

# Convergence results for the discrete regularization of linear-quadratic control problems with bang–bang solutions

Martin Seydenschwanz

Received: 10 September 2013 / Published online: 8 February 2015  
© Springer Science+Business Media New York 2015

**Abstract** We analyze a combined regularization–discretization approach for a class of linear-quadratic optimal control problems. By choosing the regularization parameter  $\alpha$  with respect to the mesh size  $h$  of the discretization we approximate the optimal bang–bang control. Under weaker assumptions on the structure of the switching function we generalize existing convergence results and prove error estimates of order  $\mathcal{O}(h^{1/(k+1)})$  with respect to the controllability index  $k$ .

**Keywords** Linear-quadratic optimal control · Bang–bang control · Regularization · Discretization

**Mathematics Subject Classification** 49J15 · 49M25 · 49N10 · 49J30

## 1 Introduction

We consider the following linear-quadratic control problem:

$$\begin{aligned} \text{(OS)} \quad & \min_{(x,u) \in X} f(x, u) \\ \text{s.t.} \quad & \dot{x}(t) = A(t)x(t) + B(t)u(t) \quad \text{a.e. on } [0, T], \\ & x(0) = a, \\ & u(t) \in U \quad \text{a.e. on } [0, T], \end{aligned}$$

where the cost functional  $f$  is defined by

---

M. Seydenschwanz (✉)  
Siemens AG, Corporate Technology – Research and Technology Center, 81739 Munich, Germany  
e-mail: martin.seydenschwanz@siemens.com

$$f(x, u) := \frac{1}{2}x(T)^\top Qx(T) + q^\top x(T) + \int_0^T \frac{1}{2}x(t)^\top W(t)x(t) + x(t)^\top S(t)u(t) + w(t)^\top x(t) + r(t)^\top u(t) dt.$$

Here,  $u(t) \in \mathbb{R}^m$  is the control, and  $x(t) \in \mathbb{R}^n$  is the state of the system at time  $t$ . Further  $Q \in \mathbb{R}^{n \times n}$ ,  $q \in \mathbb{R}^n$ . The functions  $W: [0, T] \rightarrow \mathbb{R}^{n \times n}$ ,  $S: [0, T] \rightarrow \mathbb{R}^{n \times m}$ ,  $w: [0, T] \rightarrow \mathbb{R}^n$ ,  $r: [0, T] \rightarrow \mathbb{R}^m$ ,  $A: [0, T] \rightarrow \mathbb{R}^{n \times n}$  and  $B: [0, T] \rightarrow \mathbb{R}^{n \times m}$  are assumed to be Lipschitz continuous. The set  $U \subset \mathbb{R}^m$  is defined by upper and lower bounds, i.e.,

$$U := \{u \in \mathbb{R}^m \mid b_l \leq u \leq b_u\}$$

with  $b_l, b_u \in \mathbb{R}^m$ ,  $b_l < b_u$ .

Numerical solution methods for optimal control problems have been investigated over the last decades. Most of the research has been dealing with shooting and direct approximation approaches. Discretizations based on Euler's method or more general Runge-Kutta methods are well-studied for the case that the optimal control is at least Lipschitz continuous (see e.g. [2, 8–12, 23, 25, 35]). First results on the error analysis for bang–bang controls can be found in [36]. In [6, 7] Euler discretizations for a class of linear-quadratic control problems with bang–bang solutions have been investigated. These results have been extended to a stable implicit discretization scheme in [5].

Since regularization leads to problems with smoother solutions a combined regularization-discretization approach is a good alternative to direct approximation. Therefore the regularization of the cost functional and constraints of optimal control problems have been studied over the last years (see e.g. [21, 26–28, 34]). The dependency of solutions on regularization parameters combined with discretization has been investigated in [18] for multiplier methods for optimal control problems governed by ODEs, and in [22] for elliptic problems with state constraints. First results on bang–bang controls of linear-quadratic problems without mixed state-control-term in the cost functional have been presented in [4]. The authors proved error estimates of order  $\mathcal{O}(\sqrt{h})$ , if the regularization parameter  $\alpha = \sqrt{h}$  is chosen w.r.t. the mesh size  $h$  of the discretization. But by treating the regularization and discretization separately and combining the error estimates via the triangle inequality the actual error has been overestimated. This is a known issue in the analysis of combined regularization–discretization approaches (see e.g. [18]). In [33] the error estimates from [4] have been improved to order  $\mathcal{O}(h)$  by combining proof techniques from [6, 7] and [4].

In this paper we will extend these convergence results on the discrete regularization of linear-quadratic problems with bang–bang controls under weaker assumptions of order  $k$  on the structure of the switching function (see Remarks 2 and 3). In our main result Theorem 5.1 we prove error estimates of order  $\mathcal{O}((\alpha + h)^{1/(k+1)})$  w.r.t. the mesh size  $h$  of the discretization and the regularization parameter  $\alpha$  for the control, state and adjoint state variables.

The organization of the paper is as follows. In Sect. 1 we give some preliminaries and known auxiliary results for the linear-quadratic control problem (OS). Moreover we generalize the second-order condition, which has been an important tool of the

analysis in [6, Lemma 4.1, Theorem 4.2], under weaker assumptions on the structure of the switching function to order  $k + 1$ . In Sect. 2 we present the regularization technique and prove estimates for the regularization error of order  $k$ . The combined regularization–discretization approach is introduced in Sect. 3. Section 4 is concerned with error estimates for the optimal values of the discrete control problems. The main result is then derived in Sect. 5. We present Hölder-type error estimates of order  $\mathcal{O}(h^{1/(k+1)})$  for control, state and adjoint state. In Sect. 6 we improve these error estimates for linear problems. Finally the theoretical findings are illustrated by a numerical example.

We use the following notation (cf. [4–7]):  $\mathbb{R}^n$  is the  $n$ -dimensional Euclidean space with the inner product denoted by  $\langle x, y \rangle$  and the norm  $|x| := \langle x, x \rangle^{1/2}$ . For an  $m \times n$ -matrix  $B$  we use the spectral norm  $\|B\| := \sup_{|z| \leq 1} |Bz|$ . By  $L^p(0, T; \mathbb{R}^m)$  we denote the Banach space of measurable vector functions  $u: [0, T] \rightarrow \mathbb{R}^m$  for  $1 \leq p < \infty$ , with

$$\|u\|_p := \left( \int_0^T |u(t)|^p dt \right)^{1/p} < \infty,$$

and  $L^\infty(0, T; \mathbb{R}^m)$  is the Banach space of essentially bounded vector functions  $u: [0, T] \rightarrow \mathbb{R}^m$  with the norm

$$\|u\|_\infty := \max_{1 \leq i \leq m} \operatorname{ess\,sup}_{t \in [0, T]} |u_i(t)|.$$

A function  $u$  will be said to be of *bounded variation* if the *total variation*  $V_0^T u$  of  $u$  on  $[0, T]$  is finite. For  $1 \leq p \leq \infty$  we denote by  $W_p^1(0, T; \mathbb{R}^n)$  the Sobolev spaces of absolutely continuous functions  $x: [0, T] \rightarrow \mathbb{R}^n$

$$W_p^1(0, T; \mathbb{R}^n) := \{x \in L^p(0, T; \mathbb{R}^n) \mid \dot{x} \in L^p(0, T; \mathbb{R}^n)\}$$

with

$$\|x\|_{1,p} := (|x(0)|^p + \|\dot{x}\|_p^p)^{1/p}$$

for  $1 \leq p < \infty$  and

$$\|x\|_{1,\infty} := \max \{|x(0)|, \|\dot{x}\|_\infty\}.$$

We define  $X := X_1 \times X_2$ ,  $X_1 := W_\infty^1(0, T; \mathbb{R}^n)$ ,  $X_2 := L^\infty(0, T; \mathbb{R}^m)$ , and we denote by

$$\mathcal{U} := \{u \in X_2 \mid u(t) \in U \text{ a.e. on } [0, T]\} = \{u \in X_2 \mid b_l \leq u(t) \leq b_u \text{ a.e. on } [0, T]\}$$

the set of *admissible controls*, and by

$$\mathcal{F} := \{(x, u) \in X \mid u \in \mathcal{U}, \dot{x}(t) = A(t)x(t) + B(t)u(t) \text{ a.e. on } [0, T], x(0) = a\}$$

the *feasible set* of (OS). The linear ODE

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \quad \text{a.e. on } [0, T], \\ x(0) &= a, \end{aligned} \tag{1}$$

will be called *system equation of* (OS). It is well known, that for the solution  $x$  of (1) it holds

$$\|x\|_\infty \leq c_1|a| + c_2\|u\|_1 \tag{2}$$

with constants  $c_1$  and  $c_2$  independent of  $a$  and  $u$ .

**Definition 1.1** A pair  $(x^0, u^0) \in \mathcal{F}$  is called a *minimizer* for the Problem (OS) if  $f(x^0, u^0) \leq f(x, u)$  for all  $(x, u) \in \mathcal{F}$ , and a *strict minimizer* if  $f(x^0, u^0) < f(x, u)$  for all  $(x, u) \in \mathcal{F}, (x, u) \neq (x^0, u^0)$ .

In view of convexity of the Problem (OS) we make the following assumption throughout the paper:

(AC) Let the matrices  $Q$  and  $W(t), t \in [0, T]$ , be symmetric and

$$z(T)^T Q z(T) + \int_0^T z(t)^T W(t) z(t) + 2z(t)^T S(t) v(t) dt \geq 0$$

for all  $(z, v) \in \mathcal{F} - \mathcal{F}$ , i. e.  $(z, v) \in X$  with

$$\dot{z}(t) = A(t)z(t) + B(t)v(t) \quad \text{a.e. on } [0, T],$$

$$z(0) = 0,$$

$$v(t) \in U - U \quad \text{a.e. on } [0, T].$$

The following auxiliary results concerning linear-quadratic problems of type (OS) are common knowledge and can be found in [6]. The feasible set  $\mathcal{F}$  is nonempty, closed, convex and bounded. If (AC) holds, then the cost functional is convex and continuous on  $\mathcal{F}$ . Therefore, a minimizer  $(x^0, u^0) \in W_2^1(0, T; \mathbb{R}^n) \times L^2(0, T; \mathbb{R}^m)$  of (OS) exists (cf. [13, Chap. II, Proposition 1.2]), and since  $\mathcal{U}$  is bounded we have  $(x^0, u^0) \in X = W_\infty^1(0, T; \mathbb{R}^n) \times L^\infty(0, T; \mathbb{R}^m)$ . Moreover, the cost functional is Lipschitz continuous on  $\mathcal{F}$ , i.e. there is a constant  $L_f$  such that

$$|f(x, u) - f(z, v)| \leq L_f (\|x - z\|_\infty + \|u - v\|_1) \quad \forall (x, u), (z, v) \in \mathcal{F}.$$

An immediate consequence of the compactness of  $U$ , the Lipschitz continuity of  $A$  and  $B$  as well as the solution formula for linear differential equations, is the existence of a constant  $L_x$  such that for any admissible control  $u \in \mathcal{U}$  and the associated solution  $x$  of the system Eq. (1) we have

$$\|x\|_{1,\infty} \leq L_x, \tag{3}$$

where the constant  $L_x$  is independent of  $x$ . This estimate shows that the feasible trajectories are uniformly Lipschitz with Lipschitz modulus  $L_x$ .

Let  $(x^0, u^0) \in \mathcal{F}$  be a minimizer of (OS). Then there exists an adjoint variable  $\lambda^0 \in W_\infty^1(0, T; \mathbb{R}^n)$  such that the *adjoint equation*

$$\begin{aligned} -\dot{\lambda}^0(t) &= A(t)^\top \lambda^0(t) + W(t)x^0(t) + S(t)u^0(t) + w(t) \quad \text{a.e. on } [0, T], \\ \lambda^0(T) &= Qx^0(T) + q, \end{aligned} \tag{4}$$

and the *minimum principle*

$$\left[ B(t)^\top \lambda^0(t) + S(t)^\top x^0(t) + r(t) \right]^\top (u - u^0(t)) \geq 0 \quad \forall u \in U, \quad \text{a.e. on } [0, T], \tag{5}$$

hold. Denoting the *switching function* by

$$\sigma^0(t) := B(t)^\top \lambda^0(t) + S(t)^\top x^0(t) + r(t), \tag{6}$$

it is well-known that for each  $i \in \{1, \dots, m\}$ , the minimum principle (5) implies

$$u_j^0(t) = \begin{cases} b_{l,j}, & \text{if } \sigma_j^0(t) > 0, \\ b_{u,j}, & \text{if } \sigma_j^0(t) < 0, \\ \text{undetermined,} & \text{if } \sigma_j^0(t) = 0. \end{cases} \tag{7}$$

*Remark 1* Since  $\lambda$  satisfies the adjoint equation, the parameter functions  $A, W, S$  and  $w$  are Lipschitz continuous, and  $u$  is bounded it follows that  $\dot{\lambda}$  is bounded, i.e. there exists a constant  $L_\lambda$ , independent of  $\lambda^0$ , such that for any feasible pair  $(x, u) \in \mathcal{F}$  and the associated solution  $\lambda$  of the adjoint equation we have

$$\|\lambda\|_{1,\infty} \leq L_\lambda.$$

Hence  $\lambda$  is uniformly Lipschitz continuous with Lipschitz modulus  $L_\lambda$  independent of  $N$ , which implies that the switching function  $\sigma$  is uniformly Lipschitz continuous, too. Analogously to (2) for a solution  $\lambda^0$  of (4) it holds

$$\|\lambda^0\|_\infty \leq c_1 \|x^0\|_\infty + c_2 \|u^0\|_1 + c_3 \|w\|_\infty + c_4 |q| \tag{8}$$

with constants  $c_1, c_2, c_3$  and  $c_4$  independent of  $(x^0, u^0), w$  and  $q$ .

In the case of Lipschitz continuous optimal controls the convergence analysis of discretization methods is usually based on a second-order optimality condition (see e.g. [10,23]). For bang–bang controls such conditions are available, too (see e.g. [1, 14–17, 24, 29–31]). In [6] a second-order condition which has been introduced in [14] turned out to be very useful for the analysis of Euler discretization for linear-quadratic problems. We will extend the approach used in [6, Sect. 4] and [4] to the class of problems (OS) with a mixed state-control term in the cost functional. Under weaker assumptions on the structure of the switching function we derive a condition of order  $k + 1$ , which will play an important role in the analysis of convergence for

discrete regularization approaches. This condition is closely connected to the *controllability index* in [19,32] (see Remark 3). We make the following assumption on *bang–bang regularity* (cf. [1] resp. [14,15]).

- (A1) There exists a solution  $(x^0, u^0) \in \mathcal{F}$  of (OS) such that the set  $\Sigma$  of zeros of the components  $\sigma_j^0, j = 1, \dots, m$ , of the switching function  $\sigma^0$  defined by (6) is finite and  $0, T \notin \Sigma$ , i.e.  $\Sigma = \{s_1, \dots, s_l\}$  with  $0 < s_1 < \dots < s_l < T$ .

Assumption (A1) implies bounded variation of  $u^0$ . We denote the set of active indices for the components of the switching function  $\sigma^0$  by

$$\mathcal{I}(s_i) := \{1 \leq j \leq m : \sigma_j^0(s_i) = 0\},$$

and formulate the second assumption:

- (A2)<sup>k</sup> There is a natural number  $\bar{k} \in \mathbb{N}$ , for which there exist constants  $\bar{\sigma}, \bar{\tau} > 0$ , such that for all  $i \in \{1, \dots, l\}, j \in \mathcal{I}(s_i)$  and all  $t \in [s_i - \bar{\tau}, s_i + \bar{\tau}]$  it holds

$$|\sigma_j(t)| \geq \bar{\sigma} |t - s_i|^{\bar{k}}.$$

We define  $k \in \mathbb{N}$  as the smallest natural number that fulfills this condition.

*Remark 2* The Assumption (A2)<sup>k</sup> is a weaker condition than Assumption (A2) in [4,6] (since it holds (A2)<sup>1</sup> = (A2)). In dependence of the parameter  $k \in \mathbb{N}$  we will derive generalized convergence results for the discrete regularization approach for problems of type (OS). Under the Assumptions (A1) and (A2)<sup>k</sup> we will prove Hölder-type error estimates of order  $\mathcal{O}(h^{1/(k+1)})$  for control, state and adjoint state.

*Example 1.2* We introduce a class of linear problems depending on a parameter  $k \in \mathbb{N}$ , which fulfills the Assumptions (A1) and (A2)<sup>k</sup> (cf. [7, Example 2.10], [19, Sect. 4]). In the context of the labeling in [32] these problems have a *controllability index*  $k \in \mathbb{N}$ . With  $n = k + 1, s \in \mathbb{R}^k$  and  $X = W^1_\infty(0, 1; \mathbb{R}^n) \times L^\infty(0, 1; \mathbb{R})$  we define

$$\begin{aligned} \text{(B)}_k \quad & \min_{(x,u) \in X} && x_1(1) \\ & \text{s.t.} && \dot{x}_j(t) = s_j x_{j+1}(t) + u(t) && \text{a.e. on } [0, 1], \quad j = 1, \dots, k, \\ & && \dot{x}_{k+1}(t) = u(t) && \text{a.e. on } [0, 1], \\ & && x(0) = 0_n, \\ & && -1 \leq u(t) \leq 1 && \text{a.e. on } [0, 1]. \end{aligned}$$

The adjoint equation can be written as

$$\begin{aligned} -\dot{\lambda}(t) &= (0, s_1 \lambda_1(t), s_2 \lambda_2(t), \dots, s_k \lambda_k(t))^T, \\ \lambda(1) &= (1, 0, \dots, 0)^T. \end{aligned}$$

For  $j = 1, \dots, k + 1$  we obtain

$$\lambda_j^0(t) = \left( \frac{(-1)^{j+1}}{(j-1)!} \prod_{i=1}^{j-1} s_i \right) (t-1)^{j-1}$$

and the corresponding switching function

$$\sigma^0(t) = \sum_{j=1}^{k+1} \lambda_j^0(t).$$

In order to guarantee that for  $k \in \mathbb{N}$  the Assumption (A2)<sup>k</sup> is fulfilled, we choose the parameter  $s$  such that the switching function  $\sigma$  has a zero of order  $k$  at  $t = 0.5$ , which means that  $\sigma(0.5) = 0$  and all derivatives up to order  $k - 1$  vanish at  $t = 0.5$ , too. These requirements can be written as

$$\sum_{j=1}^k \frac{(-1)^{k+j-\ell}}{(k-1-\ell)! 2^{k-1-\ell}} \prod_{i=1}^{j-1} s_i = 0, \quad \ell = 0, \dots, k-1$$

and are fulfilled if we choose  $s_j := -2(k - j + 1)$  for  $j = 1, \dots, k$ . The resulting switching function is

$$\sigma^0(t) = 2^k(t - 0.5)^k.$$

Therefore the solution of (B)<sub>k</sub> fulfills the Assumptions (A1) and (A2)<sup>k</sup> with  $k \in \mathbb{N}$ . With the help of (7) we can characterize the optimal control. For odd  $k$  it holds

$$u^0(t) = \begin{cases} 1, & 0 \leq t < 0.5, \\ -1, & 0.5 < t \leq 1. \end{cases}$$

If  $k$  is even, the optimal control is a constant function  $u^0(t) \equiv -1$ . Because of the weakened Assumption (A2)<sup>k</sup> this example shows, that the error estimates we will derive in Sects. 5 and 6 hold true for an extended class of control problems in comparison to the results in [4].

The following result is a generalization of [6, Lemma 4.1] (cf. [14, Lemma 3.3]):

**Lemma 1.3** *Let  $(x^0, u^0)$  be a minimizer for Problem (OS) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>, and let the switching function  $\sigma$  be defined by (6). Then there are constants  $\beta, \gamma, \bar{\delta} > 0$  such that for any admissible control  $u \in \mathcal{U}$ ,*

$$\int_0^T \sigma^0(t)^\top (u(t) - u^0(t)) \, dt \geq \beta \|u - u^0\|_1^{k+1}$$

if  $\|u - u^0\|_1 \leq 2\gamma\bar{\delta}$ , and

$$\int_0^T \sigma^0(t)^\top (u(t) - u^0(t)) \, dt \geq \beta \|u - u^0\|_1$$

if  $\|u - u^0\|_1 > 2\gamma\bar{\delta}$ .

*Proof* For  $0 < \delta < \bar{\tau}$  we define

$$I(\delta) := \bigcup_{1 \leq i \leq l} [s_i - \delta, s_i + \delta].$$

Let  $j \in \{1, \dots, m\}$  be arbitrary and

$$\Sigma_j := \{\tau_1, \dots, \tau_{l_j}\} \text{ with } 0 < \tau_1 < \dots < \tau_{l_j} < T$$

the set of zeros of  $\sigma_j$ . Moreover we define

$$I_-(\delta) := \bigcup_{\iota=1, \dots, l_j} [\tau_\iota - \delta, \tau_\iota + \delta] \text{ and } I_+(\delta) := [0, T] \setminus I_-(\delta).$$

Since  $\sigma$  is Lipschitz continuous, there exist constants  $\sigma_{j,\min}$  with

$$\sigma_{j,\min} := \min_{t \in I_+(\bar{\tau})} |\sigma_j(t)| > 0.$$

We choose  $0 < \bar{\delta} \leq \bar{\tau}$ , such that

$$\bar{\delta}^k \bar{\sigma} \leq \min_{1 \leq j \leq m} \sigma_{j,\min}.$$

From the Assumption (A2)<sup>k</sup> we obtain for all  $0 < \delta < \bar{\delta}$  and  $j \in \{1, \dots, m\}$

$$|\sigma_j(t)| \geq \delta^k \bar{\sigma} \quad \forall t \in [0, T] \setminus I(\delta). \tag{9}$$

Let  $u \in \mathcal{U}$  be arbitrary. From the minimum principle (5) we know that the signs of  $\sigma_j(t)$  and  $u_j(t) - u_j^0(t)$  coincide a.e. on  $[0, T]$ . It follows from (9), that

$$\begin{aligned} J &= \int_0^T \sigma^0(t)^\top (u(t) - u^0(t)) \, dt \geq \int_{[0, T] \setminus I(\delta)} \sigma^0(t)^\top (u(t) - u^0(t)) \, dt \\ &= \int_{[0, T] \setminus I(\delta)} \sum_{j=1}^m |\sigma_j(t)| |u_j(t) - u_j^0(t)| \, dt \geq \delta^k \bar{\sigma} \sum_{j=1}^m \int_{[0, T] \setminus I(\delta)} |u_j(t) - u_j^0(t)| \, dt. \end{aligned} \tag{10}$$

Furthermore, for  $1 \leq j \leq m$  it holds

$$|u_j(t) - u_j^0(t)| \leq b_{u,j} - b_{l,j} \quad \forall t \in [0, T],$$

an with

$$\gamma := 2lm \max_{1 \leq j \leq m} (b_{u,j} - b_{l,j})$$



we get

$$\sum_{j=1}^m \int_{I(\delta)} |u_j(t) - u_j^0(t)| dt \leq \gamma \delta.$$

Together with (10) this implies

$$J \geq \delta^k \bar{\sigma} \left( \|u - u^0\|_1 - \gamma \delta \right).$$

Now we choose  $\delta := \min \left\{ \bar{\delta}, \frac{1}{2\gamma} \|u - u^0\|_1 \right\}$ . If  $\delta = \bar{\delta}$ , i.e. it holds  $\|u - u^0\|_1 > 2\gamma \bar{\delta}$ , we obtain

$$J \geq \frac{\bar{\delta}^k \bar{\sigma}}{2} \|u - u^0\|_1. \tag{11}$$

If  $\delta = \frac{1}{2\gamma} \|u - u^0\|_1$ , i.e. it holds  $\|u - u^0\|_1 \leq 2\gamma \bar{\delta}$ , we have

$$J \geq \frac{\bar{\sigma}}{2^{k+1} \gamma^k} \|u - u^0\|_1^{k+1}. \tag{12}$$

With

$$\beta := \min \left\{ \frac{\bar{\delta}^k \bar{\sigma}}{2}, \frac{\bar{\sigma}}{2^{k+1} \gamma^k} \right\}$$

the assertion follows directly from (11) and (12). □

This Lemma implies a minorant of order  $k + 1$  for the optimal values of (OS) in a sufficiently small  $L^1$ -neighborhood and a linear minorant outside this neighborhood. Moreover it directly implies the following generalization of [6, Theorem 4.2]:

**Theorem 1.4** *Let  $(x^0, u^0)$  be a minimizer for Problem (OS) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>. Then there are constants  $\beta, \gamma, \bar{\delta} > 0$  such that for any feasible pair  $(x, u) \in \mathcal{F}$ ,*

$$f(x, u) - f(x^0, u^0) \geq \beta \|u - u^0\|_1^{k+1}$$

if  $\|u - u^0\|_1 \leq 2\gamma \bar{\delta}$ , and

$$f(x, u) - f(x^0, u^0) \geq \beta \|u - u^0\|_1$$

if  $\|u - u^0\|_1 > 2\gamma \bar{\delta}$ .

*Proof* Let  $(x^0, u^0)$  be a solution of (OS) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>, with corresponding adjoint variable  $\lambda^0$ . Further let  $(x, u) \in \mathcal{F}$  be arbitrary. We define  $z := x - x^0, v := u - u^0$  and obtain

$$\begin{aligned} f(x, u) - f(x^0, u^0) &= x^0(T)^\top Qz(T) + \frac{1}{2}z(T)^\top Qz(T) + q^\top z(T) \\ &\quad + \int_0^T \frac{1}{2}z(t)^\top W(t)z(t) + z(t)^\top S(t)v(t) \, dt \\ &\quad + \int_0^T [W(t)x^0(t) + S(t)u^0(t) + w(t)]^\top z(t) \, dt \\ &\quad + \int_0^T [S(t)^\top x^0(t) + r(t)]^\top v(t) \, dt. \end{aligned}$$

With Assumption (AC) it follows

$$\begin{aligned} f(x, u) - f(x^0, u^0) &\geq x^0(T)^\top Qz(T) + q^\top z(T) \\ &\quad + \int_0^T [W(t)x^0(t) + S(t)u^0(t) + w(t)]^\top z(t) \, dt \\ &\quad + \int_0^T [S(t)^\top x^0(t) + r(t)]^\top v(t) \, dt. \end{aligned} \tag{13}$$

Using partial integration we can deduce from the terminal condition of the adjoint Eq. (4)

$$\begin{aligned} x^0(T)^\top Qz(T) + q^\top z(T) &= \lambda^0(T)^\top z(T) \\ &= \int_0^T \dot{z}(t)^\top \lambda^0(t) \, dt + \int_0^T z(t)^\top \dot{\lambda}^0(t) \, dt. \end{aligned}$$

Since  $z$  solves the system Eq. (1) of (OS) with  $u = v$  and  $\lambda^0$  solves the adjoint Eq. (4) with  $(x, u) = (x^0, u^0)$  we obtain

$$\begin{aligned} x^0(T)^\top Qz(T) + q^\top z(T) &= \int_0^T [A(t)z(t) + B(t)v(t)]^\top \lambda^0(t) \, dt \\ &\quad - \int_0^T z(t)^\top [A(t)^\top \lambda^0(t) + W(t)x^0(t) + S(t)u^0(t) + w(t)] \, dt. \end{aligned}$$

Plugging this into (13) leads together with (6) to

$$\begin{aligned} f(x, u) - f(x^0, u^0) &\geq \int_0^T [B(t)^\top \lambda^0(t) + S(t)^\top x^0(t) + r(t)]^\top v(t) \, dt \\ &= \int_0^T \sigma^0(t)^\top v(t) \, dt, \end{aligned}$$

and the assertion follows directly from Lemma 1.3. □

*Remark 3* The results of order  $k + 1$  from Lemma 1.3 and Theorem 1.4 are strongly connected to the findings in [19] and [32]. In [32] the authors derived stability results for Mayer-type problems with the help of *metric regularity* and smoothness assumptions on the problem parameters. The structure of the switching function is characterized with the help of the “*controllability index*”. In the case of control constraints defined by constant upper and lower bounds

$$u(t) \in U = \{u \in \mathbb{R}^m \mid b_l \leq u \leq b_u\} \quad \text{a.e. on } [0, T]$$

the controllability index depends on the derivatives up to order  $k - 1$  of the switching function (cf. [32, Remark 1]).

Using the stability results from [32] it was shown in [19] that the direct Euler discretization for Mayer-type problems converges with order  $\mathcal{O}(h^{1/k})$  w.r.t. the controllability index  $k$  and the mesh size  $h$ . Under the Assumptions (A1) and (A2)<sup>k</sup> we will show convergence of order  $\mathcal{O}(h^{1/(k+1)})$  for a discrete regularization approach applied to linear-quadratic problems of type (OS) with bang–bang controls. Moreover we will improve these Hölder-type error estimates for linear problems to order  $\mathcal{O}(h^{1/k})$  without making further assumptions on the smoothness of the problem parameters.

## 2 Regularization

By adding the regularization term  $\frac{\alpha}{2} \|u\|_2^2$  to the cost functional of our initial problem (OS) = (OS)<sup>0</sup> we obtain the following class of linear-quadratic control problems with  $\alpha \geq 0$ :

$$\begin{aligned} \text{(OS)}^\alpha \quad & \min_{(x,u) \in X} f^\alpha(x, u) \\ \text{s.t.} \quad & \dot{x}(t) = A(t)x(t) + B(t)u(t) \quad \text{a.e. on } [0, T], \\ & x(0) = a, \\ & u(t) \in U \quad \text{a.e. on } [0, T]. \end{aligned}$$

Hereby, the cost functional is defined as

$$\begin{aligned} f^\alpha(x, u) &:= \frac{1}{2}x(T)^\top Qx(T) + q^\top x(T) + \frac{\alpha}{2} \|u\|_2^2 \\ &\quad + \int_0^T \frac{1}{2}x(t)^\top W(t)x(t) + x(t)^\top S(t)u(t) + w(t)^\top x(t) + r(t)^\top u(t) \, dt \\ &= f(x, u) + \frac{\alpha}{2} \|u\|_2^2. \end{aligned}$$

The feasible set of (OS)<sup>α</sup> is  $\mathcal{F}$ . If we choose  $\alpha > 0$ , the cost functional  $f^\alpha$  is strictly convex, and therefore the problem (OS)<sup>α</sup> has got a uniquely defined Lipschitz continuous solution  $(x^\alpha, u^\alpha)$  (cf. [9, Lemma 4]) and a corresponding adjoint state  $\lambda^\alpha \in$

$X_1$  such that the adjoint equation

$$\begin{aligned} -\dot{\lambda}^\alpha(t) &= A(t)^\top \lambda^\alpha(t) + W(t)x^\alpha(t) + S(t)u^\alpha(t) + w(t), \\ \lambda^\alpha(T) &= Qx^\alpha(T) + q \end{aligned}$$

and the minimum principle

$$\left( \alpha u^\alpha(t) + B(t)^\top \lambda^\alpha(t) + S(t)^\top x^\alpha(t) + r(t) \right)^\top (u - u^\alpha(t)) \geq 0 \quad \forall u \in U, \quad (14)$$

hold a.e. on  $[0, T]$ . For  $t \in [0, T]$  the switching function  $\sigma^\alpha$  of  $(OS)^\alpha$  is defined by

$$\sigma^\alpha(t) := \alpha u^\alpha(t) + B(t)^\top \lambda^\alpha(t) + S(t)^\top x^\alpha(t) + r(t)$$

From (14) it follows that for  $\alpha > 0$  the optimal control  $u^\alpha$  is

$$u^\alpha(t) = \text{Pr}_{[b_l, b_u]} \left[ -\frac{1}{\alpha} \left( B(t)^\top \lambda^\alpha(t) + S(t)^\top x^\alpha(t) + r(t) \right) \right]. \quad (15)$$

We will show that for  $\alpha \rightarrow 0$  the solution  $(x^\alpha, u^\alpha)$  of  $(OS)^\alpha$  converges to  $(x^0, u^0)$ . The order of convergence depends on the structure of the switching function of the initial problem (OS).

**Theorem 2.1** *Let  $(x^0, u^0)$  be a solution of (OS) that fulfills the Assumptions (A1) and  $(A2)^k$ . Then the uniquely determined solution  $(x^\alpha, u^\alpha)$  of  $(OS)^\alpha$  can be estimated by*

$$\|u^\alpha - u^0\|_1 \leq c_u \alpha^{\frac{1}{k}} \quad \text{and} \quad \|x^\alpha - x^0\|_\infty \leq c_x \alpha^{\frac{1}{k}}.$$

For the corresponding adjoint variable it holds

$$\|\lambda^\alpha - \lambda^0\|_1 \leq c_\lambda \alpha^{\frac{1}{k}}.$$

The constants  $c_u$ ,  $c_x$  and  $c_\lambda$  are independent of the regularization parameter  $\alpha$ .

*Proof* For the solution  $(x^\alpha, u^\alpha)$  of  $(OS)^\alpha$  it holds

$$f(x^\alpha, u^\alpha) + \frac{\alpha}{2} \|u^\alpha\|_2^2 = f^\alpha(x^\alpha, u^\alpha) \leq f^\alpha(x^0, u^0) = f(x^0, u^0) + \frac{\alpha}{2} \|u^0\|_2^2$$

and therefore

$$f(x^\alpha, u^\alpha) - f(x^0, u^0) \leq \frac{\alpha}{2} \left( \|u^0\|_2^2 - \|u^\alpha\|_2^2 \right). \quad (16)$$

Since  $U$  is bounded we obtain

$$\begin{aligned} \|u^0\|_2^2 - \|u^\alpha\|_2^2 &= \int_0^T |u^0(t)|^2 - |u^\alpha(t)|^2 \, dt \\ &= \int_0^T (|u^0(t)| + |u^\alpha(t)|) (|u^0(t)| - |u^\alpha(t)|) \, dt \\ &\leq \int_0^T (|u^0(t)| + |u^\alpha(t)|) |u^0(t) - u^\alpha(t)| \, dt \\ &\leq (\|u^0\|_\infty + \|u^\alpha\|_\infty) \|u^0 - u^\alpha\|_1 \\ &\leq c_1 \|u^0 - u^\alpha\|_1 \end{aligned}$$

with a constant  $c_1$  independent of  $\alpha$ . With (16) it follows

$$f(x^\alpha, u^\alpha) - f(x^0, u^0) \leq \frac{c_1}{2} \alpha \|u^0 - u^\alpha\|_1. \tag{17}$$

From Theorem 1.4 we know that there are constants  $\beta, \gamma, \bar{\delta} > 0$  with

$$f(x^\alpha, u^\alpha) - f(x^0, u^0) \geq \beta \|u^\alpha - u^0\|_1^{k+1},$$

if  $\|u^\alpha - u^0\|_1 \leq 2\gamma\bar{\delta}$  and

$$f(x^\alpha, u^\alpha) - f(x^0, u^0) \geq \beta \|u^\alpha - u^0\|_1,$$

if  $\|u^\alpha - u^0\|_1 \geq 2\gamma\bar{\delta}$ . With (17) we can deduce

$$\frac{c_1}{2} \alpha \|u^0 - u^\alpha\|_1 \geq \beta \|u^\alpha - u^0\|_1^{k+1},$$

if  $\|u^\alpha - u^0\|_1 \leq 2\gamma\bar{\delta}$  and

$$\frac{c_1}{2} \alpha \|u^0 - u^\alpha\|_1 \geq \beta \|u^\alpha - u^0\|_1,$$

if  $\|u^\alpha - u^0\|_1 \geq 2\gamma\bar{\delta}$ . Together with

$$\|u^0 - u^\alpha\|_1 \leq T (\|u^0\|_\infty + \|u^\alpha\|_\infty)$$

we obtain

$$\frac{c_1}{2\beta} \alpha \geq \|u^\alpha - u^0\|_1^k, \tag{18}$$

if  $\|u^\alpha - u^0\|_1 \leq 2\gamma\bar{\delta}$  and

$$\frac{c_1 T}{2\beta} (\|u^0\|_\infty + \|u^\alpha\|_\infty) \alpha \geq \|u^\alpha - u^0\|_1,$$

if  $\|u^\alpha - u^0\|_1 \geq 2\gamma\bar{\delta}$ . In both cases it follows for sufficiently small  $\alpha$

$$\|u^\alpha - u^0\|_1^k \leq 2\gamma\bar{\delta}.$$

From (18) we obtain the desired estimate for the control

$$\|u^\alpha - u^0\|_1 \leq c_u \alpha^{\frac{1}{k}}$$

with a constant  $c_u$  independent of  $\alpha$ . Since  $z := x^\alpha - x^0$  solves

$$\begin{aligned} \dot{z}(t) &= A(t)z(t) + B(t) \left( u^\alpha(t) - u^0(t) \right) \quad \text{a.e. on } [0, T], \\ z(0) &= 0 \end{aligned}$$

we get from (2)

$$\|x^\alpha - x^0\|_\infty = \|z\|_\infty \leq \bar{c}_1 \|u^\alpha - u^0\|_1 \leq c_x \alpha^{\frac{1}{k}}$$

with a constant  $c_x$  independent of  $\alpha$ . Moreover  $\mu := \lambda^\alpha - \lambda^0$  solves

$$\begin{aligned} -\dot{\mu}(t) &= A(t)^\top \mu(t) + W(t) \left( x^\alpha(t) - x^0(t) \right) + S(t) \left( u^\alpha(t) - u^0(t) \right) \quad \text{a.e. on } [0, T], \\ \mu(T) &= Q \left( x^\alpha(T) - x^0(T) \right). \end{aligned}$$

From (8) we can deduce the estimate

$$\|\lambda^\alpha - \lambda^0\|_\infty = \|\mu\|_\infty \leq \bar{c}_1 \|x^\alpha - x^0\|_\infty + \bar{c}_2 \|u^\alpha - u^0\|_1 \leq c_\lambda \alpha^{\frac{1}{k}}$$

with a constant  $c_\lambda$  independent of  $\alpha$ . □

Now we illustrate the theoretical findings of Theorem 2.1 with the help of Problem (B)<sub>k</sub> from Example 1.2.

*Example 2.2* The  $L^2$ -regularization of (B)<sub>k</sub> leads to:

$$\begin{aligned} \text{(B)}_k^\alpha \quad & \min_{(x,u) \in X} && x_1(1) + \frac{\alpha}{2} \|u\|_2^2 \\ \text{s.t.} &&& \dot{x}_j(t) = s_j x_{j+1}(t) + u(t) && \text{a.e. on } [0, 1], \quad j = 1, \dots, k, \\ &&& \dot{x}_{k+1}(t) = u(t) && \text{a.e. on } [0, 1], \\ &&& x(0) = 0_n, \\ &&& -1 \leq u(t) \leq 1 && \text{a.e. on } [0, 1]. \end{aligned}$$

We have already seen that with  $s_j = -2(k - j + 1)$  for  $j = 1, \dots, k$  the switching function of (B)<sub>k</sub><sup>0</sup> is given by

$$\sigma^0(t) = 2^k (t - 0.5)^k \quad \text{a.e. on } [0, 1].$$

Using the projection formula (15) the optimal control  $u^\alpha$  of the regularized problem  $(B)_k^\alpha$  with  $\alpha > 0$  can be written as

$$u^\alpha(t) = \Pr_{[-1,1]} \left[ -\frac{1}{\alpha} \sigma^0(t) \right] \quad \text{a.e. on } [0, 1].$$

We define

$$t_1 := \frac{1}{2} - \frac{1}{2} \alpha^{1/k} \quad \text{and} \quad t_2 := \frac{1}{2} + \frac{1}{2} \alpha^{1/k},$$

and for sufficiently small  $\alpha$ , we obtain  $[t_1, t_2] \subset [0, 1]$ . From the projection formula (15) it follows for odd  $k \in \mathbb{N}$

$$u^\alpha(t) = \begin{cases} 1, & 0 \leq t \leq t_1, \\ -\frac{2^k}{\alpha} (t - 0.5)^k, & t_1 < t < t_2, \\ -1, & t_2 \leq t \leq 1, \end{cases}$$

and for even  $k \in \mathbb{N}$

$$u^\alpha(t) = \begin{cases} -1, & 0 \leq t \leq t_1, \\ -\frac{2^k}{\alpha} (t - 0.5)^k, & t_1 < t < t_2, \\ -1, & t_2 \leq t \leq 1. \end{cases}$$

From this we can conclude, that for arbitrary  $k \in \mathbb{N}$  and sufficiently small  $\alpha > 0$  the optimal control  $u^0$  of  $(B)_k^0$  and the optimal control  $u^\alpha$  of  $(B)_k^\alpha$  coincide on  $[0, 1] \setminus [t_1, t_2]$ . Since it holds

$$\|u^\alpha - u^0\|_1 = 2 \left( \frac{1}{2} \alpha^{\frac{1}{k}} - \int_0^{\frac{1}{2} \alpha^{\frac{1}{k}}} \frac{2^k}{\alpha} t^k \, dt \right) = \left( 1 - \frac{1}{k+1} \right) \alpha^{\frac{1}{k}},$$

this example illustrates the theoretical results of Theorem 2.1.

### 3 The regularization–discretization approach

We choose a natural number  $N$  and define the *mesh size*  $h := T/N$ . The space  $X_2$  of controls is approximated by functions in the subspace  $X_{2,N} \subset X_2$  of piecewise constant functions represented by their values  $u_{h,i} := u(t_i)$  at the grid points  $t_i := ih, i = 0, 1, \dots, N - 1$ . Further, we approximate state and adjoint state variables by functions in the subspace  $X_{1,N} \subset X_1$  of continuous, piecewise linear functions represented by their values  $x_{h,i} := x(t_i), \lambda_{h,i} := \lambda(t_i)$  at the grid points  $t_i, i = 0, 1, \dots, N$ . We define  $\mathcal{U}_N := \mathcal{U} \cap X_{2,N}$  and  $X_N := X_{1,N} \times \mathcal{U}_N$ . In order to get a

discrete system equation we use Euler’s method:

$$\begin{aligned} x_{h,i+1} &= x_{h,i} + h [A(t_i)x_{h,i} + B(t_i)u_{h,i}], \quad i = 0, 1, \dots, N - 1, \\ x_{h,0} &= a. \end{aligned} \tag{19}$$

Discretizing the cost functional of (OS) with respect to the discretization method for the system equation and adding a discrete regularization term leads to the following class of discrete control problems depending on the regularization parameter  $\alpha$ :

$$\begin{aligned} \text{(OS)}_N^\alpha \quad & \min_{(x_h, u_h) \in X_N} f_N^\alpha(x_h, u_h) \\ \text{s.t.} \quad & x_{h,i+1} = x_{h,i} + h [A(t_i)x_{h,i} + B(t_i)u_{h,i}], \quad i = 0, 1, \dots, N - 1, \\ & x_{h,0} = a, \\ & u_{h,i} \in U, \quad i = 0, 1, \dots, N - 1. \end{aligned}$$

where the cost functional  $f_N^\alpha$  is defined by

$$\begin{aligned} f_N^\alpha(x, u) := & \frac{1}{2}x_N^\top Qx_N + q^\top x_N + h \sum_{i=0}^{N-1} \frac{\alpha}{2} u_i^\top u_i + \frac{1}{2}x_i^\top W(t_i)x_i + x_i^\top S(t_i)u_i \\ & + h \sum_{i=0}^{N-1} w(t_i)^\top x_i + r(t_i)^\top u_i. \end{aligned}$$

The feasible set of  $(\text{OS})_N^\alpha$  is independent of the regularization parameter  $\alpha$ . We denote it by  $\mathcal{F}_N$ . Analogously to (2) it holds for the solution  $x_h$  of (19)

$$\|x_h\|_\infty \leq c_1|a| + c_2\|u_h\|_1 \tag{20}$$

with constants  $c_1$  and  $c_2$  independent of  $a$  and  $u_h$ .

*Remark 4* Since we added the regularization term  $h \sum_{i=0}^{N-1} \frac{\alpha}{2} u_i^\top u_i$  to the cost functional, the problem  $(\text{OS})_N^\alpha$  fulfills a common sufficient condition for the convergence of finite-dimensional optimization methods. Therefore the combined regularization–discretization approach is numerically more stable and has advantages over direct approximation methods as presented for example in [6].

The following auxiliary results are similar to the ones concerning Euler discretization (cf. [6]).

**Definition 3.1** A pair  $(x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N$  is called a *minimizer for Problem  $(\text{OS})_N^\alpha$*  if  $f_N^\alpha(x_h^\alpha, u_h^\alpha) \leq f_N^\alpha(x_h, u_h)$  for all  $(x_h, u_h) \in \mathcal{F}_N$ , and a *strict minimizer for Problem  $(\text{OS})_N^\alpha$*  if  $f_N^\alpha(x_h^\alpha, u_h^\alpha) < f_N^\alpha(x_h, u_h)$  for all  $(x_h, u_h) \in \mathcal{F}_N, (x_h, u_h) \neq (x_h^\alpha, u_h^\alpha)$ .

Again, since  $U$  is compact there exists a constant  $L_x$  independent of  $N$  such that for any admissible control  $u_h \in \mathcal{U}$  and the associated solution  $x_h \in X_{1,N}$  of the discrete system Eq. (19) we have

$$|\dot{x}_h(t)| \leq L_x \quad \text{a.e. on } [0, T],$$



which shows that the discrete feasible trajectories are uniformly Lipschitz with Lipschitz modulus  $L_x$  independent of  $N$ , where w.l.o.g.  $L_x$  is the same constant as in (3).

Compactness of  $U$  further implies that Problem  $(OS)_N^\alpha$  has a solution  $(x_h^\alpha, u_h^\alpha) \in X_N$ , and for any solution there exists a continuous, piecewise linear multiplier  $\lambda_h^\alpha \in X_{1,N}$  such that for  $i = 0, \dots, N - 1$  the discrete adjoint equation

$$\begin{aligned}
 -\frac{\lambda_{h,i+1}^\alpha - \lambda_{h,i}^\alpha}{h} &= A(t_i)^\top \lambda_{h,i+1}^\alpha + W(t_i)x_{h,i}^\alpha + S(t_i)u_{h,i}^\alpha + w(t_i), \\
 \lambda_{h,N}^\alpha &= Qx_{h,N}^\alpha + q,
 \end{aligned}
 \tag{21}$$

and the discrete minimum principle

$$\left[ \alpha u_{h,i}^\alpha + B(t_i)^\top \lambda_{h,i+1}^\alpha + S(t_i)^\top x_{h,i}^\alpha + r(t_i) \right]^\top (u - u_{h,i}^\alpha) \geq 0 \quad \forall u \in U \tag{22}$$

are satisfied (cf. [6]). By  $\sigma_h^\alpha : [0, t_{N-1}] \rightarrow \mathbb{R}^m$  we denote the discrete switching function, which is the continuous and piecewise linear function defined by the values

$$\sigma_h^\alpha(t_i) := \alpha u_{h,i}^\alpha + B(t_i)^\top \lambda_{h,i+1}^\alpha + S(t_i)^\top x_{h,i}^\alpha + r(t_i), \quad i = 0, \dots, N - 1.$$

If  $\alpha = 0$ , from (22) we obtain for  $j = 1, \dots, m, i = 0, \dots, N - 1$ ,

$$u_{h,j}^0(t_i) = \begin{cases} b_{l,j}, & \text{if } \sigma_{h,j}^0(t_i) > 0, \\ b_{u,j}, & \text{if } \sigma_{h,j}^0(t_i) < 0, \\ \text{undetermined,} & \text{if } \sigma_{h,j}^0(t_i) = 0. \end{cases}$$

Otherwise if  $\alpha > 0$ , the following projection formula holds

$$u_{h,i}^\alpha = \text{Pr}_{[b_l, b_u]} \left[ -\frac{1}{\alpha} \left( B(t_i)^\top \lambda_{h,i+1}^\alpha + S(t_i)^\top x_{h,i}^\alpha + r(t_i) \right) \right], \quad i = 0, \dots, N - 1. \tag{23}$$

Analogously to (8), for the solution  $\lambda_h^\alpha$  of (21) we have

$$\|\lambda_h^\alpha\|_\infty \leq c_1 \|x_h^\alpha\|_\infty + c_2 \|u_h^\alpha\|_1 + c_3 \|w\|_\infty + c_4 |q| \tag{24}$$

with constants  $c_1, c_2, c_3$  and  $c_4$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $w$  and  $q$ . For the following proofs of convergence results we need some auxiliary results (cf. [6, Lemma 3.1, Lemma 3.2], [33, Satz 3.2.1, Lemma 3.2.2, Satz 3.2.3]):

**Lemma 3.2** *Let  $(x, u) \in \mathcal{F}$  be arbitrary with  $u$  having bounded variation, and let  $\hat{u} \in U_N$  be the piecewise constant approximation of  $u$  in the grid points. Then there exists a uniquely defined  $\hat{x} \in X_{1,N}$ , such that  $(\hat{x}, \hat{u}) \in \mathcal{F}_N$ , and it holds*

$$\|u - \hat{u}\|_1 \leq hV_0^T u \quad \text{and} \quad \|x - \hat{x}\|_\infty \leq c_1 hV_0^T \dot{x} \leq (c_2 + c_3 V_0^T u) h,$$

where  $c_1, c_2$  and  $c_3$  are constants independent of  $(x, u)$  and  $N$ .

**Lemma 3.3** *Let  $(x_h, u_h) \in \mathcal{F}_N$ . Then there exists a unique state  $x \in X_1$ , such that  $(x, u_h) \in \mathcal{F}$  and*

$$\|x - x_h\|_\infty \leq ch,$$

with a constant  $c$  independent of  $(x_h, u_h)$  and  $N$ .

**Lemma 3.4** *Let  $(x, u) \in \mathcal{F}$  be arbitrary with  $u$  having bounded variation, and let  $\lambda$  be the corresponding adjoint variable. Moreover let  $\hat{u} \in \mathcal{U}_N$  be the piecewise constant approximation of  $u$  in the grid points,  $\hat{x} \in X_{1,N}$  the solution of the discrete system Eq. (19) with  $u = \hat{u}$  and  $\hat{\lambda} \in X_{1,N}$  the solution of the discrete adjoint Eq. (21) with  $(x, u) = (\hat{x}, \hat{u})$ . Then it holds*

$$\|\lambda - \hat{\lambda}\|_\infty \leq c_1h + c_2h\mathbf{V}_0^T u$$

with constants  $c_1$  and  $c_2$  independent of  $(x, u)$  and  $N$ .

*Remark 5* In order to derive convergence results we introduce the following notations. Let  $\hat{u}^0 \in \mathcal{U}_N$  be the piecewise constant approximation of  $u^0$  in the grid points and  $\hat{x}^0 \in X_{1,N}$  the uniquely defined solution of the discrete system Eq. (19) with  $u = \hat{u}^0$ . Moreover let  $\hat{z}^0 \in X_{1,N}$  resp.  $z_h^\alpha \in X_{1,N}$  be the uniquely defined solutions of the system Eq. (1) with  $u = \hat{u}^0$  resp.  $u = u_h^\alpha$ . For arbitrary  $\alpha \geq 0$  it holds

$$(x^0, u^0) \in \mathcal{F}, (\hat{z}^0, \hat{u}^0) \in \mathcal{F}, (z_h^\alpha, u_h^\alpha) \in \mathcal{F}, (x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N, (\hat{x}^0, \hat{u}^0) \in \mathcal{F}_N. \tag{25}$$

### 4 Convergence of optimal values

To show convergence of the optimal values we have to prove the following auxiliary result:

**Lemma 4.1** *For  $(x_h, u_h) \in \mathcal{F}_N$  it holds*

$$|f(x_h, u_h) - f_N^\alpha(x_h, u_h)| \leq \bar{c}_1h + \bar{c}_2\alpha$$

with constants  $\bar{c}_1$  and  $\bar{c}_2$  independent of  $(x_h, u_h)$ ,  $\alpha$  and  $N$ .

*Proof* It follows from the Lipschitz continuity of  $x_h, W, S, w$  and  $r$ , that for  $(x_h, u_h) \in \mathcal{F}_N$  it holds (cf. [6, Lemma 3.3])

$$|f(x_h, u_h) - f_N^0(x_h, u_h)| \leq \bar{c}_1h \tag{26}$$

with a constant  $\bar{c}_1$  independent of  $(x_h, u_h)$ ,  $\alpha$  and  $N$ . Since  $U$  is bounded, we obtain

$$|f_N^0(x_h, u_h) - f_N^\alpha(x_h, u_h)| = \left| h \sum_{i=0}^{N-1} \frac{\alpha}{2} u_{h,i}^\top u_{h,i} \right| \leq \frac{T}{2} \|u_h\|_\infty^2 \alpha = \bar{c}_2\alpha, \tag{27}$$

with a constant  $\bar{c}_2$  independent of  $(x_h, u_h)$ ,  $\alpha$  and  $N$ . With the help of the triangle inequality we obtain the assertion from (26) and (27).  $\square$

Now we are ready to formulate the following theorem:

**Theorem 4.2** *Let  $(x^0, u^0) \in \mathcal{F}$  be a solution of (OS), with  $u^0$  having bounded variation. Then, for every solution  $(x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N$  of  $(OS)_N^\alpha$  it holds*

$$\left| f(x^0, u^0) - f_N^\alpha(x_h^\alpha, u_h^\alpha) \right| \leq \bar{c}_1 h + \bar{c}_2 \alpha$$

with constants  $\bar{c}_1$  and  $\bar{c}_2$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ .

*Proof* Since  $(x_h^\alpha, u_h^\alpha)$  is a solution of  $(OS)_N^\alpha$  we obtain with (25)

$$0 \leq f_N^\alpha(\hat{x}^0, \hat{u}^0) - f_N^\alpha(x_h^\alpha, u_h^\alpha) \leq f_N^\alpha(\hat{x}^0, \hat{u}^0) - f(x^0, u^0) + f(x^0, u^0) - f_N^\alpha(x_h^\alpha, u_h^\alpha).$$

Therefore it holds

$$\begin{aligned} f_N^\alpha(x_h^\alpha, u_h^\alpha) - f(x^0, u^0) &\leq f_N^\alpha(\hat{x}^0, \hat{u}^0) - f(x^0, u^0) \\ &\leq f_N^\alpha(\hat{x}^0, \hat{u}^0) - f(\hat{x}^0, \hat{u}^0) + f(\hat{x}^0, \hat{u}^0) - f(x^0, u^0). \end{aligned}$$

With Lemmas 3.2, 4.1 and the Lipschitz continuity of  $f$  we obtain

$$f_N^\alpha(x_h^\alpha, u_h^\alpha) - f(x^0, u^0) \leq c_1 h + c_2 \alpha \tag{28}$$

with constants  $c_1$  and  $c_2$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . Since  $(x^0, u^0)$  solves (OS) it holds

$$0 \leq f(z_h^\alpha, u_h^\alpha) - f(x^0, u^0) = f(z_h^\alpha, u_h^\alpha) - f_N^\alpha(x_h^\alpha, u_h^\alpha) + f_N^\alpha(x_h^\alpha, u_h^\alpha) - f(x^0, u^0)$$

and therefore

$$\begin{aligned} f(x^0, u^0) - f_N^\alpha(x_h^\alpha, u_h^\alpha) &\leq f(z_h^\alpha, u_h^\alpha) - f_N^\alpha(x_h^\alpha, u_h^\alpha) \\ &\leq f(z_h^\alpha, u_h^\alpha) - f(x_h^\alpha, u_h^\alpha) + f(x_h^\alpha, u_h^\alpha) - f_N^\alpha(x_h^\alpha, u_h^\alpha). \end{aligned}$$

With Lemmas 3.3, 4.1 and the Lipschitz continuity of  $f$  we obtain

$$f(x^0, u^0) - f_N^\alpha(x_h^\alpha, u_h^\alpha) \leq c_3 h + c_4 \alpha$$

with constants  $c_3$  and  $c_4$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . Together with (28) this leads to the desired assertion.  $\square$

### 5 Hölder-type error estimates

With the help of Theorem 4.2 we can now prove the following convergence result for the discrete regularization:

**Theorem 5.1** *Let  $(x^0, u^0) \in \mathcal{F}$  be a solution of (OS) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>. Then, for sufficiently small  $h$  every solution  $(x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N$  of (OS)<sub>N</sub><sup>α</sup> can be estimated by*

$$\|u_h^\alpha - u^0\|_1 \leq c_u(h + \alpha)^{\frac{1}{k+1}} \quad \text{and} \quad \|x_h^\alpha - x^0\|_\infty \leq c_x(h + \alpha)^{\frac{1}{k+1}}.$$

For the corresponding adjoint variables we get

$$\|\lambda_h^\alpha - \lambda^0\|_\infty \leq c_\lambda(h + \alpha)^{\frac{1}{k+1}}.$$

The constants  $c_u$ ,  $c_x$  and  $c_\lambda$  are independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ .

*Proof* From Theorem 1.4 we know that there exist constants  $\beta, \gamma, \bar{\delta} > 0$ , such that with (25) we have

$$f(z_h^\alpha, u_h^\alpha) - f(x^0, u^0) \geq \beta \|u_h^\alpha - u^0\|_1^{k+1}, \tag{29}$$

if  $\|u_h^\alpha - u^0\|_1 \leq 2\gamma\bar{\delta}$  and

$$f(z_h^\alpha, u_h^\alpha) - f(x^0, u^0) \geq \beta \|u_h^\alpha - u^0\|_1, \tag{30}$$

if  $\|u_h^\alpha - u^0\|_1 \geq 2\gamma\bar{\delta}$ . With Lemmas 3.3, 4.1, Theorem 4.2 and the Lipschitz continuity of  $f$  we further obtain

$$\begin{aligned} f(z_h^\alpha, u_h^\alpha) - f(x^0, u^0) &\leq |f(z_h^\alpha, u_h^\alpha) - f(x_h^\alpha, u_h^\alpha)| + |f(x_h^\alpha, u_h^\alpha) - f_N^\alpha(x_h^\alpha, u_h^\alpha)| \\ &\quad + \left| f_N^\alpha(x_h^\alpha, u_h^\alpha) - f(x^0, u^0) \right| \\ &\leq c_1 L_f \|z_h^\alpha - x_h^\alpha\|_\infty + c_2 h + c_3 \alpha \\ &\leq \bar{c}(h + \alpha) \end{aligned}$$

with a constant  $\bar{c}$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . Together with (29) and (30) it follows

$$\|u_h^\alpha - u^0\|_1 \leq \max \left\{ \bar{c}\beta^{-1}(h + \alpha), \left( \bar{c}\beta^{-1}(h + \alpha) \right)^{\frac{1}{k+1}} \right\}.$$

For sufficiently small  $h$  and  $\alpha$  it holds  $\|u_h^\alpha - u^0\|_1 \leq 2\gamma\bar{\delta}$ , and from (29) we obtain

$$\|u_h^\alpha - u^0\|_1 \leq c_u(h + \alpha)^{\frac{1}{k+1}} \tag{31}$$

with a constant  $c_u$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . From Assumption (A1) we know that  $u^0$  has bounded variation, and together with (20), (31) and Lemma 3.2 it follows for the state variables

$$\begin{aligned} \|x_h^\alpha - x^0\|_\infty &= \|x_h^\alpha - \hat{x}^0 + \hat{x}^0 - x^0\|_\infty \\ &\leq \|x_h^\alpha - \hat{x}^0\|_\infty + \|\hat{x}^0 - x^0\|_\infty \\ &\leq c_3 \|u_h^\alpha - \hat{u}^0\|_1 + c_4 h \\ &\leq c_3 \|u_h^\alpha - u^0 + u^0 - \hat{u}^0\|_1 + c_4 h \\ &\leq c_z (h + \alpha)^{\frac{1}{k+1}} \end{aligned} \tag{32}$$

with a constant  $c_z$  independent of  $(x_h^\alpha, u_h^\alpha)$  and  $N$ . For the corresponding adjoint states we obtain from (24), (31), (32), Lemmas 3.2 and 3.4

$$\begin{aligned} \|\lambda_h^\alpha - \lambda^0\|_\infty &= \|\lambda_h^\alpha - \hat{\lambda}^0 + \hat{\lambda}^0 - \lambda^0\|_\infty \\ &\leq \|\lambda_h^\alpha - \hat{\lambda}^0\|_\infty + \|\hat{\lambda}^0 - \lambda^0\|_\infty \\ &\leq c_5 \|x_h^\alpha - \hat{x}^0\|_\infty + c_6 \|u_h^\alpha - \hat{u}^0\|_1 + c_7 h \\ &\leq c_\lambda (h + \alpha)^{\frac{1}{k+1}} \end{aligned}$$

with a constant  $c_\lambda$  independent of  $(x_h^\alpha, u_h^\alpha)$  and  $N$ . □

By choosing the regularization parameter  $\alpha$  with respect to the mesh size  $h$  of the discretization as  $\alpha := c_\alpha h$  with a constant  $c_\alpha \geq 0$  independent of  $N$ , it follows directly from Theorem 5.1:

**Corollary 5.2** *Let  $(x^0, u^0) \in \mathcal{F}$  be a solution (OS) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>. Moreover we choose  $\alpha := c_\alpha h$  with a constant  $c_\alpha \geq 0$  independent of  $N$ . Then, for sufficiently small  $h$  every solution  $(x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N$  of (OS)<sub>N</sub><sup>α</sup> can be estimated by*

$$\|u_h^\alpha - u^0\|_1 \leq c_u h^{\frac{1}{k+1}} \quad \text{and} \quad \|x_h^\alpha - x^0\|_\infty \leq c_x h^{\frac{1}{k+1}}.$$

For the corresponding adjoint variables we get

$$\|\lambda_h^\alpha - \lambda^0\|_\infty \leq c_\lambda h^{\frac{1}{k+1}}.$$

The constants  $c_u$ ,  $c_x$  and  $c_\lambda$  are independent of  $(x_h^\alpha, u_h^\alpha)$  and  $N$ .

*Remark 6* For  $\alpha = 0$ , Corollary 5.2 directly gives Hölder-type error estimates for the direct Euler discretization of (OS), under the Assumptions (A1) and (A2)<sup>k</sup>. Therefore the results of this paper extend the findings of [7] and [6], too. Equivalent convergence results for an implicit discretization scheme can be found in [33].

### 6 Improved error estimates for linear problems

The error estimates of Theorem 5.1 and Corollary 5.2 can be improved for the following class of linear problems:

$$\begin{aligned}
 \text{(OL)}^\alpha \quad & \min_{(x,u) \in X} f(x, u) = q^\top x(T) + \frac{\alpha}{2} \|u\|_2^2 \\
 \text{s.t.} \quad & \dot{x}(t) = A(t)x(t) + B(t)u(t) \quad \text{a.e. on } [0, T], \\
 & x(0) = a, \\
 & u(t) \in U \quad \text{a.e. on } [0, T].
 \end{aligned}$$

For arbitrary  $\alpha \geq 0$  the adjoint equation of  $(\text{OL})^\alpha$

$$\begin{aligned}
 -\dot{\lambda}^\alpha(t) &= A(t)^\top \lambda^\alpha(t) \quad \text{a.e. on } [0, T], \\
 \lambda^\alpha(T) &= q
 \end{aligned}$$

is independent of  $(x, u)$  and  $\alpha$ . Therefore it can be solved directly, and it holds  $\lambda^0 = \lambda^\alpha$ . For  $t \in [0, T]$  the corresponding switching function is defined by

$$\sigma^\alpha(t) := B(t)^\top \lambda^\alpha(t) + \alpha u^\alpha(t) = B(t)^\top \lambda^0(t) + \alpha u^\alpha(t) = \sigma^0(t) + \alpha u^\alpha(t).$$

The discrete adjoint equation for the finite-dimensional Problem  $(\text{OL})_N^\alpha$  is given by

$$\begin{aligned}
 -\frac{\lambda_{h,i+1}^\alpha - \lambda_{h,i}^\alpha}{h} &= A(t_i)^\top \lambda_{h,i+1}^\alpha, \quad i = 0, \dots, N - 1 \\
 \lambda_{h,N}^\alpha &= q.
 \end{aligned}$$

Again, it holds  $\lambda_h^0 = \lambda_h^\alpha$ . The corresponding discrete switching function  $\sigma_h^\alpha \in X_{1,N}$  is a continuous, piecewise linear function  $\sigma_h^\alpha: [0, t_{N-1}] \rightarrow \mathbb{R}^m$ , which is uniquely defined by

$$\sigma_{h,i}^\alpha := B(t_i)^\top \lambda_{h,i+1}^\alpha + \alpha u_{h,i}^\alpha = \sigma_{h,i}^0 + \alpha u_{h,i}^\alpha, \quad i = 0, \dots, N - 1.$$

From (23) we know for  $\alpha > 0$ , that we can characterize the optimal control  $u_h^\alpha$  of  $(\text{OL})_N^\alpha$  by

$$u_{h,i}^\alpha = \text{Pr}_{[b_l, b_u]} \left[ -\frac{1}{\alpha} \sigma_{h,i}^0 \right], \quad i = 0, \dots, N - 1. \tag{33}$$

Since for our linear problems (OL) resp.  $(\text{OL})_N^\alpha$  the corresponding adjoint variables  $\lambda^0$  resp.  $\hat{\lambda}^0$  to  $(x^0, u^0)$  resp.  $(\hat{x}^0, \hat{u}^0)$  (comp. (25)) are independent of  $(x, u)$  and  $\alpha$ . Therefore it holds  $\lambda^\alpha = \lambda^0$  and  $\lambda_h^\alpha = \hat{\lambda}^0$ . From Lemma 3.4 we obtain

$$\|\lambda_h^\alpha - \lambda^0\|_\infty = \|\hat{\lambda}^0 - \lambda^0\|_\infty \leq c_\lambda h \tag{34}$$

with a constant  $c_\lambda$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . For arbitrary  $i = 0, \dots, N - 1$  it further holds

$$\begin{aligned} \left| \sigma_h^0(t_i) - \sigma^0(t_i) \right| &\leq \left| B(t_i)^\top \lambda_h^0(t_{i+1}) - B(t_i)^\top \lambda^0(t_i) \right| \\ &\leq \|B\|_\infty \left| \lambda_h^0(t_i) - \lambda^0(t_i) \right| + \|B\|_\infty \left| \lambda_h^0(t_{i+1}) - \lambda_h^0(t_i) \right| \\ &\leq \|B\|_\infty \|\lambda_h^0 - \lambda^0\|_\infty + \|B\|_\infty L_{\lambda_h^0} h \\ &= \|B\|_\infty \|\lambda_h^\alpha - \lambda^0\|_\infty + \|B\|_\infty L_{\lambda_h^0} h \\ &\leq ch \end{aligned}$$

with a constant  $c$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . With the Lipschitz continuity of  $\sigma^0$  and  $\sigma_h^0$  it follows for  $i = 0, \dots, N - 2$  and  $t \in [t_i, t_{i+1})$

$$\begin{aligned} \left| \sigma_h^0(t) - \sigma^0(t) \right| &= \left| \sigma_h^0(t) - \sigma_h^0(t_i) + \sigma_h^0(t_i) - \sigma^0(t_i) + \sigma^0(t_i) - \sigma^0(t) \right| \\ &\leq \left| \sigma_h^0(t_i) - \sigma^0(t_i) \right| + \left| \sigma_h^0(t) - \sigma_h^0(t_i) \right| + \left| \sigma^0(t_i) - \sigma^0(t) \right| \\ &\leq ch + \left( L_{\sigma_h^0} + L_{\sigma^0} \right) h \\ &\leq c_\sigma h \end{aligned} \tag{35}$$

with a constant  $c_\sigma$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . From (34) and (35) we obtain

$$\|\lambda_h^\alpha - \lambda^0\|_\infty \leq c_\lambda h \quad \text{and} \quad \max_{t \in [0, t_{N-1}]} |\sigma_h^0(t) - \sigma^0(t)| \leq c_\sigma h \tag{36}$$

with constants  $c_\lambda$  and  $c_\sigma$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . Using (36) we are now able to prove the following improved error estimates:

**Theorem 6.1** *Let  $(x^0, u^0) \in \mathcal{F}$  be a solution of (OL) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>. Then, for sufficiently small  $h$  and  $\alpha$  every solution  $(x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N$  of (OL)<sub>N</sub><sup>α</sup> can be estimated by*

$$\|u_h^\alpha - u^0\|_1 \leq c_u (h + \alpha)^{\frac{1}{k}} \quad \text{and} \quad \|x_h^\alpha - x^0\|_\infty \leq c_x (h + \alpha)^{\frac{1}{k}}.$$

For the corresponding adjoint variable it holds

$$\|\lambda_h^\alpha - \lambda^0\|_\infty \leq c_\lambda h.$$

Moreover  $u_h^\alpha$  and  $u^0$  coincide, except on a set  $M$  of measure  $\mu(M) \leq \bar{\kappa} (h + \alpha)^{\frac{1}{k}}$ . The constants  $c_u, c_x, c_\lambda$  and  $\bar{\kappa}$  are independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ .

*Proof* Let

$$\Sigma_j := \{ \tau_1, \dots, \tau_j \} \quad \text{with} \quad 0 < \tau_1 < \dots < \tau_j < T$$

be the set of zeros of  $\sigma_j^0$ . Moreover we define

$$I_-(\delta) := \bigcup_{\iota=1, \dots, l_j} [\tau_\iota - \delta, \tau_\iota + \delta] \quad \text{and} \quad I_+(\delta) := [0, T] \setminus (I_-(\delta) \cup [t_{N-1}, T]).$$

Since the switching function  $\sigma_j$  is Lipschitz continuous it follows from the Assumptions (A1) and (A2)<sup>k</sup>

$$\sigma_{j,\min}^0 := \min_{t \in I_+(\bar{\tau})} |\sigma_j^0(t)| > 0.$$

From (36) we obtain

$$|\sigma_{h,j}^0(t)| \geq |\sigma_j^0(t)| - c_\sigma h \geq \sigma_{j,\min}^0 - c_\sigma h \quad \forall t \in I_+(\bar{\tau}).$$

For sufficiently small  $h$  this implies

$$\sigma_{h,j,\min}^0 := \min_{t \in I_+(\bar{\tau})} |\sigma_{h,j}^0(t)| \geq \frac{1}{2} \sigma_{j,\min}^0 > 0.$$

For arbitrary  $\iota \in 1, \dots, l_j$  we obtain from (A2)<sup>k</sup> and (36) for  $t \in [\tau_\iota - \bar{\tau}, \tau_\iota + \bar{\tau}]$

$$|\sigma_{h,j}^0(t)| \geq |\sigma_j^0(t)| - c_\sigma h \geq \bar{\sigma} |t - \tau_\iota|^k - c_\sigma h$$

and therefore

$$\left| \frac{1}{\alpha} \sigma_{h,j}^0(t) \right| \geq \frac{1}{\alpha} |\sigma_j^0(t)| - c_\sigma \frac{h}{\alpha} \geq \frac{\bar{\sigma}}{\alpha} |t - \tau_\iota|^k - c_\sigma \frac{h}{\alpha}.$$

We define

$$\gamma_j := \max \{ -b_{l,j}, b_{u,j} \}$$

and

$$d_j(h, \alpha) := \left( \frac{1}{\bar{\sigma}} (c_\sigma h + \gamma_j \alpha) \right)^{\frac{1}{k}}.$$

It holds  $\left| \frac{1}{\alpha} \sigma_{h,j}^0(t) \right| > \gamma_j$  and therefore  $u_{h,j}^\alpha = u_j^0$ , if  $|t - \tau_\iota| > d_j(h, \alpha)$ . Now we choose the index  $i \in \{0, \dots, N - 1\}$ , such that  $\tau_i \in [t_i, t_{i+1})$ . Let  $\ell \in \mathbb{N}$  be the smallest number for which it holds

$$\ell > d_j(h, \alpha) h^{-1}.$$



Then we obtain

$$t_{i+\ell+1} - \tau_i \geq t_{i+\ell+1} - t_{i+1} = \ell h > d_j(h, \alpha)$$

and

$$\tau_i - t_{i-\ell} \geq t_i - t_{i-\ell} = \ell h > d_j(h, \alpha)$$

and

$$d_j(h, \alpha)h^{-1} < \ell \leq d_j(h, \alpha)h^{-1} + 1.$$

By setting

$$\ell_i^+ := i + \ell + 1 \quad \text{and} \quad \ell_i^- := i - \ell$$

we get for sufficiently small  $h$

$$t_{\ell_i^+} - t_{\ell_i^-} = (2\ell + 1)h \leq (2d_j(h, \alpha)h^{-1} + 3)h \leq \kappa d_j(h, \alpha)$$

with a constant  $\kappa$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . For sufficiently small  $h$  and  $\alpha$  it follows directly

$$[t_{\ell_i^-}, t_{\ell_i^+}] \subset [\tau_i - \bar{\tau}, \tau_i + \bar{\tau}].$$

By defining

$$I_{j,-} := \left( \bigcup_{i=1, \dots, l_j} [t_{\ell_i^-}, t_{\ell_i^+}] \right) \cup [t_{N-1}, T]$$

we obtain

$$\left| \frac{1}{\alpha} \sigma_{h,j}^0(t) \right| > \gamma_j \quad \forall t \in [0, T] \setminus I_{j,-}. \tag{37}$$

Now we set

$$\gamma := \max_{1 \leq j \leq m} \gamma_j, \quad d(h, \alpha) := \left( \frac{1}{\bar{\sigma}} (c_\sigma h + \gamma \alpha) \right)^{\frac{1}{\bar{k}}} \quad \text{and} \quad M := \bigcup_{j=1, \dots, m} I_{j,-}.$$

From (37), together with the projection formula (33) and the estimate (36), we can deduce that  $u^0$  and  $u_h^\alpha$  coincide on  $[0, T] \setminus M$ . The measure  $\mu(M)$  can be estimated by

$$\mu(M) \leq \kappa \sum_{j=1}^m l_j d_j(h, \alpha) \leq \kappa \sum_{j=1}^m l_j d(h, \alpha) \leq \bar{\kappa} (h + \alpha)^{\frac{1}{\bar{k}}} \tag{38}$$

with a constant  $\bar{\kappa}$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . In order to estimate  $\|u_h^\alpha - u^0\|_1$  we define  $\rho := \max_{1 \leq j \leq m} (b_{u,j} - b_{l,j})$  and it holds

$$\begin{aligned}
 \|u_h^\alpha - u^0\|_1 &= \int_0^T |u_h^\alpha(t) - u^0(t)| dt \\
 &= \int_M |u_h^\alpha(t) - u^0(t)| dt \\
 &\leq \rho\mu(M) \leq \rho\bar{\kappa}(h + \alpha)^{\frac{1}{k}} = c_u(h + \alpha)^{\frac{1}{k}}
 \end{aligned}
 \tag{39}$$

with a constant  $c_u$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . For the state variables we obtain analogously to the proof of Theorem 5.1 from (39)

$$\|x_h^\alpha - x^0\|_\infty \leq c_x(h + \alpha)^{\frac{1}{k}}
 \tag{40}$$

with a constant  $c_x$  independent of  $(x_h^\alpha, u_h^\alpha)$ ,  $\alpha$  and  $N$ . We get the desired assertion from the estimates (36), (38), (39) and (40).  $\square$

Again, by coupling  $\alpha$  with  $h$  we obtain the following convergence result directly from Theorem 6.1:

**Corollary 6.2** *Let  $(x^0, u^0) \in \mathcal{F}$  be a solution of (OL) that fulfills the Assumptions (A1) and (A2)<sup>k</sup>. Moreover we choose  $\alpha := c_\alpha h$ , with a constant  $c_\alpha > 0$  independent of  $N$ . Then, for sufficiently small  $h$ , every solution  $(x_h^\alpha, u_h^\alpha) \in \mathcal{F}_N$  of (OL)<sub>N</sub><sup>α</sup> can be estimated by*

$$\|u_h^\alpha - u^0\|_1 \leq c_u h^{\frac{1}{k}} \quad \text{and} \quad \|x_h^\alpha - x^0\|_\infty \leq c_x h^{\frac{1}{k}}.$$

For the corresponding adjoint variable it holds

$$\|\lambda_h^\alpha - \lambda^0\|_\infty \leq c_\lambda h.$$

Moreover  $u_h^\alpha$  and  $u^0$  coincide, except on a set  $M$  of measure  $\mu(M) \leq \bar{\kappa}h^{\frac{1}{k}}$ . The constants  $c_u, c_x, c_\lambda$  and  $\bar{\kappa}$  are independent of  $(x_h^\alpha, u_h^\alpha)$  and  $N$ .

*Remark 7* Similar results as in Corollary 6.2 can be found in [19] for the direct Euler discretization of Mayer type problems with controllability index  $k$  (see Remark 3).

**Table 1** Example (B)<sub>k</sub>: error discrete regularization with  $\alpha = 10h$

	$N$	20	50	100	200	500
$k = 1 :$	$e_N$	0.25	0.1	0.05	0.025	0.01
	$e_N/h$	5	5	5	5	5
$k = 3 :$	$e_N$	0.815	0.5551	0.4194	0.3202	0.2269
	$e_N/\sqrt[3]{h}$	2.212	2.045	1.947	1.873	1.801

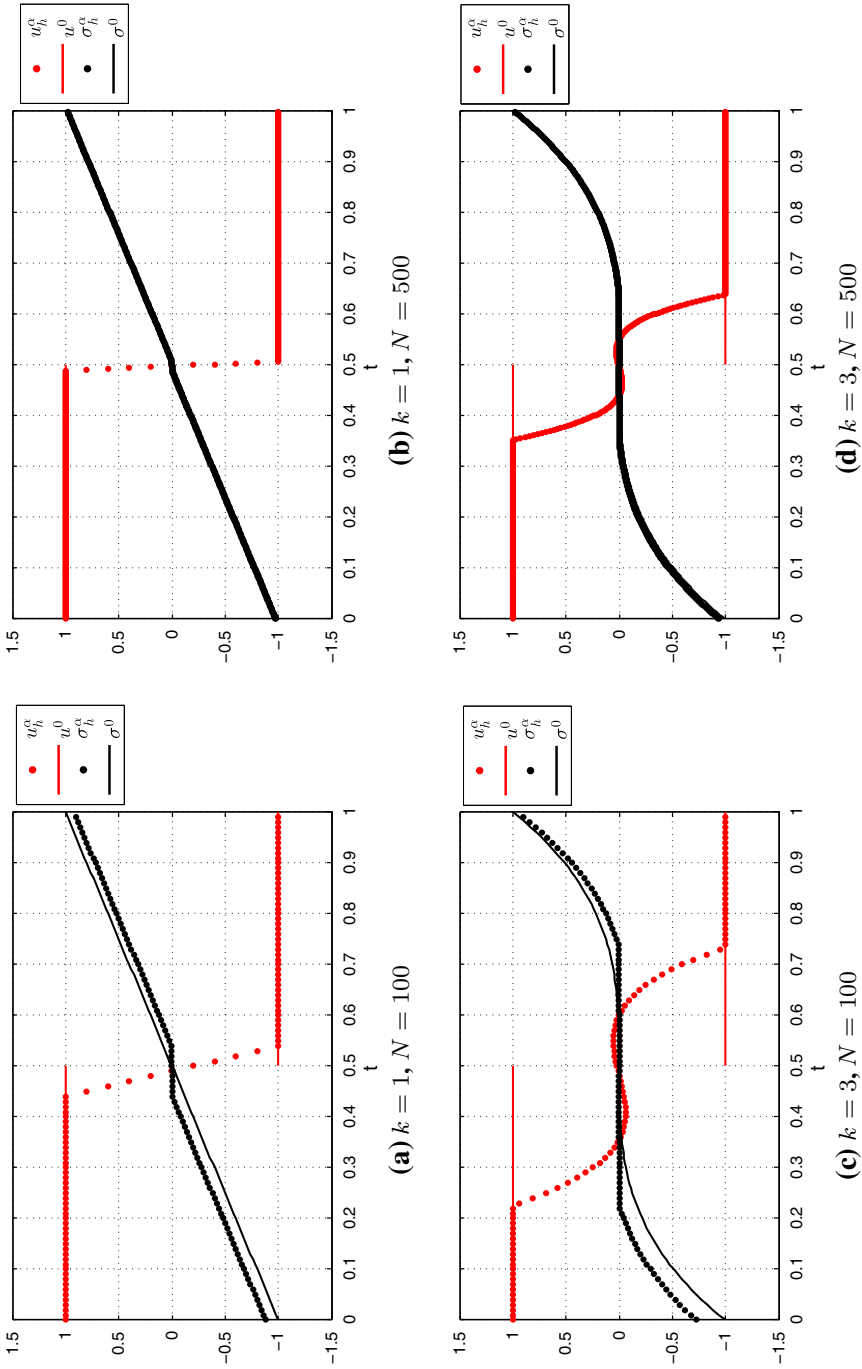


Fig. 1 Example  $(B)_k$ : solution discrete regularization with  $\alpha = 10h$

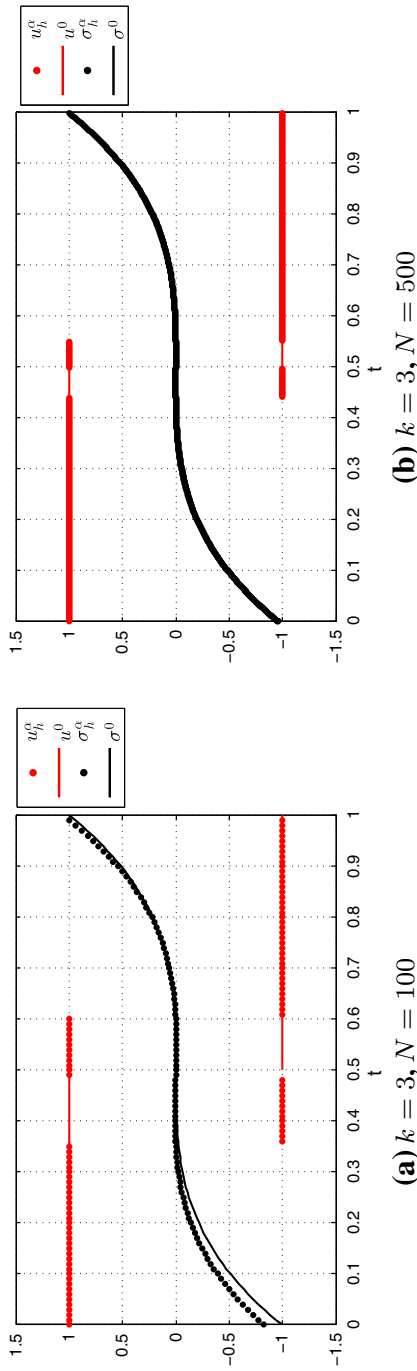


Fig. 2 Example (B)<sub>k</sub>: solution direct Euler discretization

## 7 Numerical example

We illustrate the theoretical findings of Corollary 6.2 by approximating the solution of  $(B)_k$  from Example 1.2 using discrete regularization:

*Example 7.1* We choose  $\alpha := 10h$  and define  $e_N := \|u_h^\alpha - u^0\|_1$ . The numerical results of solving  $(B)_k$  with the discrete regularization technique are displayed in Table 1 and Fig. 1. They confirm the theoretical results of Corollary 6.2.

*Remark 8* Figure 2 shows the solution of the direct Euler discretization of Example  $(B)_3$ . It seems like the optimal control has three switching points. As we can see in Fig. 1 the combined regularization–discretization approach gives a better understanding of the structure of the optimal control.

## 8 Conclusions

In this paper we proved error estimates for a combined regularization–discretization approach for a class of optimal control problems with bang–bang solutions. We were able to prove convergence of order  $\mathcal{O}(h^{1/(k+1)})$  with respect to the controllability index  $k$  for problems with mixed state-control-term in the cost functional under weaker assumptions on the structure of the switching function than in [4, 6, 7].

**Acknowledgments** The author would like to thank the anonymous reviewers for their valuable comments.

## References

1. Agrachev, A., Stefani, G., Zezza, P.L.: Strong optimality for a bang–bang trajectory. *SIAM J. Control Optim.* **41**, 991–1014 (2002)
2. Alt, W.: Discretization and mesh-independence of Newton’s method for generalized equations. In: Fiacco, A.V. (ed.) *Mathematical Programming with Data Perturbations V*. Lecture Notes in Pure and Applied Mathematics, vol. 195, pp. 1–30. Marcel Dekker (1997)
3. Alt, W., Bräutigam, N.: Finite-difference discretizations of quadratic control problems governed by ordinary elliptic differential equations. *Comput. Optim. Appl.* **43**, 133–150 (2009)
4. Alt, W., Seydenschwanz, M.: Regularization and discretization of linear-quadratic control problems. *Control Cybern.* **40**(4), 903–921 (2011)
5. Alt, W., Seydenschwanz, M.: An implicit discretization scheme for linear-quadratic control problems with bang–bang solutions. *Optim. Methods Softw.* **29**(3), 535–560 (2014)
6. Alt, W., Baier, R., Gerdts, M., Lempio, F.: Error bounds for Euler approximation of linear-quadratic control problems with bang–bang solutions. *Numer. Algebr. Control Optim.* **2**, 547–570 (2012)
7. Alt, W., Baier, R., Gerdts, M., Lempio, F.: Approximations of linear control problems with bang–bang solutions. *Optimization* **62**, 9–32 (2013)
8. Baier, R., Chahma, I.A., Lempio, F.: Stability and convergence of Euler’s method for state-constrained differential inclusions. *SIAM J. Optim.* **18**, 1004–1026 (2007)
9. Dontchev, A.L., Hager, W.W.: Lipschitzian stability in nonlinear control and optimization. *SIAM J. Control Optim.* **31**, 569–603 (1993)
10. Dontchev, A.L., Hager, W.W., Malanowski, K.: Error bounds for Euler approximation of a state and control constrained optimal control problem. *Numer. Funct. Anal. Optim.* **21**, 653–682 (2000)
11. Dontchev, A.L., Hager, W.W., Veliiov, V.M.: Second-order Runge-Kutta approximations in control constrained optimal control. *SIAM J. Numer. Anal.* **38**, 202–226 (2000)
12. Dontchev, A.L., Hager, W.W.: The Euler approximation in state constrained optimal control. *Math. Comput.* **70**, 173–203 (2001)

13. Ekeland, I., Temam, R.: *Convex Analysis and Variational Problems*. North Holland, Amsterdam-Oxford (1976)
14. Felgenhauer, U.: On stability of bang-bang type controls. *SIAM J. Control Optim.* **41**, 1843–1867 (2003)
15. Felgenhauer, U.: Optimality properties of controls with bang-bang components in problems with semilinear state equation. *Control Cybernet.* **34**, 763–785 (2005)
16. Felgenhauer, U.: The shooting approach in analyzing bang-bang extremals with simultaneous control switches. *Control Cybernet.* **37**, 307–327 (2008)
17. Felgenhauer, U., Poggiolini, L., Stefani, G.: Optimality and stability result for bang-bang optimal controls with simple and double switch behaviour. *Control Cybernet.* **38**, 1305–1325 (2009)
18. Hager, W.W.: Multiplier methods for nonlinear optimal control. *SIAM J. Numer. Anal.* **27**(4), 1061–1080 (1990)
19. Haunschmied, J., Pietrus, A., Veliov, V.M.: The Euler method for linear control systems revisited. In: Lirkov, I., Margenov, S., Wasniewski, J. (eds.) *Large-Scale Scientific Computing—LSSC 2013*. Lecture Notes in Computer Science, vol. 8353, pp. 90–97 (2014)
20. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.* **30**, 45–61 (2005)
21. Hinze, M., Meyer, C.: Variational discretization of Lavrentiev-regularized state constrained elliptic optimal control problems. *Comput. Optim. Appl.* **46**, 487–510 (2010)
22. Lorenz, D.A., Rösch, A.: Error estimates for joint Tikhonov- and Lavrentiev-regularization of constrained control problems. *Appl. Anal.* **89**(11), 1679–1691 (2010)
23. Malanowski, K., Büskens, C., Maurer, H.: Convergence of approximations to nonlinear optimal control problems. In: Fiacco, A.V. (ed.) *Mathematical Programming with Data Perturbations V'*. Lecture Notes in Pure and Applied Mathematics, vol. 195, pp. 253–284. Marcel Dekker (1997)
24. Maurer, H., Osmolowski, N.P.: Second order sufficient conditions for time optimal bang-bang control. *SIAM J. Control Optim.* **42**, 2239–2263 (2004)
25. Maurer, H., Büskens, C., Kim, J.-H.R., Kaya, C.Y.: Optimization methods for the verification of second-order sufficient conditions for bang-bang controls. *Optimal Control Appl. Methods* **26**, 129–156 (2005)
26. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. *SIAM J. Contr. Optim.* **43**, 970–985 (2004)
27. Meyer, C., Rösch, A., Tröltzsch, F.: Optimal control of PDEs with regularized pointwise state constraints. *Comput. Optim. Appl.* **33**, 209–228 (2006)
28. Neitzel, I., Tröltzsch, F.: On convergence of regularization methods for nonlinear parabolic optimal control problems with control and state constraints. *Control Cybernet.* **37**, 568–579 (2008)
29. Osmolowski, N.P., Lempio, F.: Transformation of quadratic forms to perfect squares for broken extremals. *Set-Valued Anal.* **10**, 209–232 (2002)
30. Osmolowski, N.P., Maurer, H.: Equivalence of second order optimality conditions for bang-bang control problems. Part 1: main results. *Control Cybernet.* **34**, 927–950 (2005)
31. Osmolowski, N.P., Maurer, H.: Equivalence of second order optimality conditions for bang-bang control problems. Part 2: Proofs, variational derivatives and representations. *Control Cybernet.* **36**, 5–45 (2007)
32. Quincampoix, M., Veliov, V.M.: Metric regularity and stability of optimal control problems for linear systems. *SIAM J. Control Optim.* **51**(5), 4118–4137 (2013)
33. Seydenschwanz, M.: *Diskretisierung und Regularisierung linear-quadratischer Steuerungsprobleme mit Bang-Bang Lösungen*. Dissertation, Mathematisches Institut, Friedrich-Schiller-Universität Jena (2014)
34. Tröltzsch, F., Yousept, I.: A regularization method for the numerical solution of elliptic boundary control problems with pointwise constraints. *Comput. Optim. Appl.* **42**, 43–66 (2009)
35. Veliov, V.M.: On the time-discretization of control systems. *SIAM J. Control Optim.* **35**, 1470–1486 (1997)
36. Veliov, V.M.: Error analysis of discrete approximations to bang-bang optimal control problems: the linear case. *Control Cybernet.* **34**, 967–982 (2005)