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
(revised version as of October 2021)

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 on optimal control of a quasilinear parabolic partial differential equation (PDE) with pointwise control-constraints, and additional constraints on the state variable. We prove existence of optimal controls and derive first-order necessary optimality conditions (FONs) 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 PDEs has been subject to research for several years, see e.g. [**29**, **47**]. Problems with pointwise state-constraints are particularly challenging, since one usually needs continuity of the state to fulfill a Slater-type constraint-qualification. This yields low regularity of the Lagrange multipliers, see [**6**, **7**]. For problems with nonlinear PDEs, SSCs are important because FONs are not sufficient in general. We refer e.g. to the survey [**13**] and references therein for an overview on different aspects of the topic, as well as to [**23**], to our knowledge the first contribution to SSCs for PDE-constrained optimization, and point out only a few more particular aspects. A difficulty arising in the second-order analysis of PDE-constrained optimization is the two-norm discrepancy [**12**, **32**]: differentiability of the reduced functional and coercivity of its second derivative often only hold w.r.t. different norms. 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 [**10**] or purely timedependent controls [**18**]. In [**35**] the authors obtain SSCs for semilinear parabolic PDEs in space dimension 2 and 3

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

and distributed control via a careful analysis utilizing the concept of maximal parabolic regularity. Regarding second-order necessary optimality conditions (SNCs) for pure state-constraints we only mention [**33**], as well as both SNCs and SSCs with emphasis on a possibly small gap between them in [**34, 45**] for the different setting of pointwise mixed control-state-constraints. Finally, we cite [**15, 48, 49**] for SNCs and SSCs in an abstract optimization-theoretic setting. For control problems with quasilinear PDEs, we restrict our overview to the parabolic case: Early results [**21, 22**] include existence of optimal controls and FONs for a problem with averaged-in-space and pointwise in time, or finitely many state-constraints of integral-type, respectively, yet under more restrictive regularity assumptions than in our paper. Well-posedness of the state equation and existence of optimal controls under rather general regularity assumptions on domain and coefficients has been proven in [**41**]. First- and second-order analysis has been carried out in [**4**], and convergence of the SQP method applied to the respective optimization problem has been shown in [**31**]. Optimality conditions for a similar problem with slightly more regular coefficients and domain, but unbounded nonlinearities, have been analyzed in [**8**]. In the same setting finite element discretization error estimates for the state equation have been derived in [**9**]. Optimal control of the thermistor problem, a coupled system consisting of a quasilinear parabolic and a nonlinear elliptic equation, is addressed in [**39, 40**]. We refer to the references of all mentioned papers for a more detailed history of the area. The non-trivial existence and regularity theory for solutions of the underlying PDEs poses the main difficulty of such problems.

In this paper, we first establish existence of optimal controls and FONs in the presence of state-constraints, extending the results from [**4**] and [**8**]. In particular, the assumptions from [**4, 41**] are fairly general, and include certain types of nonsmooth domains, mixed boundary conditions, and nonsmooth coefficients. One of our goals is to investigate how far first- and second-order analysis of the problem can be performed within this "rough" setting, before regularity requirements force us to switch to a different setup. Second, to our best knowledge, SSCs for state-constrained optimal control of a quasilinear parabolic PDE have not been addressed in the 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 [**12**] towards the inclusion of state-constraints, we pay particular attention on avoiding the two-norm gap.

The paper is organized as follows: In Section 2 we introduce the problem setting, prove existence of optimal controls, and derive FONs in this rather general context. In Section 3 we prove SSCs for an abstract optimization problem extending the result from [**12**]. In Section 4 we explain why our abstract result from Section 3 does not apply to the model problem as stated in Section 2. Then, we prove SSCs without 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 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 comment on changes in the case without control-constraints.

Notation. Given an interval $I \subset \mathbb{R}$ and a domain $\Omega \subset \mathbb{R}^d$ we denote the space time cylinder by $Q := I \times \Omega$. We apply standard notation for Hölder-, (Bochner-) Lebesgue- and (Bochner-)Sobolev-spaces as e.g. in [**4**]. The conjugate exponent of some integrability exponent p is denoted by p'. Since Ω stays fixed we omit it when refering to function spaces on Ω. For interpolation spaces we use standard notation, see e.g. [**3, 46**]. The domain of a densely defined operator A: $X \to Y$ between two Banach spaces X,Y, equipped with the graph norm, is denoted by $\mathrm{Dom}_X(A)$, and $\mathcal{L}(X,Y)$ is the space of bounded linear operators $X \to Y$ with the operator norm.

2. Existence of optimal controls and FONs

We introduce the model problem, state our assumptions, and collect some results from [**4**]. Following standard techniques, we derive existence of optimal controls and FONs.

2.1. Model Problem and Assumptions.

We consider the problem:

$$\text{(OCP)} \qquad \min_{y\in Y_{\mathrm{ad}},\ u\in U_{\mathrm{ad}}} J(y,u) := \frac{1}{2}\|y-y_d\|^2_{L^2(I\times\Omega)} + \frac{\gamma}{2}\|u\|^2_{L^2(\Lambda)}, \quad \text{s.t. (SE)},$$

with the quasilinear parabolic state equation (SE) given by

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

The quasilinear differential operator $\mathcal{A}$ is defined as $\mathcal{A}(y) := -\nabla\cdot\xi(y)\mu\nabla$, and boundary conditions are incorporated in the right-hand-side of (SE) and the function spaces. Boundary conditions, control space $L^s(\Lambda)$, and operator B, as well as the set of admissible controls $U_{\mathrm{ad}} \subset L^s(\Lambda)$ are introduced precisely below. The set of admissible states is clarified in each section, and given either by pointwise in space and time inequality-constraints, i.e. $Y_{\mathrm{ad}} = \{y \in C(\overline{Q})\colon\ y_a(t,x) \le y(t,x) \le y_b(t,x)\ \forall (t,x) \in \overline{Q}\}$, or, if we require a weaker type of constraints for our analysis, by pointwise in space and averaged-in-time bounds of type $Y_{\mathrm{ad}} = \{y \in L^1(I, C(\overline{\Omega}))\colon\ y_a(x) \le \int_I y(t,x)\ \mathrm{d}t \le y_b(x)\ \forall x \in \overline{\Omega}\}$. 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 [**41**]:

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 part of $\partial\Omega$, where homogeneous Dirichlet boundary conditions are prescribed. By a subscript D we indicate that the homogeneous Dirichlet boundary conditions on Γ_D are incorporated in the respective function space. Let $\Omega\cup\Gamma_N$ be Gröger regular [**25**] 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 ξ: $\mathbb{R} \to \mathbb{R}$ is twice differentiable with ξ'' being Lipschitz continuous on bounded subsets of $\mathbb{R}$. Let μ: $\Omega \to \mathbb{R}^{d\times d}$ be measurable and uniformly bounded and coercive in the following sense: $0 < \mu_\bullet := \inf_{x\in\Omega}\inf_{z\in\mathbb{R}^d\setminus\{0\}} \frac{z^T\mu(x)z}{z^Tz}$, $\mu^\bullet := \sup_{x\in\Omega}\sup_{1\le i,j\le d}|\mu_{i,j}(x)| < \infty$. We assume a coercivity condition $0 < \xi_\bullet \le$

$\xi \le \xi^\bullet$ for ξ as well. With this we define as above

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

3. We assume that there is $p \in (d,4)$ such that $-\nabla\cdot\mu\nabla + 1$: $W_D^{1,p} \to W_D^{-1,p}$ is a topological isomorphism and fix this choice of p.

4. Let $s > 2$ be fixed such that $\frac{1}{s} < \frac{1}{2}(1-\frac{d}{p})$ holds. For a measure space (Λ,ρ) we define the control space $U := L^s(\Lambda)$ and the admissible set $U_{\mathrm{ad}} = \{u \in L^s(\Lambda)\colon u_a(x) \le u(x) \le u_b(x) \quad \text{for a.a. } x \in \Lambda\}$ with $u_a, u_b \in L^\infty(\Lambda)$, $u_a \le u_b$ almost everywhere. The control operator B: $U = L^s(\Lambda) \to L^s(I, W_D^{-1,p})$ is bounded linear and admits a bounded linear extension B: $L^2(\Lambda) \to L^2(I, W_D^{-1,p})$. 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 Assumption 2.1.4, we only suppose 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], which we repeat for convenience:

Example 2.2. 1. Distributed control: It holds $\Lambda = Q$, i.e. $U = L^s(I\times\Omega)$, and B is the identity map $L^s(Q) \to L^s(I,W_D^{-1,p})$. Denoting the outer normal unit vector of $\partial\Omega$ by n_Ω, the state equation reads

$$\partial_t y + \mathcal{A}(y)y = u \quad \text{on } Q, \qquad n_\Omega\cdot\xi(y)\mu\nabla y = 0 \quad \text{on } I\times\Gamma_N, \qquad y = 0 \quad \text{on } \Gamma_D.$$

2. Neumann boundary control ($d = 2$): We choose $\Lambda = I\times\Gamma_N$, i.e. $U = L^s(I\times\Gamma_N)$, and $B = \mathrm{tr}^*$ where tr: $L^{s'}(I,W_D^{1,p'}) \to L^{s'}(I\times\Gamma_N)$ denotes the trace map. With this the state equation reads

$$\partial_t y + \mathcal{A}(y)y = 0 \quad \text{on } Q, \qquad n_\Omega\cdot\xi(y)\mu\nabla y = u \quad \text{on } I\times\Gamma_N, \qquad y = 0 \quad \text{on } \Gamma_D.$$

3. Purely timedependent controls ($d = 2,3$): We fix $b_1,\dots,b_m \in W_D^{-1,p}$, set $U = L^s(I,\mathbb{R}^m)$, and define $Bu := \sum_{i=1}^m u_i b_i$. If, for instance, $b_i = \mathrm{tr}^* f_i$ with $f_i \in L^s(\Gamma_N)$ where tr: $W_D^{1,p'} \to L^{s'}(\Gamma_N)$ denotes the trace map on Γ_N, we obtain as state equation:

$$\partial_t y + \mathcal{A}(y)y = 0 \quad \text{on } Q, \qquad n_\Omega\cdot\xi(y)\mu\nabla y = \sum_{i=1}^m u_i f_i \quad \text{on } I\times\Gamma_N, \qquad y = 0 \quad \text{on } \Gamma_D.$$

A similar construction applies to $b_i \in L^s(\Omega)$. Of course, adding a sufficiently regular, fixed nonhomogeneous Neumann boundary condition or distributed source in Example 2.2.1 or 2, respectively, is possible.

Assumptions 2.1.1 and 2.1.3 impose non-trivial conditions on the geometry of the domain, the elliptic operator $-\nabla\cdot\mu\nabla+1$, and the boundary conditions. Hence, we mention the following examples, cf. also [**4**, Remarks 2.1 and 2.3]:

Example 2.3. 1. Assumption 2.1.1 is fulfilled for any domain with a Lipschitz boundary ("strong Lipschitz domain", [**24**, Definition 1.2.1.1]) in case $\Gamma_N = \emptyset$ or $\Gamma_N = \partial\Omega$, cf. [**28**, Remark 3.3]. There are also domains without Lipschitz boundary fulfilling this assumption, e.g. a pair of crossing beams in 3D [**28**, Section 7.3]. Moreover, if Ω is a bounded domain with Lipschitz boundary, $\Gamma_N = \emptyset$ or $\Gamma_N = \partial\Omega$, and $\mu\colon \Omega \to \mathbb{R}^{d\times d}$ is symmetric-valued and uniformly, Assumption 2.1.3 is fulfilled with some $p > 3$, see [**20**, Theorem 3.12, Remark 3.17]. Therefore, Assumption 2.1 covers the classical "regular" setting of domains with Lipschitz-boundary in dimensions $d = 2, 3$ with pure Dirichlet or Neumann boundary conditions and symmetric, uniformly continuous coefficient μ.

2. The work [**25**] shows that the isomorphism property from Assumption 2.1.3 for some $p > 2$ is a consequence of Assumption 2.1.1 for any coefficient μ fulfilling Assumption 2.1.2. This is also true under more general assumptions on the domain, see [**27**]. Hence, for space dimension $d = 2$ Assumption 2.1 is guaranteed for a broad range of nonsmooth domains, mixed boundary conditions, and nonsmooth μ.

3. It is well-known that for mixed boundary conditions Assumption 2.1.3 can only be expected to hold for some $p < 4$ in general. In [**19**] for instance several real-world constellations in dimension $d = 3$ have been described that fulfill Assumption 2.1.1 and 2.1.3. Two crossing beams in 3D, e.g., equipped with constant μ and pure homogeneous Dirichlet or Neumann boundary conditions, fulfill Assumption 2.1.

Finally, note that Assumptions 2.1.1-2.1.3 are identical to Assumptions 1-3 of [**4**], i.e. the suppositions w.r.t. domain, coefficients, and boundary conditions remain unchanged. We only modify the assumptions w.r.t. the initial condition and regularity of the right-hand-side of (SE): Assumption 4 in [**4**] is related to the improved regularity analysis on Bessel-potential spaces. As pointed out in [**4**, Section 3] this analysis is not required for the first- and second-order analysis of Sections 3.1 and 4.1-4.3 of [**4**], except for [**4**, Proposition 4.7], a result concerning improved regularity of the adjoint state. We only rely on those results that are obtained completely within the $W^{-1,p}_D$-$W^{1,p}_D$-setting described in our Assumption 2.1.4, cf. also [**41**, Theorem 5.3], and do not include the improved regularity assumptions of [**4**].

2.2. Control-to-state map and reduced functional. For later reference we recall some results from [**4**]. Due to [**4**, Proposition 3.5] (see also [**41**, Corollary 5.8]) the solution map of the equation

$$\partial_t y + \mathcal{A}(y)y = v, \qquad y(0) = y_0, \tag{2.1}$$

defined by $y := \tilde{S}(v)$ if and only if (2.1) holds, is a well-defined map $\tilde{S}\colon L^s(I, W^{-1,p}_D) \to W^{1,s}(I, W^{-1,p}_D) \cap L^s(I, W^{1,p}_D)$. Hereby, $y \in W^{1,s}(I, W^{-1,p}_D) \cap L^s(I, W^{1,p}_D)$ is said to be a solution of (2.1) if and only if

$$\langle \partial_t y, \varphi\rangle_{W^{-1,p}_D, W^{1,p'}_D} + \int_\Omega \xi(y(t))\mu\nabla y(t)\nabla\varphi \,\mathrm{d}x = \langle v(t), \varphi\rangle_{W^{-1,p}_D, W^{1,p'}_D}$$

for all $\varphi \in W^{1,p'}_D$ and almost all $t \in I$, and $y(0) = y_0$ in $(W^{-1,p}_D, W^{1,p}_D)_{1/s',s}$. For well-definedness of $y(0) \in (W^{-1,p}_D, W^{1,p}_D)_{1/s',s}$ we refer e.g. to [**3**, Theorem III.4.10.2]. By composition with B we obtain the control-to-state map $S\colon L^s(\Lambda) \to W^{1,s}(I, W^{-1,p}_D) \cap L^s(I, W^{1,p}_D)$, $u \mapsto \tilde{S}(Bu)$. Given $y \in W^{1,s}(I, W^{-1,p}_D) \cap L^s(I, W^{1,p}_D)$ we recall from [**4**] the notation for the derivatives of the nonlinear term, stated in weak form:

$$\langle \mathcal{A}'(y)v, \varphi\rangle := \int_Q \xi'(y) v \mu \nabla y \nabla \varphi \; \mathrm{d}x\mathrm{d}t,$$

$$\langle \mathcal{A}''(y)[v_1, v_2], \varphi\rangle := \int_Q (\xi'(y)(v_1 \mu \nabla v_2 + v_2 \mu \nabla v_1) + \xi''(y) v_1 v_2 \mu \nabla y) \nabla \varphi \; \mathrm{d}x\mathrm{d}t,$$

with $v, v_1, v_2 \in W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$ and a test function $\varphi \in L^{s'}(I, W_D^{1,p'})$. It is possible to relax the regularity requirements on v, v_1, v_2, as done e.g. in the proof of the following differentiability properties of G [**4**, Proposition 4.4 and Lemma 4.5]:

Lemma 2.4. *Let Assumption 2.1 be satisfied. 1. The map $\tilde{S}$: $L^s(I, W_D^{-1,p}) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$ is twice continuously Fréchet differentiable with derivatives $\tilde{S}'(v)h = w$ and $\tilde{S}''(v)[h_1, h_2] = z$ given by the unique solutions of*

$$\partial_t w + \mathcal{A}(y)w + \mathcal{A}'(y)w = h, \qquad w(0) = 0, \tag{2.2}$$

$$\partial_t z + \mathcal{A}(y)z + \mathcal{A}'(y)z = \mathcal{A}''(y)[G'(v)h_1, G'(u)h_2], \qquad z(0) = 0 \tag{2.3}$$

for $y = \tilde{S}(v)$, respectively.

2. The nonautonomous operator $\mathcal{A}(y) + \mathcal{A}'(y)$ exhibits maximal parabolic regularity on $L^r(I, W_D^{-1,p})$ for $r \in (1, s]$. It holds $\tilde{S}'(v) \in \mathcal{L}(L^r(I, W_D^{-1,p}), W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}))$ for all $v \in L^s(I, W_D^{-1,p})$, $r \in (1, s]$.

We introduce the reduced objective functional j: $L^s(\Lambda) \to \mathbb{R}$, $u \mapsto J(S(u), u)$. From [**4**, Lemma 4.6] we recall that the reduced functional j is twice continuously Fréchet differentiable on $L^s(\Lambda)$ with gradient

$$\nabla j(u) = B^* \tilde{S}'(Bu)^*(y - y_d) + \gamma u. \tag{2.4}$$

A more detailed discussion of the operators B^* and $\tilde{S}'(Bu)^*$ will be provided later on, see in particular Sections 2.4 and 4.1.

2.3. Existence of optimal controls.

Let us now consider the following setting of the state-constraints, which remains unchanged until noted otherwise:

Assumption 2.5. 1. The set of admissible states is given by $Y_{\mathrm{ad}} = \{y \in C(\overline{Q}) : y_a(t,x) \le y(t,x) \le y_b(t,x) \; \forall (t,x) \in \overline{Q}\}$, with bounds $y_a, y_b \in C(\overline{Q})$ satisfying $y_a(t,x) < y_b(t,x)$ for all $(t,x) \in \overline{Q}$, $y_a(t,x) < 0 < y_b(t,x)$ for all $(t,x) \in \overline{I} \times \Gamma_D$, and $y_a(0,x) < y_0(x) < y_b(0,x)$ for $x \in \overline{\Omega}$. We allow for $y_a \equiv -\infty$ or $y_b \equiv +\infty$.

2. There is a feasible point, i.e. there is $(y, u) \in Y_{\mathrm{ad}} \times U_{\mathrm{ad}}$ 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] and [**41**, Proposition 6.4]. An analogous result for the state-constrained thermistor problem has already been obtained in [**39**].

Theorem 2.6. *Let Assumptions 2.1 and 2.5 hold. Then there exists a globally optimal control $\bar{u} \in U_{\mathrm{ad}}$ for the optimal control problem* (OCP).

Proof. The proof follows standard arguments in the calculus of variations, cf. [**4, 41**], [**29, 47**]. In particular, note that existence of an infimizing sequence is provided by the existence of a feasible point (Assumption 2.5.2). In the proof of [**41**, Proposition 6.4] it is shown that a subsequence of the corresponding sequence of states converges in $C(\overline{Q})$ to the optimal state. Note that the presence of an additional linear term in [**41**] is not essential for the respective argument which therefore can be adapted to the present case as in [**4**]. For details we refer to Appendix A. Since Y_{ad} is closed in $C(\overline{Q})$ the limit is still in Y_{ad}, i.e. it fulfills the state-constraints. $\square$

For comments on how Theorem 2.6 and all further results change in the case without or with only unilateral control-constraints we refer to Section 6.

2.4. First-order necessary optimality conditions. We now characterize local solutions of (OCP) that fulfill a Slater-type constraint-qualification by FONs. The following main difficulty is well-known in state-constrained optimal control of PDEs, cf. e.g. [**6**, **7**]: To use a Slater-type condition, we have to ensure nonempty interior of Y_{ad} in the respective space. Since Y_{ad} is defined by pointwise inequality-constraints, this excludes states in $L^q(Q)$, $1 \le q < +\infty$. We have to consider them in $C(\overline{Q})$, which infers regular Borel measures, i.e. the corresponding dual objects, in the KKT-system. To apply an abstract result for optimization problems in Banach spaces [**7**, Theorem 5.2] to our problem (OCP) we formulate an additional assumption.

Assumption 2.7. Let $\bar{u} \in U_{\text{ad}}$ be an $L^2(\Lambda)$-local solution to (OCP) with associated state $\bar{y} = S(\bar{u}) \in Y_{\text{ad}}$, i.e. there is $\epsilon > 0$ such that $j(u) \ge j(\bar{u})$ for all $u \in \mathbb{B}_\epsilon^{L^2(\Lambda)}(\bar{u}) \cap U_{\text{ad}}$ fulfilling $S(u) \in Y_{\text{ad}}$. Further, assume that the following linearized Slater-condition is fulfilled at $\bar{u}$: There is $u_{\text{Sl}} \in U_{\text{ad}}$ such that $\bar{y} + S'(\bar{u})(u_{\text{Sl}} - \bar{u}) \in \mathring{Y}_{\text{ad}}$, i.e. $y_a(t,x) < \bar{y}(t,x) + S'(\bar{u})(u_{\text{Sl}} - \bar{u})(t,x) < y_b(t,x)$ for all $(t,x) \in \overline{Q}$.

Since the L^s- is stronger than the L^2-norm, a $L^2(\Lambda)$-local is a $L^s(\Lambda)$-local solution.

Theorem 2.8. *Under Assumptions 2.1 and 2.7 and Assumption 2.5.1 there exists a regular Borel measure $\bar{\lambda} \in \mathcal{M}(\overline{Q}) = C(\overline{Q})^*$ 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*

$$\partial_t \bar{y} + \mathcal{A}(\bar{y})\bar{y} = B\bar{u}, \qquad \bar{y}(0) = y_0, \tag{2.5}$$

$$-\partial_t \bar{p} + \mathcal{A}(\bar{y})^*\bar{p} + \mathcal{A}'(\bar{y})^*\bar{p} = \bar{y} - y_d + \bar{\lambda}, \qquad \bar{p}(T) = 0, \tag{2.6}$$

$$\langle \bar{\lambda}, y - \bar{y} \rangle_{\mathcal{M}(\overline{Q}), C(\overline{Q})} \le 0 \qquad \text{for all } y \in Y_{\text{ad}}, \tag{2.7}$$

$$\langle B^*\bar{p} + \gamma\bar{u}, u - \bar{u} \rangle_{L^{s'}(\Lambda), L^s(\Lambda)} \ge 0 \qquad \text{for all } u \in U_{\text{ad}} \tag{2.8}$$

is satisfied. The so-called adjoint equation (2.6) *has to be understood in the sense outlined in the proof below, cf. also Remark 2.10.*

Before going into the details of the proof we state the specific form of the variational inequality (2.8) for the three variants of B discussed in Example 2.2:

Example 2.9. 1. In case of distributed control we obtain B^* to be the identity map $L^{s'}(I, W_D^{1,p'}) \to L^{s'}(Q)$, and (2.8) reads $\int_Q (\bar{p} + \gamma\bar{u})(u - \bar{u}) \mathrm{d}x\mathrm{d}t \ge 0$, $\forall u \in U_{\text{ad}} \subset L^s(Q)$.

2. For Neumann boundary control ($d=2$), B^* is the trace map $L^{s'}(I, W_D^{1,p'}) \to L^{s'}(I \times \Gamma_N)$ and we obtain $\int_{I\times\Gamma_N}(\bar{p}|_{I\times\Gamma_N} + \gamma\bar{u})(u-\bar{u})\mathrm{d}s\mathrm{d}t \geq 0,\ \forall u \in U_{\mathrm{ad}} \subset L^s(I\times\Gamma_N)$.

3. We obtain $B^*\bar{p} = (t \mapsto \langle b_i, \bar{p}(t)\rangle_{W_D^{-1,p}, W_D^{1,p'}})_{i=1}^m = (t \mapsto \int_{\Gamma_N} f_i \bar{p}(t)|_{\Gamma_N}\mathrm{d}s)_{i=1}^m \in L^{s'}(I, \mathbb{R}^m)$, and $\sum_{i=1}^m \int_I (\int_{\Gamma_N} f_i \cdot \bar{p}(t)|_{\Gamma_N}\mathrm{d}s + \gamma\bar{u}_i(t))(u_i(t) - \bar{u}_i(t))\mathrm{d}t \geq 0$ for all $u \in U_{\mathrm{ad}} \subset L^s(I, \mathbb{R}^m)$ in case of purely timedependent controls.

Proof of Theorem 2.8. In [**7**, Theorem 5.2] choose $U = L^s(\Lambda)$, $Z = C(\overline{Q})$, $J = j$, $G = S$, $K = U_{\mathrm{ad}}$ and $C = Y_{\mathrm{ad}}$. 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 $C(\overline{Q})$. It holds $\tilde{S}'(B\bar{u}) \in \mathcal{L}(L^r(I, W_D^{-1,p}), W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}))$ for any $r \in (1, s]$, cf. Lemma 2.4.2. Employing $W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}) \hookrightarrow C(\overline{Q})$ for $r \in (\frac{2p}{p-d}, \infty)$ [**4**, Proposition 3.3] we obtain $\tilde{S}'(B\bar{u}) \in \mathcal{L}(L^r(I, W_D^{-1,p}), C(\overline{Q}))$ for those r, and consequently

$$\tilde{S}'(B\bar{u})^* \in \mathcal{L}(\mathcal{M}(\overline{Q}), L^{r'}(I, W_D^{1,p'})) \tag{2.9}$$

for all $r' \in (1, \frac{2p}{p+d})$. Following the usual adjoint technique in optimal control, see e.g. [**47**, Chapter 6.2.1], we introduce the adjoint state $\bar{p} := \tilde{S}'(B\bar{u})^*(\bar{y} - y_d + \bar{\lambda})$. Note that $\bar{p}$ is well-defined in this way and exhibits the regularity stated in the theorem due to (2.9) and $\bar{y} - y_d + \bar{\lambda} \in \mathcal{M}(\overline{Q})$. The adjoint equation (2.6) has to be understood purely formal, in the very-weak/adjoint sense. We discuss this further in Remark 2.10 below. Combining equation (2.4) for the reduced gradient of our particular setting with the abstract variational inequality [**7**, (5.3)] and the definition of $\bar{p}$ yields (2.8). $\square$

Remark 2.10. The adjoint equation (2.6) has to be understood purely formal: In general, it is not guaranteed that $\bar{p} \in L^r(I, W_D^{1,p'})$ has a distributional time derivative or a well-defined trace on $\{T\} \times \Omega$. Hence, (2.6) really only serves as a more illustrative and intuitive notation for the precise definition of $\bar{p}$ given by $\bar{p} = \tilde{S}'(B\bar{u})^*(\bar{y} - y_d + \bar{\lambda})$. The notation as backward parabolic PDE is motivated by the fact that $\tilde{S}'(B\bar{u})^*$ restricted to the spaces $L^{r'}(I, W_D^{-1,p'})$, $r' \in (1, \infty)$, can be identified with the solution map of the respective backward nonautonomous parabolic PDE, cf. [**4**, Proposition 4.7]. Moreover, the presence of mixed boundary conditions in the state equation does not pose additional difficulties, see e.g. [**26**, **30**, **35**], in particular because the support of $\bar{\lambda}$ is disjoint from $\overline{I} \times \Gamma_D$ and $\{0\} \times \overline{\Omega}$, cf. Remark 2.11 below and Assumption 2.5.1.

Remark 2.11. Condition (2.7) can be rewritten in a more illustrative way: The Jordan decomposition $\bar{\lambda} = \bar{\lambda}^+ - \bar{\lambda}^-$ into non-negative measures $\bar{\lambda}^+, \bar{\lambda}^- \geq 0$ satisfies $\operatorname{supp} \bar{\lambda}^+ \subset \{(t,x) \in \overline{Q}\colon \bar{y}(t,x) = y_b(t,x)\}$, $\operatorname{supp} \bar{\lambda}^- \subset \{(t,x) \in \overline{Q}\colon \bar{y}(t,x) = y_a(t,x)\}$. For a proof we refer e.g. to [**11**, Proposition 2.5].

Remark 2.12. Because $\bar{\lambda}$ 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 [**11**] and parabolic [**14**] PDEs.

Due to $\frac{2p}{p+d} < 2$ and $p' < 2$ Theorem 2.8 shows rather poor temporal and spatial regularity for $\bar{p}$. This is a typical difficulty to overcome during the analysis of second-order optimality conditions for (OCP), as we will outline in Sections 4 and 5. To do so, we will either have to modify the type of state-constraints (Assumption 4.1) or assume a more regular setting for the state equation (Assumption 5.1).

3. An abstract result on SSCs

We extend the abstract framework of [**12**] towards inclusion of state-constraints, i.e. we give SSCs for an abstract optimization problem similar to the one from [**7**, Theorem 5.2], but now enriched with two norms as typical for PDE-constrained optimization. However, we prove SSCs that avoid the two-norm gap. The framework is developed having in particular the setting and the arguments from [**18**] in mind. We start by introducing the abstract problem

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

with the assumptions given below. The suppositions on the real-valued functional j and the underlying spaces U_2, U_∞, respectively, are identical to those from [**12**]. Here we extend this work towards the inclusion of a state-constraint-like constraint of type "$g(u) \in C$" that is formulated in a further Banach space Z. For instance, choosing g to be the control-to-state map allows to handle state-constraints. Since the set $K \cap g^{-1}(C)$ is nonconvex in general, this situation is not covered by the results of [**12**]. Further, j and g are differentiable w.r.t. the U_∞-norm, but not necessarily w.r.t. the weaker U_2-norm. 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 a measure space $(\Sigma, \mathrm{d}m)$. The presence of such two norms goes back to [**32**], see also the exposition in [**12**, **13**, **47**].

We briefly put our result into context: As far as we know, prior results on SSCs for state-constraints without two-norm gap required differentiability of j and g w.r.t. L^2, cf. [**45**, Section 4], [**10**, Theorem 4.3] – an assumption that can be avoided in our result. In particular we can state SSCs for the same semilinear parabolic optimal control problem as in [**18**], but without norm gap, see Example 3.3 below. In [**15**] both SNCs and SSCs for certain optimization problems in infinite dimensions are proven. The results rely on the concept of a directional curvature functional for the (possibly nonconvex) 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 [**12**], 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, explicitly tailored to situations as e.g. (OCP), [**18**], and [**12**], respectively, 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 $\|\cdot\|_2$ and $\|\cdot\|_\infty$ we denote the corresponding norms. Moreover, $\langle \cdot, \cdot \rangle_2$ is the duality product in $U_2^* \times U_2$. 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 $j\colon A \to \mathbb{R}$ is twice continuously Fréchet differentiable w.r.t. $\|\cdot\|_\infty$.

1a. The derivatives of j taken w.r.t. the space U_∞ extend to continuous linear respectively bilinear forms on U_2: $j'(u) \in \mathcal{L}(U_2, \mathbb{R})$, $j''(u) \in \mathcal{L}(U_2 \otimes U_2, \mathbb{R})$ for $u \in A$.

1b. Let $(u_k)_k \subset K$, $(v_k)_k \subset U_2$ be arbitrary sequences such that $u_k \to \bar{u}$ strongly w.r.t. 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. $j''(\bar{u})v^2 \leq \liminf_{k\to\infty} j''(u_k)v_k^2$

1biii. If $v = 0$, there is $\gamma > 0$ such that $\gamma \liminf_{k\to\infty} \|v_k\|_2^2 \leq \liminf_{k\to\infty} j''(u_k)v_k^2$.

2. Let $g\colon A \to Z$ be twice continuously Fréchet differentiable w.r.t. $\|\cdot\|_\infty$.

2a. The derivatives of g taken w.r.t. U_∞ extend to continuous linear respectively bilinear forms on U_2: $g'(u) \in \mathcal{L}(U_2, Z)$, $g''(u) \in \mathcal{L}(U_2 \otimes U_2, Z)$ for $u \in A$.

2b. Let $(u_k)_k \subset K$, $(v_k)_k \subset U_2$ be arbitrary sequences such that $u_k \to \bar{u}$ strongly w.r.t. 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 is our main abstract result and extends [**12**, Theorem 2.3] towards the inclusion of a state-constraint-like constraint of type "$g(u) \in C$". We denote by $\mathcal{R}(S, x)$ and $\mathcal{T}_S(x)$ the radial cone and the contigent cone, respectively, of a closed convex set S in a Banach space X at some $x \in S$, see e.g. [**5**, Definition 2.54].

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

$$\langle j'(\bar{u}) + g'(\bar{u})^*\bar{\zeta}, u - \bar{u}\rangle_2 \geq 0 \qquad \forall u \in K, \tag{3.1}$$

$$\langle \bar{\zeta}, z - g(\bar{u})\rangle_{Z^*,Z} \leq 0 \qquad \forall z \in C, \tag{3.2}$$

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

$$j''(\bar{u})v^2 + \langle \bar{\zeta}, g''(\bar{u})v^2\rangle_{Z^*,Z} > 0 \qquad \forall v \in C_{\bar{u}} \setminus \{0\} \tag{3.3}$$

with $C_{\bar{u}} := \mathrm{cl}_{U_2}(\mathcal{R}(K, \bar{u})) \cap \{v \in U_2\colon\ j'(\bar{u})v = 0,\ \langle g'(\bar{u})^*\bar{\zeta}, v\rangle_2 = 0,\ g'(\bar{u})v \in \mathcal{T}_C(g(\bar{u}))\}$. 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).

In the theorem and its proof we make extensive use of the continuation properties from Assumption 3.1.1a and 3.1.2a: In formula (3.1), for instance, $g'(\bar{u})^*\bar{\zeta} \in U_2^*$ is well-defined because of $g'(\bar{u}) \in \mathcal{L}(U_2, Z)$ by Assumption 3.1.2a. We follow the the proof of [**12**, Theorem 2.3], and abstract the techniques of several similar results in this context, see e.g. [**10, 34, 45**], and, in particular, [**18**].

Proof. Assume the contrary, i.e. there exist $(u_k)_k \subset K$ such that $\|u - u_k\|_2 < \frac{1}{k}$, $j(u_k) < j(\bar{u}) + \frac{1}{2k}\|u_k - \bar{u}\|_2^2$, and $g(u_k) \in C$. 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. From weak convergence and (3.1) we derive immediately: $\langle j'(\bar{u}) + g'(\bar{u})^*\bar{\zeta}, v\rangle_2 = \lim_{k\to\infty}\langle j'(\bar{u}) + g'(\bar{u})^*\bar{\zeta}, v_k\rangle_2 = \lim_{k\to\infty} \frac{1}{\rho_k}\langle j'(\bar{u}) + g'(\bar{u})^*\bar{\zeta}, u_k - \bar{u}\rangle_2 \geq 0$.

B. To show $\langle g'(\bar{u})^*\bar{\zeta}, v\rangle_2 \leq 0$ observe that $\langle \bar{\zeta}, g'(u_k^\theta)v_k\rangle_{Z^*,Z} = \frac{1}{\rho_k}\langle \bar{\zeta}, g'(u_k^\theta)(u_k - \bar{u})\rangle_{Z^*,Z} = \frac{1}{\rho_k}\langle \bar{\zeta}, g(u_k) - g(\bar{u})\rangle_{Z^*,Z} \overset{(3.2)}{\leq} 0$ 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 Assumption 3.1.2bi we obtain $\langle g'(\bar u)^*\bar\zeta, v\rangle_2 = \langle\bar\zeta, g'(\bar u)v\rangle_{Z^*,Z} = \lim_{k\to\infty}\langle\bar\zeta, g'(u_k^\theta)v_k\rangle_{Z^*,Z} \le 0$. Similarly we obtain for arbitrary but fixed $\eta\in Z^*$: $\langle\eta, \frac{1}{\rho_k}(g(u_k)-g(\bar u))\rangle_{Z^*,Z} = \langle\eta, g'(u_k^{\theta,\eta})v_k\rangle_{Z^*,Z} \to \langle\eta, g'(\bar u)v\rangle_{Z^*,Z}$ due to Assumption 3.1.2bi, i.e. $g'(\bar u)v \in$ weak-$\mathrm{cl}_Z(\mathcal{R}(C, g(\bar u))) = \mathcal{T}_C(g(\bar u))$, since C is assumed to be closed and convex: From [**5**, Proposition 2.55] we infer that $\mathcal{T}_C(g(\bar u)) = \mathrm{cl}_Z(\mathcal{R}(C,g(\bar u)))$. 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 [**12**, Theorem 2.3] we find with help of the mean value theorem that $j'(\bar u)v\le 0$ holds. Together with B. we obtain $\langle j'(\bar u)+g'(\bar u)^*\bar\zeta, v\rangle_2 \le 0$ and therefore with A.: $\langle j'(\bar u)+g'(\bar u)^*\bar\zeta, v\rangle_2 = 0$

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

As in [**12**] 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 $\frac{\rho_k^2}{2k} > j(u_k)-j(\bar u) = j'(\bar u)(u_k-\bar u)+\frac12 j''(u_k^\theta)(u_k-\bar u)^2$ with some $u_k^\theta = \theta_k u_k + (1-\theta_k)\bar u$, $(\theta_k)\subset[0,1]$. Exploiting (3.1) and (3.2) it follows

(3.4)
$$\begin{aligned}\frac{\rho_k^2}{2k} &\overset{(3.1)}{>} -\langle\bar\zeta, g'(\bar u)(u_k-\bar u)\rangle_{Z^*,Z} + \frac12 j''(u_k^\theta)(u_k-\bar u)^2 = -\langle\bar\zeta, g(u_k)-g(\bar u)\rangle_{Z^*,Z}\\ &\quad + \langle\bar\zeta, g(u_k)-g(\bar u)-g'(\bar u)(u_k-\bar u)\rangle_{Z^*,Z} + \frac12 j''(u_k^\theta)(u_k-\bar u)^2\\ &\overset{(3.2)}{\ge} \langle\bar\zeta, g(u_k)-g(\bar u)-g'(\bar u)(u_k-\bar u)\rangle_{Z^*,Z} + \frac12 j''(u_k^\theta)(u_k-\bar u)^2\\ &= \frac12\rho_k^2(\langle\bar\zeta, g''(\tilde u_k^\theta)v_k^2\rangle_{Z^*,Z} + j''(u_k^\theta)v_k^2),\end{aligned}$$

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 $\frac1k > \langle\bar\zeta, g''(\tilde u_k^\theta)v_k^2\rangle_{Z^*,Z} + j''(u_k^\theta)v_k^2$ 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 Assumptions 3.1.1bii and 3.1.2bii: $j''(\bar u)v^2 + \langle\bar\zeta, g''(\bar u)v^2\rangle_{Z^*,Z} = 0$. Since (3.3) and $v\in C_{\bar u}$ holds, we conclude $v=0$. Using Assumptions 3.1.1biii at (♣) and 3.1.2bii at (★) we finally arrive at $0<\gamma = \gamma\liminf_{k\to\infty}\|v_k\|_2^2 \overset{(♣)}{\le} \liminf_{k\to\infty} j''(u_k^\theta)v_k^2 \le \liminf_{k\to\infty}(\frac1k - \langle\bar\zeta, g''(\tilde u_k^\theta)v_k^2\rangle_{Z^*,Z}) \overset{(★)}{=} 0$, which is the desired contradiction. $\square$

We briefly indicate how Theorem 3.2 allows to extend a result from the literature:

Example 3.3. The reader may easily verify along the lines of [**12**, **18**] that the semilinear parabolic optimal control problem with pointwise constraints on the state from [**18**] fits into the framework of Assumption 3.1. Therefore, Theorem 3.2 allows to reformulate [**18**, 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 Assumption 3.1 by the stronger convergence $u_k\to\bar u$ in V, where $(V,\|\cdot\|_V)$ is a 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 $j(u)\ge j(\bar u)+\frac{\delta}{2}\|u-\bar u\|_2^2$ 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, $\bar{u}$ is at best a V-local minimizer for (P), which corresponds –on the abstract level– for $V = U_\infty$ to the result of [**18**].

The following example, although of artificial nature, illustrates that the assumptions in the formulation of Theorem 3.2 are necessary. Necessity of the assumptions on j is addressed in [**12**] and hence 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

$$\text{(E)} \quad \min_{u \in L^2([0,1])} j(u) := \int_0^1 u(t)^2 \mathrm{d}t \quad \text{s.t.} \quad -1 \le u(t) \le 1, \ [g(u)](t) \ge t, \ \forall t \in [0,1],$$

with $[g(u)](t) := \int_0^t (1 - \cos(\frac{\pi}{2} u(s)))\mathrm{d}s$. Note that j satisfies Assumption 3.1 and observe that g: $L^2([0,1]) \to C([0,1])$ is well-defined. Yet, 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 verifies that $\bar{u} \equiv 1$ is feasible for (E), and satisfies the FONs (3.1) and (3.2) with $\bar{\zeta} = -\frac{2}{\pi}\delta_1 \in C([0,1])^*$. Herein, δ_1 denotes the Dirac-measure concentrated at $t = 1$. The coercivity condition (3.3) is trivially satisfied at $(\bar{u}, \bar{\zeta})$, because $C_{\bar{u},\bar{\zeta}} = \{0\}$. Further, the second derivative of the functional at $\bar{u}$ is even $L^2([0,1])$-coercive, but 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 (E) 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 shows that a quadratic growth condition around $\bar{u}$ cannot hold. The reason is that Theorem 3.2 cannot be applied, because Assumption 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 $\langle \delta_1, g''(\hat{u}_n)v_n^2 \rangle = \frac{\pi^2}{4} \int_0^1 \cos(\frac{\pi}{2}\hat{u}_n(t))v_n^2(t)\mathrm{d}t = \frac{\pi^2}{4} \not\to 0 = \langle \delta_1, g''(\bar{u})v^2 \rangle$ which disproves Assumption 3.1.2bii. However, due to continuous Fréchet differentiability of g w.r.t. $L^\infty([0,1])$, Remark 3.4 applies: $\bar{u}$ is an $L^\infty([0,1])$-local, but not an $L^2([0,1])$-local solution of (E).

We conclude this section with an open problem. An important property of the SSCs in [**12**, Theorem 2.3] is their minimal gap to corresponding SNCs, if the admissible set K is polyhedric: Positivity of $j''(\bar{u})$ on a certain cone $C_{\bar{u}} \subset U_2$ is –together with FONs– a sufficient optimality condition for $\bar{u}$, while non-negativity of $j''(\bar{u})$ on the *same* cone is necessarily implied by local optimality of $\bar{u}$ [**12**, Theorem 2.2]. Obtaining SNCs for (P) seems to be a challenging topic and is beyond the scope of our paper. For recent results concerning no-gap second-order conditions we refer e.g. to [**34, 45**] in case of optimal control of semilinear elliptic PDEs with mixed control-state constraints, to [**16**] for optimal control of a nonsmooth quasilinear elliptic PDE, or to [**15**] 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 SSCs for (OCP): We replace the pointwise in space and time state-constraints by averaged-in-time state-constraints, see Assumption 4.1 below. For the resulting modified model problem we prove SSCs avoiding the two-norm gap while keeping the rather low regularity requirements on the state equation from Assumption 2.1. Since Assumption 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 Assumption 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 Assumption 3.1.2. This requires a careful regularity analysis of the 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 Assumptions 2.1 and 2.5 directly, and therefore motivate the introduction of averaged-in-time state-constraints. In particular, the analysis of the quasilinear problem (OCP) is quite different from the discussion of the semilinear problem mentioned in Example 3.3 due to the more complicated structure of derivatives of the nonlinearity in the differential operator. This yields slightly better regularity results in the case of semilinear PDEs.

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 is $Y_{\mathrm{ad}} = \{y \in L^1(I, C(\overline{\Omega})) : y_a(x) \leq \int_I y(t,x)\mathrm{d}t \leq y_b(x) \quad \forall x \in \overline{\Omega}\}$, with bounds $y_a, y_b \in C(\overline{\Omega})$ satisfying $y_a(x) < y_b(x)$ for all $x \in \overline{\Omega}$ and $y_a(x) < 0 < y_b(x)$ for all $x \in \Gamma_D$. We allow for $y_a \equiv -\infty$ or $y_b \equiv \infty$.

2. There is a feasible point, i.e. there is $(y,u) \in Y_{\mathrm{ad}} \times U_{\mathrm{ad}}$ 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. 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. Since the latter is only a technicality, we keep the above assumption as simple as possible. Averaged-type instead of purely pointwise constraints are common in the literature, e.g. averaged-in-space and pointwise in time bounds on the state [**21**, **23**, **36**, **38**] or its gradient [**37**]. Existence of an optimal control for (OCP) with averaged-in-time state-constraints is proven analogously to Theorem 2.6. We only state the result.

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

To address FONs we first require a suitable constraint-qualification:

Assumption 4.3. Under Assumption 4.1.1 let $\bar{u} \in U_{\mathrm{ad}}$ be an $L^2(\Lambda)$-local solution to (OCP) with associated state $\bar{y} = S(\bar{u}) \in Y_{\mathrm{ad}}$ such that that the following linearized Slater-condition is fulfilled at $\bar{u}$: There is $u_{\mathrm{Sl}} \in U_{\mathrm{ad}}$ such that $\bar{y} + S'(\bar{u})(u_{\mathrm{Sl}} - \bar{u}) \in \mathring{Y}_{\mathrm{ad}}$, i.e. $y_a(x) < \int_\Omega [\bar{y}(t,x) + S'(\bar{u})(u_{\mathrm{Sl}} - \bar{u})(t,x)]\,\mathrm{d}t < y_b(x)$ for all $x \in \overline{\Omega}$.

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

Theorem 4.4. *Under Assumptions 2.1 and 4.3 and Assumption 4.1.1 there exists a regular Borel measure $\bar{\nu} \in \mathcal{M}(\overline{\Omega}) = C(\overline{\Omega})^*$ 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*

$$\partial_t \bar{y} + \mathcal{A}(\bar{y})\bar{y} = B\bar{u}, \qquad \bar{y}(0) = y_0, \tag{4.1}$$

$$-\partial_t \bar{p} + \mathcal{A}(\bar{y})^*\bar{p} + \mathcal{A}'(\bar{y})^*\bar{p} = \bar{y} - y_d + \mathrm{d}t \otimes \bar{\nu}, \qquad \bar{p}(T) = 0, \tag{4.2}$$

$$\operatorname{supp}(\bar{\nu}^+) \subset \{\int_I \bar{y}(t,\cdot)\mathrm{d}t = y_b\}, \quad \operatorname{supp}(\bar{\nu}^-) \subset \{\int_I \bar{y}(t,\cdot)\mathrm{d}t = y_a\}, \tag{4.3}$$

$$\langle B^*\bar{p} + \gamma\bar{u}, u - \bar{u}\rangle_{L^{s'}(\Lambda),L^s(\Lambda)} \geq 0, \qquad \textit{for all } u \in U_{\mathrm{ad}}, \tag{4.4}$$

is satisfied. Here, $\bar{\nu} = \bar{\nu}^+ - \bar{\nu}^-$ denotes the Jordan-decomposition of $\bar{\nu}$, cf. Remark 2.11, and (4.2) *has to be understood in the sense outlined in the proof.*

Proof. Because the proof is completely analogous to the proof of Theorem 2.8, we mainly comment on the differences w.r.t. the new type of state-constraints: In [**7**, Theorem 5.2] we choose $Z = C(\overline{\Omega})$ and $G := \iota \circ A \circ S$: , where S: $L^s(\Lambda) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$ is the control-to-state map, ι: $W_D^{1,p} \hookrightarrow C(\overline{\Omega})$ the Sobolev embedding, and A: $\varphi \mapsto (x \mapsto \int_I \varphi(t,x)\mathrm{d}t)$ is averaging w.r.t. time, which is a bounded linear map $L^r(I, W_D^{1,p}) \to W_D^{1,p}$ for any $r \in (1,\infty)$. The choice $r = s$ shows that G is well-defined from $L^s(\Lambda)$ into $C(\overline{\Omega})$. From $A \in \mathcal{L}(L^r(I, W_D^{1,p}), W_D^{1,p})$ we conclude $A^* \in \mathcal{L}(W_D^{-1,p'}, L^{r'}(I, W_D^{-1,p'}))$, and ι^* is the embedding $\mathcal{M}(\overline{\Omega}) \hookrightarrow W_D^{-1,p'}$. For a test function $\psi \in L^r(I, W_D^{1,p})$ we compute $\langle A^*\iota^*\bar{\nu}, \psi\rangle_{L^{r'}(I,W_D^{-1,p'}),L^r(I,W_D^{1,p})} = \langle \iota^*\bar{\nu}, A\psi\rangle_{\mathcal{M}(\overline{\Omega}),C(\overline{\Omega})} = \int_{\overline{\Omega}} A\psi \mathrm{d}\bar{\nu} = \int_{\overline{Q}} \psi \mathrm{d}t\mathrm{d}\bar{\nu}$, i.e. $A^*\iota^*\bar{\nu} = \mathrm{d}t \otimes \bar{\nu} \in L^{r'}(I, W_D^{-1,p'})$ for each $r' \in (1,\infty)$. Together with $\tilde{S}'(B\bar{u}) \in \mathcal{L}(L^r(I, W_D^{-1,p}), L^r(I, W_D^{1,p}))$, for $r \in (1,s]$, which follows from Lemma 2.4.2, we find $\tilde{S}'(B\bar{u})^*A^*\iota^* \in \mathcal{L}(\mathcal{M}(\overline{\Omega}), L^{r'}(I, W_D^{1,p'}))$ for $r' \in [s',\infty)$. This shows that $\bar{p} = \tilde{S}'(B\bar{u})^*(\bar{y} - y_d + A^*\iota^*\bar{\nu}) \in L^{r'}(I, W_D^{1,p'})$, $r' \in (1,\infty)$, is well-defined. Equation (4.2) has to be understood in this sense. Finally, a short computation shows that (4.4) holds. ☐

Remark 4.5. Let in addition to Assumption 2.1 the enhanced regularity assumptions from [**4**, Assumption 4] hold that enable the improved regularity analysis of the state on the Bessel-potential space $H_D^{-\zeta,p} \hookrightarrow W_D^{-1,p}$ from [**4**, Theorem 3.20]: More precisely, $y_0 \in (H_D^{-\zeta,p}, \mathrm{Dom}_{H_D^{-\zeta,p}}(-\nabla \cdot \mu\nabla))_{1/s',s}$ and B is a bounded linear map $L^s(\Lambda) \to L^s(I, H_D^{-\zeta,p})$, where $\zeta \in (0,1)$ and $s > 2$ satisfy $\max\{1 - \frac{1}{p}, \frac{d}{p}\} < \zeta$ and $s > \max\{\frac{2}{\zeta - \frac{d}{p}}, \frac{2}{1-\zeta}\}$. The situations from Example 2.2 still fit into this setting, cf. [**4**, Examples 2.4-6], with the minor modification that $b_i \in H_D^{-\zeta,p}$ is required for purely timedependent controls. Under these stricter assumptions, [**4**, Proposition 4.7] shows that $\tilde{S}'(B\bar{u})^*$ coincides with 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 $L^{r'}(I, W^{-1,p'}) \to W^{1,r'}(I, W^{-1,p'}) \cap L^{r'}(I, W_D^{1,p'})$. Since it holds $A^*\iota^*\bar{\nu} = dt \otimes \bar{\nu} \in L^{r'}(I, W_D^{-1,p'})$ for any $r' \in (1,\infty)$ we obtain improved regularity $\bar{p} \in W^{1,r'}(I, W^{-1,p'}) \cap L^{r'}(I, W_D^{1,p'})$,

$r' \in [s', \infty)$, for the adjoint state from Theorem 4.4 in this case. Moreover, the adjoint equation (4.2) 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 S and its derivatives stated in Lemma 2.4. In this subsection we carry out a more detailed analysis w.r.t. 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 Assumption 2.1 hold and fix $u \in L^s(\Lambda)$.*

1a. 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 $q \in (1, \frac{2p}{d})$.

1b. 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 $r \in (1, \frac{2p}{p+d})$.

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 follows $S'(u_k)v_k \to S'(\bar{u})v_k$ strongly in $L^q(I, C(\overline{\Omega}))$ and $S''(u_k)v_k^2 \rightharpoonup S''(\bar{u})v^2$ 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.

As already pointed out at the end of Section 2.1 our Assumption 2.1 suffices to apply those results of [**4**] that are used in the proof below.

Proof. 1. We know $S'(u) \in \mathcal{L}((L^2(\Lambda), W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p}))$ for all $u \in L^s(\Lambda)$, cf. Lemma 2.4. Hence, 1a. follows from $W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p}) \hookrightarrow_c L^q(I, C(\overline{\Omega}))$, $q \in (1, \frac{2p}{d})$, see e.g. [**4**, Proposition 3.3]. Since this embedding is compact, $S'(u) \in \mathcal{L}(L^2(\Lambda), L^q(I, C(\overline{\Omega})))$ is also compact. For 1b. it suffices due to Lemma 2.4 to show for $r \in (1, \frac{2p}{p+d})$ and $w_1, w_2 \in W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})$ that $\|\mathcal{A}''(y)[w_1, w_2]\|_{L^r(I, W_D^{-1,p})} \lesssim \|w_1\|_{W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})} \|w_2\|_{W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})}$ holds, where $y = S(u)$. This, however, follows from the definition of $\mathcal{A}''$, Hölders inequality and the aforementioned embedding.

2. In the proof of [**4**, Proposition 4.9] it has been shown that $\tilde{S}'(Bu_k) \to \tilde{S}'(B\bar{u})$ in $\mathcal{L}(L^r(I, W^{-1,p}), W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p}))$ as long as $r \leq \frac{2p}{p-d}$; see Section 2.4 for the meaning of $\tilde{S}$. In particular, $S'(u_k) \to S'(\bar{u})$ in $\mathcal{L}(L^2(\Lambda), W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p}))$ is true, from which we conclude the first statement of part 2. For the second derivative we write $S''(u_k)v_k^2 - S''(\bar{u})v^2 = (\tilde{S}'(Bu_k) - \tilde{S}'(B\bar{u}))\mathcal{A}''(y_k)[S'(u_k)v_k]^2 + \tilde{S}'(B\bar{u})(\mathcal{A}''(y_k)[S'(u_k)v_k]^2 - \mathcal{A}''(\bar{y})[S'(\bar{u})v]^2)$, with $y_k = S(u_k)$ and $\bar{y} = S(\bar{u})$. Convergence of the operators above is in particular true for $r \in (1, \frac{2p}{p+d})$. Hence, it suffices to show that $\mathcal{A}''(y_k)[S'(u_k)v_k]^2 \rightharpoonup \mathcal{A}''(\bar{y})[S'(\bar{u})v]^2$ weakly in $L^r(I, W_D^{-1,p})$, which follows by Hölders inequality and the previous results. □

Proposition 4.6 motivates the introduction of averaged-in-time state-constraints: Assume we want to apply Theorem 3.2 to (OCP) in case of pointwise in time and space state-constraints (Assumption 2.5). Consequently, we have to verify Assumption 3.1 for $U_\infty = L^s(\Lambda)$, $U_2 = L^2(\Lambda)$, $K = U_{\text{ad}}$, $Z = C(\overline{Q})$, $C = Y_{\text{ad}}$, j being the reduced functional and $g = S$ being the control-to-state map of (OCP). 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 $u \in U_{\text{ad}}$. The proof of Proposition 4.6, however, shows this already fails to hold for the first derivative: From Lemma 2.4 we know that the extension $S'(u)$: $L^2(\Lambda) \to W^{1,2}(I, W_D^{-1,p}) \cap L^2(I, W_D^{1,p})$, 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, the application to averaged-in-space and pointwise in time state-constraints [**38**], i.e. $Y_{\mathrm{ad}} = \{y\colon\ y_a(t) \le \int_\Omega y(t,x)\omega(x)\mathrm{d}x \le y_b(t)\ \forall t \in I\}$, 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 Assumption 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 reasonable, resulting in the formulation of Assumption 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 Proposition 4.6. Moreover, 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 [**35**], cf. in particular [**35**, Proposition 3.8]: The reason 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 [**35**] 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. SSCs. Using the previously obtained results we formulate SSCs 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

$$-\partial_t \hat{p} + \mathcal{A}(\bar{y})^* \hat{p} + \mathcal{A}'(\bar{y})^* \hat{p} = \bar{y} - y_d, \qquad \hat{p}(T) = 0, \tag{4.5}$$

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 Assumption 2.1 and Assumption 4.1.1 hold, and let $\bar{u} \in U_{\mathrm{ad}}$, $\bar{y} = S(\bar{u}) \in Y_{\mathrm{ad}}$, $\bar{\nu} \in \mathcal{M}(\overline{\Omega})$ fulfill the optimality system* (4.1)-(4.4) *from Theorem 4.4. We define the critical cone by $C_{\bar{u},\bar{\nu}} := \{v \in L^2(\Lambda)$: (4.6)-(4.8) *hold*$\}$;

(4.6)
$$\int_\Lambda (\gamma\bar{u} + B^*\hat{p})v = 0, \qquad \int_\Omega \int_I z_v(t,x)\mathrm{d}t\mathrm{d}\bar{\nu} = 0,$$
(4.7)
$$\int_I z_v(t,\cdot)\mathrm{d}t \ge 0 \ \textit{on}\ \{\textstyle\int_I \bar{y}(t,\cdot)\mathrm{d}t = y_a\}, \ \int_I z_v(t,\cdot)\mathrm{d}t \le 0 \ \textit{on}\ \{\textstyle\int_I \bar{y}(t,\cdot)\mathrm{d}t = y_b\},$$
(4.8)
$$v \le 0, \qquad \textit{on}\ \{\bar{u} = u_b\}, \quad v \ge 0, \qquad \textit{on}\ \{\bar{u} = u_a\},$$

where $\bar{p}$ and $\hat{p}$ are defined by (4.2) *and* (4.5), *respectively, and $z_v = S'(\bar{u})v$. If*

$$\gamma\|v\|^2_{L^2(\Lambda)} + \int_Q ((1 - \xi''(\bar{y})\mu\nabla\bar{y}\nabla\bar{p})z_v^2 - 2\xi'(\bar{y})z_v\mu\nabla z_v\nabla\bar{p})\mathrm{d}x\mathrm{d}t > 0, \tag{4.9}$$

holds for all $v \in C_{\bar{u},\bar{\nu}} \setminus \{0\}$, 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 $u \in U_{\mathrm{ad}}$ such that $\|u - \bar{u}\|_{L^2(\Lambda)} < \epsilon$ and $y_a(x) \leq \int_I S(u)(t,x)\mathrm{d}t \leq y_b(x)$ for all $x \in \overline{\Omega}$. In particular, $\bar{u}$ is an local solution of (OCP) w.r.t. the $L^2(\Lambda)$-topology.

Proof. We apply Theorem 3.2 with $U_\infty = L^s(\Lambda)$, $U_2 = L^2(\Lambda)$, $K = U_{\mathrm{ad}}$, $Z = C(\overline{\Omega})$, 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 Assumption 3.1 have already been checked in [**4**, Theorem 4.14]. Note that the average-in-time map A 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 Assumption 3.1.2 immediately follow from Proposition 4.6. Hereby, observe that convergence of $(u_k)_k \subset U_{\mathrm{ad}}$ 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 Proposition 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 illustrates the full strength of Theorem 3.2. To prove C^2-differentiability of the control-to-state map we need controls in $L^s(\Lambda)$ with $s \gg 1$ as in Assumption 2.1, cf. [**4**], because already existence of solutions to (SE) relies on such an assumption [**41**]. 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 (4.9) with the increments v coming from $L^s(\Lambda)$. The latter condition can only hold for v coming from $L^2(\Lambda)$, cf. [**13, 32, 47**]. For the same reason, a similar situation holds for $g := A \circ S$. It is clear that g is well-defined and C^2-differentiable on $L^s(\Lambda)$. The question whether g is even well-defined on $L^2(\Lambda)$ is not clear. Although the problem necessarily requires to refer to two non-equivalent norms, a norm gap in the formulation of Theorem 4.8 can be avoided. This is the main benefit and novelty of Theorem 3.2.

5. SSCs for pointwise state-constraints

In the previous section we relaxed the type of state-constraints while keeping the regularity assumptions for the equation unchanged. Now 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. We replace Assumption 2.1 by a slightly smoother setting that allows to use stronger regularity results from [**8**]. Based on this we derive a result analogous to Proposition 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, 41**], 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 Assumption 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 Assumption 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 $\frac{1}{s} < \frac{1}{2}(1-\frac{d}{p})$. The set of admissible controls is given by $U_{\text{ad}} := \{u \in L^{2s}(I,\mathbb{R}^m) : u_a \le u \le u_b \text{ on } I\}$ with control-bounds $u_a, u_b \in L^\infty(I,\mathbb{R}^m)$, and for fixed control-functions $b_i \in L^p$, $i = 1, \ldots, m$ we define B: $L^{2s}(I,\mathbb{R}^m) \to L^{2s}(I,L^p)$, $u \mapsto \sum_{i=1}^m u_i b_i$. 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 [**18**]. The reason is the following, cf. also [**18**, 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 Example 2.2.3) of the fixed control-functions now excludes any possibility of boundary control.

Remark 5.2. From [**8**, p. 609] we recall: $C^{1,1}$-smoothness of Γ, combined with homogeneous Dirichlet boundary conditions and Lipschitz-continuity of μ implies that $-\nabla\cdot\mu\nabla + 1$: $W_D^{1,q} \to W_D^{-1,q}$ is a topological isomorphism for any $q \in (1,\infty)$. Consequently, Assumption 5.1 is indeed a tightened version of Assumption 2.1.

5.2. Improved regularity of the state. We start by recalling the following regularity result from [**8**] that will be the cornerstone of our further analysis:

Theorem 5.3 ([**8**], Theorem 2.3). *Let Assumptions 5.1.1 and 5.1.2 hold and fix $p, s \in [2,\infty)$ such that $\frac{1}{s} + \frac{d}{p} < 2$. Given $v \in L^{2s}(I,L^p)$ and $y_0 \in (L^p, W^{2,p} \cap W_D^{1,p})_{1-1/s,s} \cap (W_D^{-1,2p}, W_D^{1,2p})_{1-1/(2s),2s} \cap C(\Omega)$ there is a unique solution y to equation (2.1) 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 Assumption 5.1 .*

Note that [**8**, Remark 2], addressing 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 $p > d$.

Corollary 5.4. *Under Assumption 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 $\frac{1}{2}(1+\frac{d}{p}) < \theta < 1-\frac{1}{s}$ 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 $W^{2,p} \to W^{2,p}(\mathbb{R}^d)$ [**44**]. Thus, a standard argument utilizing the retraction-coretraction theorem ([**46**, 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 $[L^p(\mathbb{R}^d), W^{2,p}(\mathbb{R}^d)]_\theta = H^{2\theta,p}(\mathbb{R}^d)$ [**46**, Theorem 4.3.2.2] and standard Sobolev embeddings on $\mathbb{R}^d$ with $\kappa = 2\theta - \frac{d}{p} - 1 > 0$ [**46**, Theorem 2.8.1]. $\square$

5.3. Improved regularity for derivatives of the control-to-state map. We provide an improved version of Lemma 2.4 under the strengthened regularity Assumption 5.1. 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 $y \in C^{0,\rho}(I, C^{1,\kappa})$. Hence, 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 Lemma 2.4 and proposition 4.6. The approach is similar to [**4**] with the essential difference that the weaker regularity $y \in W^{1,s}(I, W_D^{1,p}) \cap L^s(I, W_D^{1,p})$ for the states in [**4**, Section 3.2] suffices to show constant domains and maximal parabolic regularity on $H_D^{-\zeta,p}$ 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 $H_D^{-\zeta,p}$ will not suffice for the derivation of SSCs for (OCP) in case of pointwise in time and space state-constraints, see Remark 4.7. The following observation is rather trivial in our case. We state it due to its importance for the following results.

Lemma 5.5. *Under Assumptions 5.1.1 and 5.1.2 let $\eta \in W^{1,\infty}$ with $\eta \geq \eta_\bullet > 0$ on Ω. Then it holds: 1.* $\mathrm{Dom}_{L^p}(-\nabla \cdot \eta\mu\nabla + 1) \cong \mathrm{Dom}_{L^p}(-\nabla \cdot \mu\nabla + 1) = W^{2,p} \cap W_D^{1,p}$*, i.e.* $-\nabla \cdot \eta\mu\nabla + 1$ *is a topological isomorphism* $W^{2,p} \cap W_D^{1,p} \to L^p$*.*

2. The map $\eta \mapsto -\nabla \cdot \eta\mu\nabla$ is bounded linear as map $W^{1,\infty} \to \mathcal{L}(W^{2,p} \cap W_D^{1,p}, L^p)$.

Similar results have been obtained in [**28**, Lemmas 6.5, 6.7, Corollary 6.8] if (SE) is considered on certain Bessel-potential spaces instead of L^p.

Proof. 1. This follows from [**24**, Theorem 2.4.2.5] for instance.

2. It holds $-\nabla \cdot \eta\mu\nabla \in \mathcal{L}(W^{2,p} \cap W_D^{1,p}, L^p)$ for any $\eta \in W^{1,\infty}$, with linear dependence on η. A short computation shows $\|-\nabla \cdot \eta\mu\nabla\varphi\|_{L^p} \lesssim \|\eta\|_{W^{1,\infty}} \|\varphi\|_{W^{2,p} \cap W_D^{1,p}}$, for any $\eta \in W^{1,\infty}$, $\varphi \in W^{2,p} \cap W_D^{1,p}$, which verifies boundedness. $\square$

The following lemma is a first step towards the analysis of the linearized state equation on L^p, where linearization takes place at some y that exhibits the regularity obtained in Theorem 5.3 for solutions of (SE). The linearized state equation is given by the parabolic PDE associated to the nonautonomous linear parabolic operator $\mathcal{A}(y) + \mathcal{A}'(y)$. Regularity of the first summand of this operator, i.e. $\mathcal{A}(y)$, is provided by the following lemma; the whole operator will be addressed in Lemma 5.7.

Lemma 5.6. *Let Assumptions 5.1.1 and 5.1.2 hold and fix $y \in W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W_D^{1,p})$. The nonautonomous linear parabolic operator $\mathcal{A}(y)$ exhibits maximal parabolic regularity on $L^r(I, L^p)$, $r \in (1, \infty)$, i.e. the solution map $(w, w_0) \mapsto z$ of the equation*

$$\partial_t z + \mathcal{A}(y)z = w, \qquad z(0) = w_0,$$

is linear and bounded as a map $L^r(I, L^p) \times (L^p, W^{2,p} \cap W^{1,p}_D)_{1/r',r} \to W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D)$. Moreover, the corresponding operators norms are bounded uniformly 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 [**42**, 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 formulas (3.16) and Lemma 3.12 in [**4**]; 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. 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. As in the proof of [**4**, Proposition 3.18] we write with help of the resolvent calculus: $(A(t)+1)R(z, A(t)+1)[(A(t) + 1)^{-1} - (A(s) + 1)^{-1}] = R(z, A(t) + 1)[A(t) - A(s)](A(s) + 1)^{-1}$. From Lemma 5.5.2 it follows that $\|A(t)-A(s)\|_{\mathcal{L}(W^{2,p}\cap W^{1,p}_D,L^p)} \le c\|\xi(y)\|_{C^\rho(I,W^{1,\infty})}|t-s|^\rho$ with $c > 0$ independent of y. Employing formula (3.16) from [**4**] there is $\theta \in (0, \pi/2)$ such that $\|R(z, A(t) + 1)\|_{\mathcal{L}(L^p)} \le c|z|^{-1}$ for all $z \in \mathbb{C} \setminus \overline{\Sigma_\theta}, t \in I$, with $\Sigma_\theta = \{z \in \mathbb{C} \setminus \{0\}$: $|\arg z| < \theta\}$, and finally it follows again from Lemma 5.5 that $\|(A(s)+1)^{-1}\|_{\mathcal{L}(L^p,W^{2,p}\cap W^{1,p}_D)} \le c\|\xi(y)\|_{L^\infty(I,W^{1,\infty})}$ with a constant c independent of y. Together, this shows the Acquistapace-Terreni condition, $\|(A(t) + 1)R(z, A(t) + 1)[(A(t)+1)^{-1} - (A(s)+1)^{-1}]\|_{\mathcal{L}(L^p)} \le C|t-s|^\rho|z|^{-1}$ for all $z \in \mathbb{C}\setminus\overline{\Sigma_\theta}, t, s \in I$, with the constant $C > 0$ depending on the $C^\rho(I, W^{1,\infty})$-norm of y. Therefore, C can be chosen uniformly for y's coming from a bounded set in $W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$ due to Corollary 5.4. □

Now, we consider maximal parabolic regularity for the linearized state equation. This extends Lemma 2.4 (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 Assumptions 5.1.1 and 5.1.2 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. Maximal parabolic regularity of $\mathcal{A}(y)$ on $L^r(I, L^p)$, $r \in (1, \infty)$, has been shown in Lemma 5.6. Corollary 5.4 together with Lemma 5.5 implies continuity of the map $I \to \mathcal{L}(W^{2,p} \cap W^{1,p}_D, L^p)$, $t \mapsto -\nabla\cdot\xi(y(t))\mu\nabla$, from which we conclude that each autonomous operator $-\nabla \cdot \xi(y(t)\mu\nabla \in \mathcal{L}(W^{2,p} \cap W^{1,p}_D, L^p)$, $t \in I$, has in fact maximal parabolic regularity on L^p. This follows from [**2**, Proposition 7.1]. Regarding the second summand, $\mathcal{A}'(y)$, we observe that the map $I \to \mathcal{L}(W^{1,\infty}, L^p)$, $t \mapsto (\psi \mapsto -\nabla\cdot\xi'(y(t))\psi\mu\nabla y(t))$ is L^s-integrable w.r.t. time: This follows from the continuity of the map $W^{1,\infty} \to \mathcal{L}(W^{2,p} \cap W^{1,p}_D, L^p)$, $\eta \mapsto -\nabla \cdot \eta\mu\nabla$, see Lemma 5.5.2, together with $\xi'(y) \in L^\infty(I, W^{1,\infty})$, the continuity of the product on $W^{1,\infty} \times W^{1,\infty}$, and the fact that $y \in L^s(I, W^{2,p} \cap W^{1,p}_D)$. Hence, we have just shown $\mathcal{A}'(y) = (t \mapsto (\psi \mapsto -\nabla \cdot \xi'(y)\psi\mu\nabla y)) \in L^s(I, \mathcal{L}(W^{1,\infty}, L^p)) \hookrightarrow L^s(I, \mathcal{L}((L^p, W^{2,p})_{\theta,\infty}, L^p))$ with some $1 - 1/s > \theta > \hat{\theta} > \frac{1}{2} + \frac{d}{2p}$. Hereby, we made use of the embedding $(L^p, W^{2,p})_{\theta,\infty} \hookrightarrow [L^p, W^{2,p}]_{\hat{\theta}} \hookrightarrow W^{1,\infty}$, cf. the proof of Corollary 5.4. From [**2**, Theorem 7.1] we conclude maximal parabolic regularity of $\mathcal{A}(y) + \mathcal{A}'(y)$ on $L^r(I, L^p)$ for $r \in (1, s)$. Similar to the proof of [**4**, Proposition 4.4] we invoke [**43**, Corollary 3.4] to get maximal parabolic regularity on $L^s(I, L^p)$. □

We point out that Lemma 5.7 and Theorem 5.3 do *not* allow immediately to conclude differentiability of the solution map of (2.1) 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 be shown with a similar argument as for Corollary 5.4. Hence, the map F: $W^{1,r}(I, L^p) \cap L^r(I, W^{2,p} \cap W^{1,p}_D) \times L^r(I, L^p) \to L^r(I, L^p) \times (L^p, W^{2,p} \cap W^{1,p}_D)_{1/r',r}, (y, v) \mapsto (\partial_t y + \mathcal{A}(y) - v, y(0) - y_0)$, is continuously Fréchet differentiable. Further, for $r \in (1, s]$ the partial derivative $\partial_y F(y, v)$ is even continuously invertible, cf. Lemma 5.7. Nevertheless, the fact that prevents us from application of the implicit function theorem is that we would 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 (2.1) 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 $\partial_y F(y, v)$ 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. Lemma 5.7.

Remark 5.8. Double time integrability on the right-hand-side of (2.1) in Theorem 5.3 is due to the technique applied in the proof of [**8**, Theorem 2.3]. For a short outline we refer to the proof of Lemma 5.10 below.

The following lemma is the first step towards an analogue to Proposition 4.6.1 under Assumption 5.1. Particularly, 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. Even in this highly regular setting $\mathcal{A}''(y)w^2$ is from $L^r(I, W^{-1,p}_D)$, i.e. a distributional object in general, which illustrates the difficulty of this term.

Lemma 5.9. *Given $y \in W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W^{1,p}_D)$ and $w \in W^{1,2}(I, L^p) \cap L^2(I, W^{2,p} \cap W^{1,p}_D)$ it holds $\|\mathcal{A}''(y)w^2\|_{L^r(I,W^{-1,p}_D)} \leq c_{y,r} \|w\|^2_{W^{1,2}(I,L^p) \cap L^2(I,W^{2,p} \cap W^{1,p}_D)}$ for $r \in (1, \infty)$. The constant $c_{y,r}$ can be chosen uniformly w.r.t. 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 $W^{1,2}(I, L^p) \cap L^2(I, W^{2,p} \cap W^{1,p}_D) \hookrightarrow C(\overline{Q})$ and $W^{1,2}(I, L^p) \cap L^2(I, W^{2,p} \cap W^{1,p}_D) \hookrightarrow L^q(I, W^{1,p})$ for every $q \in (1, \infty)$, that can be shown similarly as in Corollary 5.4. □

The following lemma 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* (2.1) *is continuous from* $L^{2s}(I, L^p)$ *to* $W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W_D^{1,p})$.

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, cf. [**8**, Theorem 3.2]. As outlined after Lemma 5.7, arguing via the implicit function theorem is not possible here. To prove continuous dependence we go through the steps in [**8**] tracking continuous dependence of the quantities under consideration.

Proof. It is well known that $L^{2s}(I, L^p) \hookrightarrow L^{2s}(I, W_D^{-1,2p})$ and that solution map of (2.1) is continuous (in fact even C^2) from $v \in L^{2s}(I, W_D^{-1,2p})$ to $y = y(v) \in W^{1,2s}(I, W^{-1,2p}) \cap L^{2s}(I, W_D^{1,2p})$: Existence is clear by [**8**, Theorem 2.1] and differentiability follows using the implicit function theorem similarly as in the proof of [**8**, Theorem 3.2]: The required invertibility property is assured by maximal parabolic regularity of $\mathcal{A}(y) + \mathcal{A}'(y)$ on $L^{2s}(I, W_D^{-1,2p})$, which is proven similarly as in the proof of [**4**, Proposition 4.4] with s and p replaced by $2s$ and $2p$, respectively, cf. also the similar proof of Lemma 5.7 in the L^p-setting. Next, following the main idea in the proof of [**8**, Theorem 2.3] we rewrite equation (2.1) as $\partial_t z - \xi \nabla \cdot \mu \nabla z = v + \nabla \xi \cdot \mu$ with $y = y(v) \in W^{1,2s}(I, W^{-1,2p}) \cap L^{2s}(I, W_D^{1,2p})$ being the solution to (2.1) and $\xi = \xi(y(v))$. It is clear that the right-hand-side measured in $L^s(I, L^p)$ depends continuously on ξ in $L^{2s}(I, W^{1,2p})$ and y in $L^{2s}(I, W^{1,2p})$, respectively, i.e. on v in $L^{2s}(I, W^{-1,2p})$ by the above consideration. Further, due to the embedding $W^{1,2s}(I, W^{-1,2p}) \cap L^{2s}(I, W_D^{1,2p}) \hookrightarrow C(\overline{Q})$, also $\xi = \xi(y(v))$ depends continuously in $C(\overline{Q})$ on v. Finally, the map $C(\overline{Q}) \to \mathcal{L}(W^{1,s}(I, L^p) \cap L^s(I, W^{2,p} \cap W_D^{1,p}), L^s(I, L^p))$, $\xi \mapsto \partial_t - \xi \nabla \cdot \mu \nabla$, is continuous. Therefore, using [**8**, Lemma 2.4] the solution $z = y$ depends continuously on ξ in $C(\overline{Q})$ and y in $W^{1,2s}(I, W^{-1,2p}) \cap L^{2s}(I, W_D^{1,2p})$, and thus on v in $L^{2s}(I, L^p)$. $\square$

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

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

1. For any $u \in L^{2s}(I, \mathbb{R}^m)$, $S'(u)$ *and* $S''(u)$ *extend to continuous linear and bilinear forms on* $L^2(I, \mathbb{R}^m)$ *with values in* $C(\overline{Q})$, *respectively.*

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

The proof has similar structure as the one of Proposition 4.6.

Proof. Differentiability of the control-to-state map and the formulas for the respective derivatives follow from Lemma 2.4. Note that Assumption 5.1 indeed suffices to invoke this result, cf. Remark 5.2. The extension property for the first derivative follows from Lemma 5.7 with $r = 2$ and the first embedding in the proof of Lemma 5.9. For the continuation of the second derivative, combine the continuation property for $S'(u)$ with Lemmas 5.7 and 5.9 and the embedding from [**4**, Proposition 3.3]. It remains to check the continuity properties: As an auxiliary result, we first show that $S'(u_k) \to S'(\bar{u})$ in $\mathcal{L}(L^r(I,\mathbb{R}^m), W^{1,r}(I,L^p)\cap L^r(I,W^{2,p}\cap W_D^{1,p}))$ for any $r \in (1,\infty)$. To do so, it suffices, by continuity of operator inversion, to show convergence $\mathcal{A}(y_k)+\mathcal{A}'(y_k) \to \mathcal{A}(\bar{y})+\mathcal{A}'(\bar{y})$ in $\mathcal{L}(W^{1,r}(I,L^p)\cap L^r(I,W^{2,p}\cap W_D^{1,p}), L^r(I,L^p))$. This can be done using Lemma 5.10, Hölders inequality and $W^{1,r}(I,L^p)\cap L^r(I,W^{2,p}\cap W_D^{1,p}) \hookrightarrow L^q(I,W^{1,\infty})$ 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 Corollary 5.4. Having at hand this auxiliary result, the continuity property for the first derivative follows similarly as in the proof of Proposition 4.6. For the second derivative we also argue similarly as in the proof of Proposition 4.6: Due to the embedding from [**4**, Proposition 3.3] for $r > \frac{2p}{p-d}$ it suffices to show that $\tilde{S}'(Bu_k) \to \tilde{S}'(B\bar{u})$ in $\mathcal{L}(L^r(I,W_D^{-1,p}), W^{1,r}(I,W_D^{-1,p})\cap L^r(I,W_D^{1,p}))$, and $\mathcal{A}''(y_k)[w_k]^2 \rightharpoonup \mathcal{A}''(y)[w]^2$ weakly in $L^r(I,W_D^{-1,p})$. We leave the details to the reader. □

5.4. SSCs. We now apply Theorem 3.2 to (OCP) under Assumptions 2.5 and 5.1 and formulate SSCs for (OCP) with pointwise in time and space state-constraint. Compared to Theorem 4.8 we crucially rely on the improved regularity results due to the strengthened regularity Assumption 5.1.

Theorem 5.12. *Let Assumption 5.1 and Assumption 2.5.1 hold, and let $\bar{u} \in L^{2s}(I,\mathbb{R}^m)$, $\bar{y} \in W^{1,s}(I,L^p)\cap L^s(I,W^{2,p}\cap W_D^{1,p})\cap Y_{\mathrm{ad}}$ and $\bar{\lambda} \in \mathcal{M}(\overline{Q})$ fulfill the FONs* (2.5)-(2.8) *from Theorem 2.8. We define the cone of critical directions by $C_{\bar{u},\bar{\lambda}} := \{v \in L^2(I,\mathbb{R}^m)\colon (5.1)-(5.3)\ hold\}$,*

$$\int_I (\gamma\bar{u}(t) + B^*\hat{p}(t))^T v(t)\mathrm{d}t = 0, \qquad \int_{\overline{Q}} z_v \mathrm{d}\bar{\lambda} = 0 \tag{5.1}$$

$$z_v(t,x) \le 0,\ \ on\ \{\bar{y} = y_b\},\ \ z_v(t,x) \ge 0,\ \ on\ \{\bar{y} = y_a\}, \tag{5.2}$$

$$v_i(t) \le 0, \qquad if\ \bar{u}_i(t) = u_{b,i}(t), \quad v_i(t) \ge 0, \qquad if\ \bar{u}_i(t) = u_{a,i}(t), \tag{5.3}$$

where $\hat{p}$ is defined by (4.5) *and $z_v = S'(\bar{u})v$. If*

$$\gamma\|v\|^2_{L^2(I,\mathbb{R}^m)} + \int_Q ((1-\xi''(\bar{y})\mu\nabla\bar{y}\nabla\bar{p})z_v^2 - 2\xi'(\bar{y})z_v\mu\nabla z_v\nabla\bar{p})\mathrm{d}x\mathrm{d}t > 0, \tag{5.4}$$

holds for all $v \in C_{\bar{u},\bar{\lambda}} \setminus \{0\}$, 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 $u \in U_{\mathrm{ad}}$ that satisfy $\|u-\bar{u}\|_{L^2(I,\mathbb{R}^m)} < \epsilon$ and $S(u) \in Y_{\mathrm{ad}}$.*

Proof. We apply Theorem 3.2 with $U_\infty = L^s(I,\mathbb{R}^m)$, $U_2 = L^2(I,\mathbb{R}^m)$, $Z = C(\overline{Q})$, $K = U_{\mathrm{ad}}$, $C = Y_{\mathrm{ad}}$. As stated in the proof of Theorem 4.8, the assumptions on the reduced functional j from Assumption 3.1.1 have been verified in [**4**] and Assumption 3.1.2 for $g = S$ is fulfilled due to Proposition 5.11. The crucial point is as in the proof of Theorem 4.8 to observe that due to L^∞-boundedness of U_{ad} $L^2(I,\mathbb{R}^m)$- implies $L^{2s}(I,\mathbb{R}^m)$- and $L^s(I,\mathbb{R}^m)$-convergence, respectively. □

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 $L^2(I,\mathbb{R}^m)$. The positivity condition (5.4) might hold for directions v from $L^2(I,\mathbb{R}^m)$, but cannot be expected to hold for directions $v \in L^s(I,\mathbb{R}^m)$. However, 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 Example 3.3 occurrence of a two-norm gap can be avoided.

6. The case without control-constraints

In the terminology of Assumptions 2.1 and 5.1 we now consider the case $u_a \equiv -\infty$ or $u_b \equiv +\infty$.

6.1. Existence of optimal controls. The argument in the proof of Theorem 2.6 relies on the possibility to extract a $L^s(\Lambda)$-bounded subsequence of the infimizing sequence of controls, see the proof of [**41**, 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. The latter can be modified by an $L^s(\Lambda)$-Tikhonov term as follows

$$J(y,u) = \frac{1}{2}\|y-y_d\|^2_{L^2(I\times\Omega)} + \frac{\gamma}{2}\|u\|^2_{L^2(\Lambda)} + \frac{\gamma_s}{s}\|u\|^s_{L^s(\Lambda)}, \tag{6.1}$$

with some $\gamma_s > 0$. Now, the techniques used in the proof of Theorem 2.6 apply again, and existence of a globally optimal control can be shown.

6.2. FONs. Theorems 2.8 and 4.4 do not rely on the existence of optimal controls; they only characterize locally optimal controls, if they exist. Hence, these results stay valid for unilateral or no control-constraints. In the case $U_{\mathrm{ad}} = L^s(\Lambda)$ the variational inequalities (2.8) and (4.4), respectively, simplify to the equality $B^*\bar{p} + \gamma\bar{u} = 0$. Moreover, the reduced gradient for the modified functional (6.1) is given by $\nabla j(\bar{u}) = B^*\bar{p} + \gamma\bar{u} + \gamma_s|\bar{u}|^{s-1}$, and (2.8) and (4.4) have to be adapted accordingly.

6.3. SSCs. The verification of Assumption 3.1 makes use of L^∞-boundedness of the admissible set U_{ad}, cf. the proofs of Propositions 4.6 and 5.11 and [**4**, Theorem 4.14]. Hence, we cannot apply Theorem 3.2. However, following Remark 3.4 it is still possible to obtain SSCs with norm gap: The quadratic growth condition will only hold on an $L^s(\Lambda)$- and $L^{2s}(I,\mathbb{R}^m)$-neighborhood of $\bar{u}$, respectively. Furthermore, when using the modified functional (6.1) the choice $\gamma = 0$ is not possible when aiming at SSCs with the technique of the present paper, even at such with two-norm gap: The condition $\gamma > 0$ is crucial for verifying Assumption 3.1.1biii for the reduced functional; see formula (30) in [**18**], or formula (5.3) in [**12**], respectively.

Appendix A. Applicability of the results of [41]

In the paper we used results from [**41**]. We now comment on their applicability in our setting, applying the notation of [**41**]. The results of [**41**, Sections 5 and 6] refer to the PDE

$$\partial_t w - \nabla\cdot\phi(w)\rho\nabla w + w = \mathcal{F}(t,w(t)), \qquad w(T_0) = w_0, \tag{A.1}$$

which differs from (SE) by the additional identity-term on the left-hand-side and the nonlinearity on the right-hand-side. The special case of right-hand-sides independent of w, as in the present paper, is contained in this setting. Existence and regularity of solutions [**41**, Theorem 5.3] and weak-to-strong continuity of the solution map [**41**, Theorem 6.3] have alreadby been applied to (SE) in [**4**]. We now give the details concerning the required adaptations. If Γ_D has nonzero measure, omitting the identity-term does not require changes in the arguments, cf. [**41**, Remark 2.12]. In the following we outline the changes needed to deal with the general case.

A.1. Uniform Hölder estimates. The crucial idea is to obtain uniform Hölder estimates for nonautonomous linear parabolic PDEs without identity-term from those with identity-term by application of a well-known scaling technique, see [**17**, Chapter XVIII.§3.1.4, Remark 2] for instance: Observe that u is a solution to

$$\partial_t u(t) - \nabla \cdot \mu(t,\cdot)\nabla u(t) = f(t), \qquad u(T_0) = u_0,$$

if and only if $\tilde{u}(t,\cdot) := e^{-t}u(t,\cdot)$ is a solution to

$$\partial_t \tilde{u}(t) - \nabla \cdot \mu(t,\cdot)\nabla \tilde{u}(t) + \tilde{u}(t) = e^{-t}f(t), \qquad \tilde{u}(T_0) = e^{-T_0}u_0.$$

Since the results of [**41**] apply to $\tilde{u}$ and scaling back $u(t,\cdot) = e^t\tilde{u}(t,\cdot)$ does not affect Hölder-continuity, both [**41**, Proposition 2.10] and [**41**, Theorem 2.13] also hold true without the identity-term, with adapted constants, of course.

A.2. Existence of solutions to (SE). We check that the proof of [**41**, Theorem 5.3] works without the identity-term. First, we ensure that the auxiliary results used in the proof stay valid: The proof of [**41**, Lemma 5.4] does not need modification. In part (i) of [**41**, Lemma 5.5] we have to argue on maximal parabolic regularity of $-\nabla \cdot \varphi\rho\nabla$ instead of $-\nabla \cdot \varphi\rho\nabla + 1$. To do so, note that $-\nabla \cdot \varphi\rho\nabla = (-\nabla \cdot \varphi\rho\nabla + 1) - 1$ inherits maximal parabolic regularity from $-\nabla \cdot \varphi\rho\nabla + 1$ due to $-1 \in L^\infty((T_0,T_1),\mathcal{L}(W^{1,q}_D, W^{1,q}_D))$ by [**2**, Theorem 7.1]. It is clear that the continuous dependence in part (ii) stays true. Then, the proof of [**41**, Theorem 5.3] can easily be adapted: Equation (5.7) now reads

$$\partial_t v - \nabla \cdot \phi(u+\psi)\rho\nabla v = \mathcal{F}(u+\psi) - \partial_t u + \nabla\phi(u+\psi)\rho\nabla u, \qquad v(T_0) = 0. \tag{A.2}$$

The technique of the proof of [**41**, Theorem 5.3] does not need further adaptation. The fixed-point argument from [**41**] continues to be applicable to (A.2), since uniform Hölder-estimates for the respective nonautonomous linear parabolic operators are also valid without the identity-term. For the same reason also the uniform estimates from [**41**, Corollaries 5.7 and 5.8] stay true, again with adapted constants.

A.3. Weak-to-strong continuity of the solution map. Weak-to-strong continuity of the solution map of (A.1) is shown in [**41**, Theorem 6.3] based on the results already discussed in Appendices A.1 and A.2. No further adaptations are needed.

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] ———, *Numerical analysis of quasilinear parabolic equations under low regularity assumptions*, Numer. Math. **143** (2019), no. 4, 749–780. MR4026371

[10] 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

[11] 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

[12] 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

[13] 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

[14] 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.

[15] 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

[16] Christian Clason, Vu Huu Nhu, and Arnd Rösch, *No-gap second-order optimality conditions for optimal control of a non-smooth quasilinear elliptic equation*, ESAIM Control Optim. Calc. Var. **27** (2021), Paper No. 62, 35. MR4274828

[17] Robert Dautray and Jacques-Louis Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 5*, Springer-Verlag, Berlin, 1992. Evolution problems. I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig. MR1156075

[18] 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

[19] 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

[20] Johannes Elschner, Joachim Rehberg, and Gunther Schmidt, *Optimal regularity for elliptic transmission problems including C^1 interfaces*, Interfaces Free Bound. **9** (2007), no. 2, 233–252. MR2314060

[21] 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

[22] ———, *Integral state-constrained optimal control problems for some quasilinear parabolic equations*, Nonlinear Anal. **39** (2000), no. 8, Ser. A: Theory Methods, 977–996. MR1735179

[23] 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

[24] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985. MR775683

[25] 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

[26] R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela, *Hölder continuity and optimal control for nonsmooth elliptic problems*, Appl. Math. Optim. **60** (2009), no. 3, 397–428. MR2551487

[27] Robert Haller-Dintelmann, Alf Jonsson, Dorothee Knees, and Joachim Rehberg, *Elliptic and parabolic regularity for second-order divergence operators with mixed boundary conditions*, Math. Methods Appl. Sci. **39** (2016), no. 17, 5007–5026. MR3573649

[28] 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

[29] 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

[30] D. Hömberg, C. Meyer, J. Rehberg, and W. Ring, *Optimal control for the thermistor problem*, SIAM J. Control Optim. **48** (2009/10), no. 5, 3449–3481. MR2599927

[31] Fabian Hoppe and Ira Neitzel, *Convergence of the SQP method for quasilinear parabolic optimal control problems*, Optim. Eng. (2020).

[32] 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

[33] 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

[34] 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

[35] 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

[36] 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

[37] 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

[38] 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

[39] 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

[40] ———, *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

[41] 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

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

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

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

[45] 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

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

[47] 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

[48] 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

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