Second order analysis of state-constrained control-affine problems

Maria Soledad Aronna, J. Frederic Bonnans, Bean San Goh

Noname manuscript No.
(will be inserted by the editor)

Second order analysis of state-constrained control-affine problems

M.S. Aronna and J.F. Bonnans and B.S. Goh

the date of receipt and acceptance should be inserted later

Abstract In this article we establish new second order necessary and sufficient optimality conditions for a class of control-affine problems with a scalar control and a scalar state constraint. These optimality conditions extend to the constrained state framework the Goh transform, which is the classical tool for obtaining an extension of the Legendre condition.

1 Introduction

Control-affine problems have been intensively studied since the 1960s and there is a wide literature on this subject. In what respect to second order conditions, the main feature of these systems that are affine in the control variable is that the second derivative of the pre-Hamiltonian function with respect to the control vanishes, and hence the classical *Legendre-Clebsch conditions* hold trivially and do not provide any useful information. Second and higher order necessary conditions for problems affine in the control, without control nor state constraints were first established in [26, 18, 17, 23, 3]. The case with control constraints and purely bang-bang solutions was investigated by [33, 4, 36, 31, 15, 35] among many others, while the class of bang-singular solutions was analyzed in e.g. [37, 6, 16].

This article is devoted to the study of Mayer-type optimal control problems governed by the dynamics

$$\dot{x}_t = f_0(x_t) + u_t f_1(x_t), \quad \text{for a.a. } t \in [0, T],$$

subject to endpoint constraints

$$\Phi(x_0, x_T) \in K_\Phi,$$

control constraints

$$u_{\min} \le u_t \le u_{\max},$$

EMAp/FGV, 22250-900 Rio de Janeiro, Brazil E-mail: soledad.aronna@fgv.br and INRIA-Saclay and Centre de Mathématiques Appliquées, Ecole Polytechnique, 91128 Palaiseau, France E-mail: Frederic.Bonnans@inria.fr and Curtin University, Sarawak Malaysia E-mail: goh2optimum@gmail.com

and a scalar state constraint of the form

$$g(x_t) \leq 0.$$

For this class of problems, we show necessary optimality conditions involving the regularity of the control and the state constraint multiplier at the junction points. Some of these necessary conditions which hold at the junction points were proved in [29]. Moreover, we provide second order necessary and sufficient optimality conditions in integral form obtained through the *Goh transformation* [18].

This investigation is strongly motivated by applications since it allows to deal with both control and state constraints, which appear naturally in realistic models. Many practical examples can be found in the existing literature, a non exhaustive list includes the prey-predator model [19], the Goddard problem in presence of a dynamic pressure limit [45,20], an optimal production and maintenance system studied in [30], and a recent optimization problem of running strategies [2]. We refer also to [13], [11], [44] and references therein.

About second order analysis in the state constrained case, we quote the early work by Russak [41,42], and more recently, Malanowski and Maurer [28], Bonnans and Hermant [9] provided second order necessary and sufficient optimality conditions and related the sufficient conditions with the convergence of the shooting algorithm in the case where the strengthened Legendre-Clebsch condition holds. Second order necessary conditions for the general nonlinear case with phase constraints were also proved in [7,24]. For control-affine problems with bounded scalar control variable, a scalar state constraint (as is the case in this article) and solutions (possibly) containing singular, bang-bang and constrained arcs, Maurer [29] proved necessary conditions (similar to those developed by McDanell and Powers [32]) that hold at the junction points of optimal solutions. In Maurer et al. [30] they extended to the state-constrained framework, a second order sufficient test for optimality given in Agrachev et al. [4] and Maurer-Osmolovskii [34] for optimal bang-bang solutions. Also in [44], the author provided a synthesis-like method to prove optimality for a class of control-affine problems with scalar control, vector state constraint and bang-constrained solutions. Some remarks about how the method in [44] could be extended to general bang-singular-constrained solutions are given in [43], but no proof of the validity of this extension is provided.

There is a literature dealing with the important case when the standard statement of Pontryagin's principle is degenerate in the sense that the costate is equal to zero, so that the principle is trivially verified and no information of the optimal solution is provided. In some cases, a nontrivial version of Pontryagin's principle can be obtained, see in particular Arutyunov [8], Rampazzo and Vinter [38]. Also second order conditions for the general nonlinear case were stated in strongly non-degenerate form, see e.g. [24,8].

Up to our knowledge, there is no result in the existing literature about second order necessary and sufficient conditions in integral (quadratic) form for control-affine problems with state constraints and solutions containing singular arcs.

The paper is organized as follows. In Section 2 we give the basic definitions and show necessary optimality conditions concerning the regularity of the optimal control and associated multipliers. Section 3 is devoted to second order necessary optimality conditions in integral form and to the Goh transformation, while second order sufficient conditions are provided in Section 4. The Appendix is consecrated to the presentation of abstract results on second order necessary conditions.

Notations. Let $\mathbb{R}^k$ denote the $k-$dimensional real space, i.e. the space of column real vectors of dimension k, and by $\mathbb{R}^{k*}$ its corresponding dual space, which consists of $k-$dimensional row real vectors. With $\mathbb{R}^k_+$ and $\mathbb{R}^k_-$ we refer to the subsets of $\mathbb{R}^k$ consisting of vectors with nonnegative, respectively nonpositive, components. We write h_t for the value of function h at time t if h is a function that depends only on t, and by $h_{i,t}$ the ith component of h evaluated at t. Let $h(t+)$ and $h(t-)$ be, respectively, the right and left limits of h at t, if they exist. Partial derivatives of a function h of (t,x) are referred as $D_t h$ or $\dot{h}$ for the derivative in time, and $D_x h$, h_x or h' for the differentiations with respect to space variables. The same convention is extended to higher order derivatives. By $L^p(0,T)^k$ we mean the Lebesgue space with domain equal to the interval $[0,T] \subset \mathbb{R}$ and with values in $\mathbb{R}^k$. The notations $W^{q,s}(0,T)^k$ and $H^1(0,T)^k$ refer to the Sobolev spaces (see Adams [1] for further details on Sobolev spaces). We let $BV(0,T)$ be the set of functions with bounded total variation. In general, when there is no place for confusion, we omit the argument $(0,T)$ when referring to a space of functions. For instance, we write L^∞ for $L^\infty(0,T)$, or $(W^{1,\infty})^{k*}$ for the space of $W^{1,\infty}-$functions from $[0,T]$ to $\mathbb{R}^{k*}$. We say that a function $h : \mathbb{R}^k \to \mathbb{R}^d$ is of class C^ℓ if it is $\ell-$times continuously differentiable in its domain.

2 Framework

2.1 The problem

Consider the control and state spaces L^∞ and $(W^{1,\infty})^n$, respectively. We say that a control-state pair $(u,x) \in L^\infty \times (W^{1,\infty})^n$ is a *trajectory* if it satisfies both the *state equation*

$$\dot{x}_t = f_0(x_t) + u_t f_1(x_t), \quad \text{for a.a. } t \in [0,T], \tag{2.1}$$

and the finitely many *endpoint constraints* of equality and inequality type

$$\Phi(x_0, x_T) \in K_\Phi. \tag{2.2}$$

Here f_0 and f_1 are twice continuously differentiable and Lipschitz continuous vector fields over $\mathbb{R}^n$, Φ is of class C^2 from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^{n_1+n_2}$, and

$$K_\Phi := \{0\}_{\mathbb{R}^{n_1}} \times \mathbb{R}^{n_2}_-, \tag{2.3}$$

where $\{0\}_{\mathbb{R}^{n_1}}$ is the singleton consisting of the zero vector of $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}_- := \{y \in \mathbb{R}^{n_2} : y_i \leq 0, \text{ for all } i = 1,\dots,n_2\}$. Given $(u, x_0) \in L^\infty \times \mathbb{R}^n$, (2.1) has a unique solution. In addition, we consider the *cost functional* $\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, then *bound control constraints*

$$u_{\min} \leq u_t \leq u_{\max}, \quad \text{for a.a. } t \in [0,T], \tag{2.4}$$

and a *scalar state constraint*

$$g(x_t) \leq 0, \quad \text{for all } t \in [0,T], \tag{2.5}$$

with ϕ and $g : \mathbb{R}^n \to \mathbb{R}$ of class C^2. Here we allow $u_{\min}$ and $u_{\max}$ to be either finite real numbers, or to take the values $-\infty$ or $+\infty$, respectively, in the sense that problems with control constraints of the form $u_t \leq u_{\max}$ or $u_{\min} \leq u_t$ are also considered in our investigation, as well as problems in the absence of control constraints. We say that the trajectory (u,x) is *feasible* if it satisfies (2.4)-(2.5). Let us then consider the optimal control problem in the Mayer form

$$\min \phi(x_0, x_T); \qquad \text{subject to (2.1)-(2.5).} \tag{P}$$

2.2 Regular extremals

We set $f(u,x) := f_0(x) + uf_1(x)$, and define the *pre-Hamiltonian function* and the *endpoint Lagrangian,* respectively, by

$$\begin{cases} H(u,x,p) := pf(u,x) = p\big(f_0(x) + uf_1(x)\big), \\ \ell^{\beta,\Psi}(x_0,x_T) := \beta\phi(x_0,x_T) + \Psi\Phi(x_0,x_T), \end{cases} \tag{2.6}$$

where $p \in \mathbb{R}^{n*}$, $\beta \in \mathbb{R}$ and $\Psi \in \mathbb{R}^{(n_1+n_2)*}$.

Any function $\mu \in BV(0,T)$ (shortly, BV) has left limit on $(0,T]$ and right limits on $[0,T)$ and, therefore, the values μ_{0+} and μ_T are well-defined. Moreover, μ has a distributional derivative that belongs to the space $\mathcal{M}(0,T)$ (shortly, $\mathcal{M}$) of finite Radon measures. Conversely, any measure $\mathrm{d}\mu \in \mathcal{M}$ can be identified with the derivative of a function μ of bounded variation such that $\mu_T \in BV_0$, i.e., μ belongs to the *space of bounded variation functions that vanish at time T*.

Let (u,x) be a feasible trajectory. We say that $\Psi \in \mathbb{R}^{(n_1+n_2)*}$ is *complementary to the endpoint constraint* if

$$\Psi \in N_\Phi\big(\Phi(x_0,x_T)\big), \tag{2.7}$$

where $N_\Phi\big(\Phi(x_0,x_T)\big)$ denotes the *normal cone to* K_Φ at the point $\Phi(x_0,x_T)$, i.e.

$$N_\Phi\big(\Phi(x_0,x_T)\big) := \{\Psi \in \mathbb{R}^{(n_1+n_2)*} : \Psi_i \geq 0,\ \Psi_i\Phi_i(x_0,x_T) = 0,\ i = n_1{+}1,\ldots,n_1{+}n_2\}. \tag{2.8}$$

A bounded variation function μ is *complementary to the state constraint* if and only if

$$\mathrm{d}\mu \geq 0, \quad \text{and} \quad \int_0^T g(x_t)\mathrm{d}\mu_t = 0. \tag{2.9}$$

For $\beta \in \mathbb{R}$, $\Psi \in \mathbb{R}^{(n_1+n_2)*}$ and $\mu \in BV_0$, the *costate equation* associated with $(\beta,\Psi,\mathrm{d}\mu)$ is given by

$$-\,\mathrm{d}p_t = p_t f_x(u_t,x_t)\mathrm{d}t + g'(x_t)\,\mathrm{d}\mu_t, \quad \text{for a.a. } t \in [0,T], \tag{2.10}$$

with endpoint conditions

$$(-p_0, p_T) = D\ell^{\beta,\Psi}(x_0,x_T). \tag{2.11}$$

Given $(\beta,\Psi,\mathrm{d}\mu) \in \mathbb{R} \times \mathbb{R}^{(n_1+n_2)*} \times \mathcal{M}$, the boundary value problem (2.10)-(2.11) has at most one solution. In addition, the *condition of minimization* of the pre-Hamiltonian H implied by the Pontryagin's Maximum Principle can be expressed as follows, for a.a. $t \in [0,T]$,

$$\begin{cases} u_t = u_{\min}, \quad \text{if } u_{\min} > -\infty \text{ and } p_t f_1(x_t) > 0, \\ u_t = u_{\max}, \quad \text{if } u_{\max} < +\infty \text{ and } p_t f_1(x_t) < 0, \\ p_t f_1(x_t) = 0, \text{ if } u_{\min} < u_t < u_{\max}. \end{cases} \tag{2.12}$$

Denote the quadruple of dual variables by $\lambda := (\beta, \Psi, p, \mathrm{d}\mu)$, element of the space

$$E^{\Lambda} := \mathbb{R}_+ \times \mathbb{R}^{(n_1+n_2)*} \times BV^{n*} \times \mathcal{M}. \tag{2.13}$$

The *Lagrangian* of the problem is

$$\mathcal{L}(u, x, \lambda) := \ell^{\beta,\Psi}(x_0, x_T) + \int_0^T p_t\big(f(u_t, x_t) - \dot{x}_t\big)\mathrm{d}t + \int_0^T g(x_t)\mathrm{d}\mu_t. \tag{2.14}$$

Note that the costate equation (2.10) expresses the stationarity of $\mathcal{L}$ with respect to the state. For a feasible trajectory $(u, x) \in L^\infty \times (W^{1,\infty})^n$, define the *set of Lagrange multipliers* as

$$\Lambda(u, x) := \left\{ \begin{array}{l} \lambda = (\beta, \Psi, p, \mathrm{d}\mu) \in E^{\Lambda} : \ (\beta, \Psi, \mathrm{d}\mu) \neq 0; \\ \text{(2.1)-(2.5) and (2.7)-(2.12) hold} \end{array} \right\}. \tag{2.15}$$

Remark 1 We could as well call $\Lambda(u, x)$ the set of Pontryagin multipliers since they satisfy condition of minimization of the Hamiltonian. But since the dynamics is an affine function of the control variable, the condition of minimization of the Hamiltonian is equivalent to the stationarity condition (2.12) which gives the classical definition of Lagrange multipliers in this setting.

In addition we set, for $\beta \in \mathbb{R}_+$,

$$\Lambda_\beta(u, x) := \{(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda(u, x)\}. \tag{2.16}$$

Since $\Lambda(u, x)$ is a cone we have that

$$\Lambda(u, x) = \begin{cases} \mathbb{R}_+ \Lambda_1(u, x), & \text{if } \Lambda_0(u, x) = \emptyset, \\ \Lambda_0(u, x) + \mathbb{R}_+ \Lambda_1(u, x), & \text{otherwise.} \end{cases} \tag{2.17}$$

Consider a nominal feasible trajectory $(\hat{u}, \hat{x}) \in L^\infty \times (W^{1,\infty})^n$. Set

$$A_t := D_x f(\hat{u}_t, \hat{x}_t), \quad \text{for } t \in [0, T]. \tag{2.18}$$

From now on, when the argument of a function is omitted, we mean that it is evaluated at the nominal pair $(\hat{u}, \hat{x})$. In particular, we write Λ to refer to $\Lambda(\hat{u}, \hat{x})$.

For $(v, z^0) \in L^\infty \times \mathbb{R}^n$, let $z[v, z^0] \in (W^{1,\infty})^n$ denote the solution of the *linearized state equation*

$$\dot{z}_t = A_t z_t + v_t f_{1,t}, \quad \text{for a.a. } t \in [0, T], \tag{2.19}$$

with initial condition

$$z_0 = z^0. \tag{2.20}$$

Let $\bar{\Phi}_E$ denote the function from $L^\infty \times \mathbb{R}^n$ to $\mathbb{R}^{n_1}$ that, to each $(u, x_0) \in L^\infty \times \mathbb{R}^n$, assigns the value $\big(\Phi_1(x_0, x_T), \ldots, \Phi_{n_1}(x_0, x_T)\big)$, where x is the solution of (2.1) associated with (u, x_0). For some results obtained in this article, we shall consider the following *qualification condition,* which corresponds to *Robinson condition [39]* (see also (A.5) and Remark 10 in the Appendix):

$$\begin{cases} \text{(i)} \ D\bar{\Phi}_E(\hat{u}, \hat{x}_0) \text{ is onto from } L^\infty \times \mathbb{R}^n \text{ to } \mathbb{R}^{n_1}, \\ \text{(ii) there exists } (\bar{v}, \bar{z}^0) \in L^\infty \times \mathbb{R}^n \text{ in the kernel of } D\bar{\Phi}_E(\hat{u}, \hat{x}_0), \\ \qquad \text{such that, for some } \varepsilon > 0, \text{ setting } \bar{z} = z[\bar{v}, \bar{z}^0], \text{ one has:} \\ \qquad \hat{u}_t + \bar{v}_t \in [u_{\min} + \varepsilon, u_{\max} - \varepsilon], \text{ for a.a. } t \in [0, T], \\ \qquad g(\hat{x}_t) + g'(\hat{x}_t)\bar{z}_t < 0, \text{ for all } t \in [0, T], \\ \qquad \Phi_i'(\hat{x}_0, \hat{x}_T)(\bar{z}^0, \bar{z}_T) < 0, \text{ if } \Phi_i(\hat{x}_0, \hat{x}_T) = 0, \text{ for } n_1 + 1 \leq i \leq n_1 + n_2. \end{cases} \tag{2.21}$$

Here the notation $\hat{u}_t + \bar{v}_t \in [u_{\min} + \varepsilon, u_{\max} - \varepsilon]$ means that, a.e., $\hat{u}_t + \bar{v}_t \in [u_{\min} + \varepsilon, +\infty)$ if $u_{\max} = +\infty$ and $u_{\min}$ is finite; $\hat{u}_t + \bar{v}_t \in (-\infty, u_{\max} - \varepsilon)$ if $u_{\min} = -\infty$ and $u_{\max}$ is finite; and $\hat{u}_t + \bar{v}_t \in \mathbb{R}$ if neither $u_{\min}$ nor $u_{\max}$ is finite.

Definition 1 A *weak minimum* for (P) is a feasible trajectory (u, x) such that $\phi(x_0, x_T) \leq \phi(\tilde{x}_0, \tilde{x}_T)$ for any feasible $(\tilde{u}, \tilde{x})$ for which $\|(\tilde{u}, \tilde{x}) - (u, x)\|_\infty$ is small enough.

Theorem 1 *Assume that* $(\hat{u}, \hat{x})$ *is a weak minimum for* (P)*. Then* (i) *the set* Λ *of Lagrange multipliers is nonempty,* (ii) *if the qualification condition* (2.21) *holds, then* Λ_0 *is empty, and* Λ_1 *is bounded and weakly* compact.*

Proof If (2.21)(i) holds, we deduce the result from proposition (8). Otherwise, there exists some nonzero Ψ_E in the image of normal space to $D\bar{\Phi}_E(\hat{u}, \hat{x}_0)$. Setting $\Psi_i = 0$ for all $n_1 < i \leq n_1 + n_2$, and $p = 0$, we obtain a (singular) Lagrange multiplier. This ends the proof.

2.3 Jump conditions

Given a function of time $h : [0, T] \to \mathbb{R}^d$ for $d \in \mathbb{N}$, we define its *jump at time* $t \in [0, T]$ by

$$[h_t] := h(t+) - h(t-), \tag{2.22}$$

when the left and right limits, $h(t-)$ and $h(t+)$, respectively, exist and are finite. Here we adopt the convention $h(0-) := h_0$ and $h(T+) := h_T$. For any function of bounded variation the associated jump function is well-defined. For a function defined almost everywhere with respect to the Lebesgue measure, we will accord that its jump at time t is the jump at t of a representative of this function for which the left and right limit exist, provided that such a representative exists. By the costate equation (2.10), we have that, for any $(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$,

$$\begin{cases} \text{(i)} \;\; [p_t] = -[\mu_t] g'(\hat{x}_t), \\ \text{(ii)} \;\; \big[H_u(t)\big] = [p_t] f_1(\hat{x}_t) = -[\mu_t] g'(\hat{x}_t) f_1(\hat{x}_t). \end{cases} \tag{2.23}$$

In addition, if $[\hat{u}_t]$ makes sense, then the jump in the derivative of the state constraint exists and satisfies

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} g(\hat{x}_t)\right] = [\hat{u}_t] g'(\hat{x}_t) f_1(\hat{x}_t). \tag{2.24}$$

Lemma 1 *Let* $(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$*. Then, if* $t \in [0, T]$ *is such that* $[H_u(t)] = 0$ *and* $g'(\hat{x}_t) f_1(\hat{x}_t) \neq 0$*, then* μ *is continuous at* t*.*

Proof It follows from item (ii) in equation (2.23).

Lemma 2 *Let* $t \in (0, T)$ *be such that* $[\hat{u}_t]$ *makes sense. Then the following conditions hold*

$$\begin{cases} \text{(i)} \quad [\hat{u}_t][H_u(t)] = 0, \\ \text{(ii)} \quad [\mu_t] \left[\frac{\mathrm{d}}{\mathrm{d}t} g(\hat{x}_t)\right] = 0. \end{cases} \tag{2.25}$$

Proof Note that

$$[\mu_t]\left[\frac{\mathrm{d}}{\mathrm{d}t}g(\hat{x}_t)\right] = [\mu_t]g'(\hat{x}_t)f_1(\hat{x}_t)[\hat{u}_t] = -[H_u(t)][\hat{u}_t], \tag{2.26}$$

where the last equality follows from (2.23). This implies that (i) is equivalent to (ii), and so we only need to prove (i), which holds trivially when $[\mu_t] = 0$. Hence, let us assume that $[\mu_t] \neq 0$. Then $g(\hat{x}_t) = 0$ in view of (2.9) and, necessarily, $t \mapsto g(\hat{x}_t)$ attains its maximum at t, so that $\left[\frac{\mathrm{d}}{\mathrm{d}t}g(\hat{x}_t)\right] \leq 0$. Since $\mathrm{d}\mu \geq 0$, it follows that $[\mu_t] \geq 0$, and therefore by (2.26), $[H_u(t)][\hat{u}_t] \geq 0$. However, the converse inequality holds in view of (2.12). The conclusion follows.

We say that the *state constraint is of first order* if

$$g'(\hat{x}_t)f_1(\hat{x}_t) \neq 0, \quad \text{when } g(\hat{x}_t) = 0. \tag{2.27}$$

Corollary 1 *Assume that the state constraint is of first order. Then if the control has a jump at time $\tau \in (0,T)$ for which $g(\hat{x}_\tau) = 0$, then μ is continuous at τ for any associated multiplier $(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$.*

Proof From the identity (2.25) in Lemma 2, if $[\hat{u}_\tau] \neq 0$, then $[H_u(t)] = 0$. The latter implies that $[\mu_\tau] = 0$, in view of the second equality in (2.26) and since the state constraint is of first order, i.e. (2.27) holds.

We refer to Remark 3 regarding the link between previous Corollary 1 and Maurer [29].

Remark 2 Let us illustrate by means of this example that the associated μ might have jumps at the initial and or at final times. In fact, consider the problem

$$\begin{aligned}
&\min x_{1,T} + x_{2,T};\\
&\qquad \dot{x}_{1,t} = u_t \in [-1,1]; \quad \dot{x}_{2,t} = x_{1,t};\\
&\qquad g(x_{1,t}) = -x_{1,t} \leq 0; \quad x_{1,0} = 1,\ x_{2,0} = 0.
\end{aligned}$$

with $T = 2$, and note that

$$\hat{u}_t := \begin{cases} -1 & \text{on } [0,1],\\ 0 & \text{on } (1,2] \end{cases}$$

is optimal for it. Over $[0,1)$ the state constraint is not active, and on $[1,2]$ it is. The Hamiltonian here is $H = p_1 u + p_2 x_1$, the costate being $p = (p_1, p_2)$. We obtain $\dot{p}_2 = 0$, so that $p_{2,t} \equiv p_{2,T} = 1$, and $-\mathrm{d}p_1 = \mathrm{d}t - \mathrm{d}\mu$ with boundary condition $p_1(1) = 1$. Furthermore, on the constraint arc $(1,2)$ we necessarily have $H_u = 0$, leading to $p_1 = 0$. We conclude that $[p_T] = 1$ and therefore, due to (2.23), $[\mu_T] = 1$, this is, μ jumps at the final time.

2.4 Arcs and junction points

The *contact* set associated with the state constraint is defined as

$$C := \{t \in [0,T] :\ g(\hat{x}_t) = 0\}. \tag{2.28}$$

For $0 \leq a < b \leq T$, we say that (a,b) is a *(maximal) active arc* for the state constraint, or a *C arc,* if C contains (a,b), but no open interval in which (a,b) is strictly contained. Note that, since $t \mapsto g(\hat{x}_t)$ is continuous, the set C consists of a countable union of arcs, which can be ordered by size. We say that $\tau \in (0,T)$ is a *junction point of the state constraint* if it is the extreme point of a C arc.

We give similar definitions for the control constraint, paying attention to the fact that the control variable is not continuous, and is defined only almost everywhere. So, we define the contact and interior sets for the control bounds as

$$\begin{cases} B_- := \{t \in [0,T] :\ \hat{u}_t = u_{\min}\}, \\ B_+ := \{t \in [0,T] :\ \hat{u}_t = u_{\max}\}, \end{cases} \tag{2.29}$$

and set $B := B_- \cup B_+$. We shall make clear that if $u_{\min} = -\infty$ then $B_- = \emptyset$ and, analogously, if $u_{\max} = +\infty$ then $B_+ = \emptyset$. These sets are defined up to a null measure set and they can be identified with their characteristic functions. We define arcs in a similar way, using representatives of the characteristic functions. That is, we say that (a,b) is a *B_-, B_+ arc* (or simply B arc if we do not want to precise in which bound $\hat{u}$ lies) if (a,b) is included, up to a null measure set, in B_-, B_+, (in B_- or B_+) respectively, but no open interval strictly containing (a,b) is. Finally, let S denote the *singular set*

$$S := \{t \in [0,T] : u_{\min} < \hat{u}_t < u_{\max} \text{ and } g(\hat{x}_t) < 0\}. \tag{2.30}$$

Junction times are in general points $\tau \in (0,T)$ at which the trajectory $(\hat{x},\hat{u})$ switches from one type of arc (B_-, B_+, C or S) to another type. We may have, for instance, *CS junctions, BC junctions, etc.* With *BC control* we refer to a control that is a concatenation of a bang arc and a constrained arc, and we extend this notation for any finite sequence, i.e. *SC*, *SBC*, etc.

Remark 3 The result in Corollary 1 above was proved by Maurer [29] at times $\tau \in (0,T)$ being junction points of the state constraint, and extended to state constraints of any order.

Consider the following *geometric hypotheses* on the control structure:

$$\begin{cases} \text{(i)} & \text{the interval } [0,T] \text{ is (up to a zero measure set) the disjoint} \\ & \text{union of finitely many arcs of type } B,\ C \text{ and } S, \\ \text{(ii)} & \text{the control } \hat{u} \text{ is at uniformly positive distance of the bounds} \\ & u_{\min} \text{ and } u_{\max},\ \text{over } C \text{ and } S \text{ arcs}, \\ \text{(iii)} & \text{the control } \hat{u} \text{ is discontinuous at CS and SC junctions.} \end{cases} \tag{2.31}$$

Remark 4 Condition (2.31)(ii) implies the discontinuity of the control at BS and SB junctions. Therefore, (2.31) yields the discontinuity of the control at any junction.

Observe that the example in Remark 2 above verifies the structural hypothesis (2.31).

Let us note that we can rewrite the condition (2.27) of first order state constraint as

$$g'(\hat{x}_t)f_1(\hat{x}_t) \neq 0, \quad \text{on } C. \tag{2.32}$$

From $g(\hat{x}_t) = 0$ on C, we get

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} g(\hat{x}_t) = g'(\hat{x}_t)\big(f_0(\hat{x}_t) + \hat{u}_t f_1(\hat{x}_t)\big), \quad \text{on } C, \tag{2.33}$$

and, whenever (2.32) holds, we obtain that

$$\hat{u}_t = -\frac{g'(\hat{x}_t)f_0(\hat{x}_t)}{g'(\hat{x}_t)f_1(\hat{x}_t)}, \quad \text{on } C. \tag{2.34}$$

In the remainder of the article we assume that $(\hat{x}, \hat{u})$ *satisfies the geometric hypotheses* (2.31) *and that the state constraint is of first order, that is,* (2.32) *holds true.*

2.5 About $\mathrm{d}\mu$

Observe that, along a constrained arc (a, b), over which $u_{\min} < \hat{u}_t < u_{\max}$ a.e., in view of the minimum condition (2.12), we have that $H_u = pf_1 = 0$, for any $(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$. Differentiating this equation with respect to time and using (2.10), we get

$$0 = \mathrm{d}H_u = p[f_1, f_0]\mathrm{d}t - g' f_1 \mathrm{d}\mu. \tag{2.35}$$

Thus, since the state constraint is of first order, $\mathrm{d}\mu$ has a density $\nu \geq 0$ over C given by the absolutely continuous function

$$\nu : [0, T] \to \mathbb{R}, \quad t \mapsto \nu_t := \frac{p_t[f_1, f_0](\hat{x}_t)}{g'(\hat{x}_t)f_1(\hat{x}_t)}, \tag{2.36}$$

where $[X, Y] := X'Y - Y'X$ denotes the *Lie bracket* associated to a pair of vector fields $X, Y : \mathbb{R}^n \to \mathbb{R}^n$.

3 Second order necessary conditions

In this section we state second order necessary conditions for weak optimality of problem (P). We start by defining the cones of critical directions and giving necessary conditions, obtained by applying the abstract result in Theorem 8 in the Appendix. Afterwards, we present the Goh transformation [18] and we state second order conditions in the transformed variables.

3.1 Critical directions

Let us extend the use of $z[v, z^0]$ to denote the solution of the linearized state equation (2.19)-(2.20) for $(v, z^0) \in L^2 \times \mathbb{R}^n$, and let us write T_Φ to refer to the *tangent cone* to $\{0\}_{\mathbb{R}^{n_1}} \times \mathbb{R}_-^{n_2}$ at the point $(\hat{x}_0, \hat{x}_T)$, given by

$$T_\Phi := \{0\}_{\mathbb{R}^{n_1}} \times \{\eta \in \mathbb{R}^{n_2};\ \eta_i \leq 0 : \text{if } \Phi_{n_1+i}(\hat{x}_0, \hat{x}_T) = 0,\ i = 1, \dots, n_2\}.$$

For $v \in L^2$ and $z_0 \in \mathbb{R}^n$, consider the *linearization of the cost and endpoint constraints*

$$\begin{cases} \phi'(\hat{x}_0, \hat{x}_T)(z_0, z_T) \leq 0, \\ \Phi'(\hat{x}_0, \hat{x}_T)(z_0, z_T) \in T_\Phi, \end{cases} \tag{3.1}$$

where $z := z[v, z_0]$. Define the *critical cone* as the set

$$\mathcal{C} := \left\{ \begin{array}{l} (v, z) \in L^\infty \times (W^{1,\infty})^n : (v, z) \text{ satisfies (2.19) and (3.1)}, \\ v_t \geq 0 \ \text{ a.e. on } B_-,\ v_t \leq 0 \ \text{ a.e. on } B_+, \quad g'(\hat{x}_t) z_t \leq 0 \ \text{ on } C \end{array} \right\}. \tag{3.2}$$

Define the *strict critical cone,*

$$\mathcal{C}_S := \left\{ (v, z) \in \mathcal{C} : v_t = 0 \ \text{ a.e. on } B_\pm,\ g'(\hat{x}_t) z_t = 0 \ \text{ on } C \right\}. \tag{3.3}$$

Note that the strict critical cone is a polyhedron of a closed subspace of $L^\infty \times (W^{1,\infty})^n$. Consider the following *weak complementarity condition*: there exists a Lagrange multiplier $(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$ (defined in (2.15)) such that

$$\begin{cases} \text{(i) } H_u(\hat{u}, \hat{x}, p) \neq 0, \quad \text{a.e. over } B, \\ \text{(ii) the support of } \mathrm{d}\mu \text{ is } C. \end{cases} \tag{3.4}$$

We have the following identity:

Proposition 1 *Assume that the weak complementarity condition* (3.4) *holds, then the critical and strict critical cones coincide, this is* $\mathcal{C} = \mathcal{C}_S$.

3.2 Radiality of critical directions

In order to prove the necessary condition of Theorem 2 below, we make use of the abstract result in Theorem 8 of the Appendix and of its related concepts. We consider in particular *radial critical directions* as given in Definition 7 of Section A.2. For our optimal control problem, an element (v, z) of $\mathcal{C}$ is said to be *radial* if and only if, for some $\sigma > 0$, the following conditions are satisfied:

$$u_{\min} \leq \hat{u}_t + \sigma v_t \leq u_{\max}, \quad \text{a.e. on } [0, T], \tag{3.5}$$

$$g(\hat{x}_t) + \sigma g'(\hat{x}_t) z_t \leq 0, \quad \text{for all } t \in [0, T]. \tag{3.6}$$

Recall hypotheses (2.31) and (2.32) on the control structure and the order of the state constraint, respectively.

Proposition 2 *Any critical direction (in* $\mathcal{C}$*) is radial.*

Proof Let $(v, z) \in \mathcal{C}$. Relation (3.5) holds for $\sigma > 0$ small enough over B arcs, and on S and C arcs by (2.31)(ii). Relation (3.6) trivially holds over C arcs and, since the state constraint is not active over B and S arcs, it does also hold over these arcs, except perhaps in the vicinity of entry or exit points to C arcs. For $t \in [0, T]$, let δ_t denote the distance between t and C. By (2.31)(ii)-(iii) and (2.32), $\frac{\mathrm{d}}{\mathrm{d}t} g(\hat{x}_t)$ has a jump at the entry and exit points of any C arc. Let us check that, for $\varepsilon > 0$ and $\sigma > 0$ small enough, we have $g(\hat{x}_t) + \sigma g'(\hat{x}_t) z_t \leq 0$ for all $t \in [0, T]$ such that $\delta_t \in (0, \varepsilon)$. For such t, reducing $\varepsilon > 0$ if necessary, we have that $g(\hat{x}_t) \leq -c_1 \delta_t$. On the other hand, since $g'(\hat{x}_t) z_t$ is Lipschitz continuous and nonpositive over C arcs, we have that $g'(\hat{x}_t) z_t \leq c_2 \delta_t$, where $c_2 > 0$ depends on v. Therefore, $g(\hat{x}_t) + \sigma g'(\hat{x}_t) z_t \leq (\sigma c_2 - c_1) \delta_t < 0$ as soon as $\sigma < c_1 / c_2$. The conclusion follows.

We next give an example of a problem in which the optimal control is not discontinuous at the junction points of C arcs, but whose associated critical cone is nevertheless radial since it reduces to $\{0\}$.

Remark 5 Let us consider the problem

$$\begin{aligned} &\min \int_0^2 x_{1,t} \mathrm{d}t, \\ &\dot{x}_{1,t} = u_t \in [-1, 1], \;\; \dot{x}_{2,t} = 1, \text{ a.e. on } [0, T], \\ &x_{1,0} = 1,\, x_{2,0} = 0; \;\; -(x_{2,t} - 1)^2 - x_{1,t} \leq 0 \text{ over } [0, T]. \end{aligned} \tag{3.7}$$

Notice that $x_{2,t} = t$ and that the state constraint is of first order since $g(x) := -(x_2 - 1)^2 - x_1$ satisfies $\frac{\mathrm{d}}{\mathrm{d}t} g(x_t) = -2(x_{2,t} - 1) - u + t$. Thus, $u_t = -2(t - 1)$ on a constraint arc. It is easy to see that the optimal control is of type $B_- C B_-$. More precisely,

$$\hat{u}_t = \begin{cases} -1, & t \in [0, 1], \\ -2(t-1), & t \in (1, 3/2], \\ -1, & t \in (3/2, 2]. \end{cases} \tag{3.8}$$

Thus, $\hat{u}$ is continuous at the junction time $t = 3/2$. Yet, since no singular arc occurs, the critical cone reduces to $\{0\}$ and, therefore, any critical direction is radial.

3.3 Statement of second order necessary conditions

Next we state a second order necessary condition in terms of the Hessian of the Lagrangian $\mathcal{L}$ (which respect to (u, x)), which is given by the quadratic form

$$Q := Q^0 + Q^E + Q^g, \tag{3.9}$$

where

$$\begin{cases} Q^0(v, z, \lambda) := \displaystyle\int_0^T \big(z_t^\top H_{xx} \, z_t + 2 v_t H_{ux} \, z_t \big) \mathrm{d}t, \\ Q^E(v, z, \lambda) := D^2 \ell^{\beta, \Psi}(z_0, z_T)^2, \\ Q^g(v, z, \lambda) := \displaystyle\int_0^T z_t^\top \, g'' \, z_t \mathrm{d}\mu_t, \end{cases} \tag{3.10}$$

We recall that the Lagrangian $\mathcal{L}$ was defined in (2.14).

Proposition 3 *For every multiplier $\lambda \in \Lambda$, we have*

$$D^2_{(u,x)^2}\mathcal{L}(\hat{u},\hat{x},\lambda)(v,z)^2 = Q(v,z,\lambda), \quad \textit{for all } (v,z) \in \mathcal{C}. \tag{3.11}$$

As a consequence of Proposition 2 and Theorem 8 in the Appendix, the following result holds.

Theorem 2 (Second order necessary condition) *Assume that $(\hat{u},\hat{x})$ is a weak minimum of problem* (P). *Then*

$$\max_{\lambda\in\Lambda} Q(v,z,\lambda) \geq 0, \quad \textit{for all } (v,z) \in \mathcal{C}. \tag{3.12}$$

3.4 Goh transformation and primitives of critical directions

For $(v,z^0) \in L^\infty \times \mathbb{R}^n$, and $z := z[v,z^0]$ being the solution of the linearized state equation (2.19)-(2.20), let us set,

$$y_t := \int_0^t v_s \mathrm{d}s; \quad \xi_t := z_t - y_t f_1(\hat{x}_t), \quad \text{for } t \in [0,T]. \tag{3.13}$$

This change of variables is called *Goh transformation* [18]. Defining $E_t := A_t f_1(\hat{x}_t) - \frac{\mathrm{d}}{\mathrm{d}t} f_1(\hat{x}_t)$ (where A was defined in (2.18)), observe that ξ is solution of

$$\dot{\xi}_t = A_t\xi_t + y_t E_t, \tag{3.14}$$

on the interval $[0,T]$, with initial condition

$$\xi_0 = z^0.$$

Consider the set of *strict primitive critical directions*

$$\mathcal{P}_S := \left\{ (y,h,\xi) \in W^{1,\infty} \times \mathbb{R} \times (W^{1,\infty})^n : y_0 = 0,\ y_T = h,\ (\dot{y}, \xi + y f_1) \in \mathcal{C}_S \right\},$$

and let $\overline{\mathcal{P}_S}$ denote its closure with respect to the $L^2 \times \mathbb{R} \times (H^1)^n$−topology. The final value y_T of y is involved in the definition of $\mathcal{P}_S$ since it becomes an independent variable when we consider its closure. We provide a characterization of $\overline{\mathcal{P}}_S$ in Theorem 3 below.

Let us note that if $(v,z) \in \mathcal{C}_S$ is a strict critical direction, then (y,ξ) given by Goh transform (3.13) satisfies the following conditions:

$$\begin{cases} \text{(i)} \ \ g'(\hat{x}_t)(\xi_t + y_t f_1(\hat{x}_t)) = 0 \text{ on } C, \\ \text{(ii) } y \text{ is constant on each } B \text{ arc}, \end{cases} \tag{3.15}$$

and if (v,z) satisfies the linearized endpoint relations (3.1), then

$$\begin{cases} \phi'(\hat{x}_0,\hat{x}_T)\big(\xi_0, \xi_T + h f_1(\hat{x}_T)\big) \leq 0, \\ \Phi'(\hat{x}_0,\hat{x}_T)\big(\xi_0, \xi_T + h f_1(\hat{x}_T)\big) \in T_\Phi, \end{cases} \tag{3.16}$$

where we set $h := y_T$. In the sequel we let $0 =: \hat{\tau}_0 < \hat{\tau}_1 < \cdots < \hat{\tau}_N := T$ denote the union of the set of junction times with $\{0,T\}$.

A B arc starting at time 0 (respectively ending at time T) is called a $B_{0\pm}$ (respectively $B_{T\pm}$) *arc.*

Proposition 4 *Any* $(y, h, \xi) \in \overline{\mathcal{P}_S}$ *verifies* (3.15)-(3.16) *and*

$$\begin{cases} \text{(i)} & y \text{ is continuous at the } BC,\ CB \text{ and } BB \text{ junctions,} \\ \text{(ii)} & y_t = 0, \text{ on } B_{0\pm} \text{ if a } B_{0\pm} \text{ arc exists,} \\ \text{(iii)} & y_t = h, \text{ on } B_{T\pm} \text{ if a } B_{T\pm} \text{ arc exists,} \\ \text{(iv)} & \lim_{t\uparrow T} y_t = h, \text{ if } T \in C. \end{cases} \tag{3.17}$$

Proof Let $(y, h, \xi) \in \overline{\mathcal{P}_S}$ be the limit of a sequence $(y^k, y^k_T, \xi^k)_k \subset \mathcal{P}_S$ in the $L^2 \times \mathbb{R} \times (H^1)^n-$topology. By the Ascoli-Arzelà theorem, if (y^k) is equi-bounded and equi-Lipschitz continuous over an interval $[a, b]$, then $y^k \to y$ uniformly over $[a, b]$ and the limit function y is Lipschitz continuous on $[a, b]$. Consequently, we obtain that y is constant over B arcs, and continuous at BC, CB and BB junctions. We can prove the other statements by an analogous reasoning.

Define the set

$$\mathcal{P}^2_S := \Big\{ (y, h, \xi) \in L^2 \times \mathbb{R} \times (H^1)^n : \ \text{(3.14)-(3.17) hold} \Big\}. \tag{3.18}$$

Then the following characterization holds.

Theorem 3 *We have that* $\mathcal{P}^2_S = \overline{\mathcal{P}_S}$.

Proof That $\overline{\mathcal{P}_S} \subset \mathcal{P}^2_S$ follows from Proposition 4. Let us prove the converse inclusion. Define the linear space

$$\mathcal{Z} := \left\{ \begin{array}{l} (y, y_T, \xi) \in W^{1,\infty} \times \mathbb{R} \times \left(\displaystyle\prod_{j=1}^{N} W^{1,\infty}(\hat\tau_{j-1}, \hat\tau_j)^n \right) \\ y_0 = 0, \ \text{(3.14) holds at each } (\hat\tau_{j-1}, \hat\tau_j), \text{ and (3.15) holds} \end{array} \right\}, \tag{3.19}$$

that is obtained from $\mathcal{P}_S$ by removing condition (3.16) and allowing ξ to be discontinuous at the junctions $\hat\tau_j$, $j = 1, \dots N-1$.

Let $(y, h, \xi) \in \mathcal{P}^2_S$. For any $\varepsilon > 0$, we now construct $(y_\varepsilon, y_{\varepsilon,T}, \xi_\varepsilon) \in \mathcal{Z}$ such that $y_{\varepsilon,T} = h$ and

$$\|y - y_\varepsilon\|_2 + \|\xi - \xi_\varepsilon\|_\infty = o(1). \tag{3.20}$$

First set

$$y_{\varepsilon,t} := y_t, \quad \xi_{\varepsilon,t} := \xi_t, \quad \text{on } B \cup C. \tag{3.21}$$

On S, define y_ε in such a way that $y_\varepsilon \in W^{1,\infty}(0, T)$, the values at the junction times are fixed in the following way:

$$\begin{cases} y_\varepsilon(\hat\tau_j+) = y(\hat\tau_j-), & \text{if } \hat\tau_j > 0 \text{ is an entry point of an } S \text{ arc,} \\ y_\varepsilon(\hat\tau_j-) = y(\hat\tau_j+), & \text{if } \hat\tau_j < T \text{ is an exit point of an } S \text{ arc,} \\ y_{\varepsilon,0} = 0, \quad \text{if } 0 \in S; \quad y_{\varepsilon,T} = h, \quad \text{if } T \in S, \end{cases} \tag{3.22}$$

and $\|y - y_\varepsilon\|_2 < \varepsilon$. Such an y_ε exists, see [6, Lemma 8.1]. Define ξ_ε over each S arc by integrating (3.14) over the respective arc with $y = y_\varepsilon$ and the initial condition

$\xi_{\varepsilon,\tau} = \xi_\tau$, where τ denotes the entry point of the arc. Then $(y_\varepsilon, y_{\varepsilon,T}, \xi_\varepsilon) \in \mathcal{Z}$ satisfies the estimate (3.20). In particular, we have

$$|\xi_\varepsilon(\hat\tau_j-) - \xi_\varepsilon(\hat\tau_j+)| = |\xi_\varepsilon(\hat\tau_j-) - \xi(\hat\tau_j)| = o(1), \quad \text{for all } j = 1,\dots,N-1, \tag{3.23}$$

$$|\phi'(\hat x_0, \hat x_T)(\xi_{\varepsilon,0}, \xi_{\varepsilon,T} + f_1(\hat x_T)y_{\varepsilon,T}) - \phi'(\hat x_0, \hat x_T)(\xi_0, \xi_T + f_1(\hat x_T)h)| = o(1), \tag{3.24}$$

$$|\Phi'(\hat x_0, \hat x_T)(\xi_{\varepsilon,0}, \xi_{\varepsilon,T} + f_1(\hat x_T)y_{\varepsilon,T}) - \Phi'(\hat x_0, \hat x_T)(\xi_0, \xi_T + f_1(\hat x_T)h)| = o(1). \tag{3.25}$$

Notice that the cone $\mathcal{P}_S$ is obtained from $\mathcal{Z}$ by adding the constraints in (3.16) and the continuity conditions

$$\xi(\hat\tau_j-) - \xi(\hat\tau_j+) = 0, \quad \text{for all } j = 1,\dots,N-1. \tag{3.26}$$

In view of Hoffman's lemma [22] and estimates (3.23)-(3.25), we get that there exists $(\tilde y_\varepsilon, \tilde y_{\varepsilon,T}, \tilde\xi_\varepsilon) \in \mathcal{P}_S$ such that

$$\|\tilde y_\varepsilon - y_\varepsilon\|_2 + \|\tilde\xi_\varepsilon - \xi_\varepsilon\|_\infty = o(1). \tag{3.27}$$

Finally, from (3.20) and (3.27) we have that

$$\|\tilde y_\varepsilon - y\|_2 + \|\tilde\xi_\varepsilon - \xi\|_\infty = o(1), \tag{3.28}$$

and hence, the density of $\mathcal{P}_S$ in $\mathcal{P}_S^2$ follows.

3.5 Goh transformation on the Hessian of Lagrangian

Next, we want to express each of the quadratic functions in (3.9)-(3.10) as functions of (y,h,ξ). For the terms that are quadratic in z, it suffices to replace z by $\xi + f_1(\hat x)y$ as given in Goh transform. With this aim, set for $(y,h,\xi) \in L^2 \times \mathbb{R} \times (H^1)^n$ and $\lambda := (\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$,

$$\begin{aligned}
\Omega_T(y,h,\xi,\lambda) &:= 2h\, H_{ux,T}\xi_T + h\, H_{ux,T} f_1(\hat x_T)h,\\
\Omega^0(y,h,\xi,\lambda) &:= \int_0^T \left(\xi_t^\top H_{xx}\xi_t + 2y_t M\xi_t + y_t R y_t\right)\mathrm{d}t,\\
\Omega^E(y,h,\xi,\lambda) &:= D^2\ell^{\beta,\Psi}(\xi_0, \xi_T + f_1(\hat x_T)h)^2,\\
\Omega^g(y,h,\xi,\lambda) &:= \int_0^T (\xi_t + f_1(\hat x_t)y)^\top g''(\hat x_t)(\xi_t + f_1(\hat x_t)y)\mathrm{d}\mu_t,\\
\Omega &:= \Omega_T + \Omega^0 + \Omega^E + \Omega^g,
\end{aligned} \tag{3.29}$$

with

$$M := f_1^\top H_{xx} - \dot H_{ux} - H_{ux}A, \tag{3.30}$$

$$R := f_1^\top H_{xx} f_1 - 2H_{ux}E - \frac{\mathrm{d}}{\mathrm{d}t}(H_{ux}f_1). \tag{3.31}$$

Here, when we omit the argument of M, R, H or its derivatives, we mean that they are evaluated at $(\hat x, \hat u, \lambda)$.

Remark 6 Easy computations show that R does not depend on u. More precisely, R is given by

$$R(\hat x_t, \lambda_t) = p_t\big[[f_0, f_1], f_1\big](\hat x_t) + g'(\hat x_t) f_1'(\hat x_t) f_1(\hat x_t)\nu_t, \tag{3.32}$$

where ν is the density of $\mathrm{d}\mu$ given in (2.36).

Proposition 5 *Let* $(v, z) \in L^2 \times (H^1)^n$ *be a solution of* (2.19) *and let* (y, ξ) *be defined by the Goh transformation* (3.13). *Then, for any* $\lambda \in \Lambda$,

$$Q(v, z, \lambda) = \Omega(y, y_T, \xi, \lambda). \tag{3.33}$$

Proof Take (v, z) and (y, ξ) as in the statement. It is straightforward to prove that

$$Q^E(v, z, \lambda) = \Omega^E(y, y_T, \xi, \lambda), \quad \text{and} \quad Q^g(v, z, \lambda) = \Omega^g(y, y_T, \xi, \lambda). \tag{3.34}$$

In order to prove the equality between Q^0 and Ω^0, let us replace each occurrence of z in Q^0, by its expression in the Goh transformation (3.13), i.e. change z by $\xi + f_1(\hat{x})y$. The first term in Q^0 can be written as

$$\int_0^T z^\top H_{xx} z \mathrm{d}t = \int_0^T (\xi + f_1 y)^\top H_{xx}(\xi + f_1 y)\mathrm{d}t. \tag{3.35}$$

Let us consider the second term in Q^0 :

$$\int_0^T v H_{ux}(\xi + f_1 y)\mathrm{d}t = \int_0^T \big(v H_{ux}\xi + v H_{ux} f_1 y\big)\mathrm{d}t. \tag{3.36}$$

Integrating by parts the first term in previous equation we get

$$\begin{aligned} \int_0^T v H_{ux}\xi \mathrm{d}t &= [y H_{ux}\xi]_0^T - \int_0^T y\big(\dot{H}_{ux}\xi + H_{ux}\dot{\xi}\big)\mathrm{d}t \\ &= y_T H_{ux,T}\xi_T - \int_0^T y\big(\dot{H}_{ux}\xi + H_{ux}A\xi + H_{ux}Ey\big)\mathrm{d}t. \end{aligned} \tag{3.37}$$

For the second term in the right hand-side of (3.36) we obtain

$$\int_0^T v H_{ux} f_1 y \mathrm{d}t = [y H_{ux} f_1 y]_0^T - \int_0^T y \left(\frac{\mathrm{d}}{\mathrm{d}t}(H_{ux} f_1) y + H_{ux} f_1 v\right) \mathrm{d}t. \tag{3.38}$$

This identity yields the following

$$\int_0^T v H_{ux} f_1 y \mathrm{d}t = \tfrac{1}{2} y_T H_{ux,T} f_1 y_T - \tfrac{1}{2}\int_0^T y^2 \frac{\mathrm{d}}{\mathrm{d}t}(H_{ux} f_1)\, \mathrm{d}t. \tag{3.39}$$

From (3.35), (3.37) and (3.39) we get the desired result.

3.6 Second order necessary condition in the new variables

We can obtain the following new necessary condition in the variables after Goh's transformation.

Theorem 4 *If* $(\hat{u}, \hat{x})$ *is a weak minimum, then*

$$\max_{\lambda \in \Lambda} \Omega(y, h, \xi, \lambda) \geq 0, \quad \textit{for all } (y, h, \xi) \in \mathcal{P}_S^2. \tag{3.40}$$

Proof Let us assume first that the qualification condition (2.21)(i) does not hold. Therefore, there exists a nonzero element Ψ_E in $[\operatorname{Im} D\bar{\Phi}_E(\hat{u},\hat{x}_0)]^\perp$. Hence, the multiplier λ composed by such Ψ_E, and having $(\beta,\Psi_I,\mathrm{d}\mu) := 0$ and the associated costate p, is a Lagrange multiplier, as well as its opposite $-\lambda$. Therefore, either $\Omega(y,h,\xi,\lambda)$ or $\Omega(y,h,\xi,-\lambda)$ is greater or equal zero, and the conclusion follows.

Let us now consider the case when (2.21)(i) holds true. Then, the corresponding abstract problem (P_A) (defined in the Appendix A.1) verifies the qualification condition (A.5)(i). Recall the definition of $\hat{\Lambda}$ in (A.10), and the corresponding set $\hat{\Lambda}_1$ (see (2.16)). In view of Theorem 6, $\hat{\Lambda}_1$ is non empty and bounded. Furthermore, due to the Banach-Alaoglu Theorem, and since the space of continuous functions in $[0,T]$ is separable, we get that $\hat{\Lambda}_1$ is weakly* sequentially compact.

Consider now $(y,h,\xi) \in \mathcal{P}_S^2$. By Theorem 3, there exists a sequence (y^k, y^k_T, ξ^k) in $\mathcal{P}_S$ such that

$$(y^k, y^k_T, \xi^k) \to (y,h,\xi), \quad \text{in the } L^2 \times \mathbb{R} \times (H^1)^n\text{-topology.} \tag{3.41}$$

By Proposition 5, for all $\lambda \in \Lambda$,

$$\Omega(y^k, y^k_T, \xi^k, \lambda) = Q(v^k, z^k, \lambda), \tag{3.42}$$

where $v^k := \frac{\mathrm{d}}{\mathrm{d}t} y^k$ and $z^k := \xi^k + f_1 y^k$. For each (v^k, z^k), due to Theorem 2, there exists $\lambda^k \in \Lambda$ for which

$$Q(v^k, z^k, \lambda^k) \geq 0. \tag{3.43}$$

Let $\hat{\lambda}_k$ be the corresponding element of $\hat{\Lambda}_1$ given by the bijection (A.11), and consider $\bar{\lambda}_k \in \Lambda$ such that $\hat{\lambda}_k = (1, \bar{\lambda}_k)$. Then

$$Q(v^k, z^k, \bar{\lambda}^k) \geq 0, \tag{3.44}$$

in view of (3.43) and since $\bar{\lambda}_k$ is obtained from λ_k by multiplying by a positive scalar. Given that $\hat{\Lambda}_1$ is weakly* sequentially compact, there exists a subsequence $(\hat{\lambda}^{k_j})_j$ weakly* convergent to $\hat{\lambda} = (1, \bar{\lambda}) \in \hat{\Lambda}_1$, where $\bar{\lambda} \in \Lambda$. Thus, $\bar{\lambda}_k \xrightharpoonup{*} \bar{\lambda}$ in Λ and we get

$$\Omega(y,h,\xi,\bar{\lambda}) = \lim_{k\to\infty} \Omega(y^k, y^k_T, \xi^k, \bar{\lambda}^k) = \lim_{k\to\infty} Q(v^k, z^k, \bar{\lambda}^k) \geq 0. \tag{3.45}$$

This concludes the proof.

4 Second order sufficient conditions

In this section we show a second order sufficient condition in terms of the uniform positivity of Ω and guaranteeing that the nominal solution $(\hat{u},\hat{x})$ is a strict Pontryagin minimum whenever this condition is satisfied.

To state the main result of this section (see Theorem 5) we need to introduce the following concepts.

Definition 2 We say that $(\hat{u},\hat{x})$ is a *Pontryagin minimum* of problem (P) if for any $M > 0$, there exists $\varepsilon_M > 0$ such that $(\hat{u},\hat{x})$ is a minimum in the set of feasible trajectories (u,x) satisfying

$$\|x - \hat{x}\|_\infty + \|u - \hat{u}\|_1 < \varepsilon_M, \quad \|u - \hat{u}\|_\infty < M. \tag{4.1}$$

A sequence $(v_k) \subset L^\infty$ is said to *converge to 0 in the Pontryagin sense* if $\|v_k\|_1 \to 0$ and there exists $M > 0$ such that $\|v_k\|_\infty \leq M$, for all $k \in \mathbb{N}$.

Definition 3 Let us define the function $\gamma : L^2 \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, given by

$$\gamma(y, h, x_0) := \int_0^T y_t^2 \mathrm{d}t + h^2 + |x_0|^2. \tag{4.2}$$

We say that $(\hat{u}, \hat{x})$ satisfies the $\gamma-$*growth condition in the Pontryagin sense* if there exists $\rho > 0$ such that, for every sequence of feasible variations $(v_k, \delta x_k)$ having (v_k) convergent to 0 in the Pontryagin sense and $\delta x_{k,0} \to 0$, one has

$$\phi(\hat{x}_0 + \delta x_{k,T}\, , \hat{x}_T + \delta x_{k,T}) - \phi(\hat{x}_0, \hat{x}_T) \geq \rho\gamma(y_k, y_{k,T}, \delta x_{k,0}), \tag{4.3}$$

for k large enough, where $y_{k,t} := \int_0^t v_{k,s} \mathrm{d}s$, for $t \in [0, T]$.

Definition 4 We say that $(\hat{u}, \hat{x})$ satisfies *strict complementarity condition for the control constraints* if the following conditions hold:

(i) $\max_{\lambda \in \Lambda} H_u(\hat{u}_t, \hat{x}_t, p_t) > 0$, for all t in the interior of B_-,
(ii) $\min_{\lambda \in \Lambda} H_u(\hat{u}_t, \hat{x}_t, p_t) < 0$, for all t in the interior of B_+,
(iii) $\max_{\lambda \in \Lambda} H_u(\hat{u}_0, \hat{x}_0, p_0) > 0$, if $0 \in B_-$, and $\max_{\lambda \in \Lambda} H_u(\hat{u}_T, \hat{x}_T, p_T) > 0$, if $T \in B_-$, $\min_{\lambda \in \Lambda} H_u(\hat{u}_0, \hat{x}_0, p_0) < 0$, if $0 \in B_+$, and $\min_{\lambda \in \Lambda} H_u(\hat{u}_T, \hat{x}_T, p_T) < 0$, if $T \in B_+$.

Consider the following *extended cones of critical directions* (compare to $\mathcal{P}_S^2$ defined in (3.18)):

$$\begin{aligned} &\hat{\mathcal{P}}^2 := \{(y, h, \xi) \in L^2 \times \mathbb{R} \times (H^1)^n : \text{(3.14)-(3.16), (3.17)(ii)-(iii) hold}\},\\ &\hat{\hat{\mathcal{P}}}^2 := \{(y, h, \xi) \in \hat{\mathcal{P}}^2 : \text{(3.17)(iv) holds}\}, \end{aligned} \tag{4.4}$$

and

$$\mathcal{P}_*^2 := \begin{cases} \hat{\hat{\mathcal{P}}}^2, \text{ if } T \in C \text{ and } [\mu(T)] > 0, \text{ for some } (\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda, \\ \hat{\mathcal{P}}^2, \text{ otherwise.} \end{cases} \tag{4.5}$$

Let us recall the definition of Legendre form (see e.g. [21]):

Definition 5 Let W be a Hilbert space. We say that a quadratic mapping $\mathcal{Q} : W \to \mathbb{R}$ is a *Legendre form* if it is sequentially weakly lower semi continuous such that, if $w_k \to w$ weakly in W and $\mathcal{Q}(w_k) \to \mathcal{Q}(w)$, then $w_k \to w$ strongly.

Theorem 5 *Suppose that the following conditions hold true:*

(i) $(\hat{u}, \hat{x})$ *satisfies strict complementarity for the control constraint and the complementarity for the state constraint* (3.4)*(ii);*
(ii) for each $\lambda \in \Lambda$, $\Omega(\cdot, \lambda)$ *is a Legendre form in the linear space* $\{(y, h, \xi) \in L^2 \times \mathbb{R} \times (H^1)^n : (3.14)$ *holds*$\}$; *and*
(iii) there exists $\rho > 0$ *such that*

$$\max_{\lambda \in \Lambda} \Omega(y, h, \xi, \lambda) \geq \rho\gamma(y, h, \xi_0), \quad \textit{for all } (y, h, \xi) \in \mathcal{P}_*^2. \tag{4.6}$$

Then $(\hat{u}, \hat{x})$ *is a Pontryagin minimum satisfying* $\gamma-$*growth.*

Remark 7 We have that $\Omega(\cdot,\lambda)$ is a Legendre form iff there exists $\alpha > 0$, such that, for every $\lambda \in \Lambda$,

$$R(\hat{x}_t, \lambda_t) + f_1(\hat{x}_t)^\top g''(\hat{x}_t) f_1(\hat{x}_t)\nu_t > \alpha, \quad \text{on } [0,T], \tag{4.7}$$

where ν is the density of $\mathrm{d}\mu$, which vanishes on $[0,T]\backslash C$, and is given by the expression (2.36) on the set C. See [21] for more details.

The remainder of this section is devoted to the proof of Theorem 5. We first show the technical results in Lemma 3, Propositions 6 and 7, and then give the proof of the Theorem 5.

For the lemma below recall the definition of the Lagrangian function $\mathcal{L}$ which was given in (2.14).

Lemma 3 *Let $(u,x) \in L^2 \times (H^1)^n$ be a solution of (2.1), and set $(\delta x, \delta u) := (x,u) - (\hat{x},\hat{u})$. Then, the following expression for the Lagrangian function holds for every multiplier $\lambda \in \Lambda$,*

$$\mathcal{L}(u,x,\lambda) = \mathcal{L}(\hat{u},\hat{x},\lambda) + \tilde{Q}(\delta u, \delta x, \lambda) + r(\delta u, \delta x, \lambda), \tag{4.8}$$

where

$$\tilde{Q}(\delta u, \delta x, \lambda) := \int_0^T H_u \delta u \mathrm{d}t + Q(\delta u, \delta x, \lambda)$$

and

$$\begin{aligned} r(\delta u, \delta x) &:= \mathcal{O}\big(|(\delta x_0, \delta x_T)|^3\big) \\ &+ \int_0^T \big\{\tfrac{1}{2} H_{uxx}(\delta u, \delta x, \delta x) + p(\hat{u} + \delta u)\mathcal{O}\big(|\delta x|^3\big)\big\}\mathrm{d}t + \int_0^T \mathcal{O}\big(|\delta x|^3\big)\mathrm{d}\mu. \end{aligned}$$

Proof Let us consider the following second order Taylor expansions, written in compact form,

$$\begin{aligned} \ell^{\beta,\Psi}(x_0,x_T) &= \ell^{\beta,\Psi} + D\ell^{\beta,\Psi}(\delta x_0, \delta x_T) + \tfrac{1}{2} D^2 \ell^{\beta,\Psi}(\delta x_0, \delta x_T)^2 + \mathcal{O}\big(|(\delta x_0, \delta x_T)|^3, \\ f_i(x_t) &= f_{i,t} + Df_i \delta x_T + \tfrac{1}{2} D^2 f_i \delta x_t^2 + \mathcal{O}\big(|\delta x|^3\big), \quad i = 0,1, \\ g(x_t) &= \tfrac{1}{2} g + Dg\delta x_t + \tfrac{1}{2} g'' \delta x_t^2 + \mathcal{O}\big(|\delta x|^3\big). \end{aligned} \tag{4.9}$$

Observe that

$$D\ell^{\beta,\Psi}(\delta x_0, \delta x_T) = [p\,\delta x]_0^T = \int_0^T p(-f_x \delta x + \dot{\delta x})\mathrm{d}t - \int_0^T g'(\hat{x})\delta x \mathrm{d}\mu. \tag{4.10}$$

Using the identities (4.9) and (4.10) in the following equation

$$\mathcal{L}(u,x,\lambda) = \ell^{\beta,\Psi}(x_0,x_T) + \int_0^T p\big(f_0 + uf_1 - \dot{x}\big)\mathrm{d}t + \int_0^T g(x)\mathrm{d}\mu, \tag{4.11}$$

we obtain the desired expression for $\mathcal{L}(u,x,\lambda)$. The result follows.

In view of the previous Lemma 3 and [6, Lemma 8.4] we can prove the following:

Proposition 6 *Let $(v_k) \subset L^\infty$ be a sequence converging to 0 in the Pontryagin sense and $(x_{k,0})$ a sequence in $\mathbb{R}^n$ converging to $\hat{x}_0$. Set $u_k := \hat{u} + v_k$, and let x_k be the corresponding solution of equation* (2.1) *with initial value equal to $x_{k,0}$. Then, for every $\lambda \in \Lambda$, one has*

$$\mathcal{L}(u_k, x_k, \lambda) = \mathcal{L}(\hat{u}, \hat{x}, \lambda) + \int_0^T H_u(t) v_{k,t} dt + Q(v_k, z_k, \lambda) + o(\gamma_k), \tag{4.12}$$

where $z_k := z[v_k, x_{k,0} - \hat{x}_0]$, $y_{k,t} := \int_0^t v_{k,s} \mathrm{d}s$ and $\gamma_k := \gamma(y_k, y_{k,T}, x_{k,0} - \hat{x}_0)$.

Remark 8 Notice that Lemma 8.4 in [6] was proved for γ depending only on (y, h) since the initial value x_0 was fixed throughout the article, while here γ depends also on the initial variation of the state. The extension of Lemmas 8.3 and 8.4 in [6] for the case with variable initial state are immediate, and the proofs are given in detail in [5].

Proposition 7 *Let $(p, \mathrm{d}\mu) \in (BV)^{n*} \times \mathcal{M}$ verify the costate equation* (2.10)-(2.11), *and let $(v, z) \in L^2 \times (H^1)^n$ satisfy the linearized state equation* (2.19). *Then,*

$$\int_0^T g'(\hat{x}_t) z_t \mathrm{d}\mu_t + D\ell^{\beta,\Psi}(z_0, z_T) = \int_0^T H_u(t) v_t \mathrm{d}t. \tag{4.13}$$

Proof Note that

$$\begin{aligned} \int_0^T g'(\hat{x}_t) z_t \mathrm{d}\mu_t &= -\int_0^T \sum_{j=1}^n z_{j,t} \mathrm{d}p_{j,t} - \int_0^T p_t f_x(\hat{u}_t, \hat{x}_t) z_t \mathrm{d}t \\ &= -\int_0^T \sum_{j=1}^n z_{j,t} \mathrm{d}p_{j,t} - \int_0^T p_t (\dot{z}_t - f_1(\hat{x}_t) v_t) \mathrm{d}t \\ &= -[pz]_{0-}^{T+} + \int_0^T H_u(t) v_t \mathrm{d}t. \end{aligned} \tag{4.14}$$

The wanted result follows from latter equation and the boundary conditions (2.11).

Proof (of Theorem 5). Let us suppose on the contrary that the conclusion does not hold. Hence, there should exist a sequence $(v_k, x_{k,0}) \subset L^\infty \times \mathbb{R}^n$ of non identically zero functions having (v_k) converging to 0 in the Pontryagin sense and $x_{k,0} \to \hat{x}_0$, and such that, setting $u_k := \hat{u} + v_k$, the corresponding solutions x_k of (2.1) are feasible and

$$\phi(x_{k,0}, x_{k,T}) \leq \phi(\hat{x}_0, \hat{x}_T) + o(\gamma_k), \tag{4.15}$$

where $y_{k,t} := \int_0^t v_{k,s} \mathrm{d}s$ and $\gamma_k := \gamma(y_k, y_{k,T}, x_{k,0} - \hat{x}_0)$. Set $z_k := z[v_k, z_{k,0}]$ with $z_{k,0} := x_{k,0} - \hat{x}_0$. Take any $\lambda = (\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$, and multiply inequality (4.15) by β (which is nonnegative), afterwards add the nonpositive term $\Psi \cdot \Phi(x_{k,0}, x_{k,T}) + \int_0^T g(x_k) \mathrm{d}\mu$ to the left hand-side of the resulting inequality, and obtain the following:

$$\mathcal{L}(u_k, x_k, \lambda) \leq \mathcal{L}(\hat{u}, \hat{x}, \lambda) + o(\gamma_k). \tag{4.16}$$

Set $(\bar{y}_k, \bar{h}_k, \bar{\xi}_{k,0}) := (y_k, y_{k,T}, z_{k,0}) / \sqrt{\gamma_k}$. Note that the elements of this sequence have unit norm in $L^2 \times \mathbb{R} \times \mathbb{R}^n$. By the Banach-Alaoglu Theorem, extracting if necessary a subsequence, we may assume that there exists $(\bar{y}, \bar{h}, \bar{\xi}_0) \in L^2 \times \mathbb{R} \times \mathbb{R}^n$ such that

$$\bar{y}_k \rightharpoonup \bar{y}, \text{ and } (\bar{h}_k, \bar{\xi}_{k,0}) \to (\bar{h}, \bar{\xi}_0), \tag{4.17}$$

where the first limit is given in the weak topology of L^2. Define $\bar{\xi}$ as the solution of (3.14) associated to $\bar{y}$ and the initial condition $\bar{\xi}_0$.

The remainder of the proof is split in two parts:

Fact 1: The weak limit $(\bar{y}, \bar{h}, \bar{\xi})$ belongs to $\mathcal{P}_*^2$.

Fact 2: The inequality (4.16) contradicts the hypothesis of uniform positivity (4.6).

Proof of Fact 1. Recall the definition of the cone $\mathcal{P}_*^2$ given in equation (4.5). The proof of Fact 1 into four parts, in which we establish: *A)* condition (3.15)(*ii*), *B)* (3.17)(ii)-(iii), *C)* (3.15)(i) and (3.17)(iv), and *D)* that $(\bar{y}, \bar{h}, \bar{\xi})$ satisfies (3.16).

A) Let us show that $(\bar{y}, \bar{h}, \bar{\xi})$ verifies (3.15)(ii), i.e. $\bar{y}$ is constant on each B arc. From (4.12), (4.16) and the equivalence between Q and Ω stated in Proposition 5, it follows that

$$-\Omega(\xi_k, y_k, h_k, \lambda) + o(\gamma_k) \geq \int_0^T H_u(t) v_{k,t} \mathrm{d}t \geq 0, \tag{4.18}$$

where ξ_k is solution of (3.14) corresponding to y_k. The last inequality in (4.18) holds in view of the minimum condition (2.12) and since $\hat{u} + v_k$ satisfies the control constraint (2.4). By the continuity of the mapping Ω over the space $L^2 \times \mathbb{R} \times (H^1)^n$ and from (4.18) we deduce that

$$0 \leq \int_0^T H_u(t) v_{k,t} \mathrm{d}t \leq O(\gamma_k).$$

Hence, since the integrand in previous inequality is nonnegative for all $k \in \mathbb{N}$, we have that

$$\lim_{k\to\infty} \int_0^T H_u(t) \varphi_t \frac{v_{k,t}}{\sqrt{\gamma_k}} dt = 0, \tag{4.19}$$

for any nonnegative C^1 function $\varphi : [0, T] \to \mathbb{R}$. Let us consider, in particular, such a function φ having its support included in a B arc (c, d). Integrating by parts in (4.19) and in view of (4.17), we obtain

$$0 = \lim_{k\to\infty} \int_c^d \frac{\mathrm{d}}{\mathrm{d}t} \left(H_u(t) \varphi_t\right) \bar{y}_{k,t} \mathrm{d}t = \int_c^d \frac{\mathrm{d}}{\mathrm{d}t} \left(H_u(t) \varphi_t\right) \bar{y}_t \mathrm{d}t.$$

Over (c, d), v_k has constant sign and therefore, $\bar{y}$ is either nondecreasing or nonincreasing. Thus, we can integrate by parts in the previous equation to get

$$\int_c^d H_u(t) \varphi_t \mathrm{d}\bar{y}_t = 0. \tag{4.20}$$

Take now any $t_0 \in (c, d)$. By the strict complementary condition for the control constraint assumed here (see Definition 4), there exists a multiplier a multiplier $\hat{\lambda} = (\hat{\beta}, \hat{\Psi}, \hat{p}, \mathrm{d}\hat{\mu}) \in \Lambda$ such that $H_u(\hat{u}_{t_0}, \hat{x}_{t_0}, \hat{p}_{t_0}) > 0$. Hence, in view of the continuity of H_u on B,[1] there exists $\varepsilon > 0$ such that $H_u(\hat{u}_t, \hat{x}_t, \hat{p}_t) > 0$ on $(t_0 - 2\varepsilon, t_0 + 2\varepsilon) \subset$

[1] Actually H_u is continuous on B since p does not jump on B.

(c,d). Choose φ such that $\operatorname{supp}\varphi \subset (t_0-2\varepsilon, t_0+2\varepsilon)$, and $H_u(\hat{u}_t,\hat{x}_t,\hat{p}_t)\varphi_t = 1$ on $(t_0-\varepsilon, t_0+\varepsilon)$. Since $\mathrm{d}\bar{y} \geq 0$, equation (4.20) yields

$$\begin{aligned} 0 &= \int_c^d H_u(\hat{u}_t,\hat{x}_t,\hat{p}_t)\varphi_t \mathrm{d}\bar{y}_t \geq \int_{t_0-\varepsilon}^{t_0+\varepsilon} H_u(\hat{u}_t,\hat{x}_t,\hat{p}_t)\varphi_t \mathrm{d}\bar{y}_t \\ &= \int_{t_0-\varepsilon}^{t_0+\varepsilon} \mathrm{d}\bar{y}_t = \bar{y}_{t_0+\varepsilon} - \bar{y}_{t_0-\varepsilon}. \end{aligned} \tag{4.21}$$

As both ε and $t_0 \in (c,d)$ are arbitrary we deduce that

$$\mathrm{d}\bar{y}_t = 0, \quad \text{on } B. \tag{4.22}$$

Hence, $(\bar{y},\bar{h},\bar{\xi})$ satisfies (3.15)(ii).

B) Assume now that a $B_{0\pm}$ arc exists, and let us prove that $\bar{y}=0$ on $B_{0\pm}$ (see the definition of $B_{0\pm}$ in the paragraph preceding Proposition 4). Let $B_{0\pm}$ be equal to $[0,t_1]$ for some $t_1>0$. Assume without loss of generality that $\hat{u}=u_{\min}$ on $[0,t_1]$. Notice that by the strict complementarity condition for the control constraint (condition (i) of the present theorem) there exists $\lambda' = (\beta',\Psi',p',\mathrm{d}\mu') \in \Lambda$ and $\varepsilon,\delta>0$ such that $H_u(\hat{u}_t,\hat{x}_t,p'_t) > \delta$ for all $t\in[0,\varepsilon]$, and thus, by considering in (4.19) a nonnegative Lipschitz continuous function $\varphi:[0,T]\to\mathbb{R}$ being equal to $1/\delta$ on $[0,t]$, we obtain $\bar{y}_{k,t} = \displaystyle\int_0^t \frac{v_{k,s}}{\sqrt{\gamma_k}}\mathrm{d}s \to 0$, since $v_k \geq 0$ on $[0,t_1]$. Hence $\bar{y}=0$ on $[0,\varepsilon]$. This last assertion, together with (4.22), imply that $\bar{y}=0$ on $B_{0\pm}$.

Suppose that T is in a boundary arc $B_{T\pm}$. Let $B_{T\pm}=[t_N,T]$. Then, we can derive that for some $\varepsilon>0$, $\bar{y}_{k,T}-\bar{y}_{k,t} = \displaystyle\int_t^T \frac{v_{k,s}}{\sqrt{\gamma_k}}\mathrm{d}s \to 0$ for all $t\in[T-\varepsilon,T]$, by an argument analogous to the one above. Thus, $\bar{y}_t=\bar{h}$ on $[T-\varepsilon,T]$, and hence, by (4.22) we get that

$$\bar{y}=\bar{h}, \text{ on } (t_N,T], \text{ if } T\in B. \tag{4.23}$$

Therefore, $(\bar{y},\bar{h},\bar{\xi})$ verifies (3.17)(ii)-(iii)

C) Let us prove that $(\bar{y},\bar{h},\bar{\xi})$ satisfies (3.15)(i) and (3.17)(iv). For all $t\in[0,T]$, a first order Taylor expansion gives

$$0 \geq g(x_{k,t}) = g(\hat{x}_t) + g'(\hat{x}_t)\delta x_{k,t} + O(|\delta x_{k,t}|^2). \tag{4.24}$$

From latter estimate and [6, Lemma 8.12], we deduce that

$$g'(\hat{x}_t)z_{k,t} \leq o(\sqrt{\gamma_k}) + O(|\delta x_{k,t}|^2), \quad \text{for all } t\in C. \tag{4.25}$$

Let $\varphi \geq 0$ be some continuous function with support in C. From (4.25), we get that

$$\begin{aligned} \int_0^T \varphi_t g'(\hat{x}_t)(\xi_{k,t}+f_1(\hat{x}_t)y_{k,t})\mathrm{d}t &= \int_0^T \varphi_t g'(\hat{x}_t)z_{k,t}\mathrm{d}t \\ &\leq \|\varphi\|_\infty \int_0^T \big(o(\sqrt{\gamma_k}) + O(|\delta x_{k,t}|^2)\big)\mathrm{d}t \leq o(\sqrt{\gamma_k}), \end{aligned} \tag{4.26}$$

where the last inequality follows from [6, Lemma 8.4]. Therefore, dividing by $\sqrt{\gamma_k}$ and passing to the limit, we obtain

$$\int_0^T \varphi_t g'(\hat{x}_t)(\bar{\xi}_t + f_1(\hat{x}_t)\bar{y}_t)\mathrm{d}t \leq 0. \tag{4.27}$$

Since φ is an arbitrary nonnegative continuous function with support in C (and C is a finite union of intervals), we deduce that

$$g'(\hat{x}_t)(\bar{\xi}_t + f_1(\hat{x}_t)\bar{y}_t) \leq 0, \quad \text{for a.a. } t \in C. \tag{4.28}$$

In particular, if $T \in C$, we get from (4.25):

$$g'(\hat{x}_T)(\bar{\xi}_T + f_1(\hat{x}_T)\bar{h}) \leq 0 \quad \text{if } T \in C. \tag{4.29}$$

Take $(\beta, \Psi, p, \mathrm{d}\mu) \in \Lambda$. By Proposition 7 and since $u_{\min} \leq \hat{u} + v_k \leq u_{\max}$, we have

$$D\ell^{\beta,\Psi}(\hat{x}_0, \hat{x}_T)(z_{k,0}, z_{k,T}) + \int_0^T g'(\hat{x}_t) z_{k,t} \mathrm{d}\mu_t = \int_0^T H_u(t) v_{k,t} \mathrm{d}t \geq 0. \tag{4.30}$$

On the other hand, a first order Taylor expansion of $\ell^{\beta,\Psi}$ and [6, Lemma 8.12] lead to

$$D\ell^{\beta,\Psi}(\hat{x}_0, \hat{x}_T)(z_{k,0}, z_{k,T}) = \ell^{\beta,\Psi}(x_{k,0}, x_{k,T}) - \ell^{\beta,\Psi}(\hat{x}_0, \hat{x}_T) + o(\sqrt{\gamma}_k). \tag{4.31}$$

Hence, by (4.30),

$$\begin{aligned} 0 &\leq \ell^{\beta,\Psi}(x_{k,0}, x_{k,T}) - \ell^{\beta,\Psi}(\hat{x}_0, \hat{x}_T) + o(\sqrt{\gamma}_k) + \int_0^T g'(\hat{x}_t) z_{k,t} \mathrm{d}\mu_t \\ &\leq \beta\phi(x_{k,0}, x_{k,T}) - \beta\phi(\hat{x}_0, \hat{x}_T) + o(\sqrt{\gamma}_k) + \int_0^T g'(\hat{x}_t) z_{k,t} \mathrm{d}\mu_t, \end{aligned} \tag{4.32}$$

where the last inequality holds since $\sum_{i=n_1+1}^{n_1+n_2} \Psi_i \Phi_i(x_{k,0}, x_{k,T}) \leq 0$. Observe now that, due to (4.15), $\beta\phi(x_{k,0}, x_{k,T}) - \beta\phi(\hat{x}_0, \hat{x}_T) \leq o(\gamma_k)$. Hence, by latter estimate and from (4.32), we deduce that

$$\frac{1}{\sqrt{\gamma_k}} \int_0^T g'(\hat{x}_t) z_{k,t} \mathrm{d}\mu_t = \int_0^T g'(\hat{x}_t)(\bar{\xi}_{k,t} + f_1(\hat{x}_t)\bar{y}_{k,t}) \mathrm{d}\mu_t \geq o(1). \tag{4.33}$$

Since $\mathrm{d}\mu_t$ has an essentially bounded density over $[0, T)$, we have that

$$\begin{aligned} 0 &\leq \liminf_{k\to\infty} \int_0^T g'(\hat{x}_t)(\bar{\xi}_{k,t} + f_1(\hat{x}_t)\bar{y}_{k,t}) \mathrm{d}\mu_t \\ &= \lim_{k\to\infty} \Big(\int_{[0,T)} g'(\hat{x}_t)(\bar{\xi}_{k,t} + f_1(\hat{x}_t)\bar{y}_{k,t}) \mathrm{d}\mu_t + g'(\hat{x}_T)(\bar{\xi}_{k,T} + f_1(\hat{x}_T)\bar{y}_{k,T})[\mu(T)] \Big) \\ &= \int_{[0,T)} g'(\hat{x}_t)(\bar{\xi}_t + f_1(\hat{x}_t)\bar{y}_t) \mathrm{d}\mu_t + g'(\hat{x}_T)(\bar{\xi}_T + f_1(\hat{x}_T)\bar{h})[\mu(T)]. \end{aligned} \tag{4.34}$$

In view of (4.28), (4.29) and the complementary condition for the state constraint (3.4)(ii), we get from (4.34) that

$$g'(\hat{x}_t)(\bar{\xi}_t + f_1(\hat{x}_t)\bar{y}_t) = 0, \quad \text{for a.a. } t \in C, \tag{4.35}$$

and

$$g'(\hat{x}_T)(\bar{\xi}_T + f_1(\hat{x}_T)\bar{h}) = 0, \quad \text{whenever } T \in C \text{ and } [\mu(T)] > 0. \tag{4.36}$$

Then, $(\bar{y}, \bar{h}, \bar{\xi})$ satisfies (3.15)(i). Furthermore, if $T \in$ and $[\mu(T)] > 0$ for some $\lambda \in \Lambda$, it holds

$$\bar{h} = -\frac{g'(\hat{x}_T)\bar{\xi}_T}{g'(\hat{x}_T)f_1(\hat{x}_T)} = \lim_{t\uparrow T} -\frac{g'(\hat{x}_t)\bar{\xi}_t}{g'(\hat{x}_t)f_1(\hat{x}_t)} = \lim_{t\uparrow T} \bar{y}(t), \tag{4.37}$$

so that (3.17)(iv) holds.

D) We shall now prove (3.16). Let $i = 1, \ldots, n_1 + n_2$, then

$$\begin{aligned} \Phi_i'(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + f_1(\hat{x}_T)\bar{h}) &= \lim_{k\to\infty} \Phi_i'(\hat{x}_0, \hat{x}_T)\frac{(\xi_{k,0}, \xi_{k,T} + f_1(\hat{x}_T)y_{k,T})}{\sqrt{\gamma_k}} \\ &= \lim_{k\to\infty} \Phi_i'(\hat{x}_0, \hat{x}_T)\frac{(z_{k,0}, z_{k,T})}{\sqrt{\gamma_k}}. \end{aligned} \tag{4.38}$$

A first order Taylor expansion of Φ_i at $(\hat{x}_0, \hat{x}_T)$ and [6, Lemma 8.12] yield

$$\Phi_i'(\hat{x}_0, \hat{x}_T)\frac{(z_{k,0}, z_{k,T})}{\sqrt{\gamma_k}} = \frac{\Phi_i(x_{k,0}, x_{k,T}) - \Phi_i(\hat{x}_0, \hat{x}_T)}{\sqrt{\gamma_k}} + o(1). \tag{4.39}$$

Thus, from (4.38)-(4.39) we get

$$\Phi_i'(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + f_1(\hat{x}_T)\bar{h}) = 0, \quad \text{for } i = 1, \ldots, n_1,$$

$$\Phi_i'(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + f_1(\hat{x}_T)\bar{h}) \leq 0, \text{ for } i = n_1 + 1, \ldots, n_1 + n_2, \text{ with } \Phi_i(\hat{x}_0, \hat{x}_T) = 0.$$

For the endpoint cost, we can obtain analogous expressions to (4.38)-(4.39), and then, from (4.15) we get

$$\phi'(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + f_1(\hat{x}_T)\bar{h}) \leq 0. \tag{4.40}$$

Hence, (3.16) is verified.

We conclude that $(\bar{y}, \bar{h}, \bar{\xi}) \in \mathcal{P}_*^2$, this is, *Fact 1* follows.

Proof of Fact 2. From equation (4.12) in Proposition 6 we obtain that

$$\Omega(y_k, h_k, \xi_k, \lambda) = \mathcal{L}(u_k, x_k, \lambda) - \mathcal{L}(\hat{u}, \hat{x}, \lambda) - \int_0^T H_u(t)v_{k,t}\mathrm{d}t + o(\gamma_k) \leq o(\gamma_k), \tag{4.41}$$

where the last inequality follows from (4.16) and since $H_u(t)v_{k,t} \geq 0$, a.e. on $[0, T]$. Hence,

$$\liminf_{k\to\infty} \Omega(y_k, h_k, \xi_k, \lambda) \leq \limsup_{k\to\infty} \Omega(y_k, h_k, \xi_k, \lambda) \leq 0. \tag{4.42}$$

Let us recall that, for each $\lambda \in \Lambda$, the mapping $\Omega(\cdot, \lambda)$ is a Legendre form in the Hilbert space $\{(y, h, \xi) \in L^2 \times \mathbb{R} \times (H^1)^n : (3.14) \text{ holds}\}$ (in view of hypothesis (iii) of the current theorem). In particular, for $\bar{\lambda} \in \Lambda$ reaching the maximum in (4.6) for the critical direction $(\bar{y}, \bar{h}, \bar{\xi})$, one has

$$\rho\gamma(\bar{y}, \bar{h}, \bar{\xi}_0) \leq \Omega(\bar{y}, \bar{h}, \bar{\xi}, \bar{\lambda}) = \liminf \Omega(\bar{y}_k, \bar{h}_k, \bar{\xi}_k, \bar{\lambda}) \leq 0, \tag{4.43}$$

where the equality holds since Ω is a Legendre form and the last inequality follows from (4.42). In view of (4.43), we get that $(\bar{y}, \bar{h}, \bar{\xi}_0) = 0$ and $\lim \Omega(\bar{y}_k, \bar{h}_k, \bar{\xi}_k, \bar{\lambda}) = 0$. Consequently, $(\bar{y}_k, \bar{h}_k, \bar{\xi}_{k,0})$ converges strongly to $(\bar{y}, \bar{h}, \bar{\xi}_0) = 0$, which is a contradiction since $(\bar{y}_k, \bar{h}_k, \bar{\xi}_{k,0})$ has unit norm in $L^2 \times \mathbb{R} \times \mathbb{R}^n$. We conclude that $(\hat{u}, \hat{x})$ is a Pontryagin minimum satisfying $\gamma-$growth in the Pontryagin sense.

Acknowledgments

We wish to thank the anonymous referees for their bibliographical advices.

This work was partially supported by the European Union under the 7th Framework Pro-gramme FP7-PEOPLE-2010-ITN Grant agreement number 264735-SADCO. The last stage of this research took place while the first author was holding a postdoctoral position at IMPA, Rio de Janeiro, with CAPES-Brazil funding.

A On second-order necessary conditions

A.1 General constraints

We will study an abstract optimization problem of the form

$$\min f(x); \quad G_E(x) = 0,\ G_I(x) \in K_I, \tag{A.1}$$

where X, Y_E, Y_I are Banach spaces, $f : X \to \mathbb{R}$, $G_E : X \to Y_E$, and $G_I : X \to Y_I$ are functions of class C^2, and K_I is a closed convex subset of Y_I with nonempty interior. The subindex E is used to refer to 'equalities' and I to 'inequalities'.

Setting

$$Y := Y_E \times Y_I, \quad G(x) := (G_E(x), G_I(x)), \quad K := \{0\}_{K_E} \times K_I, \tag{A.2}$$

we can rewrite problem (A.1) in the more compact form

$$\min f(x); \quad G(x) \in K. \tag{P_A}$$

We use $F(P_A)$ to denote the set of feasible solutions of (P_A).

Remark 9 We refer to [10] for a systematic study of problem (P_A). Here we will take advantage of the product structure (that one can find in essentially all practical applications) to introduce a *non qualified version* of second order necessary conditions specialized to the case of quasi radial directions, that extends in some sense [10, Theorem 3.50]. See Kawasaki [25] for non radial directions.

The *tangent cone* (in the sense of convex analysis) to K_I at $y \in K_I$ is defined as

$$T_{K_I}(y) := \{z \in Y_I : \mathrm{dist}(y + tz, K_I) = o(t), \ \text{ with } t \geq 0\}, \tag{A.3}$$

and the *normal cone* to K_I at $y \in K_I$ is

$$N_{K_I}(y) := \{z^* \in Y_I^* : \langle z^*, y' - y\rangle \leq 0, \text{ for all } y' \in K_I\}. \tag{A.4}$$

In what follows, we shall study a nominal feasible solution $\hat{x} \in F(P_A)$ that may satisfy or not the *qualification condition*

$$\begin{cases} \text{(i)} \ \ DG_E(\hat{x}) \text{ is onto,} \\ \text{(ii)} \text{ there exists } z \in \operatorname{Ker} DG_E(\hat{x}) \text{ such that } \ G_I(\hat{x}) + DG_I(\hat{x})z \in \operatorname{int}(K_I). \end{cases} \tag{A.5}$$

The latter condition coincides with the qualification condition in (2.21) which was introduced for the optimal control problem (P).

Remark 10 Condition (A.5) is equivalent to the Robinson qualification condition in [39]. See the discussion in [10, Section 2.3.4].

The *Lagrangian function* of problem (P_A) is defined as

$$L(x,\lambda) := \beta f(x) + \langle \lambda_E, G_E(x)\rangle + \langle \lambda_I, G_I(x)\rangle, \tag{A.6}$$

where we set $\lambda := (\beta, \lambda_E, \lambda_I) \in \mathbb{R}_+ \times Y_E^* \times Y_I^*$. Define the *set of Lagrange multipliers* associated with $x \in F(P_A)$ as

$$\Lambda(x) := \{\lambda \in \mathbb{R}_+ \times Y_E^* \times N_{K_I}(G_I(x)) : \lambda \neq 0,\ D_x L(x,\lambda) = 0\}. \tag{A.7}$$

Let $y_I \in \operatorname{int}(K_I)$, $y_I \neq G_I(\hat{x})$. We consider the following auxiliary problem, where $(x,\gamma) \in X \times \mathbb{R}$:

$$\begin{aligned} &\min_{x,\gamma} \gamma;\ f(x) - f(\hat{x}) \leq \gamma, \quad G_E(x) = 0, \quad \gamma \geq -1/2, \\ &\qquad y_I + (1+\gamma)^{-1}\big(G_I(x) - y_I\big) \in K_I. \end{aligned} \tag{AP_A}$$

Note that we recognize the idea of a *gauge function* (see e.g. [40]) in the last constraint.

Lemma 4 *Assume that $\hat{x}$ is a local solution of (P_A). Then $(\hat{x},0)$ is a local solution of (AP_A).*

Proof We easily check that $(\hat{x},0) \in F(AP_A)$. Now take $(x,\gamma) \in F(AP_A)$. Let us prove that if $-1/2 \leq \gamma < 0$, then x cannot be closed to $\hat{x}$ (in the norm of the Banach space X). Assuming that $-1/2 \leq \gamma < 0$, we get $G_E(x) = 0$, $G_I(x) \in K_I + (-\gamma)y_I \subseteq K_I$, and $f(x) < f(\hat{x})$. Since $\hat{x}$ is a local solution of (P_A), the x cannot be too closed to $\hat{x}$. The conclusion follows.

The Lagrangian function of (AP_A), in qualified form, is

$$\gamma + \beta(f(x) - f(\hat{x}) - \gamma) + \langle \lambda_E, G_E(x)\rangle + \langle \lambda_I, y_I + (1+\gamma)^{-1}(G_I(x) - y_I)\rangle. \tag{A.8}$$

or equivalently

$$L(x,\lambda) + (\beta_0 - \beta)\gamma + \big((1+\gamma)^{-1} - 1\big)\langle \lambda_I, G_I(x) - y_I\rangle. \tag{A.9}$$

Setting $\hat{\lambda} = (\beta_0, \beta, \lambda_E, \lambda_I)$, we see that the set of Lagrange multipliers of the auxiliary problem (AP_A) at $(\hat{x},0)$ is

$$\hat{\Lambda} := \left\{ \begin{array}{l} \hat{\lambda} \in \mathbb{R}_+ \times \mathbb{R}_+ \times Y_E^* \times N_{K_I}(G_I(\hat{x})) : \lambda \neq 0, \\ D_x L(\hat{x},\lambda) = 0;\ \beta + \langle \lambda_I, G_I(\hat{x}) - y_I\rangle = 1 \end{array} \right\}. \tag{A.10}$$

Proposition 8 *Suppose that* (A.5)*(i) holds. Then, the mapping*

$$(\beta, \lambda_E, \lambda_I) \mapsto \frac{(\beta + \langle \lambda_I, G_I(x) - y_I\rangle, \beta, \lambda_E, \lambda_I)}{\beta + \langle \lambda_I, G_I(x) - y_I\rangle} \tag{A.11}$$

is a bijection between $\Lambda(\hat{x})$ and $\hat{\Lambda}_1$ (recall the definition in (2.16)*).*

Proof Since (A.5)(i) holds, then we necessarily have that $(\beta, \lambda_I) \neq 0$ for all $\lambda = (\beta, \lambda_E, \lambda_I) \in \Lambda(\hat{x})$. Therefore, if $\lambda_I = 0$ then $\beta > 0$ and $\beta + \langle \lambda_I, G_I(x) - y_I\rangle > 0$. If by the contrary, $\lambda_I \neq 0$, then $\langle \lambda_I, G_I(x) - y_I\rangle > 0$ and again, $\beta + \langle \lambda_I, G_I(x) - y_I\rangle > 0$. Hence, the mapping in (A.11) is well-defined and is a bijection from $\Lambda(\hat{x})$ to $\hat{\Lambda}_1$, as we wanted to show.

Theorem 6 *Let $\hat{x}$ be a local solution of (P_A), such that $DG_E(\hat{x})$ is surjective. Then $\hat{\Lambda}_1$ is non empty and bounded.*

Proof By lemma 4, $(\hat{x},0)$ is a local solution of (AP_A). In addition the qualification condition for the latter problem at the point $(\hat{x},0)$ states as follows: there exists $(z,\delta) \in \operatorname{Ker} DG_E(\hat{x}) \times \mathbb{R}$ such that

$$\begin{gathered} Df(\hat{x})z < \delta, \quad \delta > 0, \\ G_I(\hat{x}) + DG_I(\hat{x})z - \delta(G_I(\hat{x}) - y_I) \in \operatorname{int}(K_I). \end{gathered}$$

These conditions trivially hold for $(z,\delta) = (0,1)$. Hence, in view of classical results by e.g. Robinson [39], the conclusion follows.

A.2 Second order necessary optimality conditions

Let us introduce the notation $[a,b]$ to refer to the segment $\{\rho a + (1-\rho)b;\ \text{for } \rho \in [0,1]\}$, defined for any pair of points a, b in an arbitrary vector space Z.

Definition 6 Let $y \in K$. We say that $z \in Y$ is a *radial direction* to K at y if $[y, y+\varepsilon z] \subset K$ for some $\varepsilon > 0$, and a *quasi-radial direction* if $\operatorname{dist}(y + \sigma z, K) = o(\sigma^2)$ for $\sigma > 0$.

Note that any radial direction is also quasi-radial, and both radial and quasi radial directions are tangent. With $\hat{x} \in F(P_A)$, we associate the *critical cone*

$$C(\hat{x}) := \{z \in X : Df(\hat{x})z \leq 0,\ DG_E(\hat{x})z = 0,\ DG_I(\hat{x})z \in T_K(G_I(\hat{x}))\}. \tag{A.12}$$

Definition 7 We say that $z \in C(\hat{x})$ is a *radial (quasi radial) critical direction* for problem (P_A) if $DG_I(\hat{x})z$ is a radial (quasi radial) direction to K_I at $G_I(\hat{x})$. We write $C_{QR}(\hat{x})$ for the *set of quasi radial critical directions.* The critical cone $C(\hat{x})$ is *quasi radial* if $C_{QR}(\hat{x})$ is a dense subset of $C(\hat{x})$.

It is immediate to check that $C_{QR}(\hat{x})$ is a convex cone.

We next state *primal second order necessary conditions* for the problem (P_A). Consider the following optimization problem, where $z \in X$, $w \in X$ and $\theta \in \mathbb{R}$:

$$\begin{cases} \min\limits_{w,\theta} \ \theta, \\ Df(\hat{x})w + D^2 f(\hat{x})(z,z) \leq \theta, \\ DG_E(\hat{x})w + D^2 G_E(\hat{x})(z,z) = 0, \\ DG_I(\hat{x})w + D^2 G_I(\hat{x})(z,z) - \theta(G(\hat{x}) - y_I) \in T_K(G_I(\hat{x})). \end{cases} \tag{Q_z}$$

Theorem 7 *Let $(\hat{x}, 0)$ be a local solution of (AP_A), such that $DG_E(\hat{x})$ is surjective, and let $h \in C_{QR}(\hat{x})$. Then problem (Q_z) is feasible, and has a nonnegative value.*

Proof We shall first show that (Q_z) is feasible. Since $DG_E(\hat{x})$ is surjective, there exists $w \in X$ such that $DG_E(\hat{x})w + D^2 G_E(\hat{x})(z,z) = 0$. Since $T_K(G_I(\hat{x}))$ is a cone, the last equation divided by $\theta > 0$ is equivalent to

$$\theta^{-1}\big(DG_I(\hat{x})w + D^2 G_I(\hat{x})(z,z)\big) + y_I - G(\hat{x}) \in T_K(G_I(\hat{x})). \tag{A.13}$$

Since $y_I \in \operatorname{int}(K_I)$, we have that $y_I - G(\hat{x}) \in \operatorname{int} T_K(G_I(\hat{x}))$, and therefore the last constraint of (Q_z) holds when θ is large enough. So it does the first constraint, and hence, (Q_z) is feasible.

We next have to show that we cannot have $(w, \theta_0) \in F(Q_z)$ with $\theta_0 < 0$. Let us suppose, on the contrary, that there is such a feasible solution (w, θ_0). Set $\theta := \frac{1}{2}\theta_0$. Then $Df(\hat{x})w + D^2 f(\hat{x})(z,z) < \theta$. Using (A.13) and $y_I \in \operatorname{int}(K_i)$, we can easily show that, for some $\varepsilon > 0$:

$$DG_I(\hat{x})w + D^2 G_I(\hat{x})(z,z) - \theta(G_I(\hat{x}) - y_I) + \varepsilon B \in T_K(G_I(\hat{x})). \tag{A.14}$$

Consider, for $\sigma > 0$, the path

$$x_\sigma := \hat{x} + \sigma z + \tfrac{1}{2}\sigma^2 w. \tag{A.15}$$

By a second order Taylor expansion we obtain that $G_E(x_\sigma) = o(\sigma^2)$. Since $DG_E(\hat{x})$ is onto, by Lyusternik's theorem [27], there exists a path $x'_\sigma = x_\sigma + o(\sigma^2)$, such that $G_E(x'_\sigma) = 0$. Assuming, without loss of generality, that $G_I(\hat{x}) = 0$, we get

$$G_I(x'_\sigma) = \sigma DG_I(\hat{x})z + \tfrac{1}{2}\sigma^2\left[DG_I(\hat{x})w + D^2 G_I(\hat{x})(z,z)\right] + o(\sigma^2). \tag{A.16}$$

Setting

$$\begin{cases} k_1(\sigma) := (1-\sigma)^{-1}\sigma DG_I(\hat{x})z, \\ k_2(\sigma) := \sigma\big(DG_I(\hat{x})w + D^2 G_I(\hat{x})(z,z)\big), \end{cases} \tag{A.17}$$

we can rewrite (A.16) as

$$G_I(x'_\sigma) = (1-\sigma)k_1(\sigma) + \tfrac{1}{2}\sigma k_2(\sigma) + o(\sigma^2). \tag{A.18}$$

Since z is a quasi radial critical direction, there exists $k_1'(\sigma) \in K_I$ such that

$$\sigma DG_I(\hat{x})z = k_1'(\sigma) + o(\sigma^2), \tag{A.19}$$

and so,

$$G_I(x_\sigma') \in (1-\sigma)K_I + \tfrac{1}{2}\sigma k_2(\sigma) + o(\sigma^2). \tag{A.20}$$

Using (A.14) and $G_I(\hat{x}) = 0$ we obtain

$$k_2(\sigma) + \sigma\theta y_I + \sigma\varepsilon B \in K_I. \tag{A.21}$$

Therefore, for $\sigma > 0$ small enough

$$G_I(x_\sigma') \in (1 - \tfrac{1}{2}\sigma)K_I + \tfrac{1}{2}\sigma K_I - \tfrac{1}{2}\sigma^2(\theta y_I + o(1)) \subset K_I, \tag{A.22}$$

where we have used the fact that since $0 = G_I(\bar{x}) \in K_I$, we have that (remember that $\theta < 0$): $\frac{1}{2}\sigma^2(-\theta)(y_I + \varepsilon B) \subset K_I$.

We check easily that $f(x_\sigma') < 0$, and so, we have constructed a feasible path for (AP_A), contradicting the local optimality of $(\hat{x}, 0)$.

We conclude that such a solution (w, θ_0) of (Q_z) with $\theta_0 < 0$ cannot exist and, therefore, (Q_z) has nonnegative value.

We now present dual second order necessary conditions.

Theorem 8 *Let $\hat{x}$ be a local minimum of (P_A), that satisfies the qualification condition* (A.5)*. Then, for every $z \in C_{QR}(\hat{x})$,*

$$\max_{\lambda \in \Lambda(\hat{x})} D^2_{xx} L(\hat{x}, \lambda)(z, z) \geq 0. \tag{A.23}$$

Proof Since problem (Q_z) is qualified with a finite nonnegative value, by the convex duality theory [14], its dual has a nonnegative value and a nonempty set of solutions. The Lagrangian of problem (Q_z) in qualified form ($\beta_0 = 1$) can be written as

$$D_x L(\hat{x}, \lambda)w + D^2_{xx} L(\beta, \hat{x}, \lambda)(z, z) + \big(1 - \beta + \langle \lambda_I, G(\hat{x}) - y_I \rangle\big)\theta, \tag{A.24}$$

where $\lambda = (\beta, \lambda_E, \lambda_I)$ as before, and so, the dual problem of (Q_z) can be written as

$$\operatorname*{Max}_{\lambda \in \Lambda(\hat{x})} D^2_{xx} L(\beta, \hat{x}, \lambda)(z, z); \quad \beta + \langle \lambda_I, G(\hat{x}) - y_I \rangle)\theta = 1.$$

The conclusion follows.

Remark 11 Whereas the above theorem follows from Cominetti [12] or Kawasaki [25], our proof avoids the concepts of second order tangent set and its associated calculus, used in these references. This considerably simplifies the proof.

References

1. R.A. Adams. *Sobolev spaces.* Academic Press, New York, 1975.
2. A. Aftalion and J. Bonnans. Optimization of running strategies based on anaerobic energy and variations of velocity. *SIAM J. Applied Math.*, 74(5):1615–1636, 2014.
3. A. Agrachev and Y. Sachkov. *Control theory from the geometric viewpoint.* Encyclopaedia of Mathematical Sciences, 87, Control Theory and Optimization, II. Springer-Verlag, Berlin, 2004.
4. A.A. Agrachev, G. Stefani, and P.L. Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control Optim.*, 41:991–1014, 2002.
5. M.S. Aronna. Singular solutions in optimal control: second order conditions and a shooting algorithm. Technical report, ArXiv, 2013. [Published online as arXiv:1210.7425, submitted].

6. M.S. Aronna, J.F. Bonnans, A.V. Dmitruk, and P.A. Lotito. Quadratic order conditions for bang-singular extremals. *Numer. Algebra, Control Optim., special issue dedicated to Professor Helmut Maurer on the occasion of his 65th birthday,*, 2(3):511–546, 2012.
7. A.V. Arutyunov. On necessary conditions for optimality in a problem with phase constraints. *Dokl. Akad. Nauk SSSR*, 280(5):1033–1037, 1985.
8. A.V. Arutyunov. *Optimality conditions*, volume 526 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000. Abnormal and degenerate problems, Translated from the Russian by S. A. Vakhrameev.
9. J.F. Bonnans and A. Hermant. Revisiting the analysis of optimal control problems with several state constraints. *Control Cybernet.*, 38(4A):1021–1052, 2009.
10. J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer-Verlag, New York, 2000.
11. B. Bonnard, L. Faubourg, G. Launay, and E. Trélat. Optimal control with state constraints and the space shuttle re-entry problem. *Journal of Dynamical and Control Systems*, 9(2):155–199, 2003.
12. R. Cominetti. Metric regularity, tangent sets and second order optimality conditions. *J. Appl. Math. Optim.*, 21:265–287, 1990.
13. M.R. de Pinho, M.M. Ferreira, U. Ledzewicz, and H. Schaettler. A model for cancer chemotherapy with state-space constraints. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5):e2591–e2602, 2005.
14. I. Ekeland and R. Temam. *Convex analysis and variational problems*, volume 1 of *Studies in Mathematics and its Applications*. North-Holland, Amsterdam, 1976. French edition: Analyse convexe et problèmes variationnels, Dunod, Paris, 1974.
15. U. Felgenhauer. Optimality and sensitivity for semilinear bang-bang type optimal control problems. *Int. J. Appl. Math. Comput. Sci.*, 14(4):447–454, 2004.
16. H. Frankowska and D. Tonon. Pointwise second-order necessary optimality conditions for the Mayer problem with control constraints. *SIAM J. Control Optim.*, 51(5):3814–3843, 2013.
17. R. Gabasov and F.M. Kirillova. High-order necessary conditions for optimality. *J. SIAM Control*, 10:127–168, 1972.
18. B.S. Goh. Necessary conditions for singular extremals involving multiple control variables. *J. SIAM Control*, 4:716–731, 1966.
19. B.S. Goh, G. Leitmann, and T.L. Vincent. Optimal control of a prey-predator system. *Math. Biosci.*, 19:263–286, 1974.
20. K. Graichen and N. Petit. Solving the Goddard problem with thrust and dynamic pressure constraints using saturation functions. In *17th World Congress of The International Federation of Automatic Control*, volume Proc. of the 2008 IFAC World Congress, pages 14301–14306, Seoul, 2008. IFAC.
21. M.R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.*, 1(4):525–581, 1951.
22. A. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards, Section B, Mathematical Sciences*, 49:263–265, 1952.
23. D.H. Jacobson and Speyer J.L. Necessary and sufficient conditions for optimality for singular control problems: A limit approach. *J. Math. Anal. Appl.*, 34:239–266, 1971.
24. D. Yu. Karamzin. Necessary conditions for an extremum in a control problem with phase constraints. *Zh. Vychisl. Mat. Mat. Fiz.*, 47(7):1123–1150, 2007.
25. H. Kawasaki. An envelope-like effect of infinitely many inequality constraints on second order necessary conditions for minimization problems. *Math. Program.*, 41:73–96, 1988.
26. H.J. Kelley. A second variation test for singular extremals. *AIAA Journal*, 2:1380–1382, 1964.
27. L. Lyusternik. Conditional extrema of functions. *Math. USSR-Sb*, 41:390–40, 1934.
28. K. Malanowski and H. Maurer. Sensitivity analysis for parametric control problems with control-state constraints. *Comput. Optim. Appl.*, 5(3):253–283, 1996.
29. H. Maurer. On optimal control problems with bounded state variables and control appearing linearly. *SIAM J. Control Optim.*, 15(3):345–362, 1977.
30. H. Maurer, J.-H.R. Kim, and G. Vossen. On a state-constrained control problem in optimal production and maintenance. In C. Deissenberg and R.F. Hartl, editors, *Optimal Control and Dynamic Games*, volume 7 of *Advances in Computational Management Science*, pages 289–308. Springer US, 2005.

31. H. Maurer and N.P. Osmolovskii. Second order optimality conditions for bang-bang control problems. *Control Cybernet.*, 32:555–584, 2003.
32. J.P. McDanell and W.F. Powers. Necessary conditions joining optimal singular and non-singular subarcs. *SIAM J. Control*, 9:161–173, 1971.
33. A.A. Milyutin and N. N. Osmolovskii. *Calculus of Variations and Optimal Control.* American Mathematical Society, Providence, 1998.
34. N. P. Osmolovskii and H. Maurer. Equivalence of second order optimality conditions for bang-bang control problems. I. Main results. *Control Cybernet.*, 34(3):927–950, 2005.
35. N.P. Osmolovskii and H. Maurer. *Applications to regular and bang-bang control*, volume 24 of *Advances in Design and Control.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2012. Second-order necessary and sufficient optimality conditions in calculus of variations and optimal control.
36. L. Poggiolini and M. Spadini. Strong local optimality for a bang-bang trajectory in a Mayer problem. *SIAM J. Control Optim.*, 49:140–161, 2011.
37. L. Poggiolini and G. Stefani. Bang-singular-bang extremals: sufficient optimality conditions. *J. Dyn. Control Syst.*, 17(4):469–514, 2011.
38. F. Rampazzo and R. Vinter. Degenerate optimal control problems with state constraints. *SIAM J. Control Optim.*, 39(4):989–1007 (electronic), 2000.
39. S.M. Robinson. First order conditions for general nonlinear optimization. *SIAM J. Appl. Math.*, 30:597–607, 1976.
40. R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.
41. I. B. Russak. Second-order necessary conditions for general problems with state inequality constraints. *J. Optimization Theory Appl.*, 17(112):43–92, 1975.
42. I.B. Russak. Second order necessary conditions for problems with state inequality constraints. *SIAM J. Control*, 13:372–388, 1975.
43. H. Schattler. A local field of extremals near boundary arc - interior arc junctions. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pages 945–950, Dec 2005.
44. H. Schättler. Local fields of extremals for optimal control problems with state constraints of relative degree 1. *J. Dyn. Control Syst.*, 12(4):563–599, 2006.
45. H. Seywald and E.M. Cliff. Goddard problem in presence of a dynamic pressure limit. *Journal of Guidance, Control, and Dynamics*, 16(4):776–781, 1993.