HOKKAIDO UNIVERSITY

Title	Duality in Stochastic Optimal Control and Applications
Author(s)	Mikami, Toshio; Thieullen, Michèle
Citation	, 1428, 106-119
Issue Date	2005-04
Doc URL	http://hdl.handle.net/2115/5884
Type	article (author version)
File Information	mikami1428.pdf

Instructions for use

Duality in Stochastic Optimal Control and Applications

Toshio Mikami[*] Michèle Thieullen[†]
Hokkaido University Université Paris VI

October 29, 2004

Abstract

We review a duality result and its applications for a stochastic control problem with fixed marginals obtained in [10]. This problem is the stochastic analog of the well known Monge and Monge-Kantorovich optimal transportation problems.

---

[*]Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan; mikami@math.sci.hokudai.ac.jp; phone no. 81/11/706/3444; fax no. 81/11/727/3705; Partially supported by the Grant-in-Aid for Scientific Research, No. 15340047, 15340051 and 16654031, JSPS.

[†]Corresponding author, Laboratoire de Probabilités et Modèles Aléatoires, Boite 188, Université Paris VI, 75252 Paris, France

1

1 Introduction.

In the present paper we review a duality result and its applications for a stochastic control problem with fixed marginals published in [10]. For a few proofs we do not give all details, rather we prefered to focus on the arguments; details for these proofs can be found in [10].

The problem were are interested in is defined as follows: given $\epsilon > 0$,

$$V_\epsilon(P_0, P_1) := \inf\left\{E\left[\int_0^1 L(t, X(t); \beta_X(t, X))dt\right]\right|$$
$$PX(t)^{-1} = P_t(t = 0, 1), X \in \mathcal{A}\right\}. \qquad (1.1)$$

where P_0 and P_1 are Borel probability measures on $\mathbf{R}^d$ and $L(t, x; u) : [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d \mapsto [0, \infty)$ is measurable and convex w.r.t. u. The infimum is taken over the set $\mathcal{A}$ of all $\mathbf{R}^d$-valued, continuous semimartingales $\{X(t)\}_{0 \le t \le 1}$ on a probability space $(\Omega_X, \mathbf{B}_X, P_X)$ such that there exists a Borel measurable $\beta_X : [0, 1] \times C([0, 1]) \mapsto \mathbf{R}^d$ for which
(i) $\omega \mapsto \beta_X(t, \omega)$ is $\mathcal{B}(C([0, t]))_+$-measurable for all $t \in [0, 1]$, where $\mathcal{B}(C([0, t]))$ denotes the Borel σ-field of $C([0, t])$,
(ii) $\{X(t) - X(0) - \int_0^t \beta_X(s, X)ds := \sqrt{\epsilon}W_X(t)\}_{0 \le t \le 1}$ where W_X is a $\sigma[X(s) : 0 \le s \le t]$-Brownian motion (see [7]).

Remark It would appear more natural to consider semi martingales of the form
$$X^u(t) = X_o + \int_0^t u(s)ds + W(t) \quad (t \in [0, 1]). \qquad (1.2)$$

with $\{u(t)\}_{0 \le t \le 1}$ a $(\mathbf{B}_t)$-progressively measurable stochastic process. However, if we set
$$\beta_{X^u}(t, X^u) = E[u(t)|X^u(s), 0 \le s \le t], \qquad (1.3)$$

then using conditional expectations Jensen inequality and convexity of L one obtains,

$$E\left[\int_0^1 L(t, X^u(t); u(t))dt\right] \ge E\left[\int_0^1 L(t, X^u(t); \beta_{X^u}(t, X^u))dt\right]. \qquad (1.4)$$

and therefore it is sufficient to consider drifts of the form β_X as long as one is interested in the minimizing problem $V_\epsilon(P_0, P_1)$.

2

When L depends only on u, problem V_ϵ has a counterpart in the deterministic setting. this counterpart has been intensively studied since it is the Monge-Kantorovich problem (for a complete list of references we refer the reader to [11] and [13])

$$\mathcal{T}(P_0, P_1) \quad := \quad \inf\Big\{ E\Big[\int_0^1 \ell\Big(\frac{d\phi(t)}{dt} \Big) dt \Big] \Big| P\phi(t)^{-1} = P_t (t = 0, 1),$$
$$t \mapsto \phi(t) \text{ is absolutely continuous}\Big\}. \qquad (1.5)$$

Actually the most usual (and better known) form of the Monge-Kantorovich problem is

$$T(P_0, P_1) := \inf\Big\{ E(L(Y - X)); X \sim P_0, Y \sim P_1 \Big\} \qquad (1.6)$$

where $X \sim P_0$ (resp. $Y \sim P_1$) means that the law of X (resp. Y) is P_0 (resp. P_1). It is not difficult to show that $T(P_0, P_1) = \mathcal{T}(P_0, P_1)$. In the quadratic case, that is when $L(t, x, u) = \frac{1}{2}|u|^2$, the Monge-Kantorovich problem has received much attention, in probability as well as in statistics, in particular because $\sqrt{T(P_0, P_1)}$, called Wasserstein metric, metrizes convergence in distribution on the set of probability measures on $\mathbf{R}^d$ with finite second moments. It is not difficult to show that $T(P_0, P_1) = \mathcal{T}(P_0, P_1)$. More recently the results obtained by Brenier (cf. [1], [2]) have revived the subject by enlightening its connection with fluid mechanics and geometry.

Duality results play a fundamental role in the study of Monge-Kantorovich problem. There are two duality results. For the sequel the most important for us is the duality result due to Evans ([5]):

$$T(P_0, P_1) = \sup\Big\{ \int_{\mathbf{R}^d} \psi(1, x) P_1(dx) - \int_{\mathbf{R}^d} \psi(0, x) P_0(dx) \Big\}, \qquad (1.7)$$

where the supremum is taken over all continuous viscosity solutions ψ to the following Hamilton-Jacobi equation:

$$\frac{\partial \psi(t, x)}{\partial t} + \ell^*(D_x \psi(t, x)) = 0 \quad ((t, x) \in (0, 1) \times \mathbf{R}^d) \qquad (1.8)$$

(see E Chap. 3). Here $D_x := (\partial/\partial x_i)_{i=1}^d$ and for $z \in \mathbf{R}^d$,

3

$$\ell^*(z) := \sup_{u \in \mathbf{R}^d} \{ <z, u> -\ell(u) \}$$

and $< \cdot, \cdot >$ denotes the inner product in $\mathbf{R}^d$.

The second duality result was chronologically proved before by Kantorovich and implies (1.7) (cf. for instance V):

$$T(P_0, P_1) := \sup \Big\{ \int_{\mathbf{R}^d} \psi(y) P_1(dy) + \int_{\mathbf{R}^d} \varphi(x) P_0(dx); \\ (\varphi, \psi) \in L^1(P_0) \times L^1(P_1), \varphi(x) + \psi(y) \le L(y - x) \Big\}. (1.9)$$

In the sequel we describet how it is possible to prove a duality theorem for V_ϵ in the spirit of (1.7) and describe applications. We will not give all proofs in detail; for detailed proofs we refer the reader to [10].

2 Duality Theorem

For simplicity in what follows we restrict to the case when $L(t, x, u) = L(u)$ (that is L depends only on u). However our main result (duality theorem) and its applications are valid even if L depends on (t, x) (cf. [10]). Let us recall that P_0 and P_1 are given Borel probability measures on $\mathbf{R}^d$, and $L(u) : \mathbf{R}^d \mapsto [0, \infty)$ is a measurable and convex function of u. We moreover assume that

$$V_\epsilon(P_0, P_1) < +\infty \qquad (2.1)$$

We will need assumptions on L which we denote as follows:

(A.1). L is *superlinear*: for some $\delta > 1$,

$$\liminf_{|u| \to \infty} \frac{L(u)}{|u|^\delta} > 0.$$

(A.2). (i)$L \in C^3(\mathbf{R}^d)$,
(ii) $D_u^2 L(u)$ is positive definite for all $u \in \mathbf{R}^d$,

We will look for sufficient conditions for V_ϵ to admit a minimizer, unique and/or Markovian and also for a characterization of minimizers. A duality theorem will provide such a characterization(the characterization itself will be obtained in the next section). As already mentioned we focus on the main steps and articulations of the argument.

2.1 Existence and uniqueness of a minimizer.

Results about existence and uniqueness are gathered in

Theorem 2.1 *(i) $V_\epsilon(P_0, P_1)$ admits a minimizer.*
(ii) If assumpion (A.1) holds with $\delta = 2$, $V_\epsilon(P_0, P_1)$ admits a Markovian minimizer
(iii) If L is strictly convex and assumpion (A.1) holds with $\delta = 2$, then $V_\epsilon(P_0, P_1)$ admits a unique minimizer (which is Markovian from (ii)).

Our tool for the proof of (ii) and (iii) in Theorem 2.1 is the following minimization problem with fixed marginals

$$\underline{V}_\epsilon(P_0, P_1) := \inf \int_0^1 \int_{\mathbf{R}^d} L(b(t, x)) P(t, dx) dt, \tag{2.2}$$

where the infimum is taken over all $(b(t, x), P(t, dx))$ for which $P(t, dx)$ ($0 \leq t \leq 1$) are Borel probability measures, on $\mathbf{R}^d$, such that $p(t, x) := P(t, dx)/dx$ exists for all $t \in (0, 1]$, $P(t, dx) = P_t$ ($t = 0, 1$) and the following Fokker-Planck pde

$$\frac{\partial P(t, dx)}{\partial t} = \frac{\epsilon}{2} \triangle P(t, dx) - div(b(t, x) P(t, dx)) \tag{2.3}$$

is satisfied. Let us notice that $\underline{V}_\epsilon$ is a stochastic analog of the problem onsidered by Benamou and Brenier in [3]. Then

Proposition 2.1 *(cf. [10] Lemma 3.5). Assume (A.1) with $\delta = 2$ holds. Then $V_\epsilon(P_0, P_1) = \underline{V}_\epsilon(P_0, P_1)$.*

Proof of Theorem 2.1. Proof of (i): Let (X_n) denote a minimizing sequence of processes in the set $\mathcal{A}$; this means that

$$\lim_{n \to \infty} E\left[\int_0^1 L(\beta_{X_n}(t, X_n)) dt\right] = V_\epsilon(P_0, P_1) \tag{2.4}$$

Since $X_n \in \mathcal{A}$ for all n and assumption (A.1) holds (L is superlinear), it follows that the sequence (X_n) is tight: the sufficient condition for tightness of [14] is satisfied. In particular (A.1) implies that

$$E\left[\int_0^1 |\beta_{X_n}(t, X_n))|^\delta dt\right] < +\infty \tag{2.5}$$

5

(with $\delta > 1$). Hence there exists a subsequence (X_{n_k}) weach converges weakly; let us denote its limit by $(X(t))$. The process X belongs to $\mathcal{A}$: from [14], Theorem 5, we obtain that $\frac{1}{\sqrt{\epsilon}}\{X(t) - X(0) - A(t)\}_{t\in[0,1]}$ is a standard Brownian motion and $\{A(t)\}_{t\in[0,1]}$ is absolutely continuous. Moreover $(X(t))$ satisfies

$$\lim_{k\to\infty} E\Big[\int_0^1 L(\beta_{X_{n_k}}(t, X_{n_k}))dt\Big] \qquad (2.6)$$
$$\geq \quad E\Big[\int_0^1 L\Big(\frac{dA(t)}{dt}\Big)dt\Big].$$

which implies that it is a minimizer of V_ϵ. Inequality (2.6) may be proved following the argument of [9]in the proof of Theorem 1, which is here simplified since L depends on u only.

Proof of (ii): we now assume that (A.1) holds with $\delta = 2$. Using the same argument as in the proof of (i) one can show that $\underline{V}_\epsilon(P_0, P_1)$ admits a minimizer. From Proposition 2.1 this minimizer also is a minimizer of V_ϵ (here it is actually sufficient that $V_\epsilon \geq \underline{V}_\epsilon$).

Proof of (ii): we moreover assume that L is strictly convex. From Proposition (actually it is sufficient that $V_\epsilon \leq \underline{V}_\epsilon$) it is enough to show uniqueness for $\underline{V}_\epsilon$ (cf. [10] proof of Proposition 2.2 where we use the strict convexity of L and the linearity of Fokker-Planck pde). Q.E.D.

2.2   Duality Theorem.

Theorem 2.2 *Suppose that (A.1) and (A.2) are satisfied. Then*

$$V_\epsilon(P_0, P_1) = \sup\Big\{\int_{\mathbf{R}^d} \varphi(1, y)P_1(dy) - \int_{\mathbf{R}^d} \varphi(0, x)P_0(dx)\Big\}, \qquad (2.7)$$

where the supremum is taken over all classical solutions φ, to the following HJB equation, for which $\varphi(1, \cdot) \in C_b^\infty(\mathbf{R}^d)$:

$$\frac{\partial \varphi(t, x)}{\partial t} + \frac{\epsilon}{2}\triangle\varphi(t, x) + H(D_x\varphi(t, x)) = 0 \quad ((t, x) \in (0, 1) \times \mathbf{R}^d) \qquad (2.8)$$

Proof of 2.2 The two main arguments of the proof are:
1. A property of the Legendre transform: on a Banach space if f is a lower semi continuous function not identically equal to $+\infty$, then $f^{**} = f$ where $*$ denotes Legendre transform.

2. A representation of the value function of a stochastic control problem (with sufficiently regular terminal cost) by a solution of an Hamilton-Jacobi-Bellman pde.

For point 1., we rely on results of [4] (namely Theorem 2.2.15 and Lemma 3.2.3). To apply these results, one has to prove first that $P \mapsto V(P_0, P)$ is lower semicontinuous and convex. This is proved in detail in [10] Lemmas 3.1 and 3.2. It follows that

$$V(P_0, P_1) = \sup_{f \in C_b(\mathbf{R}^d)} \left\{ \int_{\mathbf{R}^d} f(x) P_1(dx) - V_{P_0}^*(f) \right\}, \qquad (2.9)$$

where for $f \in C_b(\mathbf{R}^d)$,

$$V_{P_0}^*(f) := \sup_{P \in \mathcal{M}_1(\mathbf{R}^d)} \left\{ \int_{\mathbf{R}^d} f(x) P(dx) - V(P_0, P) \right\},$$

and $\mathcal{M}_1(\mathbf{R}^d)$ denotes the complete separable metric space, with a weak topology, of Borel probability measures on $\mathbf{R}^d$.

For point 2., we refer the reader to [6]: for $f \in C_b^\infty(\mathbf{R}^d)$,

$$V_{P_0}^*(f) = \sup\left\{ E[f(X(1))] - E\left[\int_0^1 L(t, X(t); \beta_X(t, X))dt \right] : \right.$$
$$\left. X \in \mathcal{A}, PX(0)^{-1} = P_0 \right\}$$
$$= \int_{\mathbf{R}^d} \varphi_f(0, x) P_0(dx), \qquad (2.10)$$

where φ_f denotes the unique classical solution to the HJB equation (2.3) with $\varphi(1, \cdot) = f(\cdot)$. Using both identities (2.9) and (2.10), we obtain

$$V_\epsilon(P_0, P_1) \geq \sup_{f \in C_b^\infty(\mathbf{R}^d)} \int_{\mathbf{R}^d} \varphi(1, y) P_1(dy) - \int_{\mathbf{R}^d} \varphi(0, x) P_0(dx), \qquad (2.11)$$

To prove the converse inequality we have to pass from $C_b(\mathbf{R}^d)$ to $C_b^\infty(\mathbf{R}^d)$ with the help of a mollifier sequence. Take $\Phi \in C_o^\infty([-1, 1]^d; [0, \infty))$ for which $\int_{\mathbf{R}^d} \Phi(x) dx = 1$, and for $\delta > 0$, and define

$$\Phi_\delta(x) := \delta^{-d} \Phi(x/\delta).$$

For $f \in C_b(\mathbf{R}^d)$, we set

7

$$f_\delta(x) := \int_{\mathbf{R}^d} f(y) \Phi_\delta(x - y) dy. \tag{2.12}$$

Then $f_\delta \in C_b^\infty(\mathbf{R}^d)$ and

$$\sup_{f \in C_b^\infty(\mathbf{R}^d)} \int_{\mathbf{R}^d} \varphi(1, y) P_1(dy) - \int_{\mathbf{R}^d} \varphi(0, x) P_0(dx)$$

$$\geq \int_{\mathbf{R}^d} f_\delta(x) P_1(dx) - V_{P_0}^*(f_\delta)$$

$$\geq \int_{\mathbf{R}^d} f(x) \Phi_\delta * P_1(dx) - V_{\Phi_\delta * P_0})^*(f).$$

Indeed, for any $X \in \mathcal{A}$

$$E[f_\delta(X(1))] = \int_{\mathbf{R}^d} \Phi(z) dz E[f(X(1) - \delta z)] \tag{2.13}$$

Then identity (2.9) implies that

$$\sup_{f \in C_b^\infty(\mathbf{R}^d)} \int_{\mathbf{R}^d} \varphi(1, y) P_1(dy) - \int_{\mathbf{R}^d} \varphi(0, x) P_0(dx)$$

$$\geq V(\Phi_\delta * P_0, \Phi_\delta * P_1)$$

It remains to let δ go to 0 and use the lower semi-continuity of $(P, Q) \mapsto V(P, Q)$ proved in [10]. Q.E.D.

3 Applications.

3.1 Characterization.

We first recall the following property of Legendre transform which we will use repeatedly: if L is strictly convex, superlinear (i.e. satisfies (A.1)) and smooth (for instance belongs to $C^2(\mathbf{R}^d)$) then $L^{**} = L$; $\nabla L : \mathbf{R}^d \to \mathbf{R}^d$ is a bijection from $\mathbf{R}^d$ onto itself and $\nabla H = \nabla L^{-1}$ where $H = L^*$. If moreover $D^2 L$ is positive definite, H is twice differentiable and

$$D^2 H(\nabla L(u)) = D^2 L(u)^{-1} \tag{3.1}$$

Theorem 3.1 *Suppose that (A.1) and (A.2) hold. Then for any minimizer $\{X(t)\}_{0 \leq t \leq 1}$ of $V_\epsilon(P_0, P_1)$, there exists a sequence of classical solutions $\{\varphi_n\}_{n \geq 1}$ to the HJB equation (2.8), such that $\varphi_n(1, \cdot) \in C_b^\infty(\mathbf{R}^d)$ ($n \geq 1$) and that the following holds:*

$$\beta_X(t, X) = b_X(t, X(t)) := E[\beta_X(t, X)|(t, X(t))] \tag{3.2}$$
$$= \lim_{n \to \infty} D_z H(t, X(t); D_x \varphi_n(t, X(t))) \quad dt dP X(\cdot)^{-1} - a.e..$$

Proof of Theorem 3.1 From Theorem 2.2 here exists a sequence of classical solutions $\{\varphi_n\}_{n \geq 1}$ to the HJB equation (2.8), such that $\varphi_n(1, \cdot) \in C_b^\infty(\mathbf{R}^d)$ ($n \geq 1$) and

$$\lim_{n \to \infty} \int_{\mathbf{R}^d} \varphi_n(1, y) P_1(dy) - \int_{\mathbf{R}^d} \varphi_n(0, x) P_0(dx) = V_\epsilon(P_0, P_1) \tag{3.3}$$

Therefore, for X a minimizer of V_ϵ, it holds

$$\lim_{n \to \infty} \int_{\mathbf{R}^d} \varphi_n(1, y) P_1(dy) - \int_{\mathbf{R}^d} \varphi_n(0, x) P_0(dx) = E\left[\int_0^1 L(\beta_X(t, X)) dt\right] \tag{3.4}$$

Since $X(0) \sim P_0$ (resp. $X(1) \sim P_1$) and $\{\varphi_n\}_{n \geq 1}$ solves the HJB pde (2.8), Ito formula yields

$$\lim_{n \to \infty} E \int_0^1 < \beta_X(t, X), \nabla \varphi_n(t, X(t)) > -L(\beta_X(t, X)) - H(\nabla \varphi_n(t, X(t)) dt = 0 \tag{3.5}$$

Moreover by definition of H as the Legendre transform of L, the integrand in (3.5) is positive. Hence the sequence

$$(< \beta_X(t, X), \nabla \varphi_n(t, X(t)) > -L(\beta_X(t, X)) - H(\nabla \varphi_n(t, X(t)))) \tag{3.6}$$

converges to 0 in $L^1(dt dP)$ and admits a subsequence which converges a.s. For simplicity we still denote this subsequence by (φ_n). Let (t, ω) be such that the sequence $(< \beta_X(t, X), \nabla \varphi_n(t, X(t)) > -H(\nabla \varphi_n(t, X(t)))$ converges to $L(\beta_X) = H^*(\beta_X)$. The supremum in the definition of

$$H^*(u) = \sup(< p, u > -H(p)) \tag{3.7}$$

is attained at $p^* = \nabla L(u)$. We therefore obtain that

$$\lim \nabla \varphi_n(t, X(t)) = \nabla L(\beta_X(t, X)) \tag{3.8}$$

9

or equivalently $\beta_X(t, X) = \lim \nabla H(\nabla \varphi_n(t, X(t)))$. Q.E.D.

We would like to show now that a minimizer solves a stochastic equation. We were able to prove such a result under the additional assumption:
(A.3). $D^2 L(u)$ is bounded.

The following lemma will be useful below:

Lemma 3.1 *Let $L \in C^2(\mathbf{R}^d)$ be strictly convex and superlinear such that*

$$C := \sup\{< D^2 L(u)z, z >: (u, z) \in \mathbf{R}^d \times \mathbf{R}^d, |z| = 1\} < +\infty \qquad (3.9)$$

Then

$$\forall (u, z) \in \mathbf{R}^d \times \mathbf{R}^d \quad ||z - \nabla L(u)||^2 \le C|L(u) - (< u, z > -H(z))| \quad (3.10)$$

Proof of Lemma 3.1. By definition of $H = L^*$, for all (u, z), $L(u) - (< u, z > -H(z)) \ge 0$. The assumptions of the lemma ensure that for all u, $u = \nabla H(\nabla L(u))$ and $H(p) =< p, \nabla H(p) > -L(\nabla H(p))$ for all p. We therefore have

$$L(u) - (< u, z > -H(z)) = H(z) - H(\nabla L(u)) - < \nabla H(\nabla L(u)), z - \nabla L(u) > \tag{3.11}$$

The conclusion follows from identity (3.1). Q.E.D.

Theorem 3.2 *Suppose that (A.1) holds with $\delta = 2$ as well as (A.2) and (A.3). Then for the unique minimizer $\{X(t)\}_{0 \le t \le 1}$ of $V_\epsilon(P_0, P_1)$,*
(1) there exist $f(\cdot) \in L^1(\mathbf{R}^d, P_1(dx))$ and a $\sigma[X(s) : 0 \le s \le t]$- continuous semimartingale $\{Y(t)\}_{0 \le t \le 1}$ such that

$$\{(X(t), Y(t), Z(t) := D_u L(b_X(t, X(t))))\}_{0 \le t \le 1}$$

satisfies the following FBSDE in a weak sense: for $t \in [0, 1]$,

$$X(t) = X(0) + \int_0^t D_z H(Z(s))ds + \sqrt{\epsilon}W(t), \qquad (3.12)$$

$$Y(t) = f(X(1)) - \int_t^1 L(D_z H(Z(s)))ds$$
$$- \int_t^1 < Z(s), dW(s) > .$$

(2) there exist $f_0(\cdot) \in L^1(\mathbf{R}^d, P_0(dx))$ and $\varphi(\cdot, \cdot) \in L^1([0,1] \times \mathbf{R}^d, P((t, X(t)) \in dtdx))$ such that $Y(0) = f_0(X(0))$ and such that

$$Y(t) - Y(0) = \varphi(t, X(t)) - \varphi(0, X(0)) \quad dtdPX(\cdot)^{-1} - a.e., \qquad (3.13)$$

that is, $Y(t)$ is a continuous version of $\varphi(t, X(t)) - \varphi(0, X(0)) + f_0(X(0))$.

Proof of Theorem 3.2 Let (φ_n) be a sequence satisfying the same conditions as in the proof of Theorem 3.1 and X a minimizer of V_ϵ. From Ito formula,

$$\varphi_n(t, X(t)) - \varphi_n(0, X(0)) \qquad (3.14)$$
$$= \int_0^t \{< b_X(s, X(s)), D_x\varphi_n(s, X(s)) > -H(D_x\varphi_n(s, X(s)))\} ds$$
$$+ \int_0^t < D_x\varphi_n(s, X(s)), \sqrt{\epsilon} dW(s) > .$$

We first consider convergence of the martingale part. By Doob's inequality

$$E(\sup_{0 \le t \le 1} \left| \int_0^t < D_x\varphi_n(s, X(s)) - D_u L(b_X(s, X(s))), dW(s) > \right|^2)$$
$$\le 4E(\int_0^1 |D_x\varphi_n(s, X(s)) - D_u L(b_X(s, X(s)))|^2 ds) \qquad (3.15)$$

By Lemma 3.1 it follows that

$$E(\sup_{0 \le t \le 1} \left| \int_0^t < D_x\varphi_n(s, X(s)) - D_u L(b_X(s, X(s))), dW(s) > \right|^2)$$
$$\le 4CE(\int_0^1 |L(b_X(s, X(s))) - (< b_X(s, X(s)), D_x\varphi_n(s, X(s)) >$$
$$- H(D_x\varphi_n(s, X(s))))| ds)$$

which converges to 0 by Theorem 3.1. This theorem also implies that

$$\int_0^t \{< b_X(s, X(s)), D_x\varphi_n(s, X(s)) > -H(D_x\varphi_n(s, X(s)))\} ds \qquad (3.16)$$

converges in L^1 to $\int_0^1 L(b_X(s, X(s))) ds$. We therefore obtain that $\varphi_n(1, y) - \varphi_n(0, x)$ and $\varphi_n(t, y) - \varphi_n(0, x)$ are convergent in $L^1(\mathbf{R}^d \times \mathbf{R}^d, P((X(0), X(1)) \in$

11

$dxdy))$ and $L^1(\mathbf{R}^d \times [0,1] \times \mathbf{R}^d, P((X(0),(t,X(t))) \in dxdtdy))$, respectively. The question is whether the limit is still of the separable form $\psi(1,y) - \psi(0,x)$ and $\psi(t,y) - \psi(0,x)$ respectively. From [12] this is indeed the case provided that the law of $(X(0),X(1))$ (resp. $(X(0),X(t))$) is absolutely continuous with respect to $P_0(dx)P_1(dy)$ (resp. $P_0(dx)P_t(dy)$) where P_t denotes the law of X_t. These conditions are satisfied here since (A.1) holds with $\delta = 2$ and consequently the process X has finite entropy w.r.t. the Wiener measure on $C(\mathbf{R}^d)$ with initial law P_0. Hence, from [12], Prop. 2, there exist $f \in L^1(\mathbf{R}^d, P_1(dx))$, $f_0 \in L^1(\mathbf{R}^d, P_0(dx))$, $\varphi_0 \in L^1(\mathbf{R}^d, P_0(dx))$ and $\varphi \in L^1([0,1] \times \mathbf{R}^d, P((t,X(t)) \in dtdy))$ such that

$$\lim_{n \to \infty} E[|\varphi_n(1,X(1)) - \varphi_n(0,X(0)) - \{f(X(1)) - f_0(X(0))\}|] = 0, \quad (3.17)$$

and

$$\lim_{n \to \infty} E\left[\int_0^1 |\varphi_n(t,X(t)) - \varphi_n(0,X(0)) - \{\varphi(t,X(t)) - \varphi_0(X(0))\}|dt\right] = 0. \tag{3.18}$$

It is easy to check that $(Y(t))$ defined by

$$
\begin{aligned}
Y(t) \quad := \quad & f_0(X(0)) + \int_0^t L(s,X(s);b_X(s,X(s)))ds && (3.19)\\
& + \int_0^t <D_u L(s,X(s);b_X(s,X(s))), dW(s)> .
\end{aligned}
$$

satisfies the statement of Theorem 3.2. Q.E.D.

References

[1] Brenier, Y., *Décomposition polaire et réarrangement monotone des champs de vecteurs.* C. R. Acad. Sci. Paris, Série I, 305, pp. 805- 808 (1987)

[2] Brenier, Y., *Extended Monge-Kantorovich theory* Proc. of the CIME course 2001 " Optimal Transportation and Applications", Martina France (Italy) Lecture Notes in Mathematics n. 1813, p. 91-121, (2003) Springer, Berlin.

[3] Brenier, Y., Benamou, J.D., A numerical method for the optimal mass transport problem and related problems, in: Caffarelli, L.A., Milman, M. (Eds.), Monge Ampre equation: applications to geometry and optimization, Proc. of the NSF-CBMS Conference, Deerfield Beach, FL 1997, Contemporary Mathematics 226, Amer. Math. Soc., Providence, RI, pp. 1-11 (1999).

[4] Deuschel, J.D., Stroock, D.W., . Large deviations, Pure and Applied Mathematics Vol. 137, Academic Press,Inc., Boston, MA. 1989

[5] Evans, L.C., 1998. Partial differential equations, Graduate Studies in Mathematics Vol. 19, Amer. Math. Soc., Providence, RI, USA. 1998.

[6] Fleming, W.H., Soner, H.M., . Controlled Markov Processes and Viscosity Solutions, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo. 1993

[7] Liptser, R.S., Shiryaev, A.N., Statistics of random processes I, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo (1977)

[8] Mikami, T., *Variational processes from the weak forward equation.* Commun. Math. Phys. 135, 19-40 (1990)

[9] Mikami, T *Optimal control for absolutely continuous stochastic processes and the mass transportation problem.* Elect. Comm. in Probab. 7, 199-213 (2002)

[10] Mikami, T., Thieullen, M., *Duality Theorem for Stochastic Optimal Control Problem.* Hokkaido University Preprint Series in Mathematics , n. 652, June 2004. Submitted to Stoch. Proc. Appl.

[11] Rachev, S.T., Rüschendorf, L., Mass transportation problems, Vol. I: Theory, Vol. II: Application, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo. 1998.

[12] Rüschendorf, L., Thomsen, W., *Note on the Schrödinger equation and I-projections.* Statist. Probab. Lett. 17, 369-375 (1993).

[13] Villani, C., Topics in Optimal Transportation, Graduate Studies in Mathematics Vol. 58, Amer. Math. Soc., Providence, RI. 2003

[14] Zheng, W.A., *Tightness results for laws of diffusion processes application to stochastic mechanics.* Ann. Inst. Henri Poincaré 21, pp. 103-124 (1985)

14