Comput Optim Appl (2018) 69:825–856
https://doi.org/10.1007/s10589-017-9969-7

Euler discretization for a class of nonlinear optimal control problems with control appearing linearly

Walter Alt[1] · Ursula Felgenhauer[2] · Martin Seydenschwanz[3]

Received: 21 May 2017 / Published online: 15 November 2017

Abstract We investigate Euler discretization for a class of optimal control problems with a nonlinear cost functional of Mayer type, a nonlinear system equation with control appearing linearly and constraints defined by lower and upper bounds for the controls. Under the assumption that the cost functional satisfies a growth condition we prove for the discrete solutions Hölder type error estimates w.r.t. the mesh size of the discretization. If a stronger second-order optimality condition is satisfied the order of convergence can be improved. Numerical experiments confirm the theoretical findings.

Keywords Optimal control · Bang-bang control · Euler discretization · Error estimates

Mathematics Subject Classification 49J15 · 49M25

1 Introduction

Discretization methods like Euler discretization are used for the numerical solution of optimal control problems. The accuracy of the approximate solutions obtained in

✉ Walter Alt
walteralt.unijena@gmail.com

Ursula Felgenhauer
felgenh@gmail.com

Martin Seydenschwanz
Martin.Seydenschwanz@siemens.com

1 Institut für Mathematik, Friedrich-Schiller-Universität Jena, 07740 Jena, Germany

2 Institut für Mathematik, Brandenburgische Technische Universität Cottbus-Senftenberg, 03013 Cottbus, Germany

3 Siemens AG, Research in Digitalization and Automation, 81739 Munich, Germany

this way are often satisfactory from a practical point of view. However, if the optimal control has a special structure, a discretization method may help to detect this structure and then methods based on structural assumptions can be used to determine the optimal control more accurately. Especially for bang–bang controls Euler discretization can be used to compute approximations of the switching points, and then efficient numerical approaches, such as switching time parameterization, can be employed to determine the switching times more accurately (see e.g. Kaya et al. [23], Maurer et al. [29], Osmolovskii and Maurer [34] and the papers cited therein). It is well-known that, in particular when the solution controls are of bang–bang or bang-singular type, many difficulties are encountered in getting an approximate solution. Therefore, it is of practical interest to have conditions implying error estimates for approximate solutions which ensure that the approximate controls (optimal controls for the discretized problems) converge to the optimal control of the original, continuous-time control problem. Such error estimates are closely related to estimates for solutions of perturbed optimal control problems.

Discretization and perturbation of nonlinear optimal control problems governed by ordinary differential equations are well studied for the case that the optimal control is sufficiently smooth, and the results are usually based on strong second-order optimality conditions which require coercivity of the second derivative of the Lagrangian function with respect to the control variables (see e.g. Dontchev and Hager [10,11], Dontchev et al. [12], Malanowski [24–26], Malanowski et al. [27], Alt [1–3]). For control problems with control appearing linearly such conditions are not satisfied and the optimal control may be discontinuous. Therefore, there have been only a few papers on discretization of such problems (see e.g. Alt and Mackenroth [5], Dhamo and Tröltzsch [9], Veliov [41] and the papers cited therein).

New second-order optimality conditions for optimal control problems with control appearing linearly have been developed during the last 10–15 years (see e.g. Felgenhauer [15–18,20], Maurer et al. [29], Osmolovskii and Maurer [32–34] and the papers cited therein). In case of bang–bang controls these conditions have been used in Alt et al. [4], Alt and Seydenschwanz [7], and in Seydenschwanz [40] to obtain error estimates for Euler discretization of linear-quadratic optimal control problems governed by ordinary differential equations and in Deckelnick and Hinze [8] for discretizations of elliptic control problems. For convex control problems of Mayer type with a linear system equation and bang–bang solutions Veliov [41] has shown convergence of order 1 for Euler discretization. These results have been extended in Haunschmied et al. [21] under more general conditions based on a result on stability of optimal control problems under strong bi-metric regularity of Quincampoix and Veliov [36]. Pietrus et al. [35] investigate high order discrete approximations to Mayer type problems based on second order Volterra-Fliess approximations. Felgenhauer [19] shows convergence of order 1 for a class of nonlinear optimal control problems, where the linear term in the system equation does not depend on the state variables and the solution has bang-singular-bang structure. Alt et al. [6] prove convergence of order 1 for implicit Euler discretization of a general class of convex, linear-quadratic control problems with bang–bang solutions.

In the present paper we investigate a class of optimal control with a nonlinear cost functional of Mayer type, a nonlinear system equation with control appearing linearly and constraints defined by lower and upper bounds for the controls. Under the assumption that the cost functional satisfies a growth condition of order $\kappa \geq 1$ we prove for the discrete solutions Hölder type error estimates of order $1/\kappa$ w.r.t. the mesh size of the discretization. If a stronger second-order condition for the derivative of the Lagrangian w.r.t. the control and a weakened coercivity condition for the second derivative of the Lagrangian are satisfied, the order of convergence can be improved to 1.

We use the following notations: $\mathbb{R}^n$ 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 M we denote the spectral norm by $\|M\| = \sup_{|z|\leq 1} |Mz|$. Let $t_0, t_f \in \mathbb{R}$, $t_0 < t_f$. We denote by $L^1(t_0, t_f; \mathbb{R}^m)$ the Banach space of integrable, measurable functions $u\colon [t_0, t_f] \to \mathbb{R}^m$ with

$$\|u\|_1 = \int_{t_0}^{t_f} \sum_{i=1}^{m} |u_i(t)|\,\mathrm{d}t = \sum_{i=1}^{m} \|u_i\|_1 < \infty,$$

by $L^\infty(t_0, t_f; \mathbb{R}^m)$ the Banach space of essentially bounded functions $u\colon [t_0, t_f] \to \mathbb{R}^m$ with the norm

$$\|u\|_\infty = \max_{1\leq i\leq m} \operatorname{ess\,sup}_{t\in[t_0,t_f]} |u_i(t)|,$$

and $C(t_0, t_f; \mathbb{R}^m)$ is the Banach space of continuous functions $u\colon [t_0, t_f] \to \mathbb{R}^m$ with the norm

$$\|u\|_\infty = \max_{1\leq i\leq m} \max_{t\in[t_0,t_f]} |u_i(t)|.$$

For $p \in \{1, \infty\}$ we denote by $W_p^1(t_0, t_f; \mathbb{R}^n)$ the spaces of absolutely continuous functions on $[t_0, t_f]$ with derivative in $L^p(t_0, t_f; \mathbb{R}^n)$, i.e.

$$W_p^1(t_0, t_f; \mathbb{R}^n) = \left\{x \in L^p(t_0, t_f; \mathbb{R}^n) \mid \dot{x} \in L^p(t_0, t_f; \mathbb{R}^n)\right\}$$

with

$$\|x\|_{1,1} = |x(t_0)| + \|\dot{x}\|_1, \quad \|x\|_{1,\infty} = \max\left\{\|x\|_\infty, \|\dot{x}\|_\infty\right\}.$$

Let $X = X_1 \times X_2$, where $X_1 = W_1^1(t_0, t_f; \mathbb{R}^n)$, $X_2 = L^1(t_0, t_f; \mathbb{R}^m)$. We consider the following optimal control problem:

$$\text{(OC)} \qquad \begin{aligned} &\min f(x(t_f)) \\ &\text{s.t.} \\ &\dot{x}(t) = g(x(t), u(t), t) && \text{a.e. on } [t_0, t_f], \\ &x(t_0) = a, \\ &u(t) \in U && \text{a.e. on } [t_0, t_f], \end{aligned}$$

where g is defined by

$$g(x, u, t) = g^{(1)}(x, t) + g^{(2)}(x, t)u. \tag{1.1}$$

Here, $u(t) \in \mathbb{R}^m$ is the control, and $x(t) \in \mathbb{R}^n$ is the state of a system at time $t \in [t_0, t_f]$. Further $f : \mathbb{R}^n \to \mathbb{R}$, $g^{(1)} : \mathbb{R}^n \times [t_0, t_f] \to \mathbb{R}^n$, $g^{(2)} : \mathbb{R}^n \times [t_0, t_f] \to \mathbb{R}^{n\times m}$, and the set $U \subset \mathbb{R}^m$ is defined by lower and upper bounds, i.e.,

$$U = \{u \in \mathbb{R}^m \mid b_\ell \le u \le b_u\}$$

with $b_\ell, b_u \in \mathbb{R}^m$, $b_\ell < b_u$, where all inequalities are to be understood componentwise.

The organization of the paper is as follows. In Sect. 2 we state some basic results. Section 3 introduces the Euler discretization for Problem (OC). In Sect. 4 Hölder type error estimates are derived assuming a growth condition for the cost functional. Under stronger second-order conditions we prove in Sect. 5 convergence of order 1. Section 6 discusses some numerical results.

2 Basic results

We denote by

$$\mathscr{U} = \{u \in X_2 \mid u(t) \in U \ \forall' t \in [t_0, t_f]\}$$

the set of admissible controls, and by

$$\mathscr{F} = \left\{(x, u) \in X \mid \dot{x}(t) = g(x(t), u(t), t) \ \text{ a.e. on } [t_0, t_f], \ x(t_0) = a, \ u \in \mathscr{U}\right\}$$

the feasible set of Problem (OC). For $\varepsilon > 0$ and $(x^*, u^*) \in X$

$$\mathscr{B}_\varepsilon(x^*, u^*) = \left\{(x, u) \in X \mid \|x - x^*\|_\infty < \varepsilon, \ \|u - u^*\|_1 < \varepsilon\right\}.$$

is the open ball around (x^*, u^*) with radius ε.

Definition 1 A pair $(x^*, u^*) \in \mathscr{F}$ is called a *local minimizer* of f on $\mathscr{F}$ or a *local solution* of Problem (OC), if there exists $\varepsilon > 0$ such that $f(x^*(t_f)) \le f(x(t_f))$ for all $(x, u) \in \mathscr{F} \cap \mathscr{B}_\varepsilon(x^*, u^*)$. ♢

Note that we allow discontinuous optimal controls, especially bang–bang controls. Therefore, we consider local solutions w.r.t. the L^1-norm for control functions. We suppose in the following:

(2.1) There exists $\bar{\varepsilon} > 0$ and $(x^*, u^*) \in \mathscr{F}$ such that $f(x^*(t_f)) \le f(x(t_f))$ for all $(x, u) \in \mathscr{F} \cap \mathscr{B}_{\bar{\varepsilon}}(x^*, u^*)$, i.e., (x^*, u^*) is a local solution of (OC).

Since U is bounded, there exists a constant K_u such that for all $u \in \mathscr{U}$

$$|u(t)| \le K_u \quad \text{a.e. on } [t_0, t_f]. \tag{2.2}$$

Let

$$\mathscr{N}_{\bar{\varepsilon}}(x^*) = \left\{ x \in X_1 \mid \|x - x^*\|_\infty < \bar{\varepsilon} \right\},$$

and let $\mathscr{B} \subset \mathbb{R}^n$ be a convex and open set such that

$$\mathscr{B} \supset \left\{ z \in \mathbb{R}^n \mid z = x(t) \text{ for some } \ t \in [t_0, t_f] \text{ and some } \ x \in \mathscr{N}_{\bar{\varepsilon}}(x^*) \right\}.$$

For given numbers $n_1, n_2 \in \mathbb{N}$, $n_1 \leq n_2$, we define

$$J_{n_1}^{n_2} = \{n_1, n_1 + 1, \ldots, n_2\}.$$

We suppose that the following assumptions are satisfied:

(2.3) The functions f, $g^{(1)}$, and $g^{(2)}$ are continuously differentiable w.r.t. x on $\mathscr{B}$.

(2.4) The functions f, $g^{(1)}$, and $g^{(2)}$ are Lipschitz continuous, i.e., there are constants L_f, $\tilde{L}_f$ and L_g such that

$$\begin{aligned}
&|f(x) - f(z)| \leq L_f\, |x - z|,\\
&|g^{(1)}(x,t) - g^{(1)}(z,s)| \leq L_g\, (|x - z| + |t - s|),\\
&\|g^{(2)}(x,t) - g^{(2)}(z,s)\| \leq L_g\, (|x - z| + |t - s|),
\end{aligned}$$

for all $s, t \in [t_0, t_f]$ and all $x, z \in \mathscr{B}$

(2.5) The functions f_x, $g_x^{(1)}$, and $g_x^{(2)}$ are Lipschitz continuous, i.e., there are constants $L_f^{(1)}$ and $L_g^{(1)}$ such that

$$\begin{aligned}
&|f_x(x) - f_x(z)| \leq L_f^{(1)}\, |x - z|,\\
&|g_{j,x}^{(1)}(x,t) - g_{j,x}^{(1)}(z,s)| \leq L_g^{(1)}\, (|x - z| + |t - s|), \ \ j \in J_1^n,\\
&|g_{ji,x}^{(2)}(x,t) - g_{ji,x}^{(2)}(z,s)| \leq L_g^{(1)}\, (|x - z| + |t - s|), \ \ j \in J_1^n, \ \ i \in J_1^m,
\end{aligned}$$

for all $s, t \in [t_0, t_f]$ and all $x, z \in \mathscr{B}$.

For $(x, u) \in X$ with $\|x - x^*\| \leq \bar{\varepsilon}$ it follows from (2.2) and (2.5) that

$$\begin{aligned}
|g(x(t), u(t), t)| &\leq |g(x^*(t), u^*(t), t)| + |g(x(t), u(t), t) - g(x^*(t), u^*(t), t)|\\
&\leq |g(x^*(t), u^*(t), t)| + |g^{(1)}(x(t), t) - g^{(1)}(x^*(t), t)|\\
&\quad + \|g^{(2)}(x(t), t) - g^{(2)}(x^*(t), t)\|\, |u(t)|\\
&\quad + \|g^{(2)}(x^*(t), t)\|\, |u(t) - u^*(t)|\\
&\leq |g(x^*(t), u^*(t), t)| + L_g \bar{\varepsilon}(1 + K_u) + 2K_u \|g^{(2)}(x^*(t), t)\|.
\end{aligned}$$

This implies

$$|g(x(t), u(t), t)| \leq K_g \tag{2.6}$$

with some constant K_g independent of $(x,u) \in X$ with $\|x - x^*\| \le \bar{\varepsilon}$. Moreover, for $(x,u) \in \mathscr{F}$ with $x \in \mathscr{N}_{\bar{\varepsilon}}(x^*)$ and $t, s \in [t_0, t_f]$ we have

$$|\dot{x}(t) - \dot{x}(s)| \le |g^{(1)}(x(t),t) - g^{(1)}(x(s),s)| + \|g^{(2)}(x(t),t) - g^{(2)}(x(s),s)\| |u(t)| + \|g^{(2)}(x(s),s)\| |u(t) - u(s)|.$$

By (2.2), Assumption (2.4), and (2.6) this implies

$$|\dot{x}(t) - \dot{x}(s)| \le L_g(1 + K_u)(|x(t) - x(s)| + |t - s|) + K_g\, |u(t) - u(s)|. \quad (2.7)$$

This further implies that with some constant L_x

$$|\dot{x}(t)| \le L_x \quad \forall' t \in [t_0, t_f], \quad (2.8)$$

for all $(x,u) \in \mathscr{F}$ with $x \in \mathscr{N}_{\bar{\varepsilon}}(x^*)$, which shows that the feasible trajectories $x \in \mathscr{N}_{\bar{\varepsilon}}(x^*)$ are uniformly Lipschitz with Lipschitz modulus L_x.

The *Hamiltonian* $\mathscr{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times [t_0, t_f] \to \mathbb{R}$ for Problem (OC) is defined by

$$\mathscr{H}(x,u,\lambda,t) = \lambda^{\mathsf{T}} g(x,u,t) = \sum_{j=1}^{n} \lambda_j \left[g_j^{(1)}(x,t) + \sum_{i=1}^{m} u_i g_{j,i}^{(2)}(x,t) \right].$$

We denote by

$$g_{.i}^{(2)}(x,t) = \left[g_{1i}^{(2)}(x,t), \ldots, g_{ni}^{(2)}(x,t) \right]^{\mathsf{T}}, \; i = 1, \ldots, m,$$

the i-th column vector of $g^{(2)}(x,t)$ and by

$$g_j^{(2)}(x,t) = \left[g_{j1}^{(2)}(x,t), \ldots, g_{jm}^{(2)}(x,t) \right], \; j = 1, \ldots, n,$$

the j-th row of $g^{(2)}(x,t)$. Then

$$\mathscr{H}_x(x,u,\lambda,t) = \lambda^{\mathsf{T}} g_x(x,t) = \lambda^{\mathsf{T}} \left[g_x^{(1)}(x,t) + \sum_{i=1}^{m} u_i g_{.i,x}^{(2)}(x,t) \right],$$
$$\mathscr{H}_u(x,u,\lambda,t) = \lambda^{\mathsf{T}} g^{(2)}(x,t) = \sum_{j=1}^{n} \lambda_j g_j^{(2)}(x,t).$$

Optimality conditions for Problem (OC) are well-known. Let $(x^*, u^*) \in \mathscr{F}$ be a local solution of (OC). Then there exists a function $\lambda^* \in W_1^1(t_0, t_f; \mathbb{R}^n)$ such that the adjoint equation

$$-\dot{\lambda}^*(t) = \mathscr{H}_x(x^*(t), u^*(t), \lambda^*(t), t)^{\mathsf{T}} = g_x(x^*(t), u^*(t), t)^{\mathsf{T}} \lambda^*(t) \quad (2.9)$$

is satisfied for a.a. $t \in [t_0, t_f]$ with terminal condition $\lambda^*(t_f) = f_x(x^*(t_f))^{\mathsf{T}}$, and the local minimum principle

$$\begin{aligned} &\mathscr{H}_u(x^*(t), u^*(t), \lambda^*(t), t)^{\mathsf{T}}(u - u^*(t)) \\ &\quad = \lambda^*(t)^{\mathsf{T}} g^{(2)}(x^*(t), t)(u - u^*(t)) \geq 0 \end{aligned} \tag{2.10}$$

holds for a.a. $t \in [t_0, t_f]$ and all $u \in U$.

We denote by $\sigma^* \colon [t_0, t_f] \to \mathbb{R}^m$ the *switching function* defined by

$$\sigma^*(t) = \mathscr{H}_u(x^*(t), u^*(t), \lambda^*(t), t)^{\mathsf{T}} = g^{(2)}(x^*(t), t)^{\mathsf{T}} \lambda^*(t). \tag{2.11}$$

For a strong local solution $(x^*, u^*) \in \mathscr{F}$ of Problem (OC) with associated adjoint function $\lambda^* \in X_1$, (2.10) implies for $i \in \{1, \ldots, m\}$

$$u_i^*(t) = \begin{cases} b_{l,i}, & \text{if } \sigma_i^*(t) > 0, \\ b_{u,i}, & \text{if } \sigma_i^*(t) < 0, \\ \text{undetermined}, & \text{if } \sigma_i^*(t) = 0. \end{cases} \tag{2.12}$$

Therefore, the optimal control u^* is of bang–bang type or may have singular arcs.

3 Euler Approximation

Let $N \in \mathbb{N}$, $h = h_N = (t_f - t_0)/N$ be the mesh size and $t_j = t_0 + jh$, $j \in J_0^N$, the grid points of the discretization. We approximate the space X_2 of controls by functions in the subspace $X_{2,N} \subset X_2$ 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, resp. λ_h, in the subspace $X_{1,N} \subset X_1$ of continuous, piecewise linear functions represented by their values $x_h(t_j) = x_{h,j}$, resp. $\lambda_h(t_j) = \lambda_{h,j}$, at the grid points t_j, $j \in J_0^N$. Then based on Euler's method for the discretization of the system equation we obtain the discrete optimal control problem

$$(\mathrm{OC})_N \qquad \min_{(x_h,u_h)\in X_{1,N}\times X_{2,N}} f(x_{h,N})$$

subject to

$$\begin{aligned} &x_{h,j+1} = x_{h,j} + h_N g(x_{h,j}, u_{h,j}, t_j), && j \in J_0^{N-1}, \\ &x_{h,0} = a, && \\ &u_j \in U, && j \in J_0^{N-1}. \end{aligned}$$

By $\mathscr{F}_N$ we denote the feasible set of $(\mathrm{OC})_N$.

Definition 2 A pair $(x_h^*, u_h^*) \in \mathscr{F}_N$ is called a *local minimizer* of f on $\mathscr{F}_N$ or a *local solution* of Problem $(\mathrm{OC})_N$, if there exists $\varepsilon > 0$ such that $f(x_{h,N}^*) \leq f(x_{h,N})$ for all $(x_h, u_h) \in \mathscr{F}_N \cap \mathscr{B}_\varepsilon(x_h^*, u_h^*)$. $\lozenge$

Since $\mathscr{F}_N$ is nonempty and bounded, Problem $(\mathrm{OC})_N$ has a (global) solution. Optimality conditions can be derived in the same way as in Ioffe and Tihomirov [22, Section 6.4]. For any local solution $(x_h^*, u_h^*) \in \mathscr{F}_N$ of Problem $(\mathrm{OC})_N$ there exists a multiplier λ_h^* such that the discrete adjoint equation

$$-\frac{\lambda_{h,j+1}^* - \lambda_{h,j}^*}{h_N} = H_x(x_{h,j}^*, u_{h,j}^*, \lambda_{h,j+1}^*, t_j)^\mathsf{T} = g_x(x_{h,j}^*, u_{h,j}^*, t_j)^\mathsf{T} \lambda_{h,j+1}^* \quad (3.1)$$

for $j \in J_0^{N-1}$ with terminal condition $\lambda_{h,N}^* = f_x(x_{h,N}^*)^\mathsf{T}$, and the discrete minimum principle

$$H_u(x_{h,j}^*, u_{h,j}^*, \lambda_{h,j+1}^*, t_j)(u - u_{h,j}^*) = (\lambda_{h,j+1}^*)^\mathsf{T} g^{(2)}(x_{h,j}^*, t_j)(u - u_{h,j}^*) \geq 0 \quad (3.2)$$

for $j \in J_0^{N-1}$ and all $u \in U$ are satisfied.

By λ_h^* we denote the continuous, piecewise linear function defined by the values $\lambda_h(t_j) = \lambda_{h,j}$, $i = 0, \dots, N$, and by $\sigma_h^*(t)$ we denote the continuous, piecewise constant function defined by the values

$$\sigma_h^*(t_j) := g^{(2)}(x_{h,j}^*, t_j)^\mathsf{T} \lambda_{h,j+1}^*, \quad j \in J_0^{N-1}, \quad (3.3)$$

the discrete analogue of the switching function (2.11). From (3.2) we obtain for $i = 1, \dots, m$, $j \in J_0^{N-1}$,

$$u_{h,i}^*(t_j) = \begin{cases} b_{l,i} & \text{if } \sigma_{h,i}^*(t_j) > 0, \\ b_{u,i} & \text{if } \sigma_{h,i}^*(t_j) < 0, \\ \text{undetermined} & \text{if } \sigma_{h,i}^*(t_j) = 0. \end{cases} \quad (3.4)$$

4 Error estimates for local minimizers

We first prove some auxiliary results. For a function $z\colon [t_0, t_f] \to \mathbb{R}^k$ of bounded variation and $s_1, s_2 \in [t_0, t_f]$, $s_1 < s_2$, we denote by $\mathsf{V}_{s_1}^{s_2} z$ the total variation of z on $[s_1, s_2]$.

Lemma 1 *Suppose that $u \in X_2$ has bounded variation, and let $u_h \in X_{2,N}$ be the piecewise constant function defined by the values $u_{h,j} = u(t_j)$, $j \in J_0^{N-1}$. Then*

$$\|u - u_h\|_1 \leq h_N \mathsf{V}_{t_0}^{t_f} u. \quad (4.1)$$

Proof Since for $s \in [t_j, t_{j+1}]$

$$|u(s) - u(t_j)| \leq |u(t_{j+1}) - u(s)| + |u(s) - u(t_j)| \leq \mathsf{V}_{t_j}^{t_{j+1}} u,$$

we have

$$\|u - u_h\|_1 = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |u(s) - u(t_j)|\, ds \le \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \mathsf{V}_{t_j}^{t_{j+1}} u$$
$$\le h_N \mathsf{V}_{t_0}^{t_f} u,$$

which proves (4.1). □

Remark 1 In many applications the optimal control u^* is a piecewise Lipschitz continuous function. In this case u^* has bounded variation. ♢

The following result is a special case of Sendov and Popov [39, Theorem 6.1].

Lemma 2 *If Assumptions* (2.3) *and* (2.4) *are satisfied,* $(x, u) \in \mathscr{F} \cap \mathscr{B}_\varepsilon(x^*, u^*)$, $\dot{x}$ *has bounded variation, and* x_h *is the solution of the discrete system equation*

$$x_{h,j+1} = x_{h,j} + h_N g(x_{h,j}, u(t_j), t_j), \ \ j \in J_0^{N-1}, x_{h,0} = a, \tag{4.2}$$

then

$$\max_{1 \le j \le N} |x_{h,j} - x(t_j)| \le c_1\, h_N \mathsf{V}_{t_0}^{t_f} \dot{x}, \tag{4.3}$$

where $c_1 = e^{(t_f - t_0)L_g(1+K_u)}$ *is a constant independent of* N.

Lemma 3 *Suppose that Assumptions* (2.1), (2.3), *and* (2.4) *are satisfied, and that* u^* *has bounded variation. Then for* $(x, u) \in \mathscr{F} \cap \mathscr{B}_\varepsilon(x^*, u^*)$ *we have*

$$\mathsf{V}_{t_0}^{t_f} \dot{x} \le L_g(1 + K_u)(L_x + 1)(t_f - t_0) + c_2 \mathsf{V}_{t_0}^{t_f} u \tag{4.4}$$

where c_2 *is a constant independent of* N.

Proof The variation of $\dot{x}$ can be estimated by the variation of the right hand side of the system equation. For $t, s \in [t_0, t_f]$ we have by (2.7)

$$|\dot{x}(t) - \dot{x}(s)| \le L_g(1 + K_u)(|x(t) - x(s)| + |t - s|) + K_g\, |u(t) - u(s)|.$$

Hence, by (2.8) we obtain

$$\mathsf{V}_{t_0}^{t_f} \dot{x} \le L_g(1 + K_u)(L_x + 1)(t_f - t_0) + K_g \mathsf{V}_{t_0}^{t_f} u,$$

which proves the assertion. □

For $\rho > 0$ we consider the auxiliary problem

$$
\begin{aligned}
(\mathrm{OC})_{N,\rho} \quad & \min_{(x_h,u_h)\in X_{1,N}\times X_{2,N}} f(x_{h,N}) \\
& \text{subject to} \\
& x_{h,j+1} = x_{h,j} + h_N g(x_{h,j}, u_{h,j}, t_j), \quad j \in J_0^{N-1}, \\
& x_{h,0} = a, \\
& u_j \in U, \qquad\qquad\qquad\qquad j \in J_0^{N-1}, \\
& \|x_h - x^*\|_\infty \le \rho, \ \|u_h - u^*\| \le \rho
\end{aligned}
$$

which is Problem $(\mathrm{OC})_N$ with the additional constraints

$$\|x_h - x^*\|_\infty \le \rho, \ \|u_h - u^*\|_1 \le \rho. \tag{4.5}$$

For $\rho > 0$ we denote by $\mathscr{F}_{N,\rho}$ the feasible set of Problem $(\mathrm{OC})_{N,\rho}$, i.e.

$$\mathscr{F}_{N,\rho} = \left\{(x_h, u_h) \in \mathscr{F}_N \mid \|x_h - x^*\|_\infty \le \rho, \ \|u_h - u^*\|_1 \le \rho \right\}.$$

Lemma 4 *Suppose that Assumptions* (2.1), (2.3)*, and* (2.4) *are satisfied, and that u^* has bounded variation. Further let $\rho > 0$ be arbitrary but fixed. Then for sufficiently large N Problem* $(\mathrm{OC})_{N,\rho}$ *has a solution.*

Proof Let $\hat{u}_h \in X_{2,N}$ be defined by the values $\hat{u}_{h,j} = u^*(t_j)$, $j \in J_0^{N-1}$. Then $\hat{u}_h \in \mathscr{U}$, and by Lemma 1 we have

$$\|u^* - \hat{u}_h\|_1 \le h_N \mathrm{V}_{t_0}^{t_f} u^*.$$

Let $\hat{x}_h$ be the solution of the discrete system Eq. (4.2) for $u = u^*$. By Lemmas 2 and 3 we have

$$\max_{1\le j\le N} |\hat{x}_{h,j} - x^*(t_j)| \le c_1 \left(L_g(1 + K_u)(L_x + 1)(t_f - t_0) + c_2 \mathrm{V}_{t_0}^{t_f} u^* \right) h_N.$$

This shows that $(\hat{x}_h, \hat{u}_h) \in \mathscr{F}_{N,\rho}$, and hence $\mathscr{F}_{N,\rho} \ne \emptyset$ for sufficiently large N. Since the cost functional is continuous and the feasible set is closed and bounded, a solution exists. □

The following result on the dependence of solutions of the system equation on parameters is well-known.

Lemma 5 *If Assumptions* (2.1), (2.3)*, and* (2.5) *are satisfied, there exists $\bar{\rho} \in]0, \bar{\varepsilon}]$ such that for each $u \in X_2$ with $\|u - u^*\|_1 < \bar{\rho}$, and each $\eta \in L^1(t_0, t_f; \mathbb{R}^n)$ with $\|\eta\|_1 < \bar{\rho}$ the perturbed system equation*

$$\dot{x}(t) = g(x(t), u(t), t) + \eta(t) \ \text{a.e. on } [t_0, t_f], \ x(t_0) = a,$$

has a unique solution $x = x(u, \eta)$, *and for* $x_i = x(u_i, \eta_i)$, $i = 1, 2$, *we have*

$$\|x_2 - x_1\|_{1,1} \leq c_s \left(\|u_2 - u_1\|_1 + \|\eta_2 - \eta_1\|_1\right),$$

where the constant c_s *is independent of* u *and* η.

We further need the following auxiliary result.

Lemma 6 *Suppose that Assumptions* (2.1), (2.3), (2.4), *and* (2.5) *are satisfied, and let* $\bar{\rho} > 0$ *be given by Lemma* 5. *Then there is a number* $\bar{N} \in \mathbb{N}$ *such that for* $N \geq \bar{N}$ *and any* $(x_h, u_h) \in \mathscr{F}_{N,\bar{\rho}}$ *there exists a function* $z \in X_1$ *such that* $(z, u_h) \in \mathscr{F}$ *and*

$$\|z - x_h\|_{1,1} \leq c h_N \tag{4.6}$$

with a constant c *independent of* N *and* $(x_h, u_h) \in \mathscr{F}_{N,\bar{\rho}}$.

Proof Let $(x_h, u_h) \in \mathscr{F}_{N,\bar{\rho}}$ and $N \in \mathbb{N}$ be given. Since $\|x_h - x^*\| \leq \bar{\rho} \leq \bar{\varepsilon}$ it follows from (2.6) that

$$|g(x_h(t_j), u_h(t_j), t_j)| \leq K_g. \tag{4.7}$$

By Lemma 5 the system equation of (OC) for $u = u_h$, i.e.,

$$\dot{z}(t) = g(z(t), u_h(t), t) \text{ a.e. on } [t_0, t_f], \quad z(t_0) = a,$$

has a unique solution z, i.e. $(z, u_h) \in \mathscr{F}$. We define the piecewise constant function $\bar{x}_h : [t_0, t_f] \to \mathbb{R}^n$ by $\bar{x}_h(t_j) = x_h(t_j)$ for $j \in J_0^{N-1}$. Then the discrete system Eq. (4.2) for x_h can be written in the form

$$\dot{x}_h(t) = g(x_h(t), u_h(t), t) + \eta(t) \text{ a.e. on } [t_0, t_f], \quad x_h(t_0) = a,$$

where

$$\eta(t) = g^{(1)}(\bar{x}_h(t), t) - g^{(1)}(x_h(t), t) + (g^{(2)}(\bar{x}_h(t), t) - g^{(2)}(x_h(t), t))u_h(t).$$

Since x_h solves the discrete system Eq. (4.2) we have by (4.7) for $t \in [t_j, t_{j+1}[$,

$$\begin{aligned} |\bar{x}_h(t) - x_h(t)| &= \left|(t - t_j)\frac{x(t_{j+1}) - x(t_j)}{t_{j+1} - t_j}\right| \\ &= |(t - t_j)|\,|g(x_h(t_j), u_h(t_j), t_j)| \leq h_N K_g, \end{aligned}$$

and by (2.2), Assumption (2.4), and (2.8) it follows that for $t \in [t_0, t_f]$,

$$|\eta(t)| \leq L_g(1 + K_u)|\bar{x}_h(t) - x_h(t)| \leq L_g(1 + K_u)K_g h_N.$$

We choose $\bar{N} \in \mathbb{N}$ such that

$$(t_f - t_0)L_g(1 + K_u)K_g h_N < \bar{\rho}$$

for $N \geq \bar{N}$. Then

$$\|\eta\|_1 \leq (t_f - t_0)|\eta(t)|_\infty < \bar{\rho}$$

for $N \geq \bar{N}$. By Lemma 5 this implies

$$\|x_h - z\|_{1,1} \leq c_s \,\|\eta\|_1 \leq c_s L_g(1 + K_u)K_g(t_f - t_0)h_N,$$

which proves (4.6). □

In order to obtain error estimates for local solutions we proceed similarly to Alt [1,2] (compare also Alt et al. [6]). In addition to (2.1) we use the following growth condition for the cost functional:

(4.8) There exist $\alpha > 0$, $\kappa \geq 1$ such that

$$f(x(t_f)) - f(x^*(t_f)) \geq \alpha \,\|u - u^*\|_1^\kappa$$

for all $(x, u) \in \mathscr{F} \cap \mathscr{B}_{\bar{\varepsilon}}(x^*, u^*)$.

Remark 2 The growth condition required here implies that (x^*, u^*) is a strict local solution of (OC). Such conditions are closely related to second-order optimality conditions (see e.g. Ioffe and Tihomirov [22, Chapter 7] or Maurer and Zowe [31, Theorem 5.6]). In the following section we use the stronger second-order optimality condition (5.7) implying (4.8) with $\kappa = 2$ (see Theorem 3).

Theorem 1 *Let Assumptions* (2.1), (2.3), (2.4), (2.5), *and* (4.8) *be satisfied and suppose that u^* has bounded variation. Then for each $0 < \rho < \bar{\rho}$, where $\bar{\rho} > 0$ is given by Lemma 5, Problem* $(\mathrm{OC})_{N,\rho}$ *has a global solution for sufficiently large N. Further for each such solution (x_h^*, u_h^*) the estimates*

$$\|u_h^* - u^*\|_1 \leq c_u h_N^{\frac{1}{\kappa}}, \quad \|x_h^* - x^*\|_{1,1} \leq c_x h_N^{\frac{1}{\kappa}} \tag{4.9}$$

hold with constants c_u, c_x independent of N and the solution (x_h^*, u_h^*).

Proof We choose $N \geq \bar{N}$ sufficiently large, where $\bar{N}$ is defined by Lemma 6. Then by Lemma 4 Problem $(\mathrm{OC})_{N,\rho}$ has a (global) solution. Let (x_h^*, u_h^*) be any such solution. By Lemma 6 there exists a function $z^* \in X_1$, such that $(z^*, u_h^*) \in \mathscr{F}$ and

$$\|z^* - x_h^*\|_{1,1} \leq c_1 h_N \tag{4.10}$$

with a constant c_1 independent of N and (x_h^*, u_h^*). Further, since $(x_h^*, u_h^*) \in \mathscr{F}_{N,\rho}$ we have $\|x_h^* - x^*\|_\infty \leq \rho$. Together with (4.10) and the fact that $\|z^* - x_h^*\|_\infty \leq \|z^* - x_h^*\|_{1,1}$ this implies

$$\|z^* - x^*\|_\infty \leq \|z^* - x_h^*\|_\infty + \|x_h^* - x^*\|_\infty \leq c_1 h_N + \rho$$

and therefore

$$\|z^* - x^*\|_\infty < \bar{\rho} < \bar{\varepsilon} \tag{4.11}$$

for sufficiently large N. Since $(x_h^*, u_h^*) \in \mathscr{F}_{N,\rho}$ we have $(z^*, u_h^*) \in \mathscr{F} \cap \mathscr{B}_{\bar{\varepsilon}}(x^*, u^*)$. By (4.8) we therefore have

$$f(z^*(t_f)) - f(x^*(t_f)) \geq \alpha \, \|u_h^* - u^*\|_1^\kappa.$$

Further, by (2.4) and (4.10) we have

$$f(z^*(t_f)) = f(x_h^*(t_f)) + f(z^*(t_f)) - f(x_h^*(t_f)) \leq f(x_h^*(t_f)) + L_f c_1 h_N,$$

and therefore

$$\alpha \, \|u_h^* - u^*\|_1^\kappa \leq f(x_h^*(t_f)) - f(x^*(t_f)) + L_f c_1 h_N \tag{4.12}$$

for N sufficiently large.

Let $\hat{u}_h \in X_{2,N}$ be defined as in the proof of Lemma 4. Then for sufficiently large N we have $(\hat{x}_h, \hat{u}_h) \in \mathscr{F}_{N,\rho}$ (see proof of Lemma 4) and therefore $f(\hat{x}_h(t_f)) \geq f(x_h^*(t_f))$. Further we have

$$\max_{1 \leq j \leq N} |\hat{x}_h(t_j) - x^*(t_j)| \leq c_2 h_N \tag{4.13}$$

with a constant c_2 independent of N. By (4.12), (2.4) this implies

$$\alpha \, \|u_h^* - u^*\|_1^\kappa \leq f(\hat{x}_h(t_f)) - f(x^*(t_f)) + L_f c_1 h_N \leq L_f (c_1 + c_2) h_N. \tag{4.14}$$

In the proof of Lemma 6 we have shown that the discrete system Eq. (4.2) for $x_h = x_h^*$ can be written in the form

$$\dot{x}_h(t) = g(x_h(t), u_h(t), t) + \eta(t) \text{ a.e. on } [t_0, t_f], \quad x_h(t_0) = a,$$

where $|\eta(t)| \leq c_3 h_N$ with a constant c_3 independent of N. By Lemma 5 we therefore obtain

$$\|x_h^* - x^*\|_{1,1} \leq c_4 \left(\|u_h^* - u^*\|_1 + \|\eta\|_1 \right) \leq c_4 \left(\|u_h^* - u^*\|_1 + c_3 (t_f - t_0) h_N \right), \tag{4.15}$$

where the constant c_4 is independent of u and N. □

Remark 3 Note that Theorem 1 assumes that (x_h^*, u_h^*) is a global solution of Problem $(\mathrm{OC})_{N,\rho}$. For such a solution we have $\|u_h^* - u^*\|_1 < \rho$ and $\|x_h^* - x^*\|_{1,1} < \rho$ for sufficiently large N, i.e. the additional constraints (4.5) are not active, and (x_h^*, u_h^*) is a local minimizer of Problem $(\mathrm{OC})_N$. Similar results on the existence of approximate local minimizers for control problems obtained by Euler discretization and error estimates for the discrete solutions are well-known in case that the optimal control is continuous (see e.g. Malanowski et al. [27], Dontchev and Hager [11], Dontchev et al. [12,13]). In these papers a strong second-order sufficient optimality condition is used which also implies local uniqueness of the discrete solutions. This can not be shown under the weaker condition (4.8) used here. ♢

If (x_h^*, u_h^*) is a global solution of Problem $(\mathrm{OC})_{N,\rho}$, then by Remark 3 (x_h^*, u_h^*) is a local minimizer of Problem $(\mathrm{OC})_N$. Therefore a multiplier λ_h^* exists satisfying the discrete adjoint Eq. (3.1). In order to derive an error estimate for this multiplier we need some auxiliary results. Since the adjoint equation is a linear differential equation one easily obtains the following result.

Lemma 7 *Suppose that Assumptions* (2.1), (2.3), (2.4)*, and* (2.5) *are satisfied. Let* $\bar\rho > 0$ *be given by Lemma* 5 *and* $0 < \rho \le \bar\rho$*. Then if* N *is sufficiently large we have for any solution* (x_h^*, u_h^*) *of Problem* $(\mathrm{OC})_{N,\rho}$ *and the associated adjoint function* λ_h^* *the estimate*

$$\|\lambda_h^*\|_\infty \le K_\lambda \tag{4.16}$$

with a constant K_λ *independent of* N *and the solution* (x_h^*, u_h^*)*.*

In the proof of Lemma 6 we have shown that the discrete state variables can be viewed as the solution of a perturbation of the system equation of Problem (OC). In the same way one can show that the discrete adjoint variables λ_h^* can be viewed as the solution of a perturbation of the adjoint Eq. (2.9).

Lemma 8 *Suppose that Assumptions* (2.1), (2.3), (2.4)*, and* (2.5) *are satisfied. Let* $\bar\rho > 0$ *be given by Lemma* 5 *and* $0 < \rho \le \bar\rho$*. Then, if* N *is sufficiently large, we can write the discrete adjoint equation* (3.1) *in the form*

$$-\dot\lambda_h^*(t) = g_x(x_h^*(t), u_h^*(t), t)^{\mathsf{T}} \lambda_h^*(t) + \xi_h(t) \tag{4.17}$$

for a.a. $t \in [t_0, t_f]$*, where the function* $\xi_h \colon [t_0, t_f] \to \mathbb{R}^n$ *can be estimated by*

$$|\xi_h(t)| \le c_\xi h_N \tag{4.18}$$

for a.a. $t \in [t_0, t_f]$ *with a constant* c_ξ *independent of* N *and the solution* (x_h^*, u_h^*)*.*

Now we can derive an error estimate for the discrete adjoint functions.

Theorem 2 *Let the assumptions of Theorem* 1 *be satisfied and suppose that* u^* *has bounded variation. Then for each* $0 < \rho < \bar\rho$*, where* $\bar\rho > 0$ *is given by Lemma* 5*, Problem* $(\mathrm{OC})_{N,\rho}$ *has a (global) solution for sufficiently large* N*. Further for each such solution* (x_h^*, u_h^*) *and the associated adjoint function* λ_h^* *the estimate*

$$\|\lambda_h^* - \lambda^*\|_1 \le c_\lambda h_N^{\frac{1}{\kappa}} \tag{4.19}$$

holds with a constant c_λ *independent of* N *and the solution* (x_h^*, u_h^*)*.*

Proof We define $\lambda := \lambda_h^* - \lambda^*$. By (2.9) and (4.17) we have

$$\begin{aligned}
-\dot{\lambda}(t) &= -\dot{\lambda}_h^*(t) + \dot{\lambda}^*(t) \\
&= \left[g_x^{(1)}(x_h^*(t), t) + \sum_{i=1}^{m} u_h^*(t)_i g_{.i,x}^{(2)}(x_h^*(t), t) \right]^{\mathsf{T}} \lambda_h^*(t) + \xi_h(t) \\
&\quad - \left[g_x^{(1)}(x^*(t), t) + \sum_{i=1}^{m} u^*(t)_i g_{.i,x}^{(2)}(x^*(t), t) \right]^{\mathsf{T}} \lambda^*(t) \\
&= \left[g_x^{(1)}(x_h^*(t), t) + \sum_{i=1}^{m} u_h^*(t)_i g_{.i,x}^{(2)}(x_h^*(t), t) \right. \\
&\quad \left. - g_x^{(1)}(x^*(t), t) - \sum_{i=1}^{m} u^*(t)_i g_{.i,x}^{(2)}(x^*(t), t) \right]^{\mathsf{T}} \lambda_h^*(t) + \xi_h(t) \\
&\quad + \left[g_x^{(1)}(x^*(t), t) + \sum_{i=1}^{m} u^*(t)_i g_{.i,x}^{(2)}(x^*(t), t) \right]^{\mathsf{T}} \lambda(t)
\end{aligned}$$

with terminal condition $\lambda(t_f) = f_x(x_{h,N}^*)^{\mathsf{T}} - f_x(x^*(t_f))^{\mathsf{T}}$. Since this is a linear differential equation it follows that

$$\|\lambda\|_{1,1} \le c_{\lambda,1} \left(\|x_h^* - x^*\|_{1,1} + \|u_h^* - u^*\|_1 + h_N \right)$$

with some constant $\bar{c}_{\lambda,1}$ independent of N and (x_h^*, u_h^*)). Finally, together with (4.15) we obtain

$$\|\lambda_h^* - \lambda^*\|_{1,1} \le c_{\lambda,2} \left(\|u_h^* - u^*\|_1 + h_N \right) \tag{4.20}$$

with a constant $c_{\lambda,2}$ independent of N and the solution (x_h^*, u_h^*). By Theorem 1 this implies (4.19). □

5 Improved error estimates

We can improve the error estimates of the last section to order 1, if we replace condition (4.8) by a stronger second-order sufficient optimality condition. To this end we require in addition to Assumptions (2.1), (2.3), (2.4), and (2.5):

(5.1) The functions f, $g^{(1)}$, and $g^{(2)}$ are twice continuously differentiable w.r.t. x on $\mathscr{B}$.

(5.2) The functions f_{xx}, $g_{xx}^{(1)}$, and $g_{xx}^{(2)}$ are Lipschitz continuous, i.e., there are constants $L_f^{(2)}$ and $L_g^{(1)}$ such that for all $s, t \in [t_0, t_f]$ and all $x, z \in \mathscr{B}$

$$|f_{xx}(x) - f_{xx}(z)| \leq L_f^{(2)}\,|x - z|,$$
$$|g_{j,xx}^{(1)}(x,t) - g_{j,xx}^{(1)}(z,s)| \leq L_g^{(2)}\,(|x - z| + |t - s|),\ \ j \in J_1^n,$$
$$|g_{ji,xx}^{(2)}(x,t) - g_{ji,xx}^{(2)}(z,s)| \leq L_g^{(2)}\,(|x - z| + |t - s|),\ \ j \in J_1^n,\ \ i \in J_1^m.$$

Remark 4 Assumptions (5.1), (5.2) imply Assumptions (2.3), (2.4), and (2.5). ◊

As in Alt [2, Section 6] we can formulate Problem (OC) as an abstract optimization problem of type

$$\min_{z \in X} F(z) \quad \text{s.t.} \quad z \in C, \quad G(z) \in K,$$

where $z = (x, u) \in X$, $F\colon X \to \mathbb{R}$ is defined by

$$F(z) = F(x, u) = f(x(t_f)),$$

$G\colon X \to Y := L^1(t_0, t_f; \mathbb{R}^n) \times \mathbb{R}^n$ is defined by

$$G(z)(t) = G(x, u)(t) = \begin{pmatrix} g(x(t), u(t), t) - \dot{x}(t) \\ x(t_0) - a \end{pmatrix},$$

and $C = X_1 \times \mathscr{U}$, $K = \{0_Y\}$. As shown in Alt [2] it then follows by results of Robinson [37,38] that the set

$$\begin{aligned} T(x^*, u^*) &= \{(x, u) \in C \mid G(x^*, u^*) + G'(x^*, u^*)\left((x, u) - (x^*, u^*)\right) \in K\} \\ &= \{(x, u) \mid (x, u) \in X, u \in \mathscr{U}, x(t_0) = a, \\ \dot{x} - \dot{x}^* &= g_x(x^*(\cdot), u^*(\cdot), \cdot)(x - x^*) + g_u(x^*(\cdot), u^*(\cdot), \cdot)(u - u^*)\} \end{aligned}$$

approximates the feasible set of Problem (OC) in the sense of Maurer and Zowe [31, Definition 4.1]. From Alt [2, Lemma 2.1] and Lemma 5 we therefore get the following result:

Lemma 9 *Let Assumptions* (2.1), (2.3), (2.4)*, and* (2.5) *be satisfied. Then for each* $\gamma > 0$ *there exists* $\rho(\gamma) > 0$ *such that for each* $(x, u) \in \mathscr{F}$ *with* $\|u - u^*\|_1 < \rho(\gamma)$ *there exists* $(\bar{x}, \bar{u}) \in T(x^*, u^*)$ *with*

$$\|x - \bar{x}\|_{1,1} + \|u - \bar{u}\|_1 \leq \gamma\left(\|x - x^*\|_{1,1} + \|u - u^*\|_1\right).$$

For $\lambda \in L^\infty(t_0, t_f; \mathbb{R}^n)$ we define the *Lagrange function* by

$$\mathscr{L}(x, u, \lambda) = f(x(t_f)) + \int_{t_0}^{t_f} \lambda(t)^{\mathsf{T}}\,[g(x(t), u(t), t) - \dot{x}(t)]\,\mathrm{d}t.$$

It follows from the adjoint Eq. (2.9) that

$$\begin{aligned}\mathscr{L}_x(x^*, u^*, \lambda^*)(x) &= f_x\left(x^*(t_f)\right) x(t_f)\\ &\quad + \int_{t_0}^{t_f} \lambda^*(t)^{\mathsf{T}} \left[g_x(x^*(t), u^*(t), t)x(t) - \dot{x}(t)\right] \mathrm{d}t = 0 \end{aligned}\tag{5.3}$$

for all $x \in X_1$ with $x(t_0) = 0$, and by the local minimum principle (2.10) we have

$$\begin{aligned}\mathscr{L}_u(x^*, u^*, \lambda^*)(u - u^*) &= \int_{t_0}^{t_f} \lambda^*(t)^{\mathsf{T}} g_u(x^*(t), u^*(t), t)\left(u(t) - u^*(t)\right) \mathrm{d}t\\ &= \int_{t_0}^{t_f} \sigma^*(t)^{\mathsf{T}} \left(u(t) - u^*(t)\right) \mathrm{d}t \geq 0\end{aligned}\tag{5.4}$$

for all $u \in \mathscr{U}$, where σ^* is the switching function defined by (2.11).

By $\mathscr{L}''$ we denote the second derivate of $\mathscr{L}$ w.r.t. (x, u). Since the control u appears only linearly in Problem (OC), we have

$$\mathscr{H}_{uu}(x, u, \lambda, t) = 0 \quad \text{for all } (x, u, \lambda, t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times [t_0, t_f], \tag{5.5}$$

and therefore

$$\begin{aligned}\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((x_1, u_1), (x_2, u_2)\right) &= x_1(t_f)^{\mathsf{T}} f_{xx}\left(\bar{x}(t_f)\right) x_2(t_f)\\ &\quad + \int_{t_0}^{t_f} x_1(t)^{\mathsf{T}} \mathscr{H}_{xx}(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t), t)x_2(t)\, \mathrm{d}t\\ &\quad + \int_{t_0}^{t_f} x_1(t)^{\mathsf{T}} \mathscr{H}_{xu}(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t), t)u_2(t)\, \mathrm{d}t\\ &\quad + \int_{t_0}^{t_f} u_1(t)^{\mathsf{T}} \mathscr{H}_{ux}(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t), t)x_2(t)\, \mathrm{d}t\end{aligned}$$

for all $(\bar{x}, \bar{u}, \bar{\lambda}), (x, u, \lambda) \in X \times X_1$.

If (5.1) is satisfied, then there exists a constant $C_{\mathscr{L}}$ such that

$$\begin{aligned}&|\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((x_1, u_1), (x_2, u_2)\right)|\\ &\quad \leq C_{\mathscr{L}}\left(\|x_1\|_\infty \|x_2\|_\infty + \|x_1\|_\infty \|u_2\|_1 + \|x_2\|_\infty \|u_1\|_1\right)\end{aligned}\tag{5.6}$$

for all $(\bar{x}, \bar{u}, \bar{\lambda}) \in X \times X_1$ with $\|\bar{x} - x^*\|_\infty < \bar{\rho}$, $\|\bar{\lambda} - \lambda^*\|_\infty < \bar{\rho}$, $\bar{u} \in \mathscr{U}$, and all $(x_1, u_1), (x_2, u_2) \in X$.

In case of a continuous optimal control convergence of order 1 can be shown for Euler approximation if a strong second-order optimality condition is satisfied which especially requires that the bilinear form $\mathscr{L}''(x^*, u^*, \lambda^*)$ is positive definite w.r.t. to the control function (compare e.g. Dontchev et al. [13]). By (5.5) this condition cannot be satisfied for the class of control problems considered here. We use instead a second-order condition for the switching function σ^* defined by (2.11). This condition has been introduced by Felgenhauer [15] (see also Maurer and Osmolovskii [30], Maurer et al. [29]) and has been used e.g. in Alt et al. [6], Alt and Seydenschwanz [7], Seydenschwanz [40] to investigate Euler discretization of linear quadratic control problems:

(5.7) There exists $\bar{\alpha} > 0$ such that

$$\int_{t_0}^{t_f} \sigma^*(t)^{\mathsf{T}} \left(u(t) - u^*(t)\right) \mathrm{d}t = \mathscr{L}_u(x^*, u^*, \lambda^*)(u - u^*) \geq \bar{\alpha} \, \|u - u^*\|_1^2,$$

for all $u \in \mathscr{U}$.

Remark 5 One should note that Assumption (5.7) excludes singular arcs of the optimal control, i.e., the optimal control u^* must be of bang–bang type. As shown in Alt et al. [6, Lemma 4] the assumption is satisfied if the optimal control is of bang–bang type with finitely many boundary arcs and if an additional growth condition for the switching function around its zeros holds. ◊

In Alt and Seydenschwanz [7] and Alt et al. [6] we used an additional assumption ensuring convexity of the linear-quadratic control problems considered there. Here we use the somewhat weaker assumption:

(5.8) There exists $\beta > 0$ such that $\alpha := \bar{\alpha} - \beta > 0$ and

$$\begin{aligned} &z(t_f)^{\mathsf{T}} f_{xx}\left(x^*(t_f)\right) z(t_f) + \int_{t_0}^{t_f} z(t)^{\mathsf{T}} \mathscr{H}_{xx}\left(x^*(t), u^*(t), \lambda^*(t), t\right) z(t)\, \mathrm{d}t \\ &+ 2\int_{t_0}^{t_f} z(t)^{\mathsf{T}} \mathscr{H}_{xu}\left(x^*(t), u^*(t), \lambda^*(t), t\right) v(t)\, \mathrm{d}t \\ &= \mathscr{L}''(x^*, u^*, \lambda^*)\left((z, v), (z, v)\right) \geq -\beta \, \|v\|_1^2 \end{aligned}$$

for all $(z, v) = (x, u) - (x^*, u^*)$ with $(x, u) \in T(x^*, u^*)$.

Remark 6 The condition $(z, v) = (x, u) - (x^*, u^*)$ with $(x, u) \in T(x^*, u^*)$ is equivalent to $u \in \mathscr{U}$, $z(t_0) = 0$ and

$$\dot{z}(t) = g_x(x^*(t), u^*(t), t)z(t) + g_u(x^*(t), u^*(t), t)v(t)$$

for a.a. $t \in [t_0, t_f]$. Therefore, $\|z\|_{1,1} \leq c\, \|v\|_1$ with a constant c independent of v. ◊

Example 1 It can easily be seen that Assumption (6) used in Alt et al. [6] for a class of linear quadratic control problems is equivalent to (5.8).

If the system equation is linear then the coercivity condition in (5.8) reads

$$z(t_f)^{\mathsf{T}} f_{xx}\left(x^*(t_f)\right) z(t_f) \geq -\beta \, \|v\|_1^2$$

for all $(z, v) = (x, u) - (x^*, u^*)$ with $(x, u) \in T(x^*, u^*)$. ◊

We now show, that Assumptions (5.7) and (5.8) imply the growth condition (4.8) with $\kappa = 2$. The proof is based on a result of Ioffe and Tihomirov [22, Chapter 7] concerning a general second-order sufficient optimality condition for equality constrained optimization problems. More general results can be found in Maurer [28], and Maurer and Zowe [31] (see also Alt [2]). A general result on sufficient optimality conditions for optimal control problems can be found in Felgenhauer [15]. We need some auxiliary results which are modifications of corresponding results in Sect. 3 of Alt [2].

Lemma 10 *Let Assumptions* (2.1), (5.1), (5.2), (5.7)*, and* (5.8) *be satisfied. Then there exists* $0 < \delta_1 \le \bar{\rho}$ *such that*

$$\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\,((z, v), (z, v)) \ge -\left(\beta + \frac{\alpha}{4}\right) \|v\|_1^2$$

for all $(z, v) = (x, u) - (x^*, u^*)$ *with* $(x, u) \in T(x^*, u^*)$ *and all* $(\bar{x}, \bar{u}, \bar{\lambda}) \in X \times X_1$ *with* $\|\bar{x} - x^*\|_\infty + \|\bar{u} - u^*\|_1 + \|\bar{\lambda} - \lambda^*\|_\infty < \delta_1$.

Proof Let $(z, v) = (x, u) - (x^*, u^*)$ with $(x, u) \in T(x^*, u^*)$ and $(\bar{x}, \bar{u}, \bar{\lambda}) \in X \times X_1$ with $\|\bar{x} - x^*\|_\infty + \|\bar{u} - u^*\|_1 + \|\bar{\lambda} - \lambda^*\|_\infty < \bar{\varepsilon}$. By Assumption (5.8) we have

$$\begin{aligned}
\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\,((z, v), (z, v)) &= \mathscr{L}''(x^*, u^*, \lambda^*)\,((z, v), (z, v)) \\
&\quad + \mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\,((z, v), (z, v)) - \mathscr{L}''(x^*, u^*, \lambda^*)\,((z, v), (z, v)) \\
&\ge -\beta\,\|v\|_1^2 + \mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\,((z, v), (z, v)) - \mathscr{L}''(x^*, u^*, \lambda^*)\,((z, v), (z, v)),
\end{aligned}$$

i.e.,

$$\begin{aligned}
\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\,((z, v), (z, v)) + \beta\,\|v\|_1^2 &\ge z(t_f)^{\mathsf{T}} \left[f_{xx}\left(\bar{x}(t_f)\right) - f_{xx}\left(x^*(t_f)\right)\right] z(t_f) \\
&\quad + \int_{t_0}^{t_f} z(t)^{\mathsf{T}} \left[\mathscr{H}_{xx}(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t), t) - \mathscr{H}_{xx}(x^*(t), u^*(t), \lambda^*(t), t)\right] z(t)\,\mathrm{d}t \\
&\quad + 2\int_{t_0}^{t_f} z(t)^{\mathsf{T}} \left[\mathscr{H}_{xu}(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t), t) - \mathscr{H}_{xu}(x^*(t), u^*(t), \lambda^*(t), t)\right] v(t)\,\mathrm{d}t.
\end{aligned} \tag{5.9}$$

By (5.2) the absolute value of the first term on the right hand side of this inequality can be estimated by

$$c_1 \|\bar{x} - x^*\|_\infty \|z\|_\infty^2 \tag{5.10}$$

with some constant c_1 independent of z, v, $\bar{x}$, $\bar{u}$, and $\bar{\lambda}$. Using

$$\begin{aligned}
&\mathscr{H}_{xx}(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t), t) - \mathscr{H}_{xx}(x^*(t), u^*(t), \lambda^*(t), t) \\
&\quad = \sum_{j=1}^{n} \bar{\lambda}(t)_j \left[g^{(1)}_{j,xx}(\bar{x}(t), t) - g^{(1)}_{j,xx}(x^*(t), t) \right] \\
&\quad\quad + \sum_{j=1}^{n} \left[\bar{\lambda}(t)_j - \lambda^*(t)_j \right] g^{(1)}_{j,xx}(x^*(t), t) \\
&\quad\quad + \sum_{j=1}^{n} \bar{\lambda}(t)_j \sum_{i=1}^{m} \bar{u}(t)_i \left[g^{(2)}_{ji,xx}(\bar{x}(t), t) - g^{(2)}_{ji,xx}(x^*(t), t) \right] \\
&\quad\quad + \sum_{j=1}^{n} \bar{\lambda}(t)_j \sum_{i=1}^{m} \left[\bar{u}(t)_i - u^*(t)_i \right] g^{(2)}_{ji,xx}(x^*(t), t) \\
&\quad\quad + \sum_{j=1}^{n} \left[\bar{\lambda}(t)_j - \lambda^*(t)_j \right] \sum_{i=1}^{m} u^*(t)_i g^{(2)}_{ji,xx}(x^*(t), t).
\end{aligned}$$

the absolute value of the second term on the right hand side of (5.9) can be estimated by

$$c_2 \left(\|\bar{x} - x^*\|_\infty + \|\bar{u} - u^*\|_1 + \|\bar{\lambda} - \lambda^*\|_\infty \right) \|z\|_\infty^2 \tag{5.11}$$

with some constant c_2 independent of z, v, $\bar{x}$, $\bar{u}$, and $\bar{\lambda}$. In the same way it can be shown that the absolute value of the third term on the right hand side of (5.9) can be estimated by

$$c_3 \left(\|\bar{x} - x^*\|_\infty + \|\bar{\lambda} - \lambda^*\|_\infty \right) \|z\|_\infty \|v\|_1 \tag{5.12}$$

with some constant c_3 independent of z, v, $\bar{x}$, $\bar{u}$, and $\bar{\lambda}$. Since z satisfies the linear differential equation

$$\dot{z}(t) = g_x(x^*(t), u^*(t), t) z(t) + g_u(x^*(t), u^*(t), t) \left(u(t) - u^*(t) \right) \text{ a.e. on } [t_0, t_f]$$

with initial condition $z(t_0) = 0$ we have

$$\|z\|_\infty = \|x - x^*\|_\infty \le \|x - x^*\|_{1,1} \le c_4 \|u - u^*\|_1 \tag{5.13}$$

with a constant c_4 independent of x and u. Now combining (5.10)–(5.13) the absolute value of the right hand side of (5.9) can be estimated by

$$c_5 \left(\|\bar{x} - x^*\|_\infty + \|\bar{u} - u^*\|_1 + \|\bar{\lambda} - \lambda^*\|_\infty \right) \|v\|_1^2.$$

The assertion then follows if we choose $\delta_1 > 0$ small enough. □

Lemma 11 *Let Assumptions* (2.1), (5.1), (5.2), (5.7)*, and* (5.8) *be satisfied. Then there exists* $0 < \delta_2 \leq \delta_1$ *such that*

$$\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((x - x^*, u - u^*), (x - x^*, u - u^*)\right) \geq -\left(\beta + \frac{\alpha}{2}\right) \|u - u^*\|_1^2$$

for all $(x, u) \in \mathscr{F}$ *with* $\|u - u^*\|_1 < \delta_2$ *and all* $(\bar{x}, \bar{u}, \bar{\lambda}) \in X \times X_1$ *with* $\|\bar{x} - x^*\|_\infty + \|\bar{u} - u^*\|_1 + \|\bar{\lambda} - \lambda^*\|_\infty < \delta_2$.

Proof We choose $\gamma > 0$ such that

$$\begin{aligned}&\left(\beta + \frac{\alpha}{4}\right)(1 + \gamma(1 + c_s))^2 + C_{\mathscr{L}}\gamma(1 + c_s)\left(1 + (1 + 2\gamma)(1 + c_s)\right)\\&\quad + 3C_{\mathscr{L}}\gamma^2(1 + c_s)^2 \leq \left(\beta + \frac{\alpha}{2}\right),\end{aligned} \tag{5.14}$$

where $C_{\mathscr{L}}$ is defined by (5.6) and c_s is the constant defined by Lemma 5. Let $\rho(\gamma)$ be defined by Lemma 9. We choose $0 < \delta_2 \leq \delta_1$ such that

$$(1 + c_s)\delta_2 < \max\{\bar{\rho}, \rho(\gamma)\}.$$

Then for $(x, u) \in \mathscr{F}$ with $\|u - u^*\|_1 < \delta_2$ we have by Lemma 5

$$\|x - x^*\|_{1,1} + \|u - u^*\|_1 \leq c_s\|u - u^*\|_1 + \|u - u^*\|_1 < \max\{\bar{\rho}, \rho(\gamma)\}. \tag{5.15}$$

Hence by Lemma 9 there exists $(\bar{z}, \bar{v}) \in T(x^*, u^*)$ with

$$\|x - \bar{z}\|_{1,1} + \|u - \bar{v}\|_1 \leq \gamma\left(\|x - x^*\|_{1,1} + \|u - u^*\|_1\right).$$

Together with (5.15) this implies

$$\|x - \bar{z}\|_{1,1} + \|u - \bar{v}\|_1 \leq \gamma(1 + c_s)\|u - u^*\|_1 \tag{5.16}$$

and therefore

$$\begin{aligned}(1 - \gamma(1 + c_s))\|u - u^*\|_1 \leq \|\bar{v} - u^*\|_1 \leq (1 + \gamma(1 + c_s))\|u - u^*\|_1,\\ \|\bar{z} - x^*\|_{1,1} \leq (1 + \gamma)(1 + c_s)\|u - u^*\|_1.\end{aligned} \tag{5.17}$$

Further using $(x - x^*, u - u^*) = (\bar{z} - x^* + x - \bar{z}, \bar{v} - u^* + u - \bar{v})$ we obtain

$$\begin{aligned}&\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((x - x^*, u - u^*), (x - x^*, u - u^*)\right)\\&\quad = \mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((\bar{z} - x^*, \bar{v} - u^*), (\bar{z} - x^*, \bar{v} - u^*)\right)\\&\qquad + 2\mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((\bar{z} - x^*, \bar{v} - u^*), (x - \bar{z}, u - \bar{v})\right)\\&\qquad + \mathscr{L}''(\bar{x}, \bar{u}, \bar{\lambda})\left((x - \bar{z}, u - \bar{v}), (x - \bar{z}, u - \bar{v})\right)\end{aligned} \tag{5.18}$$

Next we estimate the terms on the ride hand side of (5.18). By Lemma 10 and (5.17) we have

$$\begin{aligned}&\mathscr{L}''(\bar{x},\bar{u},\bar{\lambda})\left((\bar{z}-x^*,\bar{v}-u^*),(\bar{z}-x^*,\bar{v}-u^*)\right)\\&\quad\geq -\left(\beta+\frac{\alpha}{4}\right)\|\bar{v}-u^*\|_1^2 \geq -\left(\beta+\frac{\alpha}{4}\right)(1+\gamma(1+c_s))^2\,\|u-u^*\|_1^2.\end{aligned}\tag{5.19}$$

By (5.6), (5.16) we obtain

$$\begin{aligned}&|\mathscr{L}''(\bar{x},\bar{u},\bar{\lambda})\left((\bar{z}-x^*,\bar{v}-u^*),(x-\bar{z},u-\bar{v})\right)|\\&\quad\leq C_{\mathscr{L}}\left(\|\bar{z}-x^*\|_\infty\|x-\bar{z}\|_\infty+\|\bar{z}-x^*\|_\infty\|u-\bar{v}\|_1+\|x-\bar{z}\|_\infty\|\bar{v}-u^*\|_1\right)\\&\quad\leq C_{\mathscr{L}}\left(\|\bar{z}-x^*\|_{1,1}\|x-\bar{z}\|_{1,1}+\|\bar{z}-x^*\|_{1,1}\|u-\bar{v}\|_1+\|x-\bar{z}\|_{1,1}\|\bar{v}-u^*\|_1\right)\\&\quad\leq C_{\mathscr{L}}\gamma(1+c_s)\left(\|\bar{z}-x^*\|_{1,1}+\|\bar{v}-u^*\|_1\right)\|u-u^*\|_1.\end{aligned}$$

By (5.17) this implies

$$\begin{aligned}&|\mathscr{L}''(\bar{x},\bar{u},\bar{\lambda})\left((\bar{z}-x^*,\bar{v}-u^*),(x-\bar{z},u-\bar{v})\right)|\\&\quad\leq C_{\mathscr{L}}\gamma(1+c_s)\,(1+(1+2\gamma)(1+c_s))\,\|u-u^*\|_1^2.\end{aligned}\tag{5.20}$$

Again by (5.6) and (5.16) we obtain

$$\begin{aligned}&|\mathscr{L}''(\bar{x},\bar{u},\bar{\lambda})\,((x-\bar{z},u-\bar{v}),(x-\bar{z},u-\bar{v}))\\&\quad\leq C_{\mathscr{L}}\left(\|x-\bar{z}\|_\infty^2+2\|x-\bar{z}\|_\infty\|u-\bar{v}\|_1\right)\\&\quad\leq C_{\mathscr{L}}\left(\|x-\bar{z}\|_{1,1}^2+2\|x-\bar{z}\|_{1,1}\|u-\bar{v}\|_1\right)\\&\quad\leq 3C_{\mathscr{L}}\gamma^2(1+c_s)^2\|u-u^*\|_1^2.\end{aligned}\tag{5.21}$$

Now inserting the estmates (5.19), (5.20), (5.21) into (5.18) the assertion follows from (5.14). □

We can now show, that Assumptions (5.7) and (5.8) imply the growth condition (4.8) with $\kappa = 2$.

Theorem 3 *Let Assumptions* (2.1), (5.1), (5.2), (5.7), *and* (5.8) *be satisfied. Then*

$$f(x(t_f)) - f(x^*(t_f)) \geq \frac{3}{4}\alpha\,\|u-u^*\|_1^2$$

for all $(x, u) \in \mathscr{F}$ *with* $\|u-u^*\|_1 < \delta_2$*, where* δ_2 *is defined by Lemma* 11*. Moreover, condition* (4.8) *is satisfied with* $\kappa = 2$.

Proof For arbitrary $(x, u) \in \mathscr{F}$ with $\|u-u^*\|_1 < \delta_2$ we have

$$f\left(x(t_f)\right) - f\left(x^*(t_f)\right) = \mathscr{L}(x, u, \lambda^*) - \mathscr{L}(x^*, u^*, \lambda^*).$$

Using $x(t_0) - x^*(t_0) = 0$ and (5.3) we get by Taylor expansion

$$f\left(x(t_f)\right) - f\left(x^*(t_f)\right) = \mathscr{L}_u(x^*, u^*, \lambda^*)(u - u^*) + \frac{1}{2}\mathscr{L}''(z, v, \lambda^*)\left((x - x^*, u - u^*), (x - x^*, u - u^*)\right),$$

where $(z, v) = (1-\tau)(x^*, u^*) + \tau(x, u)$ with $\tau \in]0, 1[$. By (5.7) and Lemma 11 this implies

$$f\left(x(t_f)\right) - f\left(x^*(t_f)\right) \geq \bar{\alpha}\, \|u - u^*\|_1^2 - \frac{1}{2}\left(\beta + \frac{\alpha}{2}\right) \|u - u^*\|_1^2 = \left(\frac{3}{4}\bar{\alpha} - \frac{\beta}{4}\right) \|u - u^*\|_1^2 \geq \frac{3}{4}\alpha\, \|u - u^*\|_1^2,$$

which proves the first part of the assertion. For arbitrary $(x, u) \in \mathscr{F}_{\delta_2}$ we have $(x, u) \in \mathscr{F}$ and $\|u - u^*\|_1 < \delta_2$ which implies that condition (4.8) is satisfied with $\kappa = 2$, α replaced by $\frac{3}{4}\alpha$, and $\bar{\varepsilon}$ replaced by δ_2. □

Remark 7 Note that for the proof of Theorem 3 we only need the fact, that (x^*, u^*) is feasible and satisfies together with the unique solution λ^* of the adjoint equation the minimum principle and Assumptions (2.3), (2.4), (2.5), (5.1), (5.2), (5.7), and (5.8). Theorem 3 then shows, that (x^*, u^*) is a strict local solution, i.e., Assumptions (5.8), (5.7) can be viewed as sufficient optimality condition. In case of linear-quadratic control problems as considered in Alt et al. [6] the second derivatives of the Hamiltonian and the Lagrange function do not depend on (x, u). This allows to use a more general version of condition (5.7). ♢

For the derivation of error estimates of order 1 for the discrete solutions we proceed similarly to Dontchev and Veliov [14], Haunschmied et al. [21] and use the fact that the discrete solutions can be interpreted as solution of a perturbation of Problem (OC). This approach has also been used in Alt et al. [6] for linear-quadratic control problems. For the more general class of nonlinear control problems considered here we adapt results of Alt [2, Section 3], where Lipschitz continuity of perturbed solutions of nonlinear optimization problems has been studied.

Lemma 12 *Let Assumptions* (2.1), (5.1), (5.2), (5.7)*, and* (5.8) *be satisfied. Further let* $0 < \rho < \bar{\rho}$*, where* $\bar{\rho} > 0$ *is given by Lemma* 5*, and let* (x_h^*, u_h^*) *be a (global) solution of Problem* $(\mathrm{OC})_{N,\rho}$*. Then there is a function* $\zeta_h\colon [t_0, t_f] \to \mathbb{R}^m$ *satisfying*

$$|\zeta_h(t)| \leq K_\lambda(L_g(L_x + 1) + c_A)h_N \tag{5.22}$$

such that

$$\int_{t_0}^{t_f} \left[\lambda_h^*(t)^{\mathsf{T}} g^{(2)}(x_h^*(t), t) + \zeta_h(t)^{\mathsf{T}}\right] \left(u^*(t) - u_h^*(t)\right) \mathrm{d}t \geq 0. \tag{5.23}$$

Proof By the discrete minimum principle (3.2) we have

$$\lambda_h^*(t_{j+1})^\mathsf{T} g^{(2)}(x_h^*(t_j), t_j)(u - u_h^*(t_j)) \geq 0 \quad \forall u \in U, \;\; j \in J_0^{N-1}. \tag{5.24}$$

We define piecewise constant functions $\bar{x}_h \colon [t_0, t_f] \to \mathbb{R}^n$, $B_h \colon [t_0, t_f] \to \mathbb{R}^{n\times m}$, and $\bar{\lambda}_h \colon [t_0, t_f] \to \mathbb{R}^n$ by

$$\bar{x}_h(t) = x_h^*(t_j), \;\; B_h(t) = g^{(2)}\left(x_h^*(t_j), t_j\right), \;\; \bar{\lambda}_h(t) = \lambda_h^*(t_{j+1}),$$

for $t \in [t_j, t_{j+1}[$, $j \in J_0^{N-1}$. Then we can write the discrete switching function σ_h^* defined by (3.3) in the form

$$\sigma_h^*(t) = B_h(t)^\mathsf{T}\bar{\lambda}_h(t) = g^{(2)}(x_h^*(t), t)^\mathsf{T}\lambda_h^*(t) + \zeta_h(t) \;\; \text{for a.a.} \;\; t \in [t_0, t_f],$$

where ζ_h is defined by

$$\zeta_h(t) = B_h(t)^\mathsf{T}\bar{\lambda}_h(t)^\mathsf{T} - g^{(2)}(x_h^*(t), t)^\mathsf{T}\lambda_h^*(t).$$

Further we can write the discrete minimum principle (5.24) in the form

$$\begin{aligned}\sigma_h^*(t)^\mathsf{T}(u - u_h^*(t)) &= \left[g^{(2)}(x_h^*(t), t)^\mathsf{T}\lambda_h^*(t) + \zeta_h(t)\right]^\mathsf{T}(u - u_h^*(t))\\ &\geq 0 \quad \forall u \in U\end{aligned} \tag{5.25}$$

for a.a. $t \in [t_0, t_f]$, From (5.25) we further obtain (5.23). In order to estimate $|\zeta_h(t)|$ we use (2.4). For $t \in [t_j, t_{j+1}[$, $j \in J_0^{N-1}$, we have

$$\begin{aligned}|\zeta_h(t)| &\leq |\bar{\lambda}_h(t)|\,\|B_h(t) - g^{(2)}(x_h^*(t), t)\| + \|g^{(2)}(x_h^*(t), t)\|\,|\bar{\lambda}_h(t) - \lambda_h^*(t)|\\ &\leq K_\lambda L_g\left(|x_h^*(t_j) - x_h^*(t)| + |t_j - t|\right) + c_g|\lambda_h^*(t_{j+1}) - \lambda_h^*(t)|\end{aligned}$$

with some constant c_g independent of $(x_h^*, u_h^*) \in \mathscr{F}_{N,\rho}$. By (2.8) we have

$$|x_h^*(t_j) - x_h^*(t)| \leq L_x h_N,$$

and from the discrete adjoint equation we obtain

$$|\lambda_h^*(t_{j+1}) - \lambda_h^*(t)| \leq h_N\|A_h(t_j)\||\lambda_{h,j+1}^*| \leq c_A K_\lambda h_N,$$

which implies (5.22). □

Theorem 4 *Let Assumptions* (2.1), (5.1), (5.2), (5.7), *and* (5.8) *be satisfied and suppose that u^* has bounded variation. Then for each $0 < \rho \leq \delta_2$ Problem $(\mathrm{OC})_{N,\rho}$ has a (global) solution for sufficiently large N. Further for each such solution (x_h^*, u_h^*) and the associated adjoint function λ_h^* the estimates*

$$\|u_h^* - u^*\|_1 \leq c_u h_N, \;\; \|x_h^* - x^*\|_{1,1} \leq c_x h_N, \;\; \|\lambda_h^* - \lambda^*\|_{1,1} \leq c_\lambda h_N \tag{5.26}$$

hold with constants c_u, c_x, and c_λ independent of N and the solution (x_h^*, u_h^*).

Proof Let $0 < \rho \le \delta_2$ be given. By Theorem 3 condition (4.8) is satisfied. Therefore, by Theorem 1 Problem $(\mathrm{OC})_{N,\rho}$ has a (global) solution for sufficiently large N. Further for each such solution (x_h^*, u_h^*) the estimates

$$\|u_h^* - u^*\|_1 \le c_u h_N^{\frac{1}{2}}, \quad \|x_h^* - x^*\|_{1,1} \le c_x h_N^{\frac{1}{2}}$$

hold with constants c_u, c_x independent of N, and by (4.19) we have

$$\|\lambda_h^* - \lambda^*\|_{1,1} \le c_{\lambda,2} h_N^{\frac{1}{2}}$$

with some constant $c_{\lambda,2}$ independent of N and x_h^*, x^*, u_h^*, and u^*. For suffiently large N we therefore have

$$\begin{aligned} &\|x_h^* - x^*\|_\infty + \|u_h^* - u^*\|_1 + \|\lambda_h^* - \lambda^*\|_{1,1} \\ &\quad \le \|x_h^* - x^*\|_{1,1} + \|u_h^* - u^*\|_1 + \|\lambda_h^* - \lambda^*\|_{1,1} < \delta_2, \end{aligned}$$

and $(x_h^*, u_h^*) \in \mathscr{F}_{N,\bar\rho}$. By Lemma 6 there exists a function $z_h^* \in X_1$, such that $(z_h^*, u_h^*) \in \mathscr{F}$ and

$$\|z_h^* - x_h^*\|_{1,1} \le c_z h_N \tag{5.27}$$

with a constant c_z independent of N, which implies

$$\|z_h^* - x^*\|_\infty \le \delta_2 \tag{5.28}$$

for sufficiently large N. As in the proof of Theorem 3, using $z_h^*(t_0) - x^*(t_0) = 0$ and (5.3), we get by Taylor expansion around (x^*, u^*)

$$\begin{aligned} f\left(z_h^*(t_f)\right) - f\left(x^*(t_f)\right) &= \mathscr{L}(z_h^*, u_h^*, \lambda^*) - \mathscr{L}(x^*, u^*, \lambda^*) \\ &= \mathscr{L}_u(x^*, u^*, \lambda^*)(u_h^* - u^*) \\ &\quad + \frac{1}{2}\mathscr{L}''(z, v, \lambda^*)\left((z_h^* - x^*, u_h^* - u^*), \left(z_h^* - x^*, u_h^* - u^*\right)\right), \end{aligned}$$

where $(z, v) = (1-\tau)(x^*, u^*) + \tau(z_h^*, u_h^*)$ with $\tau \in]0, 1[$. By (5.7), (5.8), and Lemma 11 this implies

$$\begin{aligned} f\left(z_h^*(t_f)\right) - f\left(x^*(t_f)\right) &\ge \bar\alpha \, \|u_h^* - u^*\|_1^2 - \frac{1}{2}\left(\beta + \frac{\alpha}{2}\right) \|u_h^* - u^*\|_1^2 \\ &= \left(\frac{3}{4}\alpha + \frac{\beta}{2}\right) \|u_h^* - u^*\|_1^2. \end{aligned} \tag{5.29}$$

Similarly we obtain by Taylor expansion around (z_h^*, u_h^*)

$$\begin{aligned} f\left(x^*(t_f)\right) - f\left(z_h^*(t_f)\right) &= \mathscr{L}(x^*, u^*, \lambda_h^*) - \mathscr{L}(z_h^*, u_h^*, \lambda_h^*) \\ &= \mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(x^* - z_h^*) + \mathscr{L}_u(z_h^*, u_h^*, \lambda_h^*)(u^* - u_h^*) \\ &\quad + \frac{1}{2}\mathscr{L}''(z, v, \lambda_h^*)\left((x^* - z_h^*, u^* - u_h^*), (x^* - z_h^*, u^* - u_h^*)\right), \end{aligned}$$

where $(z, v) = (1-\tau)(z_h^*, u_h^*) + \tau(x^*, u^*)$ with $\tau \in]0, 1[$. By Lemma 11 this implies

$$\begin{aligned} f\left(x^*(t_f)\right) - f\left(z_h^*(t_f)\right) \geq \mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(x^* - z_h^*) + \mathscr{L}_u(z_h^*, u_h^*, \lambda_h^*)(u^* - u_h^*) \\ - \frac{1}{2}\left(\beta + \frac{\alpha}{2}\right) \|u_h^* - u^*\|_1^2. \end{aligned}$$

Combining this estimate with (5.29) we obtain

$$\frac{\alpha}{2} \|u_h^* - u^*\|_1^2 \leq -\mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(x^* - z_h^*) - \mathscr{L}_u(z_h^*, u_h^*, \lambda_h^*)(u^* - u_h^*). \tag{5.30}$$

We define $z = x^* - z_h^*$. Using integration by parts we obtain for the first term on the right hand side of (5.30)

$$\begin{aligned} -\mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(z) &= - f_x\left(z_h^*(t_f)\right) z(t_f) - \int_{t_0}^{t_f} \lambda_h^*(t)^{\mathsf{T}} g_x(z_h^*(t), u_h^*(t), t) z(t)\,\mathrm{d}t \\ &\quad + \int_{t_0}^{t_f} \lambda_h^*(t)^{\mathsf{T}} \dot{z}(t)\,\mathrm{d}t \\ &= - f_x\left(z_h^*(t_f)\right) z(t_f) - \int_{t_0}^{t_f} \lambda_h^*(t)^{\mathsf{T}} g_x(z_h^*(t), u_h^*(t), t) z(t)\,\mathrm{d}t \\ &\quad + \lambda_h^*(t_f) z(t_f) - \int_{t_0}^{t_f} \dot{\lambda}_h^*(t)^{\mathsf{T}} z(t)\,\mathrm{d}t \end{aligned}$$

By Lemma 8 we can write the discrete adjoint Eq. (3.1) in the form (4.17). Using this and the terminal condition $\lambda_h^*(t_f) = f_x(x_h^*(t_f))^{\mathsf{T}}$ we further obtain

$$\begin{aligned} -\mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(z) &= -f_x\left(z_h^*(t_f)\right) z(t_f) - \int_{t_0}^{t_f} \lambda_h^*(t)^{\mathsf{T}} g_x(z_h^*(t), u_h^*(t), t) z(t)\,\mathrm{d}t \\ &\quad + f_x(x_h^*(t_f)) z(t_f) \\ &\quad + \int_{t_0}^{t_f} \left[\lambda_h^*(t)^{\mathsf{T}} g_x(x_h^*(t), u_h^*(t), t) + \xi_h(t)^{\mathsf{T}}\right] z(t)\,\mathrm{d}t \end{aligned}$$

By (2.4), (4.16), (4.18), and (5.27) this implies

$$\begin{aligned} -\mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(x^* - z_h^*) &\leq L_f |z_h^*(t_f) - x_h^*(t_f)|\, |x^*(t_f) - z_h^*(t_f)| \\ &\quad + (t_f - t_0)\left(K_\lambda L_g \|z_h^* - x_h^*\|_\infty + \|\xi_h\|_\infty\right) \|x^* - z_h^*\|_\infty \\ &\leq \left[L_f c_z + (t_f - t_0)\left(K_\lambda L_g c_z + c_\xi\right)\right] h_N \|x^* - z_h^*\|_\infty. \end{aligned}$$

It follows from Lemma 5 that

$$\|z_h^* - x^*\|_\infty \le \|z_h^* - x^*\|_{1,1} \le c_s \|u_h^* - u^*\|_1, \tag{5.31}$$

and

$$\begin{aligned} &- \mathscr{L}_x(z_h^*, u_h^*, \lambda_h^*)(x^* - z_h^*) \\ &\quad \le \left[L_f c_z + (t_f - t_0)\left(K_\lambda L_g c_z + c_\xi\right)\right] c_s h_N \|u_h^* - u^*\|_1. \end{aligned} \tag{5.32}$$

By Lemma 12 there is a function $\zeta_h \colon [t_0, t_f] \to \mathbb{R}^m$ satisfying (5.22) such that the discrete minimum principle can be written in the form (5.23). Therefore, defining

$$\bar{\zeta}_h(t) = \lambda_h^*(t)^\mathsf{T}\left[g^{(2)}(x_h^*(t), t) - g^{(2)}(z_h^*(t), t)\right] + \zeta_h(t)$$

we obtain for the second term on the right hand side of (5.30)

$$\begin{aligned} -\mathscr{L}_u(z_h^*, u_h^*, \lambda_h^*)(u^* - u_h^*) &= -\int_{t_0}^{t_f} \lambda_h^*(t)^\mathsf{T} g^{(2)}(z_h^*(t), t)\left(u^*(t) - u_h^*(t)\right)\, \mathrm{d}t \\ &= -\int_{t_0}^{t_f} \left[\lambda_h^*(t)^\mathsf{T} g^{(2)}(x_h^*(t), t) + \zeta_h(t)^\mathsf{T}\right]\left(u^*(t) - u_h^*(t)\right)\, \mathrm{d}t \\ &\quad + \int_{t_0}^{t_f} \bar{\zeta}_h(t)^\mathsf{T}\left(u^*(t) - u_h^*(t)\right)\, \mathrm{d}t \\ &\le \int_{t_0}^{t_f} \bar{\zeta}_h(t)^\mathsf{T}\left(u^*(t) - u_h^*(t)\right)\, \mathrm{d}t. \end{aligned}$$

Further by (2.4), (4.16), (5.22), and (5.27) we have

$$|\bar{\zeta}_h(t)| \le K_\lambda L_g |x_h^*(t) - z_h^*(t)| + |\zeta_h(t)| \le K_\lambda \left(L_g(c_z + L_x + 1) + c_A\right) h_N,$$

which implies

$$\begin{aligned} &-\mathscr{L}_u(z_h^*, u_h^*, \lambda_h^*)(u^* - u_h^*) \\ &\quad \le (t_f - t_0) K_\lambda \left(L_g(c_z + L_x + 1) + c_A\right) h_N \|u_h^* - u^*\|_1. \end{aligned} \tag{5.33}$$

Finally (5.30) and the estimates (5.32) and (5.33) show that with some constant $\tilde{c}_u$ independent of N and the solution (x_h^*, u_h^*),

$$\frac{\alpha}{2}\|u_h^* - u^*\|_1^2 \le \tilde{c}_u h_N \|u_h^* - u^*\|_1,$$

which immediately implies the estimate for $\|u_h^* - u^*\|_1$ in (5.26). The estimate for $\|x_h^* - x^*\|_{1,1}$ then follows from (5.27) and (5.31), and the estimate for $\|\lambda_h^* - \lambda^*\|_{1,1}$ follows from (4.20). □

Remark 8 (compare Remark 3) Theorem 4 also assumes that (x_h^*, u_h^*) is a global solution of Problem $(\mathrm{OC})_{N,\rho}$ and we have not shown uniqueness of the discrete solutions. The reason is that for the class of control problems considered here condition (5.7) does in general not hold for the discrete control problems. The linear-quadratic control problems considered in Alt et al. [6] are convex optimization problems. Therefore, the solutions of the discrete problems are global solutions. In this case Theorem 4 implies Alt et al. [6, Theorem 14] for $\kappa = 1$. ♢

6 Numerical results

Example 2 We consider the following modification of the rocket car problem discussed in Alt et al. [4, Example 6.1] with a nonlinear and non convex cost functional and a nonlinear state equation:

$$
\begin{aligned}
(\mathrm{OCB})\quad & \min \frac{1}{2}(x_1(5)^3 + x_2(5)^2) \\
& \text{s.t.} \\
& \dot{x}_1(t) = x_2(t), \quad \dot{x}_2(t) = (1 + \varepsilon x_2(t))\, u(t) \qquad \text{a.e. on } [0,5], \\
& x_1(0) = 6, \; x_2(0) = 1, \\
& -1 \le u(t) \le 1 \qquad \text{a.e. on } [0,5].
\end{aligned}
$$

Here the function g is defined by

$$g(x_1, x_2, u) = \begin{pmatrix} x_2 \\ (1+\varepsilon x_2)u \end{pmatrix},$$

and

$$g_x(x_1, x_2, u) = \begin{pmatrix} 0 & 1 \\ 0 & \varepsilon u \end{pmatrix}, \quad g_u(x_1, x_2, u) = \begin{pmatrix} 0 \\ 1 + \varepsilon x_2 \end{pmatrix}.$$

The Hamiltonian is defined by

$$H(x_1, x_2, u, \lambda_1, \lambda_2) = \lambda_1 x_2 + \lambda_2 (1 + \varepsilon x_2) u,$$

and we have

$$H_x(x_1, x_2, u, \lambda_1, \lambda_2) = (0,\; \lambda_1 + \varepsilon \lambda_2 u)\,, \quad H_u(x_1, x_2, u, \lambda_1, \lambda_2) = \lambda_2(1 + \varepsilon x_2),$$

and

$$H_{ux}(x_1, x_2, u, \lambda_1, \lambda_2) = (0, \varepsilon\lambda_2)\,, \quad H_{xx}(x_1, x_2, u, \lambda_1, \lambda_2) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Table 1 Discretization error

N	100	200	500	1000	5000
$\|u_h^* - u^*\|_1$	3.795×10^{-2}	2.567×10^{-2}	1.603×10^{-2}	6.900×10^{-3}	9.002×10^{-4}
$\frac{\|u_h^* - u^*\|_1}{h}$	0.759	1.027	1.603	1.380	0.9002

Therefore, the condition for the quadratic form in Assumption (5.8) is here equivalent to

$$3x_1^*(5)z_1(5)^2 + z_2(5)^2 + 2\varepsilon \int_0^5 z_2(t)\lambda_2^*(t)v(t)\,\mathrm{d}t \geq -\beta\,\|v\|_1^2$$

for all $(z, v) = (x, u) - (x^*, u^*)$ with $u \in \mathscr{U}$, $z(t_0) = 0$ and

$$\begin{aligned} \dot{z}_1(t) &= z_2(t), \\ \dot{z}_2(t) &= \varepsilon u^*(t)z_2(t) + \left(1 + \varepsilon x_2^*(t)\right) v(t) \end{aligned}$$

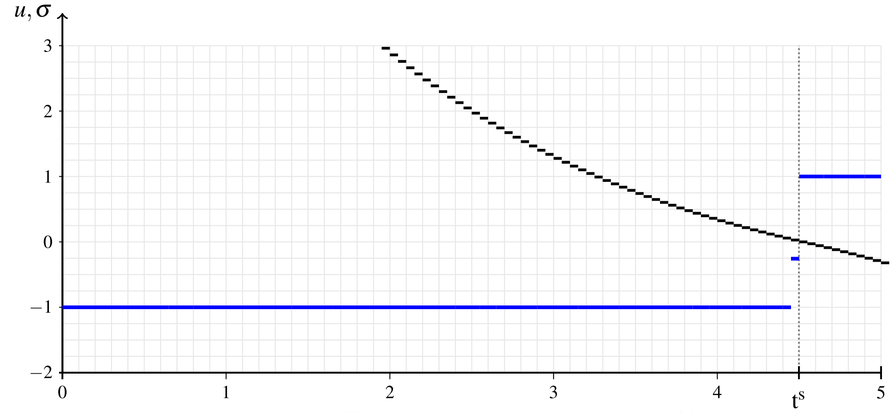
for a.a. $t \in [0, 5]$. From the numerical results we have $x_1^*(5) \approx 1.5475 > 0$. Therefore, this condition is satisfied for arbitrarily small β if ε is sufficiently small.

For $\varepsilon = \frac{1}{2}$ the optimal control is of bang–bang type with one switching point $s_1 \approx 4.49848$. The discretization errors depicted in Table 1 indicate convergence of order 1 w.r.t. the mesh size h as expected by Theorem 4. Figure 1 shows the computed optimal control and states, and the switching function for $N = 100$. ♢

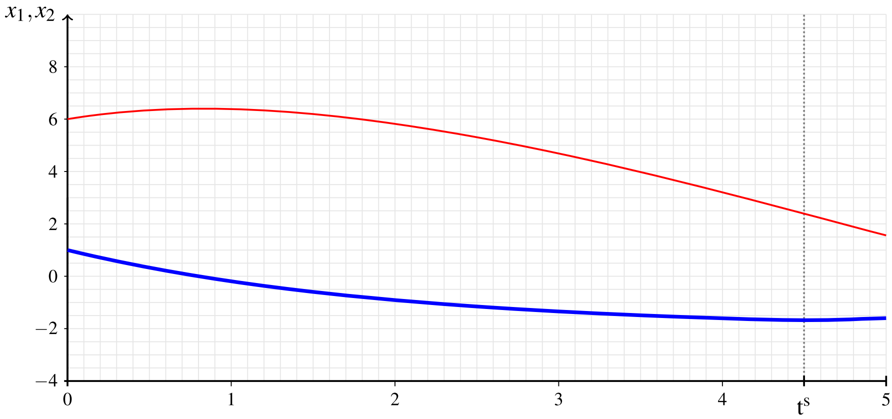
7 Conclusions

In this paper we derived error estimates for Euler approximation of a class of nonlinear optimal control problems of Mayer type with control appearing linearly. Such estimates were previously known only in case of continuous controls, for linear-quadratic problems affine w.r.t. the control, and for some special classes of control problems with a nonlinear cost functional but a linear or semilinear state equation. The results were obtained under the growth condition (4.8) for the cost functional or under the stronger second-order sufficient optimality condition (5.7) excluding singular arcs of the optimal control (see Remark 5). Felgenhauer [18] shows for scalar bang–bang controls that a second-order optimality condition for the so-called "induced finite-dimensional problem" of optimizing the switching times implies (4.8) with $\kappa = 2$. It is an open question whether (4.8) can be satisfied if the optimal control has singular arcs. But it should be noted that in this case other second-order conditions may be useful (see e.g. Felgenhauer [19], where a second-order condition in connection with the Goh transformation has been used).

A nonlinear control problem may have many local solutions and Example 2 shows that the analytical verification of conditions (4.8) and (5.7) may be difficult. Therefore, another important topic, not treated in this paper, is the numerical verification of such



Optimal control $u(t)$ (blue) and switching function $\sigma(t)$ (black).



Optimal states $x_1(t)$ (thin red line) and $x_2(t)$ (thick blue line).

Fig. 1 Optimal solution for $N = 100$ (Color figure online)

conditions. For the numerical verification of (4.8) in case of bang–bang controls one can use known results. First the control problem is solved by Euler discretization in order to obtain a good approximation for the switching times. If this works, in a second step the induced finite-dimensional problem mentioned above can be solved in order to compute the switching times more accurately. Then the test for the numerical verification of a second-order sufficient optimality condition for the induced finite-dimensional problem discussed in Maurer et al. [29] can be applied. If the test is successful, then the results of Felgenhauer [18] show that (4.8) holds with $\kappa = 2$. While the results of Maurer et al. [29] are stated for general nonlinear control problems, the results of Felgenhauer [18] so far are restricted to scalar control problems. As an alternative, tests based on Riccati differential equations can be used (see Felgenhauer [18, Section 4], Osmolovskii and Maurer [34] and the papers cited therein).

Acknowledgements The authors would like to thank the anonymous referees for their careful reading of the manuscript and their constructive and valuable suggestions and comments.

References

1. Alt, W.: On the approximation of infinite optimization problems with an application to optimal control problems. Appl. Math. Optim. **12**, 15–27 (1984)
2. Alt, W.: Local stability of solutions to differentiable optimization problems in Banach spaces. J. Optim. Theory Appl. **70**, 443–466 (1991)
3. 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, New York (1997)
4. Alt, W., Baier, R., Gerdts, M., Lempio, F.: Error bounds for Euler approximation of linear-quadratic control problems with bang–bang solutions. Numer. Algebra Control Optim. **2**(3), 547–570 (2012)
5. Alt, W., Mackenroth, U.: Convergence of finite element approximations to state constrained convex parabolic boundary control problems. SIAM J. Control Optim. **27**, 718–736 (1989)
6. Alt, W., Schneider, C., Seydenschwanz, M.: Regularization and implicit Euler discretization of linear-quadratic optimal control problems with bang-bang solutions. Appl. Math. Comput. (2016). https://doi.org/10.1016/j.amc.2016.04.028
7. 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). https://doi.org/10.1080/10556788.2013.821612
8. Deckelnick, K., Hinze, M.: A note on the approximation of elliptic control problems with bang–bang controls. Comput. Optim. Appl. **51**(2), 931–939 (2012)
9. Dhamo, V., Tröltzsch, F.: Some aspects of reachability for parabolic boundary control problems with control constraints. Comput. Optim. Appl. **50**(1), 75–110 (2011)
10. Dontchev, A.L., Hager, W.W.: Lipschitzian stability in nonlinear control and optimization. SIAM J. Control Optim. **31**(3), 569–603 (1993)
11. Dontchev, A.L., Hager, W.W.: The Euler approximation in state constrained optimal control. Math. Comput. **70**, 173–203 (2000)
12. 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)
13. Dontchev, A.L., Hager, W.W., Veliov, V.M.: Second-order Runge–Kutta approximations in control constrained optimal control. SIAM J. Numer. Anal. **38**, 202–226 (2000)
14. Dontchev, A.L., Veliov, V.M.: Metric regularity under approximations. Control Cybern. **38**(4B), 1283–1303 (2009)
15. Felgenhauer, U.: On stability of bang–bang type controls. SIAM J. Control Optim. **41**(6), 1843–1867 (2003)
16. Felgenhauer, U.: Optimality properties of controls with bang–bang components in problems with semilinear state equation. Control Cybern. **34**(3), 764–785 (2005)
17. Felgenhauer, U.: The shooting approach in analyzing bang–bang extremals with simultaneous control switches. Control Cybern. **37**(2), 307–327 (2008)
18. Felgenhauer, U.: Note on local quadratic growth estimates in bang–bang optimal control problems. Optimization **64**(3), 521–537 (2015). https://doi.org/10.1080/02331934.2013.773000
19. Felgenhauer, U.: Discretization of semilinear bang-singular-bang control problems. Comput. Optim. Appl. **64**, 295–326 (2016). https://doi.org/10.1007/s10589-015-9800-2
20. Felgenhauer, U., Poggiolini, L., Stefani, G.: Optimality and stability result for bang–bang optimal controls with simple and double switch behaviour. Control Cybern. **38**, 1305–1325 (2009)
21. Haunschmied, J.L., Pietrus, A., Veliov, V.M.: The Euler method for linear control systems revisited. In: Large-Scale Scientific Computing—9th International Conference, LSSC 2013, Sozopol, Bulgaria, pp. 90–97, 3–7 June 2013. Revised Selected Papers (2013)
22. Ioffe, A., Tihomirov, V.M.: Theorie der Extremalwertaufgaben. VEB Deutscher Verlag der Wissenschaften, Berlin (1979)
23. Kaya, C.Y., Lucas, S.K., Simakov, S.T.: Computations for bang–bang constrained optimal control using a mathematical programming formulation. Optim. Control Appl. Methods **25**, 295–308 (2004). https://doi.org/10.1002/oca.749
24. Malanowski, K.: Stability and sensitivity of solutions to optimal control problems for systems with control appearing linearly. Appl. Math. Optim. **16**, 73–91 (1987)
25. Malanowski, K.: Stability and sensitivity of solutions to nonlinear optimal control problems. Appl. Math. Optim. **32**, 111–141 (1995)

26. Malanowski, K.: Stability and sensitivity analysis for optimal control problems. A survey. Trudy Inst. Mat. i Mekh. UrO RAN **16**(5), 278–288 (2010)
27. 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, New York (1997)
28. Maurer, H.: First- and second order sufficient optimality conditions in mathematimacal programming and optimal control. Math. Program. Study **14**, 163–177 (1981)
29. 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. Optim. Control Appl. Methods **26**(3), 129–156 (2005)
30. Maurer, H., Osmolovskii, N.P.: Second order sufficient conditions for time optimal bang–bang controls. SIAM J. Control Optim. **42**(6), 2239–2263 (2004)
31. Maurer, H., Zowe, J.: First- and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. Math. Program. **16**, 98–110 (1979)
32. Osmolovskii, N.P., Maurer, H.: Equivalence of second order optimality conditions for bang–bang control problems. Part 1: main results. Control Cybern. **34**(3), 927–950 (2005)
33. Osmolovskii, N.P., Maurer, H.: Equivalence of second order optimality conditions for bang–bang control problems. Part 2: proofs, variational derivatives and representations. Control Cybern. **36**(1), 5–45 (2007)
34. Osmolovskii, N.P., Maurer, H.: Applications to Regular and Bang–Bang Control: Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control. SIAM, Philadelphia (2012)
35. Pietrus, A., Scarinci, T., Veliov, V.M.: High Order Discrete Approximations to Mayer's Problems for Linear Systems. Technical Report, TU Wien, ORCOS (2016)
36. 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)
37. Robinson, S.M.: Regularity and stability for convex multivalued functions. Math. Oper. Res. **1**, 130–143 (1976)
38. Robinson, S.M.: Stability theory for systems of inequalities, Part II: differentiable nonlinear systems. SIAM J. Numer. Anal. **13**(4), 497–513 (1976)
39. Sendov, B., Popov, V.A.: The averaged Moduli of Smoothness. Wiley, Chichester (1988)
40. Seydenschwanz, M.: Convergence results for the discrete regularization of linear-quadratic control problems with bang–bang solutions. Comput. Optim. Appl. **61**, 731–760 (2015). https://doi.org/10.1007/s10589-015-9730-z
41. Veliov, V.M.: Error analysis of discrete approximations to bang–bang optimal control problems: the linear case. Control Cybern. **34**(3), 967–982 (2005)