ESAIM: COCV 24 (2018) 811–834
https://doi.org/10.1051/cocv/2017049

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

REGULARIZATION AND DISCRETIZATION ERROR ESTIMATES FOR OPTIMAL CONTROL OF ODES WITH GROUP SPARSITY

Christopher Schneider[1] and Gerd Wachsmuth[2,a]

Abstract. It is well known that optimal control problems with L^1-control costs produce sparse solutions, *i.e.*, the optimal control is zero on whole intervals. In this paper, we study a general class of convex linear-quadratic optimal control problems with a sparsity functional that promotes a so-called *group sparsity* structure of the optimal controls. In this case, the components of the control function take the value of zero on parts of the time interval, simultaneously. These problems are both theoretically interesting and practically relevant. After obtaining results about the structure of the optimal controls, we derive stability estimates for the solution of the problem w.r.t. perturbations and L^2-regularization. These results are consequently applied to prove convergence of the Euler discretization. Finally, the usefulness of our approach is demonstrated by solving an illustrative example using a semismooth Newton method.

Mathematics Subject Classification. 49K15, 49J15, 49M15, 49M25, 65K15.

Received July 15, 2016. Revised February 21, 2017. Accepted June 29, 2017.

1. Introduction

We consider the optimal control problem

$$\begin{aligned} \text{Minimize} \quad & f(x,u) + \int_0^T \left\{ \frac{\alpha}{2} \, |u(t)|^2_{\mathbb{R}^m} + \beta \, |u(t)|_{\mathbb{R}^m} \right\} \mathrm{d}t \\ \text{with respect to} \quad & x \in W^{1,2}(0,T;\mathbb{R}^n), u \in L^2(0,T;\mathbb{R}^m) \\ \text{such that} \quad & \dot{x}(t) = A(t)\,x(t) + B(t)\,u(t) + b(t) \quad \text{for a.a. } t \in [0,T] \\ & x(0) = x_0 \\ \text{and} \quad & |u(t)|_{\mathbb{R}^m} \le u_b(t) \quad \text{for a.a. } t \in [0,T]. \end{aligned} \tag{\textbf{P}}$$

Here, $\beta \ge 0$ is a fixed weight for the L^1-control cost and $\alpha \ge 0$ is the L^2-regularization parameter. Furthermore, f is a linear-quadratic cost functional which is jointly convex in (x,u). Assumptions on the other problem data

Keywords and phrases. Optimal control, group sparsity, directional sparsity, bang-bang principle, stability analysis, discretization error estimates.

[1] Friedrich-Schiller-Universität Jena, Fakultät für Mathematik und Informatik, Professur Mathematische Optimierung, 07743 Jena, Germany.
christopher.schneider@uni-jena.de

[2] Technische Universität Chemnitz, Faculty of Mathematics, Professorship Numerical Mathematics (Partial Differential Equations), 09107 Chemnitz, Germany.
[a] Corresponding author: gerd.wachsmuth@mathematik.tu-chemnitz.de,
http://www.tu-chemnitz.de/mathematik/part_dgl/people/wachsmuth, http://orcid.org/0000-0002-3098-1503

Article published by EDP Sciences

will be made in Section 2.2. The main feature of (**P**) is that we use the Euclidean norm

$$|u|_{\mathbb{R}^m} := \left(\sum_{i=1}^{m} |u_i|^2 \right)^{\frac{1}{2}}$$

in the sparsity term of the cost functional, which is weighted by β, as well as in the control constraint $|u(t)|_{\mathbb{R}^m} \le u_b(t)$. Therefore, the control costs and the control constraints in (**P**) are isotropic and do not favor the coordinate directions. Depending on the real problem which is modeled by (**P**), this might be more adequate than the expressions

$$\beta \int_0^T \sum_{i=1}^{m} |u_i(t)| \qquad \text{and} \qquad |u_i(t)| \le u_b(t) \ \forall i = 1, \ldots, m,$$

which are usually used as a sparsity term in the cost functional and as control constraints, respectively. As an example, we mention the optimal control of a satellite moving in space. The satellite can accelerate in arbitrary directions and is constrained by a maximum thrust. In particular, this situation is isotropic and cannot be modelled by usual, coordinate-wise box constraints and the usual coordinate-wise L^1-objective. A similar example in two dimensions is discussed in Section 6.

Optimal control problems with the usual L^1-regularization term have been discussed by [2, 24, 27] for ODE constraints and [25,28,29] for PDE constraints. First and second order optimality conditions for the minimization of a general L^1-cost functional are obtained in [6]. The approach discussed in our paper is related to the group-sparsity approach in finite dimensions (see *e.g.* [7, 16]) and to [19] in which the PDE-constrained case (with $\alpha > 0$) is analyzed.

The stability of solutions of control problems governed by ODEs has been analyzed in [11] in the case that the optimal controls are Lipschitz continuous. For optimal control problems having bang-bang solutions, structure and stability of solutions has been studied *e.g.* by [3, 13, 14, 20–23]. Based on these results, the discretization of control problems with bang-bang solutions was considered by [1, 15, 26]. For an implicit discretization scheme, we refer the reader to [5].

The main contributions of our work are the following.

- We give first-order necessary and sufficient optimality conditions. In case $\beta > 0$, we observe sparsity of the control and for $\alpha = 0$, we get a bang-bang behavior of the optimal control. However, due to the isotropic control constraint, this is slightly different from the usual bang-bang structure.
- We provide a condition which yields stability of the optimal control in L^1 w.r.t. perturbations of the data in the case $\alpha = 0$. Again, we have to use novel ideas to treat with the isotropic behavior of the control constraints and the control cost.
- By employing this stability result, we prove convergence of the discretized optimal controls (with mesh size h) for the unregularized case $\alpha = 0$ as $h \searrow 0$ as well as for the regularized case $\alpha > 0$, by choosing $\alpha = c\,h$ and $h \searrow 0$.

The organization of the paper is as follows. In Section 2 we introduce notations and state the standing assumptions for problem (**P**). Section 3 deals with the existence and uniqueness of solutions as well as with first-order optimality conditions which are both necessary and sufficient, since the problem is convex. Based on these results, we derive the structure of the optimal controls in both cases $\alpha = 0$ and $\alpha > 0$. Assuming that the optimal control of the reference problem (**P**) has a bang-bang structure for $\alpha = 0$, we prove a stability result (Thm. 4.9) for the optimal controls w.r.t. standard perturbations p and the L^2-regularization parameter α in Section 4. In Section 5, we apply this stability result to prove convergence of the Euler discretization of problem (**P**). For the numerical solution of the discretized problem, we use the semismooth Newton method in Section 6. The usefulness of our group sparsity approach is then demonstrated by an illustrative example.

2. Notation and standing assumptions

2.1. Notation

We denote by $I := [0, T]$ the time interval, where $T > 0$ is the final time in (**P**). By $L^p(I;\mathbb{R}^m)$ and $W^{1,p}(I;\mathbb{R}^n)$, we denote the usual Bochner−Lebesgue spaces and Bochner−Sobolev spaces, respectively.

In the spaces $L^p(I;\mathbb{R}^m)$, $p \in [1,\infty]$, we use the norms

$$\|u\|_{L^p(I;\mathbb{R}^m)} := \left(\int_0^T |u(t)|^p_{\mathbb{R}^m} \,\mathrm{d}t\right)^{\frac{1}{p}}$$

for $p < \infty$ with the obvious modification for $p = \infty$. Here,

$$|u|_{\mathbb{R}^m} := \left(\sum_{i=1}^m |u_i|^2\right)^{\frac{1}{2}}$$

is the Euclidean norm in $\mathbb{R}^m$.

We emphasize that, in general,

$$\|u\|_{L^1(I;\mathbb{R}^m)} = \int_0^T \left(\sum_{i=1}^m |u_i(t)|^2\right)^{\frac{1}{2}} \mathrm{d}t \neq \int_0^T \sum_{i=1}^m |u_i(t)| \,\mathrm{d}t.$$

The latter norm is the traditional sparsity term which promotes coordinate-wise sparsity. This term is used in, *e.g.*, [2, 27]. Similarly, the constraint in (**P**) is not a (coordinate-wise) box constraint, but a constraint on the Euclidean norm.

We denote by c a generic constant, which may change from line to line. Similarly, we use the notation

$$a \lesssim b,$$

which means that there exists a constant $c > 0$, such that

$$a \leq c\, b$$

holds.

2.2. Standing assumptions

We define the parameters in the optimal control problem (**P**). For the right-hand side in the state equation we assume that $A\colon I \to \mathbb{R}^{n\times n}$, $B\colon I \to \mathbb{R}^{n\times m}$, and $b\colon I \to \mathbb{R}^n$ are Lipschitz continuous. The smooth part f of the objective is given by

$$f(x, u) := \frac{1}{2}\, x(T)^\top Q\, x(T) + q^\top x(T) + \int_0^T \frac{1}{2}\, x(t)^\top W(t)\, x(t) + w(t)^\top x(t) + r(t)^\top u(t) \,\mathrm{d}t. \tag{2.1}$$

The functions $W\colon I \to \mathbb{R}^{n\times n}$, $w\colon I \to \mathbb{R}^n$ and $r\colon I \to \mathbb{R}^m$ are assumed to be Lipschitz continuous as well. Moreover, the matrices Q and $W(t)$ are symmetric and positive semi-definite for all $t \in I$. Hence, the smooth part f of the objective is convex and continuous.

For simplicity, we require that the control bound $u_b \in L^\infty(I)$ satisfies $u_b(t) > 0$ (not necessarily uniform).

We define the non-smooth functionals $g\colon \mathbb{R}^m \to \mathbb{R}$ and $G\colon L^2(I;\mathbb{R}^m) \to \mathbb{R}$ by

$$g(u) := |u|_{\mathbb{R}^m}, \qquad G(u) := \int_0^T g(u(t)) \,\mathrm{d}t = \|u\|_{L^1(I;\mathbb{R}^m)}.$$

The reduced (smooth part of the) objective $F\colon L^2(I;\mathbb{R}^m) \to \mathbb{R}$ is defined by $F(u) := f(S(u),u)$, where $S : L^2(I;\mathbb{R}^m) \to W^{1,2}(I;\mathbb{R}^n)$ is the control-to-state mapping associated with (**P**).

Finally, we define the set U_{ad} of admissible controls via

$$\begin{aligned} U_{\mathrm{ad}}(t) &:= \{u \in \mathbb{R}^m : |u|_{\mathbb{R}^m} \le u_b(t)\}, \\ U_{\mathrm{ad}} &:= \{u \in L^2(I;\mathbb{R}^m) : u(t) \in U_{\mathrm{ad}}(t) \text{ for a.a. } t \in I\}. \end{aligned}$$

3. Existence and optimality conditions

In this section, we derive the first-order optimality conditions for Problem (**P**). In particular, we provide a projection formula in case $\alpha > 0$ and show that the optimal control is typically bang-bang in case $\alpha = 0$.

Since U_{ad} is bounded in $L^2(I;\mathbb{R}^m)$ and since the reduced objective of (**P**)

$$u \mapsto F(u) + \frac{\alpha}{2}\,\|u\|^2_{L^2(I;\mathbb{R}^m)} + \beta\,G(u)$$

is convex and continuous, the existence of an optimal control follows from standard arguments.

Theorem 3.1. *Problem* (**P**) *admits a solution.*

In case $\alpha > 0$, the reduced objective is strictly convex and we obtain the uniqueness of the optimal control. However, in case $\alpha = 0$, the optimal control might not be unique, but we still obtain the uniqueness of the switching function and the adjoint state, see Corollary 3.6 below.

The optimality conditions for (**P**) will contain the subdifferential of the non-smooth term $G = \|\cdot\|_{L^1(I;\mathbb{R}^m)}$ and we give a characterization of this subdifferential in the next lemma. We remark that the same result can be obtained by using ([18], Lem. 2.1) and we refer to ([19], Lem. 2.2, [9], Prop. 2.8) for similar results.

Lemma 3.2. *Let $u, v \in L^2(I;\mathbb{R}^m)$ be given. Then we have $v \in \partial G(u)$ if and only if*

$$|v(t)|_{\mathbb{R}^m} \le 1 \quad \textit{and} \quad \left(u(t) \neq 0 \Rightarrow v(t) = \frac{u(t)}{|u(t)|_{\mathbb{R}^m}}\right) \tag{3.1}$$

holds for a.a. $t \in I$. This, in turn, is equivalent to $v(t) \in \partial g(u(t))$ for a.a. $t \in I$.

Proof. "$\Leftarrow$:" Let us suppose that v satisfies (3.1). For arbitrary $w \in L^2(I;\mathbb{R}^m)$ we have

$$v(t)^\top (w(t) - u(t)) \le |v(t)|_{\mathbb{R}^m}\,|w(t)|_{\mathbb{R}^m} - v(t)^\top u(t) \le |w(t)|_{\mathbb{R}^m} - |u(t)|_{\mathbb{R}^m}$$

for a.a. $t \in I$. Integration over $t \in I$ yields

$$(v, w - u)_{L^2(I;\mathbb{R}^m)} \le G(w) - G(u).$$

Hence, $v \in \partial G(u)$.

"$\Rightarrow$:" Conversely, let $v \in \partial G(u)$ be given. By definition,

$$\int_I v^\top (w - u)\,\mathrm{d}t \le \int_I |w|_{\mathbb{R}^m} - |u|_{\mathbb{R}^m}\,\mathrm{d}t \qquad \forall w \in L^2(I;\mathbb{R}^m).$$

Using Lebesgue's differentiation theorem, this implies that

$$v(t)^\top (w - u(t)) \le |w|_{\mathbb{R}^m} - |u(t)|_{\mathbb{R}^m} \qquad \forall w \in \mathbb{R}^m$$

holds for a.a. $t \in I$. Hence, $v(t) \in \partial g(u(t))$ for a.a. $t \in I$. This yields (3.1). □

As a next step, we compute the directional derivatives of g and G.

Lemma 3.3. *For $u, v \in \mathbb{R}^m$ the directional derivative of g at u in direction v is given by*

$$g'(u;v) = \begin{cases} |v|_{\mathbb{R}^m} & \text{in case } u = 0 \\ \dfrac{u^\top v}{|u|_{\mathbb{R}^m}} & \text{in case } u \neq 0 \end{cases}$$

For $u, v \in L^2(I;\mathbb{R}^m)$ the directional derivative of G at u in direction v is given by

$$G'(u;v) = \int_0^T g'\big(u(t); v(t)\big)\,\mathrm{d}t.$$

Proof. The directional derivatives follow from standard calculations, see also ([9], Prop. 3.8). □

Using results from convex analysis, we can give an optimality condition for (**P**) involving the subdifferential of the non-smooth term G and this condition is reformulated by using Lemma 3.2.

Theorem 3.4. *Let (x^*, u^*) be a feasible point of (**P**). Then, (x^*, u^*) is an optimal solution of (**P**) if and only if there exists $\mu^* \in \partial G(u^*)$, such that*

$$\big(\sigma^*(t) + \alpha\, u^*(t) + \beta\, \mu^*(t)\big)^\top (u - u^*(t)) \geq 0 \quad \text{for all } u \in U_{\mathrm{ad}}(t) \tag{3.2a}$$

holds f.a.a. $t \in I$, where $\lambda^* \in W^{1,2}(I;\mathbb{R}^n)$ is the solution of the adjoint equation

$$-\dot{\lambda}^*(t) = A(t)^\top \lambda^*(t) + W(t)\, x^*(t) + w(t) \quad \text{for a.a. } t \in I, \tag{3.2b}$$

$$\lambda^*(T) = Q\, x^*(T) + q, \tag{3.2c}$$

and

$$\sigma^*(t) := B(t)^\top \lambda^*(t) + r(t) \tag{3.2d}$$

is the switching function.

Proof. The result follows from standard results in convex analysis applied to the reduced problem

$$\text{Minimize} \quad F(u) + \frac{\alpha}{2}\,\|u\|^2_{L^2(I;\mathbb{R}^m)} + \beta\, G(u) + I_{U_{\mathrm{ad}}}(u).$$

Here, $I_{U_{\mathrm{ad}}}$ is the indicator function of U_{ad}. Indeed, we can apply the usual sum-rule for the objective, since the first three terms are continuous. This yields the existence of $\mu^* \in \partial G(u^*)$ with

$$0 \in \partial F(u^*) + \alpha\, u^* + \beta\, \mu^* + \partial I_{U_{\mathrm{ad}}}(u^*).$$

Since the switching function, which is defined via (3.2b)–(3.2d), is (the Riesz representative in $L^2(I;\mathbb{R}^m)$ of) the derivative of F, *i.e.*,

$$\sigma^* = F'(u^*), \tag{3.3}$$

we obtain

$$\sigma^* + \alpha\, u^* + \beta\, \mu^* \in -\partial I_{U_{\mathrm{ad}}}(u^*).$$

This implies

$$(\sigma^* + \alpha\, u^* + \beta\, \mu^*, u - u^*)_{L^2(I;\mathbb{R}^m)} \geq I_{U_{\mathrm{ad}}}(u^*) - I_{U_{\mathrm{ad}}}(u) \geq 0$$

for all $u \in U_{\mathrm{ad}}$. Hence, condition (3.2a) follows. □

It is well known that the minimum principle (3.2a) can also be characterized by the normal cone of U_{ad}.

Corollary 3.5. *Let (x^*, u^*) be a feasible point of ($\mathbf{P}$) and let σ^* be the switching function defined via (3.2b)–(3.2d). Then, (x^*, u^*) is an optimal solution of ($\mathbf{P}$) if and only if there exists $\mu^* \in \partial G(u^*)$, such that*

$$\begin{aligned}\sigma^*(t) + \alpha\, u^*(t) + \beta\, \mu^*(t) &\in -\mathcal{N}_{U_{\mathrm{ad}}(t)}(u^*(t)) \\ &= \begin{cases}\{-\gamma\, u^*(t) : \gamma \ge 0\} & \text{if} \quad |u^*(t)|_{\mathbb{R}^m} = u_b(t), \\ \{0\} & \text{if} \quad |u^*(t)|_{\mathbb{R}^m} < u_b(t)\end{cases}\end{aligned} \tag{3.4}$$

holds f.a.a. $t \in I$, where $\mathcal{N}_{U_{\mathrm{ad}}(t)}(u^*(t))$ is the normal cone of the ball $U_{\mathrm{ad}}(t) = \{y \in \mathbb{R}^m : |y|_{\mathbb{R}^m} \le u_b(t)\}$ at $u^*(t)$.

As announced, we will show the uniqueness of the adjoint state and the switching function even in case $\alpha = 0$.

Corollary 3.6. *Let u_1, u_2 be two optimal controls of ($\mathbf{P}$). Then, $\lambda_1 = \lambda_2$ and $\sigma_1 = \sigma_2$, where λ_i and σ_i are the associated adjoint states and switching functions, respectively.*

Proof. Since the optimal control is unique in case $\alpha > 0$, it remains to study the case $\alpha = 0$.

We set $h := u_2 - u_1$. By Theorem 3.4, there is $\mu_1 \in \partial G(u_1)$ such that (3.2a) is satisfied with (u^*, σ^*, μ^*) replaced by (u_1, σ_1, μ_1). Since F is linear-quadratic, we have

$$\begin{aligned}F(u_1) + \beta\, G(u_1) &= F(u_2) + \beta\, G(u_2) \\ &= F(u_1) + F'(u_1)\, h + \frac{1}{2} F''(u_1)\, h^2 + \beta\, G(u_2) \\ &\ge F(u_1) + F'(u_1)\, h + \frac{1}{2} F''(u_1)\, h^2 + \beta\, G(u_1) + \beta \int_0^T \mu_1^\top h \,\mathrm{d}t \\ &\ge F(u_1) + \beta\, G(u_1).\end{aligned}$$

The last inequality follows from $F''(u_1)\, h^2 \ge 0$ and (3.2a). This chain of inequalities shows $F''(u_1)\, h^2 = 0$. We denote by $y \in W^{1,2}(I; \mathbb{R}^n)$ the difference of the states, *i.e.*, $y := S(u_2) - S(u_1)$. A straightforward calculation shows

$$F''(u_1)\, h^2 = y(T)^\top Q\, y(T) + \int_0^T y(t)^\top W(t)\, y(t) \,\mathrm{d}t.$$

From $F''(u_1)\, h^2 = 0$ and the positive semi-definiteness of Q and $W(t)$, $t \in I$, we conclude $Q\, y(T) = 0$ and $W(t)\, y(t) = 0$ for a.a. $t \in I$. The difference of the adjoint states satisfies the system

$$\begin{aligned}-(\dot\lambda_2(t) - \dot\lambda_1(t)) &= A(t)^\top (\lambda_2(t) - \lambda_1(t)) + W(t)\, y(t) \quad \text{for a.a. } t \in I, \\ \lambda_2(T) - \lambda_1(T) &= Q\, y(T),\end{aligned}$$

and, thus, we obtain $\lambda_1 = \lambda_2$. From (3.2d), we get $\sigma_1 = \sigma_2$. ☐

Now, we use the optimality condition (3.2a) in order to provide structural properties of a solution. First, we consider the case $\alpha = 0$, see also ([9], (43), (44)).

Lemma 3.7. *Let (x^*, u^*) be a solution of ($\mathbf{P}$) in case $\alpha = 0$ and $\beta > 0$. Then,*

$$|\sigma^*(t)|_{\mathbb{R}^m} > \beta \quad \Longrightarrow \quad u^*(t) = -\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}\, u_b(t), \tag{3.5a}$$

$$|\sigma^*(t)|_{\mathbb{R}^m} = \beta \quad \Longrightarrow \quad u^*(t) = -\nu\, \frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}\, u_b(t) \quad \textit{for some } \nu \in [0,1], \tag{3.5b}$$

$$|\sigma^*(t)|_{\mathbb{R}^m} < \beta \quad \Longrightarrow \quad u^*(t) = 0, \tag{3.5c}$$

and

$$\mu^*(t) = -\operatorname{Proj}_{[0,1]}\left(\frac{|\sigma^*(t)|_{\mathbb{R}^m}}{\beta}\right)\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}$$

hold for a.a. $t \in I$.

The assertion (3.5a) *also holds in case* $\alpha = \beta = 0$.

Proof. First, we prove (3.5a) under the assumptions $\alpha = 0$ and $\beta \geq 0$. Let $t \in I$ be given such that $|\sigma^*(t)|_{\mathbb{R}^m} > \beta$. By the triangle inequality we have

$$|\sigma^*(t) + \beta\,\mu^*(t)|_{\mathbb{R}^m} \geq |\sigma^*(t)|_{\mathbb{R}^m} - \beta\,|\mu^*(t)|_{\mathbb{R}^m} \geq |\sigma^*(t)|_{\mathbb{R}^m} - \beta > 0.$$

Since $\sigma^*(t) + \beta\,\mu^*(t) \neq 0$ and $\alpha = 0$, (3.4) implies

$$u^*(t) = -\frac{\sigma^*(t) + \beta\,\mu^*(t)}{|\sigma^*(t) + \beta\,\mu^*(t)|_{\mathbb{R}^m}}\,u_b(t) \neq 0. \tag{3.6}$$

Hence, Lemma 3.2 implies

$$\mu^*(t) = \frac{u^*(t)}{|u^*(t)|_{\mathbb{R}^m}}.$$

Together with (3.6), this yields the assertion.

From now on, we suppose $\alpha = 0$ and $\beta > 0$. Let $t \in I$ with $|\sigma^*(t)|_{\mathbb{R}^m} = \beta$ be given. We find

$$|\sigma^*(t) + \beta\,\mu^*(t)|_{\mathbb{R}^m} \geq |\sigma^*(t)|_{\mathbb{R}^m} - \beta\,|\mu^*(t)|_{\mathbb{R}^m} \geq |\sigma^*(t)|_{\mathbb{R}^m} - \beta \geq 0. \tag{3.7}$$

In case $|\sigma^*(t) + \beta\,\mu^*(t)|_{\mathbb{R}^m} > 0$, we can argue as above and arrive at

$$u^*(t) = -\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}\,u_b(t), \qquad \mu^*(t) = \frac{u^*(t)}{|u^*|_{\mathbb{R}^m}} = -\frac{\sigma^*(t)}{|\sigma^*|_{\mathbb{R}^m}} = -\frac{\sigma^*(t)}{\beta}.$$

Otherwise, if $|\sigma^*(t)+\beta\,\mu^*(t)|_{\mathbb{R}^m} = 0$, (3.7) implies $\sigma^*(t) = -\beta\,\mu^*(t)$ and $|\mu^*(t)|_{\mathbb{R}^m} = 1$. Together with Lemma 3.2, we arrive at

$$u^*(t) = |u^*(t)|\,\mu^*(t) = -|u^*(t)|\,\frac{\sigma^*(t)}{\beta} = -\nu\,\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}\,u_b(t)$$

with $\nu = |u^*(t)|/u_b(t) \in [0,1]$.

It remains to prove the third implication. We argue by contradiction and assume $|\sigma^*(t)|_{\mathbb{R}^m} < \beta$ and $u^*(t) \neq 0$. Now, (3.4) yields

$$\sigma^*(t) + \beta\,\mu^*(t) = \sigma^*(t) + \beta\,\frac{u^*(t)}{|u^*(t)|_{\mathbb{R}^m}} = -\gamma\,u^*(t)$$

for some $\gamma \geq 0$. This implies the contradiction

$$\beta > |\sigma^*|_{\mathbb{R}^m} = \gamma\,|u^*(t)|_{\mathbb{R}^m} + \beta \geq \beta,$$

hence $u^*(t) = 0$. Since $u_b(t) > 0$, (3.4) implies $\sigma^*(t) + \beta\,\mu^*(t) = 0$ and the formula for $\mu^*(t)$ follows. $\square$

Note that $u^*(t)$ is uniquely determined on the set $\{t \in I : |\sigma^*(t)| \neq \beta\}$, where σ^* is the uniquely determined switching function. Moreover, the multiplier μ^* is unique in case $\beta > 0$, since it is uniquely determined by the (unique) switching function. In the case $\beta = 0$, the multiplier μ^* is superfluous since it is muliplied by β, *cf.* (3.2a).

Let us discuss the structural properties of u^*, which follow from Lemma 3.7. As in the case of traditional sparsity, the behavior of the optimal control depends on the amplitude of the switching function. For the traditional sparsity term, this coupling is coordinate-wise. In contrast, our sparsity term couples all components of u^* and σ^*. Indeed, in case $|\sigma^*(t)|_{\mathbb{R}^m} < \beta$, *all* components of $u^*(t)$ are forced to be zero.

Another difference to the traditional approach is that u^* still depends on σ^* in case $|\sigma^*| > \beta$. In the traditional situation, we have $u^* = u_b$ in case $\sigma^* < -\beta$, hence u^* does not depend on σ^* (as long as σ^* stays smaller than $-\beta$).

In case $\alpha > 0$, we even obtain a projection formula for u^*, see also ([19], (3.4))

Lemma 3.8. *Let (x^*, u^*) be a solution of (**P**) and assume $\alpha > 0$. Then,*

$$u^*(t) = -\operatorname{Proj}_{[0,u_b(t)]}\left(\frac{|\sigma^*(t)|_{\mathbb{R}^m} - \beta}{\alpha}\right)\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}} \tag{3.8a}$$

holds for a.a. $t \in I$. In case $\beta > 0$, we additionally have

$$\mu^*(t) = -\operatorname{Proj}_{[0,1]}\left(\frac{|\sigma^*(t)|_{\mathbb{R}^m}}{\beta}\right)\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}} \tag{3.8b}$$

for a.a. $t \in I$.

*Moreover, if a feasible point (x^*, u^*) of (**P**) satisfies (3.8a) with σ^* determined by (3.2b)–(3.2d), then it is a solution of (**P**).*

Proof. We make a distinction by cases.

In case $|\sigma^*(t)|_{\mathbb{R}^m} \le \beta$ we can argue similarly as in the last part of the proof of Lemma 3.7 and obtain $u^*(t) = 0$. Now, (3.4) implies $\mu^*(t) = -\sigma^*(t)/\beta$.

Now let us assume $|\sigma^*(t)|_{\mathbb{R}^m} > \beta$. As in the first part of the proof of Lemma 3.7 we obtain $|\sigma^*(t) + \beta\,\mu^*(t)|_{\mathbb{R}^m} > 0$. From (3.4) we find $u^*(t) \ne 0$ and, thus, $\mu^*(t) = u^*(t)/|u^*(t)|_{\mathbb{R}^m}$. Invoking again (3.4), we find that the vectors $u^*(t)$ and $-\sigma^*(t)$ point in the same direction. This yields

$$\mu^*(t) = \frac{u^*(t)}{|u^*(t)|_{\mathbb{R}^m}} = \frac{-\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}},$$

and (3.8b) follows. Using (3.4), we find

$$\big(|\sigma^*|_{\mathbb{R}^m} - \beta\big)\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}} + \alpha\, u^*(t) \in \mathcal{N}_{U_{\mathrm{ad}}(t)}(u^*(t)).$$

This leads to the projection formula (3.8a).

Finally, let the feasible point (x^*, u^*) of (**P**) satisfy (3.8a). In case $\beta = 0$, we can check immediately that (3.4) is satisfied and this shows the optimality. In case $\beta > 0$, we can use (3.8b) as a definition of μ^* and, again, (3.4) can be verified. □

Lemmas 3.7 and 3.8 show the sparsity properties of (**P**). In particular, the relation $|\sigma^*(t)|_{\mathbb{R}^m} < \beta$ implies $|u^*(t)|_{\mathbb{R}^m} = 0$. This also demonstrates the *group* sparsity structure: in the sparse case $|\sigma^*(t)|_{\mathbb{R}^m} < \beta$ all components of $u^*(t)$ are *simultaneously* zero.

4. Stability of the solution with respect to perturbations

In this section, we establish a quantitative stability result for the solution of (**P**) with some fixed $\beta \ge 0$ under perturbations of the problem data. The main ingredient is a growth condition of the objective of (**P**) in the neighborhood of the solution and this will be provided in Theorem 4.4.

4.1. Growth condition

We begin our analysis by considering the unregularized case $\alpha = 0$ in the optimal control problem (**P**). By $(x^*, u^*, \lambda^*, \mu^*, \sigma^*)$ we denote a solution of the optimality system (3.2). The main result of this section is a growth condition for the objective $F + \beta G$.

Assumption 4.1. There exist constants $C > 0$, $\kappa \in (0, 1] \cup \{\infty\}$ and $\varepsilon_0 > 0$, such that

$$\operatorname{meas}\left\{t \in I : \big||\sigma^*(t)|_{\mathbb{R}^m} - \beta\big| \le \varepsilon\right\} \le C\,\varepsilon^\kappa \tag{4.1}$$

holds for all $\varepsilon \in [0, \infty)$ in case $\kappa \ne \infty$; and

$$\operatorname{meas}\left\{t \in I : \big||\sigma^*(t)|_{\mathbb{R}^m} - \beta\big| \le \varepsilon_0\right\} = 0 \tag{4.2}$$

in case $\kappa = \infty$.

In case $\kappa = \infty$, we will use the convention $1/\kappa = 0$.

We mention that an assumption similar to our Assumption 4.1 is commonly used to verify stability of solutions in the unregularized case $\alpha = 0$, see [1, 3, 4, 10, 13, 23, 26, 28, 29].

The satisfaction of Assumption 4.1 implies that the associated optimal control u^* is the *unique* global solution of Assumption (**P**). Indeed, from Assumption 4.1 we get $\operatorname{meas}\{t \in I : |\sigma^*(t)| = \beta\} = 0$ and, thus, u^* is uniquely determined by the uniqueness of σ^*, see Corollary 3.6 and Lemma 3.7.

Lemma 4.2. *Let u^* be an optimal control of (**P**) such that the associated switching function σ^* satisfies Assumption 4.1. Then,*

$$\begin{aligned}
&F'(u^*)\,(u - u^*) + \beta\, G'(u^*; u - u^*)\\
&\quad = \int_0^T \sigma^*(t)^\top (u(t) - u^*(t)) + \beta\, g'\big(u^*(t); u(t) - u^*(t)\big)\,\mathrm{d}t\\
&\quad \gtrsim \|u - u^*\|^{1+1/\kappa}_{L^1(I^-;\mathbb{R}^m)} + \|u_b^{-1}\,(u^*)^\top\,(u - u^*)\|^{1+1/\kappa}_{L^1(I^+)}
\end{aligned}$$

holds for all $u \in U_{\mathrm{ad}}$. Here,

$$I^+ := \{t \in I : |\sigma^*(t)|_{\mathbb{R}^m} > \beta\} \quad \textit{and} \quad I^- := \{t \in I : |\sigma^*(t)|_{\mathbb{R}^m} < \beta\}.$$

Proof. The first equality in the assertion follows from (3.3) and Lemma 3.3.

We take $\varepsilon \in (0, \infty)$ which will be specified later in case $\kappa \ne \infty$ and $\varepsilon := \varepsilon_0$ in case $\kappa = \infty$. We define

$$I_\varepsilon^- = \{t \in I : |\sigma^*(t)|_{\mathbb{R}^m} < \beta - \varepsilon\}.$$

For $t \in I^-$ we have $u^*(t) = 0$, see Lemma 3.7. Hence, $g'\big(u^*(t); u(t)\big) = |u(t)|_{\mathbb{R}^m}$ by 3.3. This implies

$$\begin{aligned}
&\int_{I^-} \sigma^*(t)^\top (u(t) - u^*(t)) + \beta\, g'\big(u^*(t); u(t) - u^*(t)\big)\,\mathrm{d}t\\
&\quad = \int_{I^-} \sigma^*(t)^\top u(t) + \beta\,|u(t)|_{\mathbb{R}^m}\,\mathrm{d}t\\
&\quad \ge \int_{I_\varepsilon^-} \sigma^*(t)^\top u(t) + \beta\,|u(t)|_{\mathbb{R}^m}\,\mathrm{d}t\\
&\quad \ge \varepsilon \int_{I_\varepsilon^-} |u(t)|_{\mathbb{R}^m}\,\mathrm{d}t\\
&\quad = \varepsilon \int_{I^-} |u(t)|_{\mathbb{R}^m}\,\mathrm{d}t - \varepsilon \int_{I^- \setminus I_\varepsilon^-} |u(t)|_{\mathbb{R}^m}\,\mathrm{d}t\\
&\quad \ge \varepsilon \int_{I^-} |u(t)|_{\mathbb{R}^m}\,\mathrm{d}t - C\,\varepsilon^{1+\kappa}\,\|u_b\|_{L^\infty(I)}.
\end{aligned}$$

Here, we used $\operatorname{meas}(I^- \setminus I^-_\varepsilon) \le C\,\varepsilon^\kappa$, which holds due to Assumption 4.1. In case $\kappa \ne \infty$, we choose the value $\varepsilon := \big(\frac{1}{2\,C\,\|u_b\|_{L^\infty(I)}}\,\|u\|_{L^1(I^-;\mathbb{R}^m)}\big)^{1/\kappa}$. This leads to

$$\int_{I^-} \sigma^*(t)^\top (u(t) - u^*(t)) + \beta\, g'\big(u^*(t); u(t) - u^*(t)\big)\,\mathrm{d}t \ge c\,\|u - u^*\|^{1+1/\kappa}_{L^1(I^-;\mathbb{R}^m)}. \tag{4.3}$$

In case $\kappa = \infty$, (4.3) follows since $\operatorname{meas}(I^- \setminus I^-_\varepsilon) = 0$.

For $t \in I^+$ we have

$$u^*(t) = -\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}\,u_b(t) \ne 0, \qquad \text{and} \qquad \sigma^*(t) = -\frac{|\sigma^*(t)|_{\mathbb{R}^m}}{u_b(t)}\,u^*(t),$$

see Lemma 3.7. Hence,

$$g'\big(u^*(t); v\big) = \frac{u^*(t)^\top v}{|u^*(t)|_{\mathbb{R}^m}}$$

We take $\varepsilon \in (0,\infty)$ which will be specified later in case $\kappa \ne \infty$ and $\varepsilon := \varepsilon_0$ in case $\kappa = \infty$. We define

$$I^+_\varepsilon = \{t \in I : |\sigma^*(t)|_{\mathbb{R}^m} > \beta + \varepsilon\}.$$

This yields

$$\begin{aligned}\int_{I^+} &\sigma^*(t)^\top (u(t) - u^*(t)) + \beta\, g'\big(u^*(t); u(t) - u^*(t)\big)\,\mathrm{d}t \\ &= \int_{I^+} (|\sigma^*(t)|_{\mathbb{R}^m} - \beta)\,u_b(t)^{-1}\,u^*(t)^\top (u^*(t) - u(t))\,\mathrm{d}t.\end{aligned}$$

Since $u^*(t)^\top\,(u^*(t) - u(t)) \ge |u^*(t)|^2_{\mathbb{R}^m} - |u^*(t)|_{\mathbb{R}^m}\,|u(t)|_{\mathbb{R}^m} \ge 0$, we find

$$\begin{aligned}\int_{I^+} &\sigma^*(t)^\top (u(t) - u^*(t)) + \beta\, g'\big(u^*(t); u(t) - u^*(t)\big)\,\mathrm{d}t \\ &\ge \varepsilon \int_{I^+_\varepsilon} u_b(t)^{-1}\,u^*(t)^\top (u^*(t) - u(t))\,\mathrm{d}t \\ &= \varepsilon \left(\int_{I^+} u_b(t)^{-1}\,u^*(t)^\top (u^*(t) - u(t))\,\mathrm{d}t - \int_{I^+\setminus I^+_\varepsilon} u_b(t)^{-1}\,u^*(t)^\top (u^*(t) - u(t))\,\mathrm{d}t\right) \\ &\ge \varepsilon \left(\int_{I^+} u_b(t)^{-1}\,u^*(t)^\top (u^*(t) - u(t))\,\mathrm{d}t - 2\,\|u_b\|_{L^\infty(I)}\,C\,\varepsilon^\kappa\right).\end{aligned}$$

In case $\kappa \ne \infty$, we choose $\varepsilon := \big(\frac{1}{4\,C\,\|u_b\|_{L^\infty(I)}}\,\|u_b^{-1}\,(u^*)^\top\,(u - u^*)\|_{L^1(I^+)}\big)^{1/\kappa}$. This yields

$$\int_{I^+} \sigma^*(t)^\top (u(t) - u^*(t)) + \beta\, g'\big(u^*(t); u(t) - u^*(t)\big)\,\mathrm{d}t \ge c\,\|u_b^{-1}\,(u^*)^\top\,(u - u^*)\|^{1+1/\kappa}_{L^1(I^+)}.$$

In case $\kappa = \infty$, the same estimate follows since $\operatorname{meas}(I^+ \setminus I^+_\varepsilon) = 0$. Together with (4.3), this yields the claim. □

Remark 4.3. It is crucial to use the directional derivative $G'(u^*;\cdot)$ in the formulation of Lemma 4.2. In fact, it is not possible to obtain a similar result by replacing $G'(u^*;\cdot)$ with some μ^* satisfying the optimality condition from Theorem 3.4, since the term

$$F'(u^*)\,(u - u^*) + \beta\,(\mu^*, u - u^*)_{L^2(I;\mathbb{R}^m)} = (\sigma^* + \beta\,\mu^*, u - u^*)_{L^2(I;\mathbb{R}^m)}$$

contains no information of $u - u^*$ on I^-, compare Lemma 3.7, which shows $\sigma^* + \beta\,\mu^* = 0$ on I^-.

Based on Lemma 4.2 we can show that a certain growth condition is satisfied for the solution u^*.

Theorem 4.4. *Let u^* be an optimal control of* (**P**) *such that the associated switching function σ^* satisfies Assumption 4.1. Then, there exists $c > 0$ such that*

$$F(u) + \beta\, G(u) \geq F(u^*) + \beta\, G(u^*) + c\,\|u - u^*\|^{2\,(1+1/\kappa)}_{L^2(I;\mathbb{R}^m)}$$

holds for all $u \in U_{\mathrm{ad}}$.

Proof. From the convexity of F and G we find

$$F(u) + \beta\, G(u) \geq F(u^*) + \beta\, G(u^*) + F'(u^*)\,(u - u^*) + \beta\, G'(u^*; u - u^*).$$

Now, Lemma 4.2 implies

$$F(u) + \beta\, G(u) - (F(u^*) + \beta\, G(u^*)) \gtrsim \|u - u^*\|^{1+1/\kappa}_{L^1(I^-;\mathbb{R}^m)} + \|u_b^{-1}\,(u^*)^\top\,(u - u^*)\|^{1+1/\kappa}_{L^1(I^+)}.$$

It remains to estimate the two norms on the right-hand side.

For a.a. $t \in I^+$, we have $|u^*(t)|_{\mathbb{R}^m} = u_b(t)$ and $|u(t)| \leq u_b(t)$. Hence,

$$2\,u^*(t)^\top(u^*(t) - u(t)) = |u^*(t)|^2_{\mathbb{R}^m} - |u(t)|^2_{\mathbb{R}^m} + |u^*(t) - u(t)|^2_{\mathbb{R}^m} \geq |u^*(t) - u(t)|^2_{\mathbb{R}^m}.$$

Thus,

$$\begin{aligned}\|u_b^{-1}\,(u^*)^\top\,(u - u^*)\|_{L^1(I^+)} &\geq \frac{1}{\|u_b\|_{L^\infty(I)}}\,\|(u^*)^\top\,(u - u^*)\|_{L^1(I^+)} \\ &\geq \frac{1}{2\,\|u_b\|_{L^\infty(I)}}\,\|u - u^*\|^2_{L^2(I^+;\mathbb{R}^m)}.\end{aligned}$$

On I^-, we use Hölder's inequality and the bound constraints to obtain

$$\begin{aligned}\|u - u^*\|^2_{L^2(I^-;\mathbb{R}^m)} &\leq \|u - u^*\|_{L^\infty(I^-;\mathbb{R}^m)}\,\|u - u^*\|_{L^1(I^-;\mathbb{R}^m)} \\ &\leq 2\,\|u_b\|_{L^\infty(I)}\,\|u - u^*\|_{L^1(I^-;\mathbb{R}^m)}.\end{aligned}$$

This shows the assertion. ☐

We remark that the rate of growth in Theorem 4.4 is worse than in the traditional, coordinate-wise setting. Indeed, in this case one obtains a growth in $L^1(I;\mathbb{R}^m)$ with rate $1 + 1/\kappa$ by using similar arguments, (*cf.* [13], Thm. 3.4) for the case $\kappa = 1$.

Hence, using Theorem 4.4 to derive stability estimates in our situation would lead to results which are not sharp. Instead, we have to work directly with Lemma 4.2.

4.2. Perturbations and stability

In this section, we consider a perturbation of problem (**P**) and show the stability of solutions under Assumption 4.1. By $(x^*, u^*, \lambda^*, \mu^*, \sigma^*)$ we denote a solution of the optimality system (3.2) of (**P**) with $\alpha = 0$ and we shall assume that Assumption 4.1 is satisfied by this solution.

To be precise, we define a perturbation to be a quadruple

$$p = (p_b, p_q, p_w, p_r) \in L^1(I;\mathbb{R}^n) \times \mathbb{R}^n \times L^1(I;\mathbb{R}^n) \times L^\infty(I;\mathbb{R}^m)$$

and this perturbation induces the perturbed problem (with $\alpha \geq 0$)

$$\begin{aligned}\text{Minimize}\quad & f_p(x,u) + \frac{\alpha}{2}\,\|u\|^2_{L^2(I;\mathbb{R}^m)} + \beta\,\|u\|_{L^1(I;\mathbb{R}^m)} \\ \text{such that}\quad & \dot{x}(t) = A(t)\,x(t) + B(t)\,u(t) + b(t) + p_b(t) \quad \text{for a.a. } t \in I \\ & x(0) = x_0 \\ \text{and}\quad & |u(t)|_{\mathbb{R}^m} \leq u_b(t) \quad \text{for a.a. } t \in I,\end{aligned} \tag{$\mathbf{P}_p^\alpha$}$$

where the perturbation of the objective is given by

$$f_p(x,u) := \frac{1}{2}\, x(T)^\top Q\, x(T) + \big(q + p_q\big)^\top x(T) \\ + \int_0^T \frac{1}{2}\, x(t)^\top W(t)\, x(t) + \big(w(t) + p_w(t)\big)^\top x(t) + \big(r(t) + p_r(t)\big)^\top u(t)\, \mathrm{d}t. \tag{4.4}$$

The solution of ($\mathbf{P}_p^\alpha$) is denoted by u_p^α. Note that $u_0^0 = u^*$, since ($\mathbf{P}$) is uniquely solvable due to Assumption 4.1. Note that u_p^α might not be unique in case $\alpha = 0$. Nevertheless, our results will hold for any solution u_p^α of ($\mathbf{P}_p^\alpha$).

We define

$$\|p\| := \|p_b\|_{L^1(I;\mathbb{R}^n)} + |p_q|_{\mathbb{R}^n} + \|p_w\|_{L^1(I;\mathbb{R}^n)} + \|p_r\|_{L^\infty(I;\mathbb{R}^m)}.$$

Similar as in Section 3 we obtain an optimality system for ($\mathbf{P}_p^\alpha$).

Theorem 4.5. *Let (x_p^α, u_p^α) be a feasible point of ($\mathbf{P}_p^\alpha$). Then, (x_p^α, u_p^α) is an optimal solution of ($\mathbf{P}_p^\alpha$) if and only if there exists $\mu_p^\alpha \in \partial G(u_p^\alpha)$, such that*

$$\big(\sigma_p^\alpha(t) + \alpha\, u_p^\alpha(t) + \beta\, \mu_p^\alpha(t)\big)^\top (u - u_p^\alpha(t)) \ge 0 \quad \text{for all } u \in U_{\mathrm{ad}}(t) \tag{4.5a}$$

holds f.a.a. $t \in I$, where $\lambda_p^\alpha \in W^{1,2}(I;\mathbb{R}^n)$ is the solution of the adjoint equation

$$-\dot\lambda_p^\alpha(t) = A(t)^\top \lambda_p^\alpha(t) + W(t)\, x_p^\alpha(t) + w(t) + p_w(t) \quad \text{for a.a. } t \in I, \tag{4.5b}$$

$$\lambda_p^\alpha(T) = Q\, x_p^\alpha(T) + q + p_q, \tag{4.5c}$$

and

$$\sigma_p^\alpha(t) := B(t)^\top \lambda_p^\alpha(t) + r(t) + p_r(t) \tag{4.5d}$$

is the switching function.

Proof. This follows from Theorem 3.4. □

We remark that Lemmas 3.7 and 3.8 carry over to the solutions u_p^α for the perturbed problems ($\mathbf{P}_p^\alpha$).

The first step towards the stability of u^* is the following lemma, which provides an estimate for the switching function in terms of the perturbation.

Lemma 4.6. *We have*

$$\|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)}^2 \lesssim \int_0^T \big(\sigma_p^\alpha(t) - \sigma^*(t)\big)^\top (u_p^\alpha(t) - u^*(t))\, \mathrm{d}t \\ + c\, \|p\|\, \|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} + c\, \|p\|^2.$$

Proof. We define $u := u_p^\alpha - u^*$, $x := x_p^\alpha - x^*$, $\lambda := \lambda_p^\alpha - \lambda^*$, and $\sigma := \sigma_p^\alpha - \sigma^*$. By the system equations for x_p^α and x^* we have

$$\dot x(t) = A(t)\, x(t) + B(t)\, u(t) + p_b(t) \quad \text{for a.a. } t \in I,$$
$$x(t_0) = 0,$$

and this implies

$$-\int_0^T \lambda(t)^\top B(t)\, u(t)\, \mathrm{d}t = -\int_0^T \lambda(t)^\top \left[\dot x(t) - A(t)\, x(t) - p_b(t)\right] \mathrm{d}t. \tag{4.6}$$

From the adjoint equations for λ_p^α and λ^* we obtain

$$\begin{aligned} -\dot\lambda(t) &= A(t)^\top \lambda(t) + W(t)\,x(t) + p_w(t) \quad \text{for a.a. } t \in I,\\ \lambda(T) &= Q\,x(T) + p_q. \end{aligned}$$

Thereby, we have

$$\begin{aligned} -\int_0^T \lambda(t)^\top \dot x(t)\,\mathrm{d}t &= -\lambda(T)^\top x(T) + \int_0^T \dot\lambda(t)^\top x(t)\,\mathrm{d}t\\ &= -x(T)^\top Q\,x(T) - p_q^\top x(T)\\ &\quad - \int_0^T \left[A(t)^\top\lambda(t) + W(t)\,x(t) + p_w(t)\right]^\top x(t)\,\mathrm{d}t. \end{aligned}$$

Together with (4.6) we obtain

$$\begin{aligned} -\int_0^T \lambda(t)^\top B(t)\,u(t)\,\mathrm{d}t &= \int_0^T \lambda(t)^\top \left[A(t)\,x(t) + p_b(t)\right]\mathrm{d}t - x(T)^\top Q\,x(T) - p_q^\top x(T)\\ &\quad - \int_0^T \left[A(t)^\top\lambda(t) + W(t)\,x(t) + p_w(t)\right]^\top x(t)\,\mathrm{d}t\\ &= -x(T)^\top Q\,x(T) - p_q^\top x(T) + \int_0^T \lambda(t)^\top p_b(t) - p_w(t)^\top x(t)\,\mathrm{d}t\\ &\quad - \int_0^T x(t)^\top W(t)\,x(t)\,\mathrm{d}t. \end{aligned}$$

Rearranging terms yields

$$\begin{aligned} &x(T)^\top Q\,x(T) + \int_0^T x(t)^\top W(t)\,x(t) - \lambda(t)^\top B(t)\,u(t)\,\mathrm{d}t\\ &\quad\le -p_q^\top x(T) + \int_0^T \lambda(t)^\top p_b(t) - p_w(t)^\top x(t)\,\mathrm{d}t\\ &\quad\le |p_q|_{\mathbb{R}^n}\,\|x\|_{L^\infty(I;\mathbb{R}^n)} + \|\lambda\|_{L^\infty(I;\mathbb{R}^n)}\,\|p_b\|_{L^1(I;\mathbb{R}^n)} + \|p_w\|_{L^1(I;\mathbb{R}^n)}\,\|x\|_{L^\infty(I;\mathbb{R}^n)}\\ &\quad\le \big(\|x\|_{L^\infty(I;\mathbb{R}^n)} + \|\lambda\|_{L^\infty(I;\mathbb{R}^n)}\big)\,\|p\|. \end{aligned} \tag{4.7}$$

From the system equation, we know that

$$\|x\|_{L^\infty(I;\mathbb{R}^n)} \lesssim \|x\|_{W^{1,1}(I;\mathbb{R}^n)} \lesssim \|u\|_{L^1(I;\mathbb{R}^m)} + \|p_b\|_{L^1(I;\mathbb{R}^n)}.$$

From the adjoint equation, (4.7), we get

$$\begin{aligned} \|\lambda\|^2_{L^\infty(I;\mathbb{R}^n)} \lesssim \|\lambda\|^2_{W^{1,1}(I;\mathbb{R}^n)} &\lesssim \|W\,x\|^2_{L^1(I;\mathbb{R}^n)} + |Q\,x(T)|^2_{\mathbb{R}^m} + \|p_w\|^2_{L^1(I;\mathbb{R}^n)} + |p_q|^2_{\mathbb{R}^n}\\ &\lesssim \int_0^T x(t)^\top W(t)\,x(t)\,\mathrm{d}t + x(T)^\top Q\,x(T) + \|p_w\|^2_{L^1(I;\mathbb{R}^n)} + |p_q|^2_{\mathbb{R}^n}\\ &\lesssim \int_0^T \lambda(t)^\top B(t)\,u(t)\,\mathrm{d}t + c\,\|p\|\,\big(\|p\| + \|\lambda\|_{L^\infty(I;\mathbb{R}^n)} + \|u\|_{L^1(I;\mathbb{R}^m)}\big). \end{aligned}$$

Hence, there is a constant $\gamma > 0$ with

$$\gamma\,\|\lambda\|^2_{L^\infty(I;\mathbb{R}^n)} \le \int_0^T \lambda(t)^\top B(t)\,u(t)\,\mathrm{d}t + c\,\|p\|\,\big(\|p\| + \|\lambda\|_{L^\infty(I;\mathbb{R}^n)} + \|u\|_{L^1(I;\mathbb{R}^m)}\big).$$

Using Young's inequality, we get

$$\gamma \, \|\lambda\|_{L^\infty(I;\mathbb{R}^n)}^2 \le \int_0^T \lambda(t)^\top B(t)\, u(t) \, \mathrm{d}t + \hat{c} \, \|p\| \, \big(\|p\| + \|u\|_{L^1(I;\mathbb{R}^m)}\big) + \frac{\gamma}{2} \, \|\lambda\|_{L^\infty(I;\mathbb{R}^n)}^2,$$

where $\hat{c} = c + c^2/(2\,\gamma)$. Hence, we can move the last addend to the left-hand side and obtain

$$\|\lambda\|_{L^\infty(I;\mathbb{R}^n)}^2 \lesssim \int_0^T \lambda(t)^\top B(t)\, u(t) \, \mathrm{d}t + \hat{c} \, \|p\| \, \big(\|p\| + \|u\|_{L^1(I;\mathbb{R}^m)}\big).$$

Now, we use (3.2d) and (4.5d) to obtain

$$\begin{aligned}
\|\sigma\|_{L^\infty(I;\mathbb{R}^n)}^2 &\lesssim \|\lambda\|_{L^\infty(I;\mathbb{R}^n)}^2 \\
&\lesssim \int_0^T \lambda(t)^\top B(t)\, u(t) \, \mathrm{d}t + \hat{c} \, \|p\| \, \big(\|p\| + \|u\|_{L^1(I;\mathbb{R}^m)}\big) \\
&\lesssim \int_0^T \big[B(t)^\top \lambda(t) + p_r(t)\big]^\top u(t) \, \mathrm{d}t - \int_0^T p_r(t)^\top u(t) \, \mathrm{d}t \\
&\quad + \hat{c} \, \|p\| \, \big(\|p\| + \|u\|_{L^1(I;\mathbb{R}^m)}\big) \\
&\lesssim \int_0^T \sigma(t)^\top u(t) \, \mathrm{d}t + \hat{c} \, \|p\|^2 + (\hat{c} + 1) \, \|p\| \, \|u\|_{L^1(I;\mathbb{R}^m)}.
\end{aligned}$$

□

Together with the first-order conditions, we get an estimate for the controls.

Lemma 4.7. *Assume that Assumption 4.1 is satisfied. Then,*

$$\begin{aligned}
\|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)}^2 + \|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|u_b^{-1} \, (u^*)^\top \, (u_p^\alpha - u^*)\|_{L^1(I^+)}^{1+1/\kappa} \\
\lesssim \|p\| \, \|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} + \|p\|^2 + \alpha^{1+\kappa}
\end{aligned}$$

holds in case $\kappa \ne \infty$; and in case $\kappa = \infty$ if α is sufficiently small, with $\alpha^{1+\kappa} = 0$ by convention.

Proof. The main ingredients in this proof are the growth condition from Lemma 4.2 for the unperturbed problem and the minimum principle (4.5a) for the perturbed problem. We use $u = u^*$ in (4.5a) and $u = u_p^\alpha$ in Lemma 4.2. This yields

$$\begin{aligned}
&\|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|u_b^{-1} \, (u^*)^\top \, (u_p^\alpha - u^*)\|_{L^1(I^+)}^{1+1/\kappa} \\
&\quad\lesssim \int_0^T \Big(B(t)^\top \, \lambda_p^\alpha(t) + r(t) + p_r(t) + \alpha \, u_p^\alpha(t) + \beta \, \mu_p^\alpha(t)\Big)^\top (u^*(t) - u_p^\alpha) \, \mathrm{d}t \\
&\qquad + \int_0^T \Big(B(t)^\top \, \lambda^*(t) + r(t)\Big)^\top (u_p^\alpha(t) - u^*) \, \mathrm{d}t + \beta \, G'(u^*; u_p^\alpha - u^*)
\end{aligned}$$

Since G is convex and $\mu_p^\alpha \in \partial G(u_p^\alpha)$, we have

$$\begin{aligned}
\int_0^T \mu_p^\alpha(t)^\top (u^*(t) - u_p^\alpha(t)) \, \mathrm{d}t &\le G(u^*) - G(u_p^\alpha), \\
G'(u^*; u_p^\alpha - u^*) &\le G(u_p^\alpha) - G(u^*).
\end{aligned}$$

Hence,

$$\int_0^T \mu_p^\alpha(t)^\top (u^*(t) - u_p^\alpha(t)) \, \mathrm{d}t + G'(u^*; u_p^\alpha - u^*) \le 0.$$

Together with the first estimate and Lemma 4.6, we find

$$\begin{aligned}\|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} + \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)}^{1+1/\kappa} \\ \lesssim \|p\|\,\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} + \|p\|^2 + \int_0^T \alpha\, u_p^\alpha(t)^\top (u^*(t) - u_p^\alpha)\,\mathrm{d}t\end{aligned}$$

It remains to estimate the term involving the integral. We find

$$\begin{aligned}\int_0^T u_p^\alpha(t)^\top (u^*(t) - u_p^\alpha)\,\mathrm{d}t &= -\,\|u_p^\alpha - u^*\|_{L^2(I;\mathbb{R}^m)}^2 + \int_0^T u^*(t)^\top (u^*(t) - u_p^\alpha)\,\mathrm{d}t \\ &\le -\,\|u_p^\alpha - u^*\|_{L^2(I;\mathbb{R}^m)}^2 + \|u_b\|_{L^\infty(I)}\,\|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)} \\ &\quad + \|u_b^{-1}\,(u^*)^\top (u_p^\alpha - u^*)\|_{L^1(I^+)}.\end{aligned}$$

This yields

$$\begin{aligned}&\|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} + \|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)}^{1+1/\kappa} \\ &\quad\lesssim \|p\|\,\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} + \|p\|^2 + \alpha\,\|u_b\|_{L^\infty(I)}\,\|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)} \\ &\qquad + \alpha\,\|u_b^{-1}\,(u^*)^\top (u_p^\alpha - u^*)\|_{L^1(I^+)}.\end{aligned}$$

In case $\kappa = \infty$, the third and fourth term on the right-hand side become smaller than the corresponding terms on the left-hand side if α is small enough, and the assertion follows. In case $\kappa \neq \infty$, we can use Young's inequality for the third and fourth term on the right-hand side and obtain

$$\begin{aligned}&\|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} + \|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)}^{1+1/\kappa} \\ &\quad\lesssim \|p\|\,\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} + \|p\|^2 + \alpha^{1+\kappa}.\end{aligned}$$

$\square$

Note that we only have an estimate for the normal component of the error $u_p^\alpha - u^*$ on I^+. This is caused by the fact that on the set $\{t \in I : |\sigma_p^\alpha|_{\mathbb{R}^m} > \beta + \alpha\, u_b(t)\}$ we only know $|u_p^\alpha(t)|_{\mathbb{R}^m} = u_b(t)$, but the direction of $u_p^\alpha(t)$ still depends on σ_p^α. This is a crucial difference to the problem with usual sparsity and box constraints, in which the control takes the value of the lower or upper bound if the absolute value of the switching function is large enough.

Hence, we have to derive a different estimate for the tangential component of the error $u_p^\alpha - u^*$ on I^+.

Lemma 4.8. *We have*

$$\|u_p^\alpha - u^*\|_{L^1(I^+;\mathbb{R}^m)} \le \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)} + \frac{\|u_b\|_{L^\infty(I)}\,\|\sigma_p^\alpha - \sigma^*\|_{L^1(I^+;\mathbb{R}^m)}}{\operatorname{ess\,inf}_{t\in I^+}|\sigma_p^\alpha(t)|_{\mathbb{R}^m}}.$$

In particular, we obtain

$$\|u_p^\alpha - u^*\|_{L^1(I^+;\mathbb{R}^m)} \le \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)} + \frac{2\,\|u_b\|_{L^\infty(I)}}{\beta}\,\|\sigma_p^\alpha - \sigma^*\|_{L^1(I^+;\mathbb{R}^m)}.$$

if $|\sigma^*(t) - \sigma_p^\alpha(t)|_{\mathbb{R}^m} \le \beta/2$ *for a.a.* $t \in I^+$.

Proof. On I^+ we have $|u^*(t)|_{\mathbb{R}^m} = u_b(t) > 0$. Hence, we can decompose the error $u_p^\alpha - u^*$ into the normal direction u^* and the tangent direction orthogonal to u^*. This yields

$$\begin{aligned}u_p^\alpha(t) - u^*(t) &= \frac{u^*(t)\,u^*(t)^\top}{u_b(t)^2}\,(u_p^\alpha(t) - u^*(t)) + \left(I - \frac{u^*(t)\,u^*(t)^\top}{u_b(t)^2}\right)(u_p^\alpha(t) - u^*(t)) \\ &= \frac{u^*(t)\,u^*(t)^\top}{u_b(t)^2}\,(u_p^\alpha(t) - u^*(t)) + \left(I - \frac{u^*(t)\,u^*(t)^\top}{u_b(t)^2}\right)u_p^\alpha(t).\end{aligned}$$

For the first term, we can use

$$\int_{I^+} \Big| \frac{u^*(t)\, u^*(t)^\top}{u_b(t)^2} \left(u_p^\alpha(t) - u^*(t)\right)\Big|_{\mathbb{R}^m} \,\mathrm{d}t = \int_{I^+} \Big| \frac{u^*(t)^\top}{u_b(t)} \left(u_p^\alpha(t) - u^*(t)\right)\Big| \,\mathrm{d}t$$
$$= \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)}.$$

By optimality of u_p^α and u^*, we find

$$u^*(t) = -\frac{\sigma^*(t)}{|\sigma^*(t)|_{\mathbb{R}^m}}\, u_b(t), \qquad u_p^\alpha(t) = -\frac{\sigma_p^\alpha(t)}{|\sigma_p^\alpha(t)|_{\mathbb{R}^m}}\, u_b(t)\, \nu(t),$$

where $0 \le \nu(t) \le 1$ (this follows from (4.5a) by deriving a projection formula similar to Lem. 3.8). This implies

$$\left(I - \frac{u^*(t)\, u^*(t)^\top}{u_b(t)^2}\right) u_p^\alpha(t) = -\frac{\nu(t)\, u_b(t)}{|\sigma_p^\alpha(t)|_{\mathbb{R}^m}} \left(I - \frac{\sigma^*(t)\,\sigma^*(t)^\top}{|\sigma^*(t)|^2_{\mathbb{R}^m}}\right) \sigma_p^\alpha(t)$$
$$= -\frac{\nu(t)\, u_b(t)}{|\sigma_p^\alpha(t)|_{\mathbb{R}^m}} \left(I - \frac{\sigma^*(t)\,\sigma^*(t)^\top}{|\sigma^*(t)|^2_{\mathbb{R}^m}}\right) (\sigma_p^\alpha(t) - \sigma^*(t)).$$

Since the term in parenthesis is a projection (with norm 1), we find

$$\left|\left(I - \frac{u^*(t)\, u^*(t)^\top}{u_b(t)^2}\right) u_p^\alpha(t)\right|_{\mathbb{R}^m} \le \frac{u_b(t)}{|\sigma_p^\alpha(t)|_{\mathbb{R}^m}}\, |\sigma_p^\alpha(t) - \sigma^*(t)|_{\mathbb{R}^m}$$

and the assertion follows. □

From Lemma 4.7 and $\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} \le 2\,T\,\|u_b\|_{L^\infty(I)}$, we obtain

$$\|\sigma_p^\alpha - \sigma^*\|^2_{L^\infty(I;\mathbb{R}^m)} \lesssim \|p\| + \alpha^{1+\kappa}.$$

Hence, $|\sigma_p^\alpha|_{\mathbb{R}^m} \ge |\sigma^*|_{\mathbb{R}^m} - |\sigma_p^\alpha - \sigma^*|_{\mathbb{R}^m} \ge \beta/2$ on I^+ for $\alpha, \|p\|$ small enough.

Together with Lemma 4.8 we obtain the main result of this section.

Theorem 4.9. *Assume that Assumption 4.1 is satisfied. If $\alpha, \|p\|$ are sufficiently small, we obtain*

$$\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} \lesssim (\alpha + \|p\|)^\kappa$$
$$\|u_p^\alpha - u^*\|_{L^2(I;\mathbb{R}^m)} \lesssim (\alpha + \|p\|)^{\kappa/2}$$
$$\|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} \lesssim (\alpha + \|p\|)^{(1+\kappa)/2}$$

in case $\kappa \ne \infty$ and

$$\|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} \lesssim \|p\|$$
$$\|u_p^\alpha - u^*\|_{L^\infty(I^-;\mathbb{R}^m)} = 0$$
$$\|u_p^\alpha - u^*\|_{L^\infty(I;\mathbb{R}^m)} \lesssim \|p\|$$

in case $\kappa = \infty$.

Proof. By Lemmas 4.7 and 4.8 we obtain

$$\alpha\,\|u_p^\alpha - u^*\|^2_{L^2(I;\mathbb{R}^m)} + \|u_p^\alpha - u^*\|^{1+1/\kappa}_{L^1(I^-;\mathbb{R}^m)} + \|\sigma_p^\alpha - \sigma^*\|^2_{L^\infty(I;\mathbb{R}^m)}$$
$$+ \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|^{1+1/\kappa}_{L^1(I^+)}$$
$$\lesssim \|p\| \left(\|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)} + \|u_b^{-1}\,(u^*)^\top\, u_p^\alpha - u^*\|_{L^1(I^+)} + \|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I^+;\mathbb{R}^m)}\right)$$
$$+ \|p\|^2 + \alpha^{1+\kappa}.$$

Now, we can use Young's inequality and obtain

$$\begin{aligned}\alpha \|u_p^\alpha - u^*\|_{L^2(I;\mathbb{R}^m)}^2 &+ \|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)}^2 \\ &+ \|u_b^{-1}\,(u^*)^\top\,(u_p^\alpha - u^*)\|_{L^1(I^+)}^{1+1/\kappa} \\ &\lesssim \|p\|^{1+\kappa} + \|p\|^2 + \alpha^{1+\kappa}\end{aligned}$$

for $\alpha, \|p\|$ small enough, with the convention $\alpha^{1+\kappa} = \|p\|^{1+\kappa} = 0$ in case $\kappa = \infty$. Using again Lemma 4.8 implies

$$\begin{aligned}\alpha \|u_p^\alpha - u^*\|_{L^2(I;\mathbb{R}^m)}^2 &+ \|u_p^\alpha - u^*\|_{L^1(I^-;\mathbb{R}^m)}^{1+1/\kappa} + \|\sigma_p^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)}^2 \\ &+ \|u_p^\alpha - u^*\|_{L^1(I^+;\mathbb{R}^m)}^{\max\{1+1/\kappa,2\}} \\ &\lesssim \|p\|^{1+\kappa} + \|p\|^2 + \alpha^{1+\kappa}\end{aligned}$$

and the assertion (up to the L^∞-rates for u in case $\kappa = \infty$) follows.

It remains to show the L^∞-rates for u in case $\kappa = \infty$. Since the switching function converges in L^∞, we get $|\sigma_p^\alpha(t)|_{\mathbb{R}^m} < \beta$ a.e. on I^- and $|\sigma_p^\alpha(t)|_{\mathbb{R}^m} > \beta + \alpha\, u_b$ a.e. on I^+ for $\alpha, \|p\|$ small enough. Hence, $u_p^\alpha(t) = u^*(t) = 0$ on I^- by Lemmas 3.7, 3.8. On I^+, $u^*(t)$, $u_p^\alpha(t)$ are the projections of $-\sigma^*(t)$, $-\sigma_p^\alpha(t)$ onto $U_{\mathrm{ad}}(t)$, respectively. Since the projection is Lipschitz-continuous with modulus 1, we find $\|u_p^\alpha - u^*\|_{L^\infty(I;\mathbb{R}^m)} \le \|\sigma^* - \sigma_p^\alpha\|_{L^\infty(I;\mathbb{R}^m)}$ and the assertion follows. □

Note that in case $\kappa = \infty$ we either have $I^+ = I$ or $I^- = I$, since $|\sigma^*|$ is continuous and does not take values in $[\beta - \varepsilon_0, \beta + \varepsilon_0]$. In the latter case, we even have $u^* = u_p^\alpha = 0$ for small values of $\|p\|$ and α.

5. Discretization

For given numbers $i_0, i_1 \in \mathbb{N}$, $i_0 \le i_1$, we define

$$J_{i_0}^{i_1} := \{i_0, i_0 + 1, \dots, i_1\}.$$

In order to discretize problem (**P**), let a natural number $N \ge 2$ be given. Then with the mesh size $h = T/N$, we define $t_j = jh$, $j \in J_0^N$, as grid points of the discretization. We approximate the space $L^2(I;\mathbb{R}^m)$ of controls by functions in the subspace $X_{2,N} \subset L^2(I;\mathbb{R}^m)$ of piecewise constant functions u_h, represented by their values $u_h(t_j) = u_{h,j}$ at the grid points t_j, $j \in J_0^{N-1}$. Further, we approximate state and adjoint state variables by functions x_h and λ_h in the subspace $X_{1,N} \subset W^{1,2}(I,\mathbb{R}^n)$ of continuous, piecewise linear functions, represented by their values $x_h(t_j) = x_{h,j}$, $\lambda_h(t_j) = \lambda_{h,j}$ at the grid points t_j, $j \in J_0^N$.

We only consider time-invariant control constraints in this section, *i.e.*, $u_b(t) \equiv u_b$. The Euler discretization of problem (**P**) then reads as follows

$$\begin{aligned} \text{Minimize}\quad & f_h(x_h,u_h) + \frac{\alpha}{2}\, h \sum_{j=0}^{N-1} |u_{h,j}|_{\mathbb{R}^m}^2 + \beta\, h \sum_{j=0}^{N-1} |u_{h,j}|_{\mathbb{R}^m} \\ \text{w.r.t.}\quad & x_h \in X_1\,, u_h \in X_2 \\ \text{such that}\quad & x_{h,j+1} = x_{h,j} + h\,[A(t_j)\,x_{h,j} + B(t_j)\,u_{h,j} + b(t_j)] \quad \text{for } j \in J_0^{N-1} \\ & x_{h,0} = x_0 \\ \text{and}\quad & |u_h(t_j)|_{\mathbb{R}^m} \le u_b \quad \text{for } j \in J_0^{N-1}. \end{aligned} \qquad (\mathbf{P}_h)$$

We are interested in solving problem ($\mathbf{P}_h$) for some fixed $\beta \ge 0$. The feasible set of problem ($\mathbf{P}_h$) is nonempty, convex, closed, and bounded. Therefore, the feasible set is compact. Since the discrete cost functional is continuous, a minimizer (x_h^α, u_h^α) exists.

There exists a multiplier $\lambda_h^\alpha \in X_{1,N}$ and subgradients $\mu_{h,j}^\alpha \in \partial|\cdot|_{\mathbb{R}^m}(u_{h,j})$, $j \in J_0^{N-1}$, such that the discrete adjoint equation

$$-\frac{\lambda_{h,j+1}^\alpha - \lambda_{h,j}^\alpha}{h} = A(t_j)^\top \lambda_{h,j+1}^\alpha + W(t_j)\, x_{h,j}^\alpha + w(t_j)\,, \quad j \in J_0^{N-1}\,,$$
$$\lambda_{h,N}^\alpha = Q\, x_{h,N}^\alpha + q$$

and the discrete minimum principle

$$\left[B(t_j)^\top \lambda_{h,j}^\alpha + r(t_j) + \alpha\, u_{h,j}^\alpha + \beta\, \mu_{h,j}^\alpha\right]^\top \left(u - u_{h,j}^\alpha\right) \geq 0 \quad \text{for all } u \in U_{\mathrm{ad}}$$

are fulfilled.

In order to estimate the discretization error $\|u_h^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)}$, we use the same approach as in ([12], Thms. 7, [17], 2, [3], and 13) and show that (x_h^α, u_h^α) solves a perturbed (infinite dimensional) problem ($\mathbf{P}_p^\alpha$) for some perturbation parameter p which can be estimated by $\|p\| \leq c\, h$.

We define piecewise constant function $A_h\colon I \to \mathbb{R}^{n\times n}$, $B_h\colon I \to \mathbb{R}^{n\times m}$, $b_h\colon I \to \mathbb{R}^n$, and $x_h\colon I \to \mathbb{R}^n$ by

$$A_h(t) := A(t_j)\,,\; B_h(t) := B(t_j)\,,\; b_h(t) := b(t_j)\,,\; x_h(t) := x_h^\alpha(t_j) \quad \text{for } t \in [t_j, t_{j+1}[\,,$$

for $j \in J_0^{N-1}$. Then we can write the system equation of problem ($\mathbf{P}_h$) as

$$\dot{x}_h^\alpha(t) = A_h(t)\, x_h(t) + B_h(t)\, u_h^\alpha(t) + b_h(t) = A(t)\, x_h^\alpha(t) + B(t)\, u_h^\alpha(t) + b(t) + p_{h,b}^\alpha(t)\,,$$

where

$$p_{h,b}^\alpha(t) = A_h(t)\, x_h(t) + B_h(t)\, u_h^\alpha(t) + b_h(t) - A(t)\, x_h^\alpha(t) - B(t)\, u_h^\alpha(t) - b(t). \tag{5.1}$$

Next, we define piecewise constant function $W_h\colon I \to \mathbb{R}^{n\times n}$, $w_h\colon I \to \mathbb{R}^n$, and $\lambda_h\colon I \to \mathbb{R}^n$ by

$$W_h(t) := W(t_j)\,,\; w_h(t) := w(t_j)\,,\; \lambda_h(t) := \lambda_h^\alpha(t_{j+1}) \quad \text{for } t \in [t_j, t_{j+1}[\,,$$

for $j \in J_0^{N-1}$. Then the discrete adjoint equation can be written as

$$\begin{aligned} -\dot{\lambda}_h^\alpha(t) &= A_h(t)^\top \lambda_h(t) + W_h(t)\, x_h(t) + w_h(t) \\ &= A(t)^\top \lambda_h^\alpha(t) + W(t)\, x_h^\alpha(t) + w(t) + p_{h,w}^\alpha(t)\,, \end{aligned}$$

where

$$p_{h,w}^\alpha(t) = A_h(t)^\top \lambda_h(t) + W_h(t)\, x_h(t) + w_h(t) - A(t)^\top \lambda_h^\alpha(t) - W(t)\, x_h^\alpha(t) - w(t). \tag{5.2}$$

By defining $r_h\colon I \to \mathbb{R}^n$, μ_h^α by

$$r_h(t) := r(t_j) \quad \text{for } t \in [t_j, t_{j+1}[\,, \qquad \mu_h^\alpha(t) := \mu_{h,j}^\alpha \quad \text{for } t \in [t_j, t_{j+1}[\,,$$

for $j \in J_0^{N-1}$, we are able to write the discrete minimum principle as

$$\left[B(t)^\top \lambda_h^\alpha(t) + r(t) + \alpha\, u_h^\alpha(t) + \beta\, \mu_h^\alpha(t) + p_{h,r}^\alpha(t)\right]^\top \left(u - u_h^\alpha(t)\right) \geq 0 \quad \text{for all } u \in U(t)\,,$$

where

$$p_{h,r}^\alpha(t) = B_h(t)^\top \lambda_h(t) + r_h(t) - B(t)^\top \lambda_h^\alpha(t) - r(t). \tag{5.3}$$

Note that we have $\mu_h^\alpha \in \partial G(u_h^\alpha)$ by construction, see also Lemma 3.2. By setting $p_{h,q}^\alpha = 0$, we have shown that (x_h^α, u_h^α) is feasible for the perturbed problem ($\mathbf{P}_p^\alpha$) with $p = (p_{h,b}^\alpha, p_{h,q}^\alpha, p_{h,w}^\alpha, p_{h,r}^\alpha)$. Also (x_h^α, u_h^α), together with the multiplier λ_h^α satisfies the first order optimality conditions (4.5) of ($\mathbf{P}_p^\alpha$). Therefore, (x_h^α, u_h^α) is a minimizer of this convex optimization problem.

Following the arguments in ([3], Sect. 6), the perturbations p defined above form a bounded subset in the space $L^1(I;\mathbb{R}^n) \times \mathbb{R}^n \times L^1(I;\mathbb{R}^n) \times L^\infty(I;\mathbb{R}^m)$. Using this fact together with the Lipschitz continuity of the problem data, one can now show that (compare again [3], Sect. 6)

$$\|p\| \lesssim h.$$

Applying Theorem 4.9 now directly gives us the following convergence result for the solutions of the discretized problems.

Theorem 5.1. *Let 4.1 be satisfied. If α and h are sufficiently small, we obtain*

$$\|u_h^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} \lesssim (\alpha + h)^\kappa$$

$$\|u_h^\alpha - u^*\|_{L^2(I;\mathbb{R}^m)} \lesssim (\alpha + h)^{\kappa/2}$$

$$\|\sigma_h^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} \lesssim (\alpha + h)^{(1+\kappa)/2}$$

in case $\kappa \neq \infty$ and

$$\|\sigma_h^\alpha - \sigma^*\|_{L^\infty(I;\mathbb{R}^m)} \lesssim h$$

$$\|u_h^\alpha - u^*\|_{L^\infty(I^-;\mathbb{R}^m)} = 0$$

$$\|u_h^\alpha - u^*\|_{L^\infty(I;\mathbb{R}^m)} \lesssim h$$

in case $\kappa = \infty$.

Remark 5.2. For the numerical solution of the unregularized problem ($\mathbf{P}$) (*i.e.*, with $\alpha = 0$) we solve discretized and L^2-regularized problems. In particular, we choose $\alpha = C\,h$ for the discretized problem ($\mathbf{P}_h$) with some constant $C > 0$. Then by Theorem 5.1 we have

$$\|u_h^\alpha - u^*\|_{L^1(I;\mathbb{R}^m)} \lesssim h^\kappa$$

for $\kappa \neq \infty$

This result is confirmed by the numerical experiment in the next section.

6. Numerical examples

In this section, we consider a specific problem of type ($\mathbf{P}$). We model the motion of a two-dimensional rocket vehicle. The vehicle, which is equipped with a turnable rocket engine on top, can accelerate and move in arbitrary directions. At each point in time, the maximum thrust is bounded and the fuel consumption is proportional to the thrust. For example, let the vehicle start at $(6, 4)$ with velocity $(-3, -5)$ and we try to reach the origin $(0, 0)$ at time $T = 7$. This leads to the optimal control problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\,|x(T)|_{\mathbb{R}^2}^2 + \frac{1}{2}\,|\dot{x}(T)|_{\mathbb{R}^2}^2 + \beta \int_0^T |u(t)|_{\mathbb{R}^2}\,\mathrm{d}t \\ \text{such that} \quad & x(0) = (6,4),\ \dot{x}(0) = (-3,-5), \\ & \ddot{x}(t) = u(t) \text{ for } t \in I, \\ \text{and} \quad & |u|_{\mathbb{R}^2} \le u_b. \end{aligned}$$

In particular, we choose $\beta = 1/2$ and $u_b = 3/2$. If we transform the second-order ODE $\ddot{x} = u$ in $\mathbb{R}^2$ to a first-order ODE in $\mathbb{R}^4$, we obtain an instance of (**P**) with the problem data

$$\beta = \frac{1}{2}, \qquad A = \begin{pmatrix} 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0&0\\ 1&0\\ 0&0\\ 0&1 \end{pmatrix}, \qquad x_0 = \begin{pmatrix} 6\\ -3\\ 4\\ -5 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix},$$
$$b = 0, \qquad u_b = 1.5, \quad T = 7, \qquad W = 0, \qquad q = 0, \qquad w = 0, \quad r = 0.$$

Numerically, we checked that the solution satisfies Assumption 4.1 with $\kappa = 1$, see also the top right plot in Figure 1.

For the numerical solution, the problem is discretized using an explicit Euler scheme according to Section 5. Further, we have chosen the regularization parameter α depending on h *via* $\alpha = h/T$, *cf.* Remark 5.2.

The discretized and regularized problems were solved by a semismooth Newton method, see ([19], Thm. 3.7; (3.4)). For the reader's convenience, we briefly sketch the method applied to the continuous problem. The adaption to the discretized problem is straightforward. The starting point is the projection formula (3.8a). Moreover, the switching function σ^* is obtained by solving the system equation of problem (**P**), the adjoint equation (3.2b)–(3.2c) and using the definition of the switching function (3.2d). The function which maps $u \mapsto \sigma$ defined in this way is denoted by Φ. Then, a control u is optimal if and only if

$$F(u)(t) := u(t) + \operatorname{Proj}_{[0,u_b]}\left(\frac{|\Phi(u)(t)|_{\mathbb{R}^m} - \beta}{\alpha}\right)\frac{\Phi(u)(t)}{|\Phi(u)(t)|_{\mathbb{R}^m}} = 0 \tag{6.1}$$

holds for a.a. $t \in I$, *cf.* Lemma 3.8, in particular (3.8a). Using the arguments in ([19], Lem. 3.2), we can show that the mapping $F\colon L^2(I;\mathbb{R}^m) \to L^2(I;\mathbb{R}^m)$ defined above is Newton-differentiable. This leads to the following iteration scheme. We choose an arbitrary initial guess $u^0 \in L^2(I;\mathbb{R}^m)$. For $k \geq 0$, we solve

$$F'(u^k)\,(u^{k+1} - u^k) = -F(u^k) \tag{6.2}$$

for the next iterate $u^{k+1} \in L^2(I;\mathbb{R}^m)$. Here, the Newton-derivative $F'(u^k)$ is given by

$$\begin{aligned}(F'(u^k)\,h)(t) = h(t) &+ \operatorname{Proj}_{[0,u_b]}\left(\frac{|(\Phi u^k)(t)|_{\mathbb{R}^m} - \beta}{\alpha}\right)\frac{(\Psi\,h)(t)}{|\Phi(u^k)(t)|_{\mathbb{R}^m}} \\ &+ \left[\frac{\beta}{\alpha}\,\chi_{I^k}(t) - u_b\,\chi_{J^k}(t)\right]\frac{\big(\Phi(u^k)(t), (\Psi\,h)(t)\big)_{\mathbb{R}^m}}{|\Phi(u^k)(t)|^3}\,\Phi(u^k)(t),\end{aligned}$$

where $\Psi : L^2(I;\mathbb{R}^m) \to L^2(I;\mathbb{R}^m)$ is the (constant) derivative of the affine mapping Φ and χ_{I^k}, χ_{J^k} are the characteristic functions of the sets $I^k, J^k \subset I$ given by

$$\begin{aligned} I^k &:= \left\{t \in I : 0 < |(\Phi u^k)(t)|_{\mathbb{R}^m} - \beta < \alpha\,u_b\right\}, \\ J^k &:= \left\{t \in I : |(\Phi u^k)(t)|_{\mathbb{R}^m} - \beta \geq \alpha\,u_b\right\}. \end{aligned}$$

It can be shown that the method is locally superlinearly convergent, (*cf.* [19], Thm. 3.7).

We will also compare our solutions to the solution of the problem with conventional sparsity, that is, the non-smooth term and the control constraint in (**P**) is replaced by

$$\hat{\beta}\int_0^T \sum_{i=1}^m |u_i(t)| \qquad \text{and} \qquad |u_i(t)| \leq u_b\ \forall i = 1, \ldots, m. \tag{6.3}$$

To get somewhat similar results, we choose $\hat{\beta} = 0.7$ in this conventional setting.

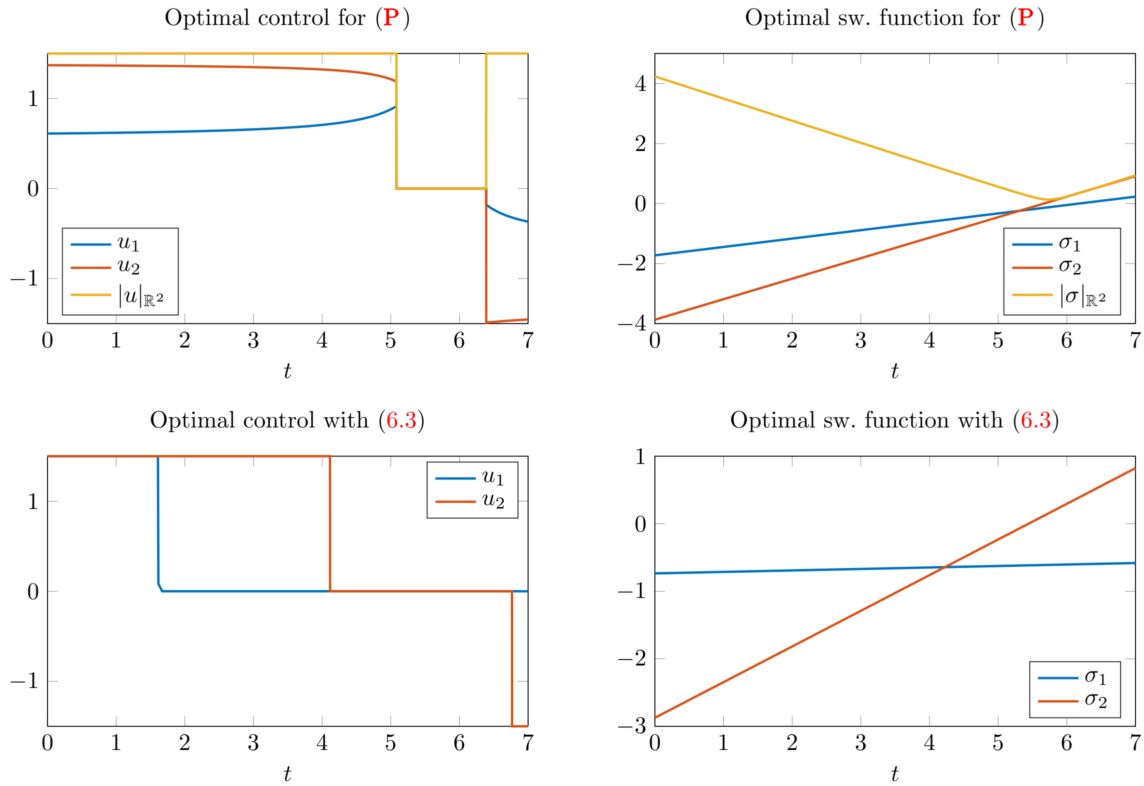

FIGURE 1. Solutions to the group sparsity problem (top row) and to a problem with conventional sparsity (bottom row).

The solution of our group sparsity approach and the conventional approach are shown in Figure 1. In this plot, we clearly see the advantages of our group sparsity approach. In fact, both controls are simultaneously zero in an interval around $t = 6$, otherwise, the vehicle uses full thrust, *i.e.*, $|u(t)|_{\mathbb{R}^2} = u_b$. In contrast, both controls act independently in the conventional setting. Moreover, in the conventional setting, the problem can be decoupled into the movement of two independent one-dimensional rocket cars moving in the x and y direction, respectively.

We show the numerical errors in Table 1 and Figures 2. In Figure 2, the dashed line corresponds to $2.764\, h^p$ with $p = 1.1185$. These results are in accordance with Theorem 5.1 which predicts convergence of order h^1, since Assumption 4.1 is satisfied with $\kappa = 1$.

For comparison with the unregularized solution, we also provide a solution of (**P**) for $\alpha = 1$ in Figure 3.

7. CONCLUSIONS AND OUTLOOK

In this paper, we used the group sparsity approach for the optimal control of an ODE. We provided stability results for the bang-bang solution and this was used to derive discretization error estimates. These estimates were confirmed by the numerical experiment.

It would be quite interesting to adopt our approach also for optimal control of PDEs, in particular in the vector-valued case. As an example, we consider the sparse optimal control of the incompressible Navier-Stokes

TABLE 1. Computational errors for the solution of (**P**). The reference value u^* is the solution on a finer temporal mesh.

h	α	$\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^2)}$
3.500000E−01	5.000000E−02	1.086318E+00
1.750000E−01	2.500000E−02	5.417342E−01
8.750000E−02	1.250000E−02	3.151054E−01
4.375000E−02	6.250000E−03	1.711812E−01
2.187500E−02	3.125000E−03	8.426044E−02
1.093750E−02	1.562500E−03	3.878615E−02
5.468750E−03	7.812500E−04	1.629570E−02
2.734375E−03	3.906250E−04	9.687963E−03
1.367187E−03	1.953125E−04	4.138973E−03
6.835937E−04	9.765625E−05	1.828597E−03
3.417969E−04	4.882812E−05	9.839766E−04
1.708984E−04	2.441406E−05	3.539314E−04

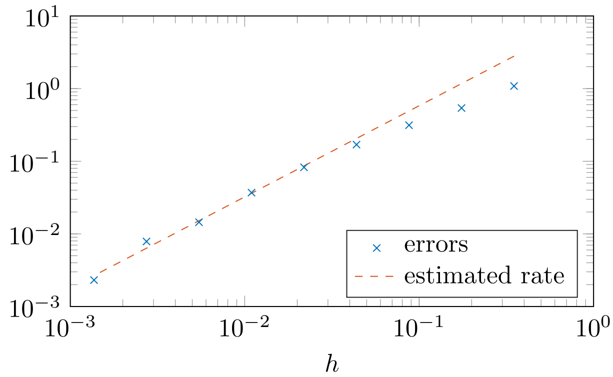

FIGURE 2. Computational errors $\|u_p^\alpha - u^*\|_{L^1(I;\mathbb{R}^2)}$ for the solution of (**P**). The reference value u^* is the solution on a finer temporal mesh.

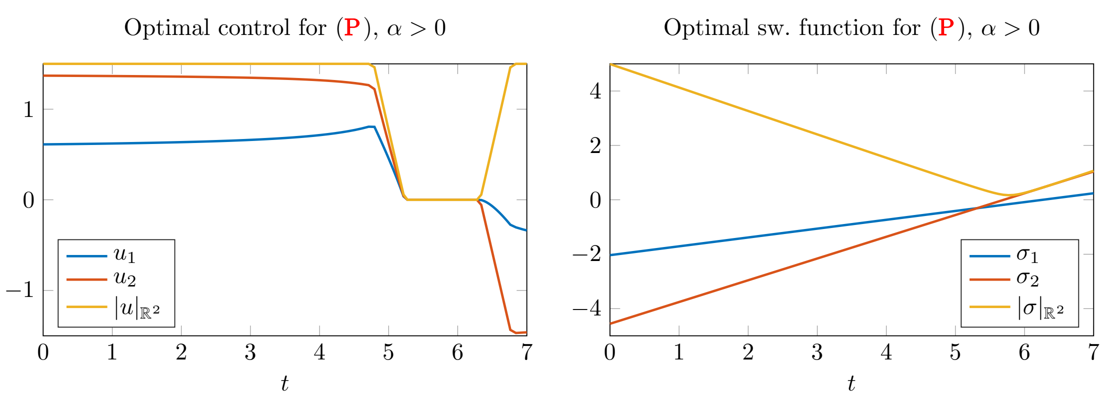

FIGURE 3. Solutions to the group sparsity problem with regularization parameter $\alpha = 1$.

equations

$$\begin{aligned} \dot{y} - \Delta y + (y \cdot \nabla)\, y + \nabla p = u & \qquad \text{in} \quad \Omega \\ \operatorname{div} y = 0 & \qquad \text{in} \quad \Omega \end{aligned}$$

complemented by suitable initial and boundary conditions, see, *e.g.*, [8]. Here, $\Omega \subset \mathbb{R}^d$ is a bounded, open set and the functions $y : I \times \Omega \to \mathbb{R}^d$, $p : I \times \Omega \to \mathbb{R}$ and $u : I \times \Omega \to \mathbb{R}^d$ have appropriate regularity. If we use an analogue to the conventional setting (6.3), we would introduce a dependence on the actual coordinate system, whereas the sparsity term $\int_{I\times\Omega} |u|_{\mathbb{R}^d} \, \mathrm{d}x \, \mathrm{d}t$ and the control constraint $|u(t,x)|_{\mathbb{R}^d} \leq u_b$ for a.a. $(t,x) \in I \times \Omega$ are independent on the chosen coordinate system. Thus, such a setting should be preferred over (6.3) and this will be subject to future research.

Acknowledgements. The authors would like to thank Daniel Wachsmuth (Universität Würzburg) and Walter Alt (FSU Jena) for fruitful discussions on the subject. We express our gratitude to the anonymous referees for valuable suggestions and comments. The first author acknowledges financial support by Carl-Zeiss-Stiftung.

References

[1] W. Alt, R. Baier, M. Gerdts and F. Lempio, Error Bounds for Euler Approximation of Linear-Quadratic Control Problems with Bang-Bang Solutions. *Numerical Algebra, Control Optimiz.* **2** (2012) 547–570.

[2] W. Alt and C. Schneider, Linear-quadratic control problems with L^1-control cost. *Optimal. Control Appl. Methods* **36** (2015) 512–534.

[3] W. Alt, C. Schneider and M. Seydenschwanz, Regularization and implicit Euler discretization of linear-quadratic optimal control problems with bang-bang solutions. *Appl. Math. Comput.* 287–288 (2016) 104–124.

[4] W. Alt and M. Seydenschwanz, Regularization and Discretization of Linear-Quadratic Control Problems. *Control and Cybernetics* **40** (2011) 903–920.

[5] W. Alt and M. Seydenschwanz, An Implicit Discretization Scheme for Linear-Quadratic Control Problems with Bang–Bang Solutions. *Optimization Methods and Software* **29** (2014) 535–560.

[6] A. Ben-Tal and J. Zowe, Second Order Optimality Conditions for the L^1-Minimization Problem. *Appl. Math. Optimiz.* **13** (1985) 45–58.

[7] P. Bühlmann and S. van de Geer, *Statistics for High-Dimensional Data: Methods, Theory and Applications.* Springer Series in Statistics. Springer (2011).

[8] E. Casas and K. Chrysafinos, Analysis of the velocity tracking control problem for the 3d evolutionary navier–stokes equations. *SIAM J. Control Optimiz.* **54** (2016) 99–128.

[9] E. Casas, R. Herzog and G. Wachsmuth, Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations. *ESAIM: COCV* **23** (2017) 263–295.

[10] K. Deckelnick and M. Hinze, A note on the approximation of elliptic control problems with bang-bang controls. *Comput. Optimiz. Appl. An Inter. J.* **51** (2012) 931–939.

[11] A.L. Dontchev and W.W. Hager, Lipschitzian Stability in Nonlinear Control and Optimization. *SIAM J. Control Optimiz.* **31** (1993) 569–603.

[12] A.L. Dontchev and V.M. Veliov, Metric Regularity under Approximations. *Control and Cybernetics* **38** (2009) 1283–1303.

[13] U. Felgenhauer, On stability of bang-bang type controls. *SIAM J. Control Optimiz.* **41** (2003) 1843–1867.

[14] U. Felgenhauer, Structural Stability Investigations of Bang-Singular-Bang Optimal Controls. *J. Optimiz. Theory Appl.* **152** (2012) 605–631.

[15] U. Felgenhauer, Discretization of semilinear bang-singular-bang control problems. *Computat. Optimiz. Appl. An International J.* **64** (2016) 295–326.

[16] T. Hastie, R. Tibshirani and M. Wainwright. *Statistical Learning with Sparsity: The Lasso and Generalizations.* CRC Press (2015).

[17] J.L. Haunschmied, A. Pietrus and V.M. Veliov, The Euler Method for Linear Control Systems Revisited. In *Proceedings of the 9th International Conference on Large-Scale Scientific Computations, Sozopol* (2013) 90–97.

[18] R. Herzog, J. Obermeier and G. Wachsmuth, Annular and sectorial sparsity in optimal control of elliptic equations. *Computational Optimization and Applications. An International Journal* **62** (2015) 157–180.

[19] R. Herzog, G. Stadler and G. Wachsmuth, Directional sparsity in optimal control of partial differential equations. *SIAM J. Control and Optimiz.* **50** (2012) 943–963.

[20] J-H.R. Kim and H. Maurer, Sensitivity Analysis of Optimal Control Problems with Bang-Bang Controls. In *Proceedings of the 42nd IEEE Conference on Decision and Control* **4** (2003) 3281–3286.

[21] N.P. Osmolovskii and H. Maurer. Equivalence of Second Order Optimality Conditions for Bang-Bang Control Problems. Part 1: Main Results. *Control and Cybernetics* **34** (2005).

[22] N.P. Osmolovskii and H. Maurer, Equivalence of Second Order Optimality Conditions for Bang-Bang Control Problems. Part 2: Proofs, Variational Derivatives and Representations. *Control and Cybernetics* **36** (2007).

[23] M. Quincampoix and V.M. Veliov, Metric Regularity and Stability of Optimal Control Problems for Linear Systems. *SIAM J. Control Optimiz.* **51** (2013) 4118–4137.

[24] C. Schneider and W. Alt, Regularization of Linear-Quadratic Control Problems with L^1-Control Cost. In C. Pötzsche, C. Heuberger, B. Kaltenbacher, and F. Rendl, System Modeling and Optimization, volume 443 of *IFIP Advances in Information and Communication Technology*, pages 296–305. 26th IFIP TC 7 Conference, CSMO 2013, Springer (2014).

[25] G. Stadler, Elliptic Optimal Control Problems with L^1-Control Cost and Applications for the Placement of Control Devices. *Comput. Optimiz. Appl.* **44** (2009) 159–181.

[26] V.M. Veliov, Error Analysis of Discrete Approximations to Bang-Bang Optimal Control Problems: The Linear Case. *Control and Cybernetics* **34** (2005) 967–982.

[27] G. Vossen and H. Maurer. On L^1-minimization in optimal control and applications to robotics. *Optimal Control Appl. Methods* **27** (2006) 301–321.

[28] D. Wachsmuth and G. Wachsmuth, Regularization error estimates and discrepancy principle for optimal control problems with inequality constraints. *Control and Cybernetics* **40** (2011) 1125–1158.

[29] G. Wachsmuth and D. Wachsmuth, Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM: COCV* **17** (2011) 858–886.