



HAL
open science

First and second order necessary conditions for stochastic optimal control problems

Joseph Frédéric Bonnans, Francisco J. Silva

► **To cite this version:**

Joseph Frédéric Bonnans, Francisco J. Silva. First and second order necessary conditions for stochastic optimal control problems. [Research Report] RR-7454, 2010. inria-00537227v1

HAL Id: inria-00537227

<https://inria.hal.science/inria-00537227v1>

Submitted on 17 Nov 2010 (v1), last revised 25 Jun 2011 (v2)

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.



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

*First and second order necessary conditions for
stochastic optimal control problems*

J. Frédéric Bonnans — Francisco J. Silva

N° 7454

Novembre 2010

Thème NUM



*R*apport
de recherche

First and second order necessary conditions for stochastic optimal control problems

J. Frédéric Bonnans* , Francisco J. Silva†

Thème NUM — Systèmes numériques
Équipes-Projets Commands

Rapport de recherche n° 7454 — Novembre 2010 — 23 pages

Abstract: In this work we consider a stochastic optimal control problem with convex control constraints. Using the variational approach, we are able to obtain first and second order expansions for the state and cost function, around a local minimum. This fact allows us to prove general first order necessary condition and, under a geometrical assumption over the constraint set, second order necessary conditions are also established.

Key-words: Stochastic optimal control, variational approach, first and second order optimality conditions, polyhedral constraints.

* INRIA-Saclay and CMAP, École Polytechnique, 91128 Palaiseau, France (Fred-eric.Bonnans@inria.fr)

† INRIA-Saclay and CMAP, École Polytechnique, 91128 Palaiseau, France (silva@cmap.polytechnique.fr)

Conditions d'optimalité du premier et second ordre en commande optimale stochastique

Résumé : Nous considérons un problème de contrôle optimal stochastique avec des contraintes convexes sur la commande. L'approche dite *variationnelle* nous permet d'obtenir un développement au premier et au second ordre pour l'état et la fonction de coût, autour d'un minimum local. Avec ces développements on peut montrer des conditions générales d'optimalité de premier ordre et, sous une hypothèse géométrique sur l'ensemble des contraintes, des conditions nécessaires du second ordre sont aussi établies.

Mots-clés : Commande optimale stochastique, approche variationnelle, conditions d'optimalité de premier et second ordre, contraintes polyédriques.

1 Introduction

Because of its wide range of applications (e.g. in mathematical finance), stochastic optimal control theory is a very active research domain. In this work we consider the following type of stochastic optimal control problem

$$\begin{aligned} & \text{Min } \mathbb{E} \left(\int_0^T \ell(t, y(t), u(t)) dt + \phi(y(T)) \right) \\ & \text{s.t. } \quad dy(t) = f(t, y(t), u(t)) dt + \sigma(t, y(t), u(t)) dW(t) \\ & \quad \quad y(0) = y_0, \quad u(t, \omega) \in U \text{ for a.a. } (t, \omega), \end{aligned} \quad (\mathcal{SP})$$

where U is a nonempty, closed and convex subset of \mathbb{R}^m and we suppose that the above stochastic differential equation (SDE) is well posed.

As in the case of deterministic optimal control problems, there are two main approaches to study problem (\mathcal{SP}) . The first one is the global approach, based in the Bellman's dynamic programming principle, which yields that the value function of (\mathcal{SP}) is the unique viscosity solution of an associated second order Hamilton-Jacobi-Bellman equation. For a complete account of this point of view, widely used in practical computations, we refer the reader to the books [8, 17, 18]. The second approach is the variational one, which consists in to analyse the local behavior of the value function under small perturbations of a local minimum. Using this technique Kushner [12, 11, 13] Bensoussan [1, 2], Bismut [3, 4, 5] and Haussmann [10] obtained natural extensions of Pontryagin maximum principle to the stochastic case, that were generalized by Peng [16]. Relations between the global and variational approach are studied in [19].

Nevertheless, to the best of our knowledge, nothing has been said about second order optimality conditions. Using the variational technique we are able to obtain first and second order expansions for the cost function, which are expressed in terms of the derivatives of the Hamiltonian of problem (\mathcal{SP}) . The main tool is a kind of generalization of Gronwall's lemma for the SDEs (proposition 1) obtained by Mou and Yong [15], which allows to expand the cost with respect to directions belonging to a more regular space than the control space. A similar idea was applied in [6] in the context of state constrained optimal control problems. By a density argument, we establish first order optimality conditions, which include the case of not necessarily local constraints. In addition, under a polyhedricity assumption (see [9, 14]), we obtain second order necessary conditions which are the natural extensions of their deterministic counterparts.

The article is organized as follows: After introducing the standard notations and assumptions in section 2, we obtain in section 3 first and second order expansions for the state and cost function. In section 4, first and second order necessary conditions are proved and explicit results are given for the case of box constraints. Finally, a discussion about a non gap second order sufficient condition is given in section 5.

2 Notations, assumptions and problem statement

Let us first fix some standard notation. For a x in a Euclidean space we will write x^i for its i -th coordinate and $|x|$ for its Euclidean norm. Let $T > 0$ and consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, on which a d -dimensional

($d \in \mathbb{N}^*$) Brownian motion $W(\cdot)$ is defined with $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$ being its natural filtration, augmented by all \mathbb{P} -null sets in \mathcal{F} . Let $(X, \|\cdot\|_X)$ be a Banach space and for $\beta \in [1, \infty)$ set

$$\begin{aligned} L^\beta(\Omega; X) &:= \left\{ v : \Omega \rightarrow X; v \text{ is } \mathcal{F} \text{-measurable and } \mathbb{E} \left(\|v(\omega)\|_X^\beta \right) < \infty \right\}, \\ L^\infty(\Omega; X) &:= \left\{ v : \Omega \rightarrow X; v \text{ is } \mathcal{F} \text{-measurable and } \text{ess sup}_{\omega \in \Omega} \|v(\omega)\|_X < \infty \right\}. \end{aligned}$$

For $\beta, p \in [1, \infty]$ and $m \in \mathbb{N}$ let us define

$$L_{\mathcal{F}}^{\beta,p} := \left\{ v \in L^\beta(\Omega; L^p([0, T]; \mathbb{R}^m)); (t, \omega) \rightarrow v(t, \omega) := v(\omega)(t) \text{ is } \mathbb{F}\text{-adapted} \right\}.$$

We endow these space with the norms

$$\|v\|_{\beta,p} := \left[\mathbb{E} \left(\|v(\omega)\|_{L^p([0, T]; \mathbb{R}^m)}^\beta \right) \right]^{\frac{1}{\beta}} \quad \text{and} \quad \|v\|_{\infty,p} := \text{ess sup}_{\omega \in \Omega} \|v(\omega)\|_{L^p([0, T]; \mathbb{R}^m)}.$$

For the sake of clarity, when the context is clear, the statement “for a.a. $t \in [0, T]$, a.s. $\omega \in \Omega$ (\mathbb{P} -a.s.)” will be simplified to “for a.a. (t, ω) ”. We will write $L_{\mathcal{F}}^p := L_{\mathcal{F}}^{p,p}$ and $\|\cdot\|_p := \|\cdot\|_{p,p}$. The spaces $L_{\mathcal{F}}^{\beta,p}$ endowed with the norms $\|\cdot\|_{\beta,p}$ are Banach spaces and for the specific case $p = 2$ the space $L_{\mathcal{F}}^2$ is a Hilbert space. We will write $\langle \cdot, \cdot \rangle_2$ for the obvious scalar product. Evidently, for $\beta \in [1, \infty]$ and $1 \leq p_1 \leq p \leq p_2 \leq \infty$, there exist positive constants $c_{\beta,p_1}, c_{\beta,p_2}, c_{p_1,\beta}, c_{p_2,\beta}$ such that

$$c_{\beta,p_1} \|v\|_{\beta,p_1} \leq \|v\|_{\beta,p} \leq c_{\beta,p_2} \|v\|_{\beta,p_2}, \quad c_{p_1,\beta} \|v\|_{p_1,\beta} \leq \|v\|_{p,\beta} \leq c_{p_2,\beta} \|v\|_{p_2,\beta}$$

For a function $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \ni (t, y, u, \omega) \rightarrow \psi(t, y, u, \omega) \in \mathbb{R}^n$ which is C^2 with respect to (y, u) , set $\psi_y(t, y, u, \omega) := D_y \psi(t, y, u, \omega)$ and $\psi_u(t, y, u, \omega) := D_u \psi(t, y, u, \omega)$. As usual, when the context is clear, we will systematically omit the ω argument in the defined functions. Now let $z \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ be variations associated with y and u respectively. The second derivatives of ψ are written in the following form

$$\begin{aligned} \psi_{yy}(t, y, u)z^2 &:= D_{yy}^2 \psi(t, y, u)(z, z); \quad \psi_{uu}(t, y, u)v^2 := D_{uu}^2 \psi(t, y, u)(v, v); \\ \psi_{yu}(t, y, u)zv &:= D_{yu}^2 \psi(t, y, u)(z, v). \end{aligned}$$

Consider the maps $f, \sigma^i : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}^n$ ($i = 1, \dots, d$). These maps will define the dynamics for our problem. Let us assume that:

(H1) [Assumptions for the dynamics] The maps $\psi = f, \sigma^i$ satisfy:

- (i) The maps are $\mathcal{B}([0, T] \times \mathbb{R}^n \times \mathbb{R}^m) \otimes \mathcal{F}_T$ -measurable.
- (ii) For all $(y, u) \in \mathbb{R}^n \times \mathbb{R}^m$ the process $[0, T] \ni t \rightarrow \psi(t, y, u) \in \mathbb{R}^n$ is \mathbb{F} -adapted.
- (iii) For almost all $(t, \omega) \in [0, T] \times \Omega$ the mapping $(y, u) \rightarrow \psi(t, y, u, \omega)$ is C^3 . Moreover, we assume that there exists a constant $L_1 > 0$ such that for almost all (t, ω)

$$\left\{ \begin{array}{l} |\psi(t, y, u, \omega)| \leq L_1 (1 + |y| + |u|), \\ |\psi_y(t, y, u, \omega)| + |\psi_u(t, y, u, \omega)| \leq L_1, \\ |\psi_{yy}(t, y, u, \omega)| + |\psi_{yu}(t, y, u, \omega)| + |\psi_{uu}(t, y, u, \omega)| \leq L_1 \\ |D^2 \psi(t, y, u, \omega) - D^2 \psi(t, y', u', \omega)| \leq L_1 (|y - y'| + |u - u'|). \end{array} \right. \quad (1)$$

Let us define $\sigma(t, y, u) := (\sigma^1(t, y, u), \dots, \sigma^d(t, y, u)) \in \mathbb{R}^{n \times d}$. For variations $z \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$, associated with y and u , set

$$\begin{aligned}\sigma_y(t, y, u)z &:= (\sigma_y^1(t, y, u)z, \dots, \sigma_y^d(t, y, u)z), \\ \sigma_{yy}(t, y, u)z^2 &:= (\sigma_{yy}^1(t, y, u)z^2, \dots, \sigma_{yy}^d(t, y, u)z^2),\end{aligned}\quad (2)$$

with analogous definitions for $\sigma_u(t, y, u)v$, $\sigma_{yu}(t, y, u)zv$ and $\sigma_{uu}(t, y, u)v^2$.

For every $\beta \geq 1$, let us define the space \mathcal{Y}^β as

$$\mathcal{Y}^\beta := \{y \in L^\beta(\Omega; C([0, T]; \mathbb{R}^n)); (t, \omega) \rightarrow y(t, \omega) := y(\omega)(t) \text{ is } \mathbb{F}\text{-adapted}\}.$$

Let $y_0 : \Omega \rightarrow \mathbb{R}^n$ be \mathcal{F}_0 measurable and such that $\mathbb{E}(|y_0|^2) < \infty$. Under **(H1)**, we have that for every $u \in L_{\mathcal{F}}^{\beta, 2}$ the SDE

$$\begin{aligned}dy(t) &= f(t, y(t), u(t))dt + \sigma(t, y(t), u(t))dW(t), \\ y(0) &= y_0\end{aligned}\quad (3)$$

is well posed. In fact (see [15, Proposition 2.1]):

Proposition 1. *Suppose that **(H1)** holds. Then, there exists $C > 0$ such that for every $u \in L_{\mathcal{F}}^{\beta, 2}$ ($\beta \geq 1$) equation (3) has a unique solution $y \in \mathcal{Y}^\beta$ with continuous trajectories a.s. and*

$$\mathbb{E} \left(\sup_{t \in [0, T]} |y(t)|^\beta \right) \leq C \mathbb{E} \left(|y_0|^\beta + \|f(\cdot, 0, u(\cdot))\|_{\beta, 1}^\beta + \|\sigma(\cdot, 0, u(\cdot))\|_{\beta, 2}^\beta \right). \quad (4)$$

Remark 1. *Note that by the first condition in (1), the right hand side of (4) is finite.*

Now, let us consider maps $\ell : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}$ and $\phi : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$. These maps will define the cost function of our problem. We assume:

(H2) [Assumptions for the cost maps] It holds that:

(i) The maps ℓ and ϕ are respectively $\mathcal{B}([0, T] \times \mathbb{R}^n \times \mathbb{R}^m) \otimes \mathcal{F}_T$ and $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{F}_T$ measurables.

(ii) For all $(y, u) \in \mathbb{R}^n \times \mathbb{R}^m$ the process $[0, T] \ni t \rightarrow \ell(t, y, u) \in \mathbb{R}$ is \mathbb{F} -adapted.

(iii) For almost all (t, ω) the maps $(y, u) \rightarrow \ell(t, y, u, \omega)$ and $y \rightarrow \phi(y, \omega)$ are \mathcal{C}^2 .

In addition, there exists $L_2 > 0$ such that:

$$\left\{ \begin{array}{l} |\ell(t, y, u, \omega)| \leq L_2 (1 + |y| + |u|)^2, \quad |\phi(y, \omega)| \leq L_2 (1 + |y|)^2, \\ |\ell_y(t, y, u, \omega)| + |\ell_u(t, y, u, \omega)| \leq L_2 (1 + |y| + |u|), \\ |\ell_{yy}(t, y, u, \omega)| + |\ell_{yu}(t, y, u, \omega)| + |\ell_{uu}(t, y, u, \omega)| \leq L_2, \\ |D^2 \ell(t, y, u, \omega) - D^2 \ell(t, y', u', \omega)| \leq L_2 (|y - y'| + |u - u'|), \\ |\phi_y(y, \omega)| \leq L_2 (1 + |y|) \\ |\phi_{yy}(y, \omega)| \leq L_2, \quad |\phi_{yy}(y, \omega) - \phi_{yy}(y', \omega)| \leq L_2 (|y - y'|). \end{array} \right. \quad (5)$$

Remark 2. *The assumptions above include the important case when the cost function is quadratic in (y, u) .*

In some of the results obtained in the sequel it will be useful to strengthen the second and fifth conditions in (5). In fact, as we will see in sections 3 and 4, under the assumption below the results obtained will be the natural extensions of the well know deterministic results.

[Lipschitz cost] There exists $C_\ell, C_\phi > 0$ such that for almost all $(t, \omega) \in [0, T] \times \Omega$ and for all $(y, u), (y', u') \in \mathbb{R}^n \times \mathbb{R}^m$ we have

$$\begin{aligned} |\ell(t, y, u, \omega) - \ell(t, y', u', \omega)| &\leq C_\ell (|u - u'| + |y - y'|), \\ |\phi(y, \omega) - \phi(y', \omega)| &\leq C_\phi |y - y'|. \end{aligned} \quad (6)$$

For every $u \in L^2_{\mathcal{F}}$ denote by $y_u \in \mathcal{Y}^2$ the solution of (3). Let us define the function $J : L^2_{\mathcal{F}} \rightarrow \mathbb{R}$ by

$$J(u) = \mathbb{E} \left[\int_0^T \ell(t, y_u(t), u(t)) dt + \phi(y_u(T)) \right]. \quad (7)$$

Note that, in view of the first condition in (5) and estimate (4) the function J is well defined. Let \mathcal{U} be a nonempty closed and convex subset of $L^2_{\mathcal{F}}$ and consider the problem

$$\text{Min } J(u) \quad \text{subject to } u \in \mathcal{U}. \quad (\mathcal{SP})$$

3 Expansions for the state and cost function

From now on we fix $\bar{u} \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^m)$ and set $\bar{y} := y_{\bar{u}}$. We also suppose that assumptions **(H1)** and **(H2)** hold. For $\psi = f, \sigma$ and $t \in [0, T]$, define

$$\begin{aligned} \psi_y(t) &= \psi_y(t, \bar{y}(t), \bar{u}(t)); \quad \psi_u(t) = \psi_u(t, \bar{y}(t), \bar{u}(t)), \quad \psi_{yu}(t) = \psi_{yu}(t, \bar{y}(t), \bar{u}(t)); \\ \psi_{yy}(t) &= \psi_{yy}(t, \bar{y}(t), \bar{u}(t)); \quad \psi_{uu}(t) = \psi_{uu}(t, \bar{y}(t), \bar{u}(t)). \end{aligned}$$

Let $\beta \in [1, \infty]$ and $v \in L^{\beta, 2}_{\mathcal{F}}$. We define $y_1[\bar{u}](v) \in \mathcal{Y}^\beta$ as the unique solution of

$$\begin{aligned} dy_1(t) &= [f_y(t)y_1(t) + f_u(t)v(t)]dt + [\sigma_y(t)y_1(t) + \sigma_u(t)v(t)]dW(t), \\ y_1(0) &= 0. \end{aligned} \quad (8)$$

The second assumption in (1) and proposition 1 yields that the mapping $v \in L^{\beta, 2}_{\mathcal{F}} \rightarrow y_1[\bar{u}](v) \in \mathcal{Y}^\beta$ is well defined. If the context is clear, for notational convenience we will write $y_1 = y_1[\bar{u}](v)$. Also, let us define $\delta y = \delta y[\bar{u}](v)$ and $d_1 = d_1[\bar{u}](v)$ by

$$\delta y := y_{\bar{u}+v} - \bar{y}, \quad d_1 := \delta y - y_1. \quad (9)$$

Our aim now is to obtain a first order expansion of J around \bar{u} . For this purpose it will be useful to obtain bounds for y_1 , δy and d_1 . The main tool for obtaining such bounds is the following corollary of proposition 1, whose proof is straightforward.

Corollary 2. *Let $A_1, A_2 \in L^\infty_{\mathcal{F}}([0, T]; \mathbb{R}^{n \times n})$, $B_1^i \in L^{\beta, 2}_{\mathcal{F}}([0, T]; \mathbb{R}^n)$ and $B_2^i \in L^\infty_{\mathcal{F}}([0, T]; \mathbb{R}^{n \times d})$ for $i = 1, 2$. Assume that there exists a constant $K > 0$ such that*

$$\|B_1^1\|_{\beta, 1} \leq K \|B_2^1\|_{\beta, 2}, \quad (10)$$

Then, omitting time from function arguments, for every $w \in L^{\beta,2}$, the SDE

$$\begin{aligned} dz &= [A_1 z + B_1^1 + B_1^2 w] dt + [A_2 z + B_2^1 + B_2^2 w] dW(t) \\ z(0) &= 0, \end{aligned} \quad (11)$$

has a unique solution in \mathcal{Y}^β and the following estimate holds

$$\mathbb{E} \left(\sup_{t \in [0, T]} |z(t)|^\beta \right) = \begin{cases} O \left(\max \left\{ \|B_2^1\|_{\beta,2}^\beta, \|w\|_{\beta,1}^\beta \right\} \right) & \text{if } B_2^2 \equiv 0, \\ O \left(\max \left\{ \|B_2^1\|_{\beta,2}^\beta, \|w\|_{\beta,2}^\beta \right\} \right) & \text{otherwise.} \end{cases}$$

Remark 3. Note that the estimates given in corollary 2 are sharp. In fact, suppose that $d = 1$ and let $w \in L^2([0, T]; \mathbb{R})$. Consider the process $z(t)$ defined by

$$z(t) := \int_0^t w(s) dW(s) \quad \text{for all } t \in [0, T].$$

We have that $\mathbb{E} \left(\sup_{t \in [0, T]} |z(t)|^\beta \right) \geq \mathbb{E}(|z(T)|^\beta) = \|w\|_2^\beta \mathbb{E}(|Z|^\beta)$, where Z is an standard normal random variable. Since, in this specific case, $\|w\|_{\beta,2}^\beta = \|w\|_2^\beta$, the conclusion follows.

Corollary 2 will be the main tool for establishing the following useful estimates:

Lemma 3. Consider y_1 defined by (8) and δy , d_1 defined in (9). For every $\beta \geq 1$ and $v \in L_{\mathcal{F}}^{2\beta,4}$, the following estimates hold:

$$\mathbb{E} \left(\sup_{t \in [0, T]} |\delta y|^\beta \right) = \begin{cases} O(\|v\|_{\beta,1}^\beta) & \text{if } \sigma_u \equiv 0, \\ O(\|v\|_{\beta,2}^\beta) & \text{otherwise.} \end{cases} \quad (12)$$

$$\mathbb{E} \left(\sup_{t \in [0, T]} |y_1|^\beta \right) = \begin{cases} O(\|v\|_{\beta,1}^\beta) & \text{if } \sigma_u \equiv 0, \\ O(\|v\|_{\beta,2}^\beta) & \text{otherwise.} \end{cases} \quad (13)$$

$$\mathbb{E} \left(\sup_{t \in [0, T]} |d_1|^\beta \right) = \begin{cases} O(\|v\|_{2\beta,2}^{2\beta}) & \text{if } \sigma_{uu} \equiv 0, \\ O(\|v\|_{2\beta,4}^{2\beta}) & \text{otherwise.} \end{cases} \quad (14)$$

Proof. For notational convenience we will suppose that $m = n = d = 1$. We have

$$\begin{aligned} d\delta y(t) &= [\tilde{f}_y(t)\delta y(t) + \tilde{f}_u(t)v(t)] dt + [\tilde{\sigma}_y(t)\delta y(t) + \tilde{\sigma}_u(t)v(t)] dW(t), \\ \delta y(0) &= 0. \end{aligned} \quad (15)$$

where, for $\psi = f, \sigma$,

$$\begin{aligned} \tilde{\psi}_y(t) &:= \int_0^1 \psi_y(\bar{y}(t) + \theta\delta y(t), \bar{u}(t) + \theta v(t)) d\theta, \\ \tilde{\psi}_u(t) &:= \int_0^1 \psi_u(\bar{y}(t) + \theta\delta y(t), \bar{u}(t) + \theta v(t)) d\theta. \end{aligned}$$

Using the second assumption in (1), estimates (12), (13) follow from corollary 2 applied to (15) and (8) respectively.

We next prove (14). We have that

$$\begin{aligned} dd_1(t) &= \left[\tilde{f}_y(t)\delta y(t) - f_y(t)y_1(t) + \left(\tilde{f}_u(t) - f_u(t) \right) v(t) \right] dt + \\ &\quad \left[\tilde{\sigma}_y(t)\delta y(t) - \sigma_y(t)y_1(t) + \left(\tilde{\sigma}_u(t) - \sigma_u(t) \right) v(t) \right] dW(t), \\ d_1(0) &= 0. \end{aligned}$$

For $\psi = f, \sigma$, we have that $[\tilde{\psi}_y(t) - \psi_y(t)] y_1(t) = O([\delta y(t) + |v(t)|] |y_1(t)|)$. Also,

$$[\tilde{\sigma}_u(t) - \sigma_u(t)] v(t) = \begin{cases} O(|\delta y(t)||v(t)|) & \text{if } \sigma_{uu} \equiv 0, \\ O([\delta y(t) + |v(t)|] |v(t)|) & \text{otherwise.} \end{cases}$$

Therefore, the following equation holds for d_1 :

$$\begin{aligned} dd_1(t) &= \left[\tilde{f}_y(t)d_1(t) + O([\delta y(t) + |v(t)|][|y_1(t)| + |v(t)|]) \right] dt + \\ &\quad \left[\tilde{\sigma}_y(t)d_1(t) + O(D(\delta y, y_1, v)) \right] dW(t), \end{aligned}$$

where

$$D(\delta y(t), y_1(t), v(t)) = \begin{cases} [|\delta y(t) + |v(t)|| |y_1(t)| + |v(t)|] - |v(t)|^2 & \text{if } \sigma_{uu} \equiv 0, \\ [|\delta y(t) + |v(t)|| |y_1(t)| + |v(t)|] & \text{otherwise.} \end{cases}$$

By (12) and (13),

$$\begin{aligned} \|\delta y\|_{\beta,2} \|\beta,2\| &= \mathbb{E} \left[\left(\int_0^T |\delta y(t)|^2 |y_1(t)|^2 dt \right)^{\frac{\beta}{2}} \right] \\ &= O \left[\mathbb{E} \left(\sup |\delta y(t)|^\beta |y_1(t)|^\beta \right) \right] \\ &= O \left(\left[\mathbb{E} \left(\sup |\delta y(t)|^{2\beta} \right) \right]^{\frac{1}{2}} \left[\mathbb{E} \left(\sup |y_1(t)|^{2\beta} \right) \right]^{\frac{1}{2}} \right) \\ &= O(\|v\|_{2\beta,2}^{2\beta}). \end{aligned} \tag{16}$$

Also, by the Cauchy Schwarz inequality and (12), (13),

$$\begin{aligned} \|y_1\|v\|_{\beta,2} &= \mathbb{E} \left[\left(\int_0^T |y_1(t)|^2 |v(t)|^2 dt \right)^{\frac{\beta}{2}} \right] = O(\|v\|_{2\beta,2}^{2\beta}), \\ \|\delta y\|v\|_{\beta,2} &= \mathbb{E} \left[\left(\int_0^T |\delta y(t)|^2 |v(t)|^2 dt \right)^{\frac{\beta}{2}} \right] = O(\|v\|_{2\beta,2}^{2\beta}), \end{aligned}$$

and (14) follows by corollary 2, since $\|v^2\|_{\beta,1}^\beta = \|v\|_{2\beta,2}^{2\beta}$ and $\|v^2\|_{\beta,2}^\beta = \|v\|_{2\beta,4}^{2\beta}$. \square

The estimates obtained in lemma 3 will provide a first order expansion of J around \bar{u} . This expansion will be expressed, as usual, in terms of an adjoint state. Let $(\bar{p}, \bar{q}) \in L^2_{\mathcal{F}}([0, T]; \mathbb{R}^n) \times (L^2_{\mathcal{F}}([0, T]; \mathbb{R}^n))^d$ be the unique solution of the following backward stochastic differential equation (BSDE) (see [1, 5])

$$\begin{aligned} dp(t) &= - \left[\ell_y(t)^\top + f_y(t)^\top p(t) + \sum_{i=1}^m \sigma_y^i(t)^\top q^i(t) \right] dt + q(t) dW(t), \\ p(T) &= \phi_y(\bar{y}(T))^\top. \end{aligned} \tag{17}$$

In the notation above σ^i and q^i denote respectively the i th column of σ and q . The following estimates hold (see [15, Proposition 3.1]):

Proposition 4. *Assume that (H1), (H2) hold and that $\bar{u} \in L_{\mathcal{F}}^{\beta,2}$. Then there exists $C_q > 0$ such that*

$$\mathbb{E} \left(\sup_{t \in [0, T]} |\bar{p}(t)|^\beta \right) + \sum_{i=1}^d \|\bar{q}^i\|_{\beta,2}^\beta \leq C_q \left(1 + \|\bar{u}\|_{\beta,2}^\beta \right).$$

Define the Hamiltonian $H : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \rightarrow \mathbb{R}$ by

$$H(t, y, u, p, q) := \ell(t, y, u) + p \cdot f(t, y, u) + \sum_{i=1}^d q^i \cdot \sigma^i(t, y, u), \quad (18)$$

and set $H_u(t) := H_u(t, \bar{y}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t))$. Define $\Upsilon_1 : L_{\mathcal{F}}^2 \rightarrow \mathbb{R}$ by

$$\Upsilon_1(v) := \mathbb{E} \left(\int_0^T H_u(t) v(t) dt \right). \quad (19)$$

In view of proposition 4, with $\beta = 2$, the function Υ_1 is well defined. The following lemma is a consequence of Itô's lemma for multidimensional Itô process (see [18]).

Lemma 5. *Let Z_1 and Z_2 be \mathbb{R}^n -valued continuous process satisfying*

$$\begin{cases} dZ_1(t) &= b_1(t)dt + \sigma_1(t)dW(t) \text{ for all } t \in [0, T], \\ dZ_2(t) &= b_2(t)dt + \sigma_2(t)dW(t) \text{ for all } t \in [0, T], \end{cases} \quad (20)$$

where b_1, b_2 are \mathbb{F} -adapted measurable process taking values in \mathbb{R}^n and σ_1, σ_2 are \mathbb{F} -adapted measurable process taking values in $\mathbb{R}^{n \times d}$. Also, let us suppose \mathbb{P} -a.s. we have that $Z_1(0) = 0$. Then

$$\mathbb{E}(Z_1(T) \cdot Z_2(T)) = \mathbb{E} \left(\int_0^T \left[Z_1(t) \cdot b_2(t) + Z_2(t) \cdot b_1(t) + \sum_{i=1}^d \sigma_1^i(t) \cdot \sigma_2^i(t) \right] dt \right).$$

Lemma 5 yields the following well known alternative expression for Υ_1 .

Lemma 6. *For every $v \in L_{\mathcal{F}}^2([0, T]; \mathbb{R}^m)$ we have that:*

$$\Upsilon_1(v) = \mathbb{E} \left(\int_0^T [\ell_y(t)y_1(t) + \ell_u(t)v(t)] dt + \phi_y(\bar{y}(T))y_1(T) \right). \quad (21)$$

Proof. Noting that

$$\phi_y(\bar{y}(T))y_1(T) = \bar{p}(T)^\top y_1(T) - \bar{p}(0)^\top y_1(0),$$

lemma 5, applied to $Z_1 = y_1$ and $Z_2 = \bar{p}$, yields $\mathbb{E}(\phi_y(\bar{y}(T))y_1(T)) = I_1 + I_2 + I_3$, where

$$\begin{aligned} I_1 &:= -\mathbb{E} \left(\int_0^T y_1(t)^\top [\ell_y(t)^\top + f_y(t)^\top \bar{p}(t) + \sum_{i=1}^d \int_0^T \sigma_y^i(t)^\top \bar{q}^i(t)] dt \right), \\ I_2 &:= \mathbb{E} \left(\int_0^T \bar{p}(t)^\top [f_y(t)y_1(t) + f_u(t)v(t)] dt \right), \\ I_3 &:= \sum_{i=1}^d \mathbb{E} \left(\int_0^T \bar{q}^i(t)^\top [\sigma_y^i(t)y_1(t) + \sigma_u^i(t)v(t)] dt \right). \end{aligned}$$

Plugging the expressions of I_1, I_2 and I_3 introduced above into the right hand side of (21) yields the result. \square

The expression above for Υ_1 allows to obtain a *first order expansion* of J around \bar{u} .

Proposition 7. *Assume that (H1), (H2) hold and let $v \in L^4_{\mathcal{F}}$. Then, $\Upsilon_1(v) = O(\|v\|_2)$ and the following expansion holds $J(\bar{u} + v) = J(\bar{u}) + \Upsilon_1(v) + r_1(v)$ with*

$$r_1(v) = \begin{cases} O(\|v\|_{4,2}^2) & \text{if } \sigma_{uu} \equiv 0, \\ O(\|v\|_4^2) & \text{otherwise.} \end{cases} \quad (22)$$

If in addition (6) holds, then

$$\Upsilon_1(v) = \begin{cases} O(\|v\|_1) & \text{if } \sigma_u \equiv 0, \\ O(\|v\|_{1,2}) & \text{otherwise,} \end{cases} \quad ; \quad r_1(v) = \begin{cases} O(\|v\|_{2,2}^2) & \text{if } \sigma_{uu} \equiv 0, \\ O(\|v\|_{2,4}^2) & \text{otherwise.} \end{cases} \quad (23)$$

Proof. Let us denote $\delta J := J(\bar{u} + v) - J(\bar{u})$. By definition

$$\delta J = \mathbb{E} \left(\int_0^T [\ell(y_{\bar{u}+v}, \bar{u} + v) - \ell(\bar{y}, \bar{u})] dt + \phi(y_{\bar{u}+v}(T)) - \phi(\bar{y}(T)) \right) = \Upsilon_1(v) + r_1(v),$$

where $r_1(v) = O(z_1(v) + z_2(v))$ and

$$z_1(v) := \mathbb{E} \left[\int_0^T |\ell_y(t)d_1(t)| dt + \phi_y(\bar{y}(T))d_1(T) \right], \quad z_2(v) := \mathbb{E} \left(\sup_{t \in [0, T]} |\delta y(t)|^2 \right) + \|v\|_2^2.$$

Now, we estimate $\Upsilon_1(v)$. By assumption (H2) and the Cauchy Schwartz inequality $\mathbb{E} \left(\int_0^T \ell_u(t)v(t) dt \right) = O(\|v\|_2)$. On the other hand, by (13)

$$\mathbb{E} \left(\int_0^T \ell_y(t)y_1(t) dt + \phi_y(\bar{y}(T))y_1(T) \right) = O \left(\left[\mathbb{E} \left(\sup_{t \in [0, T]} |y_1(t)|^2 \right) \right]^{\frac{1}{2}} \right) = O(\|v\|_2).$$

Thus $\Upsilon_1(v) = O(\|v\|_2)$. If (6) holds, then $\mathbb{E} \left(\int_0^T \ell_u(t)v(t) dt \right) = O(\|v\|_1)$, and

$$\mathbb{E} \left(\int_0^T \ell_y(t)y_1(t) dt + \phi_y(\bar{y}(T))y_1(T) \right) = O \left(\mathbb{E} \left[\sup_{t \in [0, T]} |y_1(t)| \right] \right).$$

Thus, estimates for $\Upsilon_1(v)$ in (23) follow from (13) with $\beta = 1$. Let us estimate $r_1(v)$. Assumption (H2) and (12) imply that $z_2(v) = O(\|v\|_2^2)$. On the other hand, by (H2) and the Cauchy Schwarz inequality

$$z_1(v) = O \left(\left[\mathbb{E} \left(\sup_{t \in [0, T]} |d_1(t)|^2 \right) \right]^{\frac{1}{2}} \right).$$

Thus (22) follows from estimates (14) with $\beta = 2$. If in addition (6) holds, then $z_1(v) = O \left(\mathbb{E} \left[\sup_{t \in [0, T]} |d_1(t)| \right] \right)$ and the estimates for $r_1(v)$ in (23) follows from (14) with $\beta = 1$. \square

Remark 4. The above proof shows that the hypothesis for the perturbation v can be weakened. For example, if (6) holds and $\sigma_{uu} = 0$, for all $v \in L_{\mathcal{F}}^2$ we have that $J(\bar{u} + v) = J(\bar{u}) + \Upsilon_1(v) + r_1(v)$ with $\Upsilon_1(v) = O(\|v\|_1)$ and $r_1(v) = O(\|v\|_2^2)$. Thus, in this case, the function J is differentiable at \bar{u} .

Corollary 8. Assume that **(H1)**, **(H2)** hold and let $v \in L_{\mathcal{F}}^\infty$. Then, $\Upsilon_1(v) = O(\|v\|_2)$ and $J(\bar{u} + v) = J(\bar{u}) + \Upsilon_1(v) + r_1(v)$ with $r_1(v) = O(\|v\|_\infty^2)$.

The second order linearization of $u \in L_{\mathcal{F}}^2 \mapsto y_u \in \mathcal{Y}^2$ around \bar{u} in the direction $v \in L_{\mathcal{F}}^\infty$ is defined as the unique solution $y_2 = y_2(v)$ of

$$\begin{aligned} dy_2(t) &= \left[f_y(t)y_2(t) + \frac{1}{2}f_{yy}(t)y_1(t)^2 + f_{yu}(t)y_1(t)v(t) + \frac{1}{2}f_{uu}(t)v(t)^2 \right] dt \\ &\quad + \left[\sigma_y(t)y_2(t) + \frac{1}{2}\sigma_{yy}(t)y_1(t)^2 + \sigma_{yu}(t)y_1(t)v(t) + \frac{1}{2}\sigma_{uu}(t)v(t)^2 \right] dW(t); \\ y_2(0) &= 0. \end{aligned} \tag{24}$$

Note that by the third assumption in (1) and proposition 1, we have that y_2 is well defined.

Lemma 9. Consider y_2 defined in (24) and $d_2 := \delta y - y_1 - y_2 = d_1 - y_2$. For every $\beta \geq 1$ and $v \in L_{\mathcal{F}}^\infty$, the following estimates hold:

$$\mathbb{E} \left(\sup_{t \in [0, T]} |y_2|^\beta \right) = \begin{cases} O(\|v\|_{2\beta, 2}^{2\beta}) & \text{if } \sigma_{uu} \equiv 0, \\ O(\|v\|_{2\beta, 4}^{2\beta}) & \text{otherwise.} \end{cases} \tag{25}$$

$$\mathbb{E} \left(\sup_{t \in [0, T]} |d_2|^\beta \right) = \begin{cases} O(\|v\|_{2\beta, 2}^\beta \|v\|_{4\beta, 4}^{2\beta}) & \text{if } \sigma_{uuu} \equiv 0, \\ O(\|v\|_{2\beta, 2}^\beta \|v\|_{4\beta, 4}^{2\beta} + \|v\|_{3\beta, 6}^{3\beta}) & \text{otherwise.} \end{cases} \tag{26}$$

Proof. As in the proof of lemma 3 we suppose that $m = n = d = 1$. We will use repeatedly that for every $\beta, p, q \in [1, \infty)$, we have

$$\| |v|^q \|_{\beta, p}^\beta = \|v\|_{q\beta, qp}^{q\beta} \quad \text{for all } v \in L_{\mathcal{F}}^{q\beta, qp}.$$

Proof of (25): Recall that, by **(H1)**, for $\psi = f, \sigma$ we assume that ψ_{yy}, ψ_{yu} and ψ_{uu} are bounded. Using (13),

$$\|y_1^2\|_{\beta, 2}^\beta = \mathbb{E} \left[\left(\int_0^T |y_1(t)|^4 dt \right)^{\frac{\beta}{2}} \right] = O \left[\mathbb{E} (\sup |y_1(t)|^{2\beta}) \right] = O \left(\|v\|_{2\beta, 2}^{2\beta} \right). \tag{27}$$

Analogously, the estimates associated with the term y_1v is of order $\|v\|_{2\beta, 2}^{2\beta}$. Estimate (25) follows from corollary 2 since $\|v^2\|_{\beta, 1}^\beta = \|v\|_{2\beta, 2}^{2\beta}$ and $\|v^2\|_{\beta, 2}^\beta = \|v\|_{2\beta, 4}^{2\beta}$.

Proof of (26): Recall that $d_2 = \delta y - y_1 - y_2$. We have, omitting time from the arguments,

$$dd_2(t) = \left[f_y d_2 + \frac{1}{2}f_{yy}([\delta y]^2 - y_1^2) + f_{yu}(\delta y - y_1)v + r_t(f)(\delta y, v)^2 \right] dt + \left[\sigma_y(t)d_2 + \frac{1}{2}\sigma_{yy}([\delta y]^2 - y_1^2) + \sigma_{yu}(\delta y - y_1)v + r_t(\sigma)(\delta y, v)^2 \right] dW(t).$$

where for $\psi = f, \sigma$ the map $r_t(\psi)$ is defined by

$$r_t(\psi) := \int_0^1 (1 - \theta) [\psi_{yy}(\bar{y}(t) + \theta\delta y(t), \bar{u}(t) + \theta v(t)) - \psi_{yy}(\bar{y}(t), \bar{u}(t))] d\theta.$$

Thus, since $[\delta y]^2 - y_1^2 = (\delta y + y_1)d_1$ and $D\psi$ is Lipschitz, we obtain

$$dd_2(t) = \begin{aligned} & [f_y d_2 + O(|d_1| \{|\delta y| + |y_1|\} + |d_1||v| + \alpha_t(f))] dt + \\ & [\sigma_y d_2 + O(|d_1| \{|\delta y| + |y_1|\} + |d_1||v| + \alpha_t(\sigma))] dW(t) \end{aligned} \quad (28)$$

where, for $\psi = f, \sigma$,

$$\alpha_t(\psi) := \begin{cases} |\delta y(t)|^3 + |v(t)|^3 & \text{if } \psi_{uuu} \neq 0, \\ |\delta y(t)|^3 + |\delta y(t)| |v(t)|^2 & \text{if } \psi_{uuu} \equiv 0. \end{cases}$$

Now, let us estimate the terms in the $dW(t)$ part of (28),

$$\begin{aligned} \| |d_1| |\delta y| \|_{\beta,2}^\beta &= \mathbb{E} \left[\left(\int_0^T |d_1(t)|^2 |\delta y(t)|^2 dt \right)^{\frac{\beta}{2}} \right] = O \left[\mathbb{E} (\sup |d_1(t)|^\beta |\delta y(t)|^\beta) \right] \\ &= O(\|v\|_{2\beta,2}^\beta \|v\|_{4\beta,4}^{2\beta}), \end{aligned}$$

by (12) and (14). Analogously, estimates for the terms $d_1 y_1$ and $d_1 v$ are of the same order. Let us estimate the terms appearing in $\alpha_\sigma(t)$. Using (12),

$$\| |\delta y|^3 \|_{\beta,2}^\beta = \mathbb{E} \left[\left(\int_0^T |\delta y(t)|^6 dt \right)^{\frac{\beta}{2}} \right] = O \left[\mathbb{E} (\sup |\delta y(t)|^{3\beta}) \right] = O(\|v\|_{3\beta,2}^{3\beta}). \quad (29)$$

By (12), we obtain

$$\begin{aligned} \| |\delta y| |v|^2 \|_{\beta,2}^\beta &= \mathbb{E} \left[\left(\int_0^T |\delta y(t)|^2 |v(t)|^4 dt \right)^{\frac{\beta}{2}} \right] \\ &= O \left(\mathbb{E} \left[\sup |\delta y(t)|^\beta \left(\int_0^T |v(t)|^4 dt \right)^{\frac{\beta}{2}} \right] \right) = O(\|v\|_{2\beta,2}^\beta \|v\|_{4\beta,4}^{2\beta}). \end{aligned}$$

Also, we have that $\|v^3\|_{\beta,1}^\beta = \|v\|_{3\beta,3}^{3\beta}$ and $\|v^3\|_{\beta,2}^\beta = \|v\|_{3\beta,6}^{3\beta}$. By the Cauchy Schwarz inequality,

$$\|v\|_{3\beta,3}^{3\beta} = \mathbb{E} \left[\left(\int_0^T |v(t)|^3 dt \right)^\beta \right] \leq \mathbb{E} \left[\left(\int_0^T |v(t)|^2 dt \right)^{\frac{\beta}{2}} \left(\int_0^T |v(t)|^4 dt \right)^{\frac{\beta}{2}} \right].$$

Using the Cauchy Schwarz inequality again, we get $\|v\|_{3\beta,3}^{3\beta} = O(\|v\|_{2\beta,2}^\beta \|v\|_{4\beta,4}^{2\beta})$. Therefore, estimate (26) follows from corollary 2. \square

Our aim now is to obtain a second order expansion of J around \bar{u} . Let us set $H_{(y,u)^2}(t) = H_{(y,u)^2}(t, \bar{y}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t))$ and define $\Upsilon_2 : L_{\mathcal{F}}^\infty \rightarrow \mathbb{R}$ by

$$\Upsilon_2(v) := \mathbb{E} \left(\int_0^T H_{(y,u)^2}(t)(v(t), y_1(t))^2 dt + \phi_{yy}(\bar{y}(T))(y_1(T))^2 \right).$$

As for Υ_1 a useful alternative expression for Υ_2 holds.

Lemma 10. *For every $v \in L_{\mathcal{F}}^\infty$ we have that:*

$$\begin{aligned} \frac{1}{2} \Upsilon_2(v) &= \mathbb{E} \left(\int_0^T [\ell_y(t) y_2(t) + \frac{1}{2} \ell_{(y,u)^2}(t)(y_1(t), v(t))^2] dt \right) \\ &+ \mathbb{E} \left[\phi_y(\bar{y}(T)) y_2(T) + \frac{1}{2} \phi_{yy}(\bar{y}(T))(y_1(T))^2 \right]. \end{aligned} \quad (30)$$

Proof. By definition of y_2 and \bar{p} , we have that

$$\phi_y(\bar{y}(T))y_2(T) = \bar{p}(T) \cdot y_2(T) - \bar{p}(0) \cdot y_2(0).$$

Lemma 5 yields $\mathbb{E}(\phi_y(\bar{y}(T))y_2(T)) = I'_1 + I'_2 + I'_3$, where

$$\begin{aligned} I'_1 &:= -\mathbb{E} \left(\int_0^T y_2(t)^\top \left[\ell_y(t)^\top + f_y(t)^\top \bar{p}(t) + \sum_{i=1}^d \int_0^T \sigma_y^i(t)^\top \bar{q}^i(t) \right] dt \right), \\ I'_2 &:= \mathbb{E} \left(\int_0^T \bar{p}(t)^\top \left[f_y(t) y_2(t) + \frac{1}{2} f_{(y,u)^2}(t)(y_1(t), v(t))^2 \right] dt \right), \\ I'_3 &:= \sum_{i=1}^d \mathbb{E} \left(\int_0^T \bar{q}^i(t)^\top \left[\sigma_y^i(t) y_2(t) + \frac{1}{2} \sigma_{(y,u)^2}^i(t)(y_1(t), v(t))^2 \right] dt \right). \end{aligned}$$

Plugging the expressions of I'_1, I'_2 and I'_3 introduced above into the right hand side of (30) yields the result. \square

Now we are able to obtain a *second order expansion* of J around \bar{u} .

Proposition 11. *Assume that **(H1)**, **(H2)** hold and let $v \in L_{\mathcal{F}}^\infty$. Then,*

$$J(\bar{u} + v) = J(\bar{u}) + \Upsilon_1(v) + \frac{1}{2} \Upsilon_2(v) + r_2(v), \quad (31)$$

and the following estimates hold:

$$\Upsilon_2(v) = \begin{cases} O(\|v\|_{4,2}^2) & \text{if } \sigma_{uu} \equiv 0, \\ O(\|v\|_4^2) & \text{otherwise,} \end{cases} \quad r_2(v) = \begin{cases} O(\|v\|_\infty \|v\|_{4,2}^2) & \text{if } \sigma_{uuu} \equiv 0, \\ O(\|v\|_\infty \|v\|_4^2) & \text{otherwise.} \end{cases} \quad (32)$$

If in addition (6) holds then

$$\Upsilon_2(v) = \begin{cases} O(\|v\|_{2,2}^2) & \text{if } \sigma_{uu} \equiv 0, \\ O(\|v\|_{2,4}^2) & \text{otherwise,} \end{cases} \quad r_2(v) = \begin{cases} O(\|v\|_\infty \|v\|_{2,2}^2) & \text{if } \sigma_{uuu} \equiv 0, \\ O(\|v\|_\infty \|v\|_{2,4}^2) & \text{otherwise.} \end{cases} \quad (33)$$

Proof. Let us first estimate $\Upsilon_2(v)$ by using its expression obtained in lemma 10 and lemmas 3 and 9. By (13) with $\beta = 2$,

$$\mathbb{E} \left(\sup_{t \in [0, T]} |y_1(t)|^2 + \int_0^T |v(t)|^2 dt \right) = O(\|v\|_2^2). \quad (34)$$

In view of assumption **(H2)** and (34) we obtain that

$$\mathbb{E} \left(\int_0^T \ell_{(y,u)^2}(t)(y_1(t), v(t))^2 dt + \phi_{yy}(\bar{y}(T))(y_1(T))^2 \right) = O(\|v\|_2^2). \quad (35)$$

On the other hand, assumption **(H2)** and the Cauchy Schwartz inequality yield

$$\mathbb{E} \left(\int_0^T \ell_y(t) y_2(t) dt + \phi_y(\bar{y}(T)) y_2(T) \right) = O \left(\left[\mathbb{E} \left(\sup_{t \in [0, T]} |y_2|^2 \right) \right]^{\frac{1}{2}} \right), \quad (36)$$

and the estimate for $\Upsilon_2(v)$ in (32) follows from (25). If (6) holds, then

$$\mathbb{E} \left(\int_0^T \ell_y(t) y_2(t) dt + \phi_{yy}(\bar{y}(T)) y_2(T) \right) = O \left(\mathbb{E} \left[\sup_{t \in [0, T]} |y_2| \right] \right), \quad (37)$$

and the estimate for $\Upsilon_2(v)$ in (33) follows from (25).

Now we proceed to obtain (31). As in the proof of proposition 7 we denote $\delta J := J(\bar{u} + v) - J(\bar{u})$. By definition,

$$\delta J = \mathbb{E} \left(\int_0^T [\ell(y_{\bar{u}+v}, \bar{u} + v) - \ell(\bar{y}, \bar{u})] dt + \phi(y_{\bar{u}+v}(T)) - \phi(\bar{y}(T)) \right) = I_1 + I_2,$$

where, omitting the time argument in the integral,

$$\begin{aligned} I_1 &:= \mathbb{E} \left(\int_0^T [\ell_y \delta y + \ell_u v + \frac{1}{2} \ell_{(y,u)^2} (\delta y, v)^2 + r_\ell (\delta y, v)^2] dt \right), \\ I_2 &:= \mathbb{E} \left[\phi_y(\bar{y}(T)) \delta y(T) + \frac{1}{2} \phi_{yy}(\bar{y}(T)) (\delta y(T))^2 + r_\phi(\bar{y}(T)) (\delta y(T))^2 \right]. \end{aligned} \quad (38)$$

Recalling that $\delta y = y_1 + d_1 = y_1 + y_2 + d_2$, assumption (5) in **(H2)** yields

$$\begin{aligned} I_1 &= \mathbb{E} \left(\int_0^T \ell_y(t) (y_1 + y_2) + \ell_u(t) v + \frac{1}{2} D^2 \ell(t) (y_1, v)^2 dt \right) + \mathbb{E} \left(\int_0^T \ell_y d_2 dt \right) \\ &\quad + O(z_1(v)), \end{aligned}$$

where, omitting time from function arguments,

$$z_1(v) := \mathbb{E} \left(\sup [|d_1|^2 + |d_1(t)| |y_1| + |\delta y|^3] \right) + \|v\|_1 \mathbb{E} \left(\sup |d_1| \right) + \|v\|_3^3.$$

On the other hand,

$$\begin{aligned} I_2 &= \mathbb{E} \left[\phi_y(\bar{y}(T)) (y_1(T) + y_2(T)) + \frac{1}{2} \phi_{yy}(\bar{y}(T)) (y_1(T))^2 \right] \\ &\quad + \mathbb{E} \left[\phi_y(\bar{y}(T)) d_2(T) \right] + O(z_2(v)), \end{aligned}$$

where

$$z_2(v) := \mathbb{E} \left(|\delta y(T)|^3 + |y_1(T)| |d_1(T)| + |d_1(T)|^2 \right).$$

Denoting $z(v) := z_1(v) + z_2(v)$ we get that

$$\begin{aligned} \delta J &= \mathbb{E} \left(\int_0^T [\ell_y(t) (y_1(t) + y_2(t)) + \ell_u(t) v(t) + \frac{1}{2} \ell_{(y,u)^2}(t) (y_1(t), v(t))^2] dt \right) \\ &\quad + \mathbb{E} \left[\phi_y(\bar{y}(T)) (y_1(T) + y_2(T)) + \frac{1}{2} \phi_{yy}(\bar{y}(T)) (y_1(T))^2 \right] + \zeta(v) + z(v), \end{aligned}$$

where,

$$\zeta(v) := \mathbb{E} \left(\int_0^T \ell_y(t) d_2(t) dt + \phi_y(\bar{y}(T)) d_2(T) \right). \quad (39)$$

Therefore, using (21) and (30), we get (31) with $r_2(v) := \zeta(v) + z(v)$. Now, we proceed to estimate $z(v)$. By (14) we have that

$$\mathbb{E} \left(\sup_{t \in [0, T]} |d_1(t)|^2 \right) = O(\|v\|_4^4) = O(\|v\|_\infty^2 \|v\|_2^2).$$

Estimates (13), (14) and the Cauchy Schwartz inequality yield

$$\mathbb{E} \left(\sup_{t \in [0, T]} |d_1(t)| |y_1(t)| \right) = O(\|v\|_4^2 \|v\|_2) = O(\|v\|_\infty \|v\|_2^2).$$

Analogously, using (14), we have

$$\mathbb{E} \left(\|v\|_1 \sup_{t \in [0, T]} |d_1(t)| \right) = O(\|v\|_4^2 \|v\|_{2,1}) = O(\|v\|_\infty \|v\|_2^2).$$

Estimate (12) yields $\mathbb{E} \left(\sup_{t \in [0, T]} |\delta y(t)|^3 \right) = O(\|v\|_{3,2}^3)$. But

$$\|v\|_{3,2}^3 = \mathbb{E} \left(\left[\int_0^T |v(t)|^2 dt \right]^{\frac{3}{2}} \right) = O(\|v\|_\infty \|v\|_2^2),$$

and $\|v\|_3^3 = O(\|v\|_\infty \|v\|_2^2)$. Thus, $z(v) = O(\|v\|_\infty \|v\|_2^2)$. Finally, let us estimate $\zeta(v)$. Assumption **(H2)** and the Cauchy Schwartz inequality yield that

$$\zeta(v) = O \left(\left[\mathbb{E} \left(\sup_{t \in [0, T]} |d_2(t)|^2 \right) \right]^{\frac{1}{2}} \right).$$

Hence, using (26) with $\beta = 2$,

$$\zeta(v) = \begin{cases} O(\|v\|_{4,2} \|v\|_{8,4}^2) & \text{if } \sigma_{uuu} \equiv 0, \\ O(\|v\|_{4,2} \|v\|_{8,4}^2 + \|v\|_{6,6}^3) & \text{otherwise.} \end{cases}$$

Since $O(\|v\|_{4,2} \|v\|_{8,4}^2) = O(\|v\|_\infty \|v\|_{4,2}^2)$ and $O(\|v\|_{6,6}^3) = O(\|v\|_\infty \|v\|_4^2)$, the estimate for $r_2(v)$ in (32) follows. If in addition assumption (6) holds, then by (26) with $\beta = 1$,

$$\zeta(v) = O \left(\mathbb{E} \left[\sup_{t \in [0, T]} |d_2(t)| \right] \right) = \begin{cases} O(\|v\|_2 \|v\|_4^2) & \text{if } \sigma_{uuu} \equiv 0, \\ O(\|v\|_2 \|v\|_4^2 + \|v\|_{3,6}^3) & \text{otherwise.} \end{cases}$$

Since $O(\|v\|_2 \|v\|_4^2) = O(\|v\|_\infty \|v\|_2^2)$ and $O(\|v\|_{3,6}^3) = O(\|v\|_{2,4}^2)$, the estimate for $r_2(v)$ in (33) follows. \square

Remark 5. (i) Since Υ_2 is a quadratic form and, for every $\beta, p \in [1, \infty]$, the space $L_{\mathcal{F}}^\infty$ is dense in $L_{\mathcal{F}}^{\beta, p}$, we have that: If $\Upsilon_2(v) = O(\|v\|_{\beta, p})$ then Υ_2 admits a unique continuous extension in $L^{\beta, p}$.

(ii) The proof of proposition 31 shows that the estimates $\Upsilon_2(v) = O(\|v\|_2^2)$ and $r_2(v) = O(\|v\|_\infty \|v\|_2^2)$ also hold in the case when f and σ are affine mappings, since in this case $y_2 = d_2 = 0$.

The following corollary will allow us to state second order necessary condition with respect to perturbations $v \in L_{\mathcal{F}}^\infty$.

Corollary 12. Assume that **(H1)**, **(H2)** hold and either (6) holds and $\sigma_{uu} \equiv 0$, or f and σ are affine mappings. Then, the following expansion holds:

$$J(\bar{u} + v) = J(\bar{u}) + \Upsilon_1(v) + \frac{1}{2} \Upsilon_2(v) + r(v) \quad \text{for all } v \in L_{\mathcal{F}}^\infty, \quad (40)$$

where $\Upsilon_1(v) = O(\|v\|_2)$, $\Upsilon_2(v) = O(\|v\|_2^2)$ and $r(v) = O(\|v\|_\infty \|v\|_2^2)$.

4 Necessary optimality conditions

The asymptotic expansions obtained for J in section 3 allow us to obtain first and second order necessary conditions at a local optimum $\bar{u} \in L_{\mathcal{F}}^2$ for the control constrained problem (\mathcal{SP}) . We first obtain first order optimality conditions using the procedure explained at the introduction: According to the regularity of the data of (\mathcal{SP}) and the dependence on u of the σ -term, a perturbation in an appropriate space is taken. Then, the results of the previous section yield a positivity condition of Υ_1 over a certain cone which is extended, by a density argument, to a larger one. Similar considerations apply in order to establish second order necessary conditions. Finally, we give a second order sufficient condition for the unconstrained case and we briefly discuss the difficulties arising in the constrained case.

Let us first fix some notations which are standard in optimization theory. Consider a Banach space $(X, \|\cdot\|_X)$ and a nonempty closed convex set $C \subseteq X$. For $x, x' \in X$ define the *segment* $[x, x'] := \{x + \lambda(x' - x) ; \lambda \in [0, 1]\}$. The radial, the tangent and the normal cone to C at \bar{x} are defined respectively by

$$\begin{aligned} \mathcal{R}_C(\bar{x}) &:= \{h \in X ; \exists \sigma > 0 \text{ such that } [\bar{x}, \bar{x} + \sigma h] \subseteq C\}, \\ \mathcal{T}_C(\bar{x}) &:= \{h \in X ; \exists x(\sigma) = \bar{x} + \sigma h + o(\sigma) \in C, \sigma > 0, \|o(\sigma)/\sigma\|_X \rightarrow 0\}, \\ N_{\mathcal{U}}(\bar{u}) &:= \{h^* \in X^* / \langle x^*, x \rangle_{X^*, X} \leq 0, \text{ for all } h \in \mathcal{T}_C(\bar{x})\}, \end{aligned} \quad (41)$$

where X^* denotes the dual space of X and $\langle \cdot, \cdot \rangle_{X^*, X}$ is the duality product. Recall that, since C is a closed convex set, the cone $\mathcal{T}_C(\bar{x})$ is the adherence of $\mathcal{R}_C(\bar{x})$ in X .

4.1 First order necessary conditions

Consider as in section 3 a fixed $\bar{u} \in L_{\mathcal{F}}^2$. For $\beta, p \in [1, \infty]$ and a subset $A \subseteq L_{\mathcal{F}}^{\beta, p}$ we write $\text{adh}_{\beta, p}(A)$ for the adherence of A in $L_{\mathcal{F}}^{\beta, p}$. If $A \subseteq L_{\mathcal{F}}^{\beta}$ we write $\text{adh}_{\beta}(A) := \text{adh}_{\beta, \beta}(A)$.

We have the following first order conditions for (\mathcal{SP}) .

Lemma 13. *Assume that (H1), (H2) hold and let $\bar{u} \in \mathcal{U}$ be a local solution of (\mathcal{SP}) . Then:*

$$\Upsilon_1(v) \geq 0 \text{ for all } v \in \begin{cases} \text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{4,2}) & \text{if } \sigma_{uu} \equiv 0, \\ \text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^4) & \text{otherwise.} \end{cases} \quad (42)$$

If in addition (6) holds then

$$\Upsilon_1(v) \geq 0 \text{ for all } v \in \begin{cases} \text{adh}_1(\mathcal{R}_{\mathcal{U}}(\bar{u})) & \text{if } \sigma_{uu} \equiv 0, \\ \text{adh}_{1,2}(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{2,4}) & \text{otherwise.} \end{cases} \quad (43)$$

Proof. Let $v \in \mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^4$. Proposition 7 implies that, for $\sigma > 0$ small enough, we have

$$0 \leq J(\bar{u} + \sigma v) - J(\bar{u}) = \sigma \Upsilon_1(v) + \|v\|_4^2 O(\sigma^2). \quad (44)$$

Thus, dividing by σ in (44) and letting $\sigma \downarrow 0$, we have that $\Upsilon_1(v) \geq 0$. Analogously, if $\sigma_{uu} = 0$ we have that $\Upsilon_1(v) \geq 0$ for all $v \in \mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{4,2}$. Condition

(42) follows from the fact that, by proposition 7, $v \in L_{\mathcal{F}}^4 \rightarrow \Upsilon_1(v)$ can be extended continuously to $L_{\mathcal{F}}^2$. The proof of (43) follows in the same manner, with the obvious modifications. \square

Note that the results obtained in lemma 13 are rather general, since they include the case of non local constraints. On the other hand, for some constraints the result gives no information. In fact, consider the following example.

Example 1. Let $u_0 \in L_{\mathcal{F}}^2$ and suppose that $u_0 \notin L_{\mathcal{F}}^{\beta,p}$ for any $\beta, p \in (2, \infty]$. The constraint $\mathcal{U} := \{u = \alpha u_0 / \text{for some } \alpha \in [0, 1]\}$ is such that, at $\bar{u} = 0$, the radial cone is given by $\mathcal{R}_{\mathcal{U}}(\bar{u}) = \{\lambda u_0 / \text{for } \lambda \geq 0\}$, but $\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{\beta,p} = \{0\}$.

Thus, we will assume the following assumption over the constraint set \mathcal{U} :

(H3) For every $\bar{u} \in \mathcal{U}$ we have that

$$\mathcal{T}_{\mathcal{U}}(\bar{u}) = \text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{\infty}). \quad (45)$$

We have the following proposition whose proof is straightforward.

Proposition 14. *Assume that (H1), (H2), (H3) hold and let \bar{u} be a local solution of (SP). Then*

$$\Upsilon_1(v) \geq 0 \quad \text{for all } v \in \mathcal{T}_{\mathcal{U}}(\bar{u}). \quad (46)$$

Remark 6. *Note that if $J(\cdot)$ is convex, then (46) is a sufficient condition for the (global) optimality of \bar{u} .*

Clearly, we have that **(H3)** can hold for non local constraints. As an example, it can be checked that (45) holds for and $\mathcal{U} = \{u \in L_{\mathcal{F}}^2 / \|u\|_2 \leq 1\}$ and $\bar{u} \in \mathcal{U}$. Now we consider the case when \mathcal{U} is defined by *local constraints*. Let $(t, \omega) \in [0, T] \times \Omega \rightarrow U(t, \omega) \in \mathcal{P}(\mathbb{R}^m)$ be a $B([0, T]) \times \mathcal{F}_T$ measurable multifunction satisfying that

- (i) For all a.a. t the multifunction $U(t, \cdot)$ is \mathcal{F}_t -measurable.
- (ii) For a.a. (t, ω) we have that $U(t, \omega)$ is a closed convex subset of \mathbb{R}^m .

We set

$$\mathcal{U} := \{u \in L_{\mathcal{F}}^2 ; u(t, \omega) \in U(t, \omega), \text{ a.a. } (t, \omega) \in [0, T] \times \Omega\}. \quad (47)$$

Lemma 15. *Suppose that $\bar{u} \in \mathcal{U}$, where \mathcal{U} is given by (47). Then,*

- (i) *Assumption (45) holds at \bar{u} .*
- (ii) *The tangent cone is given by*

$$\mathcal{T}_{\mathcal{U}}(\bar{u}) = \{v \in L_{\mathcal{F}}^2 ; v(t, \omega) \in \mathcal{T}_{U(t, \omega)}(\bar{u}(t, \omega)) \text{ a.a. } (t, \omega) \in [0, T] \times \Omega\}. \quad (48)$$

Proof. (i) By a diagonal argument, it suffices to prove that for every $v \in \mathcal{R}_{\mathcal{U}}(\bar{u})$ there exists a sequence $v_k \in \mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{\infty}$ such that $\|v_k - v\|_2 \rightarrow 0$. Indeed, set

$$v_k(t, \omega) := \begin{cases} v(t, \omega) & \text{if } |v(t, \omega)| \leq k \\ 0 & \text{otherwise.} \end{cases} \quad (49)$$

The convexity of $U(t, \omega)$ yields that $v_k \in \mathcal{R}_{\mathcal{U}}(\bar{u})$. Also, $v_k(t, \omega) \rightarrow v(t, \omega)$ as $k \rightarrow \infty$ for a.a. (t, ω) . The convergence in $L_{\mathcal{F}}^2$ follows by the dominated

convergence theorem.

(ii) Let $v \in \mathcal{T}_{\mathcal{U}}(\bar{u})$. By definition, for σ small enough and a.a. (t, ω)

$$\bar{u}(t, \omega) + \sigma v(t, \omega) + r_{\sigma}(t, \omega) \in U(t, \omega),$$

where $r_{\sigma}(\cdot, \cdot)/\sigma \rightarrow 0$ in $L^2_{\mathcal{F}}$ as $\sigma \downarrow 0$. Thus, extracting a subsequence if necessary, we have that $r_{\sigma}(t, \omega)/\sigma \rightarrow 0$ for a.a. (t, ω) from which we deduce that $v(t, \omega) \in \mathcal{T}_{U(t, \omega)}(\bar{u}(t, \omega))$. Conversely, let v belongs to the r.h.s. of (48) and for $\varepsilon > 0$ set

$$v_{\varepsilon} := \varepsilon^{-1} (P_{\mathcal{U}}(\bar{u} + \varepsilon v) - \bar{u}), \quad (50)$$

where $P_{\mathcal{U}}(\cdot)$ denotes the orthogonal projection in $L^2_{\mathcal{F}}$ onto \mathcal{U} . By definition of $P_{\mathcal{U}}(\cdot)$ we have that $v_{\varepsilon} \in \mathcal{R}_{\mathcal{U}}(\bar{u})$. For (t, ω) in $[0, T] \times \Omega$ set $P_{U(t, \omega)}(\cdot)$ for the orthogonal projection in \mathbb{R}^m onto $U(t, \omega)$. Definition of v_{ε} in (50) implies that for a.a. (t, ω)

$$v_{\varepsilon}(t, \omega) := \varepsilon^{-1} (P_{U(t, \omega)}(\bar{u}(t, \omega) + \varepsilon v(t, \omega)) - \bar{u}(t, \omega)).$$

Clearly, $v_{\varepsilon}(t, \omega) \in \mathcal{R}_{U(t, \omega)}(\bar{u}(t, \omega))$ and for a.a. (t, ω) we have $v_{\varepsilon}(t, \omega) \rightarrow v(t, \omega)$. Since $|v_{\varepsilon}(t, \omega)| \leq |v(t, \omega)|$, the dominated convergence theorem implies that $v_{\varepsilon} \rightarrow v$ in $L^2_{\mathcal{F}}$. Using that $v_{\varepsilon} \in \mathcal{R}_{\mathcal{U}}(\bar{u})$ we obtain that $v \in \mathcal{T}_{\mathcal{U}}(\bar{u})$. \square

Let $a, b \in \bar{\mathbb{R}}^m$ with $-\infty \leq a^i < b^i \leq +\infty$ for all $i \in \{1, \dots, m\}$ and define

$$U_{a,b} := \{x \in \mathbb{R}^m ; a^i \leq x^i \leq b^i\}. \quad (51)$$

For $u \in L^2_{\mathcal{F}}$ and every index $i \in \{1, \dots, m\}$, set

$$\begin{aligned} I_{a^i}(u) &:= \{(t, \omega) \in [0, T] \times \Omega ; u^i(t, \omega) = a^i\}, \\ I_{b^i}(u) &:= \{(t, \omega) \in [0, T] \times \Omega ; u^i(t, \omega) = b^i\}. \end{aligned}$$

The following corollary is a direct consequence of proposition 14 and lemma 15.

Corollary 16. *Assume that (H1), (H2) hold suppose that \mathcal{U} is in the form (47). Let $\bar{u} \in \mathcal{U}$ be a local solution of (SP), then*

$$H_u(t, \omega)v \geq 0 \quad \text{for all } v \in \mathcal{T}_{U(t, \omega)}(\bar{u}(t, \omega)). \quad (52)$$

In particular, if $U(t, \omega) \equiv U_{a,b}$ (defined in (51)), then for every $i \in \{1, \dots, m\}$

$$H_u^i(t, \omega) = \begin{cases} \geq 0 & \text{if } (t, \omega) \in I_{a^i}(\bar{u}), \\ \leq 0 & \text{if } (t, \omega) \in I_{b^i}(\bar{u}), \\ = 0 & \text{elsewhere.} \end{cases} \quad (53)$$

Remark 7. *Since (52) is equivalent to (46) when \mathcal{U} is in the form (47), we have that if $J(\cdot)$ is convex then (52) is a sufficient condition for the (global) optimality of \bar{u} .*

4.2 Second order necessary conditions

In order to obtain second order necessary conditions for (\mathcal{SP}) we proceed as in the previous section, i.e. we prove a general result and after, under some standard assumptions, we yield a more precise characterization for the important case of local constraints. Let us define

$$\Upsilon_1^\perp := \{v \in L_{\mathcal{F}}^2 ; \Upsilon_1(v) = 0\}. \quad (54)$$

We have the following general second order necessary conditions.

Proposition 17. *Assume that **(H1)**, **(H2)** hold and let $\bar{u} \in \mathcal{U}$ be a local solution of (\mathcal{SP}) . Then, the following second order necessary condition holds:*

$$\Upsilon_2(v) \geq 0 \text{ for all } v \in \begin{cases} \text{adh}_{4,2}(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap \Upsilon_1^\perp) & \text{if } \sigma_{uu} \equiv 0, \\ \text{adh}_4(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap \Upsilon_1^\perp) & \text{otherwise.} \end{cases} \quad (55)$$

If in addition (6) holds, or f and σ are affine mappings, then

$$\Upsilon_2(v) \geq 0 \text{ for all } v \in \begin{cases} \text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap \Upsilon_1^\perp) & \text{if } \sigma_{uu} \equiv 0, \\ \text{adh}_{2,4}(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap \Upsilon_1^\perp) & \text{otherwise.} \end{cases} \quad (56)$$

Proof. If $v \in \mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap \Upsilon_1^\perp$, then for σ small enough

$$0 \leq J(\bar{u} + \sigma v) - J(\bar{u}) = \frac{\sigma^2}{2} \Upsilon_2(v) + \sigma^3 O(\|v\|_\infty^3).$$

Dividing the above equation by σ and letting $\sigma \downarrow 0$ yields $\Upsilon_2(v) \geq 0$ and the result follows from remark 5 (i). \square

The critical cone to \mathcal{U} at \bar{u} are defined by

$$C(\bar{u}) := \{v^* \in \mathcal{T}_{\mathcal{U}}(\bar{u}) / \Upsilon_1(v) \leq 0\}. \quad (57)$$

In order to obtain more precise second order necessary conditions, we suppose standard assumptions in the second order analysis of problems with convex constraints. The first one is a natural extension of **(H3)** to the second order case.

(H4) For every $\bar{u} \in \mathcal{U}$ and $v^* \in N_{\mathcal{U}}(\bar{u})$ (recall (41)), we have that

$$\text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap (v^*)^\perp) = \text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap (v^*)^\perp). \quad (58)$$

For our second assumption, we need the following notion of polyhedricity (see [9, 14]). The set \mathcal{U} is said to be *polyhedric* at $\bar{u} \in \mathcal{U}$ if for all $v^* \in N_{\mathcal{U}}(\bar{u})$, the set $\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap (v^*)^\perp$ is dense in $\mathcal{T}_{\mathcal{U}}(\bar{u}) \cap (v^*)^\perp$ with respect to the $\|\cdot\|_2$ norm. If \mathcal{U} is polyhedric at each $u \in \mathcal{U}$ we say that \mathcal{U} is *polyhedric*.

Remark 8. *Note that, if **(H1)**- **(H3)** hold, proposition 14 yields that, at a local minimum, $-\Upsilon_1 \in N_{\mathcal{U}}(\bar{u})$ and $C(\bar{u}) = \mathcal{T}_{\mathcal{U}}(\bar{u}) \cap \Upsilon_1^\perp$. Thus, if \mathcal{U} is polyhedric and **(H4)** holds,*

$$\text{adh}_2(\mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap \Upsilon_1^\perp) = C(\bar{u}) \quad (59)$$

We state a second order necessary condition which is a natural extension of the deterministic counterpart.

Theorem 18. *Let \bar{u} be a local solution of (SP) and assume that*

- (i) *Assumptions (H1)-(H4) hold.*
- (ii) *Either (6) holds and $\sigma_{uu} = 0$ or f and σ are affine mappings.*
- (iii) *The constraint set \mathcal{U} is polyhedric.*

Then, the following second order necessary condition hold at \bar{u} :

$$\Upsilon_2(v) \geq 0 \quad \text{for all } v \in C(\bar{u}). \quad (60)$$

Proof. As in the proof of proposition 17 we have that $\Upsilon_2(v) \geq 0$ for all $v \in \mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^{\infty} \cap \Upsilon_1^{\perp}$. The result follows, by remark 5 (i), since under our assumptions $\Upsilon_2(v) = O(\|v\|_2^2)$ and (59) holds. \square

Now, let us focus our attention in local constraints, i.e. when \mathcal{U} is defined by (47).

Lemma 19. *Let \mathcal{U} be defined by (47) and let $\bar{u} \in \mathcal{U}$. It holds that*

- (i) *The normal cone $N_{\mathcal{U}}(\bar{u})$ is given by*

$$N_{\mathcal{U}}(\bar{u}) = \{v^* \in L_{\mathcal{F}}^2 / v^*(t, \omega) \in N_{U(t, \omega)}(\bar{u}(t, \omega)), \quad \text{a.a. } (t, \omega) \in [0, T] \times \Omega\}. \quad (61)$$

- (ii) *For every $v^* \in N_{\mathcal{U}}(\bar{u})$ we have that*

$$\mathcal{T}_{\mathcal{U}}(\bar{u}) \cap (v^*)^{\perp} = \{v \in \mathcal{T}_{\mathcal{U}}(\bar{u}) / v^*(t, \omega) \cdot v(t, \omega) = 0, \quad \text{a.a. } (t, \omega) \in [0, T] \times \Omega\}. \quad (62)$$

Proof. Since (ii) follows directly from (i) and lemma 15 (ii), it is enough to show (i). By lemma 15 (ii), the r.h.s. of (61) is included in $N_{\mathcal{U}}(\bar{u})$. To prove the other inclusion, let us argue by contradiction. Let $v^* \in N_{\mathcal{U}}(\bar{u})$ and suppose that it does not belong to the r.h.s. of (61). Then we can find a non null measurable set $\mathcal{A} \subseteq \mathcal{B}([0, T]) \otimes \mathcal{F}$ such that for each $(t, \omega) \in \mathcal{A}$ there exists $v(t, \omega) \in \mathcal{T}_{U(t, \omega)}(\bar{u}(t, \omega))$, which can be taken with $|v(t, \omega)| = 1$, such that $v^*(t, \omega) \cdot v(t, \omega) > \alpha$, for some $\alpha > 0$. Defining $\hat{v} \in L_{\mathcal{F}}^2$ by

$$\hat{v}(t, \omega) := \begin{cases} v(t, \omega) & \text{if } v^*(t, \omega) \cdot v(t, \omega) > \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

we see that $\hat{v} \in \mathcal{T}_{\mathcal{U}}(\bar{u})$ and $\langle v^*, \hat{v} \rangle_2 > 0$ and thus we obtain a contradiction with the fact that $v^* \in N_{\mathcal{U}}(\bar{u})$. \square

In order to verify the polyhedricity assumption in the case of local constraints, we will need in fact to assume that for a.a. (t, ω) the set $U(t, \omega)$ is a polyhedron. More precisely, let $q \in \mathbb{N}$ and suppose that there exist mappings $\Sigma : [0, T] \times \Omega \rightarrow \mathcal{P}(\{1, \dots, q\})$, $a_i : [0, T] \times \Omega \rightarrow \mathbb{R}^m$, $b_i : [0, T] \times \Omega \rightarrow \mathbb{R}^m$, where $i \in \{1, \dots, q\}$, such that Σ , a_i and b_i are $\mathcal{B}([0, T]) \times \mathcal{F}_T$ measurable and for each t we have that $\Sigma(t, \cdot)$, $a_i(t, \cdot)$ and $b_i(t, \cdot)$ are \mathcal{F}_t measurable. We suppose that

$$U(t, \omega) = \{x \in \mathbb{R}^m / \langle a_i(t, \omega), x \rangle \leq b_i(t, \omega), \text{ for } i \in \Sigma(t, \omega)\}. \quad (63)$$

We have

Lemma 20. *The set of local constraints \mathcal{U} defined in (47), with $U(t, \omega)$ given by (63), is polyhedric and satisfies (58).*

Proof. Let $\bar{u} \in \mathcal{U}$ and $v^* \in N_{\mathcal{U}}(\bar{u})$. For $v \in \mathcal{T}_{\mathcal{U}}(\bar{u}) \cap (v^*)^\perp$ and $k \geq 0$ set

$$\widehat{v}_k(t, \omega) := \begin{cases} v(t, \omega) & \text{if } |v(t, \omega)| \leq k \text{ and } \bar{u}(t, \omega) + \frac{1}{k}v(t, \omega) \in U(t, \omega), \\ 0 & \text{otherwise.} \end{cases} \quad (64)$$

Lemma 19(ii) implies that $\widehat{v}_k \in \mathcal{R}_{\mathcal{U}}(\bar{u}) \cap L_{\mathcal{F}}^\infty \cap (v^*)^\perp$. On the other hand, since $U(t, \omega)$ is a polyhedron, lemma 15(ii) implies that $v(t, \omega) \in \mathcal{T}_{\mathcal{U}}(\bar{u}(t, \omega)) = \mathcal{R}_{\mathcal{U}}(\bar{u}(t, \omega))$. Thus, as $k \uparrow \infty$, we have that $\widehat{v}_k \rightarrow v(t, \omega)$ for a.a. (t, ω) . The dominated convergence theorem, yields that $\widehat{v}_k \rightarrow v$ in $L_{\mathcal{F}}^2$, hence \mathcal{U} is polyhedric and (58) holds. \square

The following corollary is a direct consequence of theorem 18 and lemmas 19, 20.

Corollary 21. *Assume that (H1) - (H2) hold and let \bar{u} be a local solution of (SP) where \mathcal{U} is defined in (47), with $U(t, \omega)$ given by (63). Further, suppose that either (6) holds and $\sigma_{uu} = 0$ or f and σ are affine mappings. Then, the following second order necessary conditions hold at \bar{u} :*

$\Upsilon_2(v) \geq 0$, for all $v \in \mathcal{T}_{\mathcal{U}}(\bar{u})$ such that $H_u(t)v(t, \omega) = 0$ for a.a. $(t, \omega) \in [0, T] \times \Omega$.

5 On the second order sufficient condition

Let us first consider the unconstrained case, i.e. when $\mathcal{U} = L_{\mathcal{F}}^2$. Note that, in this specific case, (H3) is trivially satisfied and for every $\bar{u} \in \mathcal{U}$ it holds that $\mathcal{T}_{\mathcal{U}}(\bar{u}) = L_{\mathcal{F}}^2$. The following proposition is a consequence of corollary 12.

Proposition 22. *Assume that (H1), (H2) hold and that $\mathcal{U} = L_{\mathcal{F}}^2$. Further, let us assume that either (6) holds and $\sigma_{uu} \equiv 0$, or f and σ are affine mappings. Suppose there exist $\alpha > 0$ such that $\bar{u} \in L_{\mathcal{F}}^2$ satisfies:*

$$\Upsilon_1(v) = 0, \text{ and } \Upsilon_2(v) \geq \alpha \|v\|_2^2 \text{ for all } v \in L_{\mathcal{F}}^2. \quad (65)$$

Then, there exists $\delta > 0$ such that for all $v' \in L_{\mathcal{F}}^\infty$ with $\|v'\|_\infty \leq \delta$, we have

$$J(\bar{u} + v') \geq J(\bar{u}) + \frac{1}{2}\alpha \|v'\|_2^2. \quad (66)$$

Only very partial results are obtained when $\mathcal{U} \neq L_{\mathcal{F}}^2$. Let us recall that a quadratic form $Q : H \rightarrow \mathbb{R}$, where H is a Hilbert space, is a Legendre form if it is weakly lower semi continuous (w.l.s.c.) quadratic form over H , such that, if $h_k \rightarrow h$ weakly in H and $Q(h_k) \rightarrow Q(h)$, then $h_k \rightarrow h$ strongly. We have the following proposition, whose proof follows the lines of the parallel deterministic result (see [7, Section 3.3]):

Proposition 23. *Assume that (H1), (H2) hold and that either (6) holds and $\sigma_{uu} \equiv 0$, or f and σ are affine mappings. Suppose that at $\bar{u} \in \mathcal{U}$, the quadratic form Υ_2 is a Legendre form and there exist $\alpha > 0$ such that*

$$\Upsilon_1(v) = 0, \text{ and } \Upsilon_2(v) \geq \alpha \|v\|_2^2 \text{ for all } v \in C(\bar{u}). \quad (67)$$

Then, there exists $\delta > 0$ such that for all $u \in \mathcal{U}$ with $\|u - \bar{u}\|_\infty \leq \delta$, we have

$$J(u) \geq J(\bar{u}) + \frac{1}{2}\alpha \|u - \bar{u}\|_2^2. \quad (68)$$

In the deterministic case there is a well known sufficient condition for the associated quadratic form to be a Legendre form, which is based essentially in the fact that the application $u \in L^2([0, T]; \mathbb{R}^m) \rightarrow y_1(u)(T) \in \mathbb{R}^n$ is weakly continuous. We show with two examples that $u \in L^2_{\mathcal{F}} \rightarrow y_1(u)(T) \in L^2_{\mathcal{F}_T}(\mathbb{R}^n)$ is not weakly continuous.

Example 2 (σ dependent on u). Let us take $m = n = 1$ and let us consider the dynamics

$$dy_1(t) = u(t)dW(t) \text{ for } t \in [0, T]; \quad y_1(0) = 0.$$

Let u_n be a (deterministic) orthonormal base of $L^2([0, T]; \mathbb{R})$ and denote $y_n := y_1(u_n)$. By the dominated convergence theorem it is easy to check that u_n converges weakly to 0 in $L^2_{\mathcal{F}}$, but

$$\mathbb{E} [y_n(T)^2] = \mathbb{E} \left[\left(\int_0^T u_n(t)dW(t) \right)^2 \right] = \int_0^T u_n^2(t)dt = 1.$$

Example 3 (σ independent on u). We take $m = n = 1$ and $T = 2$. Let us consider the dynamics

$$dy_1(t) = u(t)dt \text{ for } t \in [0, T]; \quad y_1(0) = 0.$$

Let ϕ_n be an orthonormal base of the Hilbert space $L^2(\mathbb{R})$ endowed with the scalar product

$$\langle g, h \rangle_* := \int_{-\infty}^{+\infty} g(x)h(x)e^{-\frac{x^2}{2}} dx,$$

and consider the sequence $u_n \in L^2_{\mathcal{F}}$ defined by $u_n(t) := \phi_n(W(1))\mathbb{I}_{(1,2]}(t)$ and set $y_n := y_1(u_n)$. For every $f \in L^2_{\mathcal{F}}$, we have

$$\begin{aligned} \mathbb{E} \left(\int_0^2 f(t)u_n(t) dt \right) &= \mathbb{E} \left(\phi_n(W(1)) \int_1^2 f(t)dt \right), \\ &= \mathbb{E} \left[\phi_n(W(1)) \mathbb{E} \left(\int_1^2 f(t)dt | W(1) \right) \right] \rightarrow 0, \end{aligned}$$

by definition of ϕ_n . Thus u_n converges weakly to 0 in $L^2_{\mathcal{F}}$. On the other hand,

$$\mathbb{E} (y_n(T)^2) = \mathbb{E} \left(\left[\int_0^2 u_n dt \right]^2 \right) = \mathbb{E} (\phi_n(W(1))^2) = 1.$$

References

- [1] A. Bensoussan. *Lectures on stochastic control*. Lecture Notes in Math. Vol. 972, Springer-Verlag, Berlin.
- [2] A. Bensoussan. Stochastic maximum principle for distributed parameter system. *J. Franklin Inst.*, 315:387–406, 1983.
- [3] J.M. Bismut. *Analyse convexe et probabilités*. PhD thesis, Faculté des Sciences de Paris, 1973.

- [4] J.M. Bismut. Théorie probabiliste du contrôle des diffusions. *Mem. Amer. Math. Soc.*, 4:1–130, 1976.
- [5] J.M. Bismut. An introductory approach to duality in optimal stochastic control. *SIAM Rev.*, 20:62–78, 1978.
- [6] J.F. Bonnans and E. Casas. On the choice of the function spaces for some state-constrained control problems. *Numerical Functional Analysis and Optimization*, 7–4:333–348, 1984.
- [7] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer-Verlag, New York, 2000.
- [8] W.H. Fleming and H.M. Soner. *Controlled Markov processes and viscosity solutions*. Springer, New York, 1993.
- [9] A. Haraux. How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities. *J. Mathematical Society of Japan*, 29:615–631, 1977.
- [10] U.G. Haussmann. General necessary conditions for optimal control of stochastic systems. *Math. Prog. Study*, 6:34–48, 1976.
- [11] H.J. Kushner. On the existence of optimal stochastic controls. *J. Math. Anal. Appl.*, 3:463–474, 1965.
- [12] H.J. Kushner. On the stochastic maximum principle: Fixed time of control. *J. Math. Anal. Appl.*, 11:78–92, 1965.
- [13] H.J. Kushner. Necessary conditions for continuous parameter stochastic optimization problems. *SIAM J. on Control and Optimization*, 10:550–565, 1972.
- [14] F. Mignot. Contrôle dans les inéquations variationnelles. *J. Functional Analysis*, 22:25–39, 1976.
- [15] L. Mou and J. Yong. A variational formula for stochastic controls and some applications. *Pure and Applied Mathematics Quarterly*, 3:539–567, 2007.
- [16] S. Peng. A general stochastic maximum principle for optimal control problems. *SIAM J. on Control and Optimization*, 28:966–979, 1990.
- [17] H. Pham. *Optimisation et contrôle stochastique appliqués à la finance*, volume 61 of *Mathématiques & Applications*. Springer, Berlin, 2007.
- [18] J. Yong and X.Y. Zhou. *Stochastic controls, Hamiltonian systems and HJB equations*. Springer-Verlag, New York, Berlin, 2000.
- [19] X.Y. Zhou. The connection between the maximum principle and dynamic programming in stochastic control. *Stoch. & Stoch. Rep.*, 31:1–13, 1990.



Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399