HAL
archives-ouvertes.fr

Error estimates for the Euler discretization of an optimal control problem with first-order state constraints

Joseph Frederic Bonnans, Adriano Festa

▶ To cite this version:

Joseph Frederic Bonnans, Adriano Festa. Error estimates for the Euler discretization of an optimal control problem with first-order state constraints. SIAM Journal on Numerical Analysis, Society for Industrial and Applied Mathematics, 2017, 55 (2), pp.445–471. <hal-01093229v2>

HAL Id: hal-01093229
https://hal.inria.fr/hal-01093229v2

Submitted on 27 Aug 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ERROR ESTIMATES FOR THE EULER DISCRETIZATION OF AN OPTIMAL CONTROL PROBLEM WITH FIRST-ORDER STATE CONSTRAINTS *

J. FRÉDÉRIC BONNANS† AND ADRIANO FESTA‡

Abstract. We study the error introduced in the solution of an optimal control problem with first order state constraints, for which the trajectories are approximated with a classical Euler scheme.

We obtain order one approximation results in the L^∞ norm (as opposed to the order 2/3 obtained in the literature). We assume either a strong second order optimality condition, or a weaker one in the case where the state constraint is scalar, satisfies some hypotheses for junction points, and the time step is constant.

Our technique is based on some homotopy path of discrete optimal control problems that we study using perturbation analysis of nonlinear programming problems.

Key words. Optimal control, nonlinear systems, state constraints, Euler discretization, rate of convergence.

AMS subject classifications. 49M25, 65L10, 65L70, 65K10.

1. Introduction, discussion of literature.

Numerical methods for the resolution of an optimal control problem are based on a finite dimensional approximation, generally obtained through a discretization of the trajectory and a piecewise constant or polynomial control. Obtaining error estimates for such approximations is obviously an important issue.

The first works related appeared in the 1970s; they dealt with convergence of a discrete optimal control solution (see e.g. [9], [10], and [23]). Other results of convergence, provided with modern variational techniques, are also [26]; a survey of the results in this area is [11].

In this paper we will focus on the case of *pure* state constraints, a case which presents some special difficulties. In particular it is known that when the constraint qualification (see [14]) holds and the Lagrangian verifies a local condition of coercivity, the discrete problem obtained with an Euler scheme has a solution, for a sufficiently fine mesh, and the corresponding Lagrange multipliers are at distance $O(\bar{h})$, in the L^2 norm, where $\bar{h}$ is the maximal discretization step, from the continuous solution/multiplier. This important result is due to [13].

The choice of the norm is a delicate point: through the Legendre-Clebsch condition we can get typically an estimation for our variables in a L^2 norm which settles badly with the pure state constrained problem. Such problem naturally requires estimations in the L^∞ norm. This is the so called "*two-norm discrepancy*" [21].

Another sensible matter is that the cost function does not necessarily have derivatives in L^2. This suggests to work in a non linear space of Lipschitz continuous functions with bounded Lipschitz constants. In this setting the L^2 convergence implies L^∞ convergence. This is the way proposed in [13] to obtain a convergence result in the L^∞ norm. This reference obtains an error bound, of order $O(\bar{h}^{2/3})$.

We assume either (i) a strong second order optimality condition, similar to the one in [13], (but we allow a variable time step, whereas the time step was constant in

*This work was partially supported by the European Union under the 7th Framework Programme FP7-PEOPLE-2010-ITN SADCO, Sensitivity Analysis for Deterministic Controller Design.

†Inria and CMAP, Ecole Polytechnique, 91128 Palaiseau, France (frederic.bonnans@inria.fr).

‡RICAM, Austrian Academy of Science, 4040 Linz, Austria (afesta@oeaw.ac.at).

that reference), or (ii) a weaker second order optimality condition, in the case when the state constraint is scalar, structural hypotheses on arcs and junction points, and the time step is constant (the precise statements of these hypotheses are in section 2.5).

In this second case our hypotheses can be motivated in the following way. They allow to obtain the stability of the extremals (of the continuous problem) under a small perturbation, see [3]. We obtain a similar result for the discretized problem. By contrast, for a vector state constraint we are not aware of such stability results, even in the continuous case. This suggests that it might be not easy to obtain the stability of the extremals after discretization without a strong second order optimality condition. This is an interesting open question that we leave for future work, as well as the case of higher order state constraints.

1.1. Structure of the paper. In Section 2 we introduce the problem and the assumptions adopted in the paper, and we state our main result (Theorem 2.6), i.e., a $O(\bar{h})$ error estimate for the control, state, costate and multiplier. A key role in our construction is played by an homotopy path introduced in Section 3. The path links the continuous problem to the discrete one, involving the control, state, costate and multipliers, and creating a class of auxiliary problems. Through the study of the regularity of each auxiliary problem (Section 4) and checking that, under appropriate hypotheses, the application obtained (homotopy path) has bounded directional derivatives (in a sense clarified in the devoted section 5), we can establish the announced convergence estimates for the discrete problem. More precisely, due to some coercivity properties of the Hessian of the Lagrangian, we first obtain a bound in the L^2 norm from which respective estimates in the L^∞ norm easily follow. In this analysis it is used the fact that the state constraint is of first order. Section 6 is dedicated to a simple numerical test. The numerical results are in accordance with our theoretical result and they confirms the tightness of the estimate. An appendix is devoted to the analysis of hypothesis (**A5**).

1.2. Notations. By $\mathbb{R}^n$ we denote the n dimensional Euclidean space. Its dual (whose elements are row vectors) is denoted by $\mathbb{R}^{n*}$. By ∇, ∇_u, etc. we denote the gradient or partial gradient w.r.t. u, who are column vectors, by contrast to the derivatives denoted by e.g. $Dg(x)$ or $g'(x)$ depending on the context, which are identified to row vectors if g is scalar valued. The Lagrange multipliers, including costate variables, are considered as dual elements and are represented by row vectors.

By $C([0,T]$ we denote the space of real continuous functions over $[0,T]$, endowed with the supremum norm. It is known that its topological dual can be identified with the space $\mathcal{M}[0,T]$ of regular, finite Borel measures over $[0,T]$. Let $BV([0,T])$ denote the space of bounded variation functions over $[0,T]$, and let $BV_T([0,T])$ be the subspace of such functions with value 0 at time T. Any continuous linear form on $C([0,T]$ is of the form $f \mapsto \int_0^T f(t)\mathrm{d}\mu(t)$, with $\mu \in BV_T([0,T])$.

2. The continuous problem and its discretization. We consider the following pure state constrained optimal control problem

$$(\mathcal{P}) \left\{ \begin{array}{ll} \text{Minimize } \phi(y(T)); & \text{subject to} \\ \left.\begin{array}{ll} \dot{y}(t) = f(u(t), y(t)), & \text{for a.a. } t \in [0,T]; \\ y(0) = y_0; \end{array}\right\} & \\ g_i(y(t)) \leq 0, & t \in [0,T], \quad i = 1, ..., r, \end{array} \right. \tag{2.1}$$

where the initial condition $y_0 \in \mathbb{R}^n$, the control $u(t)$ and the state $y(t)$ belong to the spaces $\mathcal{U} := L^\infty(0,T;\mathbb{R}^m)$ and $\mathcal{Y} := W^{1,\infty}(0,T;\mathbb{R}^n)$, resp., and g_i is the i-th component of the vector g. Moreover we assume:

(A0) The mappings $\phi : \mathbb{R}^n \to \mathbb{R}$, $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, and $g : \mathbb{R}^n \to \mathbb{R}^r$ are of class C^2 with locally Lipschitz second order continuous derivatives.
In addition, the initial condition $y_0 \in \mathbb{R}^n$ satisfies $g_i(y_0) < 0$, $i = 1, ..., r$.

A *trajectory* of $(\mathcal{P})$ is an element (u, y) of $\mathcal{U} \times \mathcal{Y}$ solution of the state equation (2.1). We say that $(\tilde{u}, \tilde{y})$ is a *local solution* of $(\mathcal{P})$, if it minimizes $\phi(\cdot)$ over the set of feasible trajectories (u, y) satisfying $\|u - \tilde{u}\|_\infty \leq \delta$ for some $\delta > 0$. We assume that

(A1) The *nominal* trajectory $(\bar{u}, \bar{y})$ is a local solution of $(\mathcal{P})$ in $\mathcal{U} \times \mathcal{Y}$, and $\bar{u}$ is a continuous function of time.

The *first order time derivative* of the state constraint is the function

$$g^{(1)} : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^r,\ (u, y) \to g'(y)f(u, y). \tag{2.2}$$

Note that $g^{(1)}(\bar{u}, \bar{y})$ is the time derivative of $g(\bar{y})$ along a trajectory.

Denote the set of *active constraints* at time t by

$$\mathcal{A}(t) := \{i = 1, ..., r \,|\, g_i(\bar{y}(t)) = 0\}.$$

We say that the trajectory $(\bar{u}, \bar{y})$ has *regular first order state constraints* if the following holds:

(A2) There exists $\alpha_g > 0$ such that, for all $t \in [0, T]$ and $\lambda \in \mathbb{R}^{r*}$ verifying $\lambda_i = 0$ if $i \notin \mathcal{A}(t)$, the following holds:

$$|\lambda| \leq \alpha_g \left| \sum_{i \in \mathcal{A}(t)} \lambda_i \nabla_u g_i^{(1)}(\bar{u}(t), \bar{y}(t)) \right|.$$

The *Hamiltonian* function $H : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n*} \to \mathbb{R}$ is defined by:

$$H[p](u, y) := pf(u, y).$$

With this classical notation we view the Hamiltonian as a function of (u, y), parameterized by p, so that e.g. $DH[p](u, y)$ denote the derivative of the Hamiltonian w.r.t. (u, y).

For $i = 1$ to r, we define the *contact set* for the i-th constraint by

$$I_i := \{t \in [0, T];\ g_i(\bar{y}(t)) = 0\}.$$

We say also that the i-constraint is *active* at time t, if $t \in I_i$; otherwise the constraint will be inactive. A maximal open interval (a, b) of I_i (resp. of $[0, T] \setminus I_i$) is called *boundary arc* (resp. *interior arc*). The left and right endpoints of a boundary arc are called *entry* and *exit* points, respectively. We call *junction points* the union of entry and exit points.

2.1. Optimality conditions. We next introduce Pontryagin extremal in a qualified form, which is convenient in view of hypothesis **(A2)**.

DEFINITION 2.1. *A trajectory $(\bar{u},\bar{y})$ is a regular* Pontryagin extremal *of $(\mathcal{P})$, if there exist $\bar{\eta}\in BV_T([0,T],\mathbb{R}^r)$ and $p\in BV([0,T],\mathbb{R}^{n*})$, such that:*

$$\dot{\bar{y}}(t)=f(\bar{u}(t),\bar{y}(t))\ \text{a.e. on } [0,T],\qquad \bar{y}(0)=y_0, \tag{2.3}$$

$$-d\bar{p}(t)=\bar{p}(t)f_y(\bar{u}(t),\bar{y}(t))dt+\sum_{i=1}^{r}g_i'(\bar{y}(t))d\bar{\eta}_i(t),\quad \bar{p}(T)=\phi'(\bar{y}(T)), \tag{2.4}$$

$$\bar{u}(t)\in \operatorname*{argmin}_{\tilde{u}\in\mathbb{R}^m}\{\bar{p}(t)f(\tilde{u},\bar{y}(t))\}\ \text{a.e. on } [0,T], \tag{2.5}$$

$$0\geq g_i(\bar{y}(t)),\quad d\bar{\eta}_i\geq 0,\quad \int_0^T g_i(\bar{y}(t))d\bar{\eta}_i(t)=0,\quad i=1,...,r. \tag{2.6}$$

We call $\bar{\eta}$ a Pontryagin multiplier, and p a costate. Observe that condition (2.5) is equivalent to the *Hamiltonian inequality*

$$H[\bar{p}(t)](\bar{u}(t),\bar{y}(t))\leq H[\bar{p}(t)](u,\bar{y}(t)),\quad \text{for all } u\in\mathbb{R}^m,\ \text{a.e. on } [0,T]. \tag{2.7}$$

A trajectory (u,y) is a *stationary point* of $(\mathcal{P})$, if there exist $\bar{\eta}\in BV_T([0,T],\mathbb{R}^r)$ and $\bar{p}\in BV([0,T];\mathbb{R}^{n*})$ such that (2.3), (2.4), (2.6) hold, as well as

$$0=H_u[\bar{p}(t)](\bar{u}(t),\bar{y}(t))=\bar{p}(t)f_u(\bar{u}(t),\bar{y}(t))\ \text{for a.a. } t\in[0,T].$$

Then we call $\bar{\eta}$ a Lagrange multiplier. Obviously, a regular Pontryagin extremal is also a stationary point, and the converse holds if the Hamiltonian H is a convex function of the control for a.a. time.

The *linearized state equation* at $(\bar{u},\bar{y})$ is, for $v\in L^2(0,T)^m$:

$$\dot{z}(t)=f'(\bar{u}(t),\bar{y}(t))(v(t),z(t));\quad z(0)=0, \tag{2.8}$$

and we denote its solution by $z[v]$.

THEOREM 2.2. *Any qualified solution of $(\mathcal{P})$ is a regular Pontryagin extremal.*

Proof. It is known that a solution of the problem satisfies Pontryagin's principle in unqualified form, see e.g. [27]. On the other hand, by **(A2)**, $(\bar{u},\bar{y})$ satisfies the following constraints qualification [24] (cf. also [6]): there exists $\varepsilon_Q>0$ and $v^\sharp\in L^\infty$ such that $z^\sharp=z[v^\sharp]$ satisfies

$$g_i(\bar{y}(t))+g_i'[t]z^\sharp(t)<-\varepsilon_Q<0,\quad t\in[0,T],\ i=1,...,r. \tag{2.9}$$

But while the above qualification condition only guarantees the fact that the set of Lagrange multipliers is nonempty and bounded, **(A2)** implies the uniqueness of the Lagrange multiplier, see [5]. □

2.2. A key result. The next assumption is quite common in these problems and it plays a crucial role in the analysis. We assume that problem $(\mathcal{P})$ has a local solution $(\bar{u},\bar{y})$, with associated multipliers $\bar{p}$ and $\bar{\eta}$ satisfying the following condition

(A3) (*Strengthened Legendre-Clebsch condition*) There exists $\alpha>0$ such that

$$H_{uu}[\bar{p}(t)](\bar{u}(t),\bar{y}(t))\geq\alpha,\ \text{for a.a. } t\in[0,T]. \tag{2.10}$$

We recall that the continuity of the control was stated in **(A1)**. A sufficient condition for the continuity of the control, stronger than **(A3)**, is (see [4, Thm. 2]) the uniform

strong convexity of the Hamiltonian w.r.t. the control variable, i.e. there exists $\alpha > 0$, such that

$$H_{uu}[\bar{p}(t)](\hat{u}, \bar{y}(t)) \geq \alpha, \text{ for all } \hat{u} \in \mathbb{R}^m \text{ and } t \in [0,T]. \tag{2.11}$$

Observe that, when $\bar{u}$ and $\bar{\eta}$ are Lipschitz continuous, denoting by $\nu(t)$ the density of $\bar{\eta}$, the costate equation can be written in the form

$$-\dot{\bar{p}}(t) = \bar{p}(t) f_y(\bar{u}(t), \bar{y}(t)) + \sum_{i=1}^{r} \nu_i(t) g_i'(\bar{y}(t)) \quad \text{a.e. on } (0,T); \quad \bar{p}(t) = \phi'(\bar{y}(T)). \tag{2.12}$$

LEMMA 2.3. *Let* **(A0)**-**(A3)** *hold. Then both* $\bar{u}$ *and* $\bar{\eta}$ *are Lipschitz continuous, and* (2.12) *holds.*

Proof. The result is proved in [14, Thm 4.2] in the case of a convex problem, using a Lemma on 'compatible pairs'. It was generalized in [1], using the same Lemma, to non convex problems with state constraints of any order. □

In the rest of the paper we assume as standing hypothesis **(A0)**-**(A3)**.

2.3. Second Order Conditions and Alternative Formulation. Let us first recall some theoretical issues about second-order conditions. We introduce the linearized control and state space $\mathcal{V} := L^2(0,T)$ and $\mathcal{Z} := H^1(0,T;\mathbb{R}^n)$, resp. We use also the notations

$$H[t] := H[\bar{p}(t)](\bar{u}(t), \bar{y}(t)); \quad g[t] := g(\bar{y}(t)), \quad f[t] := f(\bar{u}(t), \bar{y}(t));$$

as well as for their partial derivatives, e.g. $H_u[t] := H_u[\bar{p}(t)](\bar{u}(t), \bar{y}(t))$, and other functions. Let us define the quadratic form over $\mathcal{V} \times \mathcal{Z}$, where $z = z[v]$:

$$\Omega(v) := \int_0^T \left(H_{yy}[t](v(t), z(t))^2 + \sum_{i=1}^{r} \nu_i(t) g_i''[t](z(t))^2 \right) \mathrm{d}t + \phi''(\bar{y}(T))(z(T))^2, \tag{2.13}$$

and the set $C(\bar{u})$ of *strict critical directions* is defined as those $v \in \mathcal{V}$ such that, for $z = z[v]$:

$$\dot{z} = f_u(\bar{u}, \bar{y})v + f_y(\bar{u}, \bar{y})z \text{ on } [0,T]; \;\; z(0) = 0, \tag{2.14}$$

$$g_i'(\bar{y}(t))z(t) = 0, \;\; t \in I_i \tag{2.15}$$

$$\phi'(\bar{y}(T))z(T) = 0. \tag{2.16}$$

Note that in the last relation we write an equality instead of an inequality since this is known to be equivalent for qualified solutions, which will be the case thanks to assumption (**A2**).

Let us next recall the *alternative formulation* of the optimality conditions, due to [8] and [17], and put on a sound mathematical basis by [20]. (See also [3]). The *alternative Hamiltonian* $H : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n*} \times \mathbb{R}^{r*} \to \mathbb{R}$ is defined by:

$$\tilde{H}[p^1, \bar{\eta}^1](u,y) := p^1 f(u,y) + \bar{\eta}^1 g^{(1)}(u,y). \tag{2.17}$$

Now define the *alternative costate and multiplier* of the state constraint:

$$p^1(t) := \bar{p}(t) + \sum_{i=1}^{r} \bar{\eta}_i(t) g_i'(\bar{y}(t)); \quad \bar{\eta}^1(t) := -\bar{\eta}(t), \quad t \in (0,T).$$

It is easily checked that

$$-\dot{p}^1(t) = \tilde{H}_y[p^1(t), \bar{\eta}^1(t)](\bar{u}(t), \bar{y}(t)) \quad \text{a.e. on } (0,T); \quad p^1(T) = \phi'(\bar{y}(T)). \tag{2.18}$$

At the same time, for any $u \in \mathbb{R}$, we have that

$$\tilde{H}[p^1(t), \bar{\eta}^1(t)](u, \bar{y}(t)) = (p^1(t) + \bar{\eta}^1(t) g'(\bar{y}(t)))f(u, \bar{y}(t)) = H[\bar{p}(t)](u, \bar{y}(t)).$$

Consequently the property of stationarity or minimization of the Hamiltonian w.r.t. the control holds for the original Hamiltonian H if and only if it holds for the alternative Hamiltonian $\tilde{H}$.

The corresponding alternative quadratic form, where $z = z[v]$, has the following expression:

$$\tilde{\Omega}(v) := \int_0^T \tilde{H}_{yy}[t](v(t), z(t))^2 \mathrm{d}t + \phi''(\bar{y}(T))(z(T))^2.$$

The form above involves the expression of $D^2 g^{(1)}[t]$, which is easily checked to be

$$D^2 g^{(1)}[t](v,z)^2 = g^{(3)}[t](f[t], z, z) + g'[t] f_{yy}[t](v,z)^2 + 2g''[t](z, f_y[t](v,z)). \tag{2.19}$$

The next Lemma is a variant of some results by Bonnans and Hermant [3], Malanowski and Maurer [19]. We give a short, direct proof in the case of a single constraint for convenience.

LEMMA 2.4. *We have that $\tilde{\Omega}(v) = \Omega(v)$, for all $v \in \mathcal{U}$.*

Proof. Using (2.19), we get that

$$\begin{aligned}
\int_0^T \nu(t) g''[t](z(t))^2 \mathrm{d}t &= \int_0^T g''[t](z(t))^2 \mathrm{d}\bar{\eta}(t) = \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\left(g''[t](z(t))^2\right)\bar{\eta}^1(t)\mathrm{d}t \\
&= \int_0^T \left(g^{(3)}(\bar{y}(t))(f[t], z(t), z(t)) + 2g''[t](z(t), \dot{z}(t))\right)\bar{\eta}^1(t)\mathrm{d}t \\
&= \int_0^T \left(D^2 g^{(1)}[t](v(t), z(t))^2 - g'[t] f_{yy}[t](v(t), z(t))^2\right)\bar{\eta}^1(t)\mathrm{d}t.
\end{aligned}$$

Eliminating $\bar{p}(t) = p^1(t) + \bar{\eta}^1 g'[t]$ in the expression of $\Omega(\cdot)$, we obtain that

$$\begin{aligned}
\Omega(v) &= \int_0^T \left((p^1(t) + \bar{\eta}^1(t) g'[t]) f_{yy}[t](v(t), z(t))^2 + \nu(t) g''[t](z(t))^2\right)\mathrm{d}t \\
&\quad + \phi''(\bar{y}(T))(z(T))^2 \\
&= \int_0^T \left(p^1(t) f_{yy}[t](v(t), z(t))^2 + \bar{\eta}^1(t) D^2 g^{(1)}[t](v(t), z(t))^2\right)\mathrm{d}t \\
&\quad + \phi''(\bar{y}(T))(z(T))^2,
\end{aligned}$$

as was to be proved. □

2.4. Discrete version. We introduce now the Euler discretization of the optimal control problem (2.1). Given some non zero $N \in \mathbb{N}$ and a collection of positive time steps h_k, $k = 0$ to $N-1$, such that $\sum_{k=0}^{N-1} h_k = T$, we set

$$t_k := \sum_{i=0}^{k-1} h_i, \qquad k = 0, \ldots, N; \qquad \bar{h} = \max_{k=0,\ldots,N-1} h_k,$$

and we consider the discretized problem

$$(\mathcal{P}_d)\begin{cases} \text{Minimize } \phi(y_N); & \text{subject to} \\ y_{k+1} = y_k + h_k f(u_k, y_k), & \text{for } k = 0, \dots, N-1; \\ y_0 = y_0; & \\ g(y_k) \le 0 & \text{for } k = 1, \dots, N. \end{cases} \tag{2.20}$$

We denote by $\mathcal{U}^N$ the space of discrete control variable. The associated Lagrangian function (with a proper scaling of the state constraint) is

$$\phi(y_N) + \sum_{k=0}^{N-1} p_{k+1}(y_k + h_k f(u_k, y_k) - y_{k+1}) + \sum_{k=1}^{N} h_k \nu_k g(y_k),$$

where $\nu_k g(y_k) = \sum_{i=1}^{r} \nu_{k,i} g_i(y_k)$. The first-order optimality conditions (in qualified form), for this finite dimensional optimization problem with finitely many equalities and inequalities, are

$$\begin{array}{ll} p_k & = p_{k+1} + h_k p_{k+1} f_y(u_k, y_k) + h_k \nu_k g'(y_k), \quad k = 0, \dots, N-1, \\ p_N & = \phi'(y_N) + h_N \nu_N g'(y_N), \\ 0 & = H_u[p_{k+1}](u_k, y_k), \quad k = 0, \dots, N-1, \\ g_i(y_k) & \le 0, \;\; \nu_{k,i} \ge 0; \;\; \nu_{k,i} g_i(y_k) = 0, \quad i = 1, ..., r, \quad k = 0, \dots, N. \end{array} \tag{2.21}$$

Analogously to the continuous case we define also the 'integrated' multiplier of the normal and the alternative formulation

$$\eta_k := -\sum_{j=k}^{N} h_k \nu_k, \qquad \bar{\eta} := -\eta, \tag{2.22}$$

so that $h_k \nu_k = \eta_{k+1} - \eta_k = \bar{\eta}_k - \bar{\eta}_{k+1}$, for $k = 0, \dots, N$.

DEFINITION 2.5. *We say that the discretization step is constant if*

$$h_0 = h_1 = \cdots = h_{N-1}. \tag{2.23}$$

2.5. Main result. As mentioned before, our results hold in two different cases. We need to preserve the coercivity of the Hessian of the Lagrangian over some subspace; this can be stated as hypothesis, as in [13] or obtained under structure hypotheses for a scalar state constraints. So we will assume that one of the following assumptions hold:

(A4) There exists a constant $\alpha > 0$ such that

$$\Omega(v) \ge \alpha \int_0^T |v(t)|^2 \mathrm{d}t, \qquad \text{for all } v \in \mathcal{U}$$

and all discrete steps are of the same order, i.e.

$$\max_k \left(h_k/h_{k-1} + h_{k-1}/h_k\right) = O(1). \tag{2.24}$$

The condition on Ω is known to be a sufficient condition for local optimality in $\mathcal{U}$. This follows from [22, Th. 5.6].

(A5) *(scalar constraint and finite structure)* Assume that $r = 1$, the discretization step is constant, the set I is a finite union of boundary arcs, the density ν is uniformly positive over the boundary arcs, and there exists a constant $\alpha > 0$ such that

$$\Omega(v) \geq \alpha \int_0^T |v(t)|^2 \mathrm{d}t, \quad \text{whenever } v \in C(\bar{u}). \tag{2.25}$$

The set of strict critical directions $C(\bar{u})$ is defined in (2.14)-(2.16).

This condition on Ω is known to be a sufficient condition for local optimality in $\mathcal{U}$, see [3, Th. 3.8].

Note that **(A5)** excludes *touch points*, i.e., isolated elements of I_i, for $i = 1, \ldots, r$. We denote by $\mathcal{I}^b$ the union of boundary arcs, of the form $\mathcal{I}^b := \cup_{j=1}^{N_b} [\mathcal{T}_j^{en}, \mathcal{T}_j^{ex}]$ where we have ordered the set of entry points $\mathcal{T}^{en} := \{\mathcal{T}_1^{en} < \cdots < \mathcal{T}_{N_b}^{en}\}$, and similarly for the set $\mathcal{T}^{ex}$ of exit points.

THEOREM 2.6. *Let assumptions* **(A0)**-**(A3)** *hold as well as either* **(A4)** *or* **(A5)**. *Then the discrete optimal control problem* $(\mathcal{P}_d)$ *has a local solution* (u^h, y^h) *with associated multipliers* (p^h, η^h), *such that*

$$\|y^h - \bar{y}\|_\infty + \|u^h - \bar{u}\|_\infty + \|p^h - \bar{p}\|_\infty + \|\eta^h - \bar{\eta}\|_\infty = O(\bar{h}). \tag{2.26}$$

The rest of the paper will be dedicated to the proof of this result; for that purpose we need to introduce a special auxiliary problem.

3. Homotopy path. We consider a family $(\mathcal{P}^\theta)$ of perturbed discrete optimization problems, parametrized by $\theta \in [0, 1]$. The definition is done such that for $\theta = 0$, the problem reduces to $(\mathcal{P}_d)$. As we will see in the next Section, a certain sampling of the solution of the continuous problem $(\mathcal{P})$ happens to be a solution of the optimality system of $(\mathcal{P}^1)$. The perturbed optimal control problem is, for some perturbations terms δ^p, δ^u, δ^y, δ^g, to be defined later, such that $\|\delta^\cdot\|_\infty = O(1)$,

$$(\mathcal{P}^\theta) \begin{cases} \text{Minimize } \phi(y_N^\theta) + \theta \sum_{k=0}^{N-1} h_k^2 (\delta_k^p y_k^\theta + \delta_k^u u_k^\theta); & \text{subject to} \\ y_{k+1}^\theta = y_k^\theta + h_k f(u_k^\theta, y_k^\theta) + \theta h_k^2 \delta_k^y, & \text{for } k = 0, \ldots, N-1; \\ g_i(y_k^\theta) \leq \theta h_k^2 \delta_k^g, & \text{for } k = 1, \ldots, N, \\ y_0^\theta = y_0, & i = 1, ..., r. \end{cases} \tag{3.1}$$

The corresponding optimality system is

$$\begin{array}{ll} p_k^\theta & = p_{k+1}^\theta + h_k p_{k+1}^\theta f_y(u_k^\theta, y_k^\theta) + h_k \nu_k^\theta g'(y_k^\theta) + \theta h_k^2 \delta_k^p, \\ p_N^\theta & = \phi'(y_N^\theta) + h_N \nu_N^\theta g'(y_N^\theta), \\ 0 & = H_u[p_{k+1}^\theta](u_k^\theta, y_k^\theta) + \theta h_k \delta_k^u, \end{array} \tag{3.2}$$

for $k = 1$ to $N - 1$, with the complementarity conditions, $i = 1, ..., r$,

$$g_i(y_k^\theta) - \theta h_k^2 \delta_k^g \leq 0, \;\; \nu_{k,i}^\theta \geq 0, \;\; \nu_{k,i}^\theta (g_i(y_k^\theta) - \theta h_k^2 \delta_k^g) = 0, \;\; k = 1, \ldots, N. \tag{3.3}$$

Let us set

$$f^k := (u_k^\theta, y_k^\theta); \quad H^k := H[p_{k+1}^\theta](u_k^\theta, y_k^\theta) \tag{3.4}$$

For future reference we note that the linearization of the costate equation is, denoting by $(v^\theta, z^\theta, q^\theta, \delta\nu^\theta)$ the linearizations of $(u^\theta, y^\theta, p^\theta, \nu^\theta)$:

$$\begin{aligned} q_k^\theta = \quad & q_{k+1}^\theta + h_k q_{k+1}^\theta f_y^k + h_k (v_k^\theta)^T H_{uy}^k + h_k (z_k^\theta)^T H_{yy}^k \\ & + h_k \nu_k^\theta (z_k^\theta)^T g''(y_k^\theta) + h_k \delta\nu_k^\theta g'(y_k^\theta) - \theta h_k^2 \delta_k^p, \\ q_N^\theta = \quad & (z_N^\theta)^T \phi''(y_N^\theta) + h_N \nu_N^\theta (z_N^\theta)^T g''(y_N^\theta) + h_N \delta\nu_N^\theta g'(y_N^\theta). \end{aligned} \tag{3.5}$$

The corresponding approximation of $\bar\eta$ is: (see (2.22))

$$\eta_k^\theta := -\sum_{j=k}^N h_j \nu_j^\theta, \quad k = 0, \dots, N. \tag{3.6}$$

The *sampling* of the continuous solution and the associated multipliers of the original problem (2.1) are defined by

$$\begin{cases} \hat u_k \quad := \bar u(t_k), \quad \hat y_k := \bar y(t_k), & k = 0, \dots, N-1, \\ \hat p_k \quad := \bar p(t_k), \quad \hat\nu_k := \displaystyle\int_{t_k}^{t_{k+1}} \nu(t)\mathrm{d}t, & k = 1, \dots, N, \end{cases} \tag{3.7}$$

and accordingly we can define

$$\hat\eta_k := \eta(t_{k+1}) = \sum_{j=k}^N h_k \hat\nu_k, \quad k = 1, \dots, N.$$

For $\theta = 1$ we define u_k^θ and the associated state and multipliers by

$$u_k^1 = \hat u_k,\ y_k^1 = \hat y_k,\ p_k^1 = \hat p_k,\ \eta_k^1 = \hat\eta_k, \text{ for } k = 1, ..., N. \tag{3.8}$$

We next give to the perturbation terms $(\delta_k^u, \delta_k^y, \delta_k^p, \delta_k^g)$ the unique value such that the above sampling is solution of the discretized problem for $\theta = 1$.

LEMMA 3.1. *We have that*

$$\|\delta^y\|_\infty + \|\delta^u\|_\infty + \|\delta^p\|_\infty + \|\delta^g\|_\infty = O(1). \tag{3.9}$$

Proof. Since u is Lipschitz continuous by Lemma 2.3, $t \to g(y(t))$ has a.e. a bounded second derivative. Therefore, if $\nu_k \neq 0$ there exists some $c > 0$ such that $g(y(t)) \geq -c\bar h^2$ for all $t \in [t_k, t_{k+1}]$, so that $\|\delta^g\|_\infty = O(1)$.

Next, if $w : [0, T] \to \mathbb{R}$ is C^1 with a Lipschitz continuous derivative of constant L, then by a first order Taylor expansion, we have that

$$|w(t+h) - w(t) - w'(t)h| \leq \tfrac{1}{2} L h^2.$$

By Lemma 2.3, the control is Lipschitz, and therefore, so does $\dot y(t)$; we deduce that $\|\delta^y\|_\infty = O(1)$. For the costate equation, we have that

$$\bar p(t_{k+1}) = \bar p(t_k) - \int_{t_k}^{t_{k+1}} \bar p(t) f_y[t]\mathrm{d}t - \int_{t_k}^{t_{k+1}} \nu(t) g'[t]\mathrm{d}t.$$

Now since $\bar u$, $\bar y$, and $\bar p$ are Lipschitz,

$$\left| \int_{t_k}^{t_{k+1}} \bar p(t) f_y(\bar u(t), \bar y(t))\mathrm{d}t - \int_{t_k}^{t_{k+1}} \bar p(t_{k+1}) f_y(\bar u(t_k), \bar y(t_k))\mathrm{d}t \right| = O(h_k^2),$$

and (in the parenthesis below we recognize the expression of ν_k):

$$\left| \int_{t_k}^{t_{k+1}} \nu(t) g'[t] \mathrm{d}t - \left(\int_{t_k}^{t_{k+1}} \nu(t) \mathrm{d}t \right) g'(\bar{y}(t_k)) \right| = O(h_k^2).$$

It follows that $\|\delta^p\|_\infty = O(1)$. Since p is Lipschitz and $H_u[p(t_k)](\bar{u}(t_k), \bar{y}(t_k)) = 0$ we easily deduce that $\|\delta^u\|_\infty = O(1)$. The conclusion follows. □

In the rest of the paper some of the estimates presented are valid in a special neighborhood of the continuous solution. To state them rigorously, we need the following definition: given $\varepsilon > 0$ and $\theta \in [0,1]$, we say that a solution $X^\theta := (u^\theta, y^\theta, y^\theta, \eta^\theta)$ of the optimality system (3.2) is an *ε-neighboring solution* if we have that

$$\|u^\theta - \hat{u}\|_\infty + \|y^\theta - \hat{y}\|_\infty + \|p^\theta - \hat{p}\|_\infty + \|\eta^\theta - \hat{\eta}\|_\infty \leq \varepsilon, \tag{3.10}$$

and we define

$$\theta_m := \inf\{\theta \in [0,1];\ (3.2) \text{ has an } \varepsilon\text{-neighbouring solution}\}.$$

When $\theta = 1$, the l.h.s. of (3.10) has value 0, and therefore θ_m is well-defined with value in $[0,1]$.

Through the auxiliary structure of the homotopy path problem we are now able to prove the bounds shown in Theorem 2.6. In the sequel we will analyze some technical points. In particular we will use the fact that the solution of $(\mathcal{P}^\theta)$ is uniformly Lipschitz continuous (the result is contained in Section 4) and that such solution is in a ε-neighborhood of the solution of $(\mathcal{P})$, $X^1 = (\bar{u}, \bar{y}, \bar{p}, \bar{\eta})$; this point will be discussed in Section 5. Note that we define the Lipschitz constant of a function defined for discrete times, as e.g. for u^θ, by

$$Lip(u^\theta) := \max\left(|u^k - u^{k-1}|/h_k;\ k = 0, \ldots, N-1 \right).$$

Proof. [of Th. 2.6.] We prove in Section 4 that, if $\bar{h}$ is small enough, for $\theta \in [\theta_m, 1]$, then (u^θ, y^θ) is uniquely defined, has unique associated multipliers (p^θ, η^θ), and setting $X^\theta := (u^\theta, y^\theta, p^\theta, \eta^\theta)$, (see Section 5) $\theta \to X^\theta$ has Lipschitz constant of order $\bar{h}$. It follows that, for a fix $\varepsilon > 0$, $\|X^\theta - X^1\|_\infty < \varepsilon$ when $\bar{h}$ is small enough, which gives a contradiction if $\theta_m > 0$. Therefore, X^0 is well-defined and $\|X^1 - X^0\|_\infty = O(\bar{h})$, as was to be shown. □

4. Regularity of the solutions. In this Section we present some regularity results for the solutions of the Homotopy path problem.
Given an $\varepsilon-$neighboring solution X^θ of $(\mathcal{P}^\theta)$, we will prove the uniqueness and the uniform Lipschitz continuity of X^θ. In the following, when there is no possibility of confusion, we drop θ as upper index in the notation for a better readability, keeping the complete notation in the statements of the main propositions. We need to define

$$\begin{array}{rcl} C_k^1 & := & \dfrac{p_k}{h_k} \left(f_u(u_{k-1}, y_k) - f_u(u_{k-1}, y_{k-1}) \right), \\ C_k^2 & := & p_{k+1} f_y(u_k, y_k) f_u(u_k, y_k), \\ \Delta_k^u & := & h_k \delta_k^u - h_{k-1} \delta_{k-1}^u \\ & = & (\hat{p}_k - \hat{p}_{k+1}) f_u(\hat{u}_k, \hat{y}_k) - (\hat{p}_{k-1} - \hat{p}_k) f_u(\hat{u}_{k-1}, \hat{y}_{k-1}). \\ C_k^3 & := & C_k^1 - C_k^2 - \theta h_k \delta_k^p f_u(u_k, y_k) + \theta \Delta_k^u / h_k. \end{array} \tag{4.1}$$

Observe that, since y_k is uniformly Lipschitz we have that:

$$\text{(i)}\ \ C_k^1 = \frac{h_{k-1}}{h_k} H_{uy}[p_k](u_{k-1}, y_k) f(u_{k-1}, y_{k-1}) + O(\bar{h}); \quad \text{(ii)}\ \ |C_k^3| = O(1). \tag{4.2}$$

LEMMA 4.1. (i) *Let X^θ be an ε-neighboring solution of $(\mathcal{P}^\theta)$. Then there exists $c_{\mathcal{H}} > 0$ such that, if $\varepsilon > 0$ and $\bar{h}$ are small enough, then*

$$\sum_{j=1}^{r} \nu_{k,j}^\theta \, \nabla_u g_j^{(1)}(u_k^\theta, y_k^\theta) = \mathcal{H}_k^\theta \frac{(u_k^\theta - u_{k-1}^\theta)}{h_k} + C_k^3, \tag{4.3}$$

where $\mathcal{H}_k^\theta$ satisfies

$$|\mathcal{H}_k^\theta - H_{uu}[p_k](u_k, y_k)| \leq c_{\mathcal{H}} \varepsilon, \ \ k = 0, \ldots, N-1. \tag{4.4}$$

(ii) *Let $\varepsilon' > 0$. If in addition, the time step is constant, t_k belongs to a boundary arc (t_a, t_b), and $t_{k_a} + \varepsilon' < t_{k-1} < t_k < t_{k_b} - \varepsilon'$, then the variation of C_k^3 along the homotopy path is of order $O(\bar{h})$.*

Proof. (i) Note that $h_k \delta_k^u = (\hat{p}_k - \hat{p}_{k+1}) f_u(\hat{u}_k, \hat{y}_k)$. By the optimality condition (3.2), we have that

$$\begin{array}{rcl} 0 & = & H_u[p_{k+1}](u_k, y_k) - H_u[p_k](u_{k-1}, y_{k-1}) + \theta \Delta_k^u \\ & = & p_{k+1} f_u(u_k, y_k) - p_k f_u(u_{k-1}, y_{k-1}) + \theta \Delta_k^u \\ & = & (p_{k+1} - p_k) f_u(u_k, y_k) + p_k \left[f_u(u_k, y_k) - f_u(u_{k-1}, y_{k-1}) \right] + \theta \Delta_k^u \\ & = & (p_{k+1} - p_k) f_u(u_k, y_k) + p_k \left[f_u(u_k, y_k) - f_u(u_{k-1}, y_k) \right] + h_k C_k^1 + \theta \Delta_k^u. \end{array} \tag{4.5}$$

By the mean-value Theorem, we deduce that

$$p_k (f_u(u_k, y_k) - f_u(u_{k-1}, y_k)) = \mathcal{H}_k (u_k - u_{k-1}),$$

where

$$\mathcal{H}_k := p_k \int_0^1 f_u(u_{k-1} + \sigma(u_{k-1} - u_k), y_k)(u_{k-1} - u_k) \mathrm{d}\sigma,$$

so that (4.4) holds. We conclude by combining (4.5) and the discrete costate equation in (3.2), where $\nabla_u g_i^{(1)}(u_k, y_k) = g_i'(y_k) f_u(u_k, y_k)$.

(ii) It is easily checked that C_k^1 and C_k^2 satisfy this property, as well as (it is of order of $\bar{h}$) $\theta h_k \delta_k^p f_u(u_k, y_k)$. Since the time step is constant, we have that by (4.1) $\Delta_k^u / h_k = \delta_k^u - \delta_{k-1}^u$ is of order of $\bar{h}$ over the interior of a boundary arc. The conclusion follows. □

Let us define

$$\begin{array}{rcl} \Delta_g^{k,i} := & \dfrac{g_i(y_{k+1}) - g_i(y_k)}{h_k} - \dfrac{g_i(y_k) - g_i(y_{k-1})}{h_{k-1}}. \\ \Xi_k^\theta := & g_i'(y_k) \left(f(u_{k-1}, y_k) - f(u_{k-1}, y_{k-1}) \right) \\ & + \theta g_i'(y_k)(h_k \delta_k^y - h_{k-1} \delta_{k-1}^y) + \frac{1}{2} h_k g_i''(y_k) f(u_k, y_k)^2 \\ & + \frac{1}{2} h_{k-1} g_i''(y_k) f(u_{k-1}, y_{k-1})^2 \end{array} \tag{4.6}$$

LEMMA 4.2. *We have that:*

a) the following relation holds

$$\Delta_g^{k,i} = g_i'(y_k) f_u(y_k, u_k)(u_k - u_{k-1}) + \Xi_k^\theta + O(\bar{h}^2) + O(\varepsilon |u_k - u_{k-1}|). \quad (4.7)$$

b) If in addition the time step is constant, then

$$\|\Xi^\theta - \Xi^1\|_\infty = O(\bar{h}\varepsilon). \quad (4.8)$$

Proof. Use

$$\begin{array}{rcl} g_i(y_{k+1}) - g_i(y_k) & = & g_i'(y_k)(y_{k+1} - y_k) + \frac{1}{2} g_i''(y_k)(y_{k+1} - y_k)^2 + O(h_k^3), \\ & = & h_k g_i'(y_k) f(u_k, y_k) \\ & & + \frac{1}{2} h_k^2 \left(2\theta g'(y_k) \delta_k^y + g_i''(y_k) f(u_k, y_k)^2 \right) + O(h_k^3), \\ g_i(y_{k-1}) - g_i(y_k) & = & g_i'(y_k)(y_{k-1} - y_k) + \frac{1}{2} g_i''(y_k)(y_{k-1} - y_k)^2 + O(h_k^3), \\ & = & -h_{k-1} g_i'(y_k) f(u_{k-1}, y_{k-1}) \\ & & + \frac{1}{2} h_{k-1}^2 \left(-2\theta g_i'(y_k) \delta_{k-1}^y + g_i''(y_k) f(u_{k-1}, y_{k-1})^2 \right) \\ & & + O(h_{k-1}^3). \end{array}$$

We obtain *a*) by dividing these relations by h_k and h_{k-1} respectively, adding them and observing that

$$\begin{aligned} f(u_k, y_k) - f(u_{k-1}, y_{k-1}) &= f(u_k, y_k) - f(u_{k-1}, y_k) + f(u_{k-1}, y_k) - f(u_{k-1}, y_{k-1}) \\ &= f_u(u_k, y_k)(u_k - u_{k-1}) + (f(u_{k-1}, y_k) - f(u_{k-1}, y_{k-1})) + O(\varepsilon |u_k - u_{k-1}|). \end{aligned}$$

Since $|y_k - y_{k-1}| = O(h_{k-1})$, point (b) follows using that $h_k = h_{k-1}$ and $|\delta_k^y - \delta_{k-1}^y| = O(\bar{h})$. □

Now we are ready to obtain the uniform Lipschitz estimates of the variables of the perturbed problem. A similar result, in the case of a linear quadratic optimal control problem was obtained in [12]. Let us set

$$w_k = \nu_k \nabla_u g^{(1)}(u_k, y_k) = \sum_{i=1}^r \nu_{k,i} \nabla_u g_i^{(1)}(u_k, y_k). \quad (4.9)$$

Lemma 4.3. *We have that*

$$w_k^T \mathcal{H}_k^{-1} w_k = \frac{1}{h_k} w_k \cdot (u_k - u_{k-1}) + w^T \mathcal{H}_k^{-1} C_k^3, \quad (4.10)$$

as well as, called Lip(u^θ) and Lip(p^θ) the Lipschitz constant of u^θ and p^θ,

$$Lip(u^\theta) + Lip(p^\theta) + \|\nu^\theta\|_\infty = O(1).$$

Proof. By the Legendre-Clebsch condition **(A3)**, for small enough $\varepsilon > 0$, $\mathcal{H}_k$ is uniformly invertible. Computing the scalar product of both sides of (4.3) by $w_k^T \mathcal{H}_k^{-1}$, we obtain (4.10).
When $\nu_{k,i} \neq 0$ we have that $g_{k,i}^\theta := g_i(y_k) - \theta h_k^2 \delta_{k,i}^g$ reaches a local maximum and therefore

$$0 \geq \frac{g_{k+1,i}^\theta - g_{k,i}^\theta}{h_k} - \frac{g_{k,i}^\theta - g_{k-1,i}^\theta}{h_{k-1}},$$

which amounts to

$$\Delta_g^{k,i} \leq \theta \left(\frac{h_{k+1}^2 \delta_{k+1,i}^g - h_k^2 \delta_{k,i}^g}{h_k} - \frac{h_k^2 \delta_{k,i}^g - h_{k-1}^2 \delta_{k-1,i}^g}{h_{k-1}} \right) \leq O(\bar{h}). \tag{4.11}$$

Since $|C_k^3| = O(1)$ as already noticed, it follows from (4.3) that $|u_k - u_{k-1}|/h_k = O(|\nu_k| + 1)$. Now putting it together with (4.7) and (4.11), it follows that

$$\frac{1}{h_k} w_k \cdot (u_k - u_{k-1}) + w^T \mathcal{H}^{-1} C_k^3 \leq \frac{1}{h_k} \sum_{i=1}^{r} \nu_k \Delta_g^{k,i} + O(1) + O(|\nu_k|), \tag{4.12}$$

then for the linear independence of the constraints w.r.t. the control (Assumption **(A2)**, i.e. $|w_k| \geq \alpha \|\nu_k\|$), in the l.h.s. of (4.10), we have

$$\alpha |\nu_k|^2 \leq O(|\nu_k|) + O(1). \tag{4.13}$$

By the above display, $|\nu_k| = O(1)$, and by (4.3), u_k is uniformly Lipschitz. By (3.2), so is the discretized costate. □

5. Sensitivity analysis.

5.1. Characterization of directional derivatives. In this Section we complete the proof of Theorem 2.6 showing that a solution (u^θ, y^θ) of the perturbed problem $(\mathcal{P}^\theta)$ is in a L^∞ neighborhood of $(\bar{u}, \bar{y})$, local solution of the problem $(\mathcal{P})$. The strategy consists in establishing that the path $X^\theta := (u^\theta, y^\theta, p^\theta, \eta^\theta)$ is Lipschitz, and then to show that it has directional derivatives $\delta X^\theta := (v^\theta, z^\theta, q^\theta, \delta\eta^\theta)$ satisfying $\|\delta X^\theta\|_\infty = O(\bar{h})$. This will imply that the Lipschitz constant of X^θ (in the L^∞ norm) is of order $O(\bar{h})$. We define (compare to (2.22)) $\nu_k^\theta := \eta_{k+1}^\theta - \eta_k^\theta$, $k = 0$ to N.

The fact that X^θ is Lipschitz is a consequence of Robinson's theory for strong regularity [25] and its application to nonlinear programming see e.g. [2, Sec. 5.1]. This theory gives a sufficient condition for Lipschitz stability of the local solution and associated multiplier, provided that we have the (i) linear independence of gradients of active constraints, which by (**A2**) always holds, and (ii) positivity of the Hessian of the Lagrangian over an *extended critical cone*, obtained by removing inequality constraints associated with zero components of the multiplier, as well as the condition on the linearization of the cost function. In addition, under these conditions, Jittorntrum's result [18] states that directional derivatives exist and that they are solution of the following quadratic programming problem

$$(QP) \begin{cases} \underset{v}{\text{Min}} \ \frac{1}{2}\Omega^\theta(v,z) - \theta \sum_{k=0}^{N-1} h_k^2 (\delta_k^p z_k + \delta_k^u v_k) & \text{s.t. } z = z^\theta[v] \text{ and} \\ g_i'(y_k) z_k = -\theta h_k^2 \delta_{k,i}^g & k \in I_+^{i,\theta}, \ s.t. \, \nu_{k,i}^\theta \neq 0, \\ g_i'(y_k) z_k \leq -\theta h_k^2 \delta_{k,i}^g & k \in I_0^{i,\theta}, \ s.t. \, \nu_{k,i}^\theta = 0, \end{cases} \tag{5.1}$$

Here $v \in \mathcal{V}^N$, $z^\theta[v] \in \mathcal{Z}^N$ is defined as the unique solution of the linearized state equation

$$z_{k+1}^\theta = z_k^\theta + h_k f_y(u_k, y_k)(v_k, z_k^\theta) - \theta h_k^2 \delta_k^y, \quad k = 0, ..., N-1, \quad i = 1, ..., r, \tag{5.2}$$

and the set of constraints $I_+^{i,\theta}$ and $I_0^{i,\theta}$ are defined as the inequality constraints of problem $(\mathcal{P}^\theta)$ that are active at y^θ, i.e.:

$$\begin{cases} I_+^{i,\theta} & := \{k = 0, \ldots, N; \ \nu_{k,i}^\theta > 0\}, \\ I_0^{i,\theta} & := \{k = 0, \ldots, N; \ \nu_{k,i}^\theta = g_i(y_k^\theta) = 0\}. \end{cases} \tag{5.3}$$

Finally, the Hessian of Lagrangian of the discretized problem is, with obvious notations, setting $H_k^\theta := p_{k+1}^\theta f(u_k^\theta, y_k^\theta)$, and for $z^\theta = z^\theta[v]$:

$$\Omega^\theta(v,z) := \sum_{k=0}^{N-1} h_k D^2 H_k^\theta(v_k,z_k)^2 + \sum_{k=0}^{N} h_k \nu_k^\theta D^2 g_k^\theta(z_k)^2 + \phi''(y_N^\theta)(z_N)^2. \tag{5.4}$$

We anticipate a result that will be shown in details later (Corollary 5.6) about the existence and uniqueness of the solution of (QP):

PROPOSITION 5.1. *Let* (**A0**)−(**A3**) *and, either* (**A4**) *or* (**A5**) *hold. Then problem* (QP) *has a unique solution* $\delta X^\theta := (v^\theta, z^\theta, q^\theta, \delta\eta^\theta)$.

Let us introduce the following *alternative formulation*: we underline the analogy with the alternative formulation recalled in Section 2.3. We first define the set of inequality constraints that are active at the solution of (QP):

$$I^{i,\theta} := I_+^{i,\theta} \cup \{k \in I_0^{i,\theta};\ \ g_i'(y_k^\theta) z_k^\theta = 0\}. \tag{5.5}$$

For $i = 1, \dots, r$, denote by $k[i,1] < \cdots < k[i, M_i^\theta]$ the elements of $I^{i,\theta}$, set $k[i,0] = 0$, and for $j = 0, \dots, M_i^\theta - 1$:

$$\left\{ \begin{array}{ll} \Delta t_{i,j} & := t_{k[i,j+1]} - t_{k[i,j]}, \\ b_{i,j}^E & := -\theta(h_{k[i,j+1]}^2 \delta_{k[i,j+1]}^g - h_{k[i,j]}^2 \delta_{k[i,j]}^g)/\Delta t_{i,j}. \end{array} \right. \tag{5.6}$$

Set

$$G_k(v_k,z_k) := \frac{g'(y_{k+1}) - g'(y_k)}{h_k} z_k + g'(y_{k+1}) \left(f'(u_k,y_k)(v_k,z_k) - \theta h_k \delta_k^y \right). \tag{5.7}$$

If $z = z^\theta[v]$, then

$$G_k(v_k,z_k) = \frac{g'(y_{k+1}) z_{k+1} - g'(y_k) z_k}{h_k}. \tag{5.8}$$

Since $z_0 = 0$, it follows that:

$$g'(y_k^\theta) z_k = \sum_{q=0}^{k-1} h_q G_q(v_q, z_q). \tag{5.9}$$

So the solution of (QP) satisfies the following equality constraints, denoting by $G_{i,k}$ the i-th component of G_k:

$$\sum_{q=0}^{k[i,j]-1} h_q G_{i,q}(v_q,z_q) = -\theta h_k^2 \delta_{i,k[i,j]}^g, \quad i = 1, \dots, r, \ \ j = 1, \dots M_i^\theta - 1. \tag{5.10}$$

An equivalent set of equality constraints is

$$b_{i,j}^E - \frac{1}{\Delta t_{i,j}} \sum_{k=k[i,j]}^{k[i,j+1]-1} h_k G_{i,k}(v_k,z_k) = 0, \quad i = 1, \dots, r, \ \ j = 0, \dots M_i^\theta - 1. \tag{5.11}$$

The corresponding quadratic problem is

$$(QP_E)\left\{\begin{array}{l}\underset{v}{\mathrm{Min}}\ \frac{1}{2}\Omega^\theta(v,z)-\theta\displaystyle\sum_{k=0}^{N-1}h_k^2(\delta_k^p z_k+\delta_k^u v_k)\\ \quad\text{s.t. } z=z^\theta[v] \text{ and (5.11).}\end{array}\right. \tag{5.12}$$

We call v^θ the solution of this problem, and $\delta\bar\eta^\theta$ the multiplier associated with constraints (5.11) and $\hat q$ the costate. The Lagrangian of (QP_E) is by the definition

$$\begin{array}{ll}\Omega^\theta(v,z) & +\displaystyle\sum_{k=0}^{N-1}\hat q_{k+1}(h_k f_k'(v_k,z_k)+z_k-z_{k+1})\\ & +\displaystyle\sum_{i,j}\Delta t_{i,j}\delta\bar\eta^\theta_{i,j}\left(b^E_{i,j}-\sum_{k=k[i,j]}^{k[i,j+1]-1}h_k G_k(v_k,z_k)/\Delta t_{i,j}\right).\end{array} \tag{5.13}$$

The optimality conditions of (QP_E) have the following form. The costate equation is

$$\begin{array}{ll}\hat q^\theta_k= & \hat q^\theta_{k+1}+h_k\hat q^\theta_{k+1}f_y^k+h_k(v^\theta_k)^T H^k_{uy}+h_k(z^\theta_k)^T H^k_{yy}+h_k\nu^\theta_k(z^\theta_k)^T g''(y_k)\\ & -\displaystyle\sum_{i=1}^r\delta\bar\eta^\theta_{i,j[i,k]}\left(g'(y_{k+1})-g'(y_k)+h_k g'(y_{k+1})f_y(u_k,y_k)\right)-\theta h_k^2\delta^p_k;\\ \hat q^\theta_N= & (z^\theta_N)^T\phi''(y^\theta_N)+h_N\nu^\theta_N(z^\theta_N)^T g''(y_N).\end{array} \tag{5.14}$$

Given $k\le M^\theta_i$, set

$$j[i,k]:=\min\{j\in I^{i,\theta};\ \ j\ge k+1\}. \tag{5.15}$$

Expressing the stationarity of the Lagrangian w.r.t. v we get that

$$(v_k)^T H^k_{uu}+(z_k)^T H^k_{uy}+\left(\hat q^\theta_{k+1}-\sum_{i=1}^r\delta\bar\eta^\theta_{i,j[i,k]}g_y(u_k,y_k)\right)f^k_u=0. \tag{5.16}$$

This suggests to define

$$\begin{array}{ll}\tilde q^\theta_{k+1} & :=\hat q^\theta_{k+1}-\displaystyle\sum_{i=1}^r\delta\bar\eta^\theta_{i,j[i,k]}g'(y_{k+1}),\quad k=0,\dots,N-1,\\ \delta\bar\nu^\theta_k & :=\delta\bar\eta^\theta_{i,j[i,k]}-\delta\bar\eta^\theta_{i,j[i,k-1]},\quad k=0,\dots,N-1,\end{array} \tag{5.17}$$

Then $\tilde q^\theta$ is solution of

$$\begin{array}{ll}\tilde q^\theta_k= & \tilde q^\theta_{k+1}+h_k\tilde q^\theta_{k+1}f_y^k+h_k(v^\theta_k)^T H^k_{uy}+h_k(z^\theta_k)^T H^k_{yy}+h_k(z^\theta_k)^T g''(y_k)\\ & +h_k\delta\bar\nu^\theta_k g_i'(y_k)-\theta h_k^2\delta^p_k,\\ \tilde q^\theta_N= & (z^\theta_N)^T\phi''(y^\theta_N)+h_N\nu^\theta_N(z^\theta_N)^T g''(y^\theta_N)+h_N\delta\bar\nu^\theta_N g'(y^\theta_N).\end{array} \tag{5.18}$$

In addition, we observe that if $\delta\bar\nu^\theta_k\neq 0$, then $\delta\bar\eta^\theta_{i,j[i,k]}>\delta\bar\eta^\theta_{i,j[i,k-1]}$, and therefore the i-th state constraint is active at step k. Since (QP) has a unique multiplier, we deduce that $\tilde q^\theta=q^\theta$ and $\delta\bar\eta^\theta=\delta\bar\eta^\theta_{i,j[i,k-1]}$, and

$$\delta\bar\nu^\theta_k=\delta\nu^\theta_k;\quad \delta\bar\eta^\theta_{i,j[i,k]}=\delta\eta^\theta_{i,k+1}. \tag{5.19}$$

5.2. Uniform surjectivity. We write here the *linear mappings* involved in the tangent quadratic problem, starting with

$$z^L_{k+1} = z^L_k + h_k f_y(u_k, y_k)(v_k, z^L_k), \quad k = 0, \cdots, N-1; \;\; z^L_0 = 0. \tag{5.20}$$

For any v in the space $\mathcal{V}^N$, (5.20) has a unique solution denoted by $z^L[v]$. Let ξ_k be solution of $\xi_0 = 0$ and

$$\xi_{k+1} = \xi_k + h_k f^k_y \xi_k - \theta h^2_k \delta^k_y; \quad k = 0, \ldots, N-1; \;\; \xi_0 = 0. \tag{5.21}$$

We have that

$$z_k = z^L_k + \xi_k; \;\; k = 0, \ldots, N, \quad \|\xi\|_\infty = O(\bar{h}). \tag{5.22}$$

We set

$$\begin{array}{l} \mathcal{G}^L_{i,j}(v) := (g'_i(y_{k[i,j+1]}) z^L_{k[i,j+1]}[v] - g'_i(y_{k[i,j]}) z^L_{k[i,j]}[v])/\Delta t_{i,j}, \\ j = 0, \ldots, M^\theta_i - 1; \quad i = 1, \ldots, r. \end{array} \tag{5.23}$$

The linear (homogeneous) equations corresponding to those of (QP_E) are therefore $\mathcal{G}^L(v) = 0$. Consider the following perturbation of the r.h.s. of these equations, where $\bar{b}$ is an arbitrary r.h.s.:

$$\mathcal{G}^L(v) = \bar{b}. \tag{5.24}$$

We use the following norm, for $s \in [1, \infty)$:

$$\|\bar{b}\|^s_s := \sum_{i=1}^{r} \sum_{j=0}^{M^\theta_i - 1} \Delta t_{i,j} |\bar{b}_{i,j}|^s. \tag{5.25}$$

These norms can be identified with the usual L^s norms on $[0, T]$ for piecewise constant functions and therefore we have the usual Cauchy-Schwarz and Hölder inequalities, in particular

$$\|\bar{b}\|_1 \leq \sqrt{r} T \|\bar{b}\|_\infty \tag{5.26}$$

PROPOSITION 5.2. *There exist constants C_s, $s \in [1, \infty]$, such that the linear equation $\mathcal{G}^L(v) = \bar{b}$ has, for small enough $\bar{h}$, a solution v verifying*

$$\|v\|_s \leq C_s \|\bar{b}\|_s, \quad \textit{for each } s \in [1, \infty]. \tag{5.27}$$

Before proving the Proposition 5.2 we introduce some notations. For $t \in [0, T]$, and $\varepsilon_0 > 0$, we define the set of ε_0 *active constraints* as

$$\mathcal{A}_{\varepsilon_0}(t) := \{1 \leq i \leq r; \; |t' - t| \leq \varepsilon_0, \text{ for some } t' \text{ such that } i \in \mathcal{A}(t')\}. \tag{5.28}$$

Since the control is continuous by **(A1)**, and the first order state constraint satisfies **(A2)**, we have that, for $\varepsilon_0 > 0$ small enough:

$$\left| \sum_{i \in \mathcal{A}_\varepsilon(t)} \lambda_i \nabla_u g^{(1)}_i(\bar{u}, \bar{y}) \right| \geq \tfrac{1}{2} \alpha |\lambda|, \quad \text{if } \lambda_i = 0 \text{ when } i \notin \mathcal{A}_{\varepsilon_0}(t), \text{ for all } t \in [0, T]. \tag{5.29}$$

For $i \in \{1, \dots, r\}$ we denote the set of ε_0 active constraints by $J^{i,\theta}_{\varepsilon_0}$, $i = 1, \dots, r$. This is a union of closed balls (in $[0, T]$) of radius ε_0. Since every connected component has length at least $2\varepsilon_0$, $J^{i,\theta}_{\varepsilon_0}$ is a finite union of closed intervals.

We next define the one-time-step analogue of $\mathcal{G}^L$:

$$G^L_{i,k}(v_k, z_k) := \frac{g_i'(y_{k+1}) - g_i'(y_k)}{h_k} z_k[v] + g_i'(y_{k+1})(f'(u_k, y_k)(v_k, z_k[v])). \tag{5.30}$$

Note that $G^L_{i,k}(v_k, z_k) = (g_i'(y_{k+1})z_{k+1}[v] - g_i'(y_k)z_k[v])/h_k$.

Proof. [Proof of Proposition 5.2] The idea is to compute, for each k, v_k as the minimum norm solution of the linear equations

$$G^L_{i,k}(v_k, z^L_k) = \tilde{b}_{i,k},\ i \in \mathcal{A}_\varepsilon(t_k), \tag{5.31}$$

where the variable size vector $\tilde{b}_{i,k}$, for i in $\mathcal{A}_\varepsilon(t_k)$, will be defined later, and then to set

$$\tilde{b}_{i,k} := G^L_{i,k}(v_k, z^K_k),\ \text{for } i \notin \mathcal{A}_\varepsilon(t_k). \tag{5.32}$$

Thanks to the expression of $G^L_{i,k}$ and (5.29) setting $z^L = z^L[v]$, we have that

$$|v_k| \le c_1 \left(|z^L_k| + \sum_{i \in \mathcal{A}_\varepsilon(t_k)} |\tilde{b}_{i,k}| \right). \tag{5.33}$$

Here the c_i are positive constants not depending on v or k. It follows with (5.30) that

$$|\tilde{b}_k| \le c_2 \left(|v_k| + |z^L_k[v]| \right) \le c_3 \left(|z^L_k| + \sum_{i \in \mathcal{A}_\varepsilon(t_k)} |\tilde{b}_{i,k}| \right). \tag{5.34}$$

So, by (5.20):

$$|z^L_{k+1}| \le (1 + c_4 h_k)|z^L_k| + c_5 h_k |v_k| \le (1 + c_6 h_k)|z^L_k| + c_7 h_k \sum_{i \in \mathcal{A}_\varepsilon(t_k)} |\tilde{b}_{i,k}|. \tag{5.35}$$

By the discrete Gronwall's Lemma it follows that

$$\|z^L\|_\infty \le c_8 \sum_{k=0}^{N-1} h_k \sum_{i \in \mathcal{A}_\varepsilon(t_k)} |\tilde{b}_{i,k}|, \tag{5.36}$$

and therefore, with and (5.33):

$$\|v\|^s_s \le c_9 \sum_{k=0}^{N-1} h_k \sum_{i \in \mathcal{A}_\varepsilon(t_k)} |\tilde{b}_{i,k}|^s, \tag{5.37}$$

and in the r.h.s. we recognize the L^s norm of the ε_0 active components of $\tilde{b}$. We next end the proof by fixing the $\tilde{b}_k$ in such a way that

$$\mathcal{G}^L_{i,j}(v) = \bar{b}_{i,j};\ \ j = 0, \dots, M^\theta_i - 1; \quad \|\tilde{b}\|_s = O(\|\bar{b}\|_s). \tag{5.38}$$

We will obtain the second relation by induction over k, i.e. we will prove that there exists $c > 0$ such that

$$\sum_{\ell \le k} |\tilde{b}_\ell|^s \le c \sum_{i;k[i,j]\le k} |\bar{b}_{i,k}|^s. \tag{5.39}$$

We distinguish two cases.
a) If $\{t_{k[i,j]}, \dots, t_{k[i,j+1]}\} \in J^{i,\theta}_{\varepsilon_0}$, then take

$$\tilde{b}_{i,k} = \bar{b}_{k[i,j+1]}, \quad k = k[i,j]+1, \dots, k[i,j+1]. \tag{5.40}$$

b) If $\{t_{k[i,j]}, t_{k[i,j+1]}\} \notin J^{i,\theta}_{\varepsilon_0}$, let k' be the smallest index in $k[i,j], \dots, k[i,j+1]$ such that $k \in J^{i,\theta}_{\varepsilon_0}$, whenever $k' \le k \le k[i,j+1]$; Then

$$t_{k'} + \tfrac{1}{2}\varepsilon_0 \le t_{k[i,j+1]}. \tag{5.41}$$

We choose

$$\tilde{b}_{i,k} = \begin{cases} \bar{b}_{i,j}, & k = k[i,j]+1, \dots, k', \\ \gamma, & k = k'+1, \dots, k[i,j+1], \end{cases} \tag{5.42}$$

for some γ such that

$$\sum_{k=k[i,j]+1}^{k'} h_k \tilde{b}_{i,k} + \gamma(t_{k[i,j+1]} - t_{k'}) = (t_{k[i,j+1]} - t_{k[i,j]})\bar{b}_{k[i,j+1]}, \tag{5.43}$$

so that the first relation in (5.38) holds. At the same time, since $\frac{1}{2}\varepsilon_0 \le t_{k[i,j+1]} - t_{k'}$, we have that

$$\begin{aligned} \gamma &\le \frac{2}{\varepsilon_0} \left| (t_{k[i,j+1]} - t_{k[i,j]})\bar{b}_{k[i,j+1]} - \sum_{k=k[i,j]+1}^{k'} h_k \tilde{b}_{i,k} \right| \\ &\le \frac{2}{\varepsilon_0} \left(T|\bar{b}_{k[i,j+1]}| + \sum_{k \le k'} h_k |\tilde{b}_{i,k}| \right), \end{aligned} \tag{5.44}$$

and we use the induction hypothesis (5.39) in order to estimate $\sum_{k\le k'} h_k|\tilde{b}_{i,k}|$. The conclusion follows. □

From the same argument of the previous result we can deduce also an estimate for the control of a feasible trajectory.

Note that

$$G_{i,k}(v_k, z_k) = G^L_{i,k}(v_k) + O(\bar{h}), \quad i = 1, ..., r. \tag{5.45}$$

Corollary 5.3. *Problem (QP_E) has a feasible point $\tilde{v}$, such that $\|\tilde{v}\|_\infty = O(\bar{h})$.*

Proof. In view of the Proposition above, it is enough to check that we can write the active constraints of (QP_E) in the form

$$\mathcal{G}^L_{i,k}(v) = \bar{b}_{i,k}; \quad \|\bar{b}\|_\infty = O(\bar{h}). \tag{5.46}$$

Remember that $z[v] = z^L[v] + \xi_k[v]$, with $\|\xi[v]\|_\infty = O(\|v\|_1)$, and therefore

$$\begin{aligned} \bar{b}_{i,j} = \quad & -\theta \left(\frac{g_i'(y_{k[i,j+1]}) - g_i'(y_{k[i,j]})}{\Delta t_{i,j}} + g_i'(y_{k[i,j+1]}) f_y'(u_{k[i,j]}, y_{k[i,j]}) \right) \xi_{k[i,j]} \\ & -\theta(h^2_{k+1}\delta^g_{k+1} - h^2_k \delta^g_k)/h_k, \end{aligned} \tag{5.47}$$

which is of the desired form. □

We recall a classical consequence of the coercivity of the cost function of an equality constrained quadratic problem over its feasible set. To keep the notation as simple as it possible, we formulate the problem in an abstract way. The result will be stated with the current notation as corollary (Corollary 5.6). Given two Hilbert spaces X and Y, identified with their dual, consider the optimization problem

$$\operatorname*{Min}_{x\in X}(c,x)_X+\frac{1}{2}(Hx,x)_X \quad \text{subject to} \quad Ax=b \ \text{ in } T, \tag{5.48}$$

where $(\cdot,\cdot)_X$ denotes the scalar product in X (and $\|\cdot\|:=(\cdot,\cdot)_X$) with a similar convention for Y, $H:X\to X$ is symmetric, $A\in L(X,Y)$, $c\in X$ and $b\in Y$. The Lagrangian of the problem is

$$(c,x)_X+\frac{1}{2}(Hx,x)_X+(\lambda,Ax-b)_Y. \tag{5.49}$$

The associated optimality conditions are

$$c+Hx+A^T\lambda=0; \quad Ax=b. \tag{5.50}$$

LEMMA 5.4. *Let $\alpha>0$ and $c_A>0$ be such that*

(i) *Coercivity: $\alpha\|x\|^2\le(Hx,x)_X$, for all $x\in\operatorname{Ker}A$,*

(ii) *Strong surjectivity: For any $b'\in Y$, there exists $x'\in X$ such that $Ax=b'$ and $\|x'\|\le c_A\|b'\|$.*

Then there exists $\kappa>0$, function of α and c_A, such that (5.48) *has a unique solution $\bar{x}$ and associated Lagrange multiplier λ such that*

$$\|\bar{x}\|+\|\lambda\|\le\kappa(\|b\|+\|c\|). \tag{5.51}$$

Proof. That (5.48) has a unique solution $\bar{x}$ is an easy consequence of the coercivity (which in the case of a quadratic cost implies the strong convexity over the feasible set since the latter is a vector subspace). The uniqueness of the Lagrange multiplier λ is consequence of the surjectivity of A, implied by the strong surjectivity.

The latter also implies the existence of x^0 such that $Ax^0=b$ and $\|x^0\|\le c_A\|b\|$. Then $\delta x:=\bar{x}-x^0$ is solution of

$$H\delta x+A^T\lambda=-c-Hx^0; \quad A\delta x=0. \tag{5.52}$$

Therefore, since $\delta x\in\operatorname{Ker}A$:

$$\alpha\|\delta x\|^2\le\delta x^THx=-(\delta x,c+Hx^0)_X+(A\delta x,\lambda),$$

so that $\|\delta x\|\le(\|c\|+\|Hx^0\|)/\alpha$. Since $\|x^0\|\le c_A\|b\|$, we deduce that

$$\|\bar{x}\|\le\|\delta x\|+\|x^0\|\le c_A\|b\|+\frac{1}{\alpha}\left(\|c\|+\|H\|\,c_A\|b\|\right). \tag{5.53}$$

By the surjectivity hypothesis

$$\|\lambda\|\le\frac{1}{c_A}\|A^T\lambda\|\le\frac{1}{c_A}(\|c\|+\|H\|\|x\|), \tag{5.54}$$

The conclusion follows. □

Given $v \in \mathcal{V}^N$, we denote by $\bar{v}$ the associated corresponding piecewise constant element defined by

$$\bar{v}(t) = v_k, \quad t \in (t_k, t_{k+1}), \quad \text{for all } k = 0 \text{ to } N-1. \tag{5.55}$$

Note that v and $\bar{v}$ have the same L^s norm, $s \in [1, \infty]$ norm. Setting $\bar{z} := z[\bar{v}]$ (solution of the linearized state equation (2.8) for the original (continuous time) problem, we easily check that

$$\|z^\theta - \bar{z}\|_\infty = \|v\|_\infty O(\varepsilon + \bar{h}). \tag{5.56}$$

We apply the previous result to (QP_E) using the following Lemma:

LEMMA 5.5. *We have that for any v in $\mathcal{U}^N$:*

$$\left|\Omega^\theta(v, z^L[v]) - \Omega(\bar{v})\right| = O(\varepsilon\|\bar{v}\|)^2.$$

Proof. Since by the definition (see (2.22)), $h_k^\theta \nu_k^\theta = \bar{\eta}_k^\theta - \bar{\eta}_{k+1}^\theta$ it follows that:

$$\begin{aligned}\Delta &:= \sum_{k=0}^{N} h_k \nu_k^\theta D^2 g_k^\theta(z_k)^2 = \sum_{k=0}^{N} (\bar{\eta}_k^\theta - \bar{\eta}_{k+1}^\theta) D^2 g_k^\theta(z_k)^2 \\ &= \sum_{k=1}^{N} \bar{\eta}_k^\theta (D^2 g_k^\theta(z_k)^2 - D^2 g_{k-1}^\theta(z_{k-1})^2) = \Delta_1 + \Delta_2,\end{aligned}$$

where by a Taylor expansion of $D^2 g^\theta$, for some $\hat{y}_k \in [y_{k-1}, y_k]$:

$$\Delta_1 := \sum_{k=1}^{N} \bar{\eta}_k (D^2 g_k^\theta(z_k)^2 - D^2 g_{k-1}^\theta(z_k)^2) = \sum_{k=1}^{N} h_k \bar{\eta}_k D^3 g^\theta(\hat{y}_k)(f_{k-1}^\theta, z_k, z_k),$$

using the identity $A(b,b) - A(a,a) = A(a+b, a-b)$ for any symmetric bilinear form A, and the linearized state equation

$$\begin{aligned}\Delta_2 &:= \sum_{k=1}^{N} \bar{\eta}_k^\theta (D^2 g_{k-1}^\theta(z_k)^2) - D^2 g_{k-1}^\theta(z_{k-1})^2) \\ &= \sum_{k=1}^{N} h_{k-1} \eta_k^\theta D^2 g_{k-1}^\theta(z_k + z_{k-1}, D f_{k-1}^\theta(v_{k-1}, z_{k-1})).\end{aligned}$$

We deduce that, for $\bar{h}$ small enough,

$$\Delta_1 = \int_0^T g^{(3)}(\bar{y}(t))(f[t], \bar{z}(t), \bar{z}(t))\bar{\eta}(t)\mathrm{d}t + O(\varepsilon\|\bar{v}\|)$$

and

$$\Delta_2 = 2\int_0^T g''(\bar{y}(t))(\bar{z}(t), f_y[t](\bar{z}(t), \bar{v}(t))\bar{\eta}(t)\mathrm{d}t + O(\varepsilon\|\bar{v}\|).$$

Using the identity (2.19) we can claim that $\left|\Omega^\theta(v,z) - \tilde{\Omega}(\bar{v})\right| = O(\varepsilon\|\bar{v}\|)^2$. We conclude with Lemma 2.4. ∎

COROLLARY 5.6. *Let assumptions* (**A0**)-(**A3**) *hold as well as either* (**A4**) *or* (**A5**). *We have that (QP_E) has a unique solution v^θ associated with a unique alternative multiplier $\delta\bar{\eta}^\theta$, and they satisfy*

$$\|v^\theta\|_2^2 + \|\delta\bar{\eta}^\theta\|_2^2 = O(\bar{h}). \tag{5.57}$$

Proof. The surjectivity of the constraint condition was proved in Proposition 5.2. If (**A4**) holds, then the coercivity property is an easy consequence of the stability after discretization of the Hessian of the Lagrangian (Lemma 5.5 above) and of the solution of the linearized state equation. If (**A5**) holds, the coercivity property is derived in Appendix A. □

5.3. Estimate of the directional derivative. We arrive, finally, to the main result of the Section. We recall that $\delta X^\theta := (v^\theta, z^\theta, q^\theta, \delta\eta^\theta)$ is the directional derivative (on the left) of $X^\theta := (u^\theta, y^\theta, p^\theta, \eta^\theta)$.

PROPOSITION 5.7. *We have, for a fixed $C > 0$*

$$\|v^\theta\|_\infty + \|z^\theta\|_\infty + \|q^\theta\|_\infty + \|\delta\eta^\theta\|_\infty \leq C\bar{h}.$$

Proof. (a) Applying Lemma 5.4 to (QP_E), where X and Y have norms defined by (5.27) (where $s = 2$) and (5.25), having in mind that, by the definition of the Lagrangian, we have identified the image space (of b^E) with its dual, we get that that

$$\|v^\theta\|_2 + \sum_{i,j} \Delta t_{i,j} |\delta\bar{\eta}^\theta_{i,j}|^2 \leq c_1\bar{h}. \tag{5.58}$$

Fix $\varepsilon_\eta > 0$, not depending on $\bar{h}$. Then

$$\text{If } \Delta t_{i,j} > \varepsilon_\eta, \text{ then } |\delta\bar{\eta}^\theta_{i,j}| \leq c_1\varepsilon_\eta^{-1/2}\bar{h}. \tag{5.59}$$

So, as far as $\delta\bar{\eta}^\theta$ is concerned, it remains to obtain a uniform estimate when $\Delta t_{i,j} \leq \varepsilon_\eta$. It easily follows from (5.58), the linearized state equation (5.2), and the linearized costate equations (5.14) that

$$\|z^\theta\|_\infty + \|q^\theta\|_\infty \leq c_2\bar{h}. \tag{5.60}$$

(b) By (5.6), $|\bar{b}_{i,j}| = O(\bar{h})$. Evaluating the contribution of the term containing z_k, observing $\sum_{k=k[i,j]+1}^{k[i,j+1]} h_k = \Delta t_{i,j}$ and then

$$\frac{1}{\Delta t_{i,j}} \sum_{k=k_i+1}^{k_{i+1}} h_k \nabla_u \hat{g}^{(1)}_{i,k} v_k = O(\bar{h}). \tag{5.61}$$

Eliminating v_k in (5.16), we get

$$\frac{1}{\Delta t_{i,j}} \sum_{k=k[i,j]+1}^{k[i,j+1]} h_k \nabla_u \hat{g}^{(1)}_{i,k} (H^k_{uu})^{-1} \sum_{i'=1}^{r} (\nabla_u \hat{g}^{(1)}_{i',k})^T \delta\bar{\eta}^\theta_{i',j[i',k]} = O(\bar{h}). \tag{5.62}$$

Multiplying by $\delta\bar{\eta}^\theta_{i,j[i,k]}$ on the left, summing over i and $j \in I_{i',k}$ and setting

$$w_k := \sum_{i=1}^{r} (\nabla_u \hat{g}^{(1)}_{i,k})^T \delta\bar{\eta}^\theta_{i,j[i,k]} \tag{5.63}$$

we get

$$\frac{1}{\Delta t_{i,j}} \sum_{i=1}^{r} \sum_{k=k[i,j]+1}^{k[i,j+1]} h_k w_k^T (H^k_{uu})^{-1} w_k = O(\bar{h}) |\delta\bar{\eta}^\theta|. \tag{5.64}$$

TABLE 6.1
Experimental error (DO NEW TESTS).

h	$\|y^h - \bar{y}\|_\infty$	$Ord(L^\infty)$	$\|u^h - \bar{u}\|_\infty$	$Ord(L^\infty)$
0.05	0.2909		1.1704	
0.025	0.16131	0.8506	0.6254	0.8776
0.0125	0.08313	0.9564	0.3304	0.9205
0.0063	0.04125	1.0109	0.1654	0.9982
0.0031	0.02059	1.0024	0.0829	0.9965
0.0016	0.01022	1.0105	0.0441	0.9106
0.0008	0.005187	0.9784	0.0222	0.9902
0.0004	0.002285	1.1827	0.0101	1.1362

Denote by $\hat{\eta}$ the vector of components $\delta\bar{\eta}^\theta_{i,j[i,k]}$, for $i = 1$ to r. For small enough ε depending on $\hat{\eta}$, by (**A2**)-(**A3**), we have that

$$|\hat{\eta}|^2 \leq \alpha_g^2 |w_k|^2 \leq \alpha^{-1}\alpha_g^2 w_k^T (H^k_{uu})^{-1} w_k \tag{5.65}$$

and therefore, since $\Delta t_{i,j} = \sum_{k=k[i,j]+1}^{k[i,j+1]} h_k$:

$$|\hat{\eta}|^2 \leq \frac{\alpha^{-1}\alpha_g^2}{\Delta t_{i,j}} \sum_{k=k[i,j]+1}^{k[i,j+1]} w_k^T (H^k_{uu})^{-1} w_k = O(\bar{h})|\hat{\eta}|. \tag{5.66}$$

Therefore, we get with (5.19) that

$$|\delta\bar{\eta}^\theta_{i,j[i,k]}| \leq O(\bar{h}). \tag{5.67}$$

The corresponding estimate for v^k and $\delta\eta^\theta$ follow from (5.16) and (5.19). □

6. Example. We present next an academic example which is a variant from the one in [16]. Let us consider the following optimal control problem, for some $\varepsilon > 0$: where $g := 9.8$:

$$\begin{aligned} &\text{Min } \int_0^1 \left(\tfrac{1}{2}u^2(t) + gy(t)\right)\, dt + (y(1)-1)^2/\bar{\varepsilon}, \\ &\text{s.t. } \dot{y}(t) = u(t), \quad y(0) = 1, \quad y(t) \geq 0, \end{aligned}$$

The solution of this problem can be seen as the minimum energy state of a system composed by a rope of uniformly distributed mass in a constant gravity field, with the presence of a lower constraint (for example a table). We can add a state variable say $\tilde{y}$, with zero initial condition and derivative equal to the integrand of the integral cost, and reformulate the cost as $\tilde{y}(1) + (y(1)-1)^2/\bar{\varepsilon}$, in order to comply with the format of the theoretical results. It is known that the costate associated with $\tilde{y}$ has value 1 (observe that this problem is qualified) and that the costate associated with y and the measure associated with the state constraints are invariant under this reformulation.

We solved the discrete solution using a shooting method; in Figure 6 and Table 6 are shown the results at various constant discrete steps h and for $\bar{\varepsilon} = 10^{-8}$. This test confirms the convergence results stated before.

Appendix A. Analysis of assumption (A5).

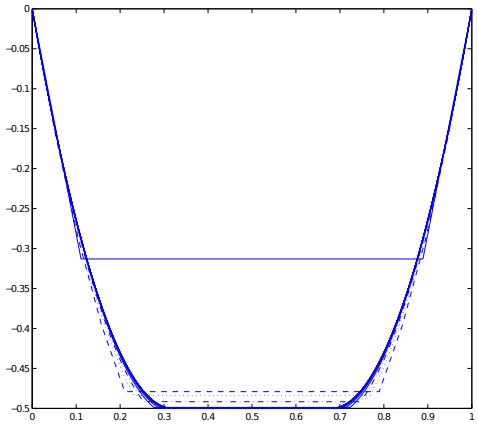

FIG. 6.1. *Test with various discrete steps.*

As shown before, a key point of the theory is the coercivity of the (QP). Under the assumptions (**A5**), we can obtain it directly showing the stability of the boundary arcs. This will be sufficient to guarantee the coercivity of the Hessian of the Lagrangian. A main point is contained in the following Lemma:

LEMMA A.1. *Let* (**A0**) − (**A3**) *and* (**A5**) *hold. Given a boundary arc* (t_a, t_b)*, let* k_a *and* k_b *be defined as*

$$k_a \text{ (resp. } k_b\text{): first index (resp. last index) for which } t_k > t_a \text{ (resp. } t_k < t_b\text{).} \quad \text{(A.1)}$$

Let $\varepsilon' > 0$*. Reducing* $\varepsilon > 0$ *small enough, we have that when* $\bar{h}$ *is small enough, the following holds:*

$$g(y_k^\theta) = 0, \text{ for all } 0 \le k \le N \text{ such that } t_{k_a} + \varepsilon' < t_k < t_{k_b} - \varepsilon'.$$

Proof. (a) By the definition $\|\eta^\theta - \bar{\eta}\|_\infty < \varepsilon$, and by (**A5**), $\bar{\eta}$ has a uniformly positive derivative over $(t_{k_a} + \varepsilon' < t_k < t_{k_b} - \varepsilon')$ minored by $c_1 > 0$. We have that for $k_a < k < k_b$:

$$\eta_k^\theta \ge \bar{\eta}(t_k) - \varepsilon \ge \bar{\eta}(t_{k_a}) + c_1(t_k - t_{k_a}) - \varepsilon \ge \eta_{k_a}^\theta + c_1(t_k - t_{k_a}) - 2\varepsilon, \quad \text{(A.2)}$$

that is,

$$c_1(t_k - t_{k_a}) - 2\varepsilon \le \eta_k^\theta - \eta_{k_a}^\theta. \quad \text{(A.3)}$$

Therefore, if $t_k - t_{k_a} > 2\varepsilon/c_1$, the above r.h.s. must be positive, proving that the constraint is active for some k such that $t_k \le t_{k_a} + 2\varepsilon/c_1$.
(b) If the conclusion does not hold, by step (a), it suffices to prove that $g(y_k^\theta)$ cannot have a negative local minimum for some $k_a < k < k_b$. We give a proof by contradiction. If this was the case, then $\Delta_g^k \ge 0$, defined in (4.6), and $\nu_k = 0$. Multiplying (4.3) by $\nabla_u g^{(1)}(u_k, y_k)^T (\mathcal{H}_k^\theta)^{-1}$ on the left and using Lemma 4.1(ii) and (4.8), we get

$$-\frac{\Delta_g^k}{h_k} + \nu_k (\nabla_u g^{(1)}(u_k, y_k))^T \mathcal{H}_k^{-1} \nabla_u g^{(1)}(u_k, y_k) = \hat{\Xi}_k^\theta, \quad \text{(A.4)}$$

where $\hat{\Xi}^\theta_k$ is such that $\|\hat{\Xi}^\theta - \hat{\Xi}^1\|_\infty = O(\varepsilon)$. We have that the l.h.s. of (A.4) is greater than a positive constant independent on $\bar{h}$, since, for $\theta = 1$, for some $K > 0$, $\Delta^k_g = 0$, $\nu_k > K$ and $|\nabla g^{(1)}_u| > K$, i.e.,

$$-\frac{\Delta^k_g}{h_k} + \nu_k(\nabla_u g^{(1)}(u_k, y_k))^T \mathcal{H}_k^{-1} \nabla_u g^{(1)}(u_k, y_k) \geq C. \tag{A.5}$$

This relation is still valid for ε and $\bar{h}$ small enough, for all $\theta \in [\theta_m, 1]$, in view of the continuity of the r.h.s. of (A.4). However, if a negative minimum of the state constraint is attained at index k, then $\nu_k = 0$ and $\Delta^k_g \geq 0$, contradicting (A.5). □

Here we prove the coercivity of Ω^θ over the feasible domain of (QP). Note that such a property is naturally preserved passing to the alternative formulation $\tilde{\Omega}^\theta$ as shown, for the continuous case in Section 2.

LEMMA A.2. *Let* **(A5)** *hold. Then $v \mapsto \Omega^\theta(v, Z^L[v]$ is uniformly (over $\bar{h}$ small enough) coercive over the feasible domain of (QP).*

Proof. (a) We first examine the continuous problem and prove that Ω is, for $\varepsilon > 0$ small enough, coercive over the following enlargement of the critical cone:

$$C_\varepsilon := \{v \in \mathcal{V} \mid g_k = 0, \text{ for all } 0 \leq k \leq N \text{ such that } t_{k_a} + \varepsilon < t_k < t_{k_b} - \varepsilon\}. \tag{A.6}$$

Indeed, otherwise we would have a sequence $\varepsilon_q \downarrow 0$ and v^q in C_{ε_q} such that $\Omega(v^q) \leq o(1)$. Extracting if necessary a subsequence, assume that v^q weakly converges to $\bar{v}$ in $\mathcal{V}$. Thanks to the Legendre condition we have that Ω is a Legendre form and therefore

$$\Omega(\bar{v}) \leq \liminf_{q \to 0} \Omega(v^q) \leq 0. \tag{A.7}$$

At the same time; by standard compactness arguments $g'(\bar{y})\bar{z} = 0$ when the constraint is active (where $\bar{z}$ is the linearized state associated with $\bar{v}$), and so, $\bar{v}$ is a critical direction. So, $\Omega(\bar{v}) \leq 0$ implies that $\bar{v} = 0$. But then $\Omega(\bar{v}) = \lim_q \Omega(v^q)$, so that v^q (of unit norm) strongly converges to $\bar{v}$, which gives the desired contradiction.
(b) Now let v belong to the feasible domain of (QP). By Lemma A.1, we know that v belongs to the set

$$\{v \in \mathcal{V}^N;\ |g'(y^\theta_k)z_k| \leq \varepsilon,\ 0 \leq k \leq N \text{ such that } t_{k_a} + \varepsilon < t_k < t_{k_b} - \varepsilon.\}. \tag{A.8}$$

Let $\bar{v}$ be the associated element of $\mathcal{V}$ and $\bar{z}$ the corresponding linearized state. Given $\varepsilon > 0$, it is easily checked that $\bar{v} \in C_\varepsilon$ when $\bar{h}$ is small enough, and so, by step (a), $\Omega(\bar{v}) \geq \frac{1}{2}\alpha\|\bar{v}\|^2$. We conclude with Lemma 5.5. □

REFERENCES

[1] J. F. BONNANS, *Lipschitz solutions of optimal control problems with state constraints of arbitrary order*, Ann. Acad. Rom. Sci. Ser. Math. Appl., 2-1 (2010), pp. 78–98.

[2] J. F. BONNANS AND A. HERMANT, *Well-posedness of the shooting algorithm for state constrained optimal control problems with a single constraint and control*, SIAM J. Control Optim., 46 (2007), pp. 1398–1430.

[3] J. F. BONNANS AND A. HERMANT, *Stability and sensitivity analysis for optimal control problems with a first-order state constraint and application to continuation methods*, ESAIM Control Optim. Calc. Var., 14 (2008), pp. 825–863.

[4] J. F. BONNANS AND A. HERMANT, *Revisiting the analysis of optimal control problems with several state constraints*, Control Cybernet., 38 (2009), pp. 1021–1052.

[5] J. F. Bonnans and A. Hermant, *Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 26 (2009), pp. 561–598.

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

[7] A. E. Bryson and Y. -C. Ho, *Applied optimal control*, Hemisphere Publishing, New-York, 1975.

[8] A. E. Bryson, W. F. Denham, and S. E. Dreyfus, *Optimal programming problems with inequality constraints*, AIAA J., 1.11 (1963), pp. 2544–2550.

[9] B. M. Budak, E. M. Berkovich, and E. N. Solov′eva, *Difference approximations in optimal control problems*, SIAM J. Control, 7 (1969), pp. 18–31.

[10] J. Cullum, *Finite-dimensional approximations of state-constrained continuous optimal control problems*, SIAM J. Control, 10 (1972), pp. 649–670.

[11] A. L. Dontchev, *Discrete approximations in optimal control*, In *Nonsmooth analysis and geometric methods in deterministic optimal control (Minneapolis, MN, 1993)*, volume 78 of *IMA Vol. Math. Appl.*, Springer, New York, pp. 59–80, 1996.

[12] A. L. Dontchev and W. W. Hager, *Lipschitzian stability for state constrained nonlinear optimal control*, SIAM J. Control Optim., 36 (1998), pp. 698–718.

[13] A. L. Dontchev and W. W. Hager, *The Euler approximation in state constrained optimal control*, Math. Comp., 70 (2001), pp. 173–203.

[14] W. W. Hager, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim., 17 (1979), pp. 321–338.

[15] W. W. Hager, *Runge-Kutta methods in optimal control and the transformed adjoint system*, Numer. Math., 87 (2000), pp. 247–282.

[16] W. W. Hager and G. D. Ianculescu, *Dual approximations in optimal control*, SIAM J. Control Optim., 22 (1984), pp. 423–465.

[17] D. H. Jacobson, M. M. Lele, and J. L. Speyer, *New necessary conditions of optimality for control problems with state-variable inequality constraints*, J. Math. Anal. Appl., 35 (1971), pp. 255–284.

[18] K. Jittorntrum, *Solution point differentiability without strict complementarity in nonlinear programming*, Math. Programming Stud., 21 (1984), pp. 127–138.

[19] K. Malanowski and H. Maurer, *Sensitivity analysis for state constrained optimal control problems*, Discrete Contin. Dynam. Systems, 4 (1998), pp. 241–272.

[20] H. Maurer, *On the minimum principle for optimal control problems with state constraints*, Schriftenreihe des Rechenzentrum 41, Universität Münster, 1979.

[21] H. Maurer, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Math. Program. Stud., 14 (1981), pp. 163–177.

[22] H. Maurer and J. Zowe, *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Program. 16 (1979), pp. 98-110.

[23] B. Sh. Mordukhovich, *On difference approximations of optimal control systems*, J. Appl. Math. Mech., 42 (1978), pp. 452–461.

[24] S. M. Robinson, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math. 30.4 (1976), pp. 597–607.

[25] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43-62.

[26] A. Schwartz and E. Polak, *Consistent approximations for optimal control problems based on Runge-Kutta integration*, SIAM J. Control Optim., 34 (1996), pp. 1235–1269.

[27] R. Vinter, *Optimal control. systems and control: foundations and applications*, Birkäuser, Boston, 2000.

[28] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.