Institut für Numerische Simulation

Rheinische Friedrich-Wilhelms-Universität Bonn

F. Hoppe, I. Neitzel

Purely time-dependent optimal control of quasilinear parabolic PDEs with sparsity enforcing penalization

Purely time-dependent optimal control of quasilinear parabolic PDEs with sparsity enforcing penalization

Fabian Hoppe and Ira Neitzel

Institut für Numerische Simulation,
Rheinische Friedrich-Wilhelms-Universität Bonn,
Friedrich-Hirzebruch-Allee 7, 53115 Bonn (Germany)

Abstract. We prove first- and second-order optimality conditions for sparse, purely time-dependent optimal control problems governed by a quasilinear parabolic PDE. In particular, we analyze sparsity patterns of the optimal controls induced by different sparsity enforcing functionals in the purely time-dependent control case and illustrate them by numerical examples. Our findings are based on results obtained by abstraction of well known techniques from the literature.

Introduction

This paper is devoted to sparse optimal control of quasilinear parabolic partial differential equations (PDEs). We focus on the purely time-dependent control setting, i.e. controls depending on time only, but not on space; cf. [**25**]. We obtain results that we expect from known results for the linear and semilinear case; see, e.g., [**13, 31**]. More precisely, we derive first-order necessary optimality conditions and associated sparsity patterns as well as second-order necessary and sufficient optimality conditions for problems of the following type:

$$(\mathrm{P}_k) \quad \min_{y,u} J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2(I\times\Omega)} + \frac{\alpha}{2}\sum_{i=1}^{m}\|u_i\|^2_{L^2(I)} + \beta j_k(u), \quad k \in \{1,...,7\},$$

$$\text{s.t.} \quad u \in U_{\mathrm{ad}} \subset L^\infty(I, \mathbb{R}^m),$$

$$(\mathrm{Eq}) \quad \text{and} \quad \left\{ \begin{aligned} \partial_t y - \nabla \cdot \xi(y)\mu\nabla y = Bu &:= \sum_{i=1}^{m} u_i b_i && \text{on } I \times \Omega, \\ y(0) = y_0 && \text{on } \Omega. \end{aligned} \right.$$

Here, $\alpha > 0$ denotes the $L^2(I, \mathbb{R}^m)$-Tikhonov regularization-parameter, and $\beta > 0$ weighs the sparsity enforcing penalization term $j_k \colon L^2(I, \mathbb{R}^m) \to \mathbb{R}$. For $k \in \{1,...,7\}$, the latter is given by one of the following functionals that are adaptations of the classical (directional)

sparsity enforcing penalizers [**31**] to the purely time-dependent-setting:

$$j_1(u) := \sum_{i=1}^m \|u_i\|_{L^1(I)},$$

$$j_2(u) := \sum_{i=1}^m \|u_i\|_{L^2(I)}, \qquad j_3(u) := \int_I \left(\sum_{i=1}^m |u_i(t)|^2 \right)^{\frac{1}{2}} \mathrm{d}t,$$

$$j_4(u) := \left(\sum_{i=1}^m \|u_i\|_{L^1(I)}^2 \right)^{\frac{1}{2}}, \quad j_5(u) := \left(\int_I \left(\sum_{i=1}^m |u_i(t)| \right)^2 \mathrm{d}t \right)^{\frac{1}{2}},$$

$$j_6(u) := \frac{1}{2} \sum_{i=1}^m \|u_i\|_{L^1(I)}^2, \qquad j_7(u) := \frac{1}{2} \int_I \left(\sum_{i=1}^m |u_i(t)| \right)^2 \mathrm{d}t.$$

For the precise setting, including boundary conditions of the state equation (Eq), conditions on the coefficient functions, the fixed actuators b_i, and the control operator B, as well as the definition of the set of admissible controls U_{ad}, we refer the reader to Section 1 below.

PDE-constrained optimization has been subject to intensive research for several decades, cf. [**33**, **41**, **58**] for instance, and has many applications. In some of them it may be desirable to determine controls that act only on a small ("sparse") part of the (space-time-)domain under consideration. Starting with the pioneering work of Stadler [**56**] on sparse optimal control of linear elliptic equations, there have been many contributions on this topic in the recent past. For an overview we refer the reader, e.g., to [**8**], and focus on literature related to the present paper in the following. Regarding literature following the original idea of Stadler to enforce sparsity by adding an L^1-penalization term to the objective functional we mention, e.g., [**12**, **16**, **18**, **20**, **21**, **55**, **59**]. These publications refer to different types of PDEs and cover several aspects, including first- and second-order optimality conditions, discretization error estimates, and additional state-constraints. When considering parabolic PDEs, it might be favorable to obtain a space-time sparsity profile of the optimal control in which space- and time-variable are treated in a different way. This leads to the concept of directional sparsity introduced in [**31**] for linear PDEs. First- and second-order optimality conditions for this concept applied to semilinear parabolic PDEs have been obtained in [**13**, **17**], for instance. The specific difficulty herein arises from the fact that sparsity enforcing penalizers are convex, but nonsmooth, whereas the remaining part of the objective functional is smooth, but —due to nonlinearity of the state equation— nonconvex. Moreover, the so-called two-norm discrepancy [**19**, **37**] occurs. Differentiability and coercivity of the second derivative of the smooth part of the objective functional hold only w.r.t. different norms. As alternative approaches to enforce sparsity we finally mention, e.g., control in measure spaces [**11**, **14**, **15**, **40**], or L^0-penalization [**22**, **38**].

To our knowledge, sparse optimal control of quasilinear PDEs has not been addressed in the literature so far. Regarding optimal control of quasilinear PDEs in general, we restrict the introduction to the parabolic setting and recent results, and refer to the introduction of [**5**, **9**] for earlier literature on quasilinear parabolic and elliptic optimal control problems. Well-posedness of the state equation and existence of optimal controls for an abstract functional that is convex, continuous, and coercive w.r.t. the control variable have been proven in [**47**] under rather general regularity assumptions on domain and coefficients. First- and second-order optimality conditions have been derived in [**5**] for control-constrained problems with usual L^2-Tikhonov functional. Based on this, convergence of the SQP-method for the respective optimization problem has been shown in [**34**]. Additional state-constraints are considered in [**35**]. Optimality conditions for a similar model problem with slightly more regular coefficients and domain, but in contrast unbounded nonlinearities, have been analyzed

in [**9**]. Optimal control of the so-called thermistor problem, a coupled system consisting of a quasilinear parabolic and a nonlinear elliptic equation, is addressed in [**45**, **46**].

Existence and regularity theory for solutions of the underlying PDEs poses the main difficulty in the analysis of such problems, in particular in the second-order analysis. The aforementioned papers have in common that they utilize the functional analytic concept of nonautonomous maximal parabolic regularity [**2**] to deal with this issue. Finally, we mention that a-priori finite element error estimates and a-posteriori Reduced Basis error estimates for the state equations from [**9**] and [**5**] have been obtained in [**10**] and [**36**], respectively.

The present paper contributes both to the fields of optimal control of quasilinear PDEs and sparse optimal control. We extend the first- and second-order analysis for sparse optimal control of semilinear parabolic PDEs from [**13**, **17**] to problems with a quasilinear parabolic state equation. Here, we focus on the purely time-dependent control setting, which may be advantageous in applications. Among the examples given in the introduction of [**25**] we mention, e.g., optimal cooling of steel profiles by controlling the intensities of the finite number of nozzles that spray water on the profile. With the techniques of this paper it would also be possible to treat directionally sparse, distributed optimal control problems along the lines of [**12**, **13**, **17**, **31**]; cf. also [**5**, Example 2.6]. However, while space-time sparsity patterns for distributed optimal control problems have already been under detailed consideration in [**13**], the purely time-dependent control setting has not been addressed systematically in the context of directional sparsity. Moreover, the chosen setup allows to include control of quasilinear parabolic PDEs by fixed Neumann boundary sources up to dimension 3, whereas distributed Neumann boundary control is only possible up to dimension 2; cf. [**5**, Example 2.4]. With respect to the state equation we rely on the low regularity assumptions of [**5**, **47**] that include certain discontinuous coefficients, nonsmooth domains, and mixed boundary conditions.

With our work, we combine two challenging aspects, namely sparsity enforcing penalization and a quasilinear state equation. In the presence of L^2-Tikhonov regularization we are able to carry out a full first- and second-order analysis, the latter one avoiding the two-norm gap. To do so, we pursue an abstract approach in the flavour of [**19**] and work out the abstract core of existing arguments for second-order conditions from [**12**, **13**, **17**], which may also facilitate the application to other problems. Due to the different nature of our nonlinearity, the second-order analysis in the bang-bang case $\alpha = 0$, i.e. the case without L^2-Tikhonov regularization, from [**13**, **16**] cannot be carried over to our setting; cf. Section 3.4 below. This illustrates that the transfer of techniques successfully applied to semilinear problems to quasilinear problems is by no means trivial. Second, we provide an extensive analysis of directional sparsity for purely time-dependent controls, which has to the best of our knowledge not been carried out so far, and include a numerical illustration of our findings. This is of particular interest due to the practical relevance of purely time-dependent controls. Besides the functionals j_1-j_5, whose structure corresponds to those already discussed in [**13**], we also propose and analyze the functionals j_6 and j_7, that have —to the best of our knowledge— not been dealt with in the context of PDE-constrained optimization so far. These two functionals are interesting, because their sparsity patterns are similar to those of j_4 and j_5, respectively, while they are advantegeous compared to j_4 and j_5 from a numerical point of view, because their proximity operator is computable.

In the next section we will state our assumptions and summarize our main results. Moreover, we provide an overview over the remaining part of the paper in which we prove our results and give a brief numerical illustration.

1. Problem setting and main results

1.1. Notation and assumptions. We introduce some notation and conventions, and state our assumptions on the control problem (P_k). Given Banach spaces X, Y we denote

by $\mathcal{L}(X, Y)$ the space of bounded linear operators from X to Y equipped with the operator norm, by $X^* := \mathcal{L}(X, \mathbb{R})$ the topological dual of X, and by $\mathbb{B}_r^X(x) \subset X$ the closed ball of radius r around $x \in X$. Moreover, $X \hookrightarrow Y$ indicates that X continuously embedds into Y. The domain of a closed linear operator $A: X \to Y$, equipped with the graph norm, is denoted by $\mathrm{Dom}_X(A)$. We apply standard notation for Bochner-Lebesgue- and Bochner-Sobolev-spaces, as well as for real and complex interpolation spaces. The following assumptions on (P_k) are close to [5], but we forego those parts that refer to the improved regularity analysis from [5] on Bessel-potential spaces and stick to the setting of [47]. Moreover, as already explained in the introduction, we restrict ourselves to the purely time-dependent control setting; cf. [5, Example 2.5].

Assumption 1.1. Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain with boundary $\partial\Omega$. $\Gamma_N \subset \partial\Omega$ is relatively open and denotes the Neumann boundary part whereas $\Gamma_D = \partial\Omega \setminus \Gamma_N$ denotes the Dirichlet boundary part. We assume that $\Omega \cup \Gamma_N$ is regular in the sense of Gröger [29] such that every chart map in the definition of regularity in the sense of Gröger can be chosen volume-preserving. The time interval $I = (0, T)$ with $T > 0$ is fixed.

In the following we denote the space time cylinder by $Q := I \times \Omega$ and apply standard notation for Hölder-, Lebesgue-, and Sobolev-spaces on Ω. The conjugate exponent of some integrability exponent p is denoted by p', and likewise for all other appearing integrability exponents. Since Ω stays fixed we omit it when referring to function spaces on Ω. Moreover, by the subscript D we indicate that the respective function space carries homogeneous Dirichlet boundary conditions on Γ_D.

Assumption 1.2.

1. The coefficient function $\xi: \mathbb{R} \to \mathbb{R}$ is twice differentiable, with ξ'' being Lipschitz continuous on bounded subsets of $\mathbb{R}$ and $0 < \xi_\bullet \leq \xi \leq \xi^\bullet$. Let $\mu: \Omega \to \mathbb{R}^{d \times d}$ be measurable, symmetric, and uniformly bounded and coercive in the following sense:

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

   We define the nonlinearity in (Eq) as follows:

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

2. 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.

This choice of p is fixed from now on. Assumptions 1.1 and 1.2 certainly impose non-trivial conditions on the geometry of the domain, the elliptic operator $-\nabla \cdot \mu\nabla + 1$, and the boundary conditions. We refer the reader to [5, Remarks 2.1 and 2.3] or [35, Example 2.3] for a discussion and examples.

Assumption 1.3. We fix regularization parameters $\alpha, \beta > 0$ and some $s > 2$ such that $\frac{1}{s} < \frac{1}{2}\left(1 - \frac{d}{p}\right)$ holds. Moreover, we choose an initial condition $y_0 \in (W_D^{-1,p}, W_D^{1,p})_{1-1/s, s}$ and the desired state $y_d \in L^\infty(I, L^p)$. The control operator is given by

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

where $b_i \in W_D^{-1,p}$, $i = 1, ..., m$, are fixed actuator functions. The set of admissible controls is given by

$$U_{\mathrm{ad}} := \{u \in L^s(I, \mathbb{R}^m): u_{a,i} \leq u_i \leq u_{b,i} \text{ a.e. on } I, \ i = 1, ..., m\},$$

with control bounds $u_a, u_b \in L^\infty(I, \mathbb{R}^m)$, $u_{a,i} \leq u_{b,i}$ a.e. on I for $i = 1, ..., m$.

Note that Assumptions 1.1-1.3 are identical to Assumptions 1-3 of [**5**], 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 (Eq), since Assumption 4 in [**5**] is related to the improved regularity analysis on Bessel-potential spaces. As pointed out in [**5**, Section 3] this analysis is not required for the first- and second-order analysis of Sections 3.1 and 4.1-4.3 of [**5**], except for [**5**, Proposition 4.7], a result concerning improved regularity of the adjoint state. We only rely on those results that are obtained completely within the $W_D^{-1,p}$-$W_D^{1,p}$-setting, cf. also [**47**, Theorem 5.3], and do not include the improved regularity assumptions of [**5**].

1.2. Main Results. Let us start the paper by mentioning the main results of our analysis of (P_k). The corresponding proofs rely on auxiliary material from Section 2 and are postponed to Section 3. Note that solutions to the state equation (Eq) have to be understood in the sense of (3.2); for details hereof we refer the reader to Section 3.1. Of course, we have to rely on well-posedness of (P_k). Due to convexity and continuity of j_k and boundedness of U_{ad}, existence of an optimal control for (P_k) is guaranteed by [**47**, Proposition 6.4]. Regarding first-order necessary optimality conditions for a $L^2(I, \mathbb{R}^m)$-local solution to (P_k) we will obtain:

Theorem 1.4 (First-order necessary optimality conditions). *Let $\bar{u}$ be a local solution to (P_k) w.r.t. the $L^2(I, \mathbb{R}^m)$-topology. Then there exists a unique, so-called adjoint state $\bar{p} \in W^{1,r}(I, L^p) \cap L^r(I, \mathrm{Dom}_{L^p}(-\nabla \cdot \mu\nabla))$, $r \in (1, \infty)$, such that*

$$-\partial_t \bar{p} - \xi(\bar{y})\nabla \cdot \mu\nabla \bar{p} = \bar{y} - y_d \qquad \text{on } Q,$$
$$\bar{p}(T) = 0 \qquad \text{on } \Omega,$$

and a unique $\bar{\lambda} \in \partial j_k(\bar{u})$ (see formulas (3.6)-(3.12)), such that the variational inequality

$$(1.1) \qquad \int_I (B^*\bar{p} + \alpha\bar{u} + \beta\bar{\lambda})^T(u - \bar{u})\mathrm{d}t \geq 0, \qquad \forall u \in U_{\mathrm{ad}},$$

is satisfied. In particular, it holds for the respective cases $k = 1, ..., 7$:

$$\bar{u}_i(t) = 0, \qquad \text{if and only if}$$

1. $|(B^*\bar{p})_i(t)| \leq \beta$,
2. $\|(B^*\bar{p})_i\|_{L^2(I)} \leq \beta$,
3. $|B^*\bar{p}(t)|_2 \leq \beta$,
4. $|(B^*\bar{p})_i(t)| \leq \beta\gamma_i$, with $\gamma_i = \frac{\|\bar{u}_i\|_{L^1(I)}}{|(\|\bar{u}_i\|_{L^1(I)})_i|_2}$, if $\bar{u} \neq 0$, and $\gamma_i = 1$, otherwise,
5. $|(B^*\bar{p})_i(t)| \leq \beta\gamma(t)$, with $\gamma(t) = \frac{|\bar{u}(t)|_1}{\||u(\cdot)|_1\|_{L^2(I)}}$, if $\bar{u} \neq 0$, and $\gamma(t) = 1$, otherwise,
6. $|(B^*\bar{p})_i(t)| \leq \beta\gamma_i$, with $\gamma_i = \|\bar{u}_i\|_{L^1(I)}$ if $\bar{u} \neq 0$, and $\gamma_i = 1$, otherwise,
7. $|(B^*\bar{p})_i(t)| \leq \beta\gamma(t)$, with $\gamma(t) = |\bar{u}(t)|_1$ if $\bar{u} \neq 0$, and $\gamma(t) = 1$, otherwise.

Speaking more illustratively, we can describe the different sparsity patterns as follows:

j_1 — **"Sparsity"**: This approach ensures both sparsity of the number of actuators and the time intervals at which they are active. However, there is no further structure in this sparsity.

j_2 — **"Sparse time-global selection of actuators"**: This approach selects a subset of the actuators that are allowed to be active. All other actuators are not used. The activity intervals of those actuators used are not sparse, in general.

j_3 — **"Sparsity in time of any control action"**: Any actuator, and then possibly all actuators, can become active only on a subset of I that is sparse.

j_4 — **"Sparse activity time for each actuator"**: An actuator i becomes active at some time point t only if its contribution at time point t is above a threshold depending on i. Therefore, the time of activity of each actuator i is sparse with a sparsity-pattern depending on i.

j_5 — **"Sparse selection of actuators at each time"**: At each time point t, an actuator i can be active only if its contribution is above a threshold depending on t. Therefore, at each time point t a certain sparse subset of actuators is selected to become active at t.

j_6 **and** j_7: These functionals result in similar sparsity patterns as j_4 and j_5, but with different thresholds that weigh the components differently.

Each of these possibilities may be of interest in certain applications. We point out that functionals j_6 and j_7 have an advantage compared to j_4 and j_5 from the perspective of fast numerical implementation, while the latter are superior in terms of interpretability. For details on that and numerical illustration of the sparsity patterns we refer the reader to Section 4. Analogous sparsity patterns are also obtained for $\alpha = 0$; cf. the results of Section 2.2. Until this point of our analysis we could indeed allow both for $\alpha > 0$ or $\alpha = 0$, to which we will refer from now on as the regular or the bang-bang case, respectively. For the formulation of the following second-order conditions, however, we have to restrict the analysis to the regular case, as will be explained in Section 3.4. For the definition of the reduced functional $\hat{J}$ we also refer to Section 3. Our second main result is the following:

Theorem 1.5 (No-gap second-order optimality conditions). *If $\bar{u} \in U_{\mathrm{ad}}$ is a local solution to* (P_k) *such that the reduced functional $\hat{J}$ fulfills*

$$(1.2) \qquad \hat{J}(u) \geq \hat{J}(\bar{u}) + \frac{c}{2}\|u - \bar{u}\|^2_{L^2(I,\mathbb{R}^m)}, \qquad \forall u \in U_{\mathrm{ad}} \ s.t. \ \|u - \bar{u}\|_{L^2(I,\mathbb{R}^m)} < \epsilon$$

with some $c \geq 0$ and $\epsilon > 0$, it holds:

$$(1.3) \qquad f''(\bar{u})v^2 + \beta j_k''(\bar{u}, v^2) \geq c\|v\|^2_{L^2(I,\mathbb{R}^m)}, \qquad \forall v \in C_{\bar{u}},$$

with $C_{\bar{u}} = \{v \in L^2(I,\mathbb{R}^m): \ v \geq 0, \ if \ \bar{u} = u_a, \ v \leq 0, \ if \ \bar{u} = u_b, \ f'(\bar{u})v + \beta j_k'(\bar{u}, v) = 0\}$, and f', f'', j_k', and j_k'' given by (3.3), (3.4), (3.13)-(3.19), *and* (3.20)-(3.26), *respectively. Conversely, let $\bar{u} \in U_{\mathrm{ad}}$ satisfy the first-order necessary optimality condition* (1.1) *and*

$$(1.4) \qquad f''(\bar{u})v^2 + \beta j_k''(\bar{u}, v^2) > 0, \qquad \forall v \in C_{\bar{u}} \setminus \{0\}.$$

Then there are $\epsilon, c > 0$ such that the quadratic growth condition (1.2) *holds true. In particular, $\bar{u}$ is a local solution to* (P_k) *w.r.t. the $L^2(I,\mathbb{R}^m)$-topology.*

Necessary and sufficient optimality conditions in the previous theorem have minimal gap. Positivity of the second derivative on the critical cone $C_{\bar{u}}$ is sufficient for local optimality, while nonnegativity on the same cone is necessary. As also observed in [13], the positivity condition (1.4) and the coercivity condition (1.3) are equivalent for the particular objective function. For no-gap second-order conditions for bang-bang problems with L^1-penalization we refer the reader to [59], and for different, non sparse, settings, e.g., to [19, 23, 24].

1.3. Overview over the remaining part of the paper. In order to reduce redundancy, we formulate as many of the results and arguments as possible on an abstract, general level, from which the concrete results can be obtained afterwards. In Section 2.1 we discuss optimality conditions for an abstract problem in the flavour of [19] by abstracting the main ideas from [12, 13, 17]. A more concrete instance hereof is dealt with in Section 2.2. There, we show that the previous results apply to a certain class of optimization problems on Lebesgue spaces with four different sparsity enforcing penalizations, and analyze the

resulting sparsity patterns of their solutions. Here, we rely again on [**12**, **13**, **17**]. In Section 3 we finally prove our main results by applying the framework of Section 2.2 to (P_k). Numerical experiments utilizing proximal methods and subgradient descent are peformed in Section 4.

2. First- and second-order optimality conditions on an abstract level

This section prepares the proofs of our main theorems in Section 3. As a first step, we analyze in Section 2.1 first- and second-order optimality conditions for a purely abstract optimization problem whose functional is given by the sum of a smooth, but nonconvex, and a convex, but nonsmooth term. More precisely, we extend the abstract framework for smooth functionals from [**19**], where, e.g., semilinear parabolic problems without sparsity have been considered, toward the inclusion of nonsmooth, but convex summands that satisfy certain properties that are typical for sparsity promoting regularizers as, e.g., j_1-j_7. The results are obtained utilizing the techniques of [**12**, **13**, **17**], and may therefore also be viewed as a summary of these earlier results on an abstract level. Thereafter, in Section 2.2, we make the problem under consideration a bit more concrete and deal with optimization problems on Lebesgue-spaces with directional sparsity enforcing penalization terms. Following [**12**, **13**, **17**] we verify that these problems fit into the framework of Section 2.1, and analyze the corresponding sparsity patterns of the solutions.

2.1. Optimality conditions for an abstract nonsmooth and nonconvex problem. Let us define the following optimization problem:

$$(\mathrm{P}_1) \qquad\qquad \min_{u \in K} \hat{J}(u) := f(u) + g(u),$$

where K is a closed convex set in a Banach space U_∞, f and g are real-valued functionals, such that f is smooth, but possibly nonconvex, and g is convex and Lipschitz, but not necessarily smooth. Since we have in mind the concrete situation of $\hat{J}$ being the reduced functional of a sparse PDE-constrained optimal control problem, we include a two-norm discrepancy [**37**]. Coercivity of second derivatives can only be expected w.r.t. a weaker norm in a Hilbert space $U_2 \supset U_\infty$. In applications we usually expect U_2 to be an L^2-space, while U_∞ is an L^s-space with $s \in (1, \infty]$. The precise setting is described in detail below. In particular, our assumptions on the smooth functional f are identical to those in [**19**], and cover a broad range of functionals arising from PDE-constrained optimization; see for instance [**5**, **19**]. Hence, we generalize the result from [**19**] for the smooth case $\hat{J} = f$ to the inclusion of a nonsmooth summand g in the functional. Our approach differs from the similar abstract approach in [**59**] by several technical aspects. Our setting includes the presence of two nonequivalent norms, but instead of working with a Banach space and its predual as in [**59**] we restrict ourselves to formulating optimality conditions w.r.t. the Hilbert space U_2. Moreover, unlike in [**59**] we do not include the convex constraint "$u \in K$" as indicator function in the nonsmooth part of the functional. We will show in Section 2.2 that our assumptions on g are typically fulfilled by penalizers promoting directional sparsity, while the applications discussed in [**59**] are primarily concerned with different, non-uniformly convex integral functionals. The proofs and assumptions of this section are inspired by [**19**] and well known techniques employed in particular in [**12**, **13**, **17**].

Assumption 2.1. Let U_2 be a Hilbert space and U_∞ a Banach space such that $U_\infty \hookrightarrow U_2$. With $\|\cdot\|_\infty$, $\|\cdot\|_2$, and $\langle \cdot, \cdot \rangle_2$ we denote the corresponding norms and the duality pairing on $U_2^* \times U_2$. Let $\emptyset \neq K \subset U_\infty$ be convex and $A \supset K$ be open in U_∞. We fix $\bar{u} \in K$.

    1. The functional $f \colon A \to \mathbb{R}$ is assumed to be twice continuously Fréchet differentiable w.r.t. $\|\cdot\|_\infty$ and to fulfill the following properties:

1a. The derivatives of f taken w.r.t. the space U_∞ extend to continuous linear, resp. bilinear, forms on U_2, i.e.

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

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

1bi. $f'(\bar{u})v = \lim_{k \to \infty} f'(u_k)v_k$

1bii. $f''(\bar{u})v^2 \le \liminf_{k \to \infty} f''(u_k)v_k^2$

1biii. If $v = 0$, there is some $\gamma > 0$ such that

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

2. The functional $g \colon U_2 \to \mathbb{R}$ is assumed to be convex and Lipschitz continuous. By g' and ∂g we denote its directional derivatives and subgradient; see, e.g., [**28**, Chapter I.5] for the definitions. We introduce the following sets

$$D_{\bar{u}} := \mathrm{cl}_{U_2} \left(\{ v \in \mathcal{R}_K(\bar{u}) \colon f'(\bar{u})v + g'(\bar{u}, v) = 0 \} \right),$$

$$C_{\bar{u}} := \mathcal{T}_K(\bar{u}) \cap \{ v \in U_2 \colon f'(\bar{u})v + g'(\bar{u}, v) = 0 \},$$

where $\mathcal{R}_K(\bar{u})$ and $\mathcal{T}_K(\bar{u})$ denote the radial and tangent cone of K at $\bar{u}$; see, e.g., [**6**, Definition 2.54] for the definition. Moreover, let $g''(\bar{u}, \cdot) \colon U_2 \to \mathbb{R}$ denote a continuous quadratic form such that:

2a. If $v \in D_{\bar{u}}$ there is a sequence $(v_k)_k \subset U_2$ such that $f'(\bar{u})v_k + g'(\bar{u}, v_k) = 0$, $v_k \to v$ in U_2, $u_k := \bar{u} + t_k v_k \in K$, $t_k \searrow 0$, $u_k \to \bar{u}$ in U_∞, and

$$g''(\bar{u}, v^2) \ge \lim_{k \to \infty} \frac{2}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right).$$

2b. If $(v_k) \subset U_2$, $(t_k) \subset \mathbb{R}_{>0}$ such that $t_k \searrow 0$, $v_k \rightharpoonup v$ weakly in U_2 with $v \in C_{\bar{u}}$, $g'(\bar{u}, v_k) \to g'(\bar{u}, v)$, and $\bar{u} + t_k v_k \in K$, it holds

$$g''(\bar{u}, v^2) \le \liminf_{k \to \infty} \frac{2}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right).$$

We start with a discussion of first-order necessary optimality conditions:

Theorem 2.2. *Let Assumption 2.1.1a hold and suppose that $\bar{u}$ is a local minimizer of* (P_1) *w.r.t. the U_2-topology. Then there is $\bar{\lambda} \in \partial g(\bar{u})$ such that*

$$(2.1) \qquad \langle f'(\bar{u}) + \bar{\lambda}, u - \bar{u} \rangle_2 \ge 0, \qquad \forall u \in K.$$

The proof works completely analogous to for instance the proof of [**13**, Theorem 2.1]. For convenience of the reader we provide the main steps.

Proof. Given $u \in K$, it holds $\hat{J}(\bar{u} + t(u - \bar{u})) - \hat{J}(\bar{u}) \ge 0$ for all sufficiently small $t \in (0, 1)$, due to local optimality of $\bar{u}$ and convexity of K. From convexity of g we infer $g(u) - g(\bar{u}) \ge t^{-1}[g(\bar{u} + t(u - \bar{u})) - g(\bar{u})]$ for all $t \in (0, 1)$, and together with differentiability of f we therefore obtain:

$$f'(\bar{u})(u - \bar{u}) + g(u) - g(\bar{u}) \ge \liminf_{t \searrow 0} t^{-1}[\hat{J}(\bar{u} + t(u - \bar{u})) - \hat{J}(\bar{u})] \ge 0, \qquad \forall u \in K.$$

In particular, $\bar{u}$ is a solution of the following optimization problem on U_2:

$$\min_{u \in K} \; f'(\bar{u})u + g(u).$$

The map $f'(\bar{u}) + g \colon U_2 \to \mathbb{R}$ is convex and continuous, and hence the claim follows from standard convex analysis; see, e.g., [**28**, Proposition I.5.6]. $\qquad \square$

Before addressing second-order optimality conditions, some comments on $D_{\bar{u}}$ and $C_{\bar{u}}$ seem to be appropriate. First, let us note:

Lemma 2.3. *Let $\bar{u}$ and $\bar{\lambda}$ satisfy (2.1). Then $C_{\bar{u}}$ is a closed, convex cone in U_2. Moreover, it holds $D_{\bar{u}} \subset C_{\bar{u}}$ and $g'(\bar{u}, v) = \langle \bar{\lambda}, v \rangle_2$ for all $v \in C_{\bar{u}}$.*

We can follow, e.g., [**12**, Proposition 3.4] to prove this.

Proof. Closedness and the cone-property of $C_{\bar{u}}$, as well as the inclusion $D_{\bar{u}} \subset C_{\bar{u}}$ are obvious consequences of the definition. Herein, note that $\mathcal{T}_K(\bar{u}) = \mathrm{cl}_{U_2}(\mathcal{R}_K(\bar{u}))$; cf. [**6**, Proposition 2.55]. Moreover, (2.1) implies

$$f'(\bar{u})v + g'(\bar{u}, v) \overset{(\star)}{\geq} \langle f'(\bar{u}) + \bar{\lambda}, v \rangle_2 \geq 0, \qquad \forall v \in \mathcal{T}_K(\bar{u}),$$

and that equality in $(\star)$, and hence $g'(\bar{u}, v) = \langle \bar{\lambda}, v \rangle_2$, holds for $v \in C_{\bar{u}}$. For $v_1, v_2 \in C_{\bar{u}}$ and $t \in (0, 1)$ we conclude, utilizing convexity of $g(\bar{u}, \cdot)$ and convexity of $\mathcal{T}_K(\bar{u})$, that

$$0 \leq f'(\bar{u})(tv_1 + (1-t)v_2) + g'(\bar{u}, tv_1 + (1-t)v_2)$$
$$\leq t[f'(\bar{u})v_1 + g'(\bar{u}, v_1)] + (1-t)[f'(\bar{u})v_2 + g'(\bar{u}, v_2)] \leq 0,$$

i.e. $tv_1 + (1-t)v_2 \in C_{\bar{u}}$. This proves convexity of $C_{\bar{u}}$. $\qquad \square$

Due to the fact that $g'(\bar{u}, \cdot)$ is not a linear form on U_2, we cannot apply the concept of polyhedricity, see, e.g., [**60**] for the definition, directly. This is different from the smooth case in [**19**], where $D_{\bar{u}} = C_{\bar{u}}$ holds for polyhedric K. In fact, we do not know whether this equality still holds true in our abstract nonsmooth setting. Since in the following sufficient conditions will be formulated on the cone $C_{\bar{u}}$, and necessary conditions on the possibly smaller cone $D_{\bar{u}}$, we do not obtain no-gap second-order conditions for the fully abstract setting. However, for sparse optimization problems on Lebesgue-spaces equality holds, cf. Section 2.2, because Assumption 2.1.2a can be verified with $D_{\bar{u}}$ replaced by $C_{\bar{u}}$ in these cases. The following necessary second-order optimality condition is the abstract version of the first part of Theorem 1.5:

Theorem 2.4. *Let Assumption 2.1.1 and 2a hold, and suppose that there are $c \geq 0$ and $r > 0$ such that*

$$\hat{J}(u) \geq \hat{J}(\bar{u}) + \frac{c}{2}\|u - \bar{u}\|_2^2, \qquad \forall u \in K \cap B_r^{U_2}(\bar{u}).$$

Then it holds

$$f''(\bar{u})v^2 + g''(\bar{u}, v^2) \geq c\|v\|_2^2, \qquad \forall v \in D_{\bar{u}}.$$

Proof. Fix $v \in D_{\bar{u}}$. Due to Assumption 2.1.2a there is a sequence $(v_k)_k \subset U_2$, such that $v_k \to v$ strongly in U_2, $u_k = \bar{u} + t_k v_k \in K$, $f'(\bar{u})v + g'(\bar{u}, v) = 0$, $t_k \searrow 0$, and $u_k \to \bar{u}$ in U_∞. It holds

$$\frac{c}{2} t_k^2 \|v_k\|_2^2 = \frac{c}{2}\|u_k - \bar{u}\|_2 \leq \hat{J}(u_k) - \hat{J}(\bar{u}) = f(u_k) - f(\bar{u}) + g(u_k) - g(\bar{u})$$

$$= f'(\bar{u})(u_k - \bar{u}) + \frac{1}{2}f''(u_k^\theta)(u_k - \bar{u})^2 + (g(u_k) - g(\bar{u}) - g'(\bar{u}, u_k - \bar{u})) + g'(\bar{u}, u_k - \bar{u}),$$

by assumption and Taylor expansion of f at $\bar{u}$ with some $u_k^\theta := (1 - \theta_k)\bar{u} + \theta_k u_k$, $\theta_k \in [0, 1]$. Exploiting that $f'(\bar{u})(u_k - \bar{u}) + g'(\bar{u}, u_k - \bar{u}) = t_k[f'(\bar{u})v_k + g'(\bar{u}, v_k)] = 0$ and dividing by t_k^2 yields:

$$(2.2) \qquad \frac{c}{2}\|v_k\|_2^2 \leq \frac{1}{2}f''(u_k^\theta)v_k^2 + \frac{1}{t_k^2}\left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k)\right).$$

Taking the limes inferior on both sides of (2.2) and utilizing Assumption 2.1.1.bii and 2a concludes the proof. $\qquad \square$

Apparently, the above proof also keeps valid with slightly modified assumptions. If Assumption 2.1.2a is weakened by replacing the limes by limes inferior, we need to strengthen Assumption 2.1.1 by demanding additional continuity of f'' as map $U_\infty \to \mathcal{L}(U_2 \otimes U_2, \mathbb{R})$. Note that this weaker version of Assumption 2.1.2 would be fulfilled in particular if $g\colon U_2 \to \mathbb{R}$ had second-order Mosco-epi-derivative $g''(u, \cdot)$; see, e.g., [**26, 48**] for the definition.

Next, we state and prove sufficient optimality conditions, that correspond —on the abstract level— to the second part of Theorem 1.5:

Theorem 2.5. *Let Assumption 2.1 hold, and suppose that there is $\bar{\lambda} \in \partial g(\bar{u})$ such that the first-order necessary optimality condition in Theorem 2.2 is satisfied. If in addition*

$$f''(\bar{u}) + g''(\bar{u}, v^2) > 0, \qquad \forall v \in C_{\bar{u}} \setminus \{0\},$$

holds, there are $c, r > 0$ such that

$$\hat{J}(u) \geq \hat{J}(\bar{u}) + \frac{c}{2}\|u - \bar{u}\|_2^2, \qquad \forall u \in B_r^{U_2}(\bar{u}).$$

In particular, $\bar{u}$ is a U_2-local solution of (P_1).

Proof. We argue by contradiction, following, e.g., the well known approach in [**13, 17, 19**]. If the statement of the theorem is not true there is a sequence $(u_k)_k \subset K$ such that $u_k \to \bar{u}$ in U_2 and

$$(2.3) \qquad \frac{1}{2k}\|u_k - \bar{u}\|_2^2 > \hat{J}(u_k) - \hat{J}(\bar{u}) = f(u_k) - f(\bar{u}) + g(u_k) - g(\bar{u}).$$

We set $t_k := \|u_k - \bar{u}\|_2$, $v_k := t_k^{-1}(u_k - \bar{u})$ and assume w.l.o.g. that $v_k \rightharpoonup v \in U_2$ weakly. The contradiction is achieved in three steps I.-III.:

I. First, we show that $v \in C_{\bar{u}}$. It clearly holds $v \in \text{weak-cl}_{U_2}(\mathcal{R}_K(\bar{u}))$. Since $\mathcal{R}_K(\bar{u})$ is convex due to convexity of K it follows that the weak and strong closure of $\mathcal{R}_K(\bar{u})$ coincide, cf. [**6**, Theorem 2.23ii], from which we deduce $v \in \mathcal{T}_K(\bar{u})$. From the first-order necessary optimality condition $\langle f'(\bar{u}) + \bar{\lambda}, u_k - \bar{u}\rangle_2 \geq 0$ together with $f'(\bar{u}) + \bar{\lambda} \in U_2^*$ and weak convergence of $(v_k)_k$ we immediately conclude $\langle f'(\bar{u}) + \bar{\lambda}, v\rangle_2 \geq 0$. The subgradient property therefore implies $f'(\bar{u})v_k + g'(\bar{u}, v_k) \geq 0$ and $f'(\bar{u})v + g'(\bar{u}, v) \geq 0$. Applying Taylor expansion to f at $\bar{u}$ we obtain from (2.3)

$$\frac{t_k^2}{2k} > \hat{J}(u_k) - \hat{J}(\bar{u}) \geq f'(u_k^\theta)(u_k - \bar{u}) + g(u_k) - g(\bar{u}),$$

where $u_k^\theta = (1 - \theta_k)\bar{u} + \theta_k u_k$, $\theta_k \in [0, 1]$. Dividing by $t_k > 0$, this leads to

$$f'(u_k^\theta)v_k + g'(\bar{u}, v_k) \leq f'(u_k^\theta)v_k + t_k^{-1}[g(u_k) - g(\bar{u})] \leq \frac{t_k}{2k} \to 0, \qquad \text{as } k \to \infty.$$

Taking the limes inferior on both sides hereof and using Assumption 2.1.1bi for the first summand on the left-hand side we obtain

$$f'(\bar{u})v + g'(\bar{u}, v) = f'(\bar{u})v + \liminf_{k \to \infty} t_k^{-1}[g(\bar{u} + t_k v_k) - g(\bar{u})] \leq 0,$$

where we have used that $g'(\bar{u}, \cdot)$ is convex and continuous as in [**17**, Proof of Theorem 4.2]. Hence this shows $v \in C_{\bar{u}}$. Moreover, we have $-f'(\bar{u})v_k \leq g'(\bar{u}, v_k) \leq \frac{t_k}{2k} - f'(u_k^\theta)v_k$, and hence $g'(\bar{u}, v_k) \to -f'(\bar{u})v = g'(\bar{u}, v)$ as $k \to \infty$.

II. Next, we prove $v = 0$. Again, we apply Taylor expansion to f in (2.3) and obtain with some $\tilde{u}_k^\theta = (1 - \tilde{\theta}_k)\bar{u} + \tilde{\theta}_k u_k$, $\tilde{\theta}_k \in [0, 1]$:

$$\frac{t_k^2}{2k} > f'(\bar{u})(u_k - \bar{u}) + \frac{1}{2}f''(\tilde{u}_k^\theta)(u_k - \bar{u})^2 + g'(\bar{u}, u_k - \bar{u}) + (g(u_k) - g(\bar{u}) - g'(\bar{u}, u_k - \bar{u}))$$

$$\geq \frac{1}{2}f''(\tilde{u}_k^\theta)(u_k - \bar{u})^2 + (g(u_k) - g(\bar{u}) - g'(\bar{u}, u_k - \bar{u})).$$

Here, we have used the first-order necessary condition in the second inequality. Dividing by t_k^2 and taking the limes inferior on both sides yields:

$$0 \geq \liminf_{k \to \infty} \left(\frac{1}{2} f''(\tilde{u}_k^\theta) v_k^2 + \frac{1}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right) \right)$$

$$\geq \liminf_{k \to \infty} \frac{1}{2} f''(\tilde{u}_k^\theta) v_k^2 + \liminf_{k \to \infty} \frac{1}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right)$$

$$\geq \frac{1}{2} f''(\bar{u}) v^2 + \frac{1}{2} g''(\bar{u}, v^2),$$

where we have applied Assumptions 2.1.1bii and 2b in the last step. Due to $v \in C_{\bar{u}}$ it follows from the assumption of the theorem that $v = 0$.

III. In this final step we arrive at the desired contradiction. From Assumption 2.1.1biii and $\|v_k\|_2 = 1$ we infer from the above considerations:

$$0 < \gamma \leq \liminf_{k \to \infty} f''(\tilde{u}_k^\theta) v_k^2 \leq \liminf_{k \to \infty} \left(\frac{1}{2k} - \frac{1}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right) \right)$$

$$\leq - \limsup_{k \to \infty} \frac{1}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right).$$

Since the term inside the limes superior is always nonnegative due to convexity of g we arrive at the desired contradiction $0 < \gamma \leq 0$. □

The crucial observation in the final step of the proof of Theorem 2.5 is the inequality

$$\liminf_{k \to \infty} f''(\tilde{u}_k^\theta) v_k^2 \leq - \limsup_{k \to \infty} \frac{1}{t_k^2} \left(g(\bar{u} + t_k v_k) - g(\bar{u}) - t_k g'(\bar{u}, v_k) \right).$$

Assumption 2.1.2biii ensures positivity of the left-hand side, while convexity of g implies non-positivity of the right-hand side. Without Assumption 2.1.2biii, we would only have nonnegativity of the left-hand side, which does not suffice to achieve a contradiction, unless the right-hand side could be shown to be negative. However, we think that the latter can only hold for g being strongly convex at $\bar{u}$ w.r.t. the U_2-norm. A strongly convex function g, however, is the sum of a convex function and a U_2-Tikhonov term $u \mapsto \frac{\gamma}{2} \|u\|_2^2$. Such a Tikhonov-term is smooth and would ensure Assumption 2.1.2b when being shifted to f. This is the reason why the application of Theorem 2.5 will be restricted to the regular case in the subsequent sections. As explained in Section 3.4, we expect that a different type of argument is needed for the the bang-bang case.

Finally, we mention that a variant of Theorem 2.5 with norm-gap can also be obtained. Let strong convergence $u_k \to \bar{u}$ in Assumption 2.1.1b hold only w.r.t. another Banach space Z such that $Z \hookrightarrow U_\infty$, e.g., $Z = U_\infty$. Under this weaker supposition, the quadratic growth condition in Theorem 2.5 holds true in a Z-neighbourhood of $\bar{u}$, and, consequently, $\bar{u}$ is a Z-local solution to (P_1).

2.2. Optimality conditions for directionally sparse optimization on Lebesgue-spaces. In this section we incorporate directionally sparse optimization problems on Lebesgue-spaces into the framework established before. We replace (P_1) by the following slightly more concrete model problem that contains (P_k) as an instance: Given a complete, σ-finite measure space (Λ, ρ) we consider

$$(P_2\text{-}X) \qquad \min_{u \in U_{\mathrm{ad}}} \hat{J}(u) := \underbrace{h(u) + \frac{\alpha}{2} \|u\|_{L^2(\Lambda)}^2}_{=:f(u)} + \underbrace{\beta j_X(u)}_{=:g(u)}, \qquad X \in \{A, B, C, D\},$$

with the following four typical (directional) sparsity enforcing functionals j_X:

A. $j_A = \|\cdot\|_{L^1(\Lambda)}$,
B. $j_B = \|\cdot\|_{L^1(\Lambda_1, L^2(\Lambda_2))}$,

  C. $j_C = \|\cdot\|_{L^2(\Lambda_1, L^1(\Lambda_2))}$,
  D. $j_D = \frac{1}{2}\|\cdot\|^2_{L^2(\Lambda_1, L^1(\Lambda_2))}$,

where in the cases B-D, (Λ, ρ) is given by the product measure space of two complete, σ-finite measure spaces (Λ_1, ρ_1) and (Λ_2, ρ_2). Moreover, let $\alpha \geq 0$, $\beta > 0$ and $f: L^s(\Lambda) \to \mathbb{R}$ fulfill Assumption 2.1.1 with $U_\infty = L^s(\Lambda)$, $s > 2$, $U_2 = L^2(\Lambda)$, and $K = U_{\mathrm{ad}} := \{u \in L^s(\Lambda): u_a \leq u \leq u_b \ \rho$-a.e. on $\Lambda\}$ with $u_a, u_b \in L^\infty(\Lambda)$.

Note that functionals j_1-j_7 from Section 1.2 are included in this setting by an appropriate choice of Λ_1 and Λ_2; cf. Section 3.2 below. In Section 3, h will be given by the first summand of the reduced functional of (P_k). Nevertheless, the particular choice of h does not matter for the arguments of the present section.

We analyze first-order optimality conditions for $(\mathrm{P}_2$-$X)$ together with the resulting sparsity patterns of the minimizers, and verify that functionals j_X, $X \in \{A, B, C, D\}$ fit into the framework of Section 2.1. Except for case D, all these results have already been obtained in [**13**,**17**,**31**,**56**] dealing with linear and semilinear problems in the more concrete setting. There, Λ is a domain in $\mathbb{R}^d$ or a space-time cylinder, equipped with the Lebesgue measure, and h is a smooth tracking-type functional originating from optimal control of a linear or semilinear PDE. For functional j_3 in the context of optimal control of an ordinary differential equation we refer the reader to [**54**]. The proofs also apply to our abstract setting, but for convenience of the reader we repeat these results in our notation. We also mention that discrete analoga of functionals A and B are well known in the machine learning as "lasso" [**57**] and "group lasso" [**61**].

To the best of our knowledge, case D has not been analyzed in the context of PDE-constrained optimization so far. It can be motivated by the successfull use of analogous functionals in the discrete setting, e.g., the so-called "exclusive lasso" [**7**] in machine learning, or the sparse regression problem in [**39**]. In particular, j_D results in sparsity patterns similar to j_C, but, as well known in the discrete case, j_D unlike j_C allows the application of proximal algorithms; cf. Section 4.1. Therefore, we may view case D as an alternative to case C that also deserves a theoretical analysis.

In the following we repeatedly make use of the fact that $h'(\bar{u}) \in L^2(\Lambda)^*$ can be identified with its Riesz-representative $\nabla h(u) \in L^2(\Lambda)$. Moreover, we use the set-valued sign function $\mathrm{sign}(z) = \{\pm 1\}$ for $z \gtrless 0$, $\mathrm{sign}(0) = [-1, 1]$.

2.2.1. *Functional A.* Both, directional derivatives and subdifferential of j_A, are well known; see for instance [**12**]. Note that the proofs that originally pertain to Lebesgue-spaces on open sets of $\mathbb{R}^d$ apply to our slightly more general setting without changes. Therefore, for some $u, v \in L^1(\Lambda)$ the directional derivatives of $\tilde{j}_A: L^1(\Lambda) \to \mathbb{R}$ are given by

$$(2.4) \qquad \tilde{j}'_A(u, v) = \int_{\Lambda^+(u)} v \, \mathrm{d}\rho - \int_{\Lambda^-(u)} v \, \mathrm{d}\rho + \int_{\Lambda^0(u)} |v| \, \mathrm{d}\rho,$$

where we use the notation $\Lambda^0(u) := \{x \in \Lambda: u(x) = 0\}$ and $\Lambda^\pm(u) := \{x \in \Lambda: u(x) \gtrless 0\}$. The subdifferential is given by

$$(2.5) \qquad \partial \tilde{j}_A(u) = \{\lambda \in L^\infty(\Lambda): \ \lambda(x) \in \mathrm{sign}(u(x)) \quad \rho\text{-a.e. on } \Lambda\}.$$

Here, recall that $L^1(\Lambda)^* = L^\infty(\Lambda)$ for any σ-finite measure space (Λ, ρ); cf. [**51**, Satz 6.16]. However, note that we will actually be concerned with $j_A = \tilde{j}_A \circ \iota$ where $\iota: L^2(\Lambda) \hookrightarrow L^1(\Lambda)$ denotes the canonical embedding. It is obvious, that the formulas for the directional derivatives remain true for j_A. Regarding the subdifferential, recall that by the chain rule, see, e.g., [**28**, Proposition 5.7], it holds $\partial j_A(u) = \partial(\tilde{j}_A \circ \iota)(u) = \iota^* \partial \tilde{j}_A(\iota u) = \iota^* \partial \tilde{j}_A(u)$, which implies that the above characterization of the subdifferential is also valid for j_A, because ι^* acts as the embedding $L^\infty(\Lambda) \hookrightarrow L^2(\Lambda)$.

Proposition 2.6.   *Let $\bar{u} \in U_{\mathrm{ad}}$ be a local solution to $(\mathrm{P}_2$-$A)$.*

1. *If $\alpha > 0$, it holds ρ-a.e. on Λ:*

$$\bar{u}(x) = 0 \quad \Leftrightarrow \quad |\bar{\nabla} h(\bar{u})(x)| \leq \beta,$$
$$\bar{\lambda} = \operatorname{Proj}_{[-1,1]}\left(-\beta^{-1} \nabla h(\bar{u})\right),$$
$$\bar{u} = \operatorname{Proj}_{[u_a,u_b]}\left[-\alpha^{-1}\left(\nabla h(\bar{u}) + \beta \bar{\lambda}\right)\right].$$

If $\alpha = 0$, it holds for ρ-a.e. on Λ:

$$|\nabla h(\bar{u})(x)| < \beta \quad \Rightarrow \quad \bar{u}(x) = 0,$$
$$\nabla h(\bar{u})(x) > \beta \quad \Rightarrow \quad \bar{u}(x) = u_a(x),$$
$$\nabla h(\bar{u})(x) < -\beta \quad \Rightarrow \quad \bar{u}(x) = u_b(x).$$

2. *$g = \beta j_A$ satisfies the properties of Assumption 2.1.2 with $D_{\bar{u}}$ replaced by $C_{\bar{u}}$ and $g''(\bar{u}, v^2) \equiv 0$ for all $v \in L^2(\Lambda)$.*

Proof. The first-order conditions and the analysis of the sparsity pattern can be found in [**12**, Corollary 3.2]. Regarding 2., Assumption 2.1.2a is verified in the proof of [**12**, Theorem 3.7], while Assumption 2.1.2b is an immediate consequence of the convexity of j.   □

2.2.2. *Functional B.* This functional has been discussed in [**13**, **17**] for the special case that Λ_1 is an interval and Λ_2 is a domain in $\mathbb{R}^d$, both equipped with the Lebesgue measure. We refer the reader to [**54**] for the particular case j_3 in the context of an optimal control problem with ODE-constraints. The results and their proofs also apply to our setting. Using the notation $\Lambda_1^0(u) = \{x_1 \in \Lambda_1 : \|u(x_1, \cdot)\|_{L^2(\Lambda_2)} = 0\}$, we obtain the directional derivatives and subgradients of $\tilde{j}_B \colon L^1(\Lambda_1, L^2(\Lambda_2)) \to \mathbb{R}$; cf. [**13**, Proposition 2.8]:

$$(2.6) \quad \tilde{j}_B'(u, v) = \int_{\Lambda_1^0(u)} \|v(x_1, \cdot)\|_{L^2(\Lambda_2)} \mathrm{d}\rho_1(x_1)$$
$$+ \int_{\Lambda_1 \setminus \Lambda_1^0(u)} \frac{1}{\|u(x_1, \cdot)\|_{L^2(\Lambda_2)}} \int_{\Lambda_2} u(x_1, x_2) v(x_1, x_2) \mathrm{d}\rho_2(x_2) \mathrm{d}\rho_1(x_1),$$
$$(2.7) \quad \partial \tilde{j}_B(u) = \{\lambda \in L^\infty(\Lambda_1, L^2(\Lambda_2)):$$
$$\lambda(x_1, \cdot) \begin{cases} \in \mathbb{B}_1^{L^2(\Lambda_2)}(0), & \text{if } x_1 \in \Lambda_1^0(u), \\ = \frac{u(x_1, \cdot)}{\|u\|_{L^1(\Lambda_1, L^2(\Lambda_2))}}, & \text{if } x_1 \notin \Lambda_1^0(u), \end{cases} \quad \rho_1\text{-a.e. on } \Lambda_1\}.$$

Here, note that $L^1(\Lambda_1, L^2(\Lambda_2))^* = L^\infty(\Lambda_1, L^2(\Lambda_2))$; cf. [**27**, Theorem 8.18.3]. As for case A, we obtain the representation of the subdifferential of j_B on $L^2(\Lambda)$ by an application of the chain-rule.

Proposition 2.7.   *Let $\bar{u} \in U_{\mathrm{ad}}$ be a local solution to (P$_2$-B).*

1. *If $\alpha > 0$, it holds ρ_1-a.e. on Λ_1 or ρ-a.e. on Λ, respectively:*

$$\|u(x_1, \cdot)\|_{L^2(\Lambda_2)} = 0 \quad \Leftrightarrow \quad \|\nabla h(\bar{u})(x_1, \cdot)\|_{L^2(\Lambda_2)} \leq \beta,$$
$$\bar{\lambda}(x_1, x_2) = \begin{cases} -\beta^{-1} \nabla h(\bar{u})(x_1, x_2), & \text{if } x_1 \in \Lambda_1^0(u), \\ \frac{u(x_1, x_2)}{\|\bar{u}(x_1, \cdot)\|_{L^2(\Lambda_2)}}, & \text{if } x_1 \in \Lambda_1 \setminus \Lambda_1^0(u), \end{cases}$$
$$\bar{u} = \operatorname{Proj}_{[u_a,u_b]}\left(-\alpha^{-1}\left(\nabla h(\bar{u}) + \beta \bar{\lambda}\right)\right).$$

If $\alpha = 0$, it holds ρ_1-a.e. on Λ_1:

$$\|\nabla h(\bar{u})(x_1, \cdot)\|_{L^2(\Lambda_2)} < \beta \quad \Rightarrow \quad u(x_1, \cdot) \equiv 0,$$
$$\bar{u}(x_1, \cdot) \equiv 0 \quad \Rightarrow \quad \|\nabla h(\bar{u})(x_1, \cdot)\|_{L^2(\Lambda_2)} \leq \beta.$$

2. *$g = \beta j_B$ satisfies Assumption 2.1.2 with $D_{\bar{u}}$ replaced by $C_{\bar{u}}$,*

$$g''(\bar{u}, v^2) := \beta \int_{\Lambda_1 \backslash \Lambda_1^0(\bar{u})} \frac{1}{\|\bar{u}(x_1, \cdot)\|_{L^2(\Lambda_2)}} \left[\|v(x_1, \cdot)\|_{L^2(\Lambda_2)}^2 \right.$$
$$\left. - \left(\int_{\Lambda_2} \frac{\bar{u}(x_1, x_2)v(x_1, x_2)}{\|\bar{u}(x_1, \cdot)\|_{L^2(\Lambda_2)}} \mathrm{d}\rho_2(x_2) \right)^2 \right] \mathrm{d}\rho_1(x_1),$$

for $\bar{u} \neq 0$, and $g''(0, v^2) \equiv 0$ otherwise.

Proof. For the first part, see [**13**, Corollary 2.9]. For the second part, Case III in the proof of [**13**, Theorem 3.3] and [**17**, Section 4] prove Assumption 2.1.2a and 2b. □

2.2.3. *Functional C.* This functional has been addressed in [**13**] for the special case that Λ_1 is an interval and Λ_2 is a domain in $\mathbb{R}^d$, both equipped with the Lebesgue measure. The proof, however, also applies to our setting and we obtain expressions for the directional derivatives and the subdifferential of $\tilde{j}_C \colon L^2(\Lambda_1, L^1(\Lambda_2)) \to \mathbb{R}$ as follows; cf. [**13**, Proposition 2.4]:

$$(2.8) \qquad \tilde{j}_C'(u, v) = \frac{1}{\|u\|_{L^2(\Lambda_1, L^1(\Lambda_2))}} \int_{\Lambda_1} \|\cdot\|_{L^1(\Lambda_2)}'(u(x_1, \cdot), v(x_1, \cdot)) \mathrm{d}\rho_1(x_1).$$

Regarding the subdifferential, first note that $L^2(\Lambda_1, L^1(\Lambda_2))^* = L^2_{\text{w-}\star}(\Lambda_1, L^\infty(\Lambda_2))$, where the latter denotes the space of weak-$\star$ measurable functions $\Lambda_1 \to L^\infty(\Lambda_2)$, equipped with the $L^2(\Lambda_1, L^\infty(\Lambda_2))$-norm; cf. [**27**, Theorem 8.20.3]. The subdifferential is given by

$$(2.9) \quad \partial \tilde{j}_C(u) = \left\{ \lambda \in L^2_{\text{w-}\star}(\Lambda_1, L^\infty(\Lambda_2)) \colon \right.$$
$$\left. \lambda(x_1, x_2) \in \mathrm{sign}(u(x_1, x_2)) \frac{\|u(x_1, \cdot)\|_{L^1(\Lambda_2)}}{\|u\|_{L^2(\Lambda_1, L^1(\Lambda_2))}} \quad \rho\text{-a.e. on } \Lambda \right\}.$$

As for A and B, the formulas for the directional derivatives also stay true for j_C instead of $\tilde{j}_C$, because if we denote by ι the embedding $L^2(\Lambda_1 \times \Lambda_2) \hookrightarrow L^2(\Lambda_1, L^1(\Lambda_2))$, it follows $\partial j_C(u) = \iota^* \partial \tilde{j}_C(u)$, where ι^* is the embedding $L^2_{\text{w-}\star}(\Lambda_1, L^\infty(\Lambda_2)) \hookrightarrow L^2(\Lambda_1 \times \Lambda_2)$. Note that this embedding is a consequence of the separability of $L^2(\Lambda_2)$, Pettis' measurability theorem [**27**, Theorem 8.15.2], and Fubini's theorem.

Proposition 2.8. *Let $\bar{u} \in U_{\text{ad}}$ be a local solution to $(P_2\text{-}C)$ and define*

$$\bar{\gamma}(x_1) = \frac{\|\bar{u}(x_1, \cdot)\|_{L^1(\Lambda_2)}}{\|\bar{u}\|_{L^2(\Lambda_1, L^1(\Lambda_2))}}, \text{ if } \bar{u} \neq 0, \qquad \gamma(x_1) = 1, \text{ else.}$$

1. *If $\alpha > 0$, it holds for ρ-a.a. $x \in \Lambda$:*
$$\bar{u}(x_1, x_2) = 0 \quad \Leftrightarrow \quad |\nabla h(\bar{u})(x_1, x_2)| \leq \beta \bar{\gamma}(x_1),$$
$$\bar{\lambda}(x_1, x_2) = \mathrm{Proj}_{[-\bar{\gamma}(x_1), \bar{\gamma}(x_1)]} \left(-\beta^{-1} \nabla h(\bar{u})(x_1, x_2) \right),$$
$$\bar{u}(x_1, x_2) = \mathrm{Proj}_{[u_a, u_b]} \left[-\alpha^{-1} \left(\nabla h(\bar{u})(x_1, x_2) + \beta \bar{\lambda}(x_1, x_2) \right) \right].$$

   If $\alpha = 0$, it holds for ρ-a.a. $x \in \Lambda$:
$$|\nabla h(\bar{u})(x)| < \beta \gamma(x_1) \quad \Rightarrow \quad \bar{u}(x) = 0,$$
$$\nabla h(\bar{u})(x) > \beta \gamma(x_1) \quad \Rightarrow \quad \bar{u}(x) = u_a(x),$$
$$\nabla h(\bar{u})(x) < -\beta \gamma(x_1) \quad \Rightarrow \quad \bar{u}(x) = u_b(x).$$

2. *If in addition $\nabla h(\bar{u}) \in L^\infty(\Lambda)$ holds, then $g = \beta j_C$ fulfills Assumption 2.1.2 with $D_{\bar{u}}$ replaced by $C_{\bar{u}}$,*
$$g''(\bar{u}, v^2) = \frac{\beta}{\|\bar{u}\|_{L^2(\Lambda_1, L^1(\Lambda_2))}} \left(\int_{\Lambda_1} \|\cdot\|_{L^1(\Lambda_2)}'(\bar{u}(x_1, \cdot), v(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1) - j_C'(\bar{u}, v)^2 \right),$$

   for $\bar{u} \neq 0$, and $g''(0, v^2) \equiv 0$ otherwise.

Proof. For the first part, see [**13**, Corollary 2.6]. For the second part, note that Assumption 2.1.2a is verified in Case II of the proof of [**13**, Theorem 3.3], while 2b is obtained as follows: For t_k, v_k as in Assumption 2.1.2b it follows from [**13**, Lemma 4.7] that

$$(2.10) \qquad \liminf_{k \to \infty} \frac{2}{t_k^2} [j_C(\bar{u} + t_k v_k) - j_C(\bar{u}) - t_k j_C'(\bar{u}, v_k)] \geq \liminf_{k \to \infty} j_C''(\bar{u}, v_k^2).$$

From [**13**, Lemma 4.6] we know that $v_k \rightharpoonup v$ in $L^2(\Lambda)$, $v \in C_{\bar{u}}$, and $j_C'(\bar{u}, v_k) \to j_C'(\bar{u}, v)$ implies that $\|\cdot\|_{L^1(\Lambda_2)}'(\bar{u}, v_k) \rightharpoonup \|\cdot\|_{L^1(\Lambda_2)}'(\bar{u}, v)$ weakly in $L^2(\Lambda_1)$. Therefore, we conclude $\liminf_{k \to \infty} j_C''(\bar{u}, v_k^2) \geq j''(\bar{u}, v^2)$, and together with (2.10) the claim follows. $\qquad \square$

The additional assumption $\nabla h(\bar{u}) \in L^\infty(\Lambda)$ is only required to verify Assumption 2.1.2a, i.e. for necessary second-order optimality conditions.

2.2.4. *Functional D.* Even though discrete versions of j_D are well known in the machine learning community, cf., e.g., [**7**, **39**], the analysis of j_D in the present infinite dimensional setting is —to the best of our knowledge— new. First, a short computation shows for any $u, v \in L^2(\Lambda_1, L^1(\Lambda_2))$ and $t > 0$:

$$(2.11) \quad \tilde{j}_D(u + tv) - \tilde{j}_D(u) = \frac{1}{2} \int_{\Lambda_1} \left(\|u(x_1, \cdot) + tv(x_1, \cdot)\|_{L^1(\Lambda_2)}^2 - \|u(x_1, \cdot)\|_{L^1(\Lambda_2)}^2 \right) d\rho_1(x_1)$$

$$= \int_{\Lambda_1} \|u(x_1, \cdot)\|_{L^1(\Lambda_2)} \cdot \left(\|u(x_1, \cdot) + tv(x_1, \cdot)\|_{L^1(\Lambda_2)} - \|u(x_1, \cdot)\|_{L^1(\Lambda_2)} \right) d\rho_1(x_1)$$

$$+ \frac{1}{2} \int_{\Lambda_1} \left(\|u(x_1, \cdot) + tv(x_1, \cdot)\|_{L^1(\Lambda_2)} - \|u(x_1, \cdot)\|_{L^1(\Lambda_2)} \right)^2 d\rho_1(x_1).$$

Dividing by t and sending $t \searrow 0$ yields:

$$(2.12) \qquad \tilde{j}_D'(u, v) = \int_{\Lambda_1} \|\cdot\|_{L^1(\Lambda_2)}'(u(x_1, \cdot), v(x_1, \cdot)) \|u(x_1, \cdot)\|_{L^1(\Lambda_2)} d\rho_1(x_1).$$

Consequently, we obtain
(2.13)
$$\partial \tilde{j}_D(u) = \left\{ \lambda \in L^2_{w\text{-}\star}(\Lambda_1, L^\infty(\Lambda_2)) \colon \lambda(x_1, x_2) \in \text{sign}(u(x_1, x_2)) \|u(x_1, \cdot)\|_{L^1(\Lambda_2)} \quad \rho\text{-a.e. on } \Lambda \right\},$$

and as for A-C these formulas remain true for j_D instead of $\tilde{j}_D$. We have:

Proposition 2.9. *Let $\bar{u} \in U_{\text{ad}}$ be a local solution to $(P_2\text{-}D)$ and define*
$$\bar{\gamma}(x_1) = \|\bar{u}(x_1, \cdot)\|_{L^1(\Lambda_2)}, \text{ if } \bar{u} \neq 0, \qquad \gamma(x_1) = 1, \text{ else.}$$

1. If $\alpha > 0$, it holds for ρ-a.a. $x \in \Lambda$:
$$\bar{u}(x_1, x_2) = 0 \quad \Leftrightarrow \quad |\nabla h(\bar{u})(x_1, x_2)| \leq \beta \bar{\gamma}(x_1),$$
$$\bar{\lambda}(x_1, x_2) = \text{Proj}_{[-\bar{\gamma}(x_1), \bar{\gamma}(x_1)]} \left(-\beta^{-1} \nabla h(\bar{u})(x_1, x_2) \right),$$
$$\bar{u}(x_1, x_2) = \text{Proj}_{[u_a, u_b]} \left[-\alpha^{-1} \left(\nabla h(\bar{u})(x_1, x_2) + \beta \bar{\lambda}(x_1, x_2) \right) \right].$$

If $\alpha = 0$, it holds for ρ-a.a. $x \in \Lambda$:
$$|\nabla h(\bar{u})(x)| < \beta \gamma(x_1) \quad \Rightarrow \quad \bar{u}(x) = 0,$$
$$\nabla h(\bar{u})(x) > \beta \gamma(x_1) \quad \Rightarrow \quad \bar{u}(x) = u_a(x),$$
$$\nabla h(\bar{u})(x) < -\beta \gamma(x_1) \quad \Rightarrow \quad \bar{u}(x) = u_b(x).$$

2. $g = \beta j_D$ fulfills Assumption 2.1.2 with $D_{\bar{u}}$ replaced by $C_{\bar{u}}$,
$$g''(\bar{u}, v^2) = \beta \int_{\Lambda_1} \|\cdot\|_{L^1(\Lambda_2)}'(\bar{u}(x_1, \cdot), v(x_1, \cdot))^2 d\rho_1(x_1),$$

for $\bar{u} \neq 0$, and $g''(0, v^2) \equiv 0$ otherwise.

Proof. Part one is verified along the lines of the proof of [**13**, Corollary 2.6] utilizing the above formula (2.13) for the subgradient. Regarding part two, we start with the verification of Assumption 2.1.2a with $D_{\bar{u}} = C_{\bar{u}}$. Let $v \in C_{\bar{u}}$ and $\bar{u} \neq 0$. As in the case II of the proof of [**13**, Theorem 3.3] we define $v_k \in L^2(\Lambda)$ by $v_k(x) = 0$ if $u(x) \in [u_a(x), u_a(x) + k^{-1}) \cup (-k^{-1}, 0) \cup (0, k^{-1}) \cup (u_b(x) - k^{-1}, u_b(x)]$, and $v_k(x) = \text{Proj}_{[-k,k]}(v(x))$, otherwise, and observe that $v_k \to v$ in $L^2(\Lambda)$, and $\bar{u} + tv_k \in U_{\text{ad}}$ if $0 < t < k^{-1}$. Moreover, it follows directly from the definition of v_k that

$$(2.14) \qquad \|\bar{u}(x_1, \cdot) + tv_k(x_1, \cdot)\|_{L^1(\Lambda_2)} = \|\bar{u}(x_1, \cdot)\|_{L^1(\Lambda_2)} + t\|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot))$$

holds ρ_1-a.e. on Λ_1 for $0 < t < k^{-2}$. With similar arguments as in [**13**] it can be shown that $f'(\bar{u})v_k + \beta j'_D(\bar{u}, v_k) = 0$, i.e. $v_k \in C_{\bar{u}}$. From (2.14) and (2.11) we conclude for those $0 < t < k^{-2}$

$$\frac{2}{t^2}[j_D(\bar{u} + tv_k) - j_D(\bar{u}) - tj'_D(\bar{u}, v_k)] = \int_{\Lambda_1} \|\cdot\|'_{L^1(\Lambda_1)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1).$$

Finally, we take $0 < t_k < k^{-2}$ and conclude due to $v_k \to v$ strongly in $L^2(\Lambda)$:

$$\lim_{k \to \infty} \frac{2}{t_k^2}[j_D(\bar{u} + t_k v_k) - j_D(\bar{u}) - t_k j'_D(\bar{u}, v_k)] = \lim_{k \to \infty} \int_{\Lambda_1} \|\cdot\|'_{L^1(\Lambda_1)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1)$$

$$= \int_{\Lambda_1} \|\cdot\|'_{L^1(\Lambda_1)}(\bar{u}(x_1, \cdot), v(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1) = j''_D(\bar{u}, v^2).$$

Hence, we have verified Assumption 2.1.2a. Next, let v_k, t_k be as in Assumption 2.1.2b. First, recall from the proof of [**13**, Lemma 4.7] that

$$\|\bar{u}(x_1, \cdot) + t_k v_k(x_1, \cdot)\|_{L^1(\Lambda_2)} \geq \left| \|\bar{u}(x_1, \cdot)\|_{L^1(\Lambda_2)} + t_k\|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot)) \right|,$$

which implies

$$\frac{1}{2} \int_{\Lambda_1} \|\bar{u}(x_1, \cdot) + t_k v_k(x_1, \cdot)\|_{L^1(\Lambda_2)}^2 \mathrm{d}\rho_1(x_1)$$

$$\geq \frac{1}{2} \int_{\Lambda_1} \left(\|\bar{u}(x_1, \cdot)\|_{L^1(\Lambda_2)} + t_k\|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot)) \right)^2 \mathrm{d}\rho_1(x_1)$$

$$= t_k \int_{\Lambda_1} \|\bar{u}(x_1, \cdot)\|_{L^1(\Lambda_2)}\|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot))\mathrm{d}\rho_1(x_1)$$

$$+ \frac{t_k^2}{2} \int_{\Lambda_1} \|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1),$$

and hence

$$(2.15) \qquad \frac{2}{t_k^2}[j_D(\bar{u} + t_k v_k) - j_D(\bar{u}) - t_k j'_D(\bar{u}, v_k)] \geq \int_{\Lambda_1} \|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1).$$

Along the lines of the proof of [**13**, Lemma 4.6] we obtain that $v_k \rightharpoonup v$, $v \in C_{\bar{u}}$, and $j'_D(\bar{u}, v_k) \to j'_D(\bar{u}, v)$ implies

$$\|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot)) \rightharpoonup \|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v_k(x_1, \cdot)), \qquad \text{weakly in } L^2(\Lambda_1).$$

Thus we conclude from (2.15):

$$\liminf_{k \to \infty} \frac{2}{t_k^2}[j_D(\bar{u} + t_k v_k) - j_D(\bar{u}) - t_k j'_D(\bar{u}, v_k)]$$

$$\geq \int_{\Lambda_1} \|\cdot\|'_{L^1(\Lambda_2)}(\bar{u}(x_1, \cdot), v(x_1, \cdot))^2 \mathrm{d}\rho_1(x_1) = j''_D(\bar{u}, v^2),$$

i.e. we have verified Assumption 2.1.2b.   □

3. Proofs of the main results

Finally we can prove Theorems 1.4 and 1.5 for (P_k) from Section 1.2. They are obtained by application of the abstract results of Sections 2.1 and 2.2 to

$$f(u) := \underbrace{\frac{1}{2}\|S(u) - y_d\|_{L^2(Q)}^2 + \frac{\alpha}{2}\|u\|_{L^2(I,\mathbb{R}^m)}^2}_{=:h(u)}, \quad g(u) := \beta j_k(u), \quad \hat{J}(u) = f(u) + g(u).$$

Here, $S\colon u \mapsto y = y(u)$ denotes the solution map of the state equation (Eq) to be specified below in Section 3.1. Before proving Theorems 1.4 and 1.5, already stated in Section 1.2, we summarize the required auxiliary results on f and g in the next subsection.

3.1. Auxiliary results regarding the smooth part f of the functional $\hat{J}$. Due to [**5**, Proposition 3.5] (see also [**47**, Corollary 5.8]) the solution map of the equation

$$(3.1) \qquad\qquad \partial_t y + \mathcal{A}(y)y = v, \qquad y(0) = y_0,$$

defined by $y := \tilde{S}(v)$ if and only if (3.1) holds, is a well defined map $\tilde{S}\colon L^s(I, W_D^{-1,p}) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$. Here, $y \in W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$ is said to be a solution of (3.1) if and only if

$$(3.2) \qquad \langle \partial_t y, \varphi \rangle_{W_D^{-1,p}, W_D^{1,p'}} + \int_\Omega \xi(y(t))\mu\nabla y(t)\nabla\varphi \, \mathrm{d}x = \langle v(t), \varphi \rangle_{W_D^{-1,p}, W_D^{1,p'}}$$

for all $\varphi \in W_D^{1,p'}$ and almost all $t \in I$, and $y(0) = y_0$ in $(W_D^{-1,p}, W_D^{1,p})_{1/s',s}$. For well definedness of $y(0) \in (W_D^{-1,p}, W_D^{1,p})_{1/s',s}$ we refer the reader, e.g., to [**3**, Theorem III.4.10.2]. By composition of $\tilde{S}$ with B we obtain the control-to-state map $S\colon L^s(\Lambda) \to W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$, $u \mapsto \tilde{S}(Bu)$. Given $y \in W^{1,s}(I, W_D^{-1,p}) \cap L^s(I, W_D^{1,p})$ we recall from [**5**] the notation for the derivatives of the nonlinear term in (Eq), stated in weak form:

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

$$\langle \mathcal{A}''(y)[v_1, v_2], \varphi \rangle := \int_Q (\xi'(y)(v_1\mu\nabla v_2 + v_2\mu\nabla v_1) + \xi''(y)v_1v_2\mu\nabla y)\nabla\varphi \, \mathrm{d}x\mathrm{d}t,$$

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 $\varphi \in L^{s'}(I, W_D^{1,p'})$. The following result holds true:

Proposition 3.1. *The functional*

$$f\colon L^s(I, \mathbb{R}^m) \to \mathbb{R}, \quad u \mapsto \frac{1}{2}\|S(u) - y_d\|_{L^2(Q)}^2 + \frac{\alpha}{2}\|u\|_{L^2(I,\mathbb{R}^m)}^2$$

is well defined, twice continuously Fréchet differentiable, and the following properties hold true:

1. Given $u, v \in L^s(I, \mathbb{R}^m)$, it holds

$$(3.3) \qquad f'(u)v = \int_I (\alpha u + B^*p)^T v\mathrm{d}t$$

$$(3.4) \qquad f''(u)v^2 = \alpha\|v\|_{L^2(I,\mathbb{R}^m)}^2 + \int_Q ([1 - \xi''(y)\nabla p \cdot \mu\nabla y]z^2 - 2\xi'(y)z\nabla p \cdot \mu\nabla z) \, \mathrm{d}x\mathrm{d}t$$

with $y = S(u)$, $z = S'(u)v$ and $p := S'(u)^*(y - y_d) \in L^{r'}(I, W_D^{1,p'})$, $r' \in [s', \infty)$. Herein, z is given by the solution to the equation

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

and exhibits regularity $z \in W^{1,r}(I, W_D^{-1,p}) \cap L^r(I, W_D^{1,p})$, $r \in (1, s]$, whenever $v \in L^r(I, W^{-1,p})$. Moreover, $\mathcal{A}(y)$ and $\mathcal{A}(y) + \mathcal{A}'(y)$ exhibit maximal parabolic regularity on $L^r(I, W_D^{-1,p})$.

2. *f satisfies the properties of Assumption 2.1.1 with $U_\infty = L^s(I, \mathbb{R}^m)$, $U_2 = L^2(I, \mathbb{R}^m)$ and $K = U_{\mathrm{ad}}$.*

Proof. These results can be found in [**5**, Sections 4.1-4.3]; see in particular Lemma 4.5, Lemma 4.6, and Proposition 4.10. As already pointed out underneath Assumption 1.3, our assumptions suffice to apply these results.   □

Next, we state an improved regularity result for the adjoint state p which is different from the regularity result [**5**, Proposition 4.7] obtained on Bessel-potential spaces because it does not need the additional assumptions from [**5**]. Note that this improved regularity result is only used to ensure $\nabla h(\bar{u}) = B^* \bar{p} \in L^\infty(I, \mathbb{R}^m)$, which is required for necessary optimality conditions for j_4 and j_5 (i.e. case C); cf. Proposition 2.8.

Proposition 3.2. *Assume $y_d \in L^\infty(I, L^q)$ with $q \in [2, \infty)$. The adjoint state p introduced in Proposition 3.1 exhibits the regularity*

$$p \in W^{1,r}(I, L^q) \cap L^r(I, \mathrm{Dom}_{L^q}(-\nabla \cdot \mu \nabla)), \qquad r \in (1, \infty),$$

and satisfies the equation

$$-\partial_t p - \xi(y) \nabla \cdot \mu \nabla p = y - y_d, \qquad p(T) = 0,$$

in the sense of distributions. Moreover, the operator $-\xi(y) \nabla \cdot \mu \nabla$ has nonautonomous maximal parabolic regularity on $L^r(I, L^q)$, and there is an embedding

(3.5) $$W^{1,r}(I, L^p) \cap L^r(I, \mathrm{Dom}_{L^p}(-\nabla \cdot \mu \nabla)) \hookrightarrow C^{0,\sigma}(I, W_D^{1,p})$$

with some $\sigma > 0$ provided that $r \in (2, \infty)$.

Proof. First, we prove that $\mathcal{A}(y)^* + \mathcal{A}'(y)^* = -\xi(y) \nabla \cdot \mu \nabla$ on $\mathrm{Dom}_{L^q}(-\nabla \cdot \mu \nabla)$ for $q \in [2, \infty)$ and any $y \in W_D^{1,p}$. Let $z \in \mathrm{Dom}_{L^q}(-\nabla \cdot \mu \nabla)$ and $\psi \in C_c^\infty(\Omega)$. Since $\psi \xi(y) \in W_D^{1,p}$ has compact support and $\mu \nabla z$ has weak divergence in L^q we obtain:

$$\int_\Omega (-\xi(y) \nabla \cdot \mu \nabla z) \psi \mathrm{d}x = \int_\Omega (-\nabla \cdot \mu \nabla z) \psi \xi(y) \mathrm{d}x = \int_\Omega \mu \nabla z \nabla(\psi \xi(y)) \mathrm{d}x$$

$$= \langle z, (\mathcal{A}(y) + \mathcal{A}'(y)) \psi \rangle_{W_D^{1,p'}, W_D^{-1,p}} = \langle (\mathcal{A}(y) + \mathcal{A}'(y))^* z, \psi \rangle_{W_D^{-1,p'}, W_D^{1,p}}.$$

The left-hand side thereof is well defined for every $\psi \in L^{q'}$, and hence the claimed identity follows from density of $C_c^\infty(\Omega)$ in $L^{q'}$. Second, the nonautonomous parabolic operator $-\xi(y) \nabla \cdot \mu \nabla$ exhibits maximal parabolic regularity on L^q, which can be seen as follows: Due to our assumptions on ξ, μ, Ω, and Γ_D we can apply [**32**, Proposition 5.4] to obtain maximal parabolic regularity of each autonomous operator $-\xi(y(t)) \nabla \cdot \mu \nabla$, $t \in I$. Due to $y \in C(\overline{Q})$, the nonautonomous operator $t \mapsto -\xi(y(t)) \nabla \cdot \mu \nabla$ is continuous as a map $I \to \mathcal{L}(\mathrm{Dom}_{L^q}(-\nabla \cdot \mu \nabla), L^q)$, from which we conclude nonautonomous maximal parabolic regularity by application of [**2**, Theorem 7.1]. Now, we can proceed similar to the proofs of [**5**, Lemma 4.6 and Proposition 4.7] to improve regularity of $p = S'(u)^*(y - y_d)$ and to obtain the adjoint equation. It remains to show (3.5). According to [**30**, Lemma 6.6] it holds $(L^p, \mathrm{Dom}_{L^p}(-\nabla \cdot \mu \nabla))_{\tau,1} \hookrightarrow W_D^{1,p}$ for each $\tau \in (\frac{1}{2}, 1)$. Hence we can apply standard Bochner-Sobolev embeddings; see, e.g., [**5**, Proposition 3.2], [**2**].   □

3.2. Auxiliary results for the nonsmooth part g of the functional $\hat{J}$. Regarding the nonsmooth part g of the functional, it suffices to observe that the seven possibilities for j_k given in the introduction can be reduced to the four generic cases A-D from Section 2.2 as follows:

$k = 1$: Case A with $(\Lambda_1, \rho_1) = (I, \mathrm{d}t)$, $(\Lambda_2, \rho_2) = (\{\bullet\}^m, \text{counting measure})$,

$k = 2$: Case B with $(\Lambda_1, \rho_1) = (\{\bullet\}^m, \text{counting measure})$, $(\Lambda_2, \rho_2) = (I, \mathrm{d}t)$,

$k = 3$: Case B with $(\Lambda_1, \rho_1) = (I, \mathrm{d}t)$, $(\Lambda_2, \rho_2) = (\{\bullet\}^m, \text{counting measure})$,

$k = 4$: Case C with (Λ_1, ρ_1), (Λ_2, ρ_2) as for $k = 2$,

$k = 5$: Case C with (Λ_1, ρ_1), (Λ_2, ρ_2) as for $k = 3$,

$k = 6$: Case D with (Λ_1, ρ_1), (Λ_2, ρ_2) as for $k = 2$,

$k = 7$: Case D with (Λ_1, ρ_1), (Λ_2, ρ_2) as for $k = 3$.

Therefore, we can translate the results for j_X, $X \in \{A, B, C, D\}$ from Section 2.2 back to j_k for $k = 1, ..., 7$. For reference, we state concrete formulas for the directional derivatives, the subgradients, and the surrogates for the second-order derivative.

Subgradients. We use the set-valued sign function as introduced in Section 2.2:

$$(3.6) \qquad \partial j_1(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda_i(t) \in \mathrm{sign}(u_i(t)) \text{ for a.a. } t \in I, \ i = 1, ..., m\right\},$$

$$(3.7) \qquad \partial j_2(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda_i \begin{cases} \in \mathbb{B}_1^{L^2(I)}(0), & \text{if } u_i \equiv 0, \\ = \frac{u_i}{|(\|u_i\|_{L^2(I)})_i|_1}, & \text{else.} \end{cases}\right\},$$

$$(3.8) \qquad \partial j_3(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda(t) \begin{cases} \in \mathbb{B}_1^{\mathbb{R}^m}(0), & \text{if } u(t) = 0, \\ = \frac{u(t)}{\||u|_2\|_{L^1(I)}}, & \text{else.} \end{cases}\right\},$$

$$(3.9) \qquad \partial j_4(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda_i(t) \in \mathrm{sign}(u_i(t)) \frac{\|u_i\|_{L^1(I)}}{|(\|u_i\|_{L^1(I)})_i|_2}\right\},$$

$$(3.10) \qquad \partial j_5(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda_i(t) \in \mathrm{sign}(u_i(t)) \frac{|u(t)|_1}{\||u|_1\|_{L^2(I)}}\right\},$$

$$(3.11) \qquad \partial j_6(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda_i(t) \in \mathrm{sign}(u_i(t))\|u_i\|_{L^1(I)}\right\},$$

$$(3.12) \qquad \partial j_7(u) = \left\{\lambda \in L^2(I, \mathbb{R}^m): \lambda_i(t) \in \mathrm{sign}(u_i(t))|u(t)|_1\right\}.$$

Directional Derivatives.

$$(3.13) \qquad j_1'(u, v) = \sum_{i=1}^m \|\cdot\|_{L^1(I)}'(u_i, v_i),$$

$$(3.14) \qquad j_2'(u, v) = \sum_{\{i: \ \|u_i\|_{L^2(I)}=0\}} \|v_i\|_{L^2(I)} + \sum_{\{i: \ \|u_i\|_{L^2(I)}\neq 0\}} \int_I \frac{u_i(t)}{\|u_i\|_{L^2(I)}} v_i(t)\mathrm{d}t,$$

$$(3.15) \qquad j_3'(u, v) = \int_{\{t: \ u(t)=0\}} |v(t)|_2\mathrm{d}t + \int_{\{t: \ u(t)\neq 0\}} \frac{u(t)^T}{|u(t)|_2} v(t)\mathrm{d}t,$$

$$(3.16) \qquad j_4'(u, v) = \sum_{i=1}^m \|\cdot\|_{L^1(I)}'(u_i, v_i) \frac{\|u_i\|_{L^1(I)}}{|(\|u_i\|_{L^1(I)})_i|_2},$$

$$(3.17) \qquad j_5'(u, v) = \int_I |\cdot|_1'(u(t), v(t)) \frac{|u(t)|_1}{\||u|_1\|_{L^2(I)}}\mathrm{d}t,$$

$$(3.18) \qquad j_6'(u, v) = \sum_{i=1}^m \|\cdot\|_{L^1(I)}'(u_i, v_i)\|u_i\|_{L^1(I)},$$

$$(3.19) \qquad j_7'(u, v) = \int_I |\cdot|_1'(u(t), v(t))|u(t)|_1\mathrm{d}t.$$

Surrogates for the Second Derivatives.

$$(3.20) \qquad j_1''(u, v^2) = 0,$$

$$(3.21) \qquad j_2''(u, v^2) = \sum_{\{i: \ u_i\neq 0\}} \frac{1}{\|u_i\|_{L^2(I)}} \left[\|v_i\|_{L^2(I)}^2 - \left(\int_I \frac{u_i(t)v_i(t)}{\|u_i\|_{L^2(I)}}\mathrm{d}t\right)^2\right],$$

$$(3.22) \qquad j_3''(u, v^2) = \int_{\{t:\, u(t) \neq 0\}} \frac{1}{|u(t)|_2} \left[|v(t)|_2^2 - \left(\frac{u(t)^T v(t)}{|u(t)|_2} \right)^2 \right],$$

$$(3.23) \qquad j_4''(u, v^2) = \frac{1}{j_4(u)} \left(\sum_{i=1}^m \|\cdot\|_{L^1(I)}' (u_i, v_i) - j_4'(u, v)^2 \right),$$

$$(3.24) \qquad j_5''(u, v^2) = \frac{1}{j_5(u)} \left(\int_I |\cdot|_1'(u(t), v(t)) \mathrm{d}t - j_5'(u, v) \right),$$

$$(3.25) \qquad j_6''(u, v^2) = \sum_{i=1}^m \|\cdot\|_{L^1(I)}' (u_i, v_i)^2,$$

$$(3.26) \qquad j_7''(u, v^2) = \int_I |\cdot|_1'(u(t), v(t))^2 \mathrm{d}t,$$

for $u \neq 0$ and $j_k(u, v^2) = 0$ for $k = 1, ..., 7$.

3.3. Main results: Proofs of Theorems 1.4 and 1.5.
First, we prove Theorem 1.4 that states first-order optimality conditions and the resulting sparsity patterns of the optimal control. As already pointed out, we obtain this result by straightforward application of Propositions 2.6, 2.7, and 2.8. Note that —unlike for existence of optimal controls— the box-constraints on the controls are not required for this result.

Proof of Theorem 1.4. Each (P_k), $k \in \{1, ..., 7\}$, can be understood as realization of $(P_2\text{-}X)$ for some $X \in \{A, B, C, D\}$; cf. Section 3.2. First-order optimality conditions and sparsity patters for the latter have been obtained in Section 2.2. The formula for the gradient of the smooth part of the functional has been stated in Propositions 3.1 and 3.2. $\quad\square$

Next, we prove Theorem 1.5 on second-order optimality conditions. Again, the proof is short, because the main work has already been done in Theorems 2.4 and 2.5, and the verification of the corresponding assumptions in Sections 2.2 and 3.1.

Proof of Theorem 1.5. Apply Theorems 2.4 and 2.5 to $U_\infty = L^s(I, \mathbb{R}^m)$, $U_2 = L^2(I, \mathbb{R}^m)$, f, g, and h specified as in Section 3.1 above, and $K = U_{\mathrm{ad}}$. Assumption 2.1.1 and 2 with $D_{\bar{u}}$ replaced by $C_{\bar{u}}$ have been verified in Proposition 3.1 and Section 2.2. The additional requirement $\nabla h(\bar{u}) = B^* \bar{p} \in L^\infty(I, \mathbb{R}^m)$ in the case B follows from Proposition 3.2. $\quad\square$

Here, the presence of box-constraints for the controls is necessary for the verification of Assumption 2.1.1, because all L^s-topologies, $s \in [1, \infty)$, are equivalent on the L^∞-bounded admissible set. Omitting the control-constraints is possible, but infers a norm-gap, as explained at the end of Section 2.1.

3.4. Limitations of the approach: The bang-bang case.
We conclude this section by an outlook to the bang-bang case $\alpha = 0$ that illustrates the limits of our second-order analysis. In fact, the present approach cannot be extended to the bang-bang case. Regarding necessary optimality conditions, a short computation shows that $C_{\bar{u}} = \{0\}$ holds for $\bar{u}$ satisfying the first-order optimality conditions of (OCP_1). Hence, the first part of the statement of Theorem 2.4 is still true, but trivial for $\alpha = 0$. For sufficient optimality conditions, Assumption 2.1.1biii is crucial. It is well known that this property for the smooth part of the functional can only be expected in the case of $\alpha > 0$, or a similar so-called Legendre-Clebsh condition; cf. [**19**]. As explained at the end of Section 2.1, it seems impossible to avoid this assumption on f by exploiting properties of g. Recall that in the case $\alpha = 0$ the second derivative of the smooth part of the functional reads as follows:

$$f''(\bar{u})v^2 = \int_Q \left([1 - \xi''(\bar{y}) \nabla \bar{p} \cdot \mu \nabla y] z_v^2 - 2\xi'(\bar{y}) z_v \nabla \bar{p} \cdot \mu \nabla z_v \right) \mathrm{d}x \mathrm{d}t, \qquad \forall v \in L^2(I, \mathbb{R}^m),$$

with $\bar{y} = S(\bar{u})$, $z_v := S'(\bar{u})v$ and $\bar{p}$ as in Proposition 3.1. Assuming appropriate higher regularity for $\bar{p}$ this can be transformed into

$$f''(\bar{u})v^2 = \int_Q (1 - \xi'(\bar{y})\nabla \cdot \mu\nabla\bar{p})z_v^2 \mathrm{d}x\mathrm{d}t.$$

The approach in [**13**, **16**] for the bang-bang case is based on a second-order sufficient optimality condition of the type

$$f''(\bar{u})v^2 + \beta j_k(\bar{u}, v^2) \geq c\|z_v\|_{L^2(Q)}^2$$

with some $c > 0$ and all v from a certain cone. For such a condition to hold in our setting, we expect that $1 - \xi'(\bar{y})\nabla \cdot \mu\nabla\bar{p} \in L^\infty(Q)$, and therefore $\nabla \cdot \mu\nabla\bar{p} \in L^\infty(Q)$ has to hold, which is a very strong assumption. Moreover, to follow the arguments of [**13**, **16**] we would need certain continuity properties like

$$|f''(u)v^2 - f''(\bar{u})v^2| \lesssim \|u - \bar{u}\|_{L^s(I,\mathbb{R}^m)}\|z_v\|_{L^2(Q)}^2.$$

Consequently, $\nabla \cdot \mu\nabla p \in L^\infty(Q)$ would be required to depend continuously on u, which is out of reach, even in a highly smooth setup and with controls measured in $L^\infty(I, \mathbb{R}^m)$.

On the other hand, to apply the approach for L^1-penalized semilinear elliptic bang-bang problems from [**59**], we would have to guarantee that there is a bounded bilinear extension $f''(\bar{u})\colon \mathcal{M}(I, \mathbb{R}^m) \times \mathcal{M}(I, \mathbb{R}^m) \to \mathbb{R}$, to the space $\mathcal{M}(I, \mathbb{R}^m)$ of $\mathbb{R}^m$-valued regular Borel measures on I. Moreover, appropriate higher regularity for the adjoint state would be needed. This is of course more delicate for parabolic problems than for elliptic ones. Moreover, it is not clear whether the results obtained in [**59**] for L^1-penalization also hold for directional sparsity functionals.

This shows that it is by no means obvious that techniques successfully applied to semilinear parabolic or semilinear elliptic problems can be transferred to the quasilinear parabolic case. We leave this as an interesting open problem.

4. Numerical illustration

Let us conclude the paper by some numerical computations that illustrate the different sparsity patterns induced by the penalizers j_k, $k \in \{1, ..., 7\}$. Before presenting these numerical examples in Section 4.2, we give a concise overview over the fast proximal method in Section 4.1, that is used to solve (P_k) for $k = 1\text{-}3, 6, 7$. For $k = 4, 5$ we apply a subgradient method to (P_k).

4.1. Proximity operators and fast proximal methods. Proximal algorithms, see, e.g., [**49**], have been applied successfully in different areas, e.g., image processing, and machine learning, but also in PDE-constrained optimization [**52**, **53**]. This class of algorithms specifically applies to problems of type $(\mathrm{P}_2\text{-}X)$, cf. Section 2.2, consisting of a nonconvex, but smooth, and a convex, but nonsmooth summand in the functional. In the context of sparse optimal control we are aware of possibly faster methods, e.g., certain Newton-type methods in function space [**31**, **44**, **50**, **56**], or algorithms on the discrete level, e.g., [**11**, Section 6] and [**31**, Section 4]. However, proximal methods usually have the advantage that they are relatively easy to implement and, compared to second-order methods, less intrusive. Let us recall from, e.g., [**52**, Algorithm 2] the basic concept of the so-called fast proximal method, formulated on behalf of $(\mathrm{P}_2\text{-}X)$. Given a fixed step size $L > 0$, an initial guess u^0, and $t^0 = 1$, set $v^0 = u^0$ and for $\ell = 1, 2, 3, ...$

$$u^\ell = \mathrm{Prox}_{\frac{\beta}{L}}\left(v^{\ell-1} - \frac{1}{L}\nabla f(v^{\ell-1})\right),$$

$$t^\ell = \frac{1}{2}\left(1 + \sqrt{1 + 4(t^{\ell-1})^2}\right), \qquad v^\ell = u^\ell + \frac{t^{\ell-1} - 1}{t^\ell}(u^\ell - u^{\ell-1}),$$

until the current iterate u^ℓ reaches a desired optimality criterion. Here, we denote by

$$(4.1) \qquad \mathrm{Prox}_\tau(v) := \mathrm{argmin}_{u \in U_{\mathrm{ad}}} \left(\frac{1}{2} \|u - v\|_{L^2(\Lambda)^2} + \tau j_X(u) \right)$$

the so-called proximity operator; see, e.g., [49] or [4, Chapter 24] for an overview. It is a crucial condition for the applicability of proximal algorithms to $(\mathrm{P}_2\text{-}X)$, that the nonsmooth part of the functional, j_X, is "proximable", i.e. we have to know how to compute (4.1) efficiently. We briefly address this issue using the notation of Section 2.2. In case A, it is well known, see, e.g., [52, Lemma 4.3] or [50, Section 3.3.2], that j_A is proximable with

$$[\mathrm{Prox}_\tau(v)](x) = \begin{cases} \min(v(x) - \tau, u_b(x)) & \text{if } v(x) > \tau, \\ 0 & \text{if } |v(x)| \le \tau, \qquad \text{for a.a. } x \in \Lambda. \\ \max(v(x) + \tau, u_a(x)) & \text{if } v(x) < -\tau, \end{cases}$$

For case B and $U_{\mathrm{ad}} = L^2(\Lambda)$ it holds

$$[\mathrm{Prox}_\tau(v)](x_1, x_2) = \max\left(0, 1 - \frac{\tau}{\|v(x_1, \cdot)\|_{L^2(\Lambda_2)}} \right) v(x_1, x_2), \qquad \text{for a.a. } x \in \Lambda;$$

cf., e.g., [50, Section 3.3.2]. We refer the reader to [49, Section 6.5.4] or [39, Theorem 3] for the same formula in the discrete case. For the proximity operator in the case B with $U_{\mathrm{ad}} := \{u \in L^2(\Lambda) : u \ge 0 \text{ a.e.}\}$, we refer the reader to [50, Section 3.3.2]. We are not aware of an explicit formula for the proximity operator in the case of bilateral box-constraints. To the best of our knowledge, the functional of case C is not "proximable". For case D and $U_{\mathrm{ad}} = L^2(\Lambda)$, however, a formula for the proximity operator is well known for the discrete analogon, see, e.g., [39, Theorem 3], and the adaption to our setting is not difficult. We obtain:

$$[\mathrm{Prox}_\tau(v)](x_1, x_2) = \mathrm{sign}(v(x_1, x_2)) \max(0, |v(x_1, x_2)| - \rho(x_1)) \quad \text{for a.a. } x \in \Lambda,$$

where $\rho(x_1) \in \mathbb{R}$ has to satisfy for every $x_1 \in \Lambda_1$:

$$(4.2) \qquad \|\max(0, |v(x_1, \cdot)| - \rho(x_1))\|_{L^1(\Lambda_2)} = \frac{\rho(x_1)}{2\tau}.$$

The efficient computation of proximity operators for case C, as well as for cases B and D in the presence of box-constraints, is certainly of interest, but beyond the scope of the present paper. Also, we do not address the convergence analysis of the fast proximal method in our precise setting.

4.2. Results. We consider the following specification of (P_k): We choose $\Omega = B_1(0) \subset \mathbb{R}^d$, $T = 8$, $\mu \equiv I_2 \in \mathbb{R}^{2 \times 2}$, $\Gamma_D = \partial\Omega$, $\alpha = \frac{\pi}{25} \cdot 10^{-2}$, $\beta = 10^{-2}$, $y_0 \equiv 0$, $\xi(s) := \frac{1}{2} + \frac{1}{1 + \exp(-20s)}$, and $y_d(t, x) := -\frac{t}{4} \sin(\frac{\pi}{2} t) \exp(-36|x - m(t)|_2^2)$ with $m(t) = \frac{2}{3}(\sin(\frac{\pi}{4} t), \cos(\frac{\pi}{4} t))^T$. The eight control actuators are given by $\langle b_i, \varphi \rangle := \int_\Omega \mathbf{1}_{B_{\frac{1}{5}}(m(i-1))}(x) \varphi(x) \mathrm{d}x$, $i = 1, \ldots, 8$. Consequently, the state equation reads as follows:

$$\partial_t y(t, x) - \nabla \cdot \left(\frac{1}{2} + \frac{1}{1 + \exp(-20y(t, x))} \right) \nabla y(t, x) = \sum_{i=1}^8 u_i(t) \cdot \mathbf{1}_{B_{\frac{1}{5}}(m(i-1))}(x),$$

$$\text{on } [0, 8] \times B_1(0),$$

$$y(t, x) = 0, \qquad \text{on } [0, 8] \times \partial B_1(0),$$

$$y(0, x) \equiv 0, \qquad \text{on } B_1(0).$$

We omit control-constraints and set $U_{\mathrm{ad}} = L^s(I, \mathbb{R}^m)$.

Space is discretized with the help of FEniCS and mshr [1, 42] using piecewise linear finite elements with 3324 DoF and mesh size $h_{\max} \approx 5.0 \cdot 10^{-2}$. For time discretization we use the implicit Euler scheme with 160 timesteps. The nonlinear problem at each timestep of the

solution of the discretized state equation is solved by the built-in nonlinear solver provided by FEniCS. For $k = 1$-$3, 6, 7$, i.e. cases A, B, and D, we solve the discretized counterpart of (P_k) by the fast proximal method described above with step size $L = 10^{-2}$. In case D, equation (4.2) is solved by bisection. Since for case C we are not aware of a proximity operator, we solve the optimal control problem for $k = 4, 5$ by a classical subgradient descent method, cf. [**43**, Chapter II.2.1.2] for instance, with the step size in iteration ℓ given by $\frac{10}{\sqrt{\ell}}$. The initial guess in all cases is $u_i^0 \equiv 0.1$.

Figure 4.2 b)-h) shows the optimal controls of (P_k) for $k = 1, ..., 7$. For comparsion we also display the non-sparse optimal control for $\beta = 0$ in Figure 4.2 a). The typical sparsity patterns described in Section 1.2 are clearly visible; cf. the description below Theorem 1.4. In c) the number of control actuators becoming active is sparse. The controls u_2, u_3 and u_5 are identical zero, but the remaining actuators are active over the whole time interval. In f) all actuators are in use at some point in time, but it seems to be the case that only two actuators can become active at the same time. In d), finally, any control action is limited to certain subintervals of I, but in them all actuators become active simultanuously. Also the difference between $k = 1$ and $k = 4$ becomes visible. In b) an actuator only becomes active if its contribution is above a certain threshold level that is the same for all actuators. This results in a number of actuators never becoming active. In e), however, this threshold level is different for each actuator, cf. the formula in Theorem 1.4, and therefore each actuator is used, but only at those times where it is needed sufficiently much. Finally, by comparison of e) with g) and f) with h) it can be seen that usage of functionals j_4 and j_6 as well as j_5 and j_7 lead to similar results w.r.t. the sparsity pattern, respectively.

In Figure 4.2 a) and b) we illustrate the convergence speed of the fast proximal and the subgradient method. We display the $L^2(I, \mathbb{R}^m)$-norm of the residuals r^ℓ of the control iterates u^ℓ, i.e.

$$r^\ell := u^\ell + \alpha^{-1}(p^\ell + \beta\lambda^\ell),$$

where p^ℓ and λ^ℓ denote the adjoint state and the subgradient of j_k associated with u^ℓ; cf. the optimality conditions in Theorem 1.4 and Section 2.2. As expected, the fast proximal method converges much faster than the subgradient method. In our opinion, this indicates that replacing functionals j_4, j_5 by j_6, j_7, respectively, may be worth considering in applications in order to allow the application of fast proximal methods.

References

[1] Martin S. Alnæs, Jan Blechta, Johan Hake, August Johansson, Benjamin Kehlet, Anders Logg, Chris Richardson, Johannes Ring, Marie E. Rognes, and Garth N. Wells, *The fenics project version 1.5*, Archive of Numerical Software **3** (2015), no. 100.

[2] H. Amann, *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] Heinz H. Bauschke and Patrick L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Second, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, Cham, 2017. With a foreword by Hédy Attouch. MR3616647

[5] 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

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

[7] Frederick Campbell and Genevera I. Allen, *Within group variable selection through the Exclusive Lasso*, Electronic Journal of Statistics **11** (2017), no. 2, 4220 –4257.

[8] Eduardo Casas, *A review on sparse solutions in optimal control of partial differential equations*, SeMA J. **74** (2017), no. 3, 319–344. MR3745942
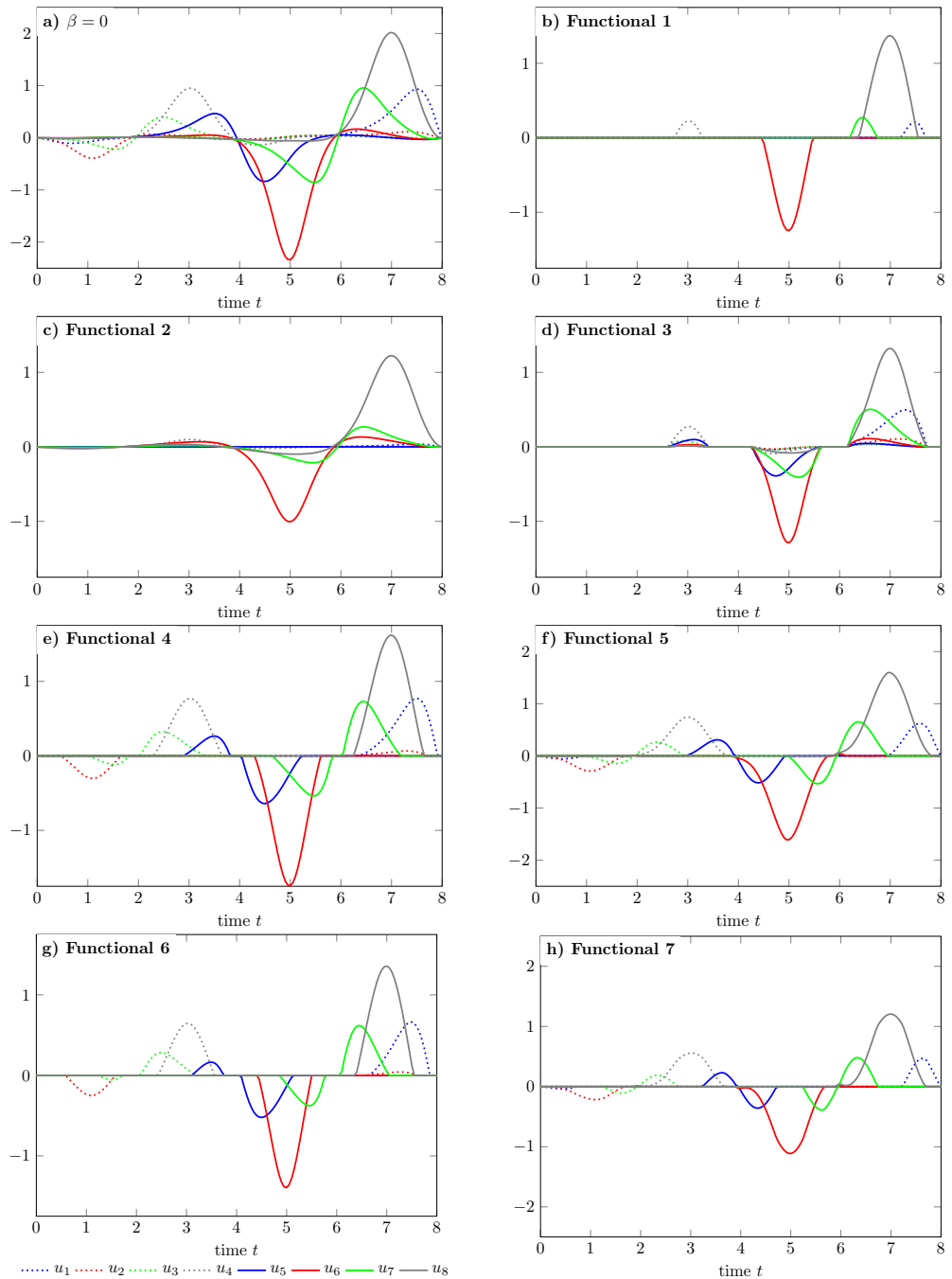
FIGURE 1. Optimal controls for $\beta = 0$ (a) and functionals j_1-j_7 (b-h).

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

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

[11] Eduardo Casas, Christian Clason, and Karl Kunisch, *Parabolic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim. **51** (2013), no. 1, 28–63. MR3032866
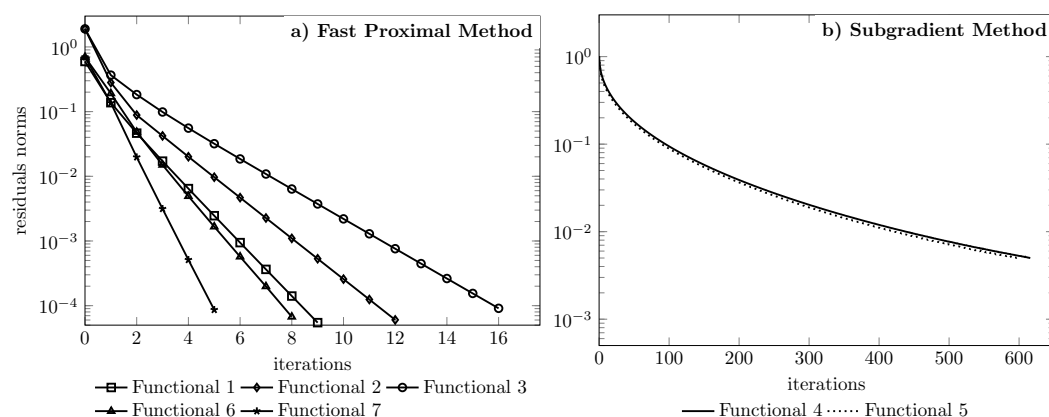
FIGURE 2. Convergence behaviour of the algorithms.

[12] Eduardo Casas, Roland Herzog, and Gerd Wachsmuth, *Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional*, SIAM J. Optim. **22** (2012), no. 3, 795–820. MR3023751

[13] ———, *Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations*, ESAIM Control Optim. Calc. Var. **23** (2017), no. 1, 263–295. MR3601024

[14] Eduardo Casas and Karl Kunisch, *Optimal control of semilinear elliptic equations in measure spaces*, SIAM J. Control Optim. **52** (2014), no. 1, 339–364. MR3162396

[15] ———, *Parabolic control problems in space-time measure spaces*, ESAIM Control Optim. Calc. Var. **22** (2016), no. 2, 355–370. MR3491774

[16] Eduardo Casas and Mariano Mateos, *Critical cones for sufficient second order conditions in PDE constrained optimization*, SIAM J. Optim. **30** (2020), no. 1, 585–603. MR4066574

[17] Eduardo Casas, Mariano Mateos, and Arnd Rösch, *Finite element approximation of sparse parabolic control problems*, Math. Control Relat. Fields **7** (2017), no. 3, 393–417. MR3669663

[18] Eduardo Casas, Christopher Ryll, and Fredi Tröltzsch, *Second order and stability analysis for optimal sparse control of the FitzHugh-Nagumo equation*, SIAM J. Control Optim. **53** (2015), no. 4, 2168–2202. MR3374636

[19] 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

[20] Eduardo Casas and Fredi Tröltzsch, *Second-order and stability analysis for state-constrained elliptic optimal control problems with sparse controls*, SIAM J. Control Optim. **52** (2014), no. 2, 1010–1033. MR3180842

[21] ———, *State-constrained semilinear elliptic optimization problems with unrestricted sparse controls*, Math. Control Relat. Fields **10** (2020), no. 3, 527–546. MR4128838

[22] Eduardo Casas and Daniel Wachsmuth, *First and second order conditions for optimal control problems with an L^0 term in the cost functional*, SIAM J. Control Optim. **58** (2020), no. 6, 3486–3507. MR4176889

[23] 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

[24] 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

[25] 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

[26] Chi Ngoc Do, *Generalized second-order derivatives of convex functions in reflexive Banach spaces*, Trans. Amer. Math. Soc. **334** (1992), no. 1, 281–301. MR1088019

[27] R. E. Edwards, *Functional analysis. Theory and applications*, Holt, Rinehart and Winston, New York-Toronto-London, 1965. MR0221256

[28] Ivar Ekeland and Roger Témam, *Convex analysis and variational problems*, English, Classics in Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999. Translated from the French. MR1727362

[29] 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

[30] 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

[31] Roland Herzog, Georg Stadler, and Gerd Wachsmuth, *Directional sparsity in optimal control of partial differential equations*, SIAM J. Control Optim. **50** (2012), no. 2, 943–963. MR2914236

[32] Matthias Hieber and Joachim Rehberg, *Quasilinear parabolic systems with mixed boundary conditions on nonsmooth domains*, SIAM J. Math. Anal. **40** (2008), no. 1, 292–305. MR2403322

[33] 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

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

[35] _____, *Optimal control of quasilinear parabolic PDEs with state constraints*, 2020. Accepted for publication in SICON. Available as INS Preprint No. 2004.

[36] _____, *A-posteriori reduced basis error-estimates for a semi-discrete in space quasilinear parabolic PDE*, Comput. Optim. Appl. (2021).

[37] 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

[38] Kazufumi Ito and Karl Kunisch, *Optimal control with $L^p(\Omega)$, $p \in [0, 1)$, control cost*, SIAM J. Control Optim. **52** (2014), no. 2, 1251–1275. MR3196933

[39] Matthieu Kowalski, *Sparse regression using mixed norms*, Appl. Comput. Harmon. Anal. **27** (2009), no. 3, 303–324. MR2559729

[40] Karl Kunisch, Konstantin Pieper, and Boris Vexler, *Measure valued directional sparsity for parabolic optimal control problems*, SIAM J. Control Optim. **52** (2014), no. 5, 3078–3108. MR3265176

[41] J.-L. Lions, *Optimal control of systems governed by partial differential equations*, Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170, Springer-Verlag, New York-Berlin, 1971. MR0271512

[42] Anders Logg, Kent-Andre Mardal, Garth N. Wells, et al., *Automated solution of differential equations by the finite element method*, Springer, 2012.

[43] Marko M. Mäkelä and Pekka Neittaanmäki, *Nonsmooth optimization*, World Scientific Publishing Co., Inc., River Edge, NJ, 1992. Analysis and algorithms with applications to optimal control. MR1177832

[44] Florian Mannel and Armin Rund, *A hybrid semismooth quasi-Newton method for nonsmooth optimal control with PDEs*, Optim. Eng. (2020).

[45] 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

[46] _____, *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

[47] 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

[48] Mohammed Moussaoui and Alberto Seeger, *Second-order subgradients of convex integral functionals*, Trans. Amer. Math. Soc. **351** (1999), no. 9, 3687–3711. MR1487628

[49] Neal Parikh and Stephen Boyd, *Proximal algorithms*, Foundations and Trends® in Optimization **1** (2014), no. 3, 127–239.

[50] K. Pieper, *Finite element discretization and efficient numerical solution of elliptic and parabolic sparse control problems*, Dissertation, Technische Universität München, 2015.

[51] Walter Rudin, *Reelle und komplexe Analysis*, R. Oldenbourg Verlag, Munich, 1999. Translated from the third English (1987) edition by Uwe Krieg. MR1736644

[52] Andreas Schindele and Alfio Borzì, *Proximal methods for elliptic optimal control problems with sparsity cost functional*, Applied Mathematics **7** (2016), 967–992.

[53] _____, *Proximal schemes for parabolic optimal control problems with sparsity promoting cost functionals*, Internat. J. Control **90** (2017), no. 11, 2349–2367. MR3704994

[54] Christopher Schneider and Gerd Wachsmuth, *Regularization and discretization error estimates for optimal control of ODEs with group sparsity*, ESAIM Control Optim. Calc. Var. **24** (2018), no. 2, 811–834. MR3816416

[55] Jürgen Sprekels and Fredi Tröltzsch, *Sparse optimal control of a phase field system with singular potentials arising in the modeling of tumor growth*, ESAIM Control Optim. Calc. Var. **27** (2021), no. suppl., Paper No. S26, 27. MR4222175

[56] Georg Stadler, *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices*, Comput. Optim. Appl. **44** (2009), no. 2, 159–181. MR2556849

[57] Robert Tibshirani, *Regression shrinkage and selection via the lasso*, J. Roy. Statist. Soc. Ser. B **58** (1996), no. 1, 267–288. MR1379242

[58] 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

[59] Daniel Wachsmuth and Gerd Wachsmuth, *Second-order conditions for non-uniformly convex integrands: quadratic growth in L^1*, 2021. Preprint. arXiv:2111.10238 [math.OC].

[60] 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

[61] Ming Yuan and Yi Lin, *Model selection and estimation in regression with grouped variables*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **68** (2006), no. 1, 49–67.