Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

Stochastic Processes and their Applications 116 (2006) 1815–1835

stochastic processes and their applications

www.elsevier.com/locate/spa

Duality theorem for the stochastic optimal control problem[☆]

Toshio Mikami[a,*], Michèle Thieullen[b]

[a] *Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan*
[b] *Laboratoire de Probabilités et Modèles Aléatoires, Université de Paris VI, 75252 Paris, France*

Received 25 May 2004; received in revised form 15 January 2006; accepted 27 April 2006
Available online 22 May 2006

Abstract

We prove a duality theorem for the stochastic optimal control problem with a convex cost function and show that the minimizer satisfies a class of forward–backward stochastic differential equations. As an application, we give an approach, from the duality theorem, to h-path processes for diffusion processes.

Keywords: Duality theorem; Stochastic control; Forward–backward stochastic differential equation

1. Introduction

Let $c(x, y) : \mathbf{R}^d \times \mathbf{R}^d \mapsto [0, \infty)$ be measurable and let P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d) :=$ the complete separable metric space, with a weak topology, of Borel probability measures on $\mathbf{R}^d$. The study of a minimizer of the following $\mathcal{T}(P_0, P_1)$ is called the Monge–Kantorovich problem (or the optimal mass transportation problem) which has been studied by many authors and which has been applied in many fields (see [1,2,6,10,12,15,30,33] and the references therein):

$$\mathcal{T}(P_0, P_1) := \inf\{E[c(\phi(0), \phi(1))] | P\phi(t)^{-1} = P_t (t = 0, 1)\}. \tag{1.1}$$

(When it is not confusing, we use the same notation P for different probability measures.)

Let $L(t, x; u) : [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d \mapsto [0, \infty)$ be continuous and convex in u. Consider the case where for $(x_0, x_1) \in \mathbf{R}^d \times \mathbf{R}^d$,

[☆] Part of this work was done while one of the authors (TM) was visiting Université de Paris VI. He would like to thank them for their hospitality.

[*] Corresponding author. Tel.: +81 11 706 3444; fax: +81 11 727 3705.
E-mail address: mikami@math.sci.hokudai.ac.jp (T. Mikami).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2006.04.014

$$c(x_0, x_1) = \inf\left\{ \int_0^1 L\left(t, \phi(t); \frac{\mathrm{d}\phi(t)}{\mathrm{d}t}\right) \mathrm{d}t \,\middle|\, \phi(t) = x_t (t = 0, 1),\right.$$
$$\left. t \mapsto \phi(t) \text{ is absolutely continuous} \right\}, \tag{1.2}$$

which is a standard variational problem in classical mechanics. Since for any $(x_0, x_1) \in \mathbf{R}^d \times \mathbf{R}^d$,

$$\mathcal{T}(\delta_{x_0}, \delta_{x_1}) = c(x_0, x_1),$$

where δ_x denotes the delta measure on x, the Monge–Kantorovich problem can be considered as a generalization of classical mechanics.

Remark 1.1. If $L = \ell(u)$, then $c(x_0, x_1) = \ell(x_1 - x_0)$ and a function $x_0 + t(x_1 - x_0)$ is a minimizer in (1.2). This can be shown using Jensen's inequality (see, e.g., [13, p. 35]).

In the last few years we have been studying the optimal mass transportation problem as stochastic mechanics in the framework of the stochastic optimal control theory. The following is a stochastic optimal control counterpart of (1.1) and (1.2) (see [24–27] and [28]):

$$V(P_0, P_1) := \inf\left\{ E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\,\mathrm{d}t\right] \,\middle|\right.$$
$$\left. PX(t)^{-1} = P_t (t = 0, 1), X \in \mathcal{A} \right\}. \tag{1.3}$$

The set $\mathcal{A}$ will be given a precise definition at the end of this section. For the moment, let us just say that $X \in \mathcal{A}$ implies that $\{W_X(t) := X(t) - X(0) - \int_0^t \beta_X(s, X)\,\mathrm{d}s\}_{0 \le t \le 1}$ is a $\sigma[X(s) : 0 \le s \le t]$-Brownian motion. The meaning of the study of $V(P_0, P_1)$ is this. Suppose that we know the probability distributions of a stochastic system at times $t = 0$ and 1. To study what happened during the time interval $(0, 1)$, we have to consider problems such as (1.3).

The duality theorem for $\mathcal{T}(P_0, P_1)$ plays a crucial role in the Monge–Kantorovich problem (see [2,20,30,26,33]). In this paper we prove the duality theorem for $V(P_0, P_1)$ and obtain the properties of minimizers of $V(P_0, P_1)$.

We explain the duality theorem for $\mathcal{T}(P_0, P_1)$ when $c(x_0, x_1) = \ell(x_1 - x_0)$ (see Remark 1.1). It is said that the duality theorem for $\mathcal{T}(P_0, P_1)$ holds if the following is true:

$$\mathcal{T}(P_0, P_1) = \sup\left\{ \int_{\mathbf{R}^d} f(x) P_1(\mathrm{d}x) - \int_{\mathbf{R}^d} Tf(x) P_0(\mathrm{d}x) \,\middle|\, f \in C_b(\mathbf{R}^d) \right\}, \tag{1.4}$$

where

$$Tf(x) := \sup\{f(y) - \ell(y - x) | y \in \mathbf{R}^d\}.$$

It is easy to see that the following holds (see Remark 1.1):

$$Tf(x) = \sup\left\{ f(\phi(1)) - \int_0^1 \ell\left(\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}\right) \mathrm{d}t \,\middle|\, \phi(0) = x,\right.$$
$$\left. t \mapsto \phi(t) \text{ is absolutely continuous} \right\}. \tag{1.5}$$

Remark 1.2. As far as the applications are concerned, one would like to replace $C_b(\mathbf{R}^d)$ on the r.h.s. of (1.4) by a smaller space so that one can take a maximizing sequence of nice functions. Indeed, [15] proved and used that the maximizer of the r.h.s. of (1.4) is locally Lipschitz continuous. For the readers' convenience, we show, in Appendix, that $C_b(\mathbf{R}^d)$ can be replaced by $C_b^\infty(\mathbf{R}^d)$ on the r.h.s. of (1.4) provided that (1.4) holds.

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

$$\psi(t,x) := \begin{cases} \sup\left\{ f(y) - (1-t)\ell\left(\dfrac{y-x}{1-t}\right) \middle| \, y \in \mathbf{R}^d \right\} & ((t,x) \in [0,1) \times \mathbf{R}^d), \\ f(x) & ((t,x) \in \{1\} \times \mathbf{R}^d). \end{cases}$$

Then $\psi(0,x) = Tf(x)$. Suppose that $\ell(u)/|u| \to \infty$ as $|u| \to \infty$. Then for any bounded, uniformly Lipschitz continuous function f, $\psi(t,x)$ is bounded and uniformly Lipschitz continuous in $[0,1] \times \mathbf{R}^d$ and is a unique continuous viscosity solution of the following:

$$\begin{aligned} &\frac{\partial \psi(t,x)}{\partial t} + \ell^*(D_x \psi(t,x)) = 0 \quad ((t,x) \in [0,1) \times \mathbf{R}^d), \\ &\psi(1,x) = f(x) \quad (x \in \mathbf{R}^d) \end{aligned} \tag{1.6}$$

(see [11, Chapters 3 and 10]), where

$$\ell^*(z) := \sup\{\langle z, u\rangle - \ell(u) | u \in \mathbf{R}^d\},$$

$\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbf{R}^d$ and $D_x := (\partial/\partial x_i)_{i=1}^d$. In particular, from Remark 1.2, if (1.4) holds, then

$$\mathcal{T}(P_0, P_1) = \sup\left\{ \int_{\mathbf{R}^d} \psi(1,x) P_1(\mathrm{d}x) - \int_{\mathbf{R}^d} \psi(0,x) P_0(\mathrm{d}x) \right\}, \tag{1.7}$$

where the supremum is taken over all bounded, uniformly Lipschitz continuous viscosity solutions of (1.6).

We explain the duality theorem for $V(P_0, P_1)$. We say that the duality theorem for $V(P_0, P_1)$ holds if the following is true (see Theorem 2.1):

$$V(P_0, P_1) = \sup\left\{ \int_{\mathbf{R}^d} \varphi(1,x) P_1(\mathrm{d}x) - \int_{\mathbf{R}^d} \varphi(0,x) P_0(\mathrm{d}x) \right\}, \tag{1.8}$$

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

$$\frac{\partial \varphi(t,x)}{\partial t} + \frac{1}{2}\triangle \varphi(t,x) + H(t,x; D_x\varphi(t,x)) = 0 \quad ((t,x) \in [0,1) \times \mathbf{R}^d). \tag{1.9}$$

Here $\triangle := \sum_{i=1}^d \partial^2/\partial x_i^2$ and for $(t,x,z) \in [0,1] \times \mathbf{R}^d \times \mathbf{R}^d$,

$$H(t,x;z) := \sup_{u \in \mathbf{R}^d} \{\langle z, u\rangle - L(t,x;u)\}. \tag{1.10}$$

Remark 1.3. In the applications of the duality theorem for $V(P_0, P_1)$, it is important that functions φ in (1.8) are sufficiently smooth so that one can consider the stochastic differential equation for $\varphi(t, X(t))$ for a minimizer $X(t)$ of $V(P_0, P_1)$. In our setting, for any $\varphi(1,\cdot) \in C_b^\infty(\mathbf{R}^d)$, the HJB equation (1.9) has a unique solution $\varphi \in C^{1,2}([0,1] \times \mathbf{R}^d) \cap C_b^{0,1}([0,1] \times \mathbf{R}^d)$

and φ has a variational expression which is a stochastic optimal control counterpart of (1.5) (see Lemma 3.3).

As is well known, the quadratic case is important and has been studied by many authors (see [34,14], and also [5,32] and the references therein).

Proposition 1.1 (*[25, Lemma 3.4]*). *Suppose that* $L = |u|^2$*, and* P_1 *is absolutely continuous w.r.t. the Lebesgue measure with* $p_1(x) := P_1(x)/\mathrm{d}x$*. Let us also assume that*

$$\int_{\mathbf{R}^d} |x|^2 (P_0(\mathrm{d}x) + P_1(\mathrm{d}x)) + \int_{\mathbf{R}^d} p_1(x) \log p_1(x)\, \mathrm{d}x < \infty.$$

Then $V(P_0, P_1)$ *is finite, there exists a unique minimizer which is an* h*-path process* $\{X_h(t)\}_{0 \le t \le 1}$ *for Brownian motion (see* [9]*) and* (1.8) *holds.*

In [25] we gave a new proof for the existence of a deterministic minimizer of $\mathcal{T}(P_0, P_1)$ when $c(x_0, x_1) = \ell(x_1 - x_0)$, by proving that the zero-noise limit of $\{X_h(t)\}_{0 \le t \le 1}$ exists, is deterministic and is a minimizer of $\mathcal{T}(P_0, P_1)$ (here we say that a stochastic process $\{X(t)\}_{0 \le t \le 1}$ is deterministic if $X(t)$ is a function of t and $X(0)$). The generalization of this result obtained by taking the zero-noise limit of the duality theorem in this paper is given in [28].

Since we fix initial and terminal distributions of the semimartingales under consideration, the known approach is not useful (see [13]). Our proof relies on the Legendre duality of a lower semicontinuous convex function of Borel probability measures on $\mathbf{R}^d$ (see the proof of Theorem 2.1).

When $D_u^2 L(t, x; u)$ and $D_u^2 L(t, x; u)^{-1}$ exist and are bounded, we show, from (1.8), that the minimizer of $V(P_0, P_1)$ satisfies a forward–backward stochastic differential equation (FBSDE for short; see Theorem 2.2). As an application, we give an approach, from the duality theorem, to h-path processes for diffusion processes (see Corollary 2.3).

We also show that the supremum in (1.8) can be taken over all bounded, uniformly Lipschitz continuous viscosity solutions to the HJB equation (1.9) (see Corollary 2.2). For the readers' convenience, we give the definition of the viscosity solution to the HJB equation (1.9).

Definition 1.1 (*Viscosity Solution*). (see e.g. [13]) (Viscosity Subsolution) $\varphi \in USC([0, 1] \times \mathbf{R}^d)$ is a viscosity subsolution of (1.9) if whenever $h \in C^{1,2}([0, 1) \times \mathbf{R}^d)$ and $\varphi - h$ takes its maximum at $(s, y) \in [0, 1) \times \mathbf{R}^d$,

$$\frac{\partial h(s, y)}{\partial s} + \frac{1}{2} \Delta h(s, y) + H(s, y; D_x h(s, y)) \ge 0.$$

(Viscosity Supersolution) $\varphi \in LSC([0, 1] \times \mathbf{R}^d)$ is a viscosity supersolution of (1.9) if whenever $h \in C^{1,2}([0, 1) \times \mathbf{R}^d)$ and $\varphi - h$ takes its minimum at $(s, y) \in [0, 1) \times \mathbf{R}^d$,

$$\frac{\partial h(s, y)}{\partial s} + \frac{1}{2} \Delta h(s, y) + H(s, y; D_x h(s, y)) \le 0.$$

(Viscosity Solution) $\varphi \in C([0, 1] \times \mathbf{R}^d)$ is a viscosity solution of (1.9) if it is both a viscosity subsolution and a viscosity supersolution of (1.9).

The construction of a semimartingale from a solution of the Fokker–Planck equation with the p-th integrable drift vector ($p > 1$) is given in [27] as an application of the duality theorem in this paper. This is a generalization of [23] where $p = 2$.

As the set $\mathcal{A}$ over which the infimum is taken in (1.3), we consider the set of all $\mathbf{R}^d$-valued, continuous semimartingales $\{X(t)\}_{0\le t\le 1}$ on a complete filtered probability space 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) $\{W_X(t) := X(t) - X(0) - \int_0^t \beta_X(s,X)\,\mathrm{d}s\}_{0\le t\le 1}$ is a $\sigma[X(s): 0\le s\le t]$-Brownian motion (see [22]).

We explain why this is appropriate. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, P)$ be a complete filtered probability space, X_0 be a $(\mathcal{F}_0)$-adapted random variable for which $PX_0^{-1} = P_0$, and $\{W(t)\}_{t\ge 0}$ denote a d-dimensional $(\mathcal{F}_t)$-Brownian motion for which $W(0)=0$ (see, e.g., [16] or [22]). For a $\mathbf{R}^d$-valued, $(\mathcal{F}_t)$-progressively measurable stochastic process $\{u(t)\}_{0\le t\le 1}$, put

$$X^u(t) = X_0 + \int_0^t u(s)\,\mathrm{d}s + W(t) \quad (t\in[0,1]). \tag{1.11}$$

If $E[\int_0^1 |u(t)|\,\mathrm{d}t]$ is finite, then $\{X^u(t)\}_{0\le t\le 1}\in\mathcal{A}$ and

$$\beta_{X^u}(t, X^u) = E[u(t)|X^u(s), 0\le s\le t] \tag{1.12}$$

(see [22, p. 270]). Besides this, by Jensen's inequality,

$$E\left[\int_0^1 L(t, X^u(t); u(t))\,\mathrm{d}t\right] \ge E\left[\int_0^1 L(t, X^u(t); \beta_{X^u}(t, X^u))\,\mathrm{d}t\right]. \tag{1.13}$$

In Section 2 we state our result which will be proved in Section 4. Technical lemmas are given in Section 3. In Appendix we give the proof of the last past of Remark 1.2 and a brief description of an h-path process $\{X_h(t)\}_{0\le t\le 1}$ for the readers' convenience.

2. Duality theorem and applications

We recall that our minimization problem is

$$V(P_0,P_1) := \inf\left\{ E\left[\int_0^1 L(t, X(t); \beta_X(t,X))\,\mathrm{d}t\right] \middle| PX(t)^{-1} = P_t (t=0,1), X\in\mathcal{A}\right\}. \tag{2.1}$$

Here $\mathcal{A}$ denotes the set of all $\mathbf{R}^d$-valued, continuous semimartingales $\{X(t)\}_{0\le t\le 1}$ on a complete filtered probability space 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) $\{W_X(t) := X(t) - X(0) - \int_0^t \beta_X(s,X)\,\mathrm{d}s\}_{0\le t\le 1}$ is a $\sigma[X(s): 0\le s\le t]$-Brownian motion.

In this paper we will use the following notation when we refer to the properties of L.

(A.0) $L(t,x;u) : [0,1]\times\mathbf{R}^d\times\mathbf{R}^d \mapsto [0,\infty)$ is continuous and is convex in u.

(A.1) There exists $\delta > 1$ such that

$$\liminf_{|u|\to\infty} \frac{\inf\{L(t,x;u) : (t,x) \in [0,1] \times \mathbf{R}^d\}}{|u|^\delta} > 0.$$

(A.2)

$$\Delta L(\varepsilon_1, \varepsilon_2) := \sup \frac{L(t,x;u) - L(s,y;u)}{1 + L(s,y;u)} \to 0 \quad \text{as } \varepsilon_1, \varepsilon_2 \to 0,$$

where the supremum is taken over all (t, x) and $(s, y), \in [0, 1]\times\mathbf{R}^d$, for which $|t-s| \le \varepsilon_1$, $|x - y| < \varepsilon_2$ and all $u \in \mathbf{R}^d$.

(A.3) (i) $L(t, x; u) \in C^3([0, 1] \times \mathbf{R}^d \times \mathbf{R}^d : [0, \infty))$,
(ii) $D_u^2 L(t, x; u)$ is positive definite for all $(t, x, u) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d$,
(iii) $\sup\{L(t, x; 0) : (t, x) \in [0, 1] \times \mathbf{R}^d\}$ is finite,
(iv) $|D_x L(t, x; u)|/(1 + L(t, x; u))$ is bounded,
(v) $\sup\{|D_u L(t, x; u)| : (t, x) \in [0, 1] \times \mathbf{R}^d, |u| \le R\}$ is finite for all $R > 0$.

(A.4) (i) $\Delta L(0, \infty)$ is finite, or (ii) $\delta = 2$ in (A.1).

Remark 2.1. (i). Take $a \in C_b^1([0, 1] \times \mathbf{R}^d) \cap C^3([0, 1] \times \mathbf{R}^d)$ for which $\inf\{a(t, x)|(t, x) \in [0, 1] \times \mathbf{R}^d\} > 0$. If $L = a(t, x)(1 + |u|^2)^{\delta/2}$ $(\delta > 1)$, then (A.1)–(A.3) and (A.4, i) hold. Take a uniformly positive definite $A(t, x) = (A_{ij}(t, x))_{i,j=1}^d$ for which $A_{ij} \in C_b^1([0, 1] \times \mathbf{R}^d) \cap C^3([0, 1] \times \mathbf{R}^d)$ $(i, j = 1, \ldots, d)$. If $L = \langle A(t, x)u, u\rangle$, then (A.1)–(A.4) hold. (ii). (A.3, i, ii) imply (A.0). (iii). (A.1) and (A.3, i, ii) imply that for any $(t, x) \in [0, 1]\times\mathbf{R}^d$, $H(t, x; \cdot) \in C^3(\mathbf{R}^d)$ and for any u and $z \in \mathbf{R}^d$,

$$\begin{aligned} &z = D_u L(t, x; u) \quad \text{if and only if } u = D_z H(t, x; z), \\ &D_u^2 L(t, x; u) = D_z^2 H(t, x; z)^{-1} \quad \text{if } u = D_z H(t, x; z) \end{aligned}$$

(see [33, 2.1.3]), where $D_u^2 := (\partial^2/\partial u_i \partial u_j)_{i,j=1}^d$.

We give a result on the existence of a minimizer of $V(P_0, P_1)$.

Proposition 2.1. *Suppose that (A.0)–(A.2) hold. Then for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ for which $V(P_0, P_1)$ is finite, $V(P_0, P_1)$ has a minimizer.*

The following is our main result.

Theorem 2.1 (*Duality Theorem*). *Suppose that* (A.1)–(A.4) *hold. Then* (1.8) *holds, namely, for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$V(P_0, P_1) = \sup\left\{\int_{\mathbf{R}^d} \varphi(1, x) P_1(\mathrm{d}x) - \int_{\mathbf{R}^d} \varphi(0, x) P_0(\mathrm{d}x)\right\} (\in [0, \infty]), \tag{2.2}$$

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

$$\frac{\partial \varphi(t, x)}{\partial t} + \frac{1}{2}\Delta\varphi(t, x) + H(t, x; D_x\varphi(t, x)) = 0 \quad ((t, x) \in [0, 1) \times \mathbf{R}^d). \tag{2.3}$$

Corollary 2.1. *Suppose that* (A.1)–(A.4) *hold. Then for any* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ *for which* $V(P_0, P_1)$ *is finite and any minimizer* $\{X(t)\}_{0\le t\le 1}$ *of* $V(P_0, P_1)$*, there exists a sequence of classical solutions* $\{\varphi_n\}_{n\ge 1}$*, of the HJB equation* (2.3)*, such that* $\varphi_n(1,\cdot) \in C_b^\infty(\mathbf{R}^d)$ $(n \ge 1)$ *and that the following holds:*

$$\begin{aligned}\beta_X(t, X) &= b_X(t, X(t)) := E[\beta_X(t, X)|(t, X(t))]\\ &= \lim_{n\to\infty} D_z H(t, X(t); D_x\varphi_n(t, X(t))) \quad \mathrm{d}t\,\mathrm{d}PX(\cdot)^{-1}\text{-a.e.}\end{aligned} \tag{2.4}$$

Since classical solutions to PDEs are viscosity solutions (see, e.g., [13]), we obtain the following.

Corollary 2.2. *Suppose that* (A.1)–(A.3) *and* (A.4, i) *hold. Then* (2.2) *holds even if the supremum is taken over all bounded, uniformly Lipschitz continuous viscosity solutions* φ *of* (2.3).

Next we study just the case where (A.4, ii) holds.

Proposition 2.2. (i) *Suppose that* (A.0)–(A.2) *and* (A.4, ii) *hold. Then for any* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ *for which* $V(P_0, P_1)$ *is finite,* $V(P_0, P_1)$ *has a Markovian minimizer.*
(ii) *Suppose in addition that for any* $(t, x) \in [0, 1] \times \mathbf{R}^d$*,* $L(t, x; u)$ *is strictly convex in* u*. Then the minimizer is unique.*

We now introduce the additional assumption:

(A.5) $D_u^2 L(t, x; u)$ is bounded,

and show that a minimizer of $V(P_0, P_1)$ satisfies a FBSDE (see (2.1) for notation).

Theorem 2.2. *Suppose that* (A.1)–(A.3), (A.4, ii) *and* (A.5) *hold. Then for any* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ *for which* $V(P_0, P_1)$ *is finite and the unique minimizer* $\{X(t)\}_{0\le t\le 1}$ *of* $V(P_0, P_1)$*, there exist* $f(\cdot) \in L^1(\mathbf{R}^d, P_1(\mathrm{d}x))$ *and a* $\sigma[X(s) : 0 \le s \le t]$*-continuous semimartingale* $\{Y(t)\}_{0\le t\le 1}$ *such that*

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

satisfies the following FBSDE: for $t \in [0, 1]$*,*

$$X(t) = X(0) + \int_0^t D_z H(s, X(s); Z(s))\,\mathrm{d}s + W_X(t), \tag{2.5}$$

$$Y(t) = f(X(1)) - \int_t^1 L(s, X(s); D_z H(s, X(s); Z(s)))\,\mathrm{d}s - \int_t^1 \langle Z(s), \mathrm{d}W_X(s)\rangle.$$

Remark 2.2. (i). (A.4, ii) and (A.5) is appropriate in our approach. Indeed, suppose that $L = |u|^\delta$ and $E[\int_0^1 |b_X(s, X(s))|^\delta \mathrm{d}s]$ is finite. Then δ should be greater than or equal to 2 so that $P((X(0), X(1)) \in \mathrm{d}x\,\mathrm{d}y)$ is absolutely continuous with respect to $P(X(0) \in \mathrm{d}x)P(X(1) \in \mathrm{d}y)$. Also δ should be less than or equal to 2 so that $\{\int_0^t < Z(s), \mathrm{d}W_X(s) >\}_{0\le t\le 1}$ is a square integrable martingale (see the proof of Theorem 2.2). (ii). The existence of a solution to (2.5) cannot be proved using the known result since assumptions in Theorem 2.2 do not imply the Lipschitz continuity of $z \mapsto L(s, x; D_z H(s, x; z))$ (see [7]). Indeed, $f(x)$ is not always smooth even when $L = |u|^2$ (see in Appendix). Besides this, (A.1) and (A.3, i, ii) imply the following (see Remark 2.1, (iii)):

$$D_z\{L(s, x; D_z H(s, x; z))\} = D_z^2 H(s, x; z) D_u L(s, x; D_z H(s, x; z)) = D_z^2 H(s, x; z) z,$$

which is not bounded.

As an application of Theorem 2.1, we consider h-path processes. We shall refer here to

(A.6) There exist bounded, uniformly continuous functions $\xi : [0, 1] \times \mathbf{R}^d \mapsto \mathbf{R}^d$ and $c : [0, 1] \times \mathbf{R}^d \mapsto [0, \infty)$ such that

$$L(t, x; u) = \frac{1}{2}|u - \xi(t, x)|^2 + c(t, x) \quad ((t, x; u) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d).$$

Let $\{\mathbf{X}(t)\}_{0 \le t \le 1}$ be a unique weak solution, to the following SDE, which can be constructed by the change of measure (see (1.11) for notation and [22]): for $t \in [0, 1]$,

$$\mathbf{X}(t) = X_0 + \int_0^t \xi(s, \mathbf{X}(s))\, \mathrm{d}s + W(t). \tag{2.6}$$

As a corollary to Theorem 2.2, we obtain an approach to the h-path process for $\{\mathbf{X}(t)\}_{0 \le t \le 1}$ by the duality theorem (see Proposition 1.1).

Corollary 2.3. *Suppose that* (A.3, i, iv) *and* (A.6) *hold. Then for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ for which $V(P_0, P_1)$ is finite and the unique minimizer $\{X(t)\}_{0 \le t \le 1}$ of $V(P_0, P_1)$, there exist $f_t \in L^1(\mathbf{R}^d, P_t(\mathrm{d}x))$ $(t = 0, 1)$ such that the following holds: for any Borel set $A \subset C([0, 1])$,*

$$P(X(\cdot) \in A) = E\left[\exp\left\{f_1(\mathbf{X}(1)) - f_0(\mathbf{X}(0)) - \int_0^1 c(t, \mathbf{X}(t))\, \mathrm{d}t\right\} : \mathbf{X}(\cdot) \in A\right]. \tag{2.7}$$

Remark 2.3. Corollary 2.3 is known (see e.g. [29]).

3. Lemmas

In this section we give technical lemmas.

The following two lemmas on the property of $V(\cdot, \cdot)$ will play a crucial role in the sequel.

Lemma 3.1. *Suppose that* (A.0)–(A.2) *hold. Then $(Q, P) \mapsto V(Q, P)$ is lower semicontinuous.*

Proof. Suppose that Q_n and P_n weakly converge to Q and P as $n \to \infty$, respectively, and that $\{V(Q_n, P_n)\}_{n \ge 1}$ is bounded. Then we can take $\{X_n(t)\}_{n \ge 1} \subset \mathcal{A}$ such that $P X_n(0)^{-1} = Q_n$ and $P X_n(1)^{-1} = P_n$ $(n \ge 1)$ and that

$$0 \le E\left[\int_0^1 L(t, X_n(t); \beta_{X_n}(t, X_n))\, \mathrm{d}t\right] - V(Q_n, P_n) \to 0 \quad \text{as } n \to \infty.$$

It is easy to see that $\{(X_n(t), \int_0^t \beta_{X_n}(s, X_n)\, \mathrm{d}s) : t \in [0, 1]\}_{n \ge 1}$ is tight in $C([0, 1]; \mathbf{R}^{2d})$ from (A.1) (see [17] or [35]).

Take a weakly convergent subsequence $\{(X_{n_k}(t), \int_0^t \beta_{X_{n_k}}(s, X_{n_k})\, ds) : t \in [0, 1]\}_{k\geq 1}$ so that

$$\lim_{k\to\infty} E\left[\int_0^1 L(t, X_{n_k}(t); \beta_{X_{n_k}}(t, X_{n_k}))\, dt\right] = \liminf_{n\to\infty} V(Q_n, P_n) < \infty. \tag{3.1}$$

Let $\{(X(t), A(t))\}_{t\in[0,1]}$ denote the limit of $\{(X_{n_k}(t), \int_0^t \beta_{X_{n_k}}(s, X_{n_k})\, ds) : t \in [0, 1]\}_{k\geq 1}$ as $k \to \infty$. Then $\{X(t) - X(0) - A(t)\}_{t\in[0,1]}$ is a $\sigma[X(s) : 0 \leq s \leq t]$-Brownian motion and $\{A(t)\}_{t\in[0,1]}$ is absolutely continuous (see [17] or [35]).

We can also prove, in the same way as in the proof of [24, (3.17)], the following: from (A.0) and (A.2),

$$\lim_{k\to\infty} E\left[\int_0^1 L(t, X_{n_k}(t); \beta_{X_{n_k}}(t, X_{n_k}))\, dt\right] \geq E\left[\int_0^1 L\left(t, X(t); \frac{dA(t)}{dt}\right) dt\right]. \tag{3.2}$$

In the same way as in (1.13), on considering the completion of $PX(\cdot)^{-1}$, the proof is over since

$$PX(t)^{-1} = \lim_{k\to\infty} PX_{n_k}(t)^{-1} \quad \text{weakly } (0 \leq t \leq 1). \quad \square$$

Lemma 3.2. *Suppose that* (A.0)–(A.2), (A.3, iii) *and* (A.4) *hold. Then for any* $P_0 \in \mathcal{M}_1(\mathbf{R}^d)$, $P \mapsto V(P_0, P)$ *is convex.*

Proof. Take $X_i \in \mathcal{A}$ $(i = 1, 2)$ for which $PX_i(0)^{-1} = P_0$ and

$$\sum_{j=1}^{2} E\left[\int_0^1 L(t, X_j(t); \beta_{X_j}(t, X_j))\, dt\right] < \infty. \tag{3.3}$$

For $i = 1, 2$, $n \geq 1$, $t \in [0, 1]$ and $\omega \in C([0, 1])$, put

$$u_{n,i}(t, \omega) := 1_{[0,n]}(|\beta_{X_i}(t, \omega)|)\beta_{X_i}(t, \omega), \tag{3.4}$$

$$X_{n,i}(t) := X_i(0) + \int_0^t u_{n,i}(s, X_i)\, ds + W_{X_i}(t), \tag{3.5}$$

where 1_A denotes the indicator function of A.

Then $\{X_{n,i}(t)\}_{0\leq t\leq 1} \in \mathcal{A}$ since $u_{n,i}$ $(i = 1, 2)$ are bounded for each $n \geq 1$ (see (1.11) and (1.12)). In particular, we can assume that on the same probability space $(\Omega, \mathcal{F}, P)$, $\{X_{n,i}(t)\}_{0\leq t\leq 1}$ $(i = 1, 2)$ are defined by the change of measures (see [22, p. 279]): for $n \geq 1$ and $t \in [0, 1]$,

$$X_{n,i}(t) = X_0 + \int_0^t \beta_{X_{n,i}}(s, X_{n,i})\, ds + W(t) \tag{3.6}$$

(see (1.11) for notation). More precisely, for any $B \in \mathcal{B}(C([0, 1]))$,

$$PX_{n,i}(\cdot)^{-1}(B) = E[M_{n,i}(1, X_0 + W(\cdot)) : X_0 + W(\cdot) \in B], \tag{3.7}$$

where

$$M_{n,i}(t, \omega) := \exp\left(\int_0^t \beta_{X_{n,i}}(s, \omega)\, d\omega(s) - \frac{\int_0^t |\beta_{X_{n,i}}(s, \omega)|^2\, ds}{2}\right).$$

By Itô's formula, we can show that for any $\lambda \in (0,1)$, $\lambda PX_{n,1}(\cdot)^{-1} + (1-\lambda)PX_{n,2}(\cdot)^{-1}$ is a distribution of $\{Z_{n,\lambda}(t)\}_{0\le t\le 1}, \in \mathcal{A}$, such that for $t \in [0,1]$,

$$\beta_{Z_{n,\lambda}}(t,\omega) = \frac{\lambda\beta_{X_{n,1}}(t,\omega)M_{n,1}(t,\omega) + (1-\lambda)\beta_{X_{n,2}}(t,\omega)M_{n,2}(t,\omega)}{\lambda M_{n,1}(t,\omega) + (1-\lambda)M_{n,2}(t,\omega)}. \tag{3.8}$$

Hence, from (A.0), (3.7) and (3.8),

$$\begin{aligned}
&E\left[\int_0^1 L(t, Z_{n,\lambda}(t); \beta_{Z_{n,\lambda}}(t, Z_{n,\lambda}))\,\mathrm{d}t\right]\\
&\quad = E\left[\int_0^1 L(t, X_0 + W(t); \beta_{Z_{n,\lambda}}(t, X_0 + W(\cdot)))\right.\\
&\qquad \left.\times\ \{\lambda M_{n,1}(t, X_0 + W(\cdot)) + (1-\lambda)M_{n,2}(t, X_0 + W(\cdot))\}\,\mathrm{d}t\right]\\
&\quad \le \lambda E\left[\int_0^1 L(t, X_{n,1}(t); \beta_{X_{n,1}}(t, X_{n,1}))\,\mathrm{d}t\right]\\
&\qquad + (1-\lambda)E\left[\int_0^1 L(t, X_{n,2}(t); \beta_{X_{n,2}}(t, X_{n,2}))\,\mathrm{d}t\right].
\end{aligned} \tag{3.9}$$

First we consider the left hand side of (3.9). In the same way as in the proof of Lemma 3.1, we can show that the liminf of the left hand side of (3.9) as $n \to \infty$ is greater than or equal to $V(P_0, \lambda PX_1(1)^{-1} + (1-\lambda)PX_2(1)^{-1})$ since, from (3.3), (3.4), (1.12), (3.7), (3.8) and (A.1), by Hölder's inequality,

$$\begin{aligned}
&E\left[\int_0^1 |\beta_{Z_{n,\lambda}}(s, Z_{n,\lambda})|^\delta\,\mathrm{d}s\right]\\
&\quad \le \lambda E\left[\int_0^1 |\beta_{X_{n,1}}(s, X_{n,1})|^\delta\,\mathrm{d}s\right] + (1-\lambda)E\left[\int_0^1 |\beta_{X_{n,2}}(s, X_{n,2})|^\delta\,\mathrm{d}s\right]\\
&\quad \le \lambda E\left[\int_0^1 |u_{n,1}(s, X_1)|^\delta\,\mathrm{d}s\right] + (1-\lambda)E\left[\int_0^1 |u_{n,2}(s, X_2)|^\delta\,\mathrm{d}s\right]\\
&\quad \le \lambda E\left[\int_0^1 |\beta_{X_1}(s, X_1)|^\delta\,\mathrm{d}s\right] + (1-\lambda)E\left[\int_0^1 |\beta_{X_2}(s, X_2)|^\delta\,\mathrm{d}s\right] < \infty.
\end{aligned}$$

Next we consider the right hand side of (3.9). We first consider the case where (A.4, i) holds. For $i = 1$ and 2, by Jensen's inequality, from (1.12),

$$\begin{aligned}
&\int_0^1 E[L(t, X_{n,i}(t); \beta_{X_{n,i}}(t, X_{n,i}))]\,\mathrm{d}t\\
&\quad \le \int_0^1 E[L(t, X_{n,i}(t); u_{n,i}(t, X_i))]\,\mathrm{d}t \to \int_0^1 E[L(t, X_i(t); \beta_{X_i}(t, X_i))]\,\mathrm{d}t
\end{aligned} \tag{3.10}$$

as $n \to \infty$ from (A.0), (A.3, iii) and (3.3), by the dominated convergence theorem. Indeed, from (A.4, i) and (3.4),

$$
\begin{aligned}
0 &\leq L(t, X_{n,i}(t); u_{n,i}(t, X_i)) \\
&\leq (1 + \Delta L(0, \infty))(L(t, X_i(t); u_{n,i}(t, X_i)) + 1) \\
&\leq (1 + \Delta L(0, \infty))\{L(t, X_i(t); \beta_{X_i}(t, X_i)) + L(t, X_i(t); 0) + 1\}.
\end{aligned}
$$

(3.9) and (3.10) imply that $P \mapsto V(P_0, P)$ is convex. Next we consider the case where (A.4, ii) holds. In this case we can let $n \to \infty$ from the beginning since $PX_i(\cdot)^{-1}$ $(i = 1, 2)$ are absolutely continuous with respect to $P(X_0 + W(\cdot))^{-1}$ (see [22]). Hence (3.9) immediately implies that $P \mapsto V(P_0, P)$ is convex. □

In the same way as for $\mathcal{A}$, we define the set of semimartingales $\mathcal{A}_t$ in $C([t, 1])$. Let us recall the following result which relies on the fact that (A.3, ii) implies that for any $(t, x) \in [0, 1] \times \mathbf{R}^d$, $L(t, x; u)$ is strictly convex in u.

Lemma 3.3 (*[13, p. 210, Remark 11.2]*). *Suppose that* (A.1) *and* (A.3) *hold. Then for any* $f \in C_b^\infty(\mathbf{R}^d)$, *the HJB equation* (2.3) *with* $\varphi(1, \cdot) = f$ *has a unique solution* φ, $\in C^{1,2}([0, 1] \times \mathbf{R}^d) \cap C_b^{0,1}([0, 1] \times \mathbf{R}^d)$, *which can be written as follows:*

$$
\varphi(t, x) = \sup_{X \in \mathcal{A}_t} \left\{ E[\varphi(1, X(1)) | X(t) = x] - E\left[\int_t^1 L(s, X(s); \beta_X(s, X)) \, \mathrm{d}s \,\middle|\, X(t) = x \right] \right\}, \tag{3.11}
$$

where for the maximizer $X \in \mathcal{A}_t$, *the following holds:*

$$
\beta_X(s, X) = D_x H(s, X(s); D_x \varphi(s, X(s))).
$$

From Remark 2.1, (iii), $H \in C^3([0, 1] \times \mathbf{R}^d \times \mathbf{R}^d)$ provided (A.1) and (A.3, i,ii) hold. The following lemma will be used in the proof of Corollary 2.2.

Lemma 3.4. *Suppose that* (A.1) *and* (A.3, i, ii, iii) *hold. Then for any* $r > 0$,

$$
\sup\{|D_z H(t, x; z)| : (t, x, z) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d, |z| < r\} < \infty. \tag{3.12}
$$

Proof. For any $r > 0$, there exists $R(r) > 0$ such that

$$
\inf\{|D_u L(t, x; u)| : (t, x, u) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d, |u| > R(r)\} \geq r \tag{3.13}
$$

since from (A.3, i, ii), for any $(t, x, u) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d$,

$$
L(t, x; 0) \geq L(t, x; u) + \langle D_u L(t, x; u), -u \rangle,
$$

from which

$$
\begin{aligned}
&\inf\{|D_u L(t, x; u)| : (t, x) \in [0, 1] \times \mathbf{R}^d\} \\
&\geq \frac{1}{|u|}\{\inf\{L(t, x; u) : (t, x) \in [0, 1] \times \mathbf{R}^d\} - \sup\{L(t, x; 0) : (t, x) \in [0, 1] \times \mathbf{R}^d\}\} \\
&\to \infty \quad (\text{as } |u| \to \infty \text{ (from (A.1) and (A.3, iii))}).
\end{aligned}
$$

The supremum in (3.12) is less than or equal to $R(r)(< \infty)$. Indeed, if this is not true, then there exists $(t, x, z) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d$ for which

$$
|z| < r, \quad |D_z H(t, x; z)| > R(r).
$$

The second inequality implies, from (3.13), that $|z| \geq r$ since

$$z = D_u L(t, x; D_z H(t, x; z))$$

from Remark 2.1, (iii). This contradicts the fact that $|z| < r$. □

Next we state and prove lemmas which will be used in the proof of Proposition 2.2.

Lemma 3.5 (*[3, p. 114] and [4]*). *Suppose that* $\{P(t, \mathrm{d}x)\}_{t \in [0,1]} \subset \mathcal{M}_1(\mathbf{R}^d)$ *such that there exists* $b(t, x) : [0, 1] \times \mathbf{R}^d \mapsto \mathbf{R}^d$ *which satisfies the following:*

$$\frac{\partial P(t, \mathrm{d}x)}{\partial t} = \frac{1}{2} \triangle P(t, \mathrm{d}x) - div(b(t, x) P(t, \mathrm{d}x))(in\ dist.\ sense), \tag{3.14}$$

$$\int_0^1 \mathrm{d}t \int_{\mathbf{R}^d} |b(t, x)|^2 P(t, \mathrm{d}x) < \infty. \tag{3.15}$$

Then there exists a unique weak solution $\{X(t)\}_{0 \leq t \leq 1}$ *to the following (see (1.11) for notation): for* $t \in [0, 1]$,

$$X(t) = X(0) + \int_0^t b(s, X(s))\, \mathrm{d}s + W(t), \tag{3.16}$$

$$P(X(t) \in \mathrm{d}x) = P(t, \mathrm{d}x). \tag{3.17}$$

Put

$$\underline{V}(P_0, P_1) := \inf \int_0^1 \int_{\mathbf{R}^d} L(t, x; b(t, x)) P(t, \mathrm{d}x)\, \mathrm{d}t, \tag{3.18}$$

where the infimum is taken over all $(b(t, x), P(t, \mathrm{d}x))$ for which $\{P(t, \mathrm{d}x)\}_{0 \leq t \leq 1} \subset \mathcal{M}_1(\mathbf{R}^d)$, (3.14) holds and $P(t, \mathrm{d}x) = P_t$ $(t = 0, 1)$.

The following which can be proved from Lemma 3.5 can be considered as a generalization of [23, Lemma 2.5] which is a stochastic control counterpart of [2] (see also [33, p. 239]) when $L(t, x; u) = |u|^2$.

Lemma 3.6. *Suppose that* (A.0)–(A.1) *and* (A.4, ii) *hold. Then for any* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$, $V(P_0, P_1) = \underline{V}(P_0, P_1)$.

Proof. We first prove

$$V(P_0, P_1) \geq \underline{V}(P_0, P_1). \tag{3.19}$$

Take $X \in \mathcal{A}$ such that $E[\int_0^1 L(t, X(t); \beta_X(t, X))\, \mathrm{d}t]$ is finite and that $PX(t)^{-1} = P_t$ $(t = 0, 1)$. Set $b_X(t, X(t)) := E[\beta_X(t, X)|(t, X(t))]$.

Then $(b_X(t, x), P(X(t) \in \mathrm{d}x))$ satisfies (3.14). Indeed, for any $f \in C_0^\infty(\mathbf{R}^d)$ and $t \in [0, 1]$, by Itô's formula,

$$\begin{aligned}
&\int_{\mathbf{R}^d} f(x) P(X(t) \in \mathrm{d}x) - \int_{\mathbf{R}^d} f(x) P(X(0) \in \mathrm{d}x) \\
&= E[f(X(t)) - f(X(0))] \\
&= \int_0^t \mathrm{d}s\, E\left[\frac{1}{2} \triangle f(X(s)) + \langle \beta_X(s, X), Df(X(s)) \rangle\right]
\end{aligned}$$

$$= \int_0^t \mathrm{d}s E\left[\frac{1}{2}\triangle f(X(s)) + \langle E[\beta_X(s,X)|(s,X(s))], Df(X(s))\rangle\right]$$

$$= \int_0^t \mathrm{d}s \int_{\mathbf{R}^d}\left(\frac{1}{2}\triangle f(x) + \langle b_X(s,x), Df(x)\rangle\right) P(X(s)\in \mathrm{d}x). \tag{3.20}$$

Hence, from (A.0), by Jensen's inequality,

$$E\left[\int_0^1 L(t,X(t);\beta_X(t,X))\,\mathrm{d}t\right]$$
$$\geq E\left[\int_0^1 L(t,X(t);b_X(t,X(t)))\,\mathrm{d}t\right]$$
$$= \int_0^1 \mathrm{d}t\int_{\mathbf{R}^d} L(t,x;b_X(t,x))P(X(t)\in \mathrm{d}x) \geq \underline{V}(P_0,P_1), \tag{3.21}$$

which implies (3.19).

Next we prove the opposite inequality of (3.19). Take $(b(t,x), P(t,\mathrm{d}x))$ for which $\{P(t,\mathrm{d}x)\}_{0\leq t\leq 1}\subset \mathcal{M}_1(\mathbf{R}^d)$, (3.14) holds and $P(t,\mathrm{d}x) = P_t(\mathrm{d}x)$ $(t=0,1)$ and for which $\int_0^1 \mathrm{d}t \int_{\mathbf{R}^d} L(t,x;b(t,x))P(t,\mathrm{d}x)$ is finite.

Then, from (A.1) and (A.4, ii), (3.15) holds. From Lemma 3.5, there exists a Markov process $\{X(t)\}_{0\leq t\leq 1}$ for which (3.16) and (3.17) hold. In particular, we have

$$\int_0^1 \mathrm{d}t\int_{\mathbf{R}^d} L(t,x;b(t,x))P(t,\mathrm{d}x)$$
$$= E\left[\int_0^1 L(t,X(t);b(t,X(t)))\,\mathrm{d}t\right] \geq V(P_0,P_1). \quad \square \tag{3.22}$$

4. Proof of our result

In this section we give the proof of our result.

Proof of Proposition 2.1. Replace (Q_n, P_n) by (P_0, P_1) in the proof of Lemma 3.1. Then the proof is over. $\square$

Since $P\mapsto V(P_0,P)$ is lower semicontinuous and convex from Lemmas 3.1 and 3.2, we can reduce the proof of Theorem 2.1 to the fact that $V(P_0,\cdot)^{**}(P) = V(P_0,P)$.

Proof of Theorem 2.1. $V(P_0,\cdot)\not\equiv\infty$. Indeed, for $P_1 = P(X_0+W(1))^{-1}$ (see (1.11) for notation), from (A.3, iii),

$$V(P_0,P_1)\leq \sup\{L(t,x;0) : (t,x)\in[0,1]\times\mathbf{R}^d\} < \infty.$$

Consider $P\mapsto V(P_0,P)$ as a function on the space of finite Borel measures on $\mathbf{R}^d$, by putting $V(P_0,P) = +\infty$ for $P\notin \mathcal{M}_1(\mathbf{R}^d)$. From Lemmas 3.1 and 3.2 and [8, Theorem 2.2.15 and Lemma 3.2.3],

$$V(P_0,P_1) = \sup_{f\in C_b(\mathbf{R}^d)}\left\{\int_{\mathbf{R}^d} f(x)P_1(\mathrm{d}x) - V^*_{P_0}(f)\right\}, \tag{4.1}$$

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

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

Take $\Phi \in C_0^\infty([-1,1]^d; [0,\infty))$ for which $\int_{\mathbf{R}^d} \Phi(x)\mathrm{d}x = 1$. For $\varepsilon > 0$, put

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

Denote by $\mathcal{V}(P_0, P_1)$ the right hand side of (2.2). We prove the following which implies (2.2):

$$V(P_0, P_1) \geq \mathcal{V}(P_0, P_1) \geq \frac{V(\Phi_\varepsilon * P_0, \Phi_\varepsilon * P_1)}{1 + \Delta L(0, \varepsilon)} - \Delta L(0, \varepsilon), \tag{4.2}$$

where $*$ denotes the convolution of two measures and should be distinguished from * in (4.1). Indeed, from (A.2), Lemma 3.1 and (4.2), we have (2.2).

The first inequality in (4.2) can be proved from (4.1) and (4.3) below: for any $f \in C_b^\infty(\mathbf{R}^d)$, from Lemma 3.3,

$$\begin{aligned} V^*_{P_0}(f) &= \sup \left\{ E[f(X(1))] - E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\, \mathrm{d}t \right] : \right. \\ &\qquad\qquad \left. X \in \mathcal{A}, PX(0)^{-1} = P_0 \right\} \\ &= \int_{\mathbf{R}^d} \varphi_f(0, x) P_0(\mathrm{d}x), \end{aligned} \tag{4.3}$$

where φ_f denotes the unique classical solution to the HJB equation (2.3) with $\varphi(1, \cdot) = f(\cdot)$.

We prove the second inequality in (4.2). For $f \in C_b(\mathbf{R}^d)$, put

$$f_\varepsilon(x) := \int_{\mathbf{R}^d} f(y) \Phi_\varepsilon(y - x)\, \mathrm{d}y. \tag{4.4}$$

Then $f_\varepsilon \in C_b^\infty(\mathbf{R}^d)$ and, from (4.3),

$$\begin{aligned} \mathcal{V}(P_0, P_1) &\geq \int_{\mathbf{R}^d} f_\varepsilon(x) P_1(\mathrm{d}x) - V^*_{P_0}(f_\varepsilon) \\ &\geq \int_{\mathbf{R}^d} f(x) \Phi_\varepsilon * P_1(\mathrm{d}x) - \frac{(V_{\Phi_\varepsilon * P_0})^*((1 + \Delta L(0, \varepsilon)) f)}{1 + \Delta L(0, \varepsilon)} - \Delta L(0, \varepsilon). \end{aligned} \tag{4.5}$$

Indeed, for any $X \in \mathcal{A}$, from (A.2),

$$\begin{aligned} &E[f_\varepsilon(X(1))] - E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\, \mathrm{d}t \right] \\ &= \int_{\mathbf{R}^d} \Phi(z)\, \mathrm{d}z E[f(X(1) + \varepsilon z)] - E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\, \mathrm{d}t \right] \\ &\leq \int_{\mathbf{R}^d} \Phi(z)\, \mathrm{d}z \left\{ E[f(X(1) + \varepsilon z)] - E\left[\int_0^1 \frac{L(t, X(t) + \varepsilon z; \beta_X(t, X))}{1 + \Delta L(0, \varepsilon)}\, \mathrm{d}t \right] \right\} \\ &\quad + \Delta L(0, \varepsilon). \end{aligned}$$

(4.1) and (4.5) imply the second inequality in (4.2). $\square$

Proof of Corollary 2.1. Identity (2.2) implies, by Itô's formula, that there exists a sequence $\{\varphi_n\}_{n\geq 1}$ of classical solutions, to the HJB equation (2.3), such that for any minimizer $\{X(t)\}_{0\leq t\leq 1}$ of $V(P_0, P_1)$,

$$E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\,\mathrm{d}t\right]$$
$$= \lim_{n\to\infty} E\left[\int_0^1 \{\langle \beta_X(t, X), D_x\varphi_n(t, X(t))\rangle - H(t, X(t); D_x\varphi_n(t, X(t)))\}\,\mathrm{d}t\right]. \tag{4.6}$$

Since (A.3, i, ii) imply that $L(t, x; u)$ is of class C^3 and is strictly convex in u for any $(t, x) \in [0, 1] \times \mathbf{R}^d$, (4.6) completes the proof.

Indeed, from (A.0), the following holds (see, e.g., [33]): for any $(t, x) \in [0, 1] \times \mathbf{R}^d$,

$$L(t, x; u) = \sup_{z\in\mathbf{R}^d}\{\langle z, u\rangle - H(t, x; z)\}. \tag{4.7}$$

Therefore (4.6) is equivalent to

$$0 = \lim_{n\to\infty} E\left[\int_0^1 |L(t, X(t); \beta_X(t, X)) - \{\langle \beta_X(t, X), D_x\varphi_n(t, X(t))\rangle - H(t, X(t); D_x\varphi_n(t, X(t)))\}|\,\mathrm{d}t\right], \tag{4.8}$$

which implies that there exists a subsequence $\{n_k\}_{k\geq 1}$ for which

$$L(t, X(t); \beta_X(t, X)) = \lim_{k\to\infty}\{\langle \beta_X(t, X), D_x\varphi_{n_k}(t, X(t))\rangle - H(t, X(t); D_x\varphi_{n_k}(t, X(t)))\} \tag{4.9}$$

$\mathrm{d}t\,\mathrm{d}PX(\cdot)^{-1}$-a.e. □

Proof of Corollary 2.2. For any $f \in C_b^\infty(\mathbf{R}^d)$, the HJB equation (2.3) with $\varphi(1, \cdot) = f(\cdot)$ has a unique solution $\varphi_f \in C^{1,2}([0, 1] \times \mathbf{R}^d) \cap C_b^{0,1}([0, 1] \times \mathbf{R}^d)$ from Lemma 3.3. In particular, φ_f is a bounded, uniformly Lipschitz continuous viscosity solution of the HJB equation (2.3) since a classical solution is a viscosity solution (see e.g. [13]). From Theorem 2.1, we only have to prove the following to complete the proof: for any bounded, uniformly Lipschitz continuous viscosity solution φ of the HJB equation (2.3) and any $X \in \mathcal{A}$

$$E[\varphi(1, X(1)) - \varphi(0, X(0))] \leq E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\,\mathrm{d}t\right]. \tag{4.10}$$

We prove (4.10). Take $C(\varphi) > 0$ so that

$$|\varphi(t, x) - \varphi(s, y)| \leq C(\varphi)(|t - s| + |x - y|) \quad ((t, x), (s, y) \in [0, 1] \times \mathbf{R}^d).$$

The following constant is finite from Lemma 3.4:

$$R_0 := \sup\{|D_z H(t, x; z)| : (t, x, z) \in [0, 1] \times \mathbf{R}^d \times \mathbf{R}^d, |z| \leq C(\varphi)\}. \tag{4.11}$$

Since φ is a viscosity subsolution of the HJB equation (2.3), φ is a viscosity subsolution of the HJB equation (2.3) with $H(t, x; z)$ replaced by

$$H_R(t, x; z) := \sup\{\langle z, u\rangle - L(t, x; u) | u \in \mathbf{R}^d, |u| \leq R\}$$

for any $R \geq R_0$. Indeed, whenever $h \in C^{1,2}([0, 1) \times \mathbf{R}^d)$ and $\varphi - h$ takes its maximum at $(s, y) \in [0, 1) \times \mathbf{R}^d$,

$$-|a|C(\varphi) \leq \varphi(s, y + a) - \varphi(s, y) \leq h(s, y + a) - h(s, y) \quad (a \in \mathbf{R}^d).$$

Therefore $|D_y h(s, y)| \leq C(\varphi)$, from which

$$\begin{aligned} &H(s, y; D_x h(s, y)) \\ &\quad = \langle D_x h(s, y), D_z H(s, y; D_x h(s, y))\rangle - L(s, y; D_z H(s, y; D_x h(s, y))) \\ &\quad = H_R(s, y; D_x h(s, y)) \quad (R \geq R_0) \end{aligned}$$

from (4.11) (see Remark 2.1, (iii)). In the same way we can prove that φ is a viscosity supersolution of the HJB equation (2.3) with H replaced by H_R for any $R \geq R_0$.

From [21, Theorem II.3], for any $X \in \mathcal{A}$,

$$E[\varphi(1, X_n(1)) - \varphi(0, X_n(0))] \leq E\left[\int_0^1 L(t, X_n(t); u_n(t, X))\,\mathrm{d}t\right], \tag{4.12}$$

where we define u_n and X_n in the same way as in (3.4) and (3.5). In the same way as in (3.10), from (A.4, i),

$$\lim_{n\to\infty} E\left[\int_0^1 L(t, X_n(t); u_n(t, X))\,\mathrm{d}t\right] = E\left[\int_0^1 L(t, X(t); \beta_X(t, X))\,\mathrm{d}t\right], \tag{4.13}$$

which completes the proof of (4.10). □

From Proposition 2.1, Lemmas 3.5 and 3.6, we prove Proposition 2.2.

Proof of Proposition 2.2. From Proposition 2.1, $V(P_0, P_1)$ has a minimizer. From Lemma 3.6, in the same way as in (3.21), we can prove that $\underline{V}(P_0, P_1)$ has a minimizer. Hence, from Lemmas 3.5 and 3.6, there exists a Markovian minimizer of $V(P_0, P_1)$.

If for any $(t, x) \in [0, 1] \times \mathbf{R}^d$, $L(t, x; u)$ is strictly convex in u, then all minimizers of $V(P_0, P_1)$ are Markovian.

Indeed, Lemma 3.6 and (3.21) imply that if X is a minimizer of $V(P_0, P_1)$, then

$$\beta_X(t, X) = b_X(t, X(t)) \quad \mathrm{d}t\,\mathrm{d}PX(\cdot)^{-1}\text{-a.e.}$$

(A.4, ii) implies that $PX(\cdot)^{-1}$ is absolutely continuous with respect to $P(X_0 + W(\cdot))^{-1}$ (see (1.11) for notation). Hence $\{X(t)\}_{0\leq t\leq 1}$ is Markovian.

In particular, from Lemmas 3.5 and 3.6, the set of all minimizers of $\underline{V}(P_0, P_1)$ is equal to that of all $\{(b_X(t, x), P(X(t) \in \mathrm{d}x))\}_{0\leq t\leq 1}$ for the Markovian minimizers $\{X(t)\}_{0\leq t\leq 1}$ of $V(P_0, P_1)$.

Hence, to prove the uniqueness of a minimizer of $V(P_0, P_1)$, we only have to prove that of b for which there exists $\{P(t, \mathrm{d}x)\}_{0\leq t\leq 1}$ such that $\{(b(t, x), P(t, \mathrm{d}x))\}_{0\leq t\leq 1}$ is a minimizer of $\underline{V}(P_0, P_1)$.

Indeed, since $PX(\cdot)^{-1}$ is absolutely continuous with respect to $P(X_0 + W(\cdot))^{-1}$ for a Markovian minimizer $\{X(t)\}_{0\leq t\leq 1}$ of $V(P_0, P_1)$, $\{b_X(t, x)\}_{0\leq t\leq 1}$ determines $PX(\cdot)^{-1}$.

Take minimizers $(b_i(t, x), P_i(t, \mathrm{d}x))$ of $\underline{V}(P_0, P_1)$ $(i = 0, 1)$. For any $\lambda \in (0, 1)$, put $p_i(t, x) := P_i(t, \mathrm{d}x)/\mathrm{d}x$ and

$$b_\lambda(t, x) := \frac{(1 - \lambda)b_0(t, x)p_0(t, x) + \lambda b_1(t, x)p_1(t, x)}{(1 - \lambda)p_0(t, x) + \lambda p_1(t, x)} \quad (0 < t \leq 1),$$

provided that the denominator is positive. Then

$$
\begin{aligned}
&\underline{V}(P_0, P_1) \\
&\le \int_0^1 \mathrm{d}t \int_{\mathbf{R}^d} L(t, x; b_\lambda(t, x))((1-\lambda)p_0(t, x) + \lambda p_1(t, x))\mathrm{d}x \\
&\le (1-\lambda)\int_0^1 \mathrm{d}s \int_{\mathbf{R}^d} L(t, x; b_0(t, x))p_0(t, x)\mathrm{d}x \\
&\quad + \lambda \int_0^1 \mathrm{d}s \int_{\mathbf{R}^d} L(t, x; b_1(t, x))p_1(t, x)\mathrm{d}x \\
&= \underline{V}(P_0, P_1).
\end{aligned} \tag{4.14}
$$

Indeed,

$$
\begin{aligned}
&\frac{\partial((1-\lambda)p_0(t, x) + \lambda p_1(t, x))}{\partial t} \\
&= \frac{1}{2}\triangle((1-\lambda)p_0(t, x) + \lambda p_1(t, x)) - \mathrm{div}(b_\lambda(t, x)((1-\lambda)p_0(t, x) + \lambda p_1(t, x)))
\end{aligned}
$$

in the dist. sense.

From (4.14), by the strict convexity of $u \mapsto L(t, x; u)$ ($(t, x) \in [0, 1] \times \mathbf{R}^d$),

$$
b_0(t, x) = b_1(t, x) \quad \text{if } p_0(t, x)p_1(t, x) > 0. \tag{4.15}
$$

Putting $b_i(t, x) = b_j(t, x)$ if $p_i(t, x) = 0$ ($i, j = 0, 1, i \neq j$), the proof is over. □

From Theorem 2.1 and Proposition 2.2, we prove Theorem 2.2.

Proof of Theorem 2.2. Take $\{\varphi_n\}_{n\ge 1}$ in (4.8). Then for $t \in [0, 1]$, by Itô's formula,

$$
\begin{aligned}
&\varphi_n(t, X(t)) - \varphi_n(0, X(0)) \\
&= \int_0^t \{\langle b_X(s, X(s)), D_x\varphi_n(s, X(s))\rangle - H(s, X(s); D_x\varphi_n(s, X(s)))\}\,\mathrm{d}s \\
&\quad + \int_0^t \langle D_x\varphi_n(s, X(s)), \mathrm{d}W_X(s)\rangle.
\end{aligned} \tag{4.16}
$$

By Doob's inequality (see [16]),

$$
\begin{aligned}
&E\left[\sup_{0\le t\le 1}\left|\int_0^t \langle D_x\varphi_n(s, X(s)), \mathrm{d}W_X(s)\rangle - \int_0^t \langle D_u L(s, X(s); b_X(s, X(s))), \mathrm{d}W_X(s)\rangle\right|^2\right] \\
&\le 4E\left[\int_0^1 |D_x\varphi_n(s, X(s)) - D_u L(s, X(s); b_X(s, X(s)))|^2\,\mathrm{d}s\right] \\
&\le 4CE\left[\int_0^1 \{L(s, X(s); b_X(s, X(s))) - \langle b_X(s, X(s)), D_x\varphi_n(s, X(s))\rangle\right. \\
&\qquad \left. + H(s, X(s); D_x\varphi_n(s, X(s)))\}\,\mathrm{d}s\right] \\
&\to 0 \quad \text{as } n \to \infty
\end{aligned} \tag{4.17}
$$

from (4.8), where

$$C := 2\sup\{\langle D_u^2 L(t,x;u)z,z\rangle : (t,x,u,z) \in [0,1]\times\mathbf{R}^d\times\mathbf{R}^d\times\mathbf{R}^d, |z|=1\}.$$

Indeed, for a smooth, strictly convex function $f : \mathbf{R}^d \mapsto [0,\infty)$ for which $f(v)/|v| \to \infty$ as $|v| \to \infty$ and $(u,z) \in \mathbf{R}^d \times \mathbf{R}^d$, by Taylor's Theorem, there exists $\theta \in (0,1)$ such that

$$\begin{aligned} & f(u) - \{\langle u,z\rangle - f^*(z)\} \\ &\quad = f^*(z) - f^*(Df(u)) - \langle Df^*(Df(u)), z - Df(u)\rangle \\ &\quad = \frac{\langle D^2 f^*(z+\theta(z-Df(u)))(z-Df(u)), z-Df(u)\rangle}{2}, \end{aligned}$$

and $D^2 f^*(z) = D^2 f(Df^*(z))^{-1}$ (see Remark 2.1, (iii)).

From (4.8), (4.16) and (4.17), $\varphi_n(1,y) - \varphi_n(0,x)$ is convergent in $L^1(\mathbf{R}^d \times \mathbf{R}^d, P((X(0),X(1)) \in \mathrm{d}x\,\mathrm{d}y))$.

From (A.4, ii), $PX(\cdot)^{-1}$ is absolutely continuous with respect to $P(X_0 + W(\cdot))^{-1}$ (see (1.11) for notation). In particular,

$$\begin{aligned} & p(t,y) := P(X(t) \in \mathrm{d}y)/\mathrm{d}y \text{ exists } (t \in (0,1]), \\ & p(0,x;t,y) := P(X(t) \in \mathrm{d}y | X(0) = x)/\mathrm{d}y \text{ exists } P_0(\mathrm{d}x)\text{-a.e. } (t \in (0,1]). \end{aligned}$$

Therefore $P((X(0),X(1)) \in \mathrm{d}x\,\mathrm{d}y)$ is absolutely continuous with respect to $P_0(\mathrm{d}x)P_1(\mathrm{d}y)$. Indeed,

$$P((X(0),X(1)) \in \mathrm{d}x\,\mathrm{d}y) = \frac{p(0,x;1,y)}{p(1,y)} P_0(\mathrm{d}x)P_1(\mathrm{d}y).$$

Hence, from [31, Prop. 2], there exist $f \in L^1(\mathbf{R}^d, P_1(\mathrm{d}x))$ and $f_0 \in L^1(\mathbf{R}^d, P_0(\mathrm{d}x))$ such that

$$\lim_{n\to\infty} E[|\varphi_n(1,X(1)) - \varphi_n(0,X(0)) - \{f(X(1)) - f_0(X(0))\}|] = 0. \tag{4.18}$$

Put

$$\begin{aligned} Y(t) := {} & f_0(X(0)) + \int_0^t L(s,X(s);b_X(s,X(s)))\,\mathrm{d}s \\ & + \int_0^t \langle D_u L(s,X(s);b_X(s,X(s))), \mathrm{d}W_X(s)\rangle. \end{aligned} \tag{4.19}$$

From (4.8) and (4.16)–(4.18), (2.5) holds. □

We prove Corollary 2.3 from Theorem 2.2.

Proof of Corollary 2.3. The assumptions in Corollary 2.3 imply those in Theorem 2.2. When (A.6) holds, from (4.8) and (4.16)–(4.18),

$$\begin{aligned} & f(X(1)) - f_0(X(0)) - \int_0^1 c(s,X(s))\,\mathrm{d}s \\ & = \int_0^1 \langle b_X(s,X(s)) - \xi(s,X(s)), \mathrm{d}X(s) - \xi(s,X(s))\,\mathrm{d}s\rangle \\ & \quad - \frac{1}{2}\int_0^1 |b_X(s,X(s)) - \xi(s,X(s))|^2\,\mathrm{d}s, \end{aligned} \tag{4.20}$$

which completes the proof (see [22]). □

Acknowledgements

We would like to thank the anonymous referees for their useful suggestions. In particular, Corollary 2.2 was suggested by them. The first author was partially supported by Grant-in-Aids for Scientific Research, No. 15340047, 15340051 and 16654031, JSPS.

Appendix

For the readers' convenience, we give the proof of Remark 1.1 and describe some properties of the h-path process $\{X_h(t)\}_{0\le t\le 1}$ introduced in Section 1.

(i) For $f \in C_b(\mathbf{R}^d)$, $f_\varepsilon \in C_b^\infty(\mathbf{R}^d)$ (see (4.4) for notation). Supposed that (1.4) holds. Then

$$\begin{aligned}\mathcal{T}(P_0, P_1) &\ge \sup\left\{\int_{\mathbf{R}^d} f(x)P_1(\mathrm{d}x) - \int_{\mathbf{R}^d} Tf(x)P_0(\mathrm{d}x)\,\middle|\, f \in C_b^\infty(\mathbf{R}^d)\right\}\\ &\ge \sup\left\{\int_{\mathbf{R}^d} f_\varepsilon(x)P_1(\mathrm{d}x) - \int_{\mathbf{R}^d} Tf_\varepsilon(x)P_0(\mathrm{d}x)\,\middle|\, f \in C_b(\mathbf{R}^d)\right\}\\ &\ge \mathcal{T}(\Phi_\varepsilon * P_0, \Phi_\varepsilon * P_1),\end{aligned} \tag{A.1}$$

where $*$ denotes the convolution. The second inequality of (A.1) is true since

$$\begin{aligned}\int_{\mathbf{R}^d} f_\varepsilon(x)P_1(\mathrm{d}x) &= \int_{\mathbf{R}^d} f(x)\Phi_\varepsilon * P_1(\mathrm{d}x),\\ Tf_\varepsilon(x) &= \sup\left\{\int_{\mathbf{R}^d}\{f(y+z) - \ell(y+z-(x+z))\}\Phi_\varepsilon(z)\,\mathrm{d}z\,\middle|\, y \in \mathbf{R}^d\right\}\\ &\le \int_{\mathbf{R}^d} Tf(x+z)\Phi_\varepsilon(z)\,\mathrm{d}z = \int_{\mathbf{R}^d} Tf(z)\Phi_\varepsilon(z-x)\,\mathrm{d}z.\end{aligned}$$

Since $\Phi_\varepsilon * P_t(\mathrm{d}x) \to P_t(\mathrm{d}x)$ weakly as $\varepsilon \to 0$ ($t = 0, 1$) and since $(P_0, P_1) \mapsto \mathcal{T}(P_0, P_1)$ is lower semicontinuous (see [33]), (A.1) implies Remark 1.1.

(ii) Suppose that $L = |u|^2$ and that (2.1) is finite. Then the probability law of $\{X_h(t)\}_{0\le t\le 1}$ is absolutely continuous with respect to that of $\{X_0 + W(t)\}_{0\le t\le 1}$ (see (1.11) for notation). In particular, P_1 is absolutely continuous with respect to the Lesbegue measure $\mathrm{d}x$ (see [22]). It is known that there exists a unique pair of nonnegative, σ-finite Borel measures (ν_0, ν_1) for which

$$\begin{cases} P_0(\mathrm{d}x) = \left(\int_{\mathbf{R}^d} g_1(x-y)\nu_1(\mathrm{d}y)\right)\nu_0(\mathrm{d}x),\\ P_1(\mathrm{d}y) = \left(\int_{\mathbf{R}^d} g_1(x-y)\nu_0(\mathrm{d}x)\right)\nu_1(\mathrm{d}y),\end{cases} \tag{A.2}$$

where for $x \in \mathbf{R}^d$ and $t > 0$,

$$g_t(x) := \frac{1}{(2\pi t)^{d/2}}\exp\left(-\frac{|x|^2}{2t}\right).$$

(A.2) is called Schrödinger's functional equation (see [18] and also [31] for the recent development). For $x \in \mathbf{R}^d$, put

$$h(t,x) := \begin{cases}\int_{\mathbf{R}^d} g_{(1-t)}(x-y)\nu_1(\mathrm{d}y) & (0 \le t < 1),\\ \dfrac{\nu_1(\mathrm{d}x)}{\mathrm{d}x} & (t = 1).\end{cases} \tag{A.3}$$

Then the h-path process $\{X_h(t)\}_{0\le t\le 1}$ is the unique weak solution to the following (see [19]): for $t \in [0, 1]$,

$$X_h(t) = X_0 + \int_0^t D_x \log h(s, X_h(s))\,\mathrm{d}s + W(t). \tag{A.4}$$

It is known that for any Borel set $A \subset C([0, 1])$,

$$P(X_h(\cdot) \in A) = E\left[\frac{h(1, X_0 + W(1))}{h(0, X_0)} : X_0 + W(\cdot) \in A\right]. \tag{A.5}$$

In particular,

$$P((X_h(0), X_h(1)) \in \mathrm{d}x\,\mathrm{d}y) = \nu_0(\mathrm{d}x) g_1(x-y)\nu_1(\mathrm{d}y). \tag{A.6}$$

From (A.2) and (A.3), $h(1, x)$ is not always smooth. But it is also known that $h \in C^{1,2}([0, 1) \times \mathbf{R}^d)$ (see [19]) and $\varphi(t, x) := \log h(t, x)$ satisfies the HJB equation (2.3). Indeed, from [19],

$$\frac{\partial h(t,x)}{\partial t} + \frac{1}{2}\triangle h(t, x) = 0 \quad ((t, x) \in (0, 1) \times \mathbf{R}^d).$$

From [25, Lemma 3.4], we also have

$$\begin{aligned} &V(P_0, P_1)\\ &= \int_{\mathbf{R}^d} |x|^2 (P_0(\mathrm{d}x) + P_1(\mathrm{d}x)) + 2\int_{\mathbf{R}^d} \left(\log \frac{P_1(\mathrm{d}x)}{\mathrm{d}x}\right) P_1(\mathrm{d}x) + \mathrm{d}\log(2\pi)\\ &\quad - 2\iint_{\mathbf{R}^d\times\mathbf{R}^d} \log\left\{\iint_{\mathbf{R}^d\times\mathbf{R}^d} \exp(\langle x, z_1\rangle + \langle y, z_0\rangle - \langle z_0, z_1\rangle)\right.\\ &\quad \left.\times\ P((X_h(0), X_h(1)) \in \mathrm{d}z_0\,\mathrm{d}z_1)\right\} P((X_h(0), X_h(1)) \in \mathrm{d}x\,\mathrm{d}y). \end{aligned} \tag{A.7}$$

References

[1] L. Ambrosio, A. Pratelli, Existence and stability results in the L^1 theory of optimal transportation, in: L.A. Caffarelli, S. Salsa (Eds.), Optimal Transportation and Applications, Martina Franca, Italy 2001, in: Lecture Notes in Math., vol. 1813, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 2003, pp. 123–160.

[2] Y. Brenier, J.D. Benamou, A numerical method for the optimal mass transport problem and related problems, in: L.A. Caffarelli, M. Milman (Eds.), Monge Ampére Equation: Applications to Geometry and Optimization (Proc. of the NSF-CBMS Conference, Deerfield Beach, FL, 1997), in: Contemporary Mathematics, vol. 226, Amer. Math. Soc., Providence, RI, 1999, pp. 1–11.

[3] P. Cattiaux, C. Léonard, Minimization of the Kullback information of diffusion processes, Ann. Inst. H. Poincaré Probab. Statist. 30 (1994) 83–132.

[4] P. Cattiaux, C. Léonard, Correction to: "Minimization of the Kullback information of diffusion processes", Ann. Inst. H. Poincaré Probab. Statist. 30 (1) (1994) 83–132; Ann. Inst. H. Poincaré Probab. Statist. 31 705–707.

[5] P. Dai Pra, A stochastic control approach to reciprocal diffusion processes, Appl. Math. Optim. 23 (1991) 313–329.

[6] A. de Acosta, Invariance principles in probability for triangular arrays of B-valued random vectors and some applications, Ann. Probab. 10 (1982) 346–373.

[7] F. Delarue, On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case, Stochastic Process. Appl. 99 (2002) 209–286.

[8] J.D. Deuschel, D.W. Stroock, Large Deviations, in: Pure and Applied Mathematics, vol. 137, Academic Press, Inc., Boston, MA, 1989.

[9] J.L. Doob, Classical Potential Theory and its Probabilistic Counterpart, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1984.

[10] R.M. Dudley, Probabilities and Metrics. Convergence of Laws on Metric Spaces, with a View to Statistical Testing, in: Lecture Notes Series, no. 45, Matematisk Institut, Aarhus Universitet, Aarhus, 1976.
[11] L.C. Evans, Partial Differential Equations, in: Graduate Studies in Mathematics, vol. 19, Amer. Math. Soc., Providence, RI, USA, 1998.
[12] L.C. Evans, Partial differential equations and Monge–Kantorovich mass transfer, in: S.T. Yau (Ed.), Current Developments in Mathematics, Cambridge, MA, 1997, Int. Press, Boston, MA, 1999, pp. 65–126.
[13] W.H. Fleming, H.M. Soner, Controlled Markov Processes and Viscosity Solutions, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1993.
[14] H. Föllmer, Random fields and diffusion processes, in: P.L. Hennequin (Ed.), École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87, in: Lecture Notes in Math., vol. 1362, Springer, Berlin, 1988, pp. 101–203.
[15] W. Gangbo, R.J. McCann, The geometry of optimal transportation, Acta Math. 177 (1996) 113–161.
[16] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland/Kodansha, Amsterdam, New York, Oxford, Tokyo, 1981.
[17] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1987.
[18] B. Jamison, Reciprocal processes, Z. Wahrsch. Verw. Gebiete 30 (1974) 65–86.
[19] B. Jamison, The Markov process of Schrödinger, Z. Wahrsch. Verw. Gebiete 32 (1975) 323–331.
[20] H.G. Kellerer, Duality theorems for marginal problems, Z. Wahrsch. Verw. Gebiete 67 (1984) 399–432.
[21] P.L. Lions, Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations, Part 2: viscosity solutions and uniqueness, Comm. P.D.E. 8 (1983) 1229–1276.
[22] R.S. Liptser, A.N. Shiryaev, Statistics of Random Processes I, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1977.
[23] T. Mikami, Variational processes from the weak forward equation, Commun. Math. Phys. 135 (1990) 19–40.
[24] T. Mikami, Optimal control for absolutely continuous stochastic processes and the mass transportation problem, Elect. Comm. in Probab. 7 (2002) 199–213.
[25] T. Mikami, Monge's problem with a quadratic cost by the zero-noise limit of h-path processes, Probab. Theory Related Fields 129 (2004) 245–260.
[26] T. Mikami, A Simple Proof of Duality Theorem for Monge–Kantorovich Problem, Kodai Math. J. 29 (2006) 1–4.
[27] T. Mikami, Semimartingales from the Fokker–Planck equation, Appl. Math. Optim. 53 (2006) 209–219.
[28] T. Mikami, M. Thieullen, Optimal Transportation Problem by Stochastic Optimal Control, in: Hokkaido University Preprint Series, vol. 690, 2005.
[29] M. Nagasawa, Stochastic Processes in Quantum Physics, in: Monographs in Mathematics, vol. 94, Birkhauser Verlag, Basel, 2000, p. 167.
[30] S.T. Rachev, L. Rüschendorf, Mass Transportation Problems, Vol. I: Theory, Vol. II: Application, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1998.
[31] L. Rüschendorf, W. Thomsen, Note on the Schrödinger equation and I-projections, Statist. Probab. Lett. 17 (1993) 369–375.
[32] M. Thieullen, Second order stochastic differential equations and non-Gaussian reciprocal diffusions, Probab. Theory Related Fields 97 (1993) 231–257.
[33] C. Villani, Topics in Optimal Transportation, in: Graduate Studies in Mathematics, vol. 58, Amer. Math. Soc., Providence, RI, 2003.
[34] J.C. Zambrini, Variational processes, in: S. Albeverio, G. Casati, D. Merlini (Eds.), Stochastic Processes in Classical and Quantum Systems, Ascona, 1985, in: Lecture Notes in Phys., vol. 262, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986, pp. 517–529.
[35] W.A. Zheng, Tightness results for laws of diffusion processes application to stochastic mechanics, Ann. Inst. Henri Poincaré 21 (1985) 103–124.