



**HAL**  
open science

## **Sensitivity analysis for relaxed optimal control problems with final-state constraints**

Joseph Frederic Bonnans, Laurent Pfeiffer, Oana Silvia Serea

► **To cite this version:**

Joseph Frederic Bonnans, Laurent Pfeiffer, Oana Silvia Serea. Sensitivity analysis for relaxed optimal control problems with final-state constraints. *Nonlinear Analysis: Theory, Methods and Applications*, 2013, 89, pp.55-80. <10.1016/j.na.2013.04.013>. <hal-00702246v2>

**HAL Id: hal-00702246**

**<https://inria.hal.science/hal-00702246v2>**

Submitted on 22 Jan 2013

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

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



HAL Authorization



# Sensitivity analysis for relaxed optimal control problems with final-state constraints

J. Frédéric Bonnans, Laurent Pfeiffer, Oana Silvia Serea

**RESEARCH  
REPORT**

**N° 7977**

May 2012

Project-Team Commands





## Sensitivity analysis for relaxed optimal control problems with final-state constraints

J. Frédéric Bonnans\*, Laurent Pfeiffer†, Oana Silvia Serea‡

Project-Team Commands

Research Report n° 7977 — version 2 — initial version May 2012 —  
revised version January 2013 — 35 pages

**Abstract:** In this article, we compute a second-order expansion of the value function of a family of relaxed optimal control problems with final-state constraints, parameterized by a perturbation variable. The sensitivity analysis is performed for controls that we call  $R$ -strong solutions. They are optimal solutions with respect to the set of feasible controls with a uniform norm smaller than a given  $R$  and having an associated trajectory in a small neighborhood for the uniform norm. In this framework, relaxation enables us to consider a wide class of perturbations and therefore to derive sharp estimates of the value function.

**Key-words:** optimal control, sensitivity analysis, relaxation, Young measures, Pontryagin's principle, strong solutions.

---

\* INRIA-Saclay and CMAP, Ecole Polytechnique, 91128 Palaiseau Cedex, France. (frederic.bonnans@inria.fr)

† INRIA-Saclay and CMAP, Ecole Polytechnique, 91128 Palaiseau Cedex, France.  
(laurent.pfeiffer@polytechnique.edu)

‡ Université de Perpignan Via Domitia, Laboratoire de Mathématiques et de Physique, EA 4217, 52 avenue Paul Alduy, 66860 Perpignan Cedex, France. (oana-silvia.serea@univ-perp.fr)

**RESEARCH CENTRE  
SACLAY – ÎLE-DE-FRANCE**

1 rue Honoré d'Estienne d'Orves  
Bâtiment Alan Turing  
Campus de l'École Polytechnique  
91120 Palaiseau

## Analyse de sensibilité pour des problèmes de contrôle optimal relaxés avec contraintes sur l'état final

**Résumé :** Dans cet article, nous calculons un développement au second ordre de la fonction valeur d'une famille de problèmes de contrôle optimal avec contraintes sur l'état final, paramétrée par une variable de perturbation. L'analyse de sensibilité est réalisée pour des contrôles nommés *R-strong solutions*. Ce sont des solutions optimales par rapport à l'ensemble des contrôles admissibles de norme infinie inférieure à  $R$  ayant une trajectoire associée dans un petit voisinage pour la norme infinie. Dans ce cadre, la relaxation nous permet de considérer une large classe de perturbations et ainsi d'obtenir des estimations précises de la fonction valeur.

**Mots-clés :** contrôle optimal, analyse de sensibilité, relaxation, mesures de Young, principe de Pontryagin, solutions fortes.

## 1 Introduction

We consider a family of relaxed optimal control problems with final-state constraints, parameterized by a perturbation variable  $\theta$ . The variable  $\theta$  can perturb the dynamic of the system, the cost function and the final-state constraints. The aim of the article is to compute a second-order expansion of the value  $V(\theta)$  of the perturbed problems, in the neighborhood of a reference value of  $\theta$ , say  $\bar{\theta}$ .

This second-order expansion is obtained by applying the methodology described in [5] and originally in [3]. The approach is the following: we begin by linearizing the family of optimization problems in the neighborhood of an optimal solution of the reference problem. The first-order and second-order linearizations provide a second-order upper estimate of the value function. Then, a first lower estimate is obtained by expanding the Lagrangian up to the second order. Considering a strong sufficient second-order condition, we show that the distance between the reference solution and solutions to the perturbed problems is of order  $|\theta - \bar{\theta}|$ . Finally, the lower estimate corresponds to the upper estimate previously obtained.

The sensitivity analysis is performed in the framework of relaxed optimal controls. Roughly speaking, at each time, the control variable is not anymore a vector in a space  $U$ , but a probability measure on  $U$ , like if we were able to use several controls simultaneously. The new control variable is now a Young measure, in reference to the pioneering work of Young [20]. Relaxation of optimal control problems with Young measures has been much studied, in particular in [8, 11, 18, 19, 20]. Any Young measure is the weak-\* limit of a sequence of classical controls, therefore, we expect that a classical optimal control problem and its relaxed version have the same value. This question is studied, for instance, in [2, 9].

Three aspects motivate the use of the relaxation. First, by considering convex combinations of controls in the sense of measures, we manage to describe in a convenient way a large class of tangential directions of the reachable set. This class of tangential directions was called *cone of variations* in the early papers of McShane [11], Gamkrelidze [10] and Warga [17, 18]. It enables to prove Pontryagin's principle with the standard methods used to derive first-order optimality conditions of optimization problems. Secondly, in this framework, we derive a metric regularity theorem for the  $L^1$ -distance using abstract results from [7]. Finally, the existence of relaxed solutions for the perturbed problem is guaranteed. Note that such solutions do not always exist in a classical framework.

The sensitivity analysis is realized locally, in a neighborhood of a local optimal solution of the reference problem. In this study, we use the notion of relaxed  $R$ -strong optimal controls. We say that a control is a relaxed  $R$ -strong optimal solution if it is optimal with respect to the Young measures having their support in a ball of radius  $R$  and having a state variable sufficiently close for the uniform norm. This notion is related to the one of bounded strong solutions [12]. In order to obtain a sharp upper estimate of  $V$ , we must derive a linearized problem from a wide class of perturbations of the control. More precisely, we must be able to perturb the reference optimal control with close controls for the  $L^1$ -distance, taking into account that they are usually not necessarily close for the  $L^\infty$ -distance. For such perturbations of the control, we use a particular linearization of the dynamics of the system, the Pontryagin linearization [12].

We obtain a lower estimate of the value function by assuming a sufficient second-order condition having the same nature as the one in [4]. We assume that a certain quadratic form is positive and that the Hamiltonian satisfies a quadratic growth condition. In order to expand the Lagrangian up to the second-order, we split the controls into two parts, one accounting for the small of the control in  $L^\infty$ -distance and the other one accounting for the large variations. We obtain an extension of the decomposition principle described in [4] and a lower estimate which corresponds to the upper estimate obtained previously.

The outline of the paper is as follows. In section 2, we prove some preliminary results and in particular, a metric regularity theorem. Note that we will always suppose that the associated qualification condition holds. In section 3, we obtain a first-order upper estimate of  $V$  and in section 4 a second-order upper estimate. In section 5, we prove the decomposition principle and we obtain the lower estimate. Two examples are discussed in section 6. The theoretical material related to Young measures is recalled in section A of the appendix, with precise references from [1, 6, 15, 16]. In section B, we justify the use of relaxation. Some technical proofs are presented in section C for completeness.

## 2 Formulation of the problem and preliminary results

### 2.1 Setting

In this section, we define the family of optimal control problems that we want to study. We also introduce the notion of *relaxed  $R$ -strong* solutions.

In the article, the perturbation parameter will be denoted by  $\theta$ . A reference value of  $\theta$ , say  $\bar{\theta}$  is given. We restrict ourselves to the case where  $\theta$  is nonnegative and  $\bar{\theta} = 0$ . Consider the control and state spaces

$$\mathcal{U} := L^\infty(0, T; \mathbb{R}^m), \quad \mathcal{Y} := W^{1, \infty}(0, T; \mathbb{R}^n), \quad (2.1)$$

and the state equation, defined in a classical framework, for the moment:

$$\begin{cases} \dot{y}_t = f(u_t, y_t, \theta), & \text{for a.a. } t \in [0, T], \\ y_0 = y^0. \end{cases} \quad (2.2)$$

For a control  $u$  in  $\mathcal{U}$  and  $\theta \geq 0$ , we denote by  $y[u, \theta]$  the trajectory satisfying the differential system (2.2). We consider the following final state constraint:

$$\Phi(y_T, \theta) \in K, \quad K := \{0\}^{n_E} \times \mathbb{R}_-^{n_I} \subset \mathbb{R}^{n_C}, \quad (2.3)$$

with  $n_C = n_E + n_I$ . The general family of optimal control problems that we consider is the following:

$$\text{Min}_{u \in \mathcal{U}} \phi(y_T[u, \theta], \theta), \quad \text{s.t. } \Phi(y_T[u, \theta], \theta) \in K. \quad (2.4)$$

All introduced functions ( $f$ ,  $\phi$ , and  $\Phi$ ) are supposed to be  $C^{2,1}$  (twice differentiable with a Lipschitz second-order derivative).

In this general setting, it is not possible to perform a sensitivity analysis of the global problem. Instead, we are interested in the local behavior of the solutions of the family of problems, in the neighborhood of a local solution of the reference problem (with  $\theta = 0$ ). Let us make clear the notion of local optimality which is used. From now on, we fix a control  $\bar{u} \in \mathcal{U}$  and its associated trajectory  $\bar{y} = y[\bar{u}, 0]$ .

**Definition 1.** Let  $R > 0$ , the control  $\bar{u}$  is said to be an  $R$ -strong optimal solution if there exists  $\eta > 0$  such that  $\bar{u}$  is solution to the following localized reference problem:

$$\text{Min}_{u \in \mathcal{U}, \|u\|_\infty \leq R} \phi(y_T[u, 0], 0), \quad \text{s.t. } \Phi(y_T[u, 0], 0) \in K, \quad \|y[u, 0] - \bar{y}\|_\infty \leq \eta. \quad (2.5)$$

Note that the control  $\bar{u}$  is a *bounded strong* solution if for all  $R > \|\bar{u}\|_\infty$ , it is an  $R$ -strong optimal solution [12, page 291].

Let us consider now a relaxed version of this definition. Let us denote by  $U_R$  the closed ball of radius  $R$  and center 0 in  $\mathbb{R}^m$ . We denote by  $\mathcal{M}_R^Y$  the space of Young measures on  $[0, T] \times U_R$ . The basic idea of relaxation is to replace *classical* controls in  $\mathcal{U}$  by Young measures, that we will call *relaxed* controls, like if we were able to take several decisions simultaneously at each time. The basic definitions related to Young measures are recalled in section A.1 of the appendix.

The dynamic associated with a Young measure  $\mu$  in  $\mathcal{M}_R^Y$  is the following:

$$\begin{cases} \dot{y}_t = \int_{U_R} f(u, y_t, \theta) d\mu_t(u), & \text{for a.a. } t \in [0, T], \\ y_0 = y^0. \end{cases} \quad (2.6)$$

This definition is compatible with (2.2) for controls in  $\mathcal{U}$ . We extend the mapping  $y[\mu, \theta]$  to Young measures and we say that  $\mu \in \mathcal{M}_R^Y$  is *feasible* for the value  $\theta$  if

$$\Phi(y_T[\mu, \theta], \theta) \in K. \quad (2.7)$$

From now on, we denote by  $\bar{\mu}$  the Young measure associated with  $\bar{u}$ .

**Definition 2.** Let  $R > 0$ , the relaxed control  $\bar{\mu}$  is said to be a relaxed  $R$ -strong optimal solution if there exists  $\eta > 0$  such that  $\bar{\mu}$  is solution to the following relaxed localized reference problem:

$$\text{Min}_{\mu \in \mathcal{M}_R^Y} \phi(y_T[\mu, 0], 0), \quad \text{s.t. } \Phi(y_T[\mu, 0], 0) \in K, \quad \|y[\mu, 0] - \bar{y}\|_\infty \leq \eta. \quad (2.8)$$

From now on, we suppose that  $\bar{\mu}$  is a relaxed  $R$ -strong optimal solution for the value  $\bar{\eta}$ . The relaxed optimal control problems that we will study are the following:

$$V^\eta(\theta) := \begin{cases} \text{Min}_{\mu \in \mathcal{M}_R^Y} & \phi(y_T[\mu, \theta], \theta), \\ \text{s.t.} & \Phi(y_T[\mu, \theta], \theta) \in K, \quad \|y[\mu, \theta] - \bar{y}\|_\infty \leq \eta. \end{cases} \quad (\mathcal{P}_\theta^{Y, \eta})$$

**Remark 3.** Note also that  $\eta$  is not fixed. For all  $0 < \eta \leq \eta'$ , for all  $\theta \geq 0$ ,

$$V^{\eta'}(\theta) \leq V^\eta(\theta).$$

By assumption, for all  $\eta \in (0, \bar{\eta}]$ ,  $V^\eta(0) = V^{\bar{\eta}}(0)$ . The role of  $\eta$  in the study is secondary, but it cannot be neglected. Indeed, all the results related to upper estimates (lemma 15 and theorem 27) are satisfied for all  $\eta \in (0, \bar{\eta}]$ . In section 5, the second-order sufficient condition ensures that for small, positive, and fixed values of  $\eta$ , there exist solutions  $\mu^\theta$  of  $(\mathcal{P}_\theta^{Y, \eta})$ , converging to  $\bar{\mu}$  for the  $L^2$ -distance (theorem 35). Thus, the associated trajectories converges uniformly. This proves that for small values of  $\eta > 0$ , for all  $0 < \eta' < \eta$ ,  $V^\eta$  and  $V^{\eta'}$  coincide on a neighborhood of 0.

## 2.2 Estimates

In our study, the addition of Young measures must be understood as the addition of measures on  $[0, T] \times U_R$ . With this definition of the addition, the space  $\mathcal{M}_R^Y$  is convex. The following lemma is a corollary of lemma 7. The distance  $d_1$  is the Wasserstein distance, defined by (A.1).

**Lemma 4.** Let  $\mu^0$  and  $\mu^1$  in  $\mathcal{M}_R^Y$ , let  $\sigma$  in  $[0, 1]$ . Then,

$$d_1(\mu^0, (1 - \sigma)\mu^0 + \sigma\mu^1) \leq \sigma d_1(\mu^0, \mu^1) \leq 2RT\sigma.$$

In the sequel, we use the notation  $g[t] := g(\bar{u}_t, \bar{y}_t, 0)$  for every function  $g$  of  $(u, y, \theta)$ . The following definition of the Pontryagin linearization is a particular linearization of the state equation. Indeed, we only linearize the dynamic with respect to the state variable. We extend the definition of [12, page 40] to Young measures.

**Definition 5.** For a given control  $\mu$ , we define the Pontryagin linearization  $\xi[\mu]$  in  $\mathcal{Y}$  as the solution of

$$\begin{cases} \dot{\xi}_t[\mu] = f_y[t]\xi_t[\mu] + \int_{U_R} f(u, \bar{y}_t, 0) d\mu_t(u) - f[t], & \text{for a.a. } t \in [0, T], \\ \xi_0[\mu] = 0. \end{cases}$$

Denote by  $\xi^\theta$  the solution of the following differential system:

$$\begin{cases} \dot{\xi}_t^\theta = f_y[t]\xi_t^\theta + f_\theta[t], & \text{for a.a. } t \in [0, T], \\ \xi_0^\theta = 0. \end{cases}$$

**Lemma 6.** The following estimates hold:

$$\|y[\mu, \theta] - \bar{y}\|_\infty = O(d_1(\mu, \bar{\mu}) + \theta), \quad (2.9)$$

$$\|y[\mu, \theta] - (\bar{y} + \xi[\mu] + \theta\xi^\theta)\|_\infty = O(d_1(\mu, \bar{\mu})^2 + \theta^2). \quad (2.10)$$

This lemma is proved in the appendix, page 28.

### 2.3 Metric regularity

For  $q \in \mathbb{N}^*$ , we set  $\Delta := \left\{ \gamma \in \mathbb{R}_+^q, \sum_{i=1}^q \gamma_i \leq 1 \right\}$ . Given  $\mu^1, \dots, \mu^q \in \mathcal{M}_R^Y$ , we denote by  $S$  the following mapping:

$$S : (\mu^0, \gamma) \in (\mathcal{M}_R^Y \times \Delta) \mapsto \left(1 - \sum_{i=1}^q \gamma_i\right) \mu^0 + \sum_{i=1}^q \gamma_i \mu^i \in \mathcal{M}_R^Y. \quad (2.11)$$

**Lemma 7.** Let  $\mu, \mu' \in \Delta$ , let  $\mu^0 \in \mathcal{M}_R^Y$ . Then,

$$d_1(S(\mu^0, \gamma), S(\mu^0, \gamma')) \leq \sum_{i=1}^q |\gamma'_i - \gamma_i| d_1(\mu^i, \mu^0) \leq 2RT \sum_{i=1}^q |\gamma'_i - \gamma_i|.$$

This lemma is proved in the appendix, page 29. We introduce the following set:

$$\mathcal{R}_T := \{\xi_T[\mu], \mu \in \mathcal{M}_R^Y\}. \quad (2.12)$$

The Pontryagin linearization being affine with respect to  $\mu$ ,  $\mathcal{R}_T$  is clearly convex. We denote by  $\mathcal{C}(\mathcal{R}_T)$  the smallest closed cone containing  $\mathcal{R}_T$ . Since  $\mathcal{R}_T$  is convex,  $\mathcal{C}(\mathcal{R}_T)$  is also convex. This set should be understood as a set of tangential directions of the reachable set (at the final time). It is a close object to the cone of variations described in [10, page 121], [18, page 132] and [11, page 457].

**Definition 8 (Qualification).** The control  $\bar{\mu}$  is qualified if there exists  $\varepsilon > 0$  such that

$$\varepsilon B \subset \Phi(\bar{y}_T, 0) + \Phi_{y_T}(\bar{y}_T, 0) \mathcal{C}(\mathcal{R}_T) - K, \quad (2.13)$$

where  $B$  is the unit ball of  $\mathbb{R}^{n_c}$ .

In the sequel, we will always assume that  $\bar{\mu}$  is qualified. Note that in remark 21, we will show that this assumption is weaker than the standard qualification assumption. The following theorem establishes a property of metric regularity for the relaxed problem. The main elements of the proof of the theorem can be found in [18, lemma 3.1].

**Theorem 9.** *If  $\bar{\mu}$  is qualified, then there exist  $\delta > 0$ ,  $\sigma > 0$ , and  $C \geq 0$  such that for all  $\theta \in [0, \sigma]$ , for all  $\mu$  satisfying  $d_1(\mu, \bar{\mu}) \leq \delta$ , there exists a control  $\mu'$  satisfying*

$$\Phi(y_T[\mu', \theta], \theta) \in K \quad \text{and} \quad d_1(\mu', \mu) \leq C \operatorname{dist}(\Phi(y_T[\mu, \theta], \theta), K). \quad (2.14)$$

*Proof.*  $\triangleright$  *First step: metric regularity of  $G_{\bar{\mu}, 0}$ .*

If (2.13) holds, it can be proved that there exists a family  $(\xi^i)_i$ ,  $i = 1, \dots, n_A$  (with  $n_A \leq n_C + 1$ ) in  $\mathcal{C}(\mathcal{R}_T)$  such that for some  $\varepsilon_1 > 0$ ,

$$\varepsilon_1 B \subset \Phi(\bar{y}_T, 0) + \Phi'(\bar{y}_T, 0)(\operatorname{conv}\{\xi_T[\mu^1], \dots, \xi_T[\mu^{n_A}]\}) - K.$$

Using the mapping  $S$  defined by (2.11), with  $q = n_A$ , we consider the mapping

$$G_{\mu, \theta} : \gamma \in \Delta \mapsto \Phi(y_T[S(\mu, \gamma), \theta], \theta) \in \mathbb{R}^{n_C}, \quad (2.15)$$

defined for all  $(\mu, \theta)$  in  $\mathcal{M}_R^Y \times \mathbb{R}_+$ . Note that  $G_{\mu, \theta}(0_{n_A}) = \Phi(y_T[\mu, \theta], \theta)$ .

Let us fix  $\mu, \theta$ . It can be shown that  $\gamma \mapsto G_{\mu, \theta}(\gamma)$  is differentiable on  $\Delta$  in the following sense: there exists a unique mapping  $\gamma \in \Delta \mapsto G'_{\mu, \theta}(\gamma) \in \mathbb{R}^{n_A}$  such that, for all  $\gamma, \gamma' \in \Delta$ ,

$$G_{\mu, \theta}(\gamma') = G_{\mu, \theta}(\gamma) + G'_{\mu, \theta}(\gamma)(\gamma' - \gamma) + o(|\gamma' - \gamma|).$$

By lemma 37, the mapping  $(\mu, \theta) \in \mathcal{M}_R^Y \times \mathbb{R}_+ \mapsto G'_{\mu, \theta}(\cdot) \in L^\infty(\Delta, \mathbb{R}^{n_A})$  is continuous (for the  $L^1$ -distance of  $\mathcal{M}_R^Y$ ) and

$$G'_{\bar{\mu}, 0}(0_{n_A})\delta\gamma = \Phi'(\bar{y}_T, 0)\left(\sum_{i=1}^{n_A} \xi_T[\mu^i]\delta\gamma_i\right). \quad (2.16)$$

An explicit formula for  $G'_{\mu, \theta}$  can be obtained with the Pontryagin linearization, see e.g. [18, equation 3.1.6]. Using (2.16), we obtain that

$$G_{\bar{\mu}, 0}(0_{n_A}) + G'_{\bar{\mu}, 0}(0_{n_A})\Delta = \Phi(\bar{y}_T, 0) + \Phi'(\bar{y}_T, 0)(\operatorname{conv}\{\xi_T[\mu^1], \dots, \xi_T[\mu^{n_A}]\}).$$

Therefore, by the Robinson-Ursescu stability theorem (see e.g. [13, 14] and also [5, theorem 2.87]),  $G_{\bar{\mu}, 0}$  is metric regular with respect to  $K$  at  $0_{n_A}$  with a constant  $C_1 > 0$  (in the sense of [5, relation (2.165)]).

$\triangleright$  *Second step: metric regularity of  $G_{\mu, \theta}$ .*

Moreover, there exist a neighborhood  $\mathcal{O}^\mu$  of  $\bar{\mu}$  (for the  $L^1$ -distance),  $\sigma > 0$ , and a neighborhood  $\mathcal{O}^\gamma$  of  $0_{n_A}$  such that for all  $(\mu, \theta, \gamma)$  in  $\mathcal{O}^\mu \times [0, \sigma] \times (\mathcal{O}^\gamma \cap \Delta)$ ,  $|G'_{\mu, \theta}(\gamma) - G'_{\bar{\mu}, 0}(\gamma)| \leq \frac{C_1}{2}$ . By [5, theorem 2.84], the whole family of functions  $G_{\mu, \theta}$  is metric regular at  $0_{n_A}$ , for all  $\mu \in \mathcal{O}^\mu$  and all  $\theta \in [0, \sigma]$ . It means in particular that there exists a constant  $C_2 \geq 0$  which is such that for all  $\mu \in \mathcal{O}^\mu$  and all  $\theta \in [0, \sigma]$ ,

$$\operatorname{dist}(0_{n_A}, G_{\mu, \theta}^{-1}(K)) \leq C_2 \operatorname{dist}(G_{\mu, \theta}(0_{n_A}), K).$$

$\triangleright$  *Third step: proof of the theorem.*

Let  $(\mu, \theta)$  in  $\mathcal{O}^\mu \times [0, \sigma]$ , since  $G_{\mu, \theta}(0_{n_A}) = \Phi(y_T[\mu, \theta], \theta)$ , there exists  $\tilde{\gamma}$  in  $G_{\mu, \theta}^{-1}(K)$  such that

$$|\tilde{\gamma}| \leq C \operatorname{dist}(G_{\mu, \theta}(0_{n_A}), K) = C \operatorname{dist}(\Phi(y_T[\mu, \theta], \theta), K).$$

Finally, we set  $\mu' = S(\mu, \tilde{\gamma})$ . This control satisfies the final-state constraint and by lemma 7,  $d_1(\mu', \mu) \leq 2RT|\tilde{\gamma}|$ . Restricting  $\mathcal{O}^\mu$  to a ball (for the  $L^1$ -distance) of radius  $\delta$  and center  $\bar{\mu}$ , we obtain the theorem with  $\delta, \sigma, \mu'$ , and  $C = 2RTC_2$ .  $\square$

**Corollary 10.** *For all  $\eta > 0$ , there exists  $\tilde{\theta} > 0$  such that for all  $\theta \in [0, \tilde{\theta}]$ , problem  $(\mathcal{P}_\theta^{Y,\eta})$  has an optimal solution.*

*Proof.* Let  $\eta > 0$ . As a consequence of the compactness of  $\mathcal{M}_R^Y$  and the weak-\* continuity of  $\mu \mapsto y[\mu, \theta]$ , every minimizing sequence has a limit point which is a solution to problem  $(\mathcal{P}_\theta^{Y,\eta})$ . Therefore, for  $\theta$  sufficiently small, we only need to prove the existence of a feasible control  $\mu$  satisfying  $\|y[\mu, \theta] - \bar{y}\|_\infty \leq \eta$ . For all  $\theta \in [0, \sigma]$ ,

$$\text{dist}(\Phi(y_T[\bar{u}, \theta], \theta), K) = O(\theta), \quad (2.17)$$

therefore, by theorem 9, there exists a feasible control  $\mu^\theta$  such that  $d_1(\bar{\mu}, \mu^\theta) = O(\theta)$ . By lemma 37,  $\|y[\mu^\theta, \theta] - \bar{y}\|_\infty = O(\theta)$ , therefore, for  $\theta$  sufficiently small,  $\|y[\mu, \theta] - \bar{y}\|_\infty \leq \eta$ . The corollary is now proved.  $\square$

## 2.4 Optimality conditions

We introduce now the *Hamiltonian* function  $H : \mathbb{R}^{n^*} \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}$  defined by

$$H[p](u, y, \theta) := pf(u, y, \theta). \quad (2.18)$$

We also define the *end-point Lagrangian*  $\Phi : \mathbb{R}^{n^*} \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  by

$$\Phi[\lambda](y_T, \theta) := \phi(y_T, \theta) + \lambda \Phi(y_T, \theta). \quad (2.19)$$

**Definition 11.** *Let  $\lambda \in \mathbb{R}^{n^*}$ . We say that  $p^\lambda$  in  $W^{1,\infty}(0, T; \mathbb{R}^{n^*})$  is the costate associated with  $\lambda$  if it satisfies the following differential equation:*

$$\begin{cases} -\dot{p}_t^\lambda = H_y[p_t^\lambda](\bar{u}_t, \bar{y}_t, 0), & \text{for a. a. } t \in [0, T], \\ p_T^\lambda = \Phi'[\lambda](\bar{y}_T, 0). \end{cases} \quad (2.20)$$

**Lemma 12.** *Given  $v \in L^\infty(0, T; \mathbb{R}^n)$  and  $z^0 \in \mathbb{R}^n$ , let  $z \in \mathcal{Y}$  be the solution of*

$$\dot{z}_t = f_y[t]z_t + v_t, \quad z_0 = 0. \quad (2.21)$$

*Then, for all  $\lambda$  in  $\mathbb{R}^{n^*}$ ,  $\Phi'[\lambda](\bar{y}_T, 0)z_T = \int_0^T p_t^\lambda v_t dt$ .*

*Proof.* The lemma is obtained with an integration by parts:

$$\begin{aligned} \Phi'[\lambda](\bar{y}_T, 0)z_T &= p_T^\lambda z_T - p_0^\lambda z_0 = \int_0^T (\dot{p}_t^\lambda z_t + p_t^\lambda \dot{z}_t) dt \\ &= \int_0^T (-p_t^\lambda f_y[t]z_t + p_t^\lambda f_y[t]z_t + p_t^\lambda v_t) dt = \int_0^T p_t^\lambda v_t dt, \end{aligned}$$

as was to be proved.  $\square$

In the sequel, the notations  $N$  and  $T$  refer to the normal and the tangent cones.

**Definition 13.** *We say that  $\lambda \in N_K(\Phi(\bar{y}_T, 0))$  is a Pontryagin multiplier if,*

$$H[p_t^\lambda](u, \bar{y}_t, 0) \geq H[p_t^\lambda](\bar{u}_t, \bar{y}_t, 0), \quad \text{for a. a. } t, \forall u \in U_R. \quad (2.22)$$

*We denote by  $\Lambda^P$  the set of Pontryagin multipliers.*

**Remark 14.** *By (2.3),  $\lambda \in N_K(\Phi(\bar{y}_T, 0))$  iff for all  $i$  in  $I$ ,  $\lambda_i \geq 0$  and  $\Phi_i(\bar{y}_T, 0) < 0 \implies \lambda_i = 0$ . Note also that (2.22) is equivalent to: for all  $\mu$  in  $\mathcal{M}_R^Y$ ,*

$$\int_0^T \int_{U_R} H[p_t^\lambda](u, \bar{y}_t) - H[p_t^\lambda](\bar{u}_t, \bar{y}_t) d\mu_t(u) dt \geq 0. \quad (2.23)$$

### 3 First-order upper estimate of the value function

In this section, we compute a first-order upper expansion of the value function. As already mentioned, the upper estimated is true for any  $\eta \in (0, \bar{\eta}]$ .

Consider the Pontryagin linearized problem

$$\begin{cases} \text{Min}_{\xi \in \mathcal{C}(\mathcal{R}_T)} & \phi'(\bar{y}_T, 0)(\xi + \xi_T^\theta, 1), \\ \text{s.t.} & \Phi'(\bar{y}_T, 0)(\xi + \xi_T^\theta, 1) \in T_K(\Phi(\bar{y}_T, 0)). \end{cases} \quad (PL_\theta)$$

**Lemma 15.** *For all  $\eta \in (0, \bar{\eta}]$ , the following upper estimate on the value function holds:*

$$\limsup_{\theta \downarrow 0} \frac{V^\eta(\theta) - V^\eta(0)}{\theta} \leq \text{Val}(PL_\theta). \quad (3.1)$$

*Proof.* Let  $\eta > 0$  and let  $(\theta_k)_k \downarrow 0$  be such that

$$\lim_{k \rightarrow \infty} \frac{V^\eta(\theta_k) - V^\eta(0)}{\theta_k} = \limsup_{\theta \downarrow 0} \frac{V^\eta(\theta) - V^\eta(0)}{\theta} \quad (3.2)$$

Let  $\xi \in F(PL_\theta)$ . There exists a sequence  $(\alpha_k, \nu^k, \xi^k)_k$  in  $\mathbb{R}_+ \times \mathcal{M}_R^Y \times \mathcal{R}_T$  such that  $\xi = \lim \alpha_k \xi^k$  and  $\xi^k = \xi_T[\nu^k]$ , for all  $k$ . Note that it may happen that  $\alpha_k \rightarrow +\infty$ . Extracting if necessary a subsequence of  $(\theta_k)_k$ , we can suppose that

$$\theta_k \alpha_k \leq 1 \quad \text{and} \quad \alpha_k^2 \leq \frac{1}{k \theta_k}. \quad (3.3)$$

We set

$$\mu^k = (1 - \theta_k \alpha_k) \bar{\mu} + \theta_k \alpha_k \nu^k.$$

Then  $(\mu^k)_k$  is a sequence of Young measures and lemma 6 implies that

$$\|y[\mu^k, \theta_k] - (\bar{y} + \xi[\mu^k] + \theta_k \xi^\theta)\|_\infty = O(d_1(\mu^k, \bar{\mu})^2 + \theta_k^2).$$

By (3.3),

$$d_1(\mu^k, \bar{\mu})^2 = O(\theta_k^2 \alpha_k^2) = O\left(\frac{\theta_k}{k}\right) = o(\theta_k), \quad (3.4)$$

thus, since  $\theta_k |\xi - \alpha_k \xi^k| = o(\theta_k)$ ,

$$\|y_T[\mu^k, \theta_k] - [\bar{y}_T + \theta_k(\xi + \xi_T^\theta)]\|_\infty = o(\theta_k). \quad (3.5)$$

We obtain the two following expansions:

$$\phi(y_T[\mu^k, \theta_k], \theta_k) = \phi(\bar{y}_T, 0) + \theta_k \phi'(\bar{y}_T, 0)(\xi + \xi_T^\theta, 1) + o(\theta_k), \quad (3.6)$$

$$\Phi(y_T[\mu^k, \theta_k], \theta_k) = \Phi(\bar{y}_T, 0) + \theta_k \Phi'(\bar{y}_T, 0)(\xi + \xi_T^\theta, 1) + o(\theta_k). \quad (3.7)$$

Since  $\Phi'(\bar{y}_T, 0)(\xi + \xi_T^\theta, 1) \in T_K(\Phi(\bar{y}_T, 0))$ , we obtain that  $\text{dist}(\Phi(y_T[\mu^k, \theta_k], \theta_k), K) = o(\theta_k)$  and by the metric regularity theorem (theorem 9), we obtain the existence of a feasible sequence  $\tilde{\mu}^k$  such that  $d_1(\tilde{\mu}^k, \mu^k) = o(\theta_k)$ . Moreover, by (3.4), for  $k$  large enough,  $\|y[\mu^k, \theta_k] - \bar{y}\|_\infty \leq \eta$ . By lemma 37, estimate (3.6) holds for  $\tilde{\mu}^k$  and therefore, for  $k$  large enough,

$$V^\eta(\theta_k) - V^\eta(0) \leq \phi(y_T[\tilde{\mu}^k, \theta_k], \theta_k) - \phi(\bar{y}_T, 0) = \theta_k \phi'(\bar{y}_T, 0)(\xi + \xi_T^\theta, 1) + o(\theta_k).$$

Finally, minimizing with respect to  $\xi$ , we find that

$$\lim_{k \rightarrow \infty} \frac{V^\eta(\theta_k) - V^\eta(0)}{\theta_k} \leq \text{Val}(PL_\theta)$$

and the lemma is now proved.  $\square$

Let us define (formally) the Lagrangian of the problem by

$$\mathcal{L}(u, y, \lambda, \theta) := \int_0^T H[p_t^\lambda](u_t, y_t, \theta) dt + \Phi[\lambda](y_T, \theta) - \int_0^T p_t^\lambda \dot{y}_t dt$$

and the dual linearized problem ( $DL_\theta$ ) by

$$\text{Max}_{\lambda \in \Lambda^P} \mathcal{L}_\theta(\bar{u}, \bar{y}, \lambda, \bar{\theta}), \quad (DL_\theta)$$

with

$$\mathcal{L}_\theta(\bar{u}, \bar{y}, \lambda, \bar{\theta}) := \int_0^T H_\theta[p_t^\lambda][t] dt + \Phi_\theta[\lambda](\bar{y}_T, 0). \quad (3.8)$$

**Theorem 16.** *Problem ( $DL_\theta$ ) is the dual of problem ( $PL_\theta$ ) and has the same value.*

*Proof.* Let us check that problem ( $PL_\theta$ ) is qualified. Since  $K \subset T_K(\Phi(\bar{y}_T, 0))$ ,

$$\varepsilon B \subset \Phi(\bar{y}_T, 0) + \Phi_{y_T}(\bar{y}_T, 0)\mathcal{C}(\mathcal{R}_T) - T_K(\Phi(\bar{y}_T, 0)). \quad (3.9)$$

It is easy to prove that  $\Phi(\bar{y}_T, 0) - T_K(\Phi(\bar{y}_T, 0))$  is a cone. Therefore, the r.h.s. of (3.9) is a cone and contains necessarily the whole space  $\mathbb{R}^{n_C}$ . Thus,

$$\varepsilon B \subset \mathbb{R}^{n_C} = \Phi(\bar{y}_T, 0) + \Phi'(\bar{y}_T, 0)(\xi_T^\theta + \mathcal{C}(\mathcal{R}_T), 1) - T_K(\Phi(\bar{y}_T, 0)),$$

which is the qualification condition for the linearized problem.

Now, let us study the dual problem, which is:

$$\text{Max}_{\lambda \in N_K(\Phi(\bar{y}_T, 0))} \inf_{\xi \in \mathcal{C}(\mathcal{R}_T)} \Phi'[\lambda](\bar{y}_T, 0)(\xi_T^\theta + \xi, 1). \quad (3.10)$$

By lemma 12, we obtain that the dual problem is

$$\text{Max}_{\lambda \in N_K(\Phi(\bar{y}_T, 0))} \inf_{\xi \in \mathcal{C}(\mathcal{R}_T)} \left\{ \Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi + \int_0^T H_\theta[p_t^\lambda][t] dt + \Phi_\theta[\lambda](\bar{y}_T, 0) \right\}. \quad (3.11)$$

We claim that for  $\lambda \in N_K(\Phi(\bar{y}_T, 0))$ ,

$$D(\lambda) := \inf_{\xi \in \mathcal{C}(\mathcal{R}_T)} \Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi = \begin{cases} 0 & \text{if } \lambda \in \Lambda^P, \\ -\infty & \text{otherwise.} \end{cases} \quad (3.12)$$

It is clear that  $D(\lambda) \in \{0, -\infty\}$  since  $\Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi$  is linear with respect to  $\xi$  and  $\mathcal{C}(\mathcal{R}_T)$  is a cone. Let  $\lambda \in \Lambda^P$ . By lemma 12, for  $\xi$  in  $\mathcal{R}_T$  with associated control  $\mu$ ,

$$\begin{aligned} & \Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi \\ &= \int_0^T \left( -p_t^\lambda f_y[t]\xi_t[\mu] + p_t^\lambda f_y[t]\xi_t[\mu] + \int_{U_R} p_t^\lambda [f(u, \bar{y}_t, 0) - f[t]] d\mu_t(u) \right) dt \\ &= \int_0^T \int_{U_R} (H[p_t^\lambda](u, \bar{y}_t) - H[p_t^\lambda](\bar{u}_t, \bar{y}_t)) d\mu_t(u) dt, \end{aligned} \quad (3.13)$$

and then,  $\Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi \geq 0$ . Let  $\xi \in \mathcal{C}(\mathcal{R}_T)$ , there exists a sequence  $(\alpha_k, \xi^k)_k$  in  $(\mathbb{R}_+ \times \mathcal{R}_T)$  such that  $\xi = \lim_k \alpha_k \xi^k$ . By (3.13),

$$\Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi = \lim_k \alpha_k \Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi^k \geq 0,$$

therefore  $D(\lambda) \geq 0$  and finally,  $D(\lambda) = 0$ . Conversely, if  $\lambda$  is not a Pontryagin multiplier, by (3.13), there exists a control  $\mu$  such that  $\Phi_{y_T}[\lambda](\bar{y}_T, 0)\xi_T[\mu] < 0$ . Consequently,  $D(\lambda) < 0$  and therefore,  $D(\lambda) = -\infty$ . This proves (3.12). Finally, combining (3.11) and (3.12), we obtain that the dual problem is equivalent to  $(DL_\theta)$  and has the same value as problem  $(PL_\theta)$  as a consequence of [5, theorem 2.165].  $\square$

Consider now the situation where there is no perturbation. The linearized problem  $(PL_\theta)$  and its dual  $(DL_\theta)$  become respectively

$$\text{Min}_{\xi \in \mathcal{C}(\mathcal{R}_T)} \phi_{y_T}(\bar{y}_T, 0)\xi, \quad \text{s.t. } \Phi_{y_T}(\bar{y}_T, 0)\xi \in T_K(\Phi(\bar{y}_T, 0)) \quad (PL)$$

and

$$\text{Max}_{\lambda \in \Lambda^P} 0, \quad (DL)$$

By lemma 15, we obtain that  $0 \leq \text{Val}(PL)$  and since  $0 \in F(PL)$ ,  $\text{Val}(PL) = 0$ . Since  $\Lambda^P$  is the set of solutions of problem  $(DL)$  and since problem  $(PL)$  has a finite value, we obtain by [5, theorem 2.165] that  $\Lambda^P$  is nonempty, convex, and compact and finally that problems  $(PL_\theta)$  and  $(DL_\theta)$  has a finite value. Therefore, estimate (3.1) writes

$$V(\theta) \leq V(0) + \theta \text{Val}(DL_\theta) + o(\theta). \quad (3.14)$$

## 4 Second-order upper estimate of the value function

In this section, we obtain a second-order upper estimate of the value function by using a ‘‘standard’’ linearization at the first order and a ‘‘Pontryagin’’ linearization at the second order. Indeed, to obtain a second-order estimate, we need to have a solution to some linearized first-order problem. Unfortunately, problem  $(PL_\theta)$  is a conic linear problem, thus, it does not have necessarily a solution. This is why we consider now a different kind of linearization, which is such that the associated linearized problem has a solution.

In this section and in the sequel, we use properties of Young measures detailed in subsection A.3.

### 4.1 Standard linearizations and estimates

We first define some operations on the space of Young measures.

**Definition 17.** Let  $\nu \in \mathcal{M}^Y$ ,  $w \in L^\infty(0, T; \mathbb{R}^m)$ , and  $\theta \in \mathbb{R}$ . We denote by

$$w \oplus \theta\nu$$

the unique Young measure  $\mu$  in  $\mathcal{M}^Y$  such that for all  $g$  in  $C^0([0, T] \times \mathbb{R}^m)$ ,

$$\int_0^T \int_{\mathbb{R}^m} g(t, u) d\mu_t(u) dt = \int_0^T \int_{\mathbb{R}^m} g(t, w_t + \theta u) d\nu_t(u).$$

If  $\theta \neq 0$ , we denote by

$$\frac{\nu \ominus w}{\theta}$$

the unique Young measure  $\mu$  in  $\mathcal{M}^Y$  which is such that for all  $g$  in  $C^0([0, T] \times \mathbb{R}^m)$ ,

$$\int_0^T \int_{\mathbb{R}^m} g(t, u) d\mu_t(u) dt = \int_0^T \int_{\mathbb{R}^m} g\left(t, \frac{u - w_t}{\theta}\right) d\nu_t(u).$$

We also denote:  $\nu \ominus w = \frac{\nu \ominus w}{1}$ .

The addition  $\oplus$  (resp. the subtraction  $\ominus$ ) must be viewed as translations on  $\mathbb{R}^m$  of vector  $w_t$  (resp.  $-w_t$ ) at each time  $t$ . The multiplication (resp. the division) by  $\theta$  must be viewed as an homothety of ratio  $\theta$  (resp.  $\frac{1}{\theta}$ ) on  $\mathbb{R}^m$ , at each time  $t$ . Note that it will always be clear from the context if the multiplication (by constants), or the division, is the operation described in the previous definition or if it the multiplication of measures by constants, which we used up to now.

**Definition 18.** For a given  $\nu$  in  $\mathcal{M}_2^Y$ , we define the standard linearization  $z[\nu]$  by

$$\begin{cases} \dot{z}_t[\nu] = f_y[t]z_t[\nu] + f_u[t]\left(\int_{\mathbb{R}^m} u \, d\nu_t(u)\right), & \text{for a.a. } t \in [0, T], \\ z_0[\nu] = 0. \end{cases}$$

We also set  $z^1[\nu] = z[\nu] + \xi^\theta$ , which is the solution of the following system:

$$\begin{cases} \dot{z}_t^1[\nu] = \int_{\mathbb{R}^m} f'[t](z_t^1[\nu], u, 1) \, d\nu_t(u), & \text{for a.a. } t \in [0, T], \\ z_0^1[\nu] = 0. \end{cases}$$

Although the Pontryagin linearization has been standard for years in the literature, we use the terminology *standard* for the linearization  $z[\nu]$  since it corresponds to the most natural way of linearizing a differential system. Note that for  $\mu \in \mathcal{M}_R^Y$ ,  $z[\mu \ominus \bar{u}]$  is the solution to

$$\begin{cases} \dot{z}_t = f_y[t]z_t + f_u[t]\left(\int_{U_R} (u - \bar{u}_t) \, d\mu_t(u)\right), \\ z_0 = 0. \end{cases} \quad (4.1)$$

**Lemma 19.** For  $\mu$  in  $\mathcal{M}_R^Y$ , the following estimates hold:

$$\|z[\mu \ominus \bar{u}] - \xi[\mu]\|_\infty = O(d_2(\bar{\mu}, \mu)^2), \quad (4.2)$$

$$\|y[\mu, \theta] - \bar{y}\|_\infty = O(\|\mu \ominus \bar{u}\|_1 + \theta), \quad (4.3)$$

$$\|y[\mu, \theta] - (\bar{y} + z[\mu \ominus \bar{u}] + \theta\xi^\theta)\|_\infty = O(\|\mu \ominus \bar{u}\|_2^2 + \theta^2). \quad (4.4)$$

The proof is given in the appendix, page 29.

**Corollary 20.** For all  $\nu$  in  $\mathcal{M}_\infty^Y$ ,

$$z[\nu] = \lim_{\theta \downarrow 0} \frac{\xi[\bar{u} \oplus \theta\nu]}{\theta}.$$

*Proof.* By estimate (4.2), for  $\theta > 0$  sufficiently small,

$$\left\|z[\nu] - \frac{\xi[\bar{u} \oplus \theta\nu]}{\theta}\right\|_\infty = \frac{1}{\theta} \|z[\theta\nu] - \xi[\bar{u} \oplus \theta\nu]\|_\infty = \frac{O(\theta^2)}{\theta} = O(\theta).$$

The corollary is now proved.  $\square$

**Remark 21.** Denote by  $\mathcal{C}$  the smallest closed convex cone containing  $\{z[\nu]_T, \nu \in L^\infty(0, T; \mathbb{R}^n)\}$ , we obtain by corollary 20 that  $\mathcal{C} \subset \mathcal{C}(\mathcal{R}_T)$ . A standard qualification condition for the problem would have been to assume that for some  $\varepsilon' > 0$ ,

$$\varepsilon' B \subset \Phi(\bar{y}_T, 0) + \Phi_{y_T}(\bar{y}_T, 0)\mathcal{C} - K.$$

This assumption is stronger than the qualification condition that we assumed.

Consider the *standard linearized problem in  $\mathcal{M}_2^Y$*

$$\text{Min}_{\nu \in \mathcal{M}_2^Y} \phi'(\bar{y}_T, 0)(z_T^1[\nu], 1), \quad \text{s.t. } \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) \in T_K(\Phi(\bar{y}_T, 0)). \quad (SPL_\theta)$$

and the *standard linearized problem in  $L^2 := L^2(0, T; \mathbb{R}^m)$*  defined by

$$\text{Min}_{\nu \in L^2} \phi'(\bar{y}_T, 0)(z_T^1[\nu], 1), \quad \text{s.t. } \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) \in T_K(\Phi(\bar{y}_T, 0)). \quad (SPL'_\theta)$$

Since  $L^2 \subset \mathcal{M}_2^Y$ ,  $\text{Val}(SPL_\theta) \leq \text{Val}(SPL'_\theta)$ . Moreover, for all  $\nu \in \mathcal{M}_2^Y$ , we can define  $v \in L^2$  by  $v_t = \int_{\mathbb{R}^m} u \, d\nu_t(u) \, dt$ . Then,  $z^1[\nu] = z^1[v]$  and therefore, the two problems have the same value.

**Definition 22.** Let  $\lambda$  in  $N_K(\Phi(\bar{y}_T, 0))$ , we say that it is a Lagrange multiplier if for almost all  $t$  in  $[0, T]$ ,  $H_u[p_t^\lambda](\bar{u}_t, \bar{y}_t) = 0$ . We denote by  $\Lambda^L$  the set of Lagrange multipliers.

Note that the inclusion  $\Lambda^P \subset \Lambda^L$  holds and that under the qualification condition (2.13),  $\Lambda^L$  is nonempty.

**Lemma 23.** The dual of problem  $(SPL'_\theta)$  is the following problem:

$$\text{Max}_{\lambda \in \Lambda^L} \mathcal{L}_\theta(\bar{u}, \bar{y}, \lambda, 0), \quad (SDL_\theta)$$

it has the same value as the primal problem. Moreover, problems  $(SPL_\theta)$  and  $(SPL'_\theta)$  have solutions and  $\text{Val}(PL_\theta) \leq \text{Val}(SPL_\theta)$ .

*Proof.* Remember the definition of the derivative of the Lagrangian, given by (3.8). By lemma 12, the dual of problem  $(SPL_\theta)$  is the following:

$$\begin{aligned} & \text{Max}_{\lambda \in N_K(\Phi(\bar{y}_T, 0))} \inf_{\nu \in L^2} \Phi'[\lambda](\bar{y}_T, 0)(z_T^1[\nu], 1) \\ &= \text{Max}_{\lambda \in N_K(\Phi(\bar{y}_T, 0))} \left\{ \mathcal{L}_\theta(\bar{u}, \bar{y}, \lambda, 0) + \inf_{v \in L^2} \int_0^T H_u[p_t^\lambda][t] v_t \, dt \right\}. \end{aligned}$$

Moreover, for all  $\lambda \in N_K(\Phi(\bar{y}_T, 0))$ , we easily check that

$$\inf_{v \in L^2} \int_0^T H_u[p_t^\lambda][t] v_t \, dt = \begin{cases} 0, & \text{if } \lambda \in \Lambda^L, \\ -\infty, & \text{otherwise.} \end{cases}$$

This proves that problem  $(SDL_\theta)$  is the dual of problem  $(SPL'_\theta)$ . Moreover, it follows directly from the inclusion  $\Lambda^P \subset \Lambda^L$  that

$$-\infty < \text{Val}(PL_\theta) = \text{Val}(DL_\theta) \leq \text{Val}(SDL_\theta) \leq \text{Val}(SPL'_\theta) = \text{Val}(SPL_\theta).$$

We also obtain from the inclusion that problem  $(SDL_\theta)$  is feasible. Since  $(SPL_\theta)$  is linear and since the value of its dual is not  $-\infty$ , it follows by [5, theorem 2.204] that both problems have the same value. Since  $(SPL'_\theta)$  has a finite value and is linear, it has solutions, which are also solutions to  $(SPL_\theta)$ .  $\square$

From now on, we suppose that the following restrictive assumption holds.

**Hypothesis 24.** The Pontryagin and the classical linearized problems have the same value:  $\text{Val}(SPL_\theta) = \text{Val}(PL_\theta)$ .

This hypothesis is satisfied in particular if the set of Lagrange multipliers is a singleton. This hypothesis is also satisfied if the Hamiltonian is convex with respect to  $u$ , since then the definitions of Lagrange and Pontryagin multipliers are equivalent.

## 4.2 Second-order upper estimate

**Definition 25.** For  $\nu \in \mathcal{M}_2^Y$ , we define the second-order linearization  $z^2[\nu]$  by

$$\begin{cases} \dot{z}_t^2[\nu] = f_y[t]z_t^2[\nu] + \frac{1}{2} \int_{\mathbb{R}^m} f''[t](u, z_t^1[\nu], 1)^2 d\nu_t(u), \\ z_0^2[\nu] = 0. \end{cases}$$

In the following problem, the notation  $T^2$  refers to the second-order tangent set [5, definition 3.28]. Given a solution  $\nu$  to problem  $(SPL_\theta)$ , consider the following associated linearized problem:

$$\begin{cases} \text{Min}_{\xi \in \mathcal{C}(\mathcal{R}_T)} \frac{1}{2} \phi''(\bar{y}_T, 0)(z_T^1[\nu], 1)^2 + \phi_{y_T}(\bar{y}_T, 0)(z_T^2[\nu] + \xi), \\ \text{s.t. } \frac{1}{2} \Phi''(\bar{y}_T, 0)(z_T^1[\nu], 1)^2 + \Phi_{y_T}(\bar{y}_T, 0)(z_T^2[\nu] + \xi) \\ \quad \in T_K^2(\Phi(\bar{y}_T, 0), \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1)). \end{cases} \quad (PQ_\theta(\nu))$$

Observe that in this linearized problem,  $\nu$  is the first-order direction of perturbation, for which we consider standard linearizations, and  $\xi$  is the second-order direction of perturbation, for which we consider a Pontryagin linearization. Let us define the mapping  $\Omega^\theta$  on  $\mathbb{R}^{n_{c^*}} \times \mathcal{M}_2^Y$  as follows:

$$\Omega^\theta[\lambda](\nu) = \int_0^T \int_{\mathbb{R}^m} H''[p_t^\lambda][t](u, z_t^1[\nu], 1)^2 d\nu_t(u) dt + \Phi''[\lambda](\bar{y}_T, 0)(z_T^1[\nu], 1)^2. \quad (4.5)$$

**Lemma 26.** The dual of problem  $(PQ_\theta(\nu))$  is the following problem,

$$\text{Max}_{\lambda \in S(DL_\theta)} \frac{1}{2} \Omega^\theta[\lambda](\nu), \quad (DQ_\theta(\nu))$$

and it has the same value as  $(PQ_\theta(\nu))$ .

*Proof.* It is proved in [5, proposition 3.34, equality 3.64] that since  $K$  is polyhedric,

$$T_K^2(\Phi(\bar{y}_T, 0), \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1)) = T_K(\Phi(\bar{y}_T, 0)) + \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1)\mathbb{R},$$

where the addition  $+$  is the Minkowski sum. Since the second-order tangent set contains the tangent cone, we obtain, like in the proof of theorem 16 that

$$\begin{aligned} \varepsilon B \subset \mathbb{R}^{n_c} = & \frac{1}{2} \Phi''(\bar{y}_T, 0)(z_T^1[\nu], 1)^2 + \Phi_{y_T}(\bar{y}_T, 0)(z_T^2[\nu] + \mathcal{C}(\mathcal{R}_T)) \\ & - T_K^2(\Phi(\bar{y}_T, 0), \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1)), \end{aligned}$$

which is the qualification condition. By [5, theorem 2.165], problem  $(PQ_\theta(\nu))$  has the same value as its dual.

Let us denote by  $N$  the polar cone of the second-order tangent set. For all  $\lambda$  in  $\mathbb{R}^{n_{c^*}}$ ,  $\lambda \in N$  iff  $\lambda \in N_K(\Phi(\bar{y}_T, 0))$  and  $\lambda \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) = 0$ . Following the proof of theorem 16, we obtain that the dual of problem  $(PQ_\theta(\nu))$  is the following problem:

$$\begin{aligned} \text{Max}_{\substack{\lambda \in \Lambda^P, \\ \lambda \Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) = 0}} \quad & \Phi''[\lambda](\bar{y}_T, 0)(z_T^1[\nu], 1)^2 + \Phi_{y_T}[\lambda](z_T^2[\nu]). \end{aligned}$$

and using lemma 12, we find that

$$\Phi_{y_T}[\lambda]z_T^2[\nu] = \frac{1}{2} \int_0^T \int_{\mathbb{R}^m} H''[p_t^\lambda][t](u, z_t^1[\nu], 1)^2 d\nu_t(u).$$

Moreover, by lemma 12 and hypothesis 24, for all  $\lambda$  in  $\Lambda^P$ ,

$$\begin{aligned} \lambda\Phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) = 0 &\iff \Phi'[\lambda](\bar{y}_T, 0)(z_T^1[\nu], 1) = \phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) \\ &\iff \int_0^T H_\theta[p_t^\lambda][t] dt + \Phi_\theta[\lambda](\bar{y}_T, 0) = \text{Val}(SPL_\theta) \\ &\iff \mathcal{L}_\theta(\bar{u}, \bar{y}, \lambda, 0) = \text{Val}(PL_\theta) \\ &\iff \lambda \in S(DL_\theta). \end{aligned}$$

The lemma is now proved.  $\square$

Consider the problem  $(PQ_\theta)$  defined by

$$\text{Min}_{\nu \in S(SPL_\theta)} \text{Val}(PQ_\theta(\nu)) = \text{Min}_{\nu \in S(SPL_\theta)} \text{Max}_{\lambda \in S(DL_\theta)} \frac{1}{2} \Omega^\theta[\lambda](\nu). \quad (PQ_\theta)$$

**Theorem 27.** For all  $\eta \in [0, \bar{\eta}]$ , the following second-order upper estimate holds:

$$\limsup_{\theta \downarrow 0} \frac{V^\eta(\theta) - (V^\eta(0) + \theta \text{Val}(SPL_\theta))}{\theta^2} \leq \text{Val}(PQ_\theta). \quad (4.6)$$

This theorem is proved in the appendix, page 30.

## 5 Lower estimate of the value function

### 5.1 A decomposition principle

In the family of optimization problems that we consider, the expression  $\Phi[\lambda](y_T, \theta)$  plays the role of a Lagrangian. The basic idea to obtain a lower estimate for the value function is to use a second-order expansion of the right-hand-side of the following inequality:

$$\phi(y_T, \theta) - \phi(\bar{y}_T, 0) \geq \Phi[\lambda](y_T, \theta) - \Phi[\lambda](\bar{y}_T, 0), \quad (5.1)$$

for a feasible trajectory  $y$  (for the perturbed problem  $(\mathcal{P}_\theta^{Y, \eta})$ ). This inequality holds since

$$\Phi(y_T, \theta) - \Phi(\bar{y}_T, 0) \in T_K(\Phi(\bar{y}_T, 0)) \quad \text{and} \quad \lambda \in N_K(\Phi(\bar{y}_T, 0)).$$

The main difficulty in computing an expansion of the difference of Lagrangians is that we cannot perform Taylor expansions with respect to the control variable, since we are interested by perturbations of the control which are not small for the  $L^\infty$ -norm. The idea to deal with this difficulty is to split the control into two intermediate controls, one accounting for the small perturbations and one accounting for the large perturbations (both for the  $L^\infty$ -norm). The decomposition principle that we obtain is an extension of [4, theorem 2.13].

In this part, we fix a sequence  $(\theta_k)_k \downarrow 0$  and a sequence  $(\mu^k, y^k)_k$  of feasible trajectories for the perturbed problems with  $\theta = \theta_k$ . We set  $\delta y^k = y^k - \bar{y}$ . We also fix  $\lambda \in S(DL_\theta)$ . In the proofs of lemma 28 and theorem 29, we omit to mention the dependence of the Hamiltonian with respect to  $p_t^\lambda$  (since the multiplier  $\lambda$  is fixed). For example, we will write  $H(u, \bar{y}_t, \theta)$  instead of  $H[p_t^\lambda](u, \bar{y}_t, \theta)$ . Moreover, we set  $R_{1,k} = d_1(\bar{\mu}, \mu^k)$ . Note that by lemma 6,  $\|\delta y_k\|_\infty = O(R_{1,k} + \theta_k)$ .

From now on, we set  $z^k := z[\mu^k \ominus \bar{u}]$  and  $z^{1,k} := z^k + \theta_k \xi^\theta$ . Note that the dynamic of  $z^k$  is given by equation (4.1). We set

$$\Delta\Phi^k = \Phi[\lambda](y_T^k, \theta_k) - \Phi[\lambda](\bar{y}_T, 0).$$

**Lemma 28.** *The following expansions hold:*

$$\Delta\Phi^k = \text{Val}(PL_\theta)\theta_k + I_1^k + I_2^k + I_3^k + I_4^k + o(\theta_k^2 + R_{1,k}^2), \quad (5.2)$$

where

$$\begin{aligned} I_1^k &= \int_0^T \int_{U_R} (H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t]) d\mu_t(u) dt, \\ I_2^k &= \int_0^T \int_{U_R} (H_y[p_t^\lambda](u, \bar{y}_t, 0) - H_y[p_t^\lambda][t]) z_t^{1,k} d\mu_t^k(u) dt, \\ I_3^k &= \int_0^T \int_{U_R} (H_\theta[p_t^\lambda](u, \bar{y}_t, 0) - H_\theta[p_t^\lambda][t]) \theta_k d\mu_t^k(u) dt, \\ I_4^k &= \frac{1}{2} \int_0^T H_{(y,\theta)^2}[p_t^\lambda][t](z_t^{1,k}, \theta_k)^2 dt + \frac{1}{2} \Phi''[\lambda](z_T^{1,k}, \theta_k)^2. \end{aligned}$$

and

$$\Delta\Phi^k = \int_0^T \int_{U_R} H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t] d\mu_t(u) dt + O(\|\delta y^k\|_\infty) + o(1). \quad (5.3)$$

The proof is given in the appendix, page 32.

In order to go further in the expansions, we need to split the control  $\mu^k$  into two controls. To that purpose, we consider a sequence  $(A^k, B^k)_k$  of measurable subsets of  $[0, T] \times U_R$  such that for all  $k$ ,  $(A^k, B^k)$  is a partition of  $[0, T] \times U_R$ . We consider the Young measures  $\mu^{A,k}$  and  $\mu^{B,k}$  which are the unique Young measures such that for all  $g$  in  $C^0([0, T] \times U_R)$ ,

$$\begin{cases} \int_0^T \int_{U_R} g(t, u) d\mu_t^{A,k}(u) dt = \int_{A^k} g(t, u) d\mu^k(t, u) + \int_{B^k} g(t, \bar{u}_t) d\mu^k(t, u), \\ \int_0^T \int_{U_R} g(t, u) d\mu_t^{B,k}(u) dt = \int_{B^k} g(t, u) d\mu^k(t, u) + \int_{A^k} g(t, \bar{u}_t) d\mu^k(t, u). \end{cases}$$

Note that if  $g$  is such that for almost all  $t$  in  $[0, T]$ ,  $g(t, \bar{u}_t) = 0$ , then

$$\int_0^T \int_{U_R} g(t, u) d\mu_t^k(u) dt = \int_0^T \int_{U_R} g(t, u) d\mu_t^{A,k}(u) dt + \int_0^T \int_{U_R} g(t, u) d\mu_t^{B,k}(u) dt.$$

For  $i = 1, 2$ , we set  $R_{i,A,k} := d_i(\bar{\mu}, \mu^{A,k})$  and  $R_{i,B,k} := d_i(\bar{\mu}, \mu^{B,k})$ . We also set  $z^{A,k} := z[\mu^{A,k} \ominus \bar{\mu}]$ , and  $z^{B,k} := z[\mu^{B,k} \ominus \bar{\mu}]$ .

Remember the definition of  $\Omega^\theta$  given by (4.5). For  $\lambda \in \mathbb{R}^{n_C^*}$ , let us denote by  $\Omega[\lambda] : \mathcal{M}_2^Y \rightarrow \mathbb{R}$  the following mapping:

$$\Omega[\lambda](\nu) = \int_0^T \int_{U_R} H_{(u,y)^2}[p_t^\lambda][t](u, z[\nu])^2 d\nu_t(u) dt + \Phi_{(y_T)^2}[\lambda](\bar{y}_T, 0)(z[\nu]_T)^2. \quad (5.4)$$

**Theorem 29** (Decomposition principle). *Assume that*

$$\mu^k(B^k) \longrightarrow 0 \quad \text{and} \quad \text{ess sup}_{k \rightarrow \infty} \{|u - \bar{u}_t|, (t, u) \in A^k\} \rightarrow 0. \quad (5.5)$$

Then,

$$z^k = z^{A,k} + o(R_{2,B,k}) \quad (5.6)$$

and the following expansions hold:

$$\begin{aligned} \Delta\Phi^k &= \text{Val}(PL_\theta)\theta_k + \frac{1}{2}\Omega^\theta[\lambda](\mu^{A,k} \ominus \bar{u}) \\ &\quad + \int_0^T \int_{U_R} H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t] d\mu_t^{B,k}(u) dt + o(R_{2,k}^2 + \theta_k^2). \end{aligned} \quad (5.7)$$

and

$$\begin{aligned} \Delta\Phi^k &= \int_0^T \int_{U_R} (H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t]) d\mu_t^{B,k}(u) dt + \text{Val}(PL_\theta)\theta_k \\ &\quad + \frac{1}{2}\Omega[\lambda](\mu^{A,k} \ominus \bar{u}) + O(\theta_k(\theta_k + R_{2,A,k})) + o(R_{2,k}^2). \end{aligned} \quad (5.8)$$

*Proof.* With the Cauchy-Schwarz inequality, we get  $R_{1,A,k} = O(R_{2,A,k})$  and since  $\mu^k(B^k) \rightarrow 0$ ,

$$\begin{aligned} R_{1,B,k} &= \int_{B^k} |u - \bar{u}_t| d\mu_t^k(t, u) dt \\ &\leq (\mu^k(B^k))^{1/2} \left[ \int_{B^k} |u - \bar{u}_t|^2 d\mu^k(t, u) \right]^{1/2} = o(R_{2,B,k}). \end{aligned} \quad (5.9)$$

Estimate (5.6) follows from (5.9) and  $z^k = z^{A,k} + z^{B,k}$ . In order to obtain expansion (5.7), we work with the terms of the expansion of lemma 28. First,

$$\begin{aligned} I_1^k &= \int_0^T \int_{U_R} (H(u, \bar{y}_t, 0) - H[t])(d\mu_t^{A,k}(u) + d\mu_t^{B,k}(u)) dt \\ &= \frac{1}{2} \int_0^T \int_{U_R} H_{uu}[t](u - \bar{u}_t)^2 d\mu_t^{A,k}(u) dt \\ &\quad + \int_0^T \int_{U_R} (H(u, \bar{y}_t, 0) - H[t]) d\mu_t^{B,k}(u) dt + o(R_{2,A,k}^2) \end{aligned} \quad (5.10)$$

and

$$\begin{aligned} I_2^k &= \int_0^T \int_{U_R} (H_y(u, \bar{y}_t, 0) - H_y[t])(z_t^{A,k} + \theta_k \xi_t^\theta) d\mu_t^{A,k}(u) dt \\ &\quad + \int_0^T \int_{U_R} (H_y(u, \bar{y}_t, 0) - H_y[t]) z_t^{B,k} d\mu_t^{A,k}(u) dt \\ &\quad + \int_0^T \int_{U_R} (H_y(u, \bar{y}_t, 0) - H_y[t]) z_t^{1,k} d\mu_t^{B,k}(u) dt \\ &= \int_0^T \int_{U_R} H_{u,y}[t](u - \bar{u}_t, z_t^{A,k} + \theta_k \xi_t^\theta) d\mu_t^{A,k}(u) dt + o(R_{2,k}^2 + \theta_k^2), \end{aligned} \quad (5.11)$$

Similarly, we prove that

$$I_3^k = \int_0^T \int_{U_R} H_{(u,\theta)}[t](u - \bar{u}_t, \theta_k) d\mu_t^{A,k} dt + o(R_{2,k}^2 + \theta_k^2), \quad (5.12)$$

$$\begin{aligned} I_4^k &= \frac{1}{2} \int_0^T \int_{U_R} H_{(y,\theta)^2}[t](z_t^{A,k} + \theta_k \xi_t^\theta, \theta_k)^2 d\mu_t^{A,k} dt \\ &\quad + \frac{1}{2} \Phi''[\lambda](\bar{y}_T, 0)(z_T^{A,k} + \theta_k \xi_T^\theta, \theta_k)^2 + o(R_{2,k}^2 + \theta_k^2). \end{aligned} \quad (5.13)$$

Finally, combining lemma 28 and estimates (5.10-5.13), we obtain expansion (5.7). Expansion (5.8) follows by replacing the second-order terms involving  $\theta_k$  by the estimate  $O(R_{2,A,k}\theta_k)$ .  $\square$

## 5.2 Study of the rate of convergence of perturbed solutions

In this part, we give estimates of the  $L^2$ -distance between a solution to the perturbed problem  $(\mathcal{P}_\theta^{Y,\eta})$  and  $\bar{\mu}$  under a strong second-order sufficient condition. The results will hold for small values of  $\eta$ .

**Definition 30.** We call critical cone  $C_2$  the following set:

$$C_2 := \{\nu \in \mathcal{M}_2^Y, \phi_{y_T}(\bar{y}_T, 0)z[\nu]_T \leq 0, \Phi_{y_T}(\bar{y}_T, 0)z[\nu]_T \in T_K(\Phi(\bar{y}_T, 0))\}. \quad (5.14)$$

In the following assumption, we denote by  $\text{ri}(S(DL_\theta))$  the relative interior of  $S(DL_\theta)$ , which is the interior of  $S(DL_\theta)$  for the topology induced by its affine hull.

**Hypothesis 31** (Second-order sufficient conditions). *There exists  $\alpha > 0$  such that*

1. *for some  $\bar{\lambda} \in \text{ri}(S(DL_\theta))$ , for almost all  $t$  in  $[0, T]$ ,*

$$H[p_t^{\bar{\lambda}}](u, \bar{y}_t, 0) - H[p_t^{\bar{\lambda}}](\bar{u}_t, \bar{y}_t, 0) \geq \alpha|u - \bar{u}_t|^2, \quad \forall u \in U_R,$$

2. *for all  $\nu$  in  $C_2 \setminus \{0\}$ ,  $\max_{\lambda \in S(DL_\theta)} \{\Omega[\lambda](\nu)\} > 0$ . Here,  $0$  is the Young measure which is equal for almost all  $t$  to the Dirac measure (centered at  $0$ ).*

As a consequence of assumption 31.1, for all  $\mu \in \mathcal{M}_R^Y$ ,

$$\int_0^T \int_{U_R} H[p_t^{\bar{\lambda}}](u, \bar{y}_t, 0) - H[p_t^{\bar{\lambda}}](\bar{u}_t, \bar{y}_t, 0) d\mu_t(u) dt \geq \alpha \|\mu \ominus \bar{\mu}\|_2^2.$$

**Remark 32.** *It is shown in [4, lemma 2.3] that, since  $S(DL_\theta)$  is compact, for all  $\tilde{\lambda} \in \text{ri}(S(DL_\theta))$ , there exists  $\beta > 0$  such that for almost all  $t$ , for all  $v$  in  $U_R$ ,*

$$H[p_t^{\tilde{\lambda}}](v, \bar{y}_t, 0) - H[p_t^{\tilde{\lambda}}](\bar{u}_t, \bar{y}_t, 0) \geq \beta \left( \max_{\lambda \in S(DL_\theta)} \{H[p_t^\lambda](v, \bar{y}_t, 0) - H[p_t^\lambda](\bar{u}_t, \bar{y}_t, 0)\} \right).$$

*It follows from this result that hypothesis 31.1 is equivalent to: there exists  $\alpha' > 0$  such that for almost all  $t$ , for all  $u \in U_R$ ,*

$$\max_{\lambda \in S(DL_\theta)} \{H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda](\bar{u}_t, \bar{y}_t, 0)\} \geq \alpha'|u - \bar{u}_t|^2.$$

The following lemma states some useful semi-continuity properties for  $\Omega$  and  $\Omega^\theta$ .

**Lemma 33.** *If hypothesis 31.1 holds, then for all bounded sequence  $(\nu^k)_k$  in  $\mathcal{M}_2^Y$  narrowly converging to some  $\nu \in \mathcal{M}_2^Y$ ,*

1. *the sequence  $(z[\nu^k])_k$  converges to  $z[\nu]$  for the  $L^\infty$ -distance*
2. *for all  $\lambda \in S(DL_\theta)$ ,  $\Omega[\lambda](\nu) \leq \liminf_{k \rightarrow \infty} \Omega[\lambda](\nu^k)$*
3. *if  $\nu = 0$  and  $\Omega[\bar{\lambda}](\nu^k) \rightarrow 0$ , then  $\|\nu^k\|_2 \rightarrow 0$ .*

This lemma is proved in the appendix, page 33.

**Lemma 34.** *If  $\eta > 0$  is sufficiently small, then for any sequence  $(\theta_k)_k \downarrow 0$ , for any sequence of solutions  $(\mu^k, y^k)_k$  to  $(\mathcal{P}_\theta^{Y,\eta})$ , with  $\theta = \theta_k$ ,*

$$R_{2,k} = d_2(\bar{\mu}, \mu^k) \rightarrow 0. \quad (5.15)$$

*Proof.* Assume, on the contrary, that there exist two sequences  $(\eta_k)_k \downarrow 0$  and  $(\theta_k)_k \downarrow 0$  and a sequence of solutions  $(\mu^k, y^k)$  to  $(\mathcal{P}_\theta^{Y,\eta})$  with  $\eta = \eta_k$  and  $\theta = \theta_k$  such that  $R_{2,k} = d_2(\bar{\mu}, \mu^k)$  does not converge to 0. It follows from inequality (5.1) and estimate (5.3) that

$$o(1) = \phi(y_T^k, \theta_k) - \phi(\bar{y}_T, 0) \geq \int_0^T \int_{U_R} (H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t]) d\mu_t^k(u) dt + o(1),$$

thus, by assumption 31.1,  $R_{2,k} \rightarrow 0$ , in contradiction with the initial assumption.  $\square$

From now on, we fix a parameter  $\eta > 0$  sufficiently small so that lemma 34 is satisfied. We are now able to build a sequence  $(A^k, B^k)_k$  which can be used in the decomposition principle. Let us set

$$A^k := \{ (t, u) \in [0, T] \times U_R, |u - \bar{u}_t| < \sqrt{R_{1,k}} \} \quad \text{and} \quad B^k := (A^k)^c. \quad (5.16)$$

We consider the sequences  $(\mu^{A,k})_k$  and  $(\mu^{B,k})_k$  associated with  $(\mu^k)_k$  and the sequence of partitions  $(A^k, B^k)_k$ . We still use the notations  $z^{A,k}$  and  $z^{B,k}$ . Then,

$$R_{1,k} = \int_0^T \int_{U_R} |u - \bar{u}_t| d\mu_t^k(u) dt \geq \sqrt{R_{1,k}} \int_0^T \int_{U_R} \mathbf{1}_{B^k}(t, u) d\mu_t^k(u) dt$$

Thus,  $\mu^k(B^k) \leq \sqrt{R_{1,k}} = O(\sqrt{R_{2,k}}) \rightarrow 0$ , by lemma 34. Moreover,

$$\operatorname{ess\,sup}_{k \rightarrow \infty} \{ |u - \bar{u}_t|, (t, u) \in A^k \} \leq \sqrt{R_{1,k}} = O(\sqrt{R_{2,k}}) \rightarrow 0.$$

As a consequence, we can apply the decomposition principle to the partition.

**Theorem 35.** *Under hypotheses 24 and 31, the following estimates on the rate of convergence of perturbed solutions hold:*

$$R_{2,k} = d_2(\bar{\mu}, \mu^k) = O(\theta_k), \quad \|y^k - \bar{y}\|_\infty = O(\theta_k). \quad (5.17)$$

*Proof.*  $\triangleright$  *First step:*  $R_{2,B,k} = O(R_{2,A,k} + \theta_k)$ .

With expansion (5.8) and the second-order upper estimate (4.6), we obtain that for all  $\lambda \in S(DL_\theta)$ ,

$$\begin{aligned} & \frac{1}{2} \Omega[\lambda](\mu^{A,k} \ominus \bar{\mu}) + \int_0^T \int_{U_R} H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t] d\mu_t^{B,k}(u) dt \\ & \leq o(R_{2,A,k}^2 + R_{2,B,k}^2) + O(\theta_k R_{2,A,k}) + O(\theta_k^2). \end{aligned} \quad (5.18)$$

Specializing (5.18) for  $\bar{\lambda}$  and since  $\Omega[\bar{\lambda}](\mu^{A,k} \ominus \bar{\mu}) = O(R_{2,A,k}^2)$ , we obtain by the second-order sufficient condition hypothesis 31.1 that

$$\alpha R_{2,B,k}^2 = O(R_{2,A,k}^2 + \theta_k^2),$$

thus,  $R_{2,B,k} = O(R_{2,A,k} + \theta_k)$ .

▷ *Second step:*  $R_{2,A,k} = O(\theta_k)$ .

Let us prove by contradiction that  $R_{2,A,k} = O(\theta_k)$ . Extracting if necessary a subsequence, we may assume that  $\theta_k = o(R_{2,A,k})$ . It follows directly that  $R_{2,B,k} = O(R_{2,A,k})$ . For all  $\lambda \in S(DL_\theta)$ , the difference of Hamiltonians is nonnegative, thus, by (5.18), for all  $\lambda \in S(DL_\theta)$ ,

$$\Omega[\lambda](\mu^{A,k} \ominus \bar{u}) \leq O(\theta_k^2) + O(\theta_k R_{2,A,k}) + o(R_{2,A,k}^2) = o(R_{2,A,k}^2). \quad (5.19)$$

Using definition 17, we set

$$\nu^k = \frac{\mu^{A,k} \ominus \bar{u}}{R_{2,A,k}} \quad \text{and} \quad z^k = z[\nu^k] = \frac{z[\mu^{A,k} \ominus \bar{u}]}{R_{2,A,k}}.$$

For all  $k$ ,  $\|\nu^k\|_2^2 = 1$ , therefore, up to a subsequence, we can suppose that  $(\nu^k)_k$  converges narrowly to  $\bar{\nu} \in \mathcal{M}_2^Y$ . By lemma 33,  $z[\nu^k]$  converges uniformly to  $z[\bar{\nu}]$ . Let us prove that

$$\phi_{y_T}(\bar{y}_T, 0)z_T[\bar{\nu}] = 0, \quad (5.20)$$

$$\Phi_{y_T}(\bar{y}_T, 0)z_T[\bar{\nu}] \in T_K(\Phi(\bar{y}_T, 0)). \quad (5.21)$$

By lemma 19, we obtain that

$$\delta y_T^k = z_T^{A,k} + z_T^{B,k} + \theta_k \xi_T^\theta + O(\theta_k^2 + R_{1,A,k}^2 + R_{1,B,k}^2) = z_T^{A,k} + o(R_{2,A,k}),$$

and finally that  $\delta y_T^k = R_{2,A,k}(z^k + o(1)) = R_{2,A,k}(z[\bar{\nu}] + o(1))$ . As a consequence,

$$\phi(y_T^k, \theta_k) - \phi(\bar{y}_T, 0) = R_{2,A,k}[\phi_{y_T}(\bar{y}_T, 0)z_T[\bar{\nu}] + o(1)], \quad (5.22)$$

$$\Phi(y_T^k, \theta_k) - \Phi(\bar{y}_T, 0) = R_{2,A,k}[\Phi_{y_T}(\bar{y}_T, 0)z_T[\bar{\nu}] + o(1)]. \quad (5.23)$$

We obtain (5.21) directly and (5.20) follows from (5.22) and from the following first-order upper estimate:

$$\phi(y_T^k, \theta_k) - \phi(\bar{y}_T, 0) \leq O(\theta_k) = o(R_{2,A,k}).$$

Therefore,  $\nu \in C_2$ . We obtain from lemma 33 and (5.19) that

$$\sup_{\lambda \in SDL_\theta} \Omega[\lambda](\nu) \leq 0.$$

By the second-order sufficient condition (hypothesis 31.2),  $\nu = 0$ . Applying (5.19) to  $\bar{\lambda}$ , we obtain by the lower semi-continuity of  $\Omega[\bar{\lambda}]$  that  $\lim_k \Omega[\bar{\lambda}](\nu^k) = 0$  and thus, by lemma 33,  $\|\nu^k\|_2 \rightarrow 0$ , in contradiction with the fact that  $\|\nu^k\|_2 = 1$  for all  $k$ . It follows that  $R_{2,A,k} = O(\theta_k)$ , thus  $R_{2,k} = O(R_{2,A,k} + R_{2,B,k}) = \theta_k$  and finally that  $\|y^k - \bar{y}\|_\infty = O(\theta_k)$ , by lemma 37.  $\square$

### 5.3 First- and second-order estimates

In this section, we prove that the first and the second-order upper estimates that we have computed in section 4 are exact expansions, for sufficiently small values of  $\eta > 0$  (so that lemma 34 holds). The first-order estimate derives directly from inequality (5.1), expansion (5.8), and theorem 35 (under hypotheses 24 and 31):

$$V^\eta(\theta_k) - V^\eta(0) = \text{Val}(PL_\theta)\theta_k + O(\theta_k^2). \quad (5.24)$$

**Theorem 36.** *Under hypotheses 24 and 31, the following second-order estimate holds:*

$$V^\eta(\theta) = V^\eta(0) + \theta \text{Val}(PL_\theta) + \theta^2 \text{Val}(PQ_\theta) + o(\theta^2). \quad (5.25)$$

Moreover, for any  $\theta_k \downarrow 0$ , we can extract a subsequence of solutions  $\mu^k$  to  $(\mathcal{P}_\theta^{Y,\eta})$  such that  $\frac{\mu^k \ominus \bar{u}}{\theta_k}$  converges narrowly to some  $\bar{\nu}$  solution of  $(PQ_\theta)$ .

*Proof.* Let  $(\theta_k)_k \downarrow 0$ . We set

$$\nu^{A,k} = \frac{\mu^{A,k} \ominus \bar{u}}{\theta_k}, \quad \nu^k = \frac{\mu^k \ominus \bar{u}}{\theta_k}.$$

By theorem 35,  $R_{2,A,k}^2 = O(\theta_k^2)$ . Therefore,  $(\nu^{A,k})_k$  is bounded for the  $L^2$ -norm and we can extract a subsequence such that  $(\nu^{A,k})$  narrowly converges to some  $\bar{\nu}$  in  $\mathcal{M}_2^Y$ . Moreover, we can show that

$$d_1(\nu^k, \nu^{A,k}) \leq \frac{\|\mu^{B,k} \ominus \bar{u}\|_1}{\theta_k} = o(1),$$

thus,  $\nu^k$  equally converges to  $\bar{\nu}$  for the narrow topology. For all  $\lambda \in S(DL_\theta)$ ,

$$\int_0^T \int_{U_R} H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t] d\mu_t^{B,k}(u) dt \geq 0,$$

thus, by inequality (5.1), by the decomposition principle (theorem 29), and by the lower semi-continuity of  $\Omega^\theta$  (lemma 33),

$$\begin{aligned} V^\eta(\theta_k) - V^\eta(0) &\geq \theta_k \text{Val}(PL_\theta) + \frac{\theta_k^2}{2} \Omega^\theta[\lambda](\nu^{A,k}) + o(\theta_k^2) \\ &\geq \theta_k \text{Val}(PL_\theta) + \frac{\theta_k^2}{2} \Omega^\theta[\lambda](\bar{\nu}) + o(\theta_k^2). \end{aligned}$$

Let us prove that  $\bar{\nu}$  is a solution to problem  $(SPL_\theta)$ . Following the proof of theorem 35, we obtain that

$$\delta y_T^k = \theta_k(z_T[\bar{\nu}] + \xi_T^\theta + o(1)),$$

and therefore that

$$\phi(y_T^k, \theta_k) - \phi(\bar{y}_T, 0) = \theta_k \phi'(\bar{y}_T, 0)(z_T[\bar{\nu}] + \xi_T^\theta, 1) + o(\theta_k), \quad (5.26)$$

$$\Phi(y_T^k, \theta_k) - \Phi(\bar{y}_T, 0) = \theta_k \Phi'(\bar{y}_T, 0)(z_T[\bar{\nu}] + \xi_T^\theta, 1) + o(\theta_k). \quad (5.27)$$

By (3.14), we obtain that

$$\phi(y_T^k, \theta_k) - \phi(\bar{y}_T, 0) \leq \text{Val}(PL_\theta)\theta_k + o(\theta_k),$$

therefore

$$\phi'(\bar{y}_T, 0)(z_T[\bar{\nu}] + \xi_T^\theta, 1) \leq \text{Val}(PL_\theta).$$

and by (5.27),

$$\Phi'(\bar{y}_T, 0)(z_T[\bar{\nu}] + \xi_T^\theta, 1) \in T_K(\Phi(\bar{y}_T, 0)).$$

This proves that  $\bar{\nu}$  is a solution to  $(SPL_\theta)$ . By lemma 26 and theorem 27, we obtain that

$$\text{Val}(PQ_\theta(\bar{\nu})) \leq \inf_{\nu \in S(SPL_\theta)} \text{Val}(PQ_\theta(\nu)),$$

thus,  $\bar{\nu}$  is a solution to problem  $(PQ_\theta)$  and the theorem is now proved. It also proves that problem  $(PQ_\theta)$  has a finite value.  $\square$

## 6 Two examples

### 6.1 A different value for the Pontryagin and the standard linearized problem

Let us consider the following dynamic in  $\mathbb{R}^2$ :

$$\begin{cases} \dot{y}_t = (u_t^3, u_t^2)^T, \text{ for a.a. } t \in [0, T], \\ y_0 = (0, 0)^T. \end{cases}$$

The control  $u$  is such that  $\|u\|_\infty \leq 1$  and we minimize  $y_{2,T}[u]$  under the constraint  $y_{1,T}[u] = \theta$ , with  $\theta \geq 0$  and  $\bar{\theta} = 0$ . The coordinate  $y_2$  correspond to the integral which would have been used in a Bolza formulation of the problem. For  $\bar{\theta} = 0$ , the problem has a unique solution  $\bar{u} = 0$ ,  $\bar{y} = (0, 0)^T$ . This solution is qualified in the sense of definition 8, since for  $v = 1$ ,  $\xi_1[v] = T$  and for  $v = -1$ ,  $\xi_1[v] = -T$ . However, the solution is not qualified in the sense of the standard definition, since the standard linearized dynamic  $z$  is equal to 0.

For  $\theta \leq T$ , the problem has infinitely many solutions, one of them being:

$$u_t^\theta = \begin{cases} 1, & \text{if } t \in (0, \theta), \\ 0, & \text{if } t \in (\theta, T). \end{cases}$$

Indeed,  $y_{1,T}[u^\theta] = \theta$ ,  $y_{2,T}[u^\theta] = \theta$  and if  $v^\theta$  is feasible, then

$$\theta = y_{1,T}[v^\theta] = \int_0^T (v_t^\theta)^3 dt \leq \int_0^T (v_t^\theta)^2 dt = y_{2,T}[v],$$

which proves that  $u^\theta$  is optimal. Moreover, if  $v^\theta$  is optimal, then the previous inequality is an equality and thus, for almost all  $t$ ,  $(v_t^\theta)^3 = (v_t^\theta)^2$ , that is to say,  $v_t^\theta \in \{0, 1\}$ . We also obtain that  $\|v^\theta - \bar{u}\|_2 = \sqrt{\theta}$  and  $\|v^\theta - \bar{u}\|_\infty = 1$ .

Now, let us compute the sets of multiplier  $\Lambda^L$  and  $\Lambda^P$  (for the reference problem). Since the dynamic does not depend on  $y$ , denoting by  $\lambda \in \mathbb{R}$  the dual variable associated with the constraint  $y_{1,T}[u] - \theta = 0$ , the costate  $p^\lambda$  is constant and given by  $p_t = (\lambda, 1)$ . The Hamiltonian is given by

$$H[\lambda](u) = u^2 + \lambda u^3.$$

As a consequence, we obtain that  $\Lambda^L = \mathbb{R} \times \{1\}$  and  $\Lambda^P = [-1, 1] \times \{1\}$ . The Lagrangian associated with our family of problem is given by

$$\mathcal{L}(u, y, \lambda, \theta) = \int_0^T (u_t^2 + \lambda u_t^3) dt + \lambda(y_{1,\theta} - \theta),$$

therefore,  $\mathcal{L}_\theta(\bar{u}, \bar{y}, \lambda, \bar{\theta}) = -\lambda$ ,  $\text{Val}(PL_\theta) = 1$ , and  $\text{Val}(SPL_\theta) = +\infty$ . In this example, the Pontryagin linearized problem enables a more accurate estimation of the value function. Since the solution  $\bar{u}$  is not qualified in a standard definition, it is not surprising that the associated linearized problem has a value equal to  $+\infty$ .

Note that the second-order theory developed in the article cannot be used to study this example, since we do not have the equality of  $\text{Val}(PL_\theta)$  and  $\text{Val}(SPL_\theta)$ . Moreover, observe that for the solution  $\lambda = -1$  of  $(DL_\theta)$ , the Hamiltonian  $H[\lambda](u) = u^2 - u^3$  has two minimizers: 0 and 1. The set of minimizers contains the support of the solutions to the perturbed problems.

## 6.2 No classical solutions for the perturbed problems

This second example shows a family of problems for which the perturbed problems do not have a classical solution. This example does not fit to the framework of the study since we consider control constraints. However, we believe it is interesting since in this case, the ratio  $(\mu^\theta \ominus \bar{u})/\theta$  converges to a purely relaxed element of  $\mathcal{M}_2^Y$  for the narrow topology. This confirms us in the idea to use relaxation to perform a sensitivity analysis of optimal control problems.

Let us consider the following dynamic in  $\mathbb{R}^2$ :

$$\begin{cases} (\dot{y}_{1,t}, \dot{y}_{2,t})^T = (u_t, y_1^2 + 2(v_t - \theta)^2 - u_t^2)^T, & \text{for a.a. } t \in [0, T], \\ (y_{1,0}, y_{2,0})^T = (0, 0)^T, \end{cases}$$

where for almost all  $t$  in  $[0, T]$ ,  $v_t \geq u_t$  and  $v_t \geq -u_t$ . The perturbation parameter  $\theta$  is nonnegative and  $\bar{\theta} = 0$ . We minimize  $y_{2,T}$ . For  $\theta = 0$ , the problem has a unique solution  $\bar{u} = (0, 0)^T$ ,  $\bar{y} = (0, 0)^T$ . The associated costate  $p = (p_1, p_2)$  is constant, given by  $p_1 = 0$  and  $p_2 = 1$ . Thus,

$$H[p](u, v, \bar{y}_t) = 2(v - \theta)^2 - u^2.$$

This Hamiltonian has been “designed” in a way to have a unique minimizer when  $\theta = 0$ , but two minimizers  $(\pm 2\theta, 2\theta)$  when  $\theta > 0$ . Let us focus on optimal solutions to the problem when  $\theta > 0$ . Let  $u, v \in L^\infty([0, T], \mathbb{R})$ , we have

$$\begin{aligned} y_{2,T}[u, v] &= \int_0^T y_{1,t}[u, v]^2 + 2(v_t - \theta)^2 - u_t^2 dt \\ &= \int_0^T y_{1,t}[u, v]^2 + 2v_t^2 - 4\theta v_t + 2\theta^2 - u_t^2 dt \\ &= \int_0^T y_{1,t}[u, v]^2 + (v_t^2 - u_t^2) + (v_t - 2\theta)^2 - 2\theta^2 dt \geq -2\theta^2 T. \end{aligned}$$

This last inequality is an equality if for almost all  $t$  in  $[0, T]$ ,  $y_{1,t}[u, v] = 0$ ,  $v_t = 2\theta$ ,  $|u_t| = v_t$ . As a consequence, the problem does not have classical solutions, but has a unique relaxed one,  $\mu^\theta = ((\delta_{-2\theta} + \delta_{2\theta})/2, 2\theta)$ . Moreover,

$$\frac{\mu^\theta \ominus \bar{u}}{\theta} = ((\delta_{-2} + \delta_2)/2, \delta_2).$$

## A Properties of Young measures

### A.1 First definitions

**Weak-\* topology on bounded measures** Let  $X$  be a closed subset of  $\mathbb{R}^m$ . We say that a real function  $\psi$  on  $[0, T] \times X$  vanishes at infinity if for all  $\varepsilon > 0$ , there exists a compact subset  $K$  of  $X$  such that for all  $(t, u)$  in  $[0, T] \times (X \setminus K)$ ,  $|\psi(t, u)| \leq \varepsilon$ . We denote by  $C^0([0, T] \times X)$  the set of continuous real functions vanishing at infinity. The set  $\mathcal{M}_b([0, T] \times X)$  of bounded measures on  $[0, T] \times X$  is the topological dual of  $C^0([0, T] \times X)$ . The associated *weak-\* topology* is metrizable since  $[0, T] \times X$  is separable.

**Young measures** Let us denote by  $P$  the projection from  $[0, T] \times X$  to  $[0, T]$ . We say that  $\mu \in \mathcal{M}_b^+([0, T] \times X)$  is a *Young measure* if  $P_\# \mu$  is the Lebesgue measure on  $[0, T]$ . We denote by  $\mathcal{M}^Y(X)$  the set of Young measures, which is weakly-\* compact [15, theorem 1].

**Disintegrability** Let us denote by  $\mathcal{P}(X)$  the set of probability measures on  $X$ . To all measurable mapping  $\nu \in L^\infty(0, T; \mathcal{P}(X))$  (see the definition in [15, page 157]), we associate a unique Young measure  $\mu$  defined by: for all  $\psi$  in  $C^0([0, T] \times X)$ ,

$$\int_{[0, T] \times X} \psi(t, u) d\mu(t, u) = \int_0^T \int_X \psi(t, u) d\nu_t(u) dt.$$

This mapping defines a bijection from  $L^\infty([0, T]; \mathcal{P}(X))$  to  $\mathcal{M}^Y(X)$ . This property is called *disintegrability*. Note that  $L^\infty(0, T; \mathcal{P}(X)) \subset L^\infty(0, T; \mathcal{M}_b(X))$ , which is the dual of  $L^1(0, T; C^0(X))$  [15, page 179]. On  $\mathcal{M}^Y(X)$ , the weak-\* topology of this dual pair is equivalent to the weak-\* topology previously defined [15, theorem 2]. In the article, we always write Young measures in a disintegrated form.

**Density** To all  $u$  in  $L([0, T]; X)$ , we associate the unique Young measure  $\mu$  defined by for almost all  $t$  in  $[0, T]$ ,  $\mu_t = \delta_{u_t}$ . The space  $L([0, T]; X)$  is dense in  $\mathcal{M}^Y(X)$  for the weak-\* topology [16, proposition 8].

**Lower semi-continuity of integrands** We say that  $\psi : [0, T] \times X \rightarrow \mathbb{R} \cup \{+\infty\}$  is a *positive normal integrand* if  $\psi$  is measurable,  $\psi \geq 0$  and if for almost all  $t$  in  $[0, T]$ ,  $\psi(t, \cdot)$  is l.s.c. If  $\psi$  is a positive normal integrand, then the mapping

$$\mu \in \mathcal{M}^Y(X) \mapsto \int_0^T \int_X \psi(t, u) d\mu_t(u) dt$$

is l.s.c. for the weak-\* topology [15, theorem 4].

**Narrow topology** We say that the measurable mapping  $\psi : [0, T] \times X \rightarrow \mathbb{R}$  is a *bounded Caratheodory integrand* if for almost all  $t$  in  $[0, T]$ ,  $\psi(t, \cdot)$  is continuous and bounded and if  $\|\psi(t, \cdot)\|_\infty$  is integrable. The *narrow topology* on  $\mathcal{M}^Y(X)$  is the weakest topology such that for all bounded Caratheodory integrand  $\psi$ ,

$$\mu \in \mathcal{M}^Y(X) \mapsto \int_0^T \int_X \psi(t, u) d\mu_t(u) du$$

is continuous. This topology is finer than the weak-\* topology.

**Wasserstein distance** We denote by  $P^1$  and  $P^2$  the two projections from  $[0, T] \times X \times X$  to  $[0, T] \times X$  defined by  $P^1(t, u, v) = (t, u)$  and  $P^2(t, u, v) = (t, v)$ . Let  $\mu^1$  and  $\mu^2$  be in  $\mathcal{M}^Y(X)$ , then  $\pi$  in  $\mathcal{M}_b^+([0, T] \times X \times X)$  is said to be a *transportation plan* between  $\mu^1$  and  $\mu^2$  if  $P_{\#}^1 \pi = \mu^1$  and  $P_{\#}^2 \pi = \mu^2$ . Note that a transportation plan is disintegrable in time, like Young measures. The set  $\Pi(\mu^1, \mu^2)$  of transportation plans between  $\mu^1$  and  $\mu^2$  is never empty, since it contains the measure  $\pi$  defined by  $\pi_t = \mu_t^1 \otimes \mu_t^2$  for a.a.  $t$ . For  $p \in [1, \infty)$ , the  $L^p$ -distance between  $\mu^1$  and  $\mu^2$  is

$$d_p(\mu^1, \mu^2) = \left[ \inf_{\pi \in \Pi(\mu^1, \mu^2)} \int_0^T \int_{X \times X} |v - u|^p d\pi_t(u, v) dt \right]^{1/p}. \quad (\text{A.1})$$

This distance is called the *Wasserstein distance* [6, section 3.4]. The set  $\Pi(\mu^1, \mu^2)$  is narrowly closed and if  $d_p(\mu^1, \mu^2)$  is finite, any minimizing sequence of the problem associated with (A.1) has a limit point by Prokhorov's theorem [15, theorem 11], thus by lower semi-continuity of

the duality product with a positive normal integrand, we obtain the existence of an optimal transportation plan.

If  $\mu^1$  is the Young measure associated to  $u^1 \in L([0, T]; X)$ , then for all  $\mu^2$  in  $\mathcal{M}^Y(X)$ , there is only one transportation plan  $\pi$  in  $\Pi(\mu^1, \mu^2)$ , which is, for almost all  $t$  in  $[0, T]$ , for all  $u$  and  $v$  in  $X$ ,  $\pi_t(u, v) = \delta_{u_t^1}(u)\mu_t^2(v)$ , therefore, for all  $p \in [1, \infty)$ ,

$$d_p(\mu^1, \mu^2) = \left[ \int_0^T \int_{U_R} |v - u_t^1|^p d\mu_t^2(v) dt \right]^{1/p}. \quad (\text{A.2})$$

Note that in this case, the mapping  $\mu^2 \mapsto d_s(\mu^1, \mu^2)$  is weakly-\* continuous. If  $\mu^1$  and  $\mu^2$  are both associated with  $u_1$  and  $u_2$  in  $L^p([0, T]; X)$ , then  $d_p(\mu^1, \mu^2) = \|u_2 - u_1\|_p$ .

## A.2 Young measures on $U_R$

We suppose here that  $X$  is equal to  $U_R$ , the ball of  $\mathbb{R}^m$  with radius  $R$  and center 0. We denote  $\mathcal{M}_R^Y = \mathcal{M}^Y(U_R)$ . The set  $U_R$  being compact,  $\mathcal{M}_R^Y$  is weakly-\* compact [15, theorem 1]. Moreover, the weak-\* topology and the narrow topology are equivalent [15, theorem 4].

**Differential systems controlled by Young measures** Let  $x^0 \in \mathbb{R}^n$ , let  $g : [0, T] \times X \rightarrow \mathbb{R}^n$  be Lipschitz continuous (with modulus  $A$ ), then for all  $\mu$  in  $\mathcal{M}_R^Y$ , the differential system

$$\dot{x}_t = \int_{U_R} f(x_t, u) d\mu_t(u), \quad x_0 = x^0$$

has a unique solution in  $C(0, T; \mathbb{R}^n)$ , denoted by  $x[\mu]$ .

**Lemma 37.** *The mapping  $\mu \in \mathcal{M}_R^Y \mapsto x[\mu] \in C(0, T; \mathbb{R}^n)$  is weakly-\* continuous and Lipschitz continuous for the  $L^1$ -distance of Young measures.*

*Proof.*  $\triangleright$  *Weak-\* continuity.*

Let  $\mu \in \mathcal{M}_R^Y$ , let  $(\mu_k)_k$  converges to  $\mu \in \mathcal{M}_R^Y$  for the weak-\* topology. The sequence  $(g^k)_k$  defined by

$$g_t^k = \int_0^t \int_{U_R} f(x_s[\mu], u) (d\mu_s^k(u) - d\mu_s(u)) ds$$

converges pointwise to 0. We can show with the Arzelà-Ascoli theorem that this convergence is uniform. For all  $t$  in  $[0, T]$ ,

$$\begin{aligned} |x_t[\mu^k] - x_t[\mu]| &\leq \int_0^t \int_{U_R} |f(x_s[\mu^k], u) - f(x_s[\mu], u)| d\mu_s^k(u) ds \\ &\quad + \left| \int_0^t \int_{U_R} f(x_s[\mu], u) (d\mu_s^k(u) - d\mu_s(u)) ds \right| \\ &= \int_0^t O(|x_s[\mu^k] - x_s[\mu]|) ds + o(1), \end{aligned}$$

where the estimate  $o(1)$  is uniform in time. The uniform convergence of  $x[\mu^k]$  follows from Gronwall's lemma.

$\triangleright$   *$L^1$ -Lipschitz continuity.*

Let  $\mu^1$  and  $\mu^2$  in  $\mathcal{M}_R^Y$ , let  $\pi$  be an optimal transportation plan between  $\mu^1$  and  $\mu^2$  for the

$L^1$ -distance. There exists a constant  $A \geq 0$  such that for all  $t$  in  $[0, T]$ ,

$$\begin{aligned} |x_t[\mu^2] - x_t[\mu^1]| &\leq \left| \int_0^t \int_{U_R \times U_R} f(x_s[\mu^2], v) - f(x_s[\mu^1], u) \, d\pi_s(u, v) \, ds \right| \\ &\leq \int_0^t \int_{U_R \times U_R} A(|x_s[\mu^2] - x_s[\mu^1]| + |v - u|) \, d\pi_s(u, v) \, ds \\ &\leq \int_0^t A|x_s[\mu^2] - x_s[\mu^1]| \, ds + Ad_1(\mu^1, \mu^2). \end{aligned}$$

The Lipschitz continuity follows from Gronwall's lemma.  $\square$

### A.3 Young measures on $\mathbb{R}^m$

We suppose here that  $X = \mathbb{R}^m$ . We equip  $\mathcal{M}^Y := \mathcal{M}^Y(\mathbb{R}^m)$  with the narrow topology. In the article, elements of  $\mathcal{M}^Y$  are denoted by  $\nu$ . For  $p$  in  $[1, \infty)$ , we denote by  $\mathcal{M}_p^Y$  the set of Young measures  $\nu$  in  $\mathcal{M}^Y$  with a finite  $L^p$ -norm, defined by  $\|\nu\|_p = d_p(0, \nu)$ , where  $d_p$  is the Wasserstein distance. We denote by  $\mathcal{M}_\infty^Y$  the set of Young measures with bounded support and we define the  $L^\infty$ -norm as follows:

$$\|\nu\|_\infty = \inf \{a \in \mathbb{R}, \nu([0, T] \times B(0, a)) = \nu([0, T] \times \mathbb{R}^m)\}.$$

Note the inclusion  $\mathcal{M}_\infty^Y \subset \mathcal{M}_2^Y \subset \mathcal{M}_1^Y$ .

**Lemma 38.** *The unit ball  $B_2^Y$  of  $\mathcal{M}_2^Y$  is narrowly compact.*

*Proof.* By Prokhorov's theorem [15, theorem 11],  $B_2^Y$  is precompact. The mapping  $(t, u) \mapsto |u|^2$  being a positive normal integrand, the  $L^2$ -norm is l.s.c. and therefore,  $B_2^Y$  is closed for the narrow topology. The conclusion is now proved.  $\square$

**Lemma 39.** *Let  $\psi : [0, T] \times X \rightarrow \mathbb{R}^m$  a measurable mapping be such that for almost all  $t$  in  $[0, T]$ ,  $\psi(t, \cdot)$  is continuous and such that*

$$\operatorname{ess\,sup}_{t \in [0, T]} |\psi(t, u)| = o_{|u| \rightarrow \infty}(|u|^2).$$

*Then, for all bounded sequence  $(\nu^k)_k$  in  $\mathcal{M}_2^Y$  converging narrowly to  $\nu \in \mathcal{M}_2^Y$ ,*

$$\int_0^T \int_{\mathbb{R}^m} \psi(t, u) \, d\nu_t^k(u) \, dt \xrightarrow[k \rightarrow \infty]{} \int_0^T \int_{\mathbb{R}^m} \psi(t, u) \, d\nu_t(u) \, dt \quad (\text{A.3})$$

*Proof.* The proof is inspired from [1, remark 5.3]. Let  $(\nu^k)_k$  be a bounded sequence in  $\mathcal{M}_2^Y$  converging narrowly to  $\nu \in \mathcal{M}_2^Y$ . Let

$$A = \max \{ \|\nu\|_2^2, \sup_k \{ \|\nu^k\|_2^2 \} \}.$$

Let  $\varepsilon > 0$ . Let  $B \geq 0$  be such that for almost all  $t$  in  $[0, T]$ , for all  $u$  in  $\mathbb{R}^m$ ,

$$\psi(t, u) \leq \varepsilon|u|^2 + B.$$

Then,  $\varepsilon|u|^2 + B - \psi(t, u)$  is a positive normal integrand. Thus,

$$\int_0^T \int_{\mathbb{R}^m} \varepsilon|u|^2 + B - \Psi(t, u) \, d\nu_t(u) \, dt \leq \liminf_{k \rightarrow \infty} \int_0^T \int_{\mathbb{R}^m} \varepsilon|u|^2 + B - \Psi(t, u) \, d\nu_t^k(u) \, dt.$$

and therefore,

$$\int_0^T \int_{\mathbb{R}^m} -\psi(t, u) \, d\nu_t(u) \, dt \leq \liminf_{k \rightarrow \infty} \int_0^T \int_{\mathbb{R}^m} -\psi(t, u) \, d\nu_t^k(u) \, dt + 2\varepsilon A^2.$$

To the limit when  $\varepsilon \downarrow 0$ , we obtain that

$$\int_0^T \int_{\mathbb{R}^m} \psi(t, u) \, d\nu_t(u) \, dt \geq \limsup_{k \rightarrow \infty} \int_0^T \int_{\mathbb{R}^m} \psi(t, u) \, d\nu_t^k(u) \, dt,$$

which proves the upper semi-continuity of the mapping (A.3). We prove similarly the lower semi-continuity.  $\square$

## B Justification of relaxation

This section aims at justifying the use of relaxation in the formulation of the problem. The results that we give are independent on the sensitivity analysis performed in the sequel. We introduce the value function associated with the notion of classical  $R$ -strong optimal solutions, denoted by  $\hat{V}^\eta(\theta)$ .

$$\hat{V}^\eta(\theta) := \begin{cases} \text{Min}_{u \in \mathcal{U}, \|u\| \leq R} & \phi(y_T[u, \theta], \theta), \\ \text{s.t.} & \Phi(y_T[u, \theta], \theta) \in K, \quad \|y[u, \theta] - \bar{y}\|_\infty \leq \eta. \end{cases} \quad (\text{B.1})$$

Note that for all  $\theta \geq 0$ , for all  $\eta > 0$ ,  $V^\eta(\theta) \leq \hat{V}^\eta(\theta)$ , since the space of Young measures contains the classical controls. The converse inequality would be true if there were no constraints. In that case, it would suffice to approximate any Young measure  $\mu$  with a sequence of classical controls converging to  $\mu$  for the weak-\* topology. In the constrained case, this sequence is not necessarily feasible. We prove in lemma 40 that if a given classical control is close in  $L^1$ -distance from  $\bar{u}$ , it can be restored (with another classical control). We obtain as a corollary that any feasible relaxed control close to  $\bar{u}$  in  $L^1$ -distance can be approximated by feasible classical controls. Using the results of convergence of the solutions of perturbed problems obtained in section 5, we prove the equality of  $V$  and  $\hat{V}$  for small values of  $\eta$  and  $\theta$ .

**Lemma 40.** *If  $\bar{\mu}$  is qualified, then there exist  $\delta_1 > 0$ ,  $\sigma > 0$ , and  $C_1 \geq 0$  such that for all classical control  $u$  with  $\|u - \bar{u}\|_1 \leq \delta_1$ , for all  $\theta \in [0, \sigma]$ , there exists a classical control  $u'$  such that*

$$\Phi(y_T[u', \theta], \theta) \in K \quad \text{and} \quad \|u' - u\|_1 \leq C_1 \text{dist}(\Phi(y_T[u, \theta], \theta), K).$$

*Proof.* Let  $\delta$ ,  $\sigma$ , and  $C$  be the constants given by the metric regularity theorem (theorem 9). Let us set  $\delta_1 = \frac{\delta}{2C+3}$ . Given  $\theta \in [0, \sigma]$ , let  $u$  be a classical control such that  $\|u - \bar{u}\|_1 \leq \delta_1$ . Set  $d = \text{dist}(\Phi(y_T[u, \theta], \theta), K)$ . Let us build a sequence  $(u^k)_k$  of classical controls with  $u^0 = u$  and which is such that for all  $k$ ,

$$\|u^{k+1} - u^k\|_1 \leq \frac{(C+1)d}{2^k} \quad \text{and} \quad \Phi(y_T[u^k, \theta], \theta) \leq \frac{d}{2^k}. \quad (\text{B.2})$$

By definition,  $\text{dist}(\Phi(y_T[u^0, \theta], \theta), K) \leq d/2^0$ . Let  $k$  in  $\mathbb{N}$ , let us suppose that we have built  $u^0, \dots, u^k$  such that (B.2) holds up to index  $k-1$ . Thus,  $\text{dist}(\Phi(y_T[u^k, \theta], \theta), K) \leq d/2^k$  and  $\|u^{j+1} - u^j\|_1 \leq (C+1)d/2^j$  for all  $j$  in  $\{0, \dots, k-1\}$ . Therefore,

$$d_1(u^k, \bar{\mu}) \leq \|u^k - u^0\|_1 + d_1(u^0, \bar{\mu}) \leq \sum_{j=0}^{k-1} \frac{(C+1)d}{2^j} + \delta_1 \leq 2(C+1)\delta_1 + \delta_1 \leq \delta.$$

By the metric regularity theorem, there exists a feasible relaxed control  $\mu$  such that  $d_1(u^k, \mu) \leq Cd/2^k$ . By the density of classical controls into  $\mathcal{M}_R^Y$ , by the weak-\* continuity of  $d_1(u^k, \cdot)$ , and by lemma 37, there exists a classical control  $u$  such that

$$\|u - u^k\|_1 \leq \frac{(C+1)d}{2^k} \quad \text{and} \quad \Phi(y_T[u, \theta], \theta) \leq \frac{d}{2^{k+1}}.$$

We set  $u^{k+1} = u$ . This justifies the existence of a sequence satisfying (B.2). Finally, we have built a sequence  $(u^k)_k$  of classical controls which converges for the  $L^1$ -norm. Let us denote by  $u'$  its limit, by lemma 37, it follows that

$$\text{dist}(\Phi(y_T[u', \theta], \theta), K) \leq \lim_{k \rightarrow \infty} \text{dist}(\Phi(y_T[u^k, \theta], \theta), K) = 0$$

and

$$\|u' - u\|_1 \leq \sum_{k=0}^{\infty} \|u^{k+1} - u^k\|_1 \leq \sum_{k=0}^{\infty} \frac{(C+1)d}{2^k} = 2(C+1)d.$$

The lemma holds with  $\delta_1$ ,  $\sigma$ ,  $u'$ , and  $C_1 = 2(C+1)$ .  $\square$

**Corollary 41.** *Let  $\mu \in \mathcal{M}_R^Y$  and  $\theta \in [0, \sigma]$  be such that  $d_1(\bar{u}, \mu) < \delta_1$  and such that  $\mu$  is feasible for  $\theta$ . Then, there exists a feasible sequence of classical controls  $(u^k)_k$  converging to  $\mu$  for the weak-\* topology.*

*Proof.* Let  $\mu \in \mathcal{M}_R^Y$  and  $\theta \in [0, \sigma]$  be as above. Let  $(u^k)_k$  be a sequence of classical controls converging to  $\mu$  for the weak-\* topology. Then, by lemma 37,  $\text{dist}(\Phi(y_T[u^k, \theta], \theta), K) \rightarrow 0$  and for  $k$  large enough,  $\|u^k - \bar{u}\|_1 \leq \delta_1$ . By lemma 40, we obtain a sequence of feasible controls  $(\tilde{u}^k)_k$  which is feasible for the value  $\theta$  and which such that  $\|u^k - \tilde{u}^k\|_1 \rightarrow 0$ . Then, it is easy to check that  $\tilde{u}^k$  converges to  $\mu$  for the weak-\* topology. This proves the corollary.  $\square$

In theorem 35, we have proved that under a second-order sufficient condition, for a small, positive, and fixed value of  $\eta$ , any sequence of solutions to problems  $(\mathcal{P}_\theta^{Y, \eta})$  converges to  $\bar{u}$  for the  $L^1$ - distance. Therefore, for small values of  $\eta$  and  $\theta$ , these solutions can be approximated by feasible classical controls and  $V^\eta(\theta) = \tilde{V}^\eta(\theta)$ .

## C Technical proofs

*Lemma 6.* For all  $t$  in  $[0, T]$ ,

$$\begin{aligned} |y[\mu, \theta]_t - \bar{y}_t| &= \left| \int_0^t \int_{U_R} (f(u, y_s[\mu, \theta], \theta) - f(\bar{u}_s, \bar{y}_s, 0)) \, d\mu_s(u) \, ds \right| \\ &= \int_0^t \int_{U_R} O(|u - \bar{u}_t|) + O(|y_s[\mu, \theta] - \bar{y}_s| + \theta) \, d\mu_s(u) \, ds + O(\theta) \\ &= O(d_1(\mu, \bar{\mu})) + O(\theta) + \int_0^t O(|y[\mu, \theta]_s - \bar{y}_s|) \, ds, \end{aligned}$$

whence estimate (2.9) by Gronwall's lemma. Now, set  $r = y[\mu, \theta] - (\bar{y} + \xi[\mu] + \theta\xi^\theta)$ , then, for all  $t$  in  $[0, T]$ ,

$$\begin{aligned} |r_t| &= \left| \int_0^t \int_{U_R} (f(u, y_s[\mu, \theta], \theta) - f(u, \bar{y}_s, 0)) \, d\mu_s(u) - f_{y,\theta}[s](\xi_s[\mu] + \theta\xi_s^\theta, \theta) \, ds \right| \\ &\leq \int_0^T |f(\bar{u}_s, y_s[\mu, \theta], \theta) - f(\bar{u}_s, \bar{y}_s, 0) - [f_y[s]\xi_s[\mu] + f_y[s]\theta\xi_s^\theta + f_\theta[s]]| \, ds \\ &= \int_0^t |f_y[s](y_s[\mu, \theta] - (\bar{y}_s + \xi_s[\mu] + \theta\xi_s^\theta))| \, ds + O(d_1(\mu, \bar{\mu})^2 + \theta^2) \\ &= \int_0^t O(|r_s|) \, ds + O(d_1(\mu, \bar{\mu})^2 + \theta^2), \end{aligned}$$

since

$$\begin{aligned} &\int_0^t \int_{U_R} [f(u, y_s[\mu, \theta], \theta) - f(u, \bar{y}_s, 0)] - [f(\bar{u}_s, y_s[\mu, \theta], \theta) - f(\bar{u}_s, \bar{y}_s, 0)] \, d\mu_s(u) \, ds \\ &= O(d_1(\mu, \bar{\mu})(\|y[\mu] - \bar{y}\|_\infty + \theta)) = O(d_1(\mu, \bar{\mu})^2 + \theta^2). \end{aligned}$$

Estimate (2.10) follows from Gronwall's lemma.  $\square$

*Lemma 7.* The result is a consequence of the dual representation of the  $L^1$ -distance given in [6, theorem 3.4.1]. Let  $\psi : [0, T] \times U_R \rightarrow \mathbb{R}$  be a bounded Caratheodory integrand which is such that for almost all  $t$ ,  $u \in U_R \mapsto \psi(t, u)$  is Lipschitz continuous with modulus 1. Then,

$$\begin{aligned} &\int_0^T \int_{U_R} \psi(t, u) (dS_t(\mu^0, \gamma')(u) - dS_t(\mu^0, \gamma)(u)) \, dt \\ &= \sum_{i=1}^q (\gamma'_i - \gamma_i) \int_{U_R} \psi(t, u) (d\mu^i(u) - d\mu^0(u)) \, dt \leq \sum_{i=1}^q |\gamma'_i - \gamma_i| d_1(\mu^0, \mu^i). \end{aligned}$$

The first inequality follows and the second one is obvious.  $\square$

*Lemma 19.* Note first that  $\|\mu \ominus \bar{\mu}\|_1 = d_1(\bar{\mu}, \mu)$ . Setting  $r = \xi[\mu] - z[\mu \ominus \bar{\mu}]$ , we obtain that for almost all  $t$  in  $[0, T]$ ,

$$\begin{aligned} \dot{r}_t &= f_y[t]r_t + \int_{U_R} [f(\bar{y}_t, u) - (f[t] + f_u[t](u - \bar{u}_t))] \, d\mu_t(u) \\ &= O(|r_t|) + \int_{U_R} O(|u - \bar{u}_t|^2) \, d\mu_t(u), \end{aligned}$$

thus, by Gronwall's lemma,  $\|r\|_\infty = O(\|\mu \ominus \bar{\mu}\|_2^2)$ , which proves estimate (4.2). Replacing  $\xi[\mu]$  by  $z[\mu \ominus \bar{\mu}]$  in estimates (2.9) and (2.10) of lemma 6, we obtain estimates (4.3) and (4.4).  $\square$

The following lemma will be used in the proof of theorem 27.

**Lemma 42.** Let  $(\theta_k)_k \downarrow 0$  and let  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_{+*}$  be a non-increasing right-continuous function converging to 0 at infinity. Then, there exists a sequence  $(c_k)_k$  of positive real numbers satisfying

$$c_k \theta_k \rightarrow 0 \quad \text{and} \quad \frac{f(c_k)}{c_k} = o(\theta_k). \quad (\text{C.1})$$

*Proof.* For all  $k$ , set

$$C_k = \{c \geq 0, f(c) \leq (\theta_k c)^2\}.$$

Since  $f$  is non-increasing and right-continuous,  $C_k$  is a closed interval of  $\mathbb{R}_{+*}$ . Set  $c_k = \inf C_k$ . The sequence  $(c_k)_k$  is well-defined and positive. Let  $C > 0$ , for  $k$  large enough,  $\theta_k < \sqrt{f(C)}/C$ , thus  $c_k \geq C$ . This proves that  $c_k \rightarrow +\infty$  and therefore that  $f(c_k) \rightarrow 0$ . Since  $c_k/2 < c_k$ ,

$$f(c_k/2) \geq \left(\frac{\theta_k c_k}{2}\right)^2,$$

therefore,  $c_k \theta_k \leq 2\sqrt{f(c_k/2)} \rightarrow 0$ . As a consequence, by right-continuity of  $f$ ,  $f(c_k)/c_k = \theta_k(\theta_k c_k) = o(\theta_k)$ . This proves the lemma.  $\square$

*Theorem 27.* We follow the proof of lemma 15. The main difficulty of the proof is that we need to combine the two different kind of linearizations: the standard one at the first order and the Pontryagin linearization at the second order. A second difficulty arises if  $\nu$  has a non-bounded support: in this case, a truncation must be realised. In the proof, we consider this case:  $\nu$  is non-bounded. Let  $\nu \in S(SPL_\theta)$ ,  $\xi \in F(PQ_\theta(\nu))$ , and  $(\theta_k)_k \downarrow 0$  be such that

$$\lim_{k \rightarrow \infty} \frac{V^\eta(\theta_k) - [V^\eta(0) + \theta_k \text{Val}(PL_\theta)]}{\theta_k^2} = \limsup_{\theta \downarrow 0} \frac{V^\eta(\theta) - [V^\eta(0) + \theta \text{Val}(PL_\theta)]}{\theta^2}.$$

Let  $(\tilde{\mu}_k, \alpha_k)_k$  be a sequence in  $\mathcal{M}_R^Y \times \mathbb{R}_+$  such that  $\xi = \lim \alpha_k \xi_T[\tilde{\mu}^k]$ . Extracting a subsequence of  $(\theta_k)_k$  if necessary, we can suppose that

$$\theta_k \alpha_k = o(1) \quad \text{and} \quad \alpha_k \theta_k^2 \leq 1.$$

For all  $c \geq 0$ , we define  $\nu^c$  and  $\omega^c$  the unique Young measures which are such that for all  $g \in C^0([0, T] \times \mathbb{R}^m)$ ,

$$\begin{cases} \int_0^T \int_{\mathbb{R}^m} g(t, u) d\nu^c dt = \int_0^T \int_{\mathbb{R}^m} \mathbf{1}_{|u - \bar{u}_t| > c} g(t, 0) + \mathbf{1}_{|u - \bar{u}_t| \leq c} g(t, u) d\nu_t^k dt, \\ \int_0^T \int_{\mathbb{R}^m} g(t, u) d\omega^c dt = \int_0^T \int_{\mathbb{R}^m} \mathbf{1}_{|u - \bar{u}_t| > c} g(t, u) + \mathbf{1}_{|u - \bar{u}_t| \leq c} g(t, 0) d\nu_t^k dt. \end{cases}$$

We set  $f(c) = \|\omega^c\|_2^2$ . It satisfies the assumptions of lemma 42. We obtain a sequence  $(c_k)_k$  satisfying (C.1) and we set  $\nu^k = \nu^{c_k}$  and  $\omega^c = \omega^{c_k}$ . Note that

$$f(c_k) = \|\omega^k\|_2^2 \geq c_k \|\omega^k\|_1,$$

therefore, by (42),

$$\|\omega^k\|_1 = o(\theta_k).$$

Now, in order to realize the first-order perturbation, we consider the measure  $\mu^{1,k} = \bar{u} \oplus \theta_k \nu^k$ . For almost all  $t$ , the support of  $\mu_t^{1,k}$  is included into the ball of center  $\bar{u}_t$  and radius  $c_k \theta_k$ . Since  $c_k \theta_k \rightarrow 0$ , for  $k$  large enough  $\mu^{1,k} \in \mathcal{M}_R^Y$  and since  $\alpha_k \theta_k^2 \leq 1$ , we can define

$$\mu^k = (1 - \alpha_k \theta_k^2) \mu^{1,k} + (\alpha_k \theta_k^2) \tilde{\mu}^k \in \mathcal{M}_R^Y.$$

We set  $y^k = y[\mu^k, \theta_k]$ . Let us show the expansion

$$\|y^k - (\bar{y} + \theta_k z^1[\nu] + \theta_k^2(z^2[\nu] + \xi))\|_\infty = o(\theta_k^2). \quad (\text{C.2})$$

We know that  $d_1(\bar{\mu}, \mu^k) = O(\theta_k)$ . Moreover,

$$\theta z[\nu] - z[\mu^k \ominus \bar{u}] = \alpha_k \theta_k^3 z[\nu] - \alpha_k \theta_k^2 z[\tilde{\mu}^k] = o(\theta_k),$$

thus, using lemma 19, we obtain that

$$\|y^k - (\bar{y} + \theta_k z^1[\nu])\|_\infty = o(\theta_k).$$

Let us set  $r^k = y^k - (\bar{y} + \theta_k z^1[\nu] + \theta_k^2(z^2[\nu] + \alpha_k \xi[\tilde{\mu}^k]))$ . Then,

$$\begin{aligned} r_t^k &= (1 - \alpha_k \theta_k^2) \int_0^t \int_{\mathbb{R}^m} f(\bar{u}_s + \theta_k u, y_s^k, \theta_k) - f[s] \, d\nu_t^k(u) \, ds \\ &\quad - \int_0^t \int_{\mathbb{R}^m} (\theta_k f'[s](u, z_s^1[\nu], 1) + \frac{1}{2} \theta_k^2 f''[s](u, z_s^1[\nu], 1)^2) \, d\nu_t(u) \, ds \\ &\quad - \theta_k^2 \int_0^t f_y[s](z_s^2[\nu] + \alpha_k \xi[\tilde{\mu}^k]_s) \, ds \\ &\quad + \alpha_k \theta_k^2 \int_0^t \int_{U_R} f(u, y_s^k, \theta_k) - (f(u, \bar{y}_s, 0) - f[s]) - f[s] \, d\tilde{\mu}_s^k(u) \, ds \\ &= \int_0^t \int_{\mathbb{R}^m} \left( f'[s](\theta_k u, (y_s^k - \bar{y}_s), \theta_k) + \frac{1}{2} f''[s](\theta_k u, y_s^k - \bar{y}_s, \theta_k)^2 \right) \, d\nu_t^k(u) \, ds \\ &\quad - \int_0^t \int_{\mathbb{R}^m} \left( f'[s](\theta_k u, \theta_k z_s^1[\nu], \theta_k) + \frac{1}{2} f''[s](\theta_k u, \theta_k z_s^1[\nu], \theta_k)^2 \right) \, d\nu_t(u) \, ds \\ &\quad - \theta_k^2 \int_0^t f_y[s](z_s^2[\nu] + \alpha_k \xi[\tilde{\mu}^k]_s) \, ds \\ &\quad + \alpha_k \theta_k^2 \int_0^t \int_{U_R} (f(u, y_s^k, \theta_k) - f(u, \bar{y}_s, \theta_k)) \, d\tilde{\mu}_s^k(u) \, ds + o(\theta_k^2) \\ &= \int_0^t \int_{\mathbb{R}^m} (\theta_k f_u[s] u + \frac{1}{2} \theta_k^2 f''[s](u, z_s^1[\nu], 1)^2) (d\nu_t^k(u) - d\nu_t(u)) \, dt \\ &\quad + \int_0^t f_y[s] r_s^k \, ds + o(\theta_k^2) \\ &= \int_0^t f_y[s] r_s^k \, ds + O(\theta_k \|\omega^k\|_1) + O(\theta_k^2 \|\omega^k\|_2^2) + o(\theta_k^2) \\ &= \int_0^t f_y[s] r_s^k \, ds + o(\theta_k^2). \end{aligned}$$

By Gronwall's lemma,  $\|r^k\|_\infty = o(\theta_k^2)$  and since  $\alpha_k \xi_T[\mu^k] \rightarrow \xi$ , expansion (C.2) holds. As a consequence, the following second-order expansion hold:

$$\begin{aligned} \phi(y_T[\mu^k], \theta_k, \theta_k) &= \phi(\bar{y}_T, 0) + \theta_k \phi'(\bar{y}_T, 0)(z_T^1[\nu], 1) \\ &\quad + \theta_k^2 \left[ \frac{1}{2} \phi''(\bar{y}_T, 0)(z_T^1[\nu], 1)^2 + \phi_{y_T}(\bar{y}_T, 0)(z_T^2[\nu] + \xi) \right] + o(\theta_k^2), \end{aligned} \quad (\text{C.3})$$

and the same expansion holds for  $\Phi(y_T[\mu^k], \theta_k, \theta_k)$ . Therefore,  $\text{dist}(\Phi(y_T^k), K) = o(\theta_k^2)$ . By the metric regularity theorem (theorem 9) and by lemma 37, there exists a sequence  $\hat{\mu}^k$  of feasible

controls such that  $d_1(\mu^k, \hat{\mu}^k) = o(\theta_k^2)$  and such that (C.3) holds for  $\phi(y_T[\hat{\mu}^k, \theta_k], \theta_k)$ . Minimizing with respect to  $\xi$ , we obtain that

$$\limsup_{\theta \downarrow 0} \frac{V^\eta(\theta) - [V^\eta(0) + \theta \text{Val}(PL_\theta)]}{\theta^2} \leq \text{Val}(PQ_\theta(\nu)).$$

Minimizing with respect to  $\nu$ , we obtain the theorem.  $\square$

*Lemma 28.* Expanding the difference of Lagrangians up to the second order, we obtain

$$\Delta\Phi^k = \Phi'[\lambda](\bar{y}_T, 0)(\delta y_T^k, \theta_k) + \frac{1}{2}\Phi''[\lambda](\delta y_T^k, \theta_k)^2 + o(\theta_k^2 + |\delta y_T^k|^2). \quad (\text{C.4})$$

Then,

$$\begin{aligned} \Phi_{y_T}[\lambda](\bar{y}_T, 0)\delta y_T^k &= [p_t^\lambda \delta y_t^k]_0^T = \int_0^T (p_t^\lambda \dot{\delta y}_t^k + \dot{p}_t^\lambda \delta y_t^k) dt \\ &= \int_0^T \left( \int_{U_R} (H(u, y_t^k, \theta_k) - H[t]) d\mu_t^k(u) - H_y[t]\delta y_t^k \right) d\mu_t^k(u) dt. \end{aligned} \quad (\text{C.5})$$

Expanding the difference of Hamiltonians, we obtain that

$$\begin{aligned} &\int_0^T \int_{U_R} (H(u, y_t^k, \theta_k) - H[t]) d\mu_t^k(u) dt \\ &= \int_0^T \int_{U_R} [H(u, y_t^k, \theta_k) - H(u, \bar{y}_t, 0)] + [H(u, \bar{y}_t, 0) - H[t]] d\mu_t^k(u) dt \\ &= \int_0^T \int_{U_R} H_{(y,\theta)^2}(u, \bar{y}_t, 0)(\delta y_t^k, \theta_k) + \frac{1}{2}H_{(y,\theta)^2}(u, \bar{y}_t, 0)(\delta y_t^k, \theta_k)^2 d\mu_t^k(u) dt \\ &\quad + \int_0^T \int_{U_R} (H(u, \bar{y}_t, 0) - H[t]) d\mu_t^k(u) dt + o(\theta_k^2 + R_{1,k}^2). \end{aligned} \quad (\text{C.6})$$

Moreover,

$$\begin{aligned} &\int_0^T \int_{U_R} |H_{(y,\theta)^2}(u, \bar{y}_t, 0)(\delta y_t^k, \theta_k)^2 - H_{(y,\theta)^2}[t](\delta y_t^k, \theta_k)^2| d\mu_t^k(u) dt \\ &= O(R_{1,k}(R_{1,k}^2 + \theta_k^2)). \end{aligned} \quad (\text{C.7})$$

and

$$R_{1,k}(R_{1,k}^2 + \theta_k^2) \leq R_{1,k}^3 + \frac{1}{2}(R_{1,k}^2 + \theta_k^2)\theta_k = o(R_{1,k}^2 + \theta_k^2). \quad (\text{C.8})$$

Finally, remember that  $\text{Val}(PL_\theta) = \int_0^T H_\theta[t] dt + \Phi_\theta[\lambda](\bar{y}_T, 0)$ . Combining expansions (C.4-C.8), we obtain that

$$\begin{aligned} \Delta\Phi^k &= \text{Val}(PL_\theta)\theta_k + \int_0^T \int_{U_R} (H[p_t^\lambda](u, \bar{y}_t, 0) - H[p_t^\lambda][t]) d\mu_t(u) dt \\ &\quad + \int_0^T \int_{U_R} (H_y[p_t^\lambda](u, \bar{y}_t, 0) - H_y[p_t^\lambda][t])\delta y_t^k d\mu_t^k(u) dt \end{aligned} \quad (\text{C.9a})$$

$$+ \int_0^T \int_{U_R} (H_\theta[p_t^\lambda](u, \bar{y}_t, 0) - H_\theta[p_t^\lambda][t])\theta_k d\mu_t^k(u) dt \quad (\text{C.9b})$$

$$+ \frac{1}{2} \left[ \int_0^T H_{(y,\theta)^2}[p_t^\lambda][t](\delta y_t^k, \theta_k)^2 dt + \Phi''[\lambda](\delta y_T^k, \theta_k)^2 \right] + o(\theta_k^2 + R_{1,k}^2). \quad (\text{C.9c})$$

We have already proved in lemma 19 the following estimate:

$$\|\delta y^k - (z^k + \theta_k \xi^\theta)\|_\infty = O(R_{1,k}^2 + \theta_k^2).$$

Therefore, we can replace  $\delta y^k$  by its standard expansion  $z^{1,k}$  in terms (C.9a) and (C.9c). The errors that we make are respectively of order  $R_{1,k}(R_{1,k}^2 + \theta_k^2)$  and  $R_{1,k}^2(R_{1,k} + \theta_k)$ . As we prove in estimate (C.8), the first term is of order  $o(R_{1,k}^2 + \theta_k^2)$ . The estimate (5.2) holds.

Expansion (5.3) follows from (C.9). We replace respectively terms (C.9a), (C.9b), and (C.9c) by the following estimates:  $O(R_{1,k}\|\delta y^k\|_\infty)$ ,  $O(R_{1,k}\theta_k)$ ,  $O(\|\delta y^k\|_\infty^2 + \theta_k^2)$ , and the estimate, since the sequence  $(R_{1,k})_k$  is bounded.  $\square$

*Lemma 33.* Let  $\nu \in \mathcal{M}_2^Y$  and let  $(\nu^k)_k$  be a bounded sequence in  $\mathcal{M}_2^Y$  narrowly converging to  $\bar{\nu}$ .

$\triangleright$  *Narrow continuity of  $\nu \mapsto z[\nu]$ .*

We set, for almost all  $t$ ,

$$v_t^k = \int_{\mathbb{R}^m} u \, d\nu_t^k(u) \quad \text{and} \quad v_t = \int_{\mathbb{R}^m} u \, d\nu_t(u).$$

It is easy to check that  $\bar{\nu}$  and  $v^k \in L^2(0, T; \mathbb{R}^m)$ . Moreover,  $z[v^k] = z[\nu^k]$  and  $z[v] = z[\nu]$ . Let us check that  $v^k$  converges to  $\bar{\nu}$  for the weak topology of  $L^2$ . Let  $h \in L^2(0, T; \mathbb{R}^m)$ , then by definition of the narrow topology,

$$\int_0^T h_t v_t^k \, dt = \int_0^T \int_{\mathbb{R}^m} h_t \, d\nu_t^k(u) \, dt \rightarrow \int_0^T \int_{\mathbb{R}^m} h_t \, d\nu_t(u) \, dt = \int_0^T h_t v_t^k \, dt. \quad (\text{C.10})$$

This proves the weak convergence of  $v^k$ . The mapping  $v \in L^2(0, T; \mathbb{R}^m) \mapsto z[v] \in H^1(0, T; \mathbb{R}^n)$  being linear continuous,  $z[v^k]$  converges for the weak topology of  $H^1$ . Since  $(z[v^k])_k$  is bounded in  $H^1$  and by the compact embedding of this space in  $C(0, T; \mathbb{R}^n)$ ,  $z[v^k]$  converges uniformly to  $z[v]$ .

$\triangleright$  *Narrow lower semi-continuity of  $\Omega[\lambda]$  and  $\Omega^\theta[\lambda]$ .*

Let  $\lambda \in \Lambda^P$ . Let us decompose  $\Omega^\theta[\lambda]$  into three terms,  $Q_0$ ,  $Q_1$ , and  $Q_2$  with

$$\begin{aligned} Q_0[\lambda](\nu) &= \int_0^T H_{y,\theta}[t](z[\nu], 1)^2 \, dt + \Phi''[\lambda](\bar{y}_T, 0)(z_T[\nu], 1)^2, \\ Q_1[\lambda](\nu) &= 2 \int_0^T \int_{\mathbb{R}^m} H_{u,\theta}[t](u, 1) + H_{u,y}[t](u, z[\nu]) \, d\nu_t(u) \, dt, \\ Q_2[\lambda](\nu) &= \int_0^T \int_{\mathbb{R}^m} H_{uu}[t](u)^2 \, d\nu_t(u) \, dt. \end{aligned}$$

Since  $z[\nu^k]$  converges uniformly to  $z[\nu]$  and since the sequence  $(\nu^k)_k$  is bounded, we obtain by lemma 39 that  $Q_0[\lambda](\nu^k)$  and  $Q_1[\lambda](\nu^k)$  converges resp. to  $Q_0[\lambda](\nu)$  and  $Q_1[\lambda](\nu)$ . Since  $\lambda \in \Lambda^P$ , the integrand  $H_{uu}[t](u)^2$  of  $Q_2[\lambda]$  is nonnegative,  $Q_2[\lambda]$  is lower semi-continuous for the narrow topology. Finally, we obtain the lower semi-continuity of  $\Omega^\theta[\lambda]$  and similarly, the one of  $\Omega[\lambda]$ .

$\triangleright$  *Strong convergence to 0.*

Suppose now that  $(\nu^k)_k$  converges narrowly to 0 and that  $\Omega[\bar{\lambda}](\nu^k) \rightarrow 0$ . Then, necessarily,  $Q_2[\bar{\lambda}](\nu^k) \rightarrow 0$ . From hypothesis 31.1, we obtain the inequality  $2\alpha\|\nu\|_2^2 \leq Q_2[\bar{\lambda}](\nu)$  and the lemma is now proved.  $\square$

## References

- [1] E. J. Balder. Lectures on Young measure theory and its applications in economics. 1998.
- [2] E. N. Barron and R. Jensen. Relaxation of constrained control problems. *SIAM J. Control Optim.*, 34(6):2077–2091, 1996.
- [3] J. F. Bonnans and R. Cominetti. Perturbed optimization in Banach spaces. I. A general theory based on a weak directional constraint qualification. *SIAM J. Control Optim.*, 34(4):1151–1171, 1996.
- [4] J. F. Bonnans and N. P. Osmolovskii. Second-order analysis of optimal control problems with control and initial-final state constraints. *J. Convex Anal.*, 17(3-4):885–913, 2010.
- [5] J. F. Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer-Verlag, New York, 2000.
- [6] C. Castaing, P. Raynaud de Fittede, and M. Valadier. *Young measures on topological spaces*, volume 571 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2004.
- [7] A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings*. Springer Monographs in Mathematics. Springer, Dordrecht, 2009.
- [8] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam, 1976.
- [9] H. Frankowska and F. Rampazzo. Relaxation of control systems under state constraints. *SIAM J. Control Optim.*, 37(4):1291–1309 (electronic), 1999.
- [10] R. V. Gamkrelidze. *Principles of optimal control theory*. Plenum Press, New York, revised edition, 1978.
- [11] E. J. McShane. Relaxed controls and variational problems. *SIAM J. Control*, 5:438–485, 1967.
- [12] A. A. Milyutin and N. P. Osmolovskii. *Calculus of variations and optimal control*, volume 180 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1998.
- [13] S. M. Robinson. Regularity and stability for convex multivalued functions. *Math. Oper. Res.*, 1(2):130–143, 1976.
- [14] C. Ursescu. Multifunctions with convex closed graph. *Czechoslovak Math. Journal*, 25(100)(3):438–441, 1975.
- [15] M. Valadier. Young measures. In *Methods of nonconvex analysis (Varenna, 1989)*, volume 1446 of *Lecture Notes in Math.*, pages 152–188. Springer, Berlin, 1990.
- [16] M. Valadier. A course on Young measures. *Rend. Istit. Mat. Univ. Trieste*, 26(suppl.):349–394 (1995), 1994. Workshop on Measure Theory and Real Analysis (Italian) (Grado, 1993).
- [17] J. Warga. Necessary conditions for minimum in relaxed variational problems. *J. Math. Anal. Appl.*, 4:129–145, 1962.
- [18] J. Warga. Relaxed variational problems. *J. Math. Anal. Appl.*, 4:111–128, 1962.

- [19] J. Warga. *Optimal control of differential and functional equations*. Academic Press, New York, 1972.
- [20] L. C. Young. *Lectures on the calculus of variations and optimal control theory*. W. B. Saunders Co., Philadelphia, 1969.



**RESEARCH CENTRE  
SACLAY – ÎLE-DE-FRANCE**

1 rue Honoré d'Estienne d'Orves  
Bâtiment Alan Turing  
Campus de l'École Polytechnique  
91120 Palaiseau

Publisher  
Inria  
Domaine de Voluceau - Rocquencourt  
BP 105 - 78153 Le Chesnay Cedex  
[inria.fr](http://inria.fr)

ISSN 0249-6399