, as well.*

(2) *Let* $(u_k)_k \subset L^{2s}(I,\mathbb{R}^m)$ *converge to* $\bar u$ *strongly in* $L^{2s}(I,\mathbb{R}^m)$ *and* $(v_k)_k \subset L^2(I,\mathbb{R}^m)$ *converge weakly in* $L^2(I,\mathbb{R}^m)$ *to some* v. *Then it holds:*

$$S'(u_k)v_k \rightharpoonup S'(\bar u)v_k \qquad \text{weakly in } C(\overline{Q}),$$
$$S''(u_k)v_k^2 \rightharpoonup S''(\bar u)v^2 \qquad \text{weakly in } C(\overline{Q}).$$

The structure of the following argument is the same as the one of the proof of Theorem 4.6.

Proof. Differentiability of the control-to-state map and the formulas for the respective derivatives follows from Theorem 2.2. Note that Theorem 5.1 indeed suffices to invoke this result, cf. Theorem 5.2.

1. The extension property for the first derivative follows from Theorem 5.7 with $r = 2$ and embedding (22). For the continuation of the second derivative, combine the continuation property for $S'(u)$ with Theorems 5.7 and 5.9 and the embedding from [**4**, Proposition 3.3].

2. It remains to check the continuity properties: As an auxiliary result, we first show that

$$S'(u_k) \to S'(\bar u) \quad \text{in } \mathcal{L}\left(L^r(I,\mathbb{R}^m), W^{1,r}(I,L^p)\cap L^r(I,W^{2,p}\cap W^{1,p}_D)\right) \tag{24}$$

for any $r \in (1,\infty)$. To do so, it suffices, by continuity of operator inversion, to show convergence

$$\mathcal{A}(y_k) + \mathcal{A}'(y_k) \to \mathcal{A}(\bar y) + \mathcal{A}'(\bar y)$$
$$\text{in } \mathcal{L}\left(W^{1,r}(I,L^p)\cap L^r(I,W^{2,p}\cap W^{1,p}_D), L^r(I,L^p)\right).$$

This can be done using Theorem 5.10, Hölders inequality and the embedding

$$W^{1,r}(I,L^p)\cap L^r(I,W^{2,p}\cap W^{1,p}_D) \hookrightarrow L^q(I,W^{1,\infty}).$$

for some q such that $\frac{1}{q}+\frac{1}{s}\le\frac{1}{r}$, which can be shown by the same technique as for Theorem 5.4. Having at hand this auxiliary result, the continuity property for the first derivative follows similarly as in the proof of Theorem 4.6.

For the second derivative we also argue similarly as in the proof of Theorem 4.6: Due to the embedding from [**4**, Proposition 3.3] for $r > \frac{2p}{p-d}$ it suffices to show that

$$G'(Bu_k) \to G'(B\bar u) \quad \text{in} \quad \mathcal{L}\left(L^r(I,W^{-1,p}_D), W^{1,r}(I,W^{-1,p}_D)\cap L^r(I,W^{1,p}_D)\right), \tag{25}$$

and $\mathcal{A}''(y_k)[w_k]^2 \rightharpoonup \mathcal{A}''(y)[w]^2$ weakly in $L^r(I,W^{-1,p}_D)$. We leave the details to the reader. □

5.4. Second-order sufficient optimality conditions. In the following we apply Theorem 3.2 to (OCP) under Theorems 2.3 and 5.1: We formulate SSCs for (OCP) for pointwise in time and space state-constraints, which is our second main result of this paper. Compared to our first main result, Theorem 4.8, we crucially rely on the improved regularity results due to the strengthened regularity Theorem 5.1.

Theorem 5.12. *Let Theorem 5.1 and Theorem 2.3 (1) hold, and let* $\bar u \in L^{2s}(I,\mathbb{R}^m)$, $\bar y \in W^{1,s}(I,L^p)\cap L^s(I,W^{2,p}\cap W^{1,p}_D)\cap Y_{ad}$ *and* $\bar\mu \in \mathcal{M}(\overline{Q})$ *fulfill the first-order necessary optimality conditions* (4)-(6) *from Theorem 2.6.*

We define the cone of critical directions by

$$C_{\bar u,\bar\mu} := \left\{v \in L^2(I,\mathbb{R}^m)\text{: } (26)-(28) \text{ hold}\right\},$$

$$\int_0^T (\gamma\bar{u}(t) + B^*\hat{p}(t))^T v(t)dt = 0, \qquad \int_{\overline{Q}} [S'(\bar{u})v]d\bar{\mu} = 0 \tag{26}$$

$$\begin{aligned} &[S'(\bar{u})v](t,x) \le 0, \qquad \textit{if } \bar{y}(t,x) = y_b(t,x), \\ &[S'(\bar{u})v](t,x) \ge 0, \qquad \textit{if } \bar{y}(t,x) = y_a(t,x), \end{aligned} \tag{27}$$

$$v_i(t) \le 0, \qquad \textit{if } \bar{u}_i(t) = u_{b,i}(t), \quad v_i(t) \ge 0, \qquad \textit{if } \bar{u}_i(t) = u_{a,i}(t), \tag{28}$$

where the regular part $\hat{p}$ of the adjoint state is defined as in (17)*. If the positivity condition*

$$\gamma\|v\|^2_{L^2(I,\mathbb{R}^m)} + \int_Q \left(|S'(\bar{u})v|^2 - \nabla\bar{p}\cdot(2\xi'(\bar{y})[S'(\bar{u})v]\mu\nabla[S'(\bar{u})v] \right. \\ \left. + \xi''(\bar{y})[S'(\bar{u})v]^2\mu\nabla\bar{y})\right) dxdt \quad > 0 \qquad \forall v \in C_{\bar{u},\bar{\mu}} \setminus \{0\}, \tag{29}$$

holds true, then $\bar{u}$ is a $L^2(I,\mathbb{R}^m)$-local minimizer for (OCP)*, and there are $\epsilon, \delta > 0$ such that the quadratic growth condition*

$$j(u) \ge j(\bar{u}) + \frac{\delta}{2}\|u - \bar{u}\|^2_{L^2(I,\mathbb{R}^m)}$$

holds for all $u \in U_{ad}$ that satisfy $\|u - \bar{u}\|_{L^2(I,\mathbb{R}^m)} < \epsilon$ and $S(u) \in Y_{ad}$.

Proof. We apply Theorem 3.2 with $U_\infty = L^s(I,\mathbb{R}^m)$, $U_2 = L^2(I,\mathbb{R}^m)$, $Z = C(\overline{Q})$, $K = U_{ad}$, $C = Y_{ad}$. As already stated in the proof of Theorem 4.8, the assumptions on the reduced functional j from Theorem 3.1 (1) have already been verified in [**4**]. Finally, Theorem 3.1 (2) for $g = S$ is fulfilled due to Theorem 5.11. Hereby, the crucial point is as in the proof of Theorem 4.8 to observe that due to L^∞-boundedness of U_{ad} convergence w.r.t. $L^2(I,\mathbb{R}^m)$ implies convergence w.r.t. $L^{2s}(I,\mathbb{R}^m)$ and $L^s(I,\mathbb{R}^m)$, respectively. □

Again, we would like to point out that the problem formulation requires two different norms: Reduced functional and control-to-state map are well-defined and C^2-Fréchet on $L^s(I,\mathbb{R}^m)$ with some $s \gg 2$, but not necessarily on $L^2(I,\mathbb{R}^m)$. However, positivity condition (29) might hold for directions v from $L^2(I,\mathbb{R}^m)$, but cannot be expected to hold for directions $v \in L^s(I,\mathbb{R}^m)$. Nevertheless, it is possible to state the quadratic growth condition in Theorem 5.12 only referring to the $L^2(I,\mathbb{R}^m)$-norm, i.e. similarly to Theorem 4.8 and Theorem 3.3 occurrence of a two-norm gap can be avoided although we might not expect this during a first look at the problem.

6. The case without control-constraints

Finally, we comment on the question how the results of the previous sections change if we do not impose (bilateral) constraints on the controls: In the terminology of Theorem 2.1 and Theorem 5.1, respectively, we consider $u_a \equiv -\infty$ and $u_b \in L^\infty(\Lambda)$, or $u_b \equiv +\infty$ and $u_a \in L^\infty(\Lambda)$.

6.1. Existence of optimal controls. The argument in the proof of Theorem 2.4 relies on the possibility to extract a $L^s(\Lambda)$-bounded subsequence of the infimizing sequence of controls, see the proof of [**34**, Theorem 6.3]. Therefore, we require either boundedness of U_{ad} in $L^s(\Lambda)$, or boundedness of the infimizing sequence has to be enforced by the choice of the objective functional. Since the functional of (OCP) is not $L^s(\Lambda)$-coercive, we are not able to prove existence of optimal controls in case of unilateral or no control-constraints. A way out is to add a $L^s(\Lambda)$-Tikhonov term: After replacing the original functional of (OCP) by the modified version

$$J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(I\times\Omega)} + \frac{\gamma}{2}\|u\|^2_{L^2(\Lambda)} + \frac{\gamma_s}{s}\|u\|^s_{L^s(\Lambda)}, \tag{30}$$

with some $\gamma_s > 0$, the techniques used in the proof of Theorem 2.4 apply again, and existence of a globally optimal control can be shown.

6.2. First-order necessary optimality conditions. Theorems 2.6 and 4.4 do not rely on the existence of optimal controls; they only characterize locally optimal controls, if they exist. Therefore, these results stay valid in case of unilateral or no control-constraints. In the particular case $U_{ad} = L^s(\Lambda)$ the variational inequalities (6) and (12), respectively, simplify to the equality $B^*\bar{p} + \gamma\bar{u} = 0$, of course. Moreover, the reduced gradient for the modified functional (30) is given by

$$\nabla j(\bar{u}) = B^*\bar{p} + \gamma\bar{u} + \gamma_s|\bar{u}|^{s-1},$$

and (6) and (12) have to be adapted accordingly.

6.3. Second-order sufficient conditions. The arguments for the verification of Theorem 3.1 make use of the L^∞-boundedness of the admissible set U_{ad}, cf. the proofs of Theorems 4.6 and 5.11 and [**4**, Theorem 4.14]. Hence, we cannot apply Theorem 3.2 in order to obtain two-norm gap free results analogous to Theorem 4.8 or Theorem 5.12 in the case without bilateral control-constraints.

However, following Theorem 3.4 it is still possible to obtain SSCs with norm gap: The quadratic growth condition in the respective modified versions of Theorems 4.8 and 5.12 will only hold on an $L^s(\Lambda)$- and $L^{2s}(I,\mathbb{R}^m)$-neighborhood of $\bar{u}$, respectively.

Furthermore, we have to point out that when using the modified functional (30) the choice $\gamma = 0$ is not possible when aiming at SSCs, even at such with two-norm gap: The condition $\gamma > 0$ is crucial for verifying Theorem 3.1 (1biii) for the reduced functional; see formula (30) in [**15**], or formula (5.3) in [**11**], respectively.

References

[1] H. Amann, *Nonautonomous parabolic equations involving measures*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **306** (2003), no. Kraev. Zadachi Mat. Fiz. i Smezh. Vopr. Teor. Funktsii. 34, 16–52, 229. MR2065496

[2] ———, *Maximal regularity for nonautonomous evolution equations*, Adv. Nonlinear Stud. **4** (2004), no. 4, 417–430. MR2100906

[3] Herbert Amann, *Linear and quasilinear parabolic problems. Vol. I*, Monographs in Mathematics, vol. 89, Birkhäuser Boston, Inc., Boston, MA, 1995. Abstract linear theory. MR1345385

[4] Lucas Bonifacius and Ira Neitzel, *Second order optimality conditions for optimal control of quasilinear parabolic equations*, Math. Control Relat. Fields **8** (2018), no. 1, 1–34. MR3810865

[5] J. Frédéric Bonnans and Alexander Shapiro, *Perturbation analysis of optimization problems*, Springer Series in Operations Research, Springer-Verlag, New York, 2000. MR1756264

[6] Eduardo Casas, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control Optim. **24** (1986), no. 6, 1309–1318. MR861100

[7] ———, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control Optim. **31** (1993), no. 4, 993–1006. MR1227543

[8] Eduardo Casas and Konstantinos Chrysafinos, *Analysis and optimal control of some quasilinear parabolic equations*, Math. Control Relat. Fields **8** (2018), no. 3-4, 607–623. MR3917455

[9] Eduardo Casas, Juan Carlos de los Reyes, and Fredi Tröltzsch, *Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM J. Optim. **19** (2008), no. 2, 616–643. MR2425032

[10] Eduardo Casas, Mariano Mateos, and Boris Vexler, *New regularity results and improved error estimates for optimal control problems with state constraints*, ESAIM Control Optim. Calc. Var. **20** (2014), no. 3, 803–822. MR3264224
[11] Eduardo Casas and Fredi Tröltzsch, *Second order analysis for optimal control problems: improving results expected from abstract theory*, SIAM J. Optim. **22** (2012), no. 1, 261–279. MR2902693
[12] Eduardo Casas and Fredi Tröltzsch, *Second order optimality conditions and their role in PDE control*, Jahresber. Dtsch. Math.-Ver. **117** (2015), no. 1, 3–44. MR3311948
[13] Constantin Christof and Boris Vexler, *New regularity results and finite element error estimates for a class of parabolic optimal control problems with pointwise state constraints*, 2020.
[14] Constantin Christof and Gerd Wachsmuth, *No-gap second-order conditions via a directional curvature functional*, SIAM J. Optim. **28** (2018), no. 3, 2097–2130. MR3832976
[15] 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), no. 1, 5–38. MR2440722
[16] Karoline Disser, Hans-Christoph Kaiser, and Joachim Rehberg, *Optimal Sobolev regularity for linear second-order divergence elliptic operators occurring in real-world problems*, SIAM J. Math. Anal. **47** (2015), no. 3, 1719–1746. MR3342697
[17] Luis A. Fernández, *State constrained optimal control for some quasilinear parabolic equations*, Optimal control of partial differential equations (Chemnitz, 1998), 1999, pp. 145–156. MR1723981
[18] ———, *Integral state-constrained optimal control problems for some quasilinear parabolic equations*, Nonlinear Anal. **39** (2000), no. 8, Ser. A: Theory Methods, 977–996. MR1735179
[19] H. Goldberg and F. Tröltzsch, *Second-order sufficient optimality conditions for a class of nonlinear parabolic boundary control problems*, SIAM J. Control Optim. **31** (1993), no. 4, 1007–1025. MR1227544
[20] Konrad Gröger, *A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations*, Math. Ann. **283** (1989), no. 4, 679–687. MR990595
[21] Robert Haller-Dintelmann and Joachim Rehberg, *Maximal parabolic regularity for divergence operators including mixed boundary conditions*, J. Differential Equations **247** (2009), no. 5, 1354–1396. MR2541414
[22] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE constraints*, Mathematical Modelling: Theory and Applications, vol. 23, Springer, New York, 2009. MR2516528
[23] Fabian Hoppe and Ira Neitzel, *Convergence of the SQP method for quasilinear parabolic optimal control problems*, Optim. Eng. (2020).
[24] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum. III. Second order conditions and augmented duality*, SIAM J. Control Optim. **17** (1979), no. 2, 266–288. MR525027
[25] B. T. Kien, V. H. Nhu, and A. Rösch, *Second-order necessary optimality conditions for a class of optimal control problems governed by partial differential equations with pure state constraints*, J. Optim. Theory Appl. **165** (2015), no. 1, 30–61. MR3327415
[26] B. T. Kien, V. H. Nhu, and N. H. Son, *Second-order optimality conditions for a semilinear elliptic optimal control problem with mixed pointwise constraints*, Set-Valued Var. Anal. **25** (2017), no. 1, 177–210. MR3606793
[27] K. Krumbiegel and J. Rehberg, *Second order sufficient optimality conditions for parabolic optimal control problems with pointwise state constraints*, SIAM J. Control Optim. **51** (2013), no. 1, 304–331. MR3032877
[28] J.-L. Lions, *Optimal control of systems governed by partial differential equations*, Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170, Springer-Verlag, New York-Berlin, 1971. MR0271512
[29] Francesco Ludovici, Ira Neitzel, and Winnifried Wollner, *A priori error estimates for state-constrained semilinear parabolic optimal control problems*, J. Optim. Theory Appl. **178** (2018), no. 2, 317–348. MR3825627
[30] Francesco Ludovici and Winnifried Wollner, *A priori error estimates for a finite element discretization of parabolic optimization problems with pointwise constraints in time on mean values of the gradient of the state*, SIAM J. Control Optim. **53** (2015), no. 2, 745–770. MR3323538

[31] Dominik Meidner, Rolf Rannacher, and Boris 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), no. 5, 1961–1997. MR2837507

[32] H. Meinlschmidt, C. Meyer, and J. Rehberg, *Optimal control of the thermistor problem in three spatial dimensions, Part 1: Existence of optimal solutions*, SIAM J. Control Optim. **55** (2017), no. 5, 2876–2904. MR3702856

[33] ———, *Optimal control of the thermistor problem in three spatial dimensions, Part 2: Optimality conditions*, SIAM J. Control Optim. **55** (2017), no. 4, 2368–2392. MR3682174

[34] Hannes Meinlschmidt and Joachim Rehberg, *Hölder-estimates for non-autonomous parabolic problems with rough data*, Evol. Equ. Control Theory **5** (2016), no. 1, 147–184. MR3485929

[35] Pierre Portal and Željko Štrkalj, *Pseudodifferential operators on Bochner spaces and an application*, Math. Z. **253** (2006), no. 4, 805–819. MR2221100

[36] Jan Prüss, *Maximal regularity for evolution equations in L_p-spaces*, Conf. Semin. Mat. Univ. Bari **285** (2002), 1–39 (2003). MR1988408

[37] Luke G. Rogers, *Degree-independent Sobolev extension on locally uniform domains*, J. Funct. Anal. **235** (2006), no. 2, 619–665. MR2225465

[38] N. H. Son, B. T. Kien, and A. Rösch, *Second-order optimality conditions for boundary control problems with mixed pointwise constraints*, SIAM J. Optim. **26** (2016), no. 3, 1912–1943. MR3549887

[39] Hans Triebel, *Interpolation theory, function spaces, differential operators*, Second, Johann Ambrosius Barth, Heidelberg, 1995. MR1328645

[40] Fredi Tröltzsch, *Optimal control of partial differential equations*, Graduate Studies in Mathematics, vol. 112, American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels. MR2583281

[41] Gerd Wachsmuth, *A guided tour of polyhedric sets: basic properties, new results on intersections, and applications*, J. Convex Anal. **26** (2019), no. 1, 153–188. MR3847219

[42] ———, *No-gap second-order conditions under n-polyhedric constraints and finitely many nonlinear constraints*, J. Convex Anal. **27** (2020), no. 2, 735–753. MR4058044