

# Interior point methods in optimal control problems of affine systems: Convergence results and solving algorithms

Paul Malisani\*

## Abstract

This paper presents an interior point method for pure-state and mixed-constrained optimal control problems for dynamics, mixed constraints, and cost function all affine in the control variable. This method relies on resolving a sequence of two-point boundary value problems of differential and algebraic equations. This paper establishes a convergence result for primal and dual variables of the optimal control problem. A primal and a primal-dual solving algorithm are presented, and a challenging numerical example is treated for illustration. Article accepted for publication at SIAM SICON.

## 1 Introduction

This paper deals with Optimal Control Problems (OCPs) with pure state and mixed constraints. These arise naturally in numerous engineering problems such as aerospace [1, 7], control of hybrid electric vehicles [18], or innate immune response [25], among other examples. Unfortunately, these problems are difficult to solve [11, 6, 17]. This difficulty mainly stems from the pure state constraints. Indeed, as shown in [11, 17], the first-order optimality conditions of these problems imply that the adjoint state of Pontryagin can be discontinuous when state constraints switch from active to inactive and vice-versa. To handle these problems, three main approaches are found in the literature. The first is a discretization-based approach that treats an OCP as a finite-dimensional optimization problem [1]. This approach is, in practice, the most widely used and the easiest to implement. However, these methods can be computationally slow and might lack precision. The second approach consists of computing the optimal trajectory without constraints and, step-by-step, computing the trajectory's structure, that is to say, the sequence of constrained and unconstrained arcs along the trajectory. Therefore, these methods assume that the optimal trajectory contains finitely many constrained arcs. Unfortunately, this is not always the case, and trajectory structures can be much more complicated, even for simple cases [21]. These methods are known as continuation or homotopy methods [6, 5, 2]. Finally, the third approach consists in adapting Interior Point Methods (IPMs), widely studied and successfully implemented in software for numerical optimization [19, 27, 28], to state and input-constrained OCPs [15, 4, 9, 16, 26]. This approach entails minimizing an augmented cost function defined as the sum of the original cost function and so-called penalty functions, which have a diverging asymptotic behavior in the vicinity of the constraints. IPMs then define a sequence of OCPs indexed by a sequence of decreasing positive parameters converging to zero. These parameters serve as weights for the penalty functions in the augmented cost. Each OCP of this sequence is then solved as a constraints-free problem whose solution strictly satisfies the constraints and asymptotically converges to the solution of the original problem. From a practical viewpoint, these methods are appealing since off-the-shelves OCP solvers such as [24] can be used. However, adapting IPMs to

---

\*IFP Energies nouvelles, Applied Mathematics Department, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France (paul.malisani@ifpen.fr).

OCPs is not straightforward and has yet to be wholly performed. Indeed, to be complete, this adaptation requires proving two things. Firstly, the optimal trajectories of the penalized problem are interior, i.e., strictly satisfy the constraints. Secondly, it requires establishing the convergence of the method to a point satisfying the first-order optimality conditions [11, 17, 6]. In other words, to prove the convergence of primal variables (namely state and control variables) and the dual variables (namely the adjoint state of Pontryagin and the inequality constraints multipliers). In [15, 9], the interiority of solutions is left as an assumption, and convergence is proved only for state and control variables. In addition, the authors assume the uniqueness of the optimal solution and strong convexity of the optimal control problem allowing them to prove convergence of the global minimum only. In [4], the authors prove the interiority of solutions and the convergence of control, state, and adjoint state variables in the case of linear control constraints for dynamical systems affine with respect to the control. To establish the proof, the authors assume the uniqueness of the optimal solution and the strong convexity of the optimal control problem. In [26], a primal-dual IPM for OCPs is presented, and convergence of control, state, adjoint state, and constraint multipliers is proved only for control-constrained problems using a strong Legendre-Clebsch type condition, which is a sufficient condition for strong convexity [3, Theorem 5.4]. In [16], the authors exhibit sufficient conditions on the penalty functions guaranteeing the interiority of solutions for state and control-constrained problems. However, the convergence of the method is only proved for control and state variables, again assuming the uniqueness of the optimal solution and strong convexity of the problem.

This paper’s contribution is proving the convergence of primal variables and constraint multipliers for pure-state and mixed constraints. In addition, the proof of convergence does not rely on the optimal solution’s uniqueness or the problem’s strong convexity assumptions. As in [4, 9], the result is established for non-linear systems and mixed constraints affine with respect to the control. To do so, we prove that using logarithmic penalties guarantees the interiority of any locally optimal solution. In addition, we prove that the derivative of the penalty functions associated with any locally optimal solutions satisfies a uniform boundedness property. Using some standard compactness argument, we can prove weak convergence of the control and strong convergence of the state. In turn, this allows us to prove weak or strong convergence (depending on the case) of the derivatives of the penalty to the constraints multipliers. Then, the strong convergence of the adjoint state stems from the convergence of the state, control, and constraint multipliers. Finally, strong convergence of state and adjoint state allows proving strong convergence of the control variable. Finally, this paper provides a primal and a primal-dual solving algorithm based on Two Point Boundary Value Problems solver.

The paper is organized as follows. Section 2 contains the problem statement, the main assumptions, and the paper’s main results. In section 3, some preliminary technical results are recalled. In section 4, we prove both the uniform boundedness properties of the derivatives of the penalty functions and the interiority of penalized trajectories when using logarithmic penalties. In section 5, we prove that the solutions of the penalized optimal control problem converge to a solution of the classical first-order optimality conditions of constrained optimal control problems. Section 6 presents a primal and a primal-dual solving algorithm. Finally, in section 7, the Robbins problem [21] is treated using the primal and the primal-dual interior point algorithm. In addition, source codes for the Robbins and the Goddard problems [23] are available at [https://ifpen-gitlab.appcollaboratif.fr/detocs/ipm\\_ocp](https://ifpen-gitlab.appcollaboratif.fr/detocs/ipm_ocp).

**Notations:** We denote  $\mathbb{R}_-$  (resp.  $\mathbb{R}_+$ ) the set of non-positive (resp. non-negative) real numbers. We denote  $\mathbb{N}_*$  (resp.  $\mathbb{R}_*$ ) the set of non-zero natural integers (resp. real numbers). Given  $p \in [1, +\infty]$ , we denote  $L^p(A; B)$  (or  $L^p$ ) the Lebesgue spaces of functions from  $A$  to  $B$  and we denote  $\|\cdot\|_{L^p}$  the corresponding  $p$ -norm. In addition, we also denote  $\text{meas}(\cdot)$  the Lebesgue measure on  $\mathbb{R}$ . Given  $p \in [1, +\infty]$ , we denote  $W^{1,p}(A; B)$  the Sobolev space of measurable functions from  $A$  to  $B$  with weak derivative in  $L^p(A; B)$ . Given  $n \in [0, +\infty]$ , we denote  $C^n(A; B)$  (or  $C^n$ ) the set of  $n$ -times continuously differentiable functions from  $A$  to  $B$ . We denote  $BV(A)$ , the set of functions with bounded variations from  $A$  to  $\mathbb{R}$ . We also denote  $\mathcal{M}(A)$  the set of Radon measures on  $A \subset \mathbb{R}$ . The topological dual of a topological vector space  $E$  is denoted  $E^*$ . Given a topological vector space  $E$ , we denote  $\sigma(E, E^*)$  the weak

topology on  $E$  and  $\sigma(E^*, E)$  the weak  $*$  topology on  $E^*$ . Let  $x_n, x \in E$ , we denote  $x_n \rightharpoonup x$  the weak convergence in  $\sigma(E, E^*)$  and let  $y_n, y \in E^*$ , we denote  $y_n \overset{*}{\rightharpoonup} y$  the weak  $*$  convergence in  $\sigma(E^*, E)$ . For  $x^* \in E^*$  and  $x \in E$ , we denote  $\langle x^*, x \rangle$ , the duality product. Given  $f \in C^{k \geq 1}(\mathbb{R}^n; \mathbb{R})$  we denote  $f'(\cdot)$  the gradient of the function. Given  $f \in C^{k \geq 1}(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^p)$ , we denote  $f'_x(x, y) := \frac{\partial f}{\partial x}(x, y) \in \mathbb{R}^{p \times n}$  (resp.  $f'_y(x, y) := \frac{\partial f}{\partial y}(x, y) \in \mathbb{R}^{p \times m}$ ) and we denote  $f'_{i,x} := \frac{\partial f_i}{\partial x}(x, y)$  (resp.  $f'_{i,y} := \frac{\partial f_i}{\partial y}(x, y)$ ). Given  $f \in C^{k \geq 1}(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R})$ , we denote  $f'_{x,i}(x, y) := (f'_x(x, y))_i$  (resp.  $f'_{y,i}(x, y) := (f'_y(x, y))_i$ ). We also denote  $f''_{xy}(x, y) := \frac{\partial^2 f}{\partial y \partial x}(x, y)$ . Let  $G : X \mapsto Y$  with  $X, Y$  Banach spaces, we denote  $DG(x)$  the derivative of the mapping  $G$  at point  $x \in X$ . The finite dimensional euclidean norm is denoted  $\|\cdot\|$  and the scalar (resp. matrix) product between  $x, y \in \mathbb{R}^n$  (resp.  $x \in \mathbb{R}^{m \times n}, y \in \mathbb{R}^n$ ) is denoted  $x.y$ . Given a set  $E$ , we denote  $|E|$  its cardinal. We also denote  $B_N(x, r)$  the closed ball of radius  $r$  centered in  $x$  for the topology induced by norm  $N$ . We denote  $x[u, x^0]$  (or  $x[u]$  if  $x^0$  is fixed) the solution of the differential equations  $\dot{x} = f(x, u)$  with initial condition  $x^0$ . Finally, we denote  $\text{const}(\cdot)$  a positive finite constant depending on the parameters in argument.

## 2 Problem statement and main result

### 2.1 Optimal control problem

The problem we are interested in consists of finding a solution  $(x, u)$  of the following Constrained Optimal Control Problem (COCP)

$$\min_{(u, x) \in \mathbf{U} \times \mathbf{X}} J(x, u) := \varphi(x(T)) + \int_0^T \ell_1(x(t)) + \ell_2(x(t)).u(t) dt \quad (1a)$$

$$:= \varphi(x(T)) + \int_0^T \ell(x(t), u(t)) dt \quad (1b)$$

$$\dot{x}(t) = f_1(x(t)) + f_2(x(t)).u(t) := f(x(t), u(t)) \quad (1c)$$

$$0 = h(x(0), x(T)) \quad (1d)$$

$$0 \geq g(x(t)); \forall t \quad (1e)$$

$$0 \geq a(x(t)).u(t) + b(x(t)) := c(x(t), u(t)) \text{ a.e.} \quad (1f)$$

$$\mathbf{U} := L^\infty([0, T]; \mathbb{R}^m) \quad (1g)$$

$$\mathbf{X} := W^{1, \infty}([0, T]; \mathbb{R}^n) \quad (1h)$$

where the time horizon  $T > 0$  is fixed.

**Definition 1.** We denote  $\mathbf{V}^{\text{ad}} \subset \mathbf{U} \times \mathbb{R}^n$  the set of admissible controls and initial conditions as follows

$$\mathbf{V}^{\text{ad}} := \{(u, x^0) \in \mathbf{U} \times \mathbb{R}^n \text{ s.t. eqs. (1c) to (1f) holds}\} \quad (2)$$

The set  $\mathbf{V}^{\text{ad}}$  is endowed with the following norm

$$\|(u, x^0)\|_{\mathbf{V}^{\text{ad}}} := \|u\|_{L^1} + \|x^0\| \quad (3)$$

And we denote  $\mathring{\mathbf{V}}^{\text{ad}}(n)$  the following set

$$\mathring{\mathbf{V}}^{\text{ad}}(n) := \left\{ (u, x^0) \in \mathbf{V}^{\text{ad}} \text{ s.t. } \begin{cases} g(x[u, x^0](t)) < 0, \forall t \\ \text{ess sup}_t c(x[u, x^0], u) \leq -\frac{1}{n} \end{cases} \right\} \quad (4)$$

## 2.2 Main assumptions and technical definitions

**Assumption 1.** The functions  $\ell : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ ,  $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ ,  $g : \mathbb{R}^n \mapsto \mathbb{R}^{n_g}$ ,  $c : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{n_c}$  are at least twice continuously differentiable.

**Assumption 2.** Any locally optimal solution  $(x[\bar{u}, \bar{x}^0], \bar{u})$  such that  $(\bar{u}, \bar{x}^0) \in \mathbb{V}^{\text{ad}}$  satisfies the following interiority accessibility assumption

$$(\bar{u}, \bar{x}^0) \in \text{cl}_{\|\cdot\|_{\mathbb{V}^{\text{ad}}}} \left\{ \liminf_n \mathring{\mathbb{V}}^{\text{ad}}(n) \right\} := \mathbb{V}_\infty^{\text{ad}} \quad (5)$$

where  $\text{cl}_{\|\cdot\|_{\mathbb{V}^{\text{ad}}}}$  stands for the closure in the  $\|\cdot\|_{\mathbb{V}^{\text{ad}}}$ -topology.

**Assumption 3.** The set of admissible initial-final states  $h^{-1}(\{0\}) \subset \mathbb{R}^n \times \mathbb{R}^n$  from eq. (1d) is closed and bounded.

**Assumption 4.** There exists  $R_v \in (0, +\infty)$  such that

$$\|(u, x^0)\|_{\mathbb{V}^{\text{ad}}} \leq R_v, \quad \forall (u, x^0) \in \mathbb{V}^{\text{ad}} \quad (6)$$

and for all  $R_v \in (0, +\infty)$ , there exists  $R_x \in (0, +\infty)$  such that

$$\|x[u, x^0]\|_{L^\infty} \leq R_x, \quad \forall \|(u, x^0)\|_{\mathbb{V}^{\text{ad}}} \leq R_v \quad (7)$$

**Definition 2** (Set of near state-saturated times and near-saturated indices). For all  $(u, x^0) \in \mathbb{V}^{\text{ad}}$  from definition 1 and  $\forall \delta \geq 0$  we define the set of near state-saturated times (resp. mixed-saturated times), denoted  $S_{u, x^0}^g$  (resp.  $S_{u, x^0}^c$ ), as follows

$$S_{u, x^0}^g(\delta) := \left\{ t \in [0, T] \text{ s.t. } \max_i g_i(x[u, x^0](t)) \geq -\delta \right\} \quad (8)$$

$$S_{u, x^0}^c(n) := \left\{ t \in [0, T] \text{ s.t. } \max_i c_i(x[u, x^0](t), u(t)) \geq -\frac{1}{n} \right\} \quad (9)$$

In addition, we define the set of near state-saturated indices (resp. mixed-saturated indices), denoted  $I_{u, x^0}^g$  (resp.  $I_{u, x^0}^c$ ), as follows

$$I_{u, x^0}^g(t, \delta) := \left\{ i \in \{1, \dots, n_g\} \text{ s.t. } g_i(x[u, x^0](t)) \geq -\delta \right\} \quad (10)$$

$$I_{u, x^0}^c(t, n) := \left\{ i \in \{1, \dots, n_c\} \text{ s.t. } c_i(x[u, x^0](t), u(t)) \geq -\frac{1}{n} \right\} \quad (11)$$

**Assumption 5.** For all  $(u, x^0) \in \mathbb{V}^{\text{ad}}$ , the mixed constraints eq. (1f) satisfy the following qualification condition. There exists  $\gamma > 0$  and  $n \in \mathbb{N}_*$  such that

$$\gamma \|\xi\| \leq \left\| c'_{I_{u, x^0}^c(t, n), u}(x[u, x^0](t), u(t))^\top \cdot \xi \right\|, \quad \forall \xi \in \mathbb{R}^{|I_{u, x^0}^c(t, n)|}, \text{ a.a. } t \in [0, T] \quad (12)$$

**Assumption 6.** The set of singular multipliers for problem eq. (1) is empty.

**Remark 1.** Sufficient conditions on pure-state and mixed constraints such that assumption 6 holds are given in [6]. In addition, these assumptions guarantee the existence of optimal solutions of problem eq. (1) (see [10]).

**Definition 3** (State-constraint measure). For all  $(u, x^0) \in B_{L^\infty}(0, R_u) \times B_{\|\cdot\|}(0, R_x)$  and for all  $E \subset \mathbb{R}$ , we denote  $m[u, x^0, g_i]$  the push-forward  $g_i$ -measure of  $E$  defined as follows

$$m[u, x^0, g_i](E) := \text{meas} \left( (g_i \circ x[u, x^0])^{-1}(E) \right) \quad (13)$$

### 2.3 First-order necessary conditions of stationarity

In this section, the first order necessary conditions of optimality for Problem eq. (1). To do so, let us introduce the infamous pre-Hamiltonian.

**Definition 4** (pre-Hamiltonian). *The pre-Hamiltonian  $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}$  of the optimal control problem is defined by*

$$H(x, u, p) := \ell(x, u) + p.f(x, u) \quad (14)$$

**Definition 5** (Stationary point). *The trajectory  $(\bar{x}, \bar{u})$  with associated multipliers  $(\bar{p}, \bar{\mu}, \bar{\nu}, \bar{\lambda}) \in \text{BV}([0, T]; \mathbb{R})^n \times \mathcal{M}([0, T])^{n_g} \times L^\infty([0, T]; \mathbb{R}^{n_c}) \times \mathbb{R}^{n_h}$ , is a stationary point for Problem (1) if it satisfies*

$$\dot{\bar{x}}(t) = f(\bar{x}(t), \bar{u}(t)) \quad (15a)$$

$$-d\bar{p}(t) = \left[ H'_x(\bar{x}(t), \bar{u}(t), \bar{p}(t)) + \sum_{i=1}^{n_c} c'_{i,x}(\bar{x}(t), \bar{u}(t))\bar{\nu}_i(t) \right] dt + \sum_{i=1}^{n_g} g'_i(\bar{x}(t))d\bar{\mu}_i(t) \quad (15b)$$

$$0 = H'_u(\bar{x}(t), \bar{u}(t), \bar{p}(t)) + \sum_{i=1}^{n_c} c'_{i,u}(\bar{x}(t), \bar{u}(t))\bar{\nu}_i(t) \quad (15c)$$

$$0 = h(\bar{x}(0), \bar{x}(T)) \quad (15d)$$

$$0 = \bar{p}(0) + h'_{x(0)}(\bar{x}(0), \bar{x}(T))^\top \bar{\lambda} \quad (15e)$$

$$0 = \bar{p}(T) - \varphi'(\bar{x}(T)) - h'_{x(T)}(\bar{x}(0), \bar{x}(T))^\top \bar{\lambda} \quad (15f)$$

$$0 = \int_0^T g_i(\bar{x}(t))d\bar{\mu}_i(t), \quad i = 1, \dots, n_g \quad (15g)$$

$$0 = \int_0^T c_i(\bar{x}(t), \bar{u}(t))\bar{\nu}_i(t)dt, \quad i = 1, \dots, n_c \quad (15h)$$

$$0 = \bar{\lambda}^\top .h(\bar{x}(0), \bar{x}(T)) \quad (15i)$$

$$0 \leq d\bar{\mu}_i(t), \quad i = 1, \dots, n_g \quad (15j)$$

$$0 \leq \bar{\nu}_i(t), \quad i = 1, \dots, n_c \quad (15k)$$

$$0 = \bar{\mu}_i(T), \quad i = 1, \dots, n_g \quad (15l)$$

It is a well-established results [17, 6] that any local solution of problem eq. (1) is a stationary point as defined in definition 5. Unfortunately, solving Problem eq. (15) is a difficult task. Indeed, the dual variable  $d\mu$  associated with the state constraints appearing in eqs. (15b), (15g), (15j) and (15l) are Radon measures, therefore, in full generality, they can be decomposed in an absolutely continuous measure with respect to the Lebesgue measure, a discrete and finally a singular part. Computing these measures' discrete and singular parts can be dramatically complex.

### 2.4 Penalized Optimal Control Problem (POCP)

To solve problem eq. (1), we use an interior point method based on log-barrier functions defined as follows

**Definition 6** (log-barrier function). *The log-barrier function  $\psi : \mathbb{R} \mapsto \mathbb{R}$  is defined as follows*

$$\psi(x) := \begin{cases} -\log(-x) & \forall x < 0 \\ +\infty & \text{otherwise} \end{cases} \quad (16)$$

The corresponding penalized optimal control problem is defined as follows

$$\min_{(x,u) \in X \times U} J_\epsilon(x, u) := J(x, u) + \epsilon \int_0^T \left[ \sum_{i=1}^{n_g} \psi \circ g_i(x(t)) + \sum_{i=1}^{n_c} \psi \circ c_i(x(t), u(t)) \right] dt \quad (17a)$$

$$\dot{x}(t) = f(x(t), u(t)) \quad (17b)$$

$$0 = h(x(0), x(T)) \quad (17c)$$

The pre-Hamiltonian associated with this penalized problem is defined here after

**Definition 7** (Penalized pre-Hamiltonian). *The penalized pre-Hamiltonian  $H^\psi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}$  of POCP eq. (17) is defined by*

$$H^\psi(x, u, p, \epsilon) := H(x, u, p) + \epsilon \left( \sum_{i=1}^{n_g} \psi(g_i(x)) + \sum_{i=1}^{n_c} \psi(c_i(x, u)) \right) \quad (18)$$

**Definition 8** (Penalized stationary point). *The trajectory  $(\bar{x}_\epsilon, \bar{u}_\epsilon)$  with associated multipliers  $(\bar{p}_\epsilon, \bar{\lambda}_\epsilon) \in W^{1,1}([0, T]; \mathbb{R}^n) \times \mathbb{R}^{n_h}$ , is a penalized stationary point for Problem (17) if it satisfies*

$$\dot{\bar{x}}_\epsilon(t) = f(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t)) \quad (19a)$$

$$\dot{\bar{p}}_\epsilon(t) = -H_x^{\psi'}(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t), \bar{p}_\epsilon(t), \epsilon) \quad (19b)$$

$$0 = H_u^{\psi'}(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t), \bar{p}_\epsilon(t), \epsilon) \quad (19c)$$

$$0 = h(\bar{x}_\epsilon(0), \bar{x}_\epsilon(T)) \quad (19d)$$

$$0 = \bar{p}_\epsilon(0) + h'_{x(0)}(\bar{x}_\epsilon(0), \bar{x}_\epsilon(T))^\top \cdot \bar{\lambda}_\epsilon \quad (19e)$$

$$0 = \bar{p}_\epsilon(T) - \varphi'(\bar{x}_\epsilon(T)) - h'_{x(T)}(\bar{x}_\epsilon(0), \bar{x}_\epsilon(T))^\top \cdot \bar{\lambda}_\epsilon \quad (19f)$$

**Remark 2.** *Handling mixed constraints using an interior point method rather than a Pontryagin minimization has two main advantages. Firstly, it allows for using off-the-shelves index-1 BVPDAE solvers. Secondly, it also allows for implementing primal-dual methods in optimal control as described in [26], which are numerically efficient as will be illustrated later in section 7 and in the numerical examples available at <https://ifpen-gitlab.appcollaboratif.fr/detoc>*

## 2.5 Contribution of the paper

The main contribution of the paper is a convergence theorem for interior point methods optimal control problem with logarithmic penalty functions.

**Theorem 1.** *Let  $(\epsilon_n)$  be a sequence of decreasing positive parameters with  $\epsilon_n \rightarrow 0$ . The associated sequence of penalized stationary points  $(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}, \bar{p}_{\epsilon_n}, \bar{\lambda}_{\epsilon_n})_n$  as defined in definition 8 contains a subsequence converging to a*

stationary point  $(\bar{x}, \bar{u}, \bar{p}, \bar{\mu}, \bar{\nu}, \bar{\lambda})$  of the original problem as defined in definition 5 as follows

$$\left\| \bar{u}_{\epsilon_{n_k}} - \bar{u} \right\|_{L^1} \rightarrow 0 \quad (20a)$$

$$\left\| \bar{x}_{\epsilon_{n_k}} - \bar{x} \right\|_{L^\infty} \rightarrow 0 \quad (20b)$$

$$\left\| h(\bar{x}_{\epsilon_{n_k}}(0), \bar{x}_{\epsilon_{n_k}}(T)) - h(\bar{x}(0), \bar{x}(T)) \right\| \rightarrow 0 \quad (20c)$$

$$\left| J_{\epsilon_{n_k}}(\bar{x}_{\epsilon_{n_k}}, \bar{u}_{\epsilon_{n_k}}) - J(\bar{x}, \bar{u}) \right| \rightarrow 0 \quad (20d)$$

$$\left\| \bar{\lambda}_{\epsilon_{n_k}} - \bar{\lambda} \right\| \rightarrow 0 \quad (20e)$$

$$\left\| \bar{p}_{\epsilon_{n_k}} - \bar{p} \right\|_{L^1} \rightarrow 0 \quad (20f)$$

$$\epsilon_{n_k} \psi' \circ c_i(\bar{x}_{\epsilon_{n_k}}, \bar{u}_{\epsilon_{n_k}}) \stackrel{*}{\rightharpoonup} \bar{\nu}_i, \quad i = 1, \dots, n_c \quad (20g)$$

$$\epsilon_{n_k} \psi' \circ g_i(\bar{x}_{\epsilon_{n_k}}) dt \stackrel{*}{\rightharpoonup} d\bar{\mu}_i, \quad i = 1, \dots, n_g \quad (20h)$$

### 3 Preliminary results

This section gathers useful definitions and preliminary results, which will be recurrently used throughout the paper. The proofs of these results are given in appendix A.

**Proposition 1** (State Lipschitz continuity). *Let  $u_1, u_2 \in U$  and  $x_1^0, x_2^0 \in \mathbb{R}^n$ , there exists  $\text{const}(f) < +\infty$  such that*

$$\|x[u_1, x_1^0] - x[u_2, x_2^0]\|_{L^\infty} \leq \text{const}(f)(\|u_1 - u_2\|_{L^1} + \|x_1^0 - x_2^0\|) \quad (21)$$

*Proof.* See appendix A.1 □

**Proposition 2.** *Let  $(u_n, x_n^0)_n$  converging to  $(\bar{u}, \bar{x}^0)$  in the weak topology  $\sigma(L^2 \times \mathbb{R}^n, L^2 \times \mathbb{R}^n)$  then we have*

$$\|x[u_n, x_n^0] - x[\bar{u}, \bar{x}^0]\|_{L^\infty} \rightarrow 0 \quad (22)$$

and for all  $\alpha \in C^0(\mathbb{R}^n; \mathbb{R}^{p \times m})$ ,  $\beta \in C^0(\mathbb{R}^n; \mathbb{R}^p)$  we have

$$\alpha(x[u_n, x_n^0]) \cdot u_n + \beta(x[u_n, x_n^0]) \stackrel{*}{\rightharpoonup} \alpha(x[\bar{u}, \bar{x}^0]) \cdot \bar{u} + \beta(x[\bar{u}, \bar{x}^0]) \quad (23)$$

*Proof.* See appendix A.2 □

**Remark 3.** *The affine property of the dynamics and the mixed constraints eqs. (1c) and (1f) are crucial in the proof of proposition 2. Indeed, if the dynamics, for example, is not affine with respect to the control, the sequence  $(x[u_n])_n$  uniformly converges to a limit which can be different from  $x[\bar{u}]$ . Take, for example,  $\dot{x}(t) = u(t)^2$  and  $u_n := \sin(nt)$ . Then  $u_n \rightharpoonup \bar{u} = 0$  and  $x_n(t) \rightarrow t/2 \neq x[\bar{u}](t)$ .*

**Proposition 3.** *For all  $u \in U$  satisfying assumption 4, for all  $x^0 \in \mathbb{R}^n$  bounded, let  $E \subseteq g_i \circ x[u, x^0]([0, T]) \subset \mathbb{R}$  be a Lebesgue-measurable set, the state-constraint measure from definition 3 is lower bounded as follows*

$$m[u, x^0, g_i](E) \geq \text{const}(f, g) \text{meas}(E) \quad (24)$$

*Proof.* See appendix A.3 □

**Proposition 4.** For all  $\delta > 0$ ,  $\exists(G_\delta, C_\delta) > 0$  such that,  $\forall(u, x^0) \in V_\infty^{\text{ad}}$ ,  $\exists(v, y^0) \in B_{\|\cdot\|_{V^{\text{ad}}}}((u, x^0), \delta) \cap V^{\text{ad}}$  satisfying the following condition

$$\sup_t g_i(x[v, y^0](t)) \leq -2G_\delta, \quad i = 1, \dots, n_g \quad (25)$$

$$\text{ess sup}_t c_i(x[v, y^0](t), v(t)) \leq -2C_\delta, \quad i = 1, \dots, n_c \quad (26)$$

and we also have

$$g_i(x[v, y^0](t)) \leq g_i(x[u, x^0](t)) - G_\delta, \quad \forall t \in S_{u, x^0}^g(G_\delta), \quad i = 1, \dots, n_g \quad (27)$$

$$c_i(x[v, y^0](t), v(t)) \leq c_i(x[u, x^0](t), u(t)) - C_\delta, \quad \text{a.a. } t \in S_{u, x^0}^c(C_\delta), \quad i = 1, \dots, n_c \quad (28)$$

*Proof.* See appendix A.4. □

## 4 Uniform boundedness and interiority analysis of penalized optimal solutions

### 4.1 State constraints analysis

In the following, we present an interior point optimal control problem to handle pure state constraints, which writes:

$$\min_{(x, u)} J_\epsilon^1(x, u) = \varphi(x(T)) + \int_0^T \left[ \ell(x(t), u(t)) + \epsilon \sum_{i=1}^{n_g} \psi(g_i(x(t))) \right] dt \quad (29a)$$

$$\dot{x}(t) = f(x(t), u(t)) \quad (29b)$$

$$h(x(0), x(T)) = 0 \quad (29c)$$

$$c(x(t), u(t)) \leq 0 \quad (29d)$$

**Lemma 1.**  $\forall \epsilon > 0$ , any locally optimal solution  $(x[u_\epsilon, x_\epsilon^0], u_\epsilon)$  of Problem eq. (29) satisfies

$$g_i(x[u_\epsilon, x_\epsilon^0](t)) < 0, \quad \forall t \in [0, T], \quad i = 1, \dots, n_g \quad (30)$$

and  $\exists K_g < +\infty$  such that  $\forall \epsilon \in (0, \epsilon_0)$  we have

$$\|\epsilon \psi' \circ g_i(x[u_\epsilon, x_\epsilon^0])\|_{L^1} \leq K_g, \quad i = 1, \dots, n_g \quad (31)$$

*Proof.* It is sufficient to prove for  $n_g = 1$ , that is to say, for just one state constraint. Assume  $(u_\epsilon, x_\epsilon^0) \in V_\infty^{\text{ad}}$  is a local optimal solution satisfying

$$\sup_t g(x[u_\epsilon, x_\epsilon^0](t)) = 0 \quad (32)$$

From proposition 4,  $\forall \delta > 0$ ,  $\exists(v_\delta, x_\delta^0) \in B_{\|\cdot\|_{V^{\text{ad}}}}((u_\epsilon, x_\epsilon^0), \delta) \cap \overset{\circ}{V}^{\text{ad}}$  and  $G_\delta > 0$  such that

$$g(x[v_\delta, x_\delta^0](t)) \leq -2G_\delta, \quad \forall t \in [0, T] \quad (33a)$$

$$g(x[v_\delta, x_\delta^0](t)) \leq g(x[u_\epsilon, x_\epsilon^0](t)) - G_\delta, \quad \forall t \in S_{u_\epsilon, x_\epsilon^0}^g(G_\delta) \quad (33b)$$



with  $S_{u_\epsilon, x_\epsilon^0}^g(G_\delta) \neq \emptyset$ . In the following, to alleviate the notations, we denote

$$z_\delta := (v_\delta, x_\delta^0) \quad (34)$$

$$z_\epsilon := (u_\epsilon, x_\epsilon^0) \quad (35)$$

$$\Delta z := z_\delta - z_\epsilon \quad (36)$$

$$\Delta x := x[z_\delta] - x[z_\epsilon] \quad (37)$$

$$\Delta g := g(x[z_\delta]) - g(x[z_\epsilon]) \quad (38)$$

Now, one can exhibit an upper-bound on the difference  $J_\epsilon^1(x[z_\delta], v_\delta) - J_\epsilon^1(x[z_\epsilon], u_\epsilon)$  as follows

$$J_\epsilon^1(x[z_\delta], v_\delta) - J_\epsilon^1(x[z_\epsilon], u_\epsilon) = \Delta_1 + \epsilon \Delta_2 \quad (39)$$

where

$$\Delta_1 := \varphi(x[z_\delta](T)) - \varphi(x[z_\epsilon](T)) + \int_0^T [\ell(x[z_\delta](t), v_\delta(t)) - \ell(x[z_\epsilon](t), u_\epsilon(t))] dt \quad (40)$$

$$\Delta_2 := \int_0^T [\psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t))] dt \quad (41)$$

Now, let us upper-bound  $\Delta_1$

$$\Delta_1 \leq \int_0^T \text{const}(\ell) (\|x[z_\delta](t) - x[z_\epsilon](t)\| + \|v_\delta(t) - u_\epsilon(t)\|) dt + \text{const}(\varphi) \|\Delta x\|_{L^\infty} \quad (42)$$

$$\leq \text{const}(\ell, T, \varphi) \|\Delta x\|_{L^\infty} + \text{const}(\ell) \|v_\delta - u_\epsilon\|_{L^1} \leq \text{const}(\ell, f, g, \varphi, T, R_v, R_x) \quad (43)$$

Now, let us upper-bound  $\Delta_2$  from eq. (41). To do so, let us introduce the following useful subsets of  $[0; T]$

$$E_1 := (g \circ x[z_\epsilon])^{-1}((-\infty, -G_\delta]) \quad (44)$$

$$E_2(\rho) := (g \circ x[z_\epsilon])^{-1}((-G_\delta, -\rho]) \quad (45)$$

$$E_3(\rho) := (g \circ x[z_\epsilon])^{-1}((-G_\delta, -\rho)) \quad (46)$$

Given eq. (32), for all  $\rho \in [0, G_\delta]$ , these sets are not empty and  $\forall t \notin (E_1 \cup E_2(\rho))$  we have  $\psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) < 0$  which yields

$$\Delta_2 \leq \int_{E_1} \psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) dt + \int_{E_2(\rho)} \psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) dt \quad (47)$$

By convexity of the log penalty, i.e.  $\psi$ , we have

$$\int_{E_1} \psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) dt \leq \int_{E_1} \psi'(G_\delta) \|\Delta g\|_{L^\infty} dt := \text{const}(T, f, g, G_\delta) \quad (48)$$

In addition,  $\forall t \in E_2(\rho)$ , we have

$$\int_{E_2(\rho)} \psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) dt = \int_{E_2(\rho)} \left( \int_0^1 \psi'(g(x[z_\epsilon](t) + s\Delta g(t))) ds \right) \Delta g(t) dt \quad (49)$$

Since  $\forall t \in E_2(\rho)$ ,  $\Delta g(t) < -G_\delta$ , we also have

$$\int_{E_2(\rho)} \psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) dt \leq -G_\delta \int_{E_2(\rho)} \left( \int_0^1 \psi'(g(x[z_\epsilon](t)) + s\Delta g(t)) ds \right) dt \quad (50)$$

From the mean value theorem,  $\forall t \in E_2(\rho)$ ,  $\exists \sigma_t$  such that

$$\psi'(g(x[z_\epsilon](t)) + \sigma_t \Delta g(t)) = \int_0^1 \psi'(g(x[z_\epsilon](t)) + s\Delta g(t)) ds \quad (51)$$

Since for all  $t \in E_2(\rho)$ , we have  $g(x[z_\epsilon](t)) - g(x[z_\delta](t)) \geq G_\delta$  and since  $\psi'$  is strictly increasing we have  $\sigma_t \in (0, 1)$  and

$$\psi' \circ g(x[z_\delta](t)) < \psi'(g(x[z_\epsilon](t)) + \sigma_t \Delta g(t)) < \psi' \circ g(x[z_\epsilon](t)) \quad (52)$$

From the intermediate value theorem,  $\exists \bar{\sigma} \in (0, 1)$  such that  $\forall t \in E_2(\rho)$  we have

$$\psi'(g(x[z_\epsilon](t)) + \sigma_t \Delta g(t)) \geq (1 - \bar{\sigma})\psi' \circ g(x[z_\epsilon](t)) + \bar{\sigma}\psi' \circ g(x[z_\delta](t)) \quad (53)$$

Gathering eqs. (50), (51) and (53) yields

$$\begin{aligned} \int_{E_2(\rho)} \psi \circ g(x[z_\delta](t)) - \psi \circ g(x[z_\epsilon](t)) dt \\ \leq -G_\delta \int_{E_2(\rho)} \left( (1 - \bar{\sigma})\psi' \circ g(x[z_\epsilon](t)) + \bar{\sigma}\psi' \circ g(x[z_\delta](t)) \right) dt \\ \leq -G_\delta \left( (1 - \bar{\sigma}) \int_{E_2(\rho)} \psi' \circ g(x[z_\epsilon](t)) dt + \text{const}(\bar{\sigma}, \psi, T, g, G_\delta) \right) \end{aligned} \quad (54)$$

Gathering eqs. (48) and (54) we have

$$\epsilon \Delta_2 \leq \text{const}(T, f, g, G_\delta, \epsilon_0, \psi, \bar{\sigma}) - \epsilon G_\delta (1 - \bar{\sigma}) \int_{E_2(\rho)} \psi' \circ g(x[z_\epsilon](t)) dt \quad (55)$$

Now, let us prove that any optimal solution is strictly interior with respect to the state constraint. The proof is by contradiction. Using definition 3, one can make the following change in measure

$$\int_{E_3(\rho)} \psi' \circ g(x[z_\epsilon](t)) dt = \int_{-G_\delta}^{-\rho} \psi'(s) m[z_\epsilon, g](ds) \quad (56)$$

Then, using proposition 3 and eqs. (45) and (46) yields

$$\int_{E_2(\rho)} \psi' \circ g(x[z_\epsilon](t)) dt \geq \int_{E_3(\rho)} \psi' \circ g(x[z_\epsilon](t)) dt \geq \text{const}(f, g) (\psi(\rho) - \psi(G_\delta)) \quad (57)$$

Gathering eqs. (39), (43), (55) and (57) yields that  $\forall \rho > 0$  we have

$$J_\epsilon^1(z_\delta) - J_\epsilon^1(z_\epsilon) \leq \text{const}(\ell, f, g, \varphi, T, \epsilon_0, \psi, G_\delta, R_u, \bar{\sigma}) - \epsilon G_\delta \text{const}(f, g, \bar{\sigma}) (\psi(\rho) - \psi(G_\delta)) \quad (58)$$

For  $\rho$  small enough, this yields  $J_\epsilon^1(z_\delta) < J_\epsilon^1(z_\epsilon)$  and contradicts the local optimality of  $z_\epsilon$  and proves eq. (30). Now, to prove eq. (31), let us ensure that the left-hand side of this equation is well-defined. From eq. (30) we have  $(g(x[z_\epsilon]))^{-1}(\{0\}) = \emptyset$ , thus

$$[0, T] = \lim_{\rho \rightarrow 0} ((g(x[z_\epsilon]))^{-1}((-\infty, \rho))) \quad (59)$$

Hence, using definition 3, one has

$$\|\psi' \circ g(x[z_\epsilon])\|_{L^1} = \int_0^T \psi' \circ g(x[z_\epsilon](t)) dt := \lim_{\rho \rightarrow 0} \int_{-\infty}^{-\rho} \psi'(s) m[z_\epsilon, g](ds) \quad (60)$$

which is well-defined. Now, let us prove eq. (31) by contradiction and assume that

$$\forall K_g > 0, \exists \epsilon > 0 \text{ s.t. } \|\epsilon \psi' \circ g(x[z_\epsilon])\|_{L^1} > K_g \quad (61)$$

Then, from eqs. (44), (45) and (59), one has

$$\lim_{\rho \rightarrow 0} \epsilon \int_{-G_\delta}^{-\rho} \psi'(s) m[z_\epsilon, g](ds) > K_g - \epsilon \int_{-\infty}^{-G_\delta} \psi'(s) m[z_\epsilon, g](ds) > K_g - \frac{\epsilon_0 T}{G_\delta} \quad (62)$$

Gathering eqs. (43), (55) and (62) yields

$$\Delta_1 + \epsilon \Delta_2 \leq \text{const}(\ell, f, g, \varphi, T, \epsilon_0, \psi, G_\delta, R_v, R_x, \bar{\sigma}) - G_\delta(1 - \bar{\sigma})K_g \quad (63)$$

Since  $G_\delta(1 - \bar{\sigma}) > 0$ ,  $\exists K_g > 0$  such that  $\Delta_1 + \epsilon \Delta_2 < 0$  which contradicts the optimality of  $z_\epsilon$ , proves eq. (31) and concludes the proof.  $\square$

## 4.2 Mixed constraints interiority analysis

**Lemma 2.** *There exists a constant  $K_c < +\infty$  such that for all  $\epsilon > 0$  and for any  $(x[u_\epsilon, x_\epsilon^0], u_\epsilon)$  locally optimal solution of Problem eq. (17) the following holds*

$$\|\epsilon \psi' \circ c_i(x[u_\epsilon, x_\epsilon^0], u_\epsilon)\|_{L^1} \leq K_c, \quad i = 1, \dots, n_c \quad (64)$$

*Proof.* It is sufficient to prove the case where  $n_c = 1$ , i.e., when there is a single mixed constraint. From proposition 4,  $\forall \delta > 0, \exists (v_\delta, x_\delta^0) \in B_{\|\cdot\|_{\text{Vad}}}((u_\epsilon, x_\epsilon^0), \delta) \cap \text{V}^{\text{ad}}$  and  $C_\delta > 0$  such that

$$c(x[v_\delta, x_\delta^0](t), v_\delta(t)) \leq -2C_\delta, \quad \text{a.a. } t \in [0, T] \quad (65a)$$

$$c(x[v_\delta, x_\delta^0](t), v_\delta(t)) \leq c(x[u_\epsilon, x_\epsilon^0](t), u_\epsilon(t)) - C_\delta, \quad \forall t \in S_{u_\epsilon, x_\epsilon^0}^c(C_\delta) \quad (65b)$$

with  $S_{u_\epsilon, x_\epsilon^0}^c(C_\delta) \neq \emptyset$ . In the following, to alleviate the notations, we denote

$$z_\delta := (v_\delta, x_\delta^0) \quad (66)$$

$$z_\epsilon := (u_\epsilon, x_\epsilon^0) \quad (67)$$

$$\Delta z := z_\delta - z_\epsilon \quad (68)$$

$$\Delta x := x[z_\delta] - x[z_\epsilon] \quad (69)$$

$$\Delta g := g(x[z_\delta]) - g(x[z_\epsilon]) \quad (70)$$

$$\Delta c := c(x[z_\delta], v_\delta) - c(x[z_\epsilon], u_\epsilon) \quad (71)$$

In addition, From lemma 1, and by continuity of the mapping  $z \mapsto x[z]$  one can chose  $\delta > 0$  such that the following holds

$$\sup_t g(x[z_\delta](t)) < 0 \quad (72)$$

$$\|\epsilon\psi' \circ g(x[z_\delta])\|_{L^1} \leq 2 \|\epsilon\psi' \circ g(x[z_\epsilon])\|_{L^1} \leq 2K_g \quad (73)$$

Now, one can exhibit an upper-bound on the difference  $J_\epsilon(x[z_\delta], v_\delta) - J_\epsilon(x[z_\epsilon], u_\epsilon)$  as follows

$$J_\epsilon(x[z_\delta], v_\delta) - J_\epsilon(x[z_\epsilon], u_\epsilon) = \Delta_1 + \Delta_2 + \epsilon\Delta_3 \quad (74)$$

where

$$\Delta_1 := \varphi(x[z_\delta](T)) - \varphi(x[z_\epsilon](T)) + \int_0^T [\ell(x[z_\delta](t), v_\delta(t)) - \ell(x[z_\epsilon](t), u_\epsilon(t))] dt \quad (75)$$

$$\Delta_2 := \epsilon \int_0^T \sum_i [\psi \circ g_i(x[z_\delta](t)) - \psi \circ g_i(x[z_\epsilon](t))] dt \quad (76)$$

$$\Delta_3 := \int_0^T [\psi \circ c(x[z_\delta](t), v_\delta(t)) - \psi \circ c(x[z_\epsilon](t), u_\epsilon(t))] dt \quad (77)$$

Now, let us upper-bound  $\Delta_1$

$$\begin{aligned} \Delta_1 &\leq \int_0^T \text{const}(\ell) (\|x[z](t) - x[z_\epsilon](t)\| + \|v(t) - u_\epsilon(t)\|) dt + \text{const}(\varphi) \|\Delta x\|_{L^\infty} \\ &\leq \text{const}(\ell, f, g, \varphi, T, R_v, R_x) \end{aligned} \quad (78)$$

Now, let us upper-bound  $\Delta_2$ .

$$\Delta_2 = \epsilon \int_0^T \sum_i [\psi \circ g_i(x[z_\delta](t)) - \psi \circ g_i(x[z_\epsilon](t))] dt = \epsilon \sum_i \int_0^T \int_0^1 \psi' [g_i(x[z_\epsilon](t)) + s\Delta g_i(t)] \Delta g_i(t) ds dt \quad (79)$$

From the mean value theorem, eqs. (72) and (73),  $\exists \theta_t \in [0, 1]$  such that

$$\Delta_2 = \epsilon \sum_i \int_0^T \psi' \circ g_i(x[z_\epsilon](t) + \theta_t \Delta g(t)) \Delta g(t) dt \leq \sum_i 2K_g \|\Delta g\|_{L^\infty} \leq \text{const}(f, g, K_g, R_x, T) \quad (80)$$

Now, let us upper-bound  $\Delta_3$  defined in eq. (77). To do so, let us introduce the following useful subsets of  $[0, T]$

$$E_1 := (c(x[z_\epsilon], u_\epsilon))^{-1}((-\infty, -C_\delta]) \quad (81)$$

$$E_2 := (c(x[z_\epsilon], u_\epsilon))^{-1}((-C_\delta, 0]) \quad (82)$$

Let us decompose  $\Delta_3$  as follows  $\Delta_3 := \Delta_{3,1} + \Delta_{3,2}$ , with

$$\Delta_{3,i} := \int_{E_i} \psi \circ c(x[z_\delta](t), v_\delta(t)) - \psi \circ c(x[z_\epsilon](t), u_\epsilon(t)) dt, \quad i = 1, 2 \quad (83)$$

By convexity of the log penalty, i.e.  $\psi$ , we have

$$\Delta_{3,1} \leq \int_{E_1} \psi'(C_\delta) \|\Delta c\|_{L^\infty} dt \leq \text{const}(T, f, c, \psi, C_\delta) \|z_\delta - z_\epsilon\|_{\text{Vad}} \leq \text{const}(T, f, c, \psi, C_\delta, R_v) \quad (84)$$

In addition, we have

$$\Delta_{3,2} = \int_{E_2} \left( \int_0^1 \psi'(c(x[z_\epsilon](t), u_\epsilon(t)) + s\Delta c(t)) \Delta c(t) ds \right) dt \quad (85)$$

Since  $\forall t \in E_2$ ,  $\Delta c(t) < -C_\delta$ , we also have

$$\Delta_{3,2} \leq -C_\delta \int_{E_2} \left( \int_0^1 \psi'(c(x[z_\epsilon](t), u_\epsilon(t)) + s\Delta c(t)) ds \right) dt \quad (86)$$

From the mean value theorem,  $\forall t \in E_2$ ,  $\exists \sigma_t \in [0, 1]$  such that

$$\psi'(c(x[z_\epsilon](t), u_\epsilon(t)) + \sigma_t \Delta c(t)) = \int_0^1 \psi'(c(x[z_\epsilon](t), u_\epsilon(t)) + s\Delta c(t)) ds \quad (87)$$

Since for all  $t \in E_2$ , we have  $c(x[z_\epsilon](t), u_\epsilon(t)) - c(x[z_\delta](t), v_\delta(t)) \geq C_\delta$  and since  $\psi'$  is strictly increasing we have  $\sigma_t \in (0, 1)$  and

$$\psi' \circ c(x[z_\delta](t), v_\delta(t)) < \psi'(c(x[z_\epsilon](t), u_\epsilon(t)) + \sigma_t \Delta c(t)) < \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) \quad (88)$$

From the intermediate value theorem,  $\exists \bar{\sigma} \in (0, 1)$  such that  $\forall t \in E_2$  we have

$$\psi'(c(x[z_\epsilon](t), u_\epsilon(t)) + \sigma_t \Delta c(t)) \geq (1 - \bar{\sigma}) \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) + \bar{\sigma} \psi' \circ c(x[z_\delta](t), v_\delta(t)) \quad (89)$$

Gathering eqs. (86), (87) and (89) yields

$$\Delta_{3,2} \leq -C_\delta \int_{E_2} \left( (1 - \bar{\sigma}) \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) + \bar{\sigma} \psi' \circ c(x[z_\delta](t), v_\delta(t)) \right) dt \quad (90)$$

$$\leq -C_\delta \left( (1 - \bar{\sigma}) \int_{E_2} \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) dt + \text{const}(\bar{\sigma}, \psi, T, c, C_\delta) \right) \quad (91)$$

Gathering eqs. (84) and (91) we have

$$\epsilon \Delta_3 \leq \text{const}(T, f, c, C_\delta, \epsilon_0, \psi, \bar{\sigma}, R_v) - \epsilon C_\delta (1 - \bar{\sigma}) \int_{E_2} \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) dt \quad (92)$$

Gathering eqs. (78), (80) and (92) yields

$$\begin{aligned} J_\epsilon(z_\delta) - J_\epsilon(z_\epsilon) &= \Delta_1 + \Delta_2 + \epsilon \Delta_3 \leq \\ &\text{const}(\ell, f, g, \varphi, T, K_g, R_v, R_x, C_\delta, c, \epsilon_0, \psi, \bar{\sigma}) - \epsilon C_\delta (1 - \bar{\sigma}) \int_{E_2} \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) dt \end{aligned} \quad (93)$$

Now let us prove eq. (64) by contradiction and assume that

$$\forall K_c > 0, \exists \epsilon > 0 \text{ s.t. } \|\epsilon \psi' \circ c(x[z_\epsilon], u_\epsilon)\|_{L^1} > K_c \quad (94)$$

From the definition of  $E_1$  and  $E_2$ , we have

$$\|\epsilon \psi' \circ c(x[z_\epsilon], u_\epsilon)\|_{L^1} = \int_{E_1} \epsilon \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) dt + \int_{E_2} \epsilon \psi' \circ c(x[z_\epsilon](t), u_\epsilon(t)) dt \quad (95)$$

which, in turns yields

$$\int_{E_2} \epsilon \psi'(c(x[z_\epsilon](t), u_\epsilon(t))) dt > K_c - \epsilon_0 \psi'(C_\delta) T \quad (96)$$

gathering eqs. (93) and (96) yields

$$J_\epsilon(z_\delta) - J_\epsilon(z_\epsilon) \leq \text{const}(\ell, f, g, \varphi, T, K_g, R_v, R_x, C_\delta, c, \epsilon_0, \psi, \bar{\sigma}) - C_\delta(1 - \bar{\sigma})K_c \quad (97)$$

For  $K_c$  large enough,  $J_\epsilon(z_\delta) - J_\epsilon(z_\epsilon) < 0$ , which contradicts the optimality of  $(x[z_\epsilon], u_\epsilon)$ , proves eq. (64) and concludes the proof.  $\square$

In addition, using lemma 1 and lemma 2, one can also prove a uniform boundedness property for the adjoint state  $\bar{p}_\epsilon$  from definition 8.

**Corollary 1.** *Let  $(\bar{u}_\epsilon, \bar{x}_\epsilon)$  be a locally optimal solution of Problem eq. (17) and let  $(\bar{p}_\epsilon, \bar{\lambda}_\epsilon)$  be the corresponding constraint multipliers, then there exists  $K_p < \infty$  such that  $\|\bar{p}_\epsilon\|_{L^\infty} \leq K_p$*

*Proof.* First, using eq. (19b) one has

$$\begin{aligned} \|\bar{p}_\epsilon(T) - \bar{p}_\epsilon(s)\| &\leq \int_s^T (\|\ell'_x(\bar{x}_\epsilon, \bar{u}_\epsilon)\|_{L^\infty} + \|f'_x(\bar{x}_\epsilon, \bar{u}_\epsilon)\|_{L^\infty} \|\bar{p}_\epsilon(t)\|) dt + \sum_i \|g'_i(\bar{x}_\epsilon)\|_{L^\infty} \|\epsilon \psi' \circ g_i(\bar{x}_\epsilon)\|_{L^1} \\ &\quad + \sum_i \|c'_{i,x}(\bar{x}_\epsilon, \bar{u}_\epsilon)\|_{L^\infty} \|\epsilon \psi' \circ c_i(\bar{x}_\epsilon, \bar{u}_\epsilon)\|_{L^1} \quad (98) \end{aligned}$$

From the continuity of  $\ell'_x, f'_x, g'_i, c'_{i,x}$  and since  $\bar{x}_\epsilon$  and  $\bar{u}_\epsilon$  are bounded, we have  $\|\ell'_x(\bar{x}_\epsilon, \bar{u}_\epsilon)\|_{L^\infty} < \text{const}(\ell, f)$  and  $\|f'_x(\bar{x}_\epsilon, \bar{u}_\epsilon)\|_{L^\infty} < \text{const}(f)$ . In addition, the terms on the right-hand side of eq. (19f) are bounded which yields that  $\|\bar{p}_\epsilon(T)\| \leq \text{const}(f, h)$ . The derivatives of the penalty functions being uniformly  $L^1$ -bounded one can use Grönwall Lemma which proves that  $\exists K_p < +\infty$  such that  $\forall s \in [0, T]$  we have  $\|\bar{p}_\epsilon(s)\| \leq K_p$  which concludes the proof.  $\square$

**Lemma 3.** *There exists a constant  $K_c < +\infty$  such that for all  $\epsilon > 0$ , any  $(\bar{x}_\epsilon, \bar{u}_\epsilon)$  locally optimal solution of Problem eq. (17) satisfies*

$$c(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t)) \leq -\epsilon/K_c, \text{ a.e.} \quad (99)$$

*Proof.* For all  $K_c > 0$ , assume that there exists  $E \subseteq [0, T]$  of strictly positive measure such that  $I_{\bar{u}_\epsilon, \bar{x}_\epsilon(0)}^c(t, K_c/\epsilon) \neq \emptyset$  for all  $t \in E$ . Now, let us denote  $C(t) := a_{I_{\bar{u}_\epsilon, \bar{x}_\epsilon(0)}^c(t, K_c/\epsilon)}(\bar{x}_\epsilon)$  and let us define  $v$  as follows

$$v(t) := \begin{cases} -C(t)^\top [C(t)C(t)^\top]^{-1} .e_v, \forall t \in E \\ 0 \text{ otherwise} \end{cases} \quad (100)$$

where  $\mathbb{R}^{|I_{\bar{u}_\epsilon, \bar{x}_\epsilon(0)}^c(t, K_c/\epsilon)|} \ni e_v := (\epsilon/K_c \ \dots \ \epsilon/K_c)^\top$ . Since  $\bar{u}_\epsilon$  is a locally optimal solution, we have for almost all  $t \in E$

$$\begin{aligned} H^\psi(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t) + v(t), \bar{p}_\epsilon(t), \epsilon) - H^\psi(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t), \bar{p}_\epsilon(t), \epsilon) &= H(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t) + v(t), \bar{p}_\epsilon(t)) - H(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t), \bar{p}_\epsilon(t)) \\ &\quad + \epsilon \sum_{i \in I_{\bar{u}_\epsilon, \bar{x}_\epsilon(0)}^c(t, K_c/\epsilon)} \log \left( \frac{c_i(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t))}{c_i(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t) + v(t))} \right) \quad (101) \end{aligned}$$

From the mean value theorem and the definition of  $v$ ,  $\exists s \in [0, 1]$  such that

$$H^\psi(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t) + v(t), \bar{p}_\epsilon(t), \epsilon) - H^\psi(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t), \bar{p}_\epsilon(t), \epsilon) \leq H'_u(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t) + sv(t), \bar{p}_\epsilon(t)).v(t) - \epsilon \sum_{i \in I_{\bar{u}_\epsilon, \bar{x}_\epsilon(0)}^c(t, K_c/\epsilon)} \log(2) \quad (102)$$

Since  $\ell, f$  are at least  $C^2$  and from corollary 1 we have

$$H^\psi(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t) + v(t), \bar{p}_\epsilon(t), \epsilon) - H^\psi(\bar{x}_\epsilon(t), \bar{u}_\epsilon(t), \bar{p}_\epsilon(t), \epsilon) \leq \epsilon \left( \frac{\text{const}(\ell, f, K_p, T, R_u)}{K_c} - \sum_{i \in I_{\bar{u}_\epsilon, \bar{x}_\epsilon(0)}^c(t, K_c/\epsilon)} \log(2) \right) \quad (103)$$

which is negative for  $K_c$  large enough and contradicts the local optimality of  $\bar{u}_\epsilon$  and proves the result.  $\square$

## 5 Convergence of interior point methods in optimal control with logarithmic penalty functions

### 5.1 Convergence of state variable and initial-final conditions

Let us denote  $(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n})_n$  a sequence of locally optimal solutions of eq. (17). The associated sequence  $(\bar{u}_{\epsilon_n}, \bar{x}_{\epsilon_n}(0))_n$  being  $L^2 \times \mathbb{R}^n$ -bounded, it contains a weakly converging subsequence  $(\bar{u}_{\epsilon_{n_k}}, \bar{x}_{\epsilon_{n_k}}(0))_k$  satisfying i.e.

$$\lim_{k \rightarrow +\infty} \bar{u}_{\epsilon_{n_k}} \rightharpoonup \bar{u} \quad (104)$$

$$\lim_{k \rightarrow +\infty} \left\| \bar{x}_{\epsilon_{n_k}}(0) - \bar{x}^0 \right\| = 0 \quad (105)$$

From proposition 2, we also have

$$\lim_{k \rightarrow +\infty} \left\| x[\bar{u}_{\epsilon_{n_k}}, \bar{x}_{\epsilon_{n_k}}(0)] - x[\bar{u}, \bar{x}^0] \right\|_{L^\infty} = 0 \quad (106)$$

$$\lim_{k \rightarrow +\infty} \left\| h \left( x[\bar{u}_{\epsilon_{n_k}}, \bar{x}_{\epsilon_{n_k}}(0)](0), x[\bar{u}_{\epsilon_{n_k}}, \bar{x}_{\epsilon_{n_k}}(0)](T) \right) - h \left( x[\bar{u}, \bar{x}^0](0), x[\bar{u}, \bar{x}^0](T) \right) \right\| = 0 \quad (107)$$

which proves eqs. (20b) and (20c).

### 5.2 Convergence of initial-final constraints multipliers

Let  $(\lambda_{\epsilon_n})_n$  be the sequence of multipliers associated with the initial-final constraints eq. (19d). These multipliers being bounded there exists a converging subsequence to some  $\bar{\lambda}$ , which writes  $\lim_{k \rightarrow +\infty} \left\| \bar{\lambda}_{\epsilon_{n_k}} - \bar{\lambda} \right\| = 0$ .

### 5.3 Convergence of state penalties

In this paragraph, we prove that the derivative of the state-constraint penalty converges to a Radon measure  $\bar{\mu} \in \mathcal{M}([0, T])^{n_g}$ . To do so, let us denote

$$L^1([0, T]; \mathbb{R}^+) \ni \theta_{\epsilon_n}^{g_i} := \epsilon_n \psi' \circ g_i(\bar{x}_{\epsilon_n}) = -\frac{\epsilon_n}{g_i(\bar{x}_{\epsilon_n})} \quad (108)$$

Identifying any element of the sequence  $(\theta_{\epsilon_n}^{g_i})_n$  with a linear form on continuous functions  $\Theta_{\epsilon_n}^{g_i} \in \mathcal{M}([0, T])$  defined as follows  $\Theta_{\epsilon_n}^{g_i} : v \in C^0([0, T], \mathbb{R}) \mapsto \int_0^T v(t) \theta_{\epsilon_n}^{g_i}(t) dt$ . From lemma 1 and eq. (31) we have  $|\Theta_{\epsilon_{n_k}}^{g_i}(v)| \leq K_g \|v\|_{L^\infty}$ , thus

$$\forall \epsilon_n, \Theta_{\epsilon_n}^{g_i} \in B_{\mathcal{M}([0, T])}(0, K_g)$$

From the weak \* compactness of the unit ball of  $\mathcal{M}([0, T])$  (see [8, Theorem 3.16]), there exists a subsequence  $(\Theta_{\epsilon_{n_k}}^{g_i})_{k \in \mathbb{N}}$  and a measure  $\bar{\mu}_i \in \mathcal{M}([0, T])$  with  $\bar{\mu}_i(T) = 0$  such that  $\lim_{k \rightarrow +\infty} \Theta_{\epsilon_{n_k}}^{g_i} \xrightarrow{*} \bar{\mu}_i$ ,  $i = 1, \dots, n_g$  which proves eq. (20h). Now, let us prove that  $\bar{\mu}$  satisfies conditions eqs. (15g) and (15j). From lemma 1 and eq. (30) and from eq. (108), we have  $\theta_{\epsilon_{n_k}}^{g_i} > 0$ ,  $\forall t \in [0, T]$  and  $\forall \epsilon_{n_k} > 0$ . Therefore,  $\forall \phi \in C^0([0, T]; \mathbb{R}^+)$  one has  $\int \phi d\bar{\mu}_i = \lim_k \int \phi \theta_{\epsilon_{n_k}}^{g_i} dt \geq 0$ , which proves eq. (15j). Finally, let us prove that  $\bar{\mu}$  satisfies the complementarity condition eq. (15g). From eq. (108), we have  $g_i(\bar{x}_{\epsilon_{n_k}}(t)) \theta_{\epsilon_{n_k}}^{g_i}(t) = -\epsilon_{n_k}$  hence

$$\lim_{k \rightarrow +\infty} \int_0^T g_i(\bar{x}_{\epsilon_{n_k}}(t)) \theta_{\epsilon_{n_k}}^{g_i}(t) dt = \lim_{k \rightarrow +\infty} -\epsilon_{n_k} T = 0 \quad (109)$$

From the continuity of  $g_i$ , the sequence  $(g_i(\bar{x}_{\epsilon_{n_k}}))_k$  uniformly converges to  $g_i(\bar{x})$ . In addition, from lemma 1 and eq. (31), the sequence  $\theta_{\epsilon_{n_k}}^{g_i}(t)$  is uniformly  $L^1$ -bounded, hence

$$\lim_{k \rightarrow +\infty} \left| \int_0^T (g_i(\bar{x}(t)) - g_i(\bar{x}_{\epsilon_{n_k}}(t))) \theta_{\epsilon_{n_k}}^{g_i}(t) dt \right| \leq \lim_{k \rightarrow +\infty} \|g_i(\bar{x}) - g_i(\bar{x}_{\epsilon_{n_k}})\|_{L^\infty} \|\theta_{\epsilon_{n_k}}^{g_i}\|_{L^1} = 0 \quad (110)$$

Gathering eqs. (109) and (110) yields

$$\lim_{k \rightarrow +\infty} \int_0^T g_i(\bar{x}(t)) \theta_{\epsilon_{n_k}}^{g_i}(t) dt = \lim_{k \rightarrow +\infty} \int_0^T g_i(\bar{x}_{\epsilon_{n_k}}(t)) \theta_{\epsilon_{n_k}}^{g_i}(t) dt = 0 \quad (111)$$

which in turns gives

$$\int_0^T g_i(\bar{x}(t)) d\bar{\mu}_i(t) = \lim_{k \rightarrow +\infty} \int_0^T g_i(\bar{x}(t)) \theta_{\epsilon_{n_k}}^{g_i}(t) dt = 0 \quad (112)$$

and proves that  $\bar{\mu}$  satisfies the complementarity condition eq. (15g).

## 5.4 Convergence of mixed-constraint penalties

In this paragraph, we prove that the derivative of the mixed-constraint penalty converges to an essentially bounded function  $\bar{\nu} \in L^\infty([0, T]; \mathbb{R}_+^{n_g})$ . To do so, let us denote

$$\eta_{\epsilon_n, i} := \epsilon_n \psi' \circ c_i(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}) \quad (113)$$

**Proposition 5.** *Let  $(\bar{x}_\epsilon, \bar{u}_\epsilon)$  be a locally optimal solution of eq. (17), then the following holds*

$$\eta_{\epsilon_n} \in L^\infty([0, T]; \mathbb{R}_+^{n_c}) \quad (114)$$

where  $\eta_{\epsilon_n, i} := (\eta_{\epsilon_n, 1} \ \dots \ \eta_{\epsilon_n, n_c})^\top$



*Proof.* The proof of this result consists in proving that the mapping

$$\Psi_\epsilon : L^1 \ni w \mapsto \int_0^T \eta_{\epsilon_n}(t)w(t)dt \in \mathbb{R} \quad (115)$$

is a continuous linear form on  $L^1$ . From lemma 3,  $c(\bar{x}_\epsilon, \bar{u}_\epsilon)$  is strictly negative (not active) almost everywhere. Therefore the Hamiltonian minimization condition of the Pontryagin maximum principle writes  $H'_u(\bar{x}_\epsilon, \bar{u}_\epsilon, \bar{p}_\epsilon) + \sum_{i=1}^{n_c} \eta_{\epsilon, i} a_i(\bar{x}_\epsilon) = 0$  for almost all time. Then, for all  $v \in U$ , one has

$$\left| \int_0^T \sum_{i=1}^{n_c} \eta_{\epsilon_n, i}(t) a_i(\bar{x}_\epsilon(t)) \cdot v(t) dt \right| \leq \|H'_u(\bar{x}_\epsilon, \bar{u}_\epsilon, \bar{p}_\epsilon)\|_{L^\infty} \|v\|_{L^1} \leq \text{const}(f, \ell, h) \|v\|_{L^1} \quad (116)$$

Let us denote  $C(t) := a_{I_{\bar{u}_\epsilon, \bar{x}_\epsilon^0}(t, n)}^c(\bar{x}_\epsilon)$  and for all  $w \in L^1([0, T]; \mathbb{R}^{n_c})$ , let us define  $v \in U$  as follows

$$v(t) := \begin{cases} C(t)^\top [C(t)C(t)^\top]^{-1} w_{I_{\bar{u}_\epsilon, \bar{x}_\epsilon^0}(t, n)}^c & \text{if } I_{\bar{u}_\epsilon, \bar{x}_\epsilon^0}(t, n) \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (117)$$

Since  $C(t)$  is  $L^\infty$ -bounded, there exists  $M > 0$  such that  $\|v\|_{L^1} \leq M \|w\|_{L^1}$ . In addition, let us define  $\lambda_\epsilon \in L^1([0, T]; \mathbb{R}^{n_c})$  as follows

$$\lambda_{\epsilon, i}(t) := \begin{cases} \eta_{\epsilon_n, i} & \text{if } i \in I_{\bar{u}_\epsilon, \bar{x}_\epsilon^0}(t, n) \\ 0 & \text{otherwise} \end{cases} \quad (118)$$

Gathering eqs. (116) to (118) yields

$$\left| \int_0^T \sum_{i=1}^{n_c} \eta_{\epsilon_n, i}(t) a_i(\bar{x}_\epsilon(t)) \cdot v(t) dt \right| = \left| \int_0^T \lambda_\epsilon(t) \cdot w(t) dt \right| \quad (119)$$

Gathering eqs. (116) and (119) we have  $\left| \int_0^T \lambda_\epsilon(t) \cdot w(t) dt \right| \leq \text{const}(f, \ell, h, M) \|w\|_{L^1}$  and using the density of  $L^\infty([0, T]; \mathbb{R}^{n_c})$  in  $L^1([0, T]; \mathbb{R}^{n_c})$  proves that  $\Psi_\epsilon$  is continuous linear form over  $L^1([0, T]; \mathbb{R}^{n_c})$  and concludes the proof.  $\square$

From lemma 3 and proposition 5,  $\eta_{\epsilon_n, i} \in B_{L^\infty}(0, K_c)$ , thus there exists a subsequence and a function  $\bar{\nu}_i \in L^\infty([0, T]; \mathbb{R}_+)$  such that

$$\lim_{k \rightarrow +\infty} \eta_{\epsilon_{n_k}, i} \xrightarrow{*} \bar{\nu}_i \quad (120)$$

which proves eq. (20g). Now, let us prove that  $\bar{\nu}$  satisfies conditions eqs. (15h) and (15k). From lemma 2 and lemma 3, we have  $\eta_{\epsilon_{n_k}, i} > 0, \forall t \in [0, T]$  and  $\forall \epsilon_{n_k} > 0$  which proves that  $\bar{\nu}$  satisfies the non negativity condition eq. (15k). Finally, let us prove that  $\bar{\nu}$  satisfies the complementarity condition eq. (15h). First, we have

$$\lim_{k \rightarrow \infty} \left\langle \eta_{\epsilon_{n_k}, i}, c_i(\bar{x}_{\epsilon_{n_k}}, \bar{u}_{\epsilon_{n_k}}) \right\rangle = \lim_{k \rightarrow \infty} -\epsilon_{n_k} T = 0 \quad (121)$$

Using proposition 2, yields

$$\lim_{s \rightarrow \infty} \lim_{r \rightarrow \infty} \left\langle \eta_{\epsilon_{n_s}, i}, c_i(\bar{x}_{\epsilon_r}, \bar{u}_{\epsilon_r}) \right\rangle = \lim_{r \rightarrow \infty} \lim_{s \rightarrow \infty} \left\langle \eta_{\epsilon_{n_s}, i}, c_i(\bar{x}_{\epsilon_r}, \bar{u}_{\epsilon_r}) \right\rangle = \langle \bar{\nu}_i, c_i(\bar{x}, \bar{u}) \rangle \quad (122)$$

which, in turn, gives

$$\begin{aligned}
\langle \bar{\nu}_i, c_i(\bar{x}, \bar{u}) \rangle &= \liminf_s \liminf_r \langle \eta_{\epsilon_{n_s}, i}, c_i(\bar{x}_{\epsilon_r}, \bar{u}_{\epsilon_r}) \rangle \\
&\leq \lim_{k \rightarrow \infty} \langle \eta_{\epsilon_{n_k}, i}, c_i(\bar{x}_{\epsilon_{n_k}}, \bar{u}_{\epsilon_{n_k}}) \rangle = 0 \\
&\leq \limsup_s \limsup_r \langle \eta_{\epsilon_{n_s}, i}, c_i(\bar{x}_{\epsilon_r}, \bar{u}_{\epsilon_r}) \rangle = \langle \bar{\nu}_i, c_i(\bar{x}, \bar{u}) \rangle \quad (123)
\end{aligned}$$

which proves that  $\bar{\nu}$  satisfies the complementarity condition eq. (15h).

## 5.5 Convergence of Pontryagin adjoint

Let  $q_n \in \text{BV}([0, T])^n$  be the solution of

$$\begin{aligned}
-dq_n(t) &= [\ell'_x(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) + f'_x(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)).q_n(t)] dt \quad (124a) \\
&\quad + \sum_{i=1}^{n_c} c'_{x,i}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) \bar{\nu}_i(t) dt + \sum_{i=1}^{n_g} g'_i(\bar{x}_{\epsilon_n}(t)) d\bar{\mu}_i(t)
\end{aligned}$$

$$q_n(0) = -h'_{x(0)}(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T))^\top \cdot \bar{\lambda} \quad (124b)$$

$$q_n(T) = \varphi'(\bar{x}_{\epsilon_n}(T)) + h'_{x(T)}(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T))^\top \cdot \bar{\lambda} \quad (124c)$$

Then, we have

$$\begin{aligned}
q_n(T) - \bar{p}(T) + \bar{p}(t) - q_n(t) &= \int_t^T \ell'_x(\bar{x}(s), \bar{u}(s)) - \ell'_x(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) ds \\
&\quad + \int_t^T f'_x(\bar{x}(s), \bar{u}(s)) \cdot \bar{p}(s) - f'_x(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) \cdot q_n(s) ds \\
&\quad + \sum_{i=1}^{n_c} \int_t^T [c'_{i,x}(\bar{x}(s), \bar{u}(s)) - c'_{i,x}(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s))] \bar{\nu}_i(s) ds \\
&\quad + \sum_{i=1}^{n_g} \int_t^T [g'_i(\bar{x}(s)) - g'_i(\bar{x}_{\epsilon_n}(s))] d\bar{\mu}_i(s) \quad (125)
\end{aligned}$$

Using proposition 2, we have  $\lim_n \|q_n(T) - \bar{p}(T)\| = 0$  and

$$\lim_n \|\bar{p}(t) - q_n(t)\| = \lim_n \left\| \int_t^T f'_x(\bar{x}(s), \bar{u}(s)) \cdot \bar{p}(s) - f'_x(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) \cdot q_n(s) ds \right\| \quad (126)$$

$$= \lim_n \left\| \int_t^T f'_x(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) \cdot (\bar{p}(s) - q_n(s)) ds \right\| \quad (127)$$

$$\begin{aligned}
&\quad + \lim_n \left\| \int_t^T [f'_x(\bar{x}(s), \bar{u}(s)) - f'_x(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s))] \cdot \bar{p}(s) ds \right\| \\
&\leq \text{const}(f) \int_t^T \|\bar{p}(s) - q_n(s)\| ds \quad (128)
\end{aligned}$$

Therefore,  $q_n$  pointwise converges to  $\bar{p}$  and since both are bounded we have

$$\lim_{n \rightarrow +\infty} \|q_n - \bar{p}\|_{L^1} = 0 \quad (129)$$

In addition, using eq. (108), we have

$$\begin{aligned} \|q_n(t) - \bar{p}_{\epsilon_n}(t)\| &\leq \left\| \int_t^T f'_x(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) \cdot (q_n(s) - \bar{p}_{\epsilon_n}(s)) ds \right. \\ &\quad + \sum_{i=1}^{n_g} \int_t^T g'_i(\bar{x}_{\epsilon_n}(s)) (d\bar{\mu}_i(s) - \theta_{\epsilon_n}^{g_i}(s) ds) \\ &\quad \left. + \sum_{i=1}^{n_c} \int_t^T c'_{i,x}(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) (\bar{\nu}_i(s) - \eta_{\epsilon_n,i}(s)) ds \right\| \\ &\leq \text{const}(f) \int_t^T \|q_n(s) - \bar{p}_{\epsilon_n}(s)\| ds \\ &\quad + \sum_{i=1}^{n_g} \left\| \int_t^T g'_i(\bar{x}_{\epsilon_n}(s)) (d\bar{\mu}_i(s) - \theta_{\epsilon_n}^{g_i}(s) ds) \right\| \\ &\quad + \sum_{i=1}^{n_c} \left\| \int_t^T c'_{i,x}(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) [\bar{\nu}_i(s) - \eta_{\epsilon_n,i}(s)] ds \right\| \end{aligned} \quad (130)$$

$$\leq \text{const}(f) \int_t^T \|q_n(s) - \bar{p}_{\epsilon_n}(s)\| ds \quad (131)$$

Now, let us define  $h_n \in L^1([0, T]; \mathbb{R}_+)$  as follows

$$h_n(t) := \sum_{i=1}^{n_g} \left\| \int_t^T g'_i(\bar{x}_{\epsilon_n}(s)) (d\mu_i(s) - \theta_{\epsilon_n}^{g_i}(s) ds) \right\| + \sum_{i=1}^{n_c} \left\| \int_t^T c'_{i,x}(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) [\bar{\nu}_i(s) - \eta_{\epsilon_n,i}(s)] ds \right\| \quad (132)$$

thus  $\|q_n(t) - \bar{p}_{\epsilon_n}(t)\| \leq \text{const}(f) \int_t^T \|q_n(s) - \bar{p}_{\epsilon_n}(s)\| ds + h_n(t)$ . From Grönwall inequality [12, Lemma A.1, p.651], we have  $\|q_n(t) - \bar{p}_{\epsilon_n}(t)\| \leq \text{const}(f, T) \int_t^T h_n(s) ds$ . From the  $L^\infty$ -convergence of  $\bar{x}_{\epsilon_n}$  and the weak \* convergence of  $\theta_{\epsilon_n}^{g_i}$  we have

$$\lim_{n \rightarrow \infty} \int_t^T g'_i(\bar{x}_{\epsilon_n}(s)) (d\mu_i(s) - \theta_{\epsilon_n}^{g_i}(s) ds) = 0 \quad (133)$$

From proposition 2 and eq. (120), we have

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \langle \eta_{\epsilon_n,i} - \bar{\nu}_i, c'_{i,x}(\bar{x}_{\epsilon_m}, \bar{u}_{\epsilon_m}) \rangle = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \eta_{\epsilon_n,i} - \bar{\nu}_i, c'_{i,x}(\bar{x}_{\epsilon_m}, \bar{u}_{\epsilon_m}) \rangle = 0 \quad (134)$$

which, in turn, gives

$$\begin{aligned} 0 &= \liminf_n \liminf_m \langle \eta_{\epsilon_n,i} - \bar{\nu}_i, c'_{i,x}(\bar{x}_{\epsilon_m}, \bar{u}_{\epsilon_m}) \rangle \leq \lim_{n \rightarrow \infty} \langle \eta_{\epsilon_n,i} - \bar{\nu}_i, c'_{i,x}(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}) \rangle \\ &\leq \limsup_n \limsup_m \langle \eta_{\epsilon_n,i} - \bar{\nu}_i, c'_{i,x}(\bar{x}_{\epsilon_m}, \bar{u}_{\epsilon_m}) \rangle = 0 \end{aligned} \quad (135)$$

which yields  $\lim_{n \rightarrow \infty} \int_t^T c'_{i,x}(\bar{x}_{\epsilon_n}(s), \bar{u}_{\epsilon_n}(s)) [\bar{\nu}_i(s) - \eta_{\epsilon_n,i}(s)] ds = 0$ . Thus,  $h_n$  pointwise converges to 0. In addition, from the boundedness of  $(h_n)_n$  there exists a subsequence such that

$$\lim_{k \rightarrow +\infty} \left\| q_{n_k}(t) - \bar{p}_{\epsilon_{n_k}}(t) \right\| \leq \text{const}(f, T) \int_t^T \lim_{k \rightarrow +\infty} h_{n_k}(s) ds = 0 \quad (136)$$

$q_{n_k}$  pointwise converges to  $\bar{p}_{\epsilon_{n_k}}$  and since both are bounded, from Lebesgue-Vitali's Theorem, we have

$$\lim_{k \rightarrow +\infty} \left\| q_{n_k} - \bar{p}_{\epsilon_{n_k}} \right\|_{L^1} = 0$$

Gathering with eq. (129) yields  $\left\| \bar{p} - \bar{p}_{\epsilon_{n_k}} \right\|_{L^1} \rightarrow 0$  which proves eq. (20f).

## 5.6 Convergence of control variable and cost function

From, the convexity of  $\ell(x, u)$  with respect to  $u$  and from the strict convexity of the penalty function we have

$$\forall n > 0, \text{ for a.e. } t \in [0, T], H_{uu}^{\psi}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t), \bar{p}_{\epsilon_n}(t), \epsilon_n) > 0 \quad (137)$$

From the implicit function theorem [22, Theorem 9.27, pp. 224-225], for almost all time, there exists a mapping  $\lambda_t$  such that

$$\bar{u}_{\epsilon_n}(t) := \lambda_t(\bar{x}_{\epsilon_n}(t), \bar{p}_{\epsilon_n}(t), \epsilon_n) \quad (138)$$

By continuity of  $\lambda_t$  and from the strong  $L^\infty$  (resp.  $L^1$ ) convergence of  $\bar{x}_{\epsilon_n}$  (resp.  $\bar{p}_{\epsilon_n}$ ),  $\bar{u}_{\epsilon_n}$  pointwise converges to some  $z \in \text{BV}([0, T])^m$ . Now, since the sequence  $(\bar{u}_{\epsilon_n})_n$  weakly converges to  $\bar{u}$ , from Mazur's lemma [20, lemma 10.19, pp. 350], there exists a function  $N : \mathbb{N} \mapsto \mathbb{N}$  and a sequence of sets of real positive numbers  $(\{\alpha[n]_k : k = n, \dots, N(n)\})_n$  satisfying  $\sum_{k=n}^{N(n)} \alpha[n]_k = 1$  and such the sequence  $(v_n)_n$  defined as follows

$$v_n := \sum_{k=n}^{N(n)} \alpha[n]_k u_{\epsilon_k} \quad (139)$$

converges in  $L^2$ -norm to  $\bar{u}$ . Therefore, there exists a subsequence denoted  $(v_m)_m$  converging almost everywhere to  $\bar{u}$ . Now, for almost all  $t \in [0, T]$ , we have

$$\|\bar{u}(t) - z(t)\| = \lim_{m \rightarrow +\infty} \|v_m(t) - z(t)\| \leq \lim_{m \rightarrow +\infty} \sum_{k=m}^{N(m)} \alpha[m]_k \|\bar{u}_{\epsilon_m}(t) - z(t)\| = 0 \quad (140)$$

which proves that there exists a subsequence  $(\bar{u}_{\epsilon_{n_k}})_k$  which converges almost everywhere to  $\bar{u}$  and since  $(\bar{u}_{\epsilon_n})_n$  is  $L^\infty$ -bounded the subsequence converges in  $L^1$ -norm to  $\bar{u}$  which proves eq. (20a) and eq. (20d) from  $\ell \in C^1$ .

## 5.7 Convergence of stationary conditions on the Hamiltonian

Let us prove that the limit point of the solution of eq. (19) is also a solution of eq. (15c). From eqs. (20a) and (20b)

$$\begin{aligned} \lim_{n \rightarrow +\infty} \left\| H_u^{\psi'}(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}, \bar{p}_{\epsilon_n}, \epsilon_n) - H'_u(\bar{x}, \bar{u}, \bar{p}) - b(\bar{x})^\top \cdot \bar{p} \right\|_{L^1} \leq \\ \lim_{n \rightarrow +\infty} \left[ \left\| \ell'_u(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}) - \ell'_u(\bar{x}, \bar{u}) \right\|_{L^1} + \left\| f_2(\bar{x}_{\epsilon_n})^\top \cdot \bar{p}_{\epsilon_n} - f_2(\bar{x})^\top \cdot \bar{p} \right\|_{L^1} \right. \\ \left. + \sum_{i=1}^{n_c} \left\| \epsilon_n \psi' \circ c_i(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}) b_i(\bar{x}_{\epsilon_n}) - \bar{v}_i b_i(\bar{x}) \right\|_{L^1} \right] \quad (141) \end{aligned}$$

Each term of the sum converges in  $L^1$ -norm to 0. Thus, taking a subsequence if necessary, eq. (19c) converges to 0 almost everywhere and proves eq. (15c).

## 6 Solving Algorithms

### 6.1 Primal solving algorithm

In theorem 1 we have proved that any sequence of solutions of eq. (19) contains a converging subsequence. In the following, we denote  $S_P(\epsilon) := (\bar{x}_\epsilon, \bar{p}_\epsilon, \bar{u}_\epsilon, \bar{\lambda}_\epsilon)$  any solution of eq. (19). Now, the primal solving algorithm naturally writes as follows

- 1: Define  $\epsilon_0 > 0$ ,  $\alpha \in (0, 1)$ ,  $\text{tol} = o(1)$ ,  $k = 0$
- 2: **while**  $\epsilon_k > \text{tol}$  **do**
- 3:    $S_P(\epsilon_{k+1}) \leftarrow$  solution of eq. (19) initialized with  $S_P(\epsilon_k)$
- 4:    $\epsilon_{k+1} \leftarrow \alpha \epsilon_k$
- 5:    $k \leftarrow k + 1$
- 6: **end while**
- 7: **return**  $S_P(\epsilon_k)$

**Algorithm 1:** Primal algorithm for optimal control problems

### 6.2 Primal-dual solving algorithm

Before describing the primal-dual solving algorithm, we need the following convergence result, which is a direct consequence of theorem 1.

**Theorem 2.** *Let  $(\epsilon_n)$  be a sequence of decreasing positive parameters with  $\epsilon_n \rightarrow 0$  and let*

$$(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}, \bar{p}_{\epsilon_n}, \bar{\theta}_{\epsilon_n}, \bar{\eta}_{\epsilon_n}, \bar{\lambda}_{\epsilon_n})_n \in \mathbb{W}^{1,\infty}([0, T]; \mathbb{R}^n) \times \mathbb{U} \times \mathbb{W}^{1,1}([0, T]; \mathbb{R}^n) \times \mathbb{L}^1([0, T]; \mathbb{R}_+^{n_g}) \times \mathbb{L}^\infty([0, T]; \mathbb{R}_+^{n_c}) \times \mathbb{R}^{n_h} \quad (142)$$

be a solution of the following Primal Dual TPBVP

$$\dot{\bar{x}}_{\epsilon_n}(t) = f(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) \quad (143a)$$

$$\dot{\bar{p}}_{\epsilon_n}(t) = -H'_x(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t), \bar{p}_{\epsilon_n}(t)) - \sum_{i=1}^{n_g} \bar{\theta}_{\epsilon_n, i}(t) g'_i(\bar{x}_{\epsilon_n}(t)) \quad (143b)$$

$$- \sum_{i=1}^{n_c} \bar{\eta}_{\epsilon_n, i}(t) c'_{i,x}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t))$$

$$0 = H'_u(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t), \bar{p}_{\epsilon_n}(t)) + \sum_{i=1}^{n_c} \bar{\eta}_{\epsilon_n, i}(t) c'_{i,u}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) \quad (143c)$$

$$0 = \bar{\theta}_{\epsilon_n, i}(t) - g_i(\bar{x}_{\epsilon_n}(t)) - \sqrt{\bar{\theta}_{\epsilon_n, i}(t)^2 + g_i(\bar{x}_{\epsilon_n}(t))^2 + 2\epsilon_n} \quad (143d)$$

$$0 = \bar{\eta}_{\epsilon_n, i}(t) - c_i(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) - \sqrt{\bar{\eta}_{\epsilon_n, i}(t)^2 + c_i(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t))^2 + 2\epsilon_n} \quad (143e)$$

$$0 = h(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T)) \quad (143f)$$

$$0 = \bar{p}_{\epsilon_n}(0) + h'_{x(0)}(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T))^\top \cdot \bar{\lambda}_{\epsilon_n} \quad (143g)$$

$$0 = \bar{p}_{\epsilon_n}(T) - \varphi'(\bar{x}_{\epsilon_n}(T)) - h'_{x(T)}(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T))^\top \cdot \bar{\lambda}_{\epsilon_n} \quad (143h)$$

Then  $(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}, \bar{p}_{\epsilon_n}, \bar{\theta}_{\epsilon_n}, \bar{\eta}_{\epsilon_n})_n$  contains a subsequence converging to a stationary point of the original problem  $(x[\bar{u}, \bar{x}^0], \bar{u}, \bar{p}, \bar{\mu}, \bar{\nu}, \bar{\lambda})$  as follows

$$\begin{aligned} \left\| \bar{u}_{\epsilon_{n_k}} - \bar{u} \right\|_{L^1} \rightarrow 0, \quad \left\| \bar{x}_{\epsilon_{n_k}} - x[\bar{u}, \bar{x}^0] \right\|_{L^\infty} \rightarrow 0, \quad |J(\bar{x}_{\epsilon_{n_k}}, \bar{u}_{\epsilon_{n_k}}) - J(x[\bar{u}, \bar{x}^0], \bar{u})| \rightarrow 0 \\ \left\| \bar{\lambda}_{\epsilon_{n_k}} - \bar{\lambda} \right\| \rightarrow 0, \quad \left\| \bar{p}_{\epsilon_{n_k}} - \bar{p} \right\|_{L^1} \rightarrow 0, \quad \bar{\eta}_{\epsilon_n} \xrightarrow{*} \bar{\nu}, \quad \bar{\theta}_{\epsilon_n} dt \xrightarrow{*} d\bar{\mu} \end{aligned} \quad (144a)$$

*Proof.* From lemma 1 and lemma 3, we have  $g_i(\bar{x}_{\epsilon_n}(t)) < 0$  and  $c_i(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) < 0$  for all  $\epsilon_n > 0$ . Therefore eq. (143d) is equivalent to  $\bar{\theta}_{\epsilon_n, i}(t) = -\epsilon_n/g_i(\bar{x}_{\epsilon_n}(t))$  and eq. (143d) is equivalent to  $\bar{\eta}_{\epsilon_n, i}(t) = -\epsilon_n/c_i(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t))$ . Combining with eqs. (143b) and (143c) proves that any solution of eq. (143) is also solution of eq. (19) and using theorem 1 concludes the proof.  $\square$

Let us denote  $S_{PD}(\epsilon) := (\bar{x}_\epsilon, \bar{u}_\epsilon, \bar{p}_\epsilon, \bar{\theta}_\epsilon, \bar{\eta}_\epsilon, \bar{\lambda}_\epsilon)$  any solution of eq. (143), then the primal-dual algorithm writes as follows.

- 1: Define  $\epsilon_0 > 0$ ,  $\alpha \in (0, 1)$ ,  $\text{tol} = o(1)$ ,  $k = 0$
- 2: **while**  $\epsilon_k > \text{tol}$  **do**
- 3:    $S_{PD}(\epsilon_{k+1}) \leftarrow$  solution of eq. (143) initialized with  $S_{PD}(\epsilon_k)$
- 4:    $\epsilon_{k+1} \leftarrow \alpha \epsilon_k$
- 5:    $k \leftarrow k + 1$
- 6: **end while**
- 7: **return**  $S_{PD}(\epsilon_k)$

**Algorithm 2:** Primal-dual algorithm for optimal control problems

Even though algorithms 1 and 2 are equivalent, the primal-dual algorithm can explore non-admissible trajectories without becoming singular. For example, the primal-dual algorithm can be initialized with non-admissible trajectories and still be numerically tractable which is of course, not the case with the primal algorithm.

## 7 Numerical example: Robbin's problem

The numerical example and the Differential Algebraic Equations (DAEs) solver used in this example are freely available at [https://ifpen-gitlab.appcollaboratif.fr/detocs/ipm\\_ocp](https://ifpen-gitlab.appcollaboratif.fr/detocs/ipm_ocp). The solver is a two point boundary differential algebraic equations solver adapted from [13] to solve index-1 differential algebraic equations.

$$\min_u \int_0^6 x(t) dt \quad (145a)$$

$$x'''(t) = u(t) \quad (145b)$$

$$x(0) = 1 \quad (145c)$$

$$x'(0), x''(0) = 0 \quad (145d)$$

$$0 \geq -x(t) \quad (145e)$$

$$u \in [-1, 1] \quad (145f)$$

This problem is challenging since the optimal solution exhibits a Fuller-like phenomenon both on the adjoint state  $\bar{p}$  and on the control  $\bar{u}$ . For this problem, one can check that conditions eqs. (19e) and (19f) are equivalent to dropping the end initial-final constraint multiplier and add the constraint  $\bar{p}(T) = 0$ .

## 7.1 Resolution using the primal algorithm

The parameterization of the primal algorithm for the Robbins problem is as follows

$$\epsilon_0 = 0.1, \quad \alpha = 0.8, \quad \text{tol} = 10^{-8}, \quad S_P(\epsilon_0)(t) := (1, 0, 0, 0, 0, 0, 0) \quad (146)$$

Using this setting, the execution time is 0.97s.

## 7.2 Resolution using the primal-dual algorithm

The parameterization of the primal-dual algorithm for the Robbins problem is as follows

$$\epsilon_0 = 0.1, \quad \alpha = 0.5, \quad \text{tol} = 10^{-9}, \quad S_{PD}(\epsilon_0)(t) := (1, 0, 0, 0, 0, 0, 0, 0, 0) \quad (147)$$

One can see that the decay rate of the primal-dual method is smaller than the one used in the primal case. Both parameters have been set to the lower value achieving convergence. In addition, the tolerance can also be set lower using the primal-dual version of the algorithm. The execution time with the primal-dual method is 0.20s thanks to the smaller decay rate.

# A Proofs of section 3

## A.1 Proof of proposition 1

From assumption 4,  $x[u, x^0]$  is valued in a compact subset of  $\mathbb{R}^n$ , In addition,  $f$  being  $C^2$  there exists  $\text{const}(f) < +\infty$  such that for all  $(u_1, x_1^0), (u_2, x_2^0) \in L^\infty \times \mathbb{R}^n$

$$\| \dot{x}[u_1, x_1^0](t) - \dot{x}[u_2, x_2^0](t) \| \leq \text{const}(f) (\| x[u_1, x_1^0](t) - x[u_2, x_2^0](t) \| + \| u_1(t) - u_2(t) \|) \quad (148)$$

Using Grönwall inequality [12, Lemma A.1, p.651] again, there exists  $\text{const}(f) < +\infty$  such that  $\| x[u_1, x_1^0] - x[u_2, x_2^0] \|_{L^\infty} \leq \text{const}(f) (\| u_1 - u_2 \|_{L^1} + \| x_1^0 - x_2^0 \|)$ .

## A.2 Proof of proposition 2

To alleviate the notation, we denote  $x_n := x[u_n, x_n^0]$  and  $\bar{x} := x[\bar{u}, \bar{x}^0]$ . using these notations, we have

$$x_n(t_2) - x_n(t_1) := \int_{t_1}^{t_2} f_1(x_n(t)) + f_2(x_n(t)) \cdot u_n(t) dt \quad (149)$$

First,  $\forall t_1, t_2 \in [0, T]$ , From Hölder inequality, we have

$$\| x_n(t_2) - x_n(t_1) \| \leq \sup_n \| f_1(x_n) + f_2(x_n) \cdot u_n \|_{L^2} \sqrt{|t_1 - t_2|} \quad (150)$$

Therefore, the sequence  $(x_n)_n$  is bounded and equicontinuous. From Arzela-Ascoli [14, Theorem 1.3.8, p.33], it contains a uniformly converging subsequence to some  $\hat{x}$ . Let  $(x_k)_k$  be the uniformly converging sequence, one has

$$\hat{x}(t) - \bar{x}(t) = \lim_{k \rightarrow +\infty} x_k(t) - \bar{x}(t) = \int_0^t f_1(\hat{x}(s)) - f_1(\bar{x}(s)) + (f_2(\hat{x}(s)) - f_2(\bar{x}(s))) \cdot \bar{u}(s) ds \quad (151)$$

Since  $\hat{x}(0) = \bar{x}^0$ , the term inside the integral in eq. (151) is always zero, thus  $\hat{x} = \bar{x}$ . Now, let us prove that the whole sequence  $(x_n)_n$  uniformly converges to  $\bar{x}$  by contradiction. Assume that, there exists a subsequence  $(x_k)_k$  such that  $\exists K > 0$  and  $\epsilon > 0$  satisfying  $\|x_k - \bar{x}\|_{L^\infty} \geq \epsilon$  for all  $k \geq K$ . One can extract a sub-subsequence  $(x_{k_j})_j$  uniformly converging to some  $x_1$  with  $\|x_1 - \bar{x}\|_{L^\infty} > 0$ . However,  $V^{\text{ad}}$  being weakly compact in the topology  $\sigma(L^2 \times \mathbb{R}^n, L^2 \times \mathbb{R}^n)$ , one can extract from  $(u_{k_j}, x_{k_j}^0)_j$  a weakly convergent subsequence. By definition, this sequence weakly converges to  $(\bar{u}, \bar{x}^0)$  and proves that  $(x_{k_j})_j$  contains a subsequence converging to  $\bar{x}$  which contradicts the initial assumption and proves the uniform convergence of  $(x_n)_n$ . As a consequence,  $\alpha(x_n)$  and  $\beta(x_n)$  uniformly converges to  $\alpha(\bar{x})$  and  $\beta(\bar{x})$  respectively. In addition, the sequence  $(\alpha(x_n).u_n + \beta(x_n))_n$  is uniformly  $L^\infty$ -bounded, thus contains a weakly  $*$  converging subsequence to some  $\theta$ . Assume that  $\theta$  is not equal to  $\alpha(\bar{x}).\bar{u} + \beta(\bar{x})$  and Let  $\varphi \in L^2([0, T]; \mathbb{R}^{n_c}) \cap L^1([0, T]; \mathbb{R}^{n_c})$  we have

$$\langle \varphi, \theta - \alpha(\bar{x}).\bar{u} - \beta(\bar{x}) \rangle = \lim_{k \rightarrow +\infty} \langle \varphi, \alpha(x_{n_k}).u_{n_k} + \beta(x_{n_k}) - \alpha(\bar{x}).\bar{u} - \beta(\bar{x}) \rangle = \lim_{k \rightarrow +\infty} \langle \varphi, \alpha(\bar{x})(u_{n_k} - \bar{u}) \rangle = 0 \quad (152)$$

Since  $L^2([0, T]; \mathbb{R}^{n_c}) \cap L^1([0, T]; \mathbb{R}^{n_c})$  is dense in  $L^1([0, T]; \mathbb{R}^{n_c})$  this contradicts the initial assumption and proves that the weak  $*$  limit is  $\alpha(\bar{x}).\bar{u} + \beta(\bar{x})$ . To prove that the whole sequence weakly  $*$  converges, we use the same argument as the one we used to prove uniform convergence of the state, which concludes the proof.

### A.3 Proof of proposition 3

Using assumption 4,  $\forall u \in U$  we have

$$|g_i(x[u](t)) - g_i(x[u](s))| \leq \text{const}(g) \|x[u](t) - x[u](s)\| \leq \text{const}(f, g)|t - s| \quad (153)$$

To prove the proposition, we only need to prove the lower bound holds on any interval  $(\alpha_1, \alpha_2) \subseteq E$ . From the continuity of  $g_i$ ,  $\exists t_1, t_2$  such that  $g_i(x[u](t_1)) = \alpha_1$ ,  $g_i(x[u](t_2)) = \alpha_2$  and such that  $(t_1, t_2) \subseteq g_i(x[u])^{-1}((\alpha_1, \alpha_2))$  and

$$\begin{aligned} m[u, g_i]((\alpha_1, \alpha_2)) &\geq |t_1 - t_2| \geq \text{const}(f, g) |g_i(x[u](t_1)) - g_i(x[u](t_2))| \\ &\geq \text{const}(f, g) |\alpha_1 - \alpha_2| \end{aligned} \quad (154)$$

### A.4 Proof of proposition 4

Let  $\delta > 0$ , and for all  $(u, x^0) \in V_\infty^{\text{ad}}$  let us denote

$$\gamma_\delta(u, x^0) := \inf_{v \in B_{\|\cdot\|_{V^{\text{ad}}}}((u, x^0), \delta) \cap V^{\text{ad}}} \left\{ \sup_t g(x[v, y^0](t)) \right\} \quad (155)$$

From assumption 2, we have  $\gamma_\delta(u, x^0) < 0$ . Then,  $\forall (u, x^0) \in V_\infty^{\text{ad}}$ ,  $\exists (v, y^0) \in B_{\|\cdot\|_{V^{\text{ad}}}}((u, x^0), \delta) \cap V^{\text{ad}}$  such that

$$\sup_t g(x[v, y^0](t)) \leq \gamma_\delta(u, x^0) \leq \sup_{(u, x^0) \in V_\infty^{\text{ad}}} \gamma_\delta(u, x^0) := -2G_\delta < 0 \quad (156)$$

In addition, if  $S_{u, x^0}^g(G_\delta) \neq \emptyset$ , then  $\forall t \in S_{u, x^0}^g(G_\delta)$  we have

$$g(x[v, y^0](t)) - g(x[u, x^0](t)) \leq -2G_\delta + G_\delta = -G_\delta \quad (157)$$

Now, let us denote

$$\kappa_\delta(u, x^0) := \inf_{v \in B_{\|\cdot\|_{V^{\text{ad}}}}((u, x^0), \delta) \cap V^{\text{ad}}} \left\{ \text{ess sup}_t c(x[v, y^0](t), v(t)) \right\} \quad (158)$$



From assumption 2, for all  $(u, x^0) \in V_\infty^{\text{ad}}$ ,  $\exists (v_n, y_n^0)_n \in \hat{V}^{\text{ad}}(n)$  converging to  $(u, x^0)$ . Thus, for all  $\delta > 0$ ,  $\exists N_\delta[u, x^0] > 0$ , such that  $\forall n \geq N_\delta[u, x^0]$ , we have  $(v_n, y_n^0) \in B_{\|\cdot\|_{V^{\text{ad}}}}((u, x^0), \delta)$ , which yields  $\kappa_\delta(u, x^0) \leq -1/N_\delta(u, x^0)$ . Now,  $\forall (u, x^0) \in V_\infty^{\text{ad}}$ ,  $\exists (v, y^0) \in B_{\|\cdot\|_{V^{\text{ad}}}}((u, x^0), \delta) \cap V^{\text{ad}}$  such that

$$\operatorname{ess\,sup}_t c(x[v, y^0](t), v(t)) \leq \kappa_\delta(u, x^0) \leq \sup_{(u, x^0) \in V_\infty^{\text{ad}}} -\frac{1}{N_\delta[u, x^0]} := -2C_\delta < 0 \quad (159)$$

In addition, if  $S_{u, x^0}^c(C_\delta) \neq \emptyset$ , then  $\forall t \in S_{u, x^0}^c(C_\delta)$  we have

$$c(x[v, y^0](t), v(t)) - c(x[u, x^0](t), u(t)) \leq -2C_\delta + C_\delta = -C_\delta \quad (160)$$

## References

- [1] J.T. Betts. *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, Philadelphia, PA, 2005.
- [2] R. Bonalli, B. Hérissé, and E. Trélat. Analytical initialization of a continuation-based indirect method for optimal control of endo-atmospheric launch vehicle systems. *IFAC-PapersOnLine*, 50(1):482–487, 2017.
- [3] J.F. Bonnans, Xavier Dupuis, and Laurent Pfeiffer. Second-order sufficient conditions for strong solutions to optimal control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 20(3):704–724, 2014.
- [4] J.F. Bonnans and T. Guilbaud. Using logarithmic penalties in the shooting algorithm for optimal control problems. *Optimal Control Applications and Methods*, 24:257–278, 2003.
- [5] 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 Journal on Control and Optimization*, 46(4):1398–1430, 2007.
- [6] 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(2):561–598, 2009.
- [7] A.E. Bryson and Y.C. Ho. *Applied Optimal Control*. Taylor & Francis, 1975.
- [8] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.
- [9] K. Graichen and N. Petit. Incorporating a class of constraints into the dynamics of optimal control problems. *Optimal Control Applications and Methods*, 30:537–561, 2009.
- [10] T. Haberkorn and E. Trélat. Convergence results for smooth regularizations of hybrid nonlinear optimal control problems. *SIAM Journal on Control and Optimization*, 49(4):1498–1522, 2011.
- [11] R. F. Hartl, Suresh Sethi, and Raymond Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37(2):181–218, 1995.
- [12] H. Khalil. *Non Linear Systems*. Prentice Hall, 2002.
- [13] Jacek Kierzenka and Lawrence F. Shampine. A bvp solver based on residual control and the matlab pse. *ACM Trans. Math. Softw.*, 27:299–316, 2001.

- [14] A.J. Kurdila and M. Zabaranin. *Convex Functional Analysis*. Birkhäuser Boston, 2005.
- [15] L. Lasdon, A. Waren, and R. Rice. An interior penalty method for inequality constrained optimal control problems. *IEEE Transactions on Automatic Control*, 12:388–395, 1967.
- [16] P. Malisani, F. Chaplais, and N. Petit. An interior penalty method for optimal control problems with state and input constraints of nonlinear systems. *Optimal Control Applications and Methods*, 37:3–33, 2014.
- [17] A. Maurer and J. Zowe. First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming*, 16:98–110, 1979.
- [18] K. Namwook, C. Sukwon, and P. Huei. Optimal control of hybrid electric vehicles based on pontryagin’s minimum principle. *IEEE Transactions on Control Systems Technology*, 19(5):1279–1287, 2011.
- [19] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2000.
- [20] M. Renardy and R. Rogers. *An introduction to partial differential equations*, volume 13 of *Texts in applied mathematics*. Springer, New York, second edition edition, 2004.
- [21] H. Robbins. Junction phenomena for optimal control with state-variable inequality constraints of third order. *Journal of Optimization Theory and Applications*, 31:85–99, 1980.
- [22] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition, 1976.
- [23] Hans Seywald and Eugene M Cliff. Goddard problem in presence of a dynamic pressure limit. *Journal of Guidance, Control, and Dynamics*, 16(4):776–781, 1993.
- [24] K. Soetaert, J. Cash, F. Mazzia, Ascher U.M., G. Bader, J. Christiansen, and R.R. Russel. Solvers for boundary value problems of differential equations. <https://cran.r-project.org/web/packages/bvpSolve/index.html>.
- [25] R.F. Stengel, R. Ghigliazza, Kulkarni N., and O. Laplace. Optimal control of innate immune response. *Optimal Control Applications and Methods*, 23:91–104, 2002.
- [26] M. Weiser. Interior point methods in function space. *SIAM Journal on Control and Optimization*, 44(5):1766–1786, 2005.
- [27] S.J. Wright. *Primal-Dual Interior-Point Methods*. SIAM publications, 1997.
- [28] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106:25–57, 2006.