

FILIPPOV APPROACH IN STOCHASTIC MAXIMUM PRINCIPLE WITHOUT DIFFERENTIABILITY ASSUMPTIONS

MOKHTAR HAFAYED

ABSTRACT. In this article, we establish necessary conditions for optimality in stochastic control of systems governed by stochastic differential equations with nonsmooth coefficients. The approach used is based on the approximation of the nonsmooth coefficient by smooth one which generate a sequence of smooth control problems. Ekeland's variational principle is then applied to obtain a sequence of nearly optimal controls which satisfy necessary conditions for near optimality. By using the generalized notion of Filippov's solutions and the stable convergence, we obtain an explicit formula for the adjoint process and the inequality between the Hamiltonians, on a good extension of the initial filtered probability space.

1. INTRODUCTION

We study a stochastic control problem where the system is governed by a nonlinear stochastic differential equation (SDE for short) of the form

$$\begin{aligned}dX_t &= b(t, X_t, u_t)dt + \sigma(t, X_t)dB_t, \\ X_0 &= x.\end{aligned}\tag{1.1}$$

Where B_t is a d -dimensional Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. The finite horizon cost function to be minimized over admissible controls is given by

$$J(u) = \mathbb{E}(g(X_T))\tag{1.2}$$

where u is an admissible control and X_T is a diffusion process solution of (1.1) at the terminal time T . A control $\hat{u} \in \mathcal{U}_{\text{ad}}$ is called optimal if it satisfies $J(\hat{u}) = \inf_{u \in \mathcal{U}_{\text{ad}}} \{J(u)\}$. The corresponding state trajectory \hat{X} and (\hat{X}, \hat{u}) are called an optimal state trajectory and optimal pair respectively.

The stochastic maximum principle (SMP in short) has been and remains an important tool in the many areas in which optimal control plays a role. Pontryagin *et al* [17] announced the maximum principle for the first time. Kushner [14] employed the spike variation and Neustadt's variational principle to derive a stochastic maximum principle. On the other hand, Haussmann [12] extensively investigated the necessary conditions of stochastic optimal state feed-back controls based on the

2000 *Mathematics Subject Classification*. 60H10, 34F05.

Key words and phrases. Stochastic differential equation; generalized Filippov's solutions; optimal control; maximum principle; Ekeland's variational principle.

©2010 Texas State University - San Marcos.

Submitted April 15, 2010. Published July 15, 2010.

Girsanov's transformation.

The case of stochastic systems with nonsmooth coefficients has been treated in [3, 4, 15, 19]. Bahlali *et al* [3] employed the Krylov's inequality to derive a stochastic maximum principle with nonsmooth coefficients and nondegenerate diffusion. Necessary conditions for optimality for degenerate diffusion with nonsmooth coefficient is established by Bahlali *et al* [4].

The necessary conditions for optimality for diffusion with nonsmooth drift, has been solved by Mezerdi [15] by using Clarke's generalized gradient and stable convergence of probability measures. A difficulty is treating the case where the diffusion coefficient σ contains the control variable u . Among those works one can see [2, 5, 16]. Peng [16] introduced the second-order adjoint equation and obtained the maximum principle in which the control enters both the drift and the diffusion coefficients where the set of controls is not necessarily convex.

A good account and an extensive list of references on the maximum principle and optimal stochastic control can be founded in Yong *et al* [20]. Filippov [10] has developed a solution concept for ordinary differential equations (ODEs in short) with a discontinuous right-hand side. When a function V is locally Lipschitz continuous, the associated Filippov differential inclusion is equal to the Clarke's generalized gradient of V .

The main contribution of the present paper is to extend the stochastic maximum principle to the case where the drift and the diffusion coefficients are nonsmooth in the sense that they are only Lipschitz continuous and satisfy a linear growth condition. Our approach is to express a generalized derivative of b and σ using a Filippov differential inclusion type argument in terms of well defined smooth approximations, and stable convergence of probability measures to characterize the first order adjoint equation. A similar type of stochastic maximum principle has primarily been derived in Mezerdi [15] with non-differentiable drift using Clarke's generalized gradient.

The novelty in our maximum principle is based on the advantage of the Filippov's approach which allows to express the generalized gradient in terms of the underlying approximating sequence b_x^n and σ_x^n constructed in Section 5, a property that is not explicit in Clarke's approach.

The rest of the paper is organized as follows. In the section 2, we present the formulation of the problem. Section 3 is devoted to the classical maximum principle. In section 4, we give some proprieties of Filippov notion. Section 5 contains our main result where we give a generalized stochastic maximum principle for our stochastic control problems.

2. PROBLEM FORMULATION AND PRELIMINARIES

Throughout this paper, we assume $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is a filtered probability space and B_t an Brownian motion with values in \mathbb{R}^d . Let \mathbb{A} be a Borelian subset of \mathbb{R}^n , u_t is called an admissible control if it is measurable and \mathcal{F}_t -adapted with values in \mathbb{A} . We denote \mathcal{U}_{ad} the space of admissible controls, B_t^j the j^{th} column of B_t and σ^j the j^{th} column of the matrix σ . Let $b : [0, T] \times \mathbb{R}^d \times \mathbb{A} \rightarrow \mathbb{R}^d$ and the diffusion matrix $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ are Borelian functions such that: for all $(t, x, y, u) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{A}$, there exists positive constants K and c such that

$$|\sigma(t, x) - \sigma(t, y)| + |b(t, x, u) - b(t, y, u)| \leq K|x - y|, \quad (2.1)$$

$$|\sigma(t, x)| + |b(t, x, u)| \leq c(1 + |x|), \quad (2.2)$$

$$b(t, x, \cdot) : \mathbb{A} \rightarrow \mathbb{R}^d \text{ is continuous.} \quad (2.3)$$

From assumptions (2.1) and (2.2) it is easy to see that equation (1.1) satisfies the usual Itô conditions; therefore it has a unique strong solution such that for any $q \geq 1$:

$$\mathbb{E}[\sup_{t \leq T} |X_t|^q] < +\infty.$$

The cost function to be minimized define by (1.2) satisfies the following condition

$$g : \mathbb{R}^d \rightarrow \mathbb{R} \text{ is continuously differentiable,} \quad (2.4)$$

such that $|g(X)| \leq c[1+|X|]$ and $|g_x(X)| \leq M$; where g_x denote the gradient of g at x . Finally throughout this paper, we assume that an optimal control \hat{u} is supposed to exist.

3. CLASSICAL STOCHASTIC MAXIMUM PRINCIPLE

In the regular case, the control problem based to defined an admissible control \hat{u} which minimizes a cost $J(u)$. The conditions must satisfied by the control \hat{u} which is supposed to exists, are called the stochastic maximum principle. In this case we assume

$$b(t, \cdot, u), \sigma^j(t, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ is continuously differentiable.} \quad (3.1)$$

To obtain these necessary conditions for optimality, we compare \hat{u} with controls which are strong perturbations defined by

$$u_h(t) = \begin{cases} v & \text{if } t \in [t_0, t_0 + h], \\ \hat{u} & \text{otherwise.} \end{cases}$$

We define the Hamiltonian $H(t, x_t, u_t, p_t) := p_t b(t, x_t, u_t)$ where $p_t b(t, x_t, u_t)$ is the scalar product in \mathbb{R}^d .

Lemma 3.1. (1) Let X_h the corresponding trajectory of u_h then

$$\mathbb{E}(\sup_{t \leq T} |X_t^h - \hat{X}_t|)^2 \leq Kh^2.$$

(2) Let $\Phi(t)$ be the solution of the linear stochastic differential equation

$$\begin{aligned} d\Phi(t) &= b_x(t, \hat{X}_t, \hat{u}_t)\Phi(t)dt + \sum_{1 \leq j \leq d} \sigma_x^j(t, \hat{X}_t)\Phi(t)dB_t^j, \\ \Phi_0 &= b(t, \hat{X}_t, v) - b(t, \hat{X}_t, \hat{u}), \end{aligned} \quad (3.2)$$

where b_x and σ_x^j are the derivatives of b and σ^j ($j = 1, \dots, d$) in the state variable x . Then

$$\lim_{h \rightarrow 0} \mathbb{E}(|\frac{X_T^h - \hat{X}_T}{h} - \Phi(T)|^2) = 0.$$

(3) $\frac{d}{dh} \{J(u_h)\}|_{h=0} = \mathbb{E}[H(t, \hat{X}_t, \hat{u}, p_t)] - \mathbb{E}[H(t, \hat{X}_t, v, p_t)].$

See Bensoussan [5] or Mezerdi [15] for the a detailed proof of the above lemma. Under the differentiability assumptions (3.1), the regular version of the stochastic maximum principle given by the following lemma.

Lemma 3.2. *Let (\hat{X}, \hat{u}) be an optimal pair. Then there exists an adapted process $p(t)$ satisfying*

$$p(t) = -\mathbb{E}[\Phi^*(T, t)g_x(\hat{X}_T)|\mathcal{F}_t], \quad (3.3)$$

$$H(t, \hat{X}_t, \hat{u}, p_t) = \max_{v \in \mathbb{A}} H(t, \hat{X}_t, v, p_t) dt \quad a.e. \mathbb{P} \text{ a.s.}, \quad (3.4)$$

where $\Phi^*(T, t)$ is the transpose of $\Phi(T, t)$ solution to (3.2).

See Mezerdi [15] or Yong *et al* [20] for the detailed proof of the above lemma.

We call $p(t)$ the adjoint process, (3.3) the adjoint equation and (3.4) the maximum condition. A control u^ε is called a near optimal if for all $\varepsilon > 0$, there exists u^ε such that

$$J(u^\varepsilon) \leq \inf\{J(u) : u \in \mathcal{U}_{\text{ad}}\} + \varepsilon.$$

In this part we establish necessary conditions of near optimality satisfied by a sequence of nearly optimal strict controls. This result is based on Ekeland's variational principle, which is given as follows.

Lemma 3.3 (Ekeland's Lemma). *Let (E, d) be a complete metric space and $f : E \rightarrow \mathbb{R}$ be lower semicontinuous and bounded from below. Given $\varepsilon > 0$ and $u^\varepsilon \in E$ satisfies $f(u^\varepsilon) \leq \inf(f) + \varepsilon$. Then for any $\lambda > 0$, there exists $v \in E$ such that*

- (i) $f(v) \leq f(u^\varepsilon)$.
- (ii) $d(u^\varepsilon, v) \leq \lambda$.
- (ii) $f(v) < f(w) + \frac{\varepsilon}{\lambda}d(v, w)$ for all $w \neq v$.

See Ekeland [9] for a proof of the above lemma. To apply Ekeland's variational principle to our problem we must define a distance d on the space of admissible controls such that $(\mathcal{U}_{\text{ad}}, d)$ becomes a complete metric space, we pose: for any $u, v \in \mathcal{U}_{\text{ad}}$

$$d(u, v) = \mathbb{P} \otimes dt\{(w, t) \in \Omega \times [0, T] : u(w, t) \neq v(w, t)\}.$$

Where $\mathbb{P} \otimes dt$ is the product measure of \mathbb{P} with the Lebesgue measure dt on $[0, T]$. In the next lemma we give some properties.

Lemma 3.4. (1) $(\mathcal{U}_{\text{ad}}, d)$ is a complete metric space.

(2) Let $u, v \in \mathcal{U}_{\text{ad}}$, then the following estimate holds

$$\mathbb{E}(\sup_{t \leq T} |X_t^u - X_t^v|^2) \leq K[d(u, v)]^{1/2},$$

where X^u and X^v are the corresponding trajectories of u and v .

Proof. The proof of (1) can be found in Yong *et al* [20, pp. 146-147]. Item (2) is proved using Lipschitz assumptions on the coefficients, Burkholder-Davis-Gandy and Cronwall's inequalities. \square

Denote by X^ε the unique solution of the stochastic differential equation

$$\begin{aligned} dX_t^\varepsilon &= b(t, X_t^\varepsilon, u_t^\varepsilon)dt + \sigma(t, X_t^\varepsilon)dB_t, \\ X_0^\varepsilon &= x. \end{aligned} \quad (3.5)$$

Let $\Phi^\varepsilon(t, s)$ be the fundamental solution of the linear stochastic differential equation ($t \geq s$)

$$\begin{aligned} d\Phi^\varepsilon(t) &= b_x(t, X_t^\varepsilon, u_t^\varepsilon)\Phi(t)dt + \sum_{1 \leq j \leq d} \sigma_x^j(t, X_t^\varepsilon)\Phi(t)dB_t^j, \\ \Phi^\varepsilon(t, t) &= I_d, \end{aligned} \quad (3.6)$$

and the adjoint process associated to u^ε is given by

$$p^\varepsilon(t) = -\mathbb{E}[\Phi^{*\varepsilon}(T, t)g_x(X_T^\varepsilon)|\mathcal{F}_t].$$

Lemma 3.5 (Approximate maximum principle). *For each $\varepsilon > 0$ there exist u_t^ε and an adapted process $p^\varepsilon(t)$, given by (3.6), such that for all $v \in \mathbb{A}$,*

$$\mathbb{E}[\langle p^\varepsilon(t), b(t; X_t^\varepsilon, v) \rangle] \leq \mathbb{E}[\langle p^\varepsilon(t), b(t; X_t^\varepsilon, u_t^\varepsilon) \rangle] + \varepsilon dt - a.e.$$

Proof. Since u^ε is optimal for the cost $J_\varepsilon(u) = J(u) + \varepsilon d(u, u^\varepsilon)$, then we apply results of the last section to derive the adjoint process and the inequality between Hamiltonians. Notice that if u_h^ε denotes a strong perturbation of u^ε then $d(u_h^\varepsilon, u^\varepsilon) = h$ (see Mezerdi [15], Bensoussan [5]). □

4. FILIPPOV’S SET-VALUED MAP AND GENERALIZED GRADIENT

4.1. Filippov’s set-valued map. We give in this section some basic notions and concepts concerning generalized Filippov’s set-valued map which is described briefly by the following.

Let us consider a function $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ to which we associate the following set-valued map called Filippov’s regularization of b ,

$$F_b(x) := \bigcap_{\lambda(N)=0} \bigcap_{\delta>0} \overline{\text{co}}b((x + \delta B) - N), \tag{4.1}$$

where $\overline{\text{co}}(A)$ means the closure of the convex hull of A . The first intersection $\bigcap_{\lambda(N)=0}$ is taken over all sets of \mathbb{R}^n , being negligible with respect to Lebesgue measure λ and B is the closed unit ball.

Let us consider a function $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ to which we associate the following ordinary differential equation,

$$x'(t) = b(x(t)), \quad t \geq 0, \quad x(0) = x. \tag{4.2}$$

Without regularity assumptions on f (Lipschitz continuity), it is well known that neither existence, nor uniqueness hold true in general.

An absolutely continuous solution $t \in [0, +\infty) \mapsto x(t) \in \mathbb{R}^n$ is a Filippov’s solution of the ODE (4.2) if and only if it is solution of the differential inclusion

$$x'(t) \in F_b(x), \quad t \geq 0, \quad X(0) = x. \tag{4.3}$$

The set valued map F_b is upper semi continuous with compact convex values. This implies that the differential inclusion (4.3) has a nonempty set of (local) solution (Aubin [1]). In the following proposition we summarized some proprieties.

Proposition 4.1. *Let $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a measurable and bounded function. Then we have:*

- (1) *There exists a negligible-set N_b under the Lebesgue measure such that for any $x \in \mathbb{R}^n$:*

$$F_b(x) = \bigcap_{\delta>0} \overline{\text{co}}b((x + \delta B) - N_b), \tag{4.4}$$

- (2) *For almost all $x \in \mathbb{R}^n$, we have $b(x) \in F_b(x)$.*
- (3) *The set valued map F_b is the smallest upper semi continuous set-valued map F with closed convex values such that $b(x) \in F(x)$, for almost all $x \in \mathbb{R}^n$.*
- (4) *The map $b \mapsto F_b$ is single-valued if and only if there exists a continuous function g which coincides almost everywhere with b . In this case we have $F_b(x) = \{g(x)\}$. for almost all $x \in \mathbb{R}^n$.*

- (5) If a function \tilde{b} coincide almost everywhere with b then $F_b(x) = F_{\tilde{b}}(x)$ for all $x \in \mathbb{R}^n$.
- (6) There exists a function \bar{b} which is equal almost everywhere to b and such that

$$F_b(x) = \cap_{\delta>0} \overline{\text{co}} \bar{b}(x + \delta B).$$

- (7) We have $F_b(x) = \cap_{b=\bar{b}} \cap_{\delta>0} \overline{\text{co}} \bar{b}(x + \delta B)$, where the first intersection is taken over all functions \bar{b} being equal to b almost everywhere.

See Bukhdahn *et al* [8] for a proof of the above proposition.

As an example, in the one dimensional case ($n = 1$), we have $b : \mathbb{R} \rightarrow \mathbb{R}$ for which one can check that: for all $x \in \mathbb{R}^n$:

$$F_b(x) = [\underline{m}_b(x), \overline{m}_b(x)]$$

where

$$\underline{m}_b(x) := \sup_{\delta>0} (\text{ess inf}_{[x-\delta, x+\delta]} b), \quad \overline{m}_b(x) := \inf_{\delta>0} (\text{ess sup}_{[x-\delta, x+\delta]} b).$$

In the case where $b(x) = \text{sgn}(x)$, then we have $F_b(0) = [-1, 1]$.

4.2. Connection between Filippov's approach and Clarke's generalized gradient. We give in this subsection the connection between Filippov's differential inclusion and the Clarke's generalized gradient.

The Clarke's generalized gradient. Let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz continuous. We define the generalized gradient of V as

$$\partial_c V(x) = \overline{\text{co}} \left\{ \lim_{x_i \rightarrow x} \nabla V(x_i), x_i \notin \Omega_V \cup N \right\}, \quad (4.5)$$

where Ω_V is the set of Lebesgue measure zero where ∇V does not exist and N is an arbitrary set of measure zero.

Lemma 4.2. *The map $F : \{b : \mathbb{R}^m \rightarrow \mathbb{R}^n\} \rightarrow \{g : \mathbb{R}^m \rightarrow 2^{\mathbb{R}^n}\}$ has the following properties:*

(1) Assume that $b : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is locally bounded. Then $\exists N_b \subset \mathbb{R}^m$, $\lambda(N_b) = 0$ such that $\forall N \subset \mathbb{R}^m$, $\lambda(N) = 0$.

$$F_b(x) = \overline{\text{co}} \left\{ \lim_{x_i \rightarrow x} b(x_i), x_i \notin N_b \cup N \right\}.$$

(2) Assume that $b, f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is locally bounded; then

$$F_{(b+f)}(x) \subset F_b(x) + F_f(x).$$

(3) Assume that $b_j : \mathbb{R}^m \rightarrow \mathbb{R}^{n_j}$ where $j \in \{1, 2, \dots, N\}$ are locally bounded; then

$$F_{\left(\prod_{j=1}^{j=N} b_j\right)}(x) \subset \prod_{j=1}^{j=N} F_{b_j}(x).$$

(4) Let $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is C^1 , rank $Dg(x) = n$ and $b : \mathbb{R}^n \rightarrow \mathbb{R}^p$ be locally bounded; then

$$F_{b \circ g}(x) = F_b(g(x)).$$

(5) Let $g : \mathbb{R}^m \rightarrow \mathbb{R}^{p \times n}$ (i.e. matrix valued) be C^0 and $b : \mathbb{R}^n \rightarrow \mathbb{R}^p$ be locally bounded; then $F_{gb}(x) = g(x)F_b(x)$, where $gb(x) := g(x)b(x)$.

(6) Let $V : \mathbb{R}^m \rightarrow \mathbb{R}$ be locally Lipschitz continuous, then $F_{\nabla V}(x) = \partial_c V(x)$.

The proof of the above lemma can be found in Paden *et al* [18].

Remark 4.3. (i) Since V is locally Lipschitz, ∇V is defined almost everywhere and locally bounded (Rademacher's Theorem). By using Lemma 3.5 we have $F_{\nabla V}(x) = \partial_c V(x)$.

(ii) In particular, if V is Fréchet-differentiable at x , then

$$F_{\nabla V}(x) = \partial_c V(x) = \{V'(x)\}.$$

5. MAIN RESULTS

In this section we establish generalized stochastic maximum principle for diffusion without differentiability assumptions on the coefficients b , σ^j satisfies the assumptions (2.1) and (2.2), so we are going to weaken the differentiability assumptions on this coefficients. This method is described briefly by the following.

Let E be a Banach space, E^* its dual and let $f : V \rightarrow \mathbb{R}^d$ where V is a closed subset of E , satisfying the following conditions:

- (H1) There exist $\lambda_n > 0$, and $f^n : B \rightarrow \mathbb{R}^d$ Gâteaux-differentiable in the ball $(y + \lambda_n B)$ and $f^n(y) = f(y)$.
- (H2) There exists $\varepsilon_n > 0$ such that $\frac{\varepsilon_n}{\lambda_n} \rightarrow 0$ as $n \rightarrow +\infty$, f^n is continuous and $|f^n(x) - f(x)| \leq \varepsilon_n$ for $x \in (y + \lambda_n B)$.

We shall approximate the drift b and the diffusion σ^j by a sequence of smooth functions b^n and $\sigma^{j,n}$ by using the following regularization

Let $\xi_n : \mathbb{R} \rightarrow \mathbb{R}$ be a positive C^∞ application vanishing out of the interval $[-\varepsilon_n, \varepsilon_n]$ such that $\int_{\mathbb{R}} \xi_n(x) dx = 1$ and $\lim_{n \rightarrow +\infty} \varepsilon_n \rightarrow 0$. We denote $\rho_n = \prod_{j=1}^{j=d} \xi_n(x^j)$ where $x = (x^1, \dots, x^d) \in \mathbb{R}^d$. ρ_n is a C^∞ function with compact support.

We define the following smooth functions $b^n = b * \rho_n$ and $\sigma^{j,n} = \sigma^j * \rho_n$ obtained by the convolution of all the components of b and σ^j with ρ_n . We give in the next theorem the following properties satisfied by these functions.

Lemma 5.1. (1) $b^n : [0, T] \times \mathbb{R}^d \times \mathbb{A} \rightarrow \mathbb{R}^d$ and $\sigma^{j,n} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ are Borelian.
 (2) b^n and $\sigma^{j,n}$ are k -Lipschitzian in the second variable x and has linear growth.
 (3) b^n and $\sigma^{j,n}$ are a C^∞ in x , and for all $(t, x, u) \in [0, T] \times \mathbb{R}^d \times \mathbb{A}$: $|b^n(t, x, u) - b(t, y, u)| \leq k\varepsilon_n$, and $|\sigma^n(t, x) - \sigma(t, y)| \leq k\varepsilon_n$.

The statements in the above lemma are a classical facts; see Frankowska [11] and Mezerdi [15].

Note that b^n and $\sigma^{j,n}$ satisfies conditions (H1) and (H2) with $\lambda_n = \sqrt{\varepsilon_n}$. So we can define

$$\begin{aligned} \partial_c b(t, y, u) &= \bigcap_{n \geq 0} \overline{c_0} \cup_{k \geq n} [b_x^k(t, x, u) : x \in (y + \lambda_n B)]. \\ \partial_c \sigma^j(t, y) &= \text{cap}_{n \geq 0} \overline{c_0} \cup_{k \geq n} [\sigma_x^{j,k}(t, x) : x \in (y + \lambda_n B)]. \end{aligned}$$

Let X^n be the solution of

$$\begin{aligned} dX_t^n &= b^n(t, X_t^n, u_t^n) dt + \sigma^n(t, X_t^n) dB_t, \\ X_0^n &= x. \end{aligned} \tag{5.1}$$

Let $\Phi_n(s, t)$ denote the fundamental solution of the linear equation ($s \geq t$),

$$\begin{aligned} d\Phi_n(t) &= b_x^n(t, X_t^n, u_t^n) \Phi_n(t) dt + \sum_{1 \leq j \leq d} \sigma_x^{j,n}(t, X_t^n) \Phi_n(t) dB_t^j, \\ \Phi_n(s, s) &= I_d. \end{aligned} \tag{5.2}$$

The following lemma will play an interesting role below.

Lemma 5.2. *There exists $u^n \in \mathcal{U}_{\text{ad}}$ and $\lambda_n = \sqrt{\delta_n}$ such that*

- (i) $d(u^n, u^*) \leq \lambda_n$.
- (ii) $\mathbb{E}[H_n(t; X_t^n, u_t^n, p_t^n)] \geq \mathbb{E}[H_n(t; X_t^n, v, p_t^n)] - \lambda_n$, for all $v \in A$, $dt - a.e.$ for all $v \in \mathbb{A}$. The associate adjoint process is given by

$$p_n(t) = -\mathbb{E}\{\Phi_n^*(T, t)g_x(X_T^n)|\mathcal{F}_t\}.$$

Proof. Since u_t^n is optimal for the cost $J_n(u) + (\delta_n)^{1/2}d(u, u^n)$ then, we proceed as in Lemma 3.2 to derive a maximum principle for u_t^n . The rest of the proof is similar to the approximate maximum principle, see Yong *et al* [20]. \square

Notice that since the gradient g_x is continuous, then we have $g_x(X_T^n) \rightarrow g_x(\hat{X}_T)$ \mathbb{P} -a.s. as $n \rightarrow +\infty$.

5.1. Weak limit solution with stable convergence. The limit of Φ_n is proved by using the stable convergence of probability measure introduced by Jacod *et al* [13]. This convergence is contained between convergence in law and convergence in probability. We shall make use of the notion of good extension of a filtered probability space.

Definition 5.3. The space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{\mathbb{P}})$ is a good extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ if the following conditions are satisfied

- (1) $\bar{\Omega} = \Omega \times \hat{\Omega}$ where $\hat{\Omega}$ is an auxiliary space.
- (2) $\mathcal{F} \in \bar{\mathcal{F}}$; i.e., $A \times \hat{\Omega} \in \bar{\mathcal{F}}$, for $A \in \mathcal{F}$, $\mathcal{F}_t \in \bar{\mathcal{F}}_t$.
- (3) $\bar{\mathbb{P}}(A \times \hat{\Omega}) = \mathbb{P}(A)$ for $A \in \mathcal{F}$.
- (4) Each $(\mathcal{F}_t - \mathbb{P})$ martingale is a $(\bar{\mathcal{F}}, -\bar{\mathbb{P}})$ martingale.

Clearly, since b^n and $\sigma^{j,n}$ are k -Lipshitz in x and continuously differentiable, then the matrix of partial derivatives b_x^n and $\sigma_x^{j,n}$ are bounded by the Lipschitz constant k .

Let us define the canonical spaces associated to the processes $b_x^n(t, X_t^n, u_t^n)$, $\sigma_x^{j,n}(t, X_t^n)$, $(\sigma_x^n)^*(\sigma_x^n)$ and $\Phi_n(t)$:

- (1) Let Ω_1 , the canonical space of $b_x^n(t, X_t^n, u_t^n)$, define by the following: Let $D_1 = \{\beta_1 : [0, T] \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d \text{ measurable such that } \|\beta_1\| \leq c\}$.

It is clear that b_x^n take values in D_1 which is uniformly integrable subset of $\mathbb{L}^1([0, T], \mathbb{R}^d \otimes \mathbb{R}^d)$, hence it is a relatively compact subset with respect to the weak topology $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ (Dunford-Pettis Theorem). Let $\Omega_1 = \bar{D}_1$ (weak or strong closure of D_1 , because D_1 is convex).

We define \mathcal{F}^1 the filtration of the coordinates generated by the subsets of the form

$$A = \{\beta_1 \in \Omega_1 : \int_0^t \langle \beta_1, f(s) \rangle ds \leq c, \text{ where } c \in \mathbb{R}, f \in L^\infty([0, T], \mathbb{R}^d \otimes \mathbb{R}^d)\}.$$

$(\Omega_1, \mathcal{F}^1, \mathcal{F}_t^1)$ is the canonical space associated to the process $b_x^n(t, X_t^n, u_t^n)$.

- (2) Let $D_2 = \{\beta_2^j : [0, T] \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d \text{ measurable such that } \|\beta_2^j\| \leq c\}$. It is clear that $\sigma_x^{j,n}(t, X_t^n)$ take values in D_2 which is uniformly integrable subset of $\mathbb{L}^1([0, T])$, hence it is a relatively compact subset with respect to the weak topology $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ (Dunford-Pettis Theorem). Let $\Omega_2 = \bar{D}_2$ (weak or strong closure of D_2 , because D_2 is convex). Then Ω_2 are compact metrisible spaces.

We define \mathcal{F}^2 the filtration of the coordinates generated by the subsets of the form

$$B = \{\beta_2^j : \int_0^t \langle \beta_2^j, h(s) \rangle ds \leq c, \text{ where } c \in \mathbb{R}, h \in L^\infty([0, T], \mathbb{R}^d \otimes \mathbb{R}^d)\}.$$

where $(\Omega_2, \mathcal{F}^2, \mathcal{F}_t^2)$ is the canonical space associated to the process $\sigma_x^{j,n}(t, X_t^n)$.

(3) Let $\Omega_3 = \{a \in \mathbb{L}^2([0, T]) \text{ such that } \|a\| \leq c\}$, $(\Omega_3, \mathcal{F}^3, \mathcal{F}_t^3)$ is the canonical space associated to the process $a_n = \sigma_x^{j,n}(t, X_t^n) \cdot \sigma_x^{j,n}(t, X_t^n)$.

(4) Φ_n has a continuous trajectories, then $\Phi_n : (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}) \rightarrow \Omega_4$ where Ω_4 is the space of continuous functions from $[0, T]$ to $\mathbb{R}^d \otimes \mathbb{R}^d$ equipped with the topology of uniform convergence and \mathcal{F}_t^4 the filtration coordinates. By this definitions, we introduce the product space

$$\bar{\Omega} = \Omega \times \Omega_1 \times \Omega_2 \times \Omega_3 \times \Omega_4,$$

equipped with the filtration

$$\bar{\mathcal{F}}_t = \cap_{s \geq t} \mathcal{F}_s \otimes \mathcal{F}_s^1 \otimes \mathcal{F}_s^2 \otimes \mathcal{F}_s^3 \otimes \mathcal{F}_s^4.$$

We associate with $(b_x^n(\cdot, X^n, u^n), \sigma_x^{j,n}(\cdot, X^n), a_n(\cdot, X^n), \Phi_n)$ the randomized variable $\bar{\mathbb{P}}_n$ which is a probability measure defined on $(\bar{\Omega}, \bar{\mathcal{F}})$ by

$$\bar{\mathbb{P}}_n(w, w_1, w_2, w_3, w_4) = \mathbb{P}(w) \delta_{b_x^n}(dw_1) \delta_{\sigma_x^{j,n}}(dw_2) \delta_{a_n}(dw_3) \delta_{\Phi_n}(dw_4),$$

where δ_x the Dirac measure at x and Φ_n solution of (5.2).

Theorem 5.4. *The space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{\mathbb{P}}_n)$ is a good extension of the space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, moreover the canonical process $\Phi_t(w, w_1, w_2, w_3, w_4) = w_4(t)$ is a solution of the stochastic differential equation*

$$d\Phi(t) = \beta_1(t)\Phi(t)dt + \sum_{1 \leq j \leq d} \beta_2^j(t)\Phi(t)dB_t^j$$

$$\Phi(0) = I_d,$$

on the space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{\mathbb{P}}_n)$, where \hat{B}_t is a Brownian motion which is independent of B_t .

Proof. Let $\mathbb{E}, \bar{\mathbb{E}}_n$ and $\bar{\mathbb{E}}$ the expectation with respect to the randomized probability $\mathbb{P}, \bar{\mathbb{P}}_n$ and $\bar{\mathbb{P}}$ respectively. It is sufficient to verify that all $(\mathcal{F}_t - \mathbb{P})$ martingale is $(\bar{\mathcal{F}}_t, \bar{\mathbb{P}}_n)$ martingale. \square

The sequence $\bar{\mathbb{P}}_n$ converge with respect to stable convergence to a limit $\bar{\mathbb{P}}$ if and only if:

$$\lim_{n \rightarrow +\infty} \bar{\mathbb{P}}_n[g(w, w_1, w_2, w_3, w_4)] = \bar{\mathbb{P}}[g(w, w_1, w_2, w_3, w_4)].$$

For every function $g : \bar{\Omega} \rightarrow \mathbb{R}$ measurable bounded such that: $g(w, \dots, \dots)$ is continuous for all $w \in \Omega$.

To prove that the sequence $\bar{\mathbb{P}}_n$ is relatively compact with respect to stable convergence, it is sufficient to prove that the projections of $\bar{\mathbb{P}}_n$ on $\Omega_1, \Omega_2, \Omega_3$ and Ω_4 are relatively compact in the topology of narrow convergence.

Lemma 5.5. *(i) Let Φ_n be the solution of (5.2) then there exists a positive constant M such that for all $n \in \mathbb{N}$ and $s, t \in [0, T]$,*

$$\mathbb{E}(\|\Phi_n(t) - \Phi_n(s)\|^4) \leq M|t - s|^2.$$

(ii) The sequence $\bar{\mathbb{P}}_n$ is relatively compact with respect to the topology of stable convergence

Proof. Statement (i) follows from the Schwartz and Burkholder-Davis-Gandy inequalities.

(ii) Since Ω_1 (resp. Ω_2) is compact, then the sequence of the projections of $\bar{\mathbb{P}}_n$ on Ω_1 (resp. Ω_2) is tight, then relatively compact (Prokhorov's Theorem). Moreover the projections of Φ_n on Ω_4 coincides with the distributions of Φ_n which satisfy (i) of Lemma 5.5, then Φ_n is relatively compact with respect to the topology of stable convergence. \square

Theorem 5.6. *Let $\bar{\mathbb{P}}$ be a limit of $\bar{\mathbb{P}}_n$ (in the sense of stable convergence), then $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{\mathbb{P}})$ is a good extension of the space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Moreover the canonical process $\Phi_t(w, w_1, w_2, w_3, w_4) = w_4(t)$ satisfies*

$$d\Phi(t) = \beta_1(t)\Phi(t)dt + \sum_{1 \leq j \leq d} \beta_2^j(t)\Phi(t)dB_t^j + \sum_{1 \leq j \leq d} \hat{\beta}_2^j(t)\Phi(t)d\hat{B}_t^j \tag{5.3}$$

$$\Phi(s, s) = I_d,$$

Proof. For doing this we need the techniques in Mezerdi [15] and in Jacod *et al* [13], so it is sufficient to prove that all $(\mathcal{F}_t - \mathbb{P})$ martingale is $(\bar{\mathcal{F}}_t - \bar{\mathbb{P}})$ martingale. Let M_t be a $(\mathcal{F}_t - \mathbb{P})$ martingale, and \mathcal{Z} a bounded random variable, $\bar{\mathcal{F}}_s$ measurable such that $(w_1, w_2, w_3, w_4) \rightarrow \mathcal{Z}(w_1, w_2, w_3, w_4)$ is continuous.

According to Theorem 5.4, the space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{\mathbb{P}}_n)$ is a good extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. ($s \leq t$). Then $\bar{\mathbb{E}}_n[M_t\mathcal{Z}] = \bar{\mathbb{E}}_n[M_s\mathcal{Z}]$ for all $n \in \mathbb{N}$. Since $\bar{\mathbb{P}}$ is a limit of $\bar{\mathbb{P}}_n$ we have

$$\bar{\mathbb{E}}[M_t\mathcal{Z}] = \lim_{n \rightarrow +\infty} \bar{\mathbb{E}}_n[M_t\mathcal{Z}] = \lim_{n \rightarrow +\infty} \bar{\mathbb{E}}_n[M_s\mathcal{Z}] = \bar{\mathbb{E}}[M_s\mathcal{Z}].$$

The set of random variables \mathcal{Z} are $\bar{\mathcal{F}}_s$ measurable bounded continuous in (w_1, w_2, w_3, w_4) generates a σ -field contained between $\bar{\mathcal{F}}_{s-}$ and $\bar{\mathcal{F}}_s$. Then $\bar{\mathbb{E}}[(M_t - M_s)|\bar{\mathcal{F}}_{s-}] = 0$, the right continuity of M_t implies that $\bar{\mathbb{E}}[(M_t - M_s)|\bar{\mathcal{F}}_s] = 0$.

Φ_t takes values in $\mathbb{R}^d \otimes \mathbb{R}^d$ then Φ_t is a solution of (5.3). Applying similar techniques as those in [13] where it is sufficient to prove that for all $(h_1, h_2) \in \mathbb{R}^d \times \mathbb{R}^d$

$$M_t(h_1, h_2) = h_1 B_t + h_1(\Phi_s - \Phi_0 - \int_0^t \beta_1(s)\Phi_s ds).$$

Here M_t is a $(\bar{\mathcal{F}}_t - \bar{\mathbb{P}})$ martingale and have a quadratic variation of the form

$$A_t(h_1, h_2) = h_1^2 t + 2h_1 h_2 \int_0^t \beta_1(s)\Phi_s ds + h_2^2 \int_0^t a(s)\Phi_s^2 ds.$$

Note that M_t is an $(\bar{\mathcal{F}}_t - \bar{\mathbb{P}}_n)$ martingale for all $n \in \mathbb{N}$ and $(w, w_1, w_2) \rightarrow M_t(h_1, h_2)$ is continuous.

To pass to the limit, we must show that M_t is sufficiently integrable. Because $\beta_1(t)$, and $\beta_2^j(t)$ are bounded and $\mathbb{E}[(\sup_{t \leq T} |\Phi_s^i|)^p] < +\infty$. We deduce that $\sup_n \bar{\mathbb{E}}_n[|M_t(h_1, h_2)|^2] < +\infty, \forall p \geq 1$. Then if \mathcal{Z} is a bounded $\bar{\mathcal{F}}_t$ -measurable random variable continuous in (w_1, w_2, w_3, w_4) we have $\bar{\mathbb{E}}_n[(M_t - M_s)\mathcal{Z}] \rightarrow \bar{\mathbb{E}}[(M_t - M_s)\mathcal{Z}]$ as $n \rightarrow +\infty$. Hence $M_t(h_1, h_2)$ is an $(\bar{\mathcal{F}}_t - \bar{\mathbb{P}})$ martingale.

The extra term $\sum_{1 \leq j \leq d} \hat{\beta}_2^j(t)\Phi_t d\hat{B}_t^j$ comes from the Itô decomposition Theorem for martingales adapted to a filtration supporting a Brownian motion. \square

The same method can be performed for $M_t^2(h_1, h_2) - A_t(h_1, h_2)$. Now we are ready to state our main result.

Theorem 5.7. *Let \hat{u} be an optimal control and \hat{X} corresponding trajectory, then there exists a probability $\bar{\mathbb{P}}$ on the space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t)$ such that*

- (1) $\bar{\mathbb{E}}[H(t, \hat{X}_t, \hat{u}_t, \bar{p}_t)] = \max_{v \in \mathbb{A}} \bar{\mathbb{E}}[H(t, \hat{X}_t, v, \bar{p}_t)], dt\text{-a.e.}$
- (ii) $\bar{p}_t = -\bar{\mathbb{E}}[\Phi^*(T, t)g_x(\hat{X}_T)|\bar{\mathcal{F}}_t]$. where $\Phi^*(T, t)$ is the transpose of $\Phi(T, t)$ given by (5.3).

Proof. According to Lemma 5.2 there exists a control u_t^n such that $d(u_t^n, \hat{u}_t) \leq \lambda_n$. So it is sufficient to prove that

$$\lim_{n \rightarrow +\infty} \mathbb{E}[H(t, X_t^n, u_t^n, p_n(t))] = \bar{\mathbb{E}}[H(t, \hat{X}_t, \hat{u}_t, \bar{p}(t))].$$

□

5.2. Filippov approach and the support of a limit $\bar{\mathbb{P}}$. Our goal in this subsection is to prove a stochastic maximum principle for the optimal controls without differentiability assumptions.

Let $\bar{\mathbb{P}}$ be a stable limit of $(\bar{\mathbb{P}}_n)$, now we give in this section the connection between the support of $\bar{\mathbb{P}}$ and the generalized Filippov's set of b and σ^j at (\hat{X}_t, \hat{u}_t) .

Let $\tilde{\Omega} = \Omega \times \Omega_1 \times \Omega_2 \times \Omega_3$, $\tilde{\mathcal{F}}_t = \cap_{s \geq t} \mathcal{F}_s \otimes \mathcal{F}_s^1 \otimes \mathcal{F}_s^2 \otimes \mathcal{F}_s^3$ and $\tilde{\mathbb{P}}$ is the projection $\bar{\mathbb{P}}$ on $\tilde{\Omega}$ then we have

$$\tilde{\mathbb{P}}(dw, dw_1, dw_2, dw_3) = \bar{\mathbb{P}}(dw, dw_1, dw_2, dw_3, \Omega_4),$$

where $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}})$ is a good extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and $\tilde{\mathbb{P}}$ is a stable limit of $\bar{\mathbb{P}}_n$, where $\bar{\mathbb{P}}_n$ denotes the projection de $\bar{\mathbb{P}}_n$ on $\tilde{\Omega}$. Moreover if we consider equation (5.3) on the space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}})$, then it has a strong unique solution. If we denote by $\tilde{\Phi}_t(w, w_1, w_2, w_3)$ the solution on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}})$ then $\tilde{\Phi}_t = \Phi_t \bar{\mathbb{P}}$ -a.s., and

$$\bar{\mathbb{P}}(dw, dw_1, dw_2, dw_3, dw_4) = \tilde{\mathbb{P}}(dw, dw_1, dw_2, dw_3) \delta_{\tilde{\Phi}_t(w, w_1, w_2, w_3)}(dw_4).$$

If D denotes a support of the probability $\tilde{\mathbb{P}}(dw, dw_1, dw_2, dw_3)$, according to Jacod *et al* [13] there exists a subsequence such that ($\mathbb{P} - a.s$).

D_w^1 : the set of limit points of the subsequence of $b_x^n(t, X^n(w), u^n(w))$ where w is fixed.

D_w^2 : the set of limit points of the subsequence of $\sigma_x^{j,n}(t, X^n(w))$ where w is fixed.

The Filippov differential inclusion allows us to express the generalized derivative of b and σ in terms of well defined smooth approximations. This advantage enables to give the following theorem.

Theorem 5.8. (i) *For almost all w , there exists a subsequence $b_x^n(t, X_t^n(w), \hat{u}_t(w))$ and $\sigma_x^{j,n}(t, X_t^n(w))$ such that*

$$\begin{aligned} \lim_{n \rightarrow +\infty} b_x^n(t, X_t^n(w), \hat{u}_t(w)) &= \beta_1(t) \quad \text{in } \mathbb{L}^1(dt). \\ \lim_{n \rightarrow +\infty} \sigma_x^{j,n}(t, X_t^n(w)) &= \beta_2^j(t) \quad \text{in } \mathbb{L}^1(dt). \end{aligned}$$

(ii) *For almost every $t \in [0, T]$. we have $\beta_1(t) \in F_{\nabla b}(t, \hat{X}_t, \hat{u}_t)$ and $\beta_2^j(t) \in F_{\nabla \sigma^j}(t, \hat{X}_t)$.*

Proof. (i) According to the definition of D_w^1 there exists a subsequences $b_x^n(t, X_t^n(w), u_t^n(w))$ which converges weakly in $\mathbb{L}^1(dt)$ to $\beta_1(t)$. Moreover,

$$\mathbb{E}\left[\int_0^T |b_x^n(s, X_t^n, u_t^n) - b_x^n(t, X_t^n, \hat{u}_t)| dt\right] \leq Md(u^n, \hat{u}_t).$$

Then there exists a subsequence such that

$$\left[\int_0^T |b_x^n(s, X_t^n, u_t^n) - b_x^n(t, X_t^n, \hat{u}_t)| dt\right] \rightarrow 0 \quad \text{as } n \rightarrow +\infty, \quad \tilde{\mathbb{P}}\text{-a.s.},$$

using a similar proof as for $\lim_{n \rightarrow +\infty} \sigma_x^{j,n}(t, X_t^n(w)) = \beta_2^j(t)$.

(ii) Let $\beta_1(t) \in D_w^1$, according to (i) Theorem 5.8, there exists a subsequence $b_x^n(t, X_t^n, \hat{u}_t)$ which converges to $\beta_1(t)$ in $\mathbb{L}^1(dt)$. Moreover, we have

$$\lim_{n \rightarrow +\infty} \mathbb{E}[\sup_{t \leq T} |X_t^n - \hat{X}_t|^2] = 0,$$

so we can extract a subsequences such that $\sup_{t \leq T} |X_t^n - \hat{X}_t| \leq \lambda_n \mathbb{P}$ -a.s., then there exists $n \in \mathbb{N}$ such that

$$b_x^n(t, X_t^n, u_t^n) \in \cup_{k \geq n} [b_x^k(t, X_t, \hat{u}_t) : X_t \in (\hat{X}_t + \lambda_n B)].$$

According to Mazur's Lemma, there exists a sequence of convex combinations which converges strongly in $\mathbb{L}^1(dt)$ to $\beta_1(t)$. Then we have

$$\beta_1(t) \in \cap_{n \geq 0} \overline{\text{co}} \cup_{k \geq n} [b_x^k(t, x, u) : x \in (y + \lambda_n B)] \quad dt - a.e.$$

According to the property (6) of Lemma 4.2 we have

$$\beta_1(t) \in F_{\nabla b}(t, \hat{X}_t, \hat{u}_t) \quad dt - a.e.$$

Applying a similar proof for $\beta_2^j(t) \in F_{\nabla \sigma^j}(t, \hat{X}_t)$. This completes the proof. \square

Remark 5.9. Using the same method of proof, we obtain a more general non-smoothness result for the stochastic maximum principle without differentiability assumptions in which the control enters both the drift and the diffusion coefficients where the set of controls is necessarily convex.

Acknowledgements. The author thanks the anonymous referee who offered many useful remarks and suggestions that improved the first version of this manuscript. The author would like to thank Prof. Julio G. Dix, Texas State University - San Marcos, and Prof. Y. Ouknine, Marrakech University for their valuable remarks and kind help.

REFERENCES

- [1] J. P. Aubin, A. Cellina; *Differential inclusions*. Grundlehren Mathematischen Wissenschaften, Volume 264, Springer-Verlag, Berlin (1984).
- [2] V. I. Arkin, M. T. Saksonov; *Necessary optimality conditions for stochastic differential equations*, Soviet. Math. Dokl. 20 (1979), pp. 1-5.
- [3] K. Bahlali, B. Mezerdi and Y. Ouknine; *The maximum principle for optimal control of diffusion with non-smooth coefficients*. Stochastics. Vol. 57, (1996), pp. 303-316.
- [4] K. Bahlali, B. Djehiche, B. Mezerdi; *On the stochastic maximum principle in optimal control of degenerate diffusions with Lipschitz coefficients*, Appl. Math. and Optim., Vol. 56 (2007), pp. 364-378.
- [5] A. Bensoussan; *Lectures on stochastic contr.* In Lect. Notes in Math. 972, Springer-Varlag (1983), pp. 1-62.
- [6] J. M. Bismut; *An introductory approach to duality in Optimal Stochastic Control*, SAIM, Rev., Vol. 20, no. 1, Jan. (1978).

- [7] V. Borkar; *Controlled diffusion processes. Probabilite surveys*. Vol. 2 (2005), pp. 213-244.
- [8] R. Buckdahn, Y. Ouknine, M. Quincampoix; *On limiting values of stochastic differential equations with small noise intensity tending to zero*. Bull. Sci. Math. 133 (2009), pp. 229-237.
- [9] I. Ekeland; *On the variational principle*, J. Math. Anal. Appl. 47 (1974), pp. 443-474.
- [10] A. F. Filippov; *Differential equations with discontinuous right-hand sides*. Mathematics and Its applications: Soviet Series, 18. Dordrecht.etc., Kluwer Academic Publishers. (1988).
- [11] H. Frankowska; *The first ordre necessary conditions for optimality for nonsmooth variational and control problems*, SAIM J. Control optimal., Vol. 22, no. 1 (1984), pp. 1-12.
- [12] U. G. Haussmann; *A stochastic maximum principle for optimal of diffusions*, Pitman Research Notes in Math. Series, 151.(1986).
- [13] J. Jacod and J. Memin; *Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité*. Seminar on probability XV, Lecture Notes in Math. 850 (1981), pp. 529-546. Springer-Verlag, Berlin.
- [14] H. J. Kushner; *Necessary conditions for continuous parameter stochastic optimization problems*, SAIM J. Control Optimal., Vol 10 (1972), pp. 550-565.
- [15] B. Mezerdi; *Necessary conditions for optimality for a diffusion with a non-smooth drift*. Stochastics, Vol. 24 (1988), pp. 305-326.
- [16] S. Peng; *A general stochastic maximum principle for optimal control problems*. SIAM J. Contr. Optim. 28, no. 4 (1990), pp. 966-979.
- [17] L. S. Pontryagin, V. G. Boltanski, R. V. Gamkrelidze; *The mathematical theory of optimal processes*. Interscience New York, (1962)
- [18] B. E. Paden, S. S. Sastry; *A calculus for Computing Filippov's Differential Inclusion with Application to the Variable Structure Control of Robot Manipulators*. IEEE (1987), Vol. 34, No. 1.
- [19] J. Warga; *Necessary conditions without differentiability assumptions in optimal control*, jour. of Diff. Equations 18 (1975), pp. 41-62.
- [20] J. Yong and X. Y. Zhou; *Stochastic Controls. Hamiltonian Systems and HJB Equations*. Springer-Verlag. New York, (1999).

MOKHTAR HAFAYED

LABORATORY OF APPLIED MATHEMATICS, UNIVERSITY OF MED-KHIDER, PO BOX 145, BISKRA (7000), ALGERIA

E-mail address: hafa.mokh@yahoo.com