HOKKAIDO UNIVERSITY

Title	Stochastic control with fixed marginal distributions
Author(s)	Mikami, Toshio
Citation	Hokkaido University Preprint Series in Mathematics, 800, 1-48
Issue Date	2006-08-24
DOI	https://doi.org/10.14943/83950
Doc URL	https://hdl.handle.net/2115/69608
Type	departmental bulletin paper
File Information	re800.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Stochastic control with fixed marginal distributions *

Toshio Mikami†
Department of Mathematics
Hokkaido University

August 24, 2006

Running title: Stochastic control with fixed marginal distributions

Abstract

We briefly describe the so-called Monge-Kantorovich Problem (MKP for short) which is often referred to as an optimal mass transportation problem and study the stochastic optimal control problem (SOCP for short) with fixed initial and terminal distributions. In particular, we study the so-called Duality Theorem for the SOCP where continuous semimartingales under consideration have a variable diffusion matrix and then discuss the relation between the MKP and the SOCP. We also study the so-called Nelson's Problem, as the SOCP with fixed marginal distributions at each time, to which we give a new approach from the Duality Theorem. We finally consider a class of deterministic variational problems with fixed marginal distributions which is related to the SOCP by extending a class of measures under consideration.

JEL classification: C61
MSC : primary 93E20; secondary 80A20

*to be submitted to Adv. Math. Econ.

†Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan; mikami@math.sci.hokudai.ac.jp; fax no. 81/11/727/3705. Partially supported by the Grant-in-Aid for Scientific Research, No.15340047, 18204009 and 18340033, JSPS.

1 Introduction.

When a diffusion matrix of continuous semimartingales under consideration is an identity, we studied the SOCP with fixed initial and terminal distributions (see [30, 31]) and that with fixed marginal distributions at each time to consider Nelson's Problem (see [28]). In this paper we generalize them to the case where continuous semimartingales under consideration have a variable diffusion matrix. These can be considered as the problems on the **random least action principle**.

In Sect. 2 we consider the SOCP with fixed initial and terminal distributions. In Sect. 2.1, we describe the problem. In Sect. 2.2, we give the Duality Theorem for it in the frameworks of classical and viscosity solutions of the Hamilton-Jacobi-Bellman (HJB for short) PDEs. A typical minimizer of our SOCP is the so-called h-path processes about which we also discuss. In the framework of viscosity solutions of the HJB PDEs, our assumption is weaker than that in the framework of classical solutions and our result is a generalization of [30] even when a diffusion matrix under consideration is an identity. We remark that a classical solutions of the HJB PDE is a viscosity solution of it, but not vice versa. In Sect. 2.3, we show that the zero-noise limit of the Duality Theorem for our SOCP yields that for the MKP, which is a generalization of [31] even when a diffusion matrix under consideration is an identity. In this paper we only consider the typical MKP which can be formally considered as the SOCP with fixed initial and terminal distributions and with a zero diffusion matrix.

Sect. 3 is devoted to the proof of Theorem 2.1 in Sect. 2.2.

In Sect. 4 we study the so-called Nelson's Problem as the SOCP with fixed marginal distributions at each time. In Sect. 4.1, we first state a positive answer to Nelson's Problem by the continuum limit of the Duality Theorem for the SOCP with fixed initial and terminal distributions. This also yields the SOCP with fixed marginal distributions at each time. In Sect. 4.2, we study the Duality Theorem for the SOCP with fixed marginal distributions at each time. This gives a new and simple approach to Nelson's Problem. In Sects. 4.3-4.4, we consider a class of deterministic variational problems with fixed marginal distributions, by extending a class of measures under consideration, which is related to the SOCP. In particular, we prove the existence and the uniqueness of a minimizer of and the Duality Theorem for them.

In the rest of this section, to discuss the MKP, we briefly describe Monge's Problem, Kantorovich's approach, the Duality Theorem, a formal derivation of a solution to Monge's Problem and the study in a one-dimensional case (see [11, 12, 15, 20, 26, 27, 29, 31, 36, 38-40] and the references therein).

1.1 The Monge-Kantorovich Problem.

The following problem is known as the origin of Monge's Problem (see [32]):

What is the best way to move a sand pile from one place to another ?

We discretize the problem to describe the mathematical formulation. For $n \geq 1$ and $2n$ different points $\{x_1, \cdots, x_n, y_1, \cdots, y_n\} \subset \mathbf{R}^3$, consider a bijection $\varphi : \{x_1, \cdots, x_n\} \mapsto \{y_1, \cdots, y_n\}$. Suppose that we have to pay the cost $|\varphi(x_i) - x_i|$ to move a (discretized) sand from x_i to $\varphi(x_i)$. Then the total cost is $\sum_{i=1}^n |\varphi(x_i) - x_i|$. To minimize this cost, we consider the following minimization problem:

$$\begin{aligned} & \inf\Big\{\sum_{i=1}^n |\varphi(x_i) - x_i| \Big| \{\varphi(x_1), \cdots, \varphi(x_n)\} = \{y_1, \cdots, y_n\}\Big\} \\ = \; & n \inf\Big\{\int_{\mathbf{R}^3} |\varphi(x) - x| \Big(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}(dx)\Big)\Big| \\ & \qquad \varphi_{\#}\Big(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}(dx)\Big) = \frac{1}{n}\sum_{i=1}^n \delta_{y_i}(dy)\Big\}, \end{aligned} \tag{1.1}$$

where $\delta_x(dy)$ denotes the delta measure on $\{x\}$ and $\varphi_{\#}P := P\varphi^{-1}$ for $P \in \mathcal{M}_1(\mathbf{R}^3)$:= the set of all Borel probability measures on $\mathbf{R}^3$, with a weak topology. More generally,
(**Monge's Problem**) Let $c(\cdot, \cdot) \in C(\mathbf{R}^d \times \mathbf{R}^d : [0, \infty))$. For P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$, study the following minimization problem:

$$T_M(P_0, P_1) := \inf\Big\{\int_{\mathbf{R}^d} c(x, \varphi(x)) P_0(dx) \Big| \varphi_{\#}P_0 = P_1\Big\}. \tag{1.2}$$

In Monge's Problem, $\int_{\mathbf{R}^d} c(x, \varphi(x)) P_0(dx)$ is nonlinear in φ, which makes the problem difficult. We describe a part of Kantorovich's idea to overcome this difficulty, from which the problem is called the Monge-Kantorovich Problem nowadays. (In Sect. 1.2 we explain how it works.) We easily obtain

Proposition 1.1 *For P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$\begin{aligned} T_M(P_0,P_1) &= \inf\Big\{\int_{\mathbf{R}^d\times\mathbf{R}^d} c(x,y)(Id\times\varphi)_{\#}P_0(dxdy)\Big|\varphi_{\#}P_0 = P_1\Big\} \\ &\geq \inf\Big\{\int_{\mathbf{R}^d\times\mathbf{R}^d} c(x,y)\mu(dxdy)\Big|\mu\in\mathcal{A}(P_0,P_1)\Big\} \\ &= : T_K(P_0,P_1), \end{aligned} \tag{1.3}$$

where $\mathcal{A}(P_0,P_1) := \{\mu \in \mathcal{M}_1(\mathbf{R}^{2d})|\mu(dx\times\mathbf{R}^d) = P_0(dx), \mu(\mathbf{R}^d\times dx) = P_1(dx)\}$.

Notice that $\int_{\mathbf{R}^d\times\mathbf{R}^d} c(x,y)\mu(dxdy)$ is linear in μ. Since $\mathcal{A}(P_0,P_1)$ is a compact subset of $\mathcal{M}_1(\mathbf{R}^{2d})$ with a weak topoplogy for P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ and since $c(\cdot,\cdot)\in C(\mathbf{R}^d\times\mathbf{R}^d : [0,\infty))$, it is easy to see that $T_K(P_0,P_1)$ has a minimizer, provided $T_K(P_0,P_1)$ is finite. This leads to
(**Kantorovich's Approach**) For any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ for which $T_K(P_0,P_1)$ is finite, prove the existence and the uniqueness of the minimizer μ of $T_K(P_0,P_1)$ for which

$$\mu(dxdy) = (Id\times\varphi)_{\#}P_0(dxdy)(= P_0(dx)\delta_{\varphi(x)}(dy)) \tag{1.4}$$

for some Borel measurable $\varphi : \mathbf{R}^d \mapsto \mathbf{R}^d$. If this is possible, then φ is the unique minimizer of $T_M(P_0,P_1)$.

1.2 Duality Theorem and the MKP.

In this section we discuss the so-called Duality Theorem for $T_K(P_0,P_1)$ which plays a crucial role in the study of the MKP. (Notice that $P \mapsto T_K(P_0,P)$ is convex and lower semicontinuous.) As an application, we give a formal derivation of a solution to the MKP.

Theorem 1.1 (Duality Theorem) *(see e.g. [20, 27] and also Sect. 2.3). For P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$\begin{aligned} & T_K(P_0,P_1) \\ = \ & \sup\Big\{\int_{\mathbf{R}^d}\varphi(1,y)P_1(dy) - \int_{\mathbf{R}^d}\varphi(0,x)P_0(dx)\Big| \\ & \qquad \varphi(t,\cdot)\in C_b(\mathbf{R}^d)(t=0,1), \varphi(1,y)-\varphi(0,x)\leq c(x,y)\Big\}. \end{aligned} \tag{1.5}$$

Remark 1.1 *It is easy to see that the l.h.s* $\geq$ *the r.h.s. in (1.5) since, for* $\mu \in \mathcal{A}(P_0, P_1)$ *and* $\varphi(t, \cdot) \in C_b(\mathbf{R}^d)(t = 0, 1)$,

$$\begin{aligned} & \int_{\mathbf{R}^d} \varphi(1, y) P_1(dy) - \int_{\mathbf{R}^d} \varphi(0, x) P_0(dx) \\ = & \int_{\mathbf{R}^d \times \mathbf{R}^d} (\varphi(1, y) - \varphi(0, x)) \mu(dxdy). \end{aligned} \tag{1.6}$$

Putting $T\varphi(y) := \inf\{c(x, y) + \varphi(x) | x \in \mathbf{R}^d\}$, (1.5) can be written as follows:

$$T_K(P_0, P_1) = \sup\left\{ \int_{\mathbf{R}^d} T\varphi(y) P_1(dy) - \int_{\mathbf{R}^d} \varphi(x) P_0(dx) \middle| \varphi \in C_b(\mathbf{R}^d) \right\}. \tag{1.7}$$

Indeed, if $\varphi(1, y) - \varphi(0, x) \leq c(x, y)$, then

$$\varphi(1, y) \leq T\varphi(0, y) \leq \varphi(0, x) + c(x, y). \tag{1.8}$$

Suppose that $c(x, y) = \ell(y - x)$ and

$$\begin{aligned} & u(t, y; \varphi) \\ := & \begin{cases} \inf\left\{ t\ell\left(\dfrac{y - x}{t}\right) + \varphi(x) \middle| x \in \mathbf{R}^d \right\} & ((t, y) \in (0, 1] \times \mathbf{R}^d), \\ \varphi(y) & ((t, y) \in \{0\} \times \mathbf{R}^d). \end{cases} \end{aligned} \tag{1.9}$$

Then $T\varphi(y) = u(1, y; \varphi)$. Suppose, in addition, that ℓ is strictly convex and $\ell(u)/|u| \to \infty$ as $|u| \to \infty$. Then for any bounded, uniformly Lipschitz continuous function φ, $u(t, y; \varphi)$ is a unique bounded, uniformly Lipschitz continuous viscosity solution of

$$\partial_t u(t, x; \varphi) + \ell^*(D_x u(t, x; \varphi)) = 0 \quad ((t, x) \in (0, 1] \times \mathbf{R}^d) \tag{1.10}$$

(see e.g. [10]), where $D_x := (\partial/\partial x_i)_{i=1}^d$, $\ell^*(z) := \sup\{< u, z > -\ell(u) | u \in \mathbf{R}^d\}$ and $< \cdot, \cdot >$ denotes the inner product in $\mathbf{R}^d$. Notice that

$$
\begin{aligned}
u(t,y;\varphi) &= \inf\Big\{\int_0^t \ell(\dot{x}(s))ds + \varphi(x(0))\Big| x(t)=y\Big\} \\
&= \sup\Big\{u(1,x;\varphi) - (1-t)\ell\Big(\frac{x-y}{1-t}\Big)\Big| x\in\mathbf{R}^d\Big\} \\
&= \sup\Big\{u(1,x(1);\varphi) - \int_t^1 \ell(\dot{x}(s))ds\Big| x(t)=y\Big\} \qquad (1.11)
\end{aligned}
$$

$((t,y)\in[0,1)\times\mathbf{R}^d)$ by Jensen's inequality, where $\dot{x}(t)$ denotes the derivative of $t\mapsto x(t)$ (see e.g. [10]).

For the readers' convenience, we give the definition of the viscosity solution to the PDE (1.10).

Definition 1.1 (Viscosity solution) *(see e.g. [10]).*
(Viscosity subsolution) $\varphi\in USC([0,1]\times\mathbf{R}^d)$ ($USC :=$ upper semicontinuous) is a viscosity subsolution of (1.10) if whenever $h\in C^{1,1}((0,1]\times\mathbf{R}^d)$ and $\varphi-h$ takes its maximum at $(s,y)\in(0,1]\times\mathbf{R}^d$,

$$\partial_s h(s,y) + \ell^*(D_y h(s,y)) \le 0. \qquad (1.12)$$

(Viscosity supersolution) $\varphi\in LSC([0,1]\times\mathbf{R}^d)$ ($LSC :=$ lower semicontinuous) is a viscosity supersolution of (1.10) if whenever $h\in C^{1,1}((0,1]\times\mathbf{R}^d)$ and $\varphi-h$ takes its minimum at $(s,y)\in(0,1]\times\mathbf{R}^d$,

$$\partial_s h(s,y) + \ell^*(D_y h(s,y)) \ge 0. \qquad (1.13)$$

(Viscosity solution) $\varphi\in C([0,1]\times\mathbf{R}^d)$ is a viscosity solution of (1.10) if it is both a viscosity subsolution and a viscosity supersolution of (1.10).

We formally explain how to derive a solution to the MKP from the Duality Theorem in the setting of (1.9) where $c(x,y)=\ell(y-x)$. It is known that there exists a maximizer $(\varphi(1,y),\varphi(0,x))$ in the Duality Theorem 1.1 (see [38]). For any minimizer μ of $T_K(P_0,P_1)$,

$$\varphi(1,y) = \varphi(0,x) + \ell(y-x) = \min\{\ell(y-z)+\varphi(0,z)|z\in\mathbf{R}^d\} \qquad (1.14)$$

$\mu(dxdy)-a.s.$, from (1.6)-(1.8). If ℓ and $\varphi(0,x)$ are differentiable, then

$$-D\ell(y-x)+D\varphi(0,x)=0 \quad \mu(dxdy)-a.s.. \tag{1.15}$$

Since $(D\ell)^{-1}=D\ell^*$ (see e.g.[40]), $\mu(dxdy)=P_0(dx)\delta_{x+D\ell^*(D\varphi(0,x))}(dy)$. More precisely, the following is known

Theorem 1.2 (see [15]) *Suppose that $c(x,y)=\ell(y-x)$, $\ell(u)=\ell(|u|)$, ℓ is strictly convex and $\ell(u)/|u|\to\infty$ as $|u|\to\infty$. Then for any P_0 and $P_1\in\mathcal{M}_1(\mathbf{R}^d)$ for which $P_0(dx)<<dx$ and $T_K(P_0,P_1)$ is finite, there exists a locally Lipschitz continuous fuction φ such that $T_K(P_0,P_1)$ has the unique minimizer $P_0(dx)\delta_{x+D\ell^*(D\varphi(x))}(dy)$.*

1.3 The MKP in a one-dimensional case.

In this section we consider the MKP in a one-dimensional case (see e.g. [36]). We refer the readers to [23] for an application to Markov optimal control problems. For any P_0 and $P_1\in\mathcal{M}_1(\mathbf{R})$,

$$\begin{aligned} F_i(x) &:= P_i((-\infty,x]) \quad (i=0,1,x\in\mathbf{R}), & (1.16)\\ F_1^{-1}(r) &:= \inf\{y\in\mathbf{R}|F_1(y)\geq r\} \quad (0<r<1). & (1.17)\end{aligned}$$

We give the proof to the following known fact for the readers' convenience.

Theorem 1.3 *Suppose that $d=1$ and that $\ell=|u|^p$ (resp. $\ell=-|u|^p$) ($p\geq 1$). Then for any P_0 and $P_1\in\mathcal{M}_1(\mathbf{R})$ for which $T_K(P_0,P_1)$ is finite, $T_K(P_0,P_1)$ has a (unique if $p>1$) minimizer $(F_0^{-1}\times F_1^{-1})_{\#}(du)$ (resp. $(F_0^{-1}\times(F_1^{-1}\circ(1-Id)))_{\#}(du)$), where we consider du only on $[0,1]$. In particular, if $F_0\in C(\mathbf{R})$, then $T_M(P_0,P_1)=T_K(P_0,P_1)$ and $T_M(P_0,P_1)$ has a (unique if $p>1$) minimizer $F_1^{-1}\circ F_0$ (resp. $F_1^{-1}\circ(1-F_0)$).*

(Proof) For $\mu\in\mathcal{A}(P_0,P_1)$ and x, $y\in\mathbf{R}$,

$$\begin{aligned} & |\{u\in[0,1]|F_0^{-1}(u)\leq x, F_1^{-1}(u)\leq y\}| \\ = \ & \min(F_0(x),F_1(y)) \\ \geq \ & \mu((-\infty,x]\times(-\infty,y]) \\ = \ & \mu((-\infty,x])-\mu((-\infty,x]\times(y,\infty)) \end{aligned}$$

$$
\begin{aligned}
&\geq \max(0, \mu((-\infty, x]) - \mu((y, \infty))) \\
&= \max(0, F_0(x) + F_1(y) - 1) \\
&= |\{u \in [0,1] | F_0^{-1}(u) \leq x, F_1^{-1}(1-u) \leq y\}|, \qquad (1.18)
\end{aligned}
$$

where $|A|$ denotes the Lebesgue measure of A for any measurable $A \subset \mathbf{R}$. If $F_0 \in C(\mathbf{R})$, then $(F_0)_{\#}P_0$ is uniformly distributed on $[0,1]$. Therefore the following completes the proof: for $p > 1$

$$
\begin{aligned}
&\int_{\mathbf{R}^2} |x-y|^p \mu(dxdy) \\
&= \int_{x \leq y} p(p-1)|x-y|^{p-2} \mu((-\infty, x] \times (y, \infty)) dxdy \\
&\quad + \int_{y \leq x} p(p-1)|x-y|^{p-2} \mu((x, \infty) \times (-\infty, y]) dxdy, \qquad (1.19)
\end{aligned}
$$

$$
\begin{aligned}
&\int_{\mathbf{R}^2} |x-y| \mu(dxdy) \\
&= \int_{\mathbf{R}} \{\mu((-\infty, x] \times (x, \infty)) + \mu((x, \infty) \times (-\infty, x])\} dx, \qquad (1.20)
\end{aligned}
$$

$$
\begin{aligned}
\mu((-\infty, x] \times (y, \infty)) &= P_0((-\infty, x]) - \mu((-\infty, x] \times (-\infty, y]), \qquad (1.21) \\
\mu((x, \infty) \times (-\infty, y]) &= P_1((-\infty, y]) - \mu((-\infty, x] \times (-\infty, y]). \square \qquad (1.22)
\end{aligned}
$$

Remark 1.2 *(i) When* $c(x,y) = \ell(y-x) = -|y-x|^p$ *(*$p \geq 1$*),*

$$
T_K(P_0, P_1) = -\sup\Big\{ \int_{\mathbf{R}^d \times \mathbf{R}^d} |y-x|^p \mu(dxdy) \Big| \mu \in \mathcal{A}(P_0, P_1) \Big\}. \qquad (1.23)
$$

(ii) If $F_1(\mathbf{R}) \cap (0,1) \not\subset F_0(\mathbf{R})$*, then* $(F_1^{-1} \circ F_0)_{\#} P_0 \neq P_1$*. Indeed, for* $r \in F_1(\mathbf{R}) \cap (0,1) \backslash F_0(\mathbf{R})$*,*

$$
\begin{aligned}
&P_0(\{x \in \mathbf{R} | F_1^{-1} \circ F_0(x) \leq F_1^{-1}(r)\}) = P_0(\{x \in \mathbf{R} | F_0(x) \leq r\}) \\
&< \quad r = F_1(F_1^{-1}(r)) \qquad (1.24)
\end{aligned}
$$

since $F_1^{-1}(\tilde{r}) > F_1^{-1}(r)$ *for* $r \in F_1(\mathbf{R}) \cap (0,1)$ *and* $\tilde{r} \in (r, 1]$*.*

2 Stochastic Optimal Control Problem.

Let $\ell : \mathbf{R}^d \mapsto [0, \infty)$ be convex. Then for any absolutely continuous function $\varphi : [0, 1] \mapsto \mathbf{R}^d$,

$$\ell(\varphi(1) - \varphi(0)) \le \int_0^1 \ell(\dot{\varphi}(t))dt \tag{2.1}$$

by Jensen's inequality, where the equality holds if $\dot{\varphi}(t) = \varphi(1) - \varphi(0)$. This implies that if $c(x, y) = \ell(y - x)$, then the following holds (see also (1.7)-(1.11)):

$$\begin{aligned} T_K(P_0, P_1) \;=\; \inf\Big\{ E\Big[\int_0^1 \ell(\dot{\varphi}(t))dt\Big] \Big| \varphi(t) \text{ is absolutely continuous a.s.,} \\ P\varphi(t)^{-1} = P_t(t = 0, 1)\Big\}. \end{aligned} \tag{2.2}$$

(When it is not confusing, we use the same notation P for different probability measures.) This implies that the MKP can or should be studied in the framework of the stochastic optimal control theory. In Sect. 2.1 we introduce the corresponding stochastic optimal control problem for which we state the Duality Theorem and its application in Sect. 2.2. This is a generalization of [30] where a diffusion matrix is an identity. The proof of Theorem 2.1 is given in Sect. 3. In Sect. 2.3 we show that the zero-noise limit of the Duality Theorem for the SOCP yields that for the MKP.

2.1 SOCP with fixed initial and terminal distributions.

Let $\sigma(t, x) = (\sigma_{ij}(t, x))_{i,j=1}^d$ ($(t, x) \in [0, 1] \times \mathbf{R}^d$) be a uniformly nondegenerate $d \times d$-matrix function for which each σ_{ij} is uniformly Lipschitz continuous in x uniformly in t. Let $\mathcal{A}$ denote the set of all $\mathbf{R}^d$-valued, continuous semimartingales $\{X(t)\}_{0 \le t \le 1}$ on a (possibly different) 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 $t \in [0, 1]$,
(ii) $X(t) = X(0) + \int_0^t \beta_X(s, X)ds + \int_0^t \sigma(s, X(s))dW_X(s)$ $(0 \le t \le 1)$.
Here $\mathcal{B}(C([0, t]))$, $\mathcal{B}(C([0, t]))_+$ and W_X denote the Borel σ-field of $C([0, t])$, $\cap_{s>t}\mathcal{B}(C([0, s]))$ and a $(\mathcal{F}_t^X)$-Brownian motion respectively, and $\mathcal{F}_t^X := \sigma[X(s) :$

$0 \le s \le t]$ (see e.g. [21]). Let $L : [0,1] \times \mathbf{R}^d \times \mathbf{R}^d \mapsto [0,\infty)$ be convex in u and be continuous. The following can be considered as the stochastic optimal control version of the MKP: for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$V(P_0,P_1) := \inf\Big\{E\Big[\int_0^1 L(t,X(t);\beta_X(t,X))dt\Big]\Big| \\ X \in \mathcal{A}, PX(t)^{-1} = P_t(t=0,1)\Big\}. \tag{2.3}$$

We explain why the set $\mathcal{A}$ is appropriate as the set over which the infimum is taken in our SOCP (2.3). 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 21]). For a $\mathbf{R}^d$-valued, $(\mathcal{F}_t)$-progressively measurable stochastic process $\{u(t)\}_{0\le t\le 1}$, consider the solution to the following:

$$X^u(t) = X_0 + \int_0^t u(s)ds + \int_0^t \sigma(s,X^u(s))dW(s) \quad (t\in[0,1]). \tag{2.4}$$

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

$$\beta_{X^u}(t,X^u) = E[u(t)|\mathcal{F}_t^{X^u}]. \tag{2.5}$$

(see Lemma 3.1 in Sect. 3). Besides, by Jensen's inequality,

$$E\Big[\int_0^1 L(t,X^u(t);u(t))dt\Big] \ge E\Big[\int_0^1 L(t,X^u(t);\beta_{X^u}(t,X^u))dt\Big]. \tag{2.6}$$

Remark 2.1 *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 the problems such as (2.3).*

In [30] where $\sigma(t,x)$ is an identity matrix, we proved the following, so-called, Duality Theorem for $V(P_0,P_1)$, as a stochastic optimal control counterpart of that for the MKP:

$$V(P_0, P_1) = \sup\Big\{\int_{\mathbf{R}^d} \varphi(1,x)P_1(dx) - \int_{\mathbf{R}^d} \varphi(0,x)P_0(dx)\Big\}, \tag{2.7}$$

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

$$\frac{\partial\varphi(t,x)}{\partial t} + \frac{1}{2}\sum_{i=1}^{d} \frac{\partial^2\varphi(t,x)}{\partial x_i^2} + H(t,x; D_x\varphi(t,x)) = 0 \tag{2.8}$$

($(t,x) \in [0,1) \times \mathbf{R}^d$). Here for $(t,x,z) \in [0,1] \times \mathbf{R}^d \times \mathbf{R}^d$,

$$H(t,x;z) := \sup\{< z, u > -L(t,x;u) | u \in \mathbf{R}^d\}. \tag{2.9}$$

In Sect. 2.2 we show that (2.7) with (2.8) replaced by

$$\frac{\partial\varphi(t,x)}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x) \frac{\partial^2\varphi(t,x)}{\partial x_i \partial x_j} + H(t,x; D_x\varphi(t,x)) = 0 \tag{2.10}$$

holds, where $a_{ij}(t,x) := \sum_{k=1}^{d} \sigma_{ik}(t,x)\sigma_{jk}(t,x)$. In the setting of [30] the HJB Eqn (2.8) has a unique classical solution. So does (2.10) in our setting only when $L = L_1(t,x) + L_2(t,u)$ (see Theorem 2.2 in Sect. 2.2). Otherwise the HJB Eqn (2.10) only has a bounded continuous viscosity solution, from which the proof of the Duality Theorem in [30] does not work. In this paper we construct and make use of a sequence of bounded continuous functions which approximate a minimal bounded continuous viscosity solution to the HJB Eqn (2.10) and which are viscosity solutions to the following HJB Eqns: for $n \geq 1$,

$$\frac{\partial\varphi_n(t,x)}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x) \frac{\partial^2\varphi_n(t,x)}{\partial x_i \partial x_j} + H_n(t,x; D_x\varphi_n(t,x)) = 0 \tag{2.11}$$

($(t,x) \in [0,1) \times \mathbf{R}^d$). Here for $(t,x,z) \in [0,1] \times \mathbf{R}^d \times \mathbf{R}^d$,

$$H_n(t,x;z) := \sup\{< z, u > -L(t,x;u) | u \in \mathbf{R}^d, |u| \leq n\} \tag{2.12}$$

(see Remark 2.2 in Sect. 2.2).

For the readers' convenience, we give the definition of the viscosity solution to the HJB Eqn (2.10).

Definition 2.1 (Viscosity solution) *(see e.g. [13]).*
(Viscosity subsolution) $\varphi \in USC([0,1] \times \mathbf{R}^d)$ *is a viscosity subsolution of (2.10) if whenever* $h \in C^{1,2}([0,1) \times \mathbf{R}^d)$ *and* $\varphi - h$ *takes its maximum at* $(s,y) \in [0,1) \times \mathbf{R}^d$,

$$\frac{\partial h(s,y)}{\partial s} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 h(s,y)}{\partial x_i \partial x_j} + H(s,y;D_x h(s,y)) \geq 0. \tag{2.13}$$

(Viscosity supersolution) $\varphi \in LSC([0,1] \times \mathbf{R}^d)$ *is a viscosity supersolution of (2.10) if whenever* $h \in C^{1,2}([0,1) \times \mathbf{R}^d)$ *and* $\varphi - h$ *takes its minimum at* $(s,y) \in [0,1) \times \mathbf{R}^d$,

$$\frac{\partial h(s,y)}{\partial s} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 h(s,y)}{\partial x_i \partial x_j} + H(s,y;D_x h(s,y)) \leq 0. \tag{2.14}$$

(Viscosity solution) $\varphi \in C([0,1] \times \mathbf{R}^d)$ *is a viscosity solution of (2.10) if it is both a viscosity subsolution and a viscosity supersolution of (2.10).*

2.2 Duality Theorem for the SOCP.

In this section we describe the Duality Theorem for the SOCP and its application.

We first describe the assumptions. The following is the assumption on the regularity of the diffusion matrix.
(A.0). (i) $\sigma(t,x) = (\sigma_{ij}(t,x))_{i,j=1}^d$ $((t,x) \in [0,1] \times \mathbf{R}^d)$ is a uniformly nondegenerate $d \times d$-matrix function, (ii) $\sigma_{ij} \in C_b^1([0,1] \times \mathbf{R}^d)$ $(i,j = 1, \cdots, d)$.

We describe the assumptions on L.
(A.1). (i) $L \in C([0,1] \times \mathbf{R}^d \times \mathbf{R}^d : [0,\infty))$, (ii) $u \mapsto L(t,x;u)$ is convex.
(A.2). There exists $\gamma > 1$ such that

$$\liminf_{|u| \to \infty} \frac{\inf\{L(t,x;u) : (t,x) \in [0,1] \times \mathbf{R}^d\}}{|u|^\gamma} > 0. \tag{2.15}$$

(A.3).

$$\Delta L(\varepsilon_1, \varepsilon_2) := \sup \frac{L(t,x;u) - L(s,y;u)}{1 + L(s,y;u)} \to 0 \quad \text{as } \varepsilon_1,\, \varepsilon_2 \to 0, \tag{2.16}$$

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 over all $u \in \mathbf{R}^d$.
(A.4). (i) $\partial L(t,x;u)/\partial t$ and $D_x L(t,x;u)$ are bounded on $[0,1] \times \mathbf{R}^d \times B_R$ for all $R > 0$, where $B_R := \{x \in \mathbf{R}^d | |x| \le R\}$, (ii) $\Delta L(0,\infty)$ is finite.

Remark 2.2 *(i) (A.1,i) and (A.4,ii) imply that L is bounded on $[0,1] \times \mathbf{R}^d \times B_R$ for all $R > 0$. (ii) For any $n \ge 1$, $f \in UC_b(\mathbf{R}^d)$ ($UC :=$ uniformly continuous) and $(t,x) \in [0,1] \times \mathbf{R}^d$,*

$$\begin{aligned} \varphi_n(t,x;f) \;\; := \;\; \sup\Big\{ E\Big[f(X(1)) - \int_t^1 L(s, X(s); \beta_X(s,X)) ds \Big] \Big| \\ X(t) = x, X \in \mathcal{A}_t, |\beta_X(s,X)| \le n \Big\}, \end{aligned} \tag{2.17}$$

where we define $\mathcal{A}_t$ in the same way as in (2.3). Then (A.0,ii), (A.1,i) and (A.4) imply that $\varphi_n(t,x;f)$ is a unique bounded continuous viscosity solution of the HJB Eqn (2.11) with $\varphi_n(1,x;f) = f(x)$ (see [13, p. 188, Corollary 7.1, p. 223, Corollary 3.1 and p.249, Theorem 9.1). We do not know if a bounded continuous viscosity solution of the HJB Eqn (2.10) with $\varphi(1,x) = f(x)$ is unique even if $f \in C_b^\infty(\mathbf{R}^d)$ (see Lemma 3.6 in Sect.3).

We give a result on the existence of a minimizer of $V(P_0, P_1)$ which can be proved by a standard argument and of which the proof is omitted (see [17, Proposition 2.1] or [30]).

Proposition 2.1 *Suppose that (A.0)-(A.3) 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 a generalization of [30] and the proof is given in Sect. 3.

Theorem 2.1 (Duality Theorem) *Suppose that (A.0)-(A.4) hold. Then, for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$V(P_0, P_1) = \sup\Big\{\int_{\mathbf{R}^d} \varphi(1,x)P_1(dx) - \int_{\mathbf{R}^d} \varphi(0,x)P_0(dx)\Big\} (\in [0,\infty]), \quad (2.18)$$

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

$$\frac{\partial \varphi(t,x)}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 \varphi(t,x)}{\partial x_i \partial x_j} + H(t,x; D_x\varphi(t,x)) = 0 \qquad (2.19)$$

$((t,x) \in [0,1) \times \mathbf{R}^d)$.

We introduce the following to replace φ in (2.18) by classical solutions to the HJB Eqn (2.19).
(A.5). (i) "σ is an identity", or "$\sigma_{ij} \in C_b^{1,2}([0,1] \times \mathbf{R}^d)$ $(i,j = 1, \cdots, d)$ and there exist functions L_1 and L_2 so that $L = L_1(t,x) + L_2(t,u)$". (ii) $L(t,x;u) \in C^1([0,1] \times \mathbf{R}^d \times \mathbf{R}^d : [0,\infty))$ and is strictly convex in u, (iii) $L \in C_b^{1,2,0}([0,1] \times \mathbf{R}^d \times B_R)$ for any $R > 0$.

Remark 2.3 (i) *Take $A_i \in C_b^1([0,1] \times \mathbf{R}^d)$ for which $\inf\{A_i(t,x) | (t,x) \in [0,1] \times \mathbf{R}^d\} > 0$ $(i = 1,2)$. If $L = A_1(t,x) + A_2(t,x)|u|^\gamma$ $(\gamma > 1)$, then (A.1)-(A.4) hold.* (ii) *(A.2) and (A.5,ii) imply that for any $(t,x) \in [0,1] \times \mathbf{R}^d$, $H(t,x;\cdot) \in C^1(\mathbf{R}^d)$ and for any u and $z \in \mathbf{R}^d$,*

$$z = D_u L(t,x;u) \quad \textit{if and only if} \quad u = D_z H(t,x;z). \qquad (2.20)$$

In addition, if $L(t,x;\cdot) \in C^2(\mathbf{R}^d)$, then

$$D_u^2 L(t,x;u) = D_z^2 H(t,x;z)^{-1} \quad \textit{if } u = D_z H(t,x;z) \qquad (2.21)$$

(see [40, 2.1.3]), where $D_u^2 := (\partial^2/\partial u_i \partial u_j)_{i,j=1}^d$.

If (A.0) and (A.5) hold, then the HJB Eqn (2.19) with terminal function $f \in C_b^3(\mathbf{R}^d)$ has a unique classical solution $\varphi(t,x;f)$ in $C_b^{1,2}([0,1] \times \mathbf{R}^d)$. In particular, $\varphi(t,x;f) = \varphi_n(t,x;f)$ for all $(t,x) \in [0,1] \times \mathbf{R}^d$, provided

$$n \geq \sup\{|D_z H(t,x; D_x\varphi(t,x;f))| | (t,x) \in [0,1] \times \mathbf{R}^d\}. \qquad (2.22)$$

This can be proved almost in the same way as in [13, p. 208, Lemma 11.3] (see also [13, pp. 169-170, Theorems 4.2 and 4.4]). Hence, in the same way as in [30, Theorem 2.1 and Corollary 2.1], we obtain the following of which the proof is omitted.

Theorem 2.2 *Suppose that (A.0)-(A.5) hold. Then the Duality Theorem (2.18) holds even if the supremum is taken over all bounded classical solutions φ of (2.19).*

Corollary 2.1 *Suppose that (A.0)-(A5) 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 $V(P_0, P_1)$, the following holds:*

$$\beta_X(t, X) = b_X(t, X(t)) := E[\beta_X(t, X)|(t, X(t))] \quad dtdPX(\cdot)^{-1} - a.e.. \tag{2.23}$$

Next we consider the case where (A.2) holds for $\gamma = 2$. We omit the proof of the following which can be obtained from Corollary 4.1 in Sect. 4 in the same way as in [30, Proposition 2.2].

Proposition 2.2 *Suppose that (A.0)-(A.5) hold and that $\gamma = 2$ in (A.2). Then for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ for which $V(P_0, P_1)$ is finite, the minimizer of $V(P_0, P_1)$ is unique and is Markovian.*

We introduce an additional assumption:
(A.6). For any $(t, x) \in [0, 1] \times \mathbf{R}^d$, $L(t, x; \cdot) \in C^2(\mathbf{R}^d)$. $D_u^2 L(t, x; u)$ is bounded and is uniformly nondegenerate on $[0, 1] \times \mathbf{R}^d \times \mathbf{R}^d$.

From Theorem 2.2, in the same way as in [30, Theorem 2.2], we can show that a minimizer of $V(P_0, P_1)$ satisfies a forward backward stochastic differential equation (FBSDE for short). We omit the proof.

Theorem 2.3 *Suppose that (A.0)-(A.1) and (A.3)-(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 $f(\cdot) \in L^1(\mathbf{R}^d, P_1(dx))$ and a $(\mathcal{F}_t^X)$-continuous semimartingale $\{Y(t)\}_{0\le t\le 1}$ such that $\{(X(t), Y(t), Z(t))\}_{0\le t\le 1}$ $(Z(t) := D_u L(t, X(t); b_X(t, X(t))))$ satisfies the following FBSDE: for $t \in [0, 1]$,*

$$
\begin{aligned}
X(t) &= X(0) + \int_0^t D_z H(s, X(s); Z(s))ds + \int_0^t \sigma(s, X(s))dW_X(s), \\
Y(t) &= f(X(1)) - \int_t^1 L(s, X(s); D_z H(s, X(s); Z(s)))ds \\
&\qquad - \int_t^1 < Z(s), \sigma(s, X(s))dW_X(s) > . \qquad (2.24)
\end{aligned}
$$

We consider h-path processes as an application of Theorem 2.3. We shall refer here to
(A.7). σ =identity. There exist functions $\xi \in C_b^{1,2}([0,1] \times \mathbf{R}^d : \mathbf{R}^d)$ and $c \in C_b^{1,2}([0,1] \times \mathbf{R}^d : [0,\infty))$ such that

$$
L(t,x;u) = \frac{1}{2}|u - \xi(t,x)|^2 + c(t,x) \quad ((t,x;u) \in [0,1] \times \mathbf{R}^d \times \mathbf{R}^d). \quad (2.25)
$$

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 [21] and also (2.4) for notation):

$$
\mathbf{X}(t) = X_0 + \int_0^t \xi(s, \mathbf{X}(s))ds + W(t) \quad (t \in [0,1]). \qquad (2.26)
$$

Then as a corollary to Theorem 2.3, we have

Corollary 2.2 *([30, Corollary 2.3]) Suppose that (A.7) holds. 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 $f_t \in L^1(\mathbf{R}^d, P_t(dx))$ $(t = 0, 1)$ such that the following holds: for any $A \in \mathcal{B}(C([0,1]))$,*

$$
\begin{aligned}
P(X(\cdot) \in A) &= E\Big[\exp\Big\{f_1(\mathbf{X}(1)) - f_0(\mathbf{X}(0)) \\
&\qquad - \int_0^1 c(t, \mathbf{X}(t))dt\Big\} : \mathbf{X}(\cdot) \in A\Big]. \qquad (2.27)
\end{aligned}
$$

Remark 2.4 *$\{X(t)\}_{0 \le t \le 1}$ in Corollary 2.2 is called the h-path process for $\{\mathbf{X}(t)\}_{0 \le t \le 1}$ with initial and terminal distributions P_0 and P_1. Corollary 2.2 is known (see e.g. [7, 14, 22, 33, 37, 41]).*

2.3 Zero-noise limit of the SOCP.

In this section we show that the zero-noise limit of the Duality Theorem for the SOCP yields that for the MKP.

When $a(t,x) = \varepsilon\times$identity ($\varepsilon > 0$), we write $V(P_0,P_1) = V_\varepsilon(P_0,P_1)$ and denote by $\mathcal{V}_\varepsilon(P_0,P_1)$ the right hand side of (2.18). Then the following $(\tilde{A})$ implies (A.1)-(A.4) and that $V_\varepsilon(P_0,P_1) = \mathcal{V}_\varepsilon(P_0,P_1)$ from Theorem 2.1.
$(\tilde{A})$ $L(\cdot) : \mathbf{R}^d \mapsto [0,\infty)$ is convex. $\liminf_{|u|\to\infty} L(u)/|u|^\gamma > 0$ for some $\gamma > 1$.

By $\mathcal{T}(P_0,P_1)$ and $*$, we denote the right hand side of (1.5) and the convolution of two measures respectively. Put also $g_\varepsilon(x) := (2\pi\varepsilon)^{-d}\exp(-|x|^2/(2\varepsilon))$. Then we have the following which generalizes [31] where we assumed that $L \in C^1(\mathbf{R}^d : [0,\infty))$ and is strictly convex.

Theorem 2.4 *Suppose that $(\tilde{A})$ holds. Then for any P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$\mathcal{T}(P_0,P_1) \le T_K(P_0,P_1) \le \liminf_{\varepsilon\to 0} V_\varepsilon(P_0, g_\varepsilon * P_1), \tag{2.28}$$

$$\mathcal{V}_\varepsilon(P_0, g_\varepsilon * P_1) \le \mathcal{T}(P_0,P_1) \quad (\varepsilon > 0). \tag{2.29}$$

In particular, $T_K(P_0,P_1) = \mathcal{T}(P_0,P_1)$.

(Proof) The first inequality in (2.28) can be proved from Remark 1.1 in Sect. 1. When $a(t,x) = \varepsilon\times$identity, for any $X \in \mathcal{A}$, by Jensen's inequality,

$$\begin{aligned} &\int_0^1 L(\beta_X(t,X))dt \\ \ge\ & L\Big(\int_0^1 \beta_X(t,X)dt\Big) = L(X(1) - X(0) - \sqrt{\varepsilon}W_X(1)), \end{aligned} \tag{2.30}$$

which implies the second inequality in (2.28) by Skhorohod's theorem and Fatou's lemma since $g_\varepsilon * P_1 \to P_1$ as $\varepsilon \to 0$ weakly.

Next we prove (2.29). When $a(t,x) = \varepsilon\times$identity and $(\tilde{A})$ holds, (A.0)-(A.4) hold. In particular, we can use Lemma 3.6 in Sect. 3. For $f \in C_b^\infty(\mathbf{R}^d)$, take $\varphi(\cdot,\cdot;f)$ defined by (3.17). For x, $y \in \mathbf{R}^d$, put $Q = \delta_x$ in (3.18) and

$$X^\varepsilon_{x,y}(t) := x + t(y-x) + \sqrt{\varepsilon}W(t) \tag{2.31}$$

(see (2.4) for notation). Then, from Lemma 3.6, $\varphi(0,\cdot;f) \in C_b(\mathbf{R}^d)$ and

$$E[f(y+\sqrt{\varepsilon}W(1))]-\varphi(0,x;f) = E[f(X^{\varepsilon}_{x,y}(1))]-\varphi(0,x;f) \leq L(y-x). \quad (2.32)$$

Since $E[f(\cdot+\sqrt{\varepsilon}W(1))] \in C_b(\mathbf{R}^d)$,

$$\begin{aligned} & \int_{\mathbf{R}^d} f(y) g_\varepsilon * P_1(dy) - \int_{\mathbf{R}^d} \varphi(0,x;f) P_0(dx) \\ = & \int_{\mathbf{R}^d} E[f(y+\sqrt{\varepsilon}W(1))] P_1(dy) - \int_{\mathbf{R}^d} \varphi(0,x;f) P_0(dx) \\ \leq & \mathcal{T}(P_0,P_1). \square \end{aligned} \quad (2.33)$$

3 Proof of Theorem 2.1.

In this section we prove Theorem 2.1. First we give technical temmas.

We prove (2.4)-(2.5) for the sake of completeness.

Lemma 3.1 *Suppose that (A.0) hold. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete filtered probability space, X_0 be a $(\mathcal{F}_0)$-adapted random variable, and $\{W(t)\}_{t\geq 0}$ denote a d-dimensional $(\mathcal{F}_t)$-Brownian motion for which $W(0)=0$. For a $\mathbf{R}^d$-valued, $(\mathcal{F}_t)$-progressively measurable stochastic process $\{u(t)\}_{0\leq t\leq 1}$, consider a solution to the following:*

$$X^u(t) = X_0 + \int_0^t u(s)ds + \int_0^t \sigma(s, X^u(s))dW(s) \quad (t\in[0,1]). \quad (3.1)$$

If $E[\int_0^1 |u(t)|dt]$ is finite, then $\{X^u(t)\}_{0\leq t\leq 1} \in \mathcal{A}$ with

$$\beta_{X^u}(t, X^u) = E[u(t)|\mathcal{F}_t^{X^u}], \quad dtdP-a.e.. \quad (3.2)$$

(Proof) Since $E[\int_0^1 |u(t)|dt]$ is finite, there exists a Borel measurable $\beta_{X^u} : [0,1]\times C([0,1]) \mapsto \mathbf{R}^d$ for which $\omega \mapsto \beta_{X^u}(t,\omega)$ is $\mathcal{B}(C([0,t]))_+$-measurable for all $t\in[0,1]$ and for which $\beta_{X^u}(t,X^u) = E[u(t)|\mathcal{F}_t^{X^u}]$ (see [21, pp. 114 and 270]). To complete the proof, we prove that $Y(t) := X^u(t) - X_0 - \int_0^t \beta_{X^u}(s,X^u)ds$ is a $(\mathcal{F}_t^{X^u})$-martingale with quadratic variational processes $(\int_0^t a_{ij}(s,X^u(s))ds)_{i,j=1}^d$. Indeed, if this is true, then the martingale representation theorem (see e.g. [16]) implies that

$$X^u(t) = X_0 + \int_0^t \beta_{X^u}(s, X^u)ds + \int_0^t \sigma(s, X^u(s))dW_{X^u}(s) \quad (0 \le t \le 1). \tag{3.3}$$

It is easy to see that $Y(t)$ is $\mathcal{F}_t^{X^u}$-adapted. Since $\mathcal{F}_s^{X^u} \subset \mathcal{F}_s$, for $t \ge s \ge 0$,

$$\begin{aligned}
& E[Y(t) - Y(s)|\mathcal{F}_s^{X^u}] \\
= \; & E\Big[\int_s^t (u(\gamma) - \beta_{X^u}(\gamma, X^u))d\gamma + \int_s^t \sigma(\gamma, X^u(\gamma))dW(\gamma)\Big|\mathcal{F}_s^{X^u}\Big] \\
= \; & \int_s^t E[u(\gamma) - E[u(\gamma)|\mathcal{F}_\gamma^{X^u}]|\mathcal{F}_s^{X^u}]d\gamma \\
& \quad + E\Big[E\Big[\int_s^t \sigma(\gamma, X^u(\gamma))dW(\gamma)\Big|\mathcal{F}_s\Big]\Big|\mathcal{F}_s^{X^u}\Big] = 0.
\end{aligned} \tag{3.4}$$

For any $f \in C_b^2(\mathbf{R})$ and $i, j = 1, \cdots, d$, by the Itô formula (see e.g. [16]),

$$\begin{aligned}
& E[f(Y_i(t) - Y_i(s))f(Y_j(t) - Y_j(s))|\mathcal{F}_s^{X^u}] \\
= \; & E\Big[\int_s^t \{f(Y_i(\gamma) - Y_i(s))f'(Y_j(\gamma) - Y_j(s))(u_j(\gamma) - \beta_{X^u,j}(\gamma, X^u)) \\
& \quad + f'(Y_i(\gamma) - Y_i(s))f(Y_j(\gamma) - Y_j(s))(u_i(\gamma) - \beta_{X^u,i}(\gamma, X^u)) \\
& \quad + a_{ij}(\gamma, X^u(\gamma))f'(Y_i(\gamma) - Y_i(s))f'(Y_j(\gamma) - Y_j(s)) \\
& \quad + \delta_{ij}a_{ij}(\gamma, X^u(\gamma))f(Y_i(\gamma) - Y_i(s))f''(Y_j(\gamma) - Y_j(s))\}d\gamma\Big|\mathcal{F}_s^{X^u}\Big] \\
= \; & E\Big[\int_s^t \{a_{ij}(\gamma, X^u(\gamma))f'(Y_i(\gamma) - Y_i(s))f'(Y_j(\gamma) - Y_j(s)) \\
& \quad + \delta_{ij}a_{ij}(\gamma, X^u(\gamma))f(Y_i(\gamma) - Y_i(s))f''(Y_j(\gamma) - Y_j(s))\}d\gamma\Big|\mathcal{F}_s^{X^u}\Big],
\end{aligned} \tag{3.5}$$

where $u(s) = (u_i(s))_{i=1}^d$, $\beta_{X^u}(s, X^u) = (\beta_{X^u,i}(s, X^u))_{i=1}^d$, $\delta_{ij} = 1$ if $i = j$ and $= 0$ if $i \ne j$. Indeed, since $Y(t)$ is $\mathcal{F}_t^{X^u}$-adapted, for $\gamma \ge s$,

$$\begin{aligned}
& E[f(Y_i(\gamma) - Y_i(s))f'(Y_j(\gamma) - Y_j(s))(u_j(\gamma) - \beta_{X^u,j}(\gamma, X^u))|\mathcal{F}_s^{X^u}] \\
= \; & E[E[f(Y_i(\gamma) - Y_i(s))f'(Y_j(\gamma) - Y_j(s)) \\
& \quad \times (u_j(\gamma) - E[u_j(\gamma)|\mathcal{F}_\gamma^{X^u}])|\mathcal{F}_\gamma^{X^u}]|\mathcal{F}_s^{X^u}] \\
= \; & E[f(Y_i(\gamma) - Y_i(s))f'(Y_j(\gamma) - Y_j(s)) \\
& \quad \times E[(u_j(\gamma) - E[u_j(\gamma)|\mathcal{F}_\gamma^{X^u}])|\mathcal{F}_\gamma^{X^u}]|\mathcal{F}_s^{X^u}] = 0.
\end{aligned} \tag{3.6}$$

On the set $\{\sup_{s\le\gamma\le t}|Y_j(\gamma)-Y_j(s)|\le n\}$ $(n\ge 1)$, taking f such that $f(x)=x$ if $|x|\le n$,

$$E[(Y_i(t)-Y_i(s))(Y_j(t)-Y_j(s))|\mathcal{F}_s^{X^u}]=E\Big[\int_s^t a_{ij}(\gamma,X^u(\gamma))d\gamma\Big|\mathcal{F}_s^{X^u}\Big]. \quad (3.7)$$

Since $\sup_{s\le\gamma\le t}|Y_j(\gamma)-Y_j(s)|$ is finite a.s., the proof is complete.□

The following two lemmas on the property of $V(\cdot,\cdot)$ will play a crucial role in the sequel and can be proved in the same way as in [30, Lemmas 3.1 and 3.2] from Lemmas 3.1 (see Remark 2.2, (i)). We omit the proof.

Lemma 3.2 *Suppose that (A.0)-(A.3) hold. Then $(Q,P)\mapsto V(Q,P)$ is lower semicontinuous.*

Lemma 3.3 *Suppose that (A.0)-(A.3) and (A.4,ii) hold. Then for any $P_0\in\mathcal{M}_1(\mathbf{R}^d)$, $P\mapsto V(P_0,P)$ is convex.*

We recall the following result.

Lemma 3.4 *(see (2.17) and [13, pp. 185-188]). Suppose that (A.0,ii) and (A.4) hold. Then for any $n\ge 1$, $f\in UC_b(\mathbf{R}^d)$, $t\in[0,1]$ and $Q\in\mathcal{M}_1(\mathbf{R}^d)$,*

$$\begin{aligned}\int_{\mathbf{R}^d}\varphi_n(t,x;f)Q(dx) &= \sup\Big\{E\Big[f(X(1))-\int_t^1 L(s,X(s);\beta_X(s,X))ds\Big]\Big| \\ &\qquad X\in\mathcal{A}_t,|\beta_X(s,X)|\le n,PX^{-1}(t)=Q\Big\} \quad (3.8)\end{aligned}$$

(A.0,ii) implies that for any $t\in[0,1]$, $X\in\mathcal{A}_t$ and $n\ge 1$, the following has a unique solution: for $u\in[t,1]$,

$$X_n(u)=X(t)+\int_t^u 1_{\{|\beta_X(s,X)|\le n\}}\beta_X(s,X)ds+\int_t^u\sigma(s,X_n(s))dW_X(s). \quad (3.9)$$

$X_n\in\mathcal{A}_t$ and $\beta_{X_n}(s,X_n)=E[1_{\{|\beta_X(s,X)|\le n\}}\beta_X(s,X)|\mathcal{F}_s^{X_n}]$ from Lemma 3.1. We also have

Lemma 3.5 *Suppose that (A.0,ii) holds. Then for any $t \in [0,1]$ and $X \in \mathcal{A}_t$, there exists a subsequence $\{X_{n(k)}\}_{k\geq 1} \subset \mathcal{A}_t$ (see (3.9)) such that*

$$\lim_{k\to\infty} \sup_{t\leq s\leq 1} |X_{n(k)}(s) - X(s)| = 0, \quad a.s.. \tag{3.10}$$

(Proof) For the sake of simplicity, we assume that $t = 0$. Putting

$$\tau_m := \inf\Big\{t > 0 \Big| \int_0^t |\beta_X(s,X)|ds > m\Big\} (\to \infty \text{ as } m \to \infty), \tag{3.11}$$

$$\lim_{n\to\infty} E\Big[\sup_{0\leq t\leq \min(1,\tau_m)} |X_n(t) - X(t)|^2\Big] = 0 \quad \text{for } m \geq 1. \tag{3.12}$$

This is true, since

$$\begin{aligned} X(t) - X_n(t) &= \int_0^t 1_{\{|\beta_X(s,X)|>n\}} \beta_X(s,X)ds \\ &\quad + \int_0^t (\sigma(s,X(s)) - \sigma(s,X_n(s)))dW_X(s), \end{aligned} \tag{3.13}$$

from which, by the Gronwall inequality and a standard argument, for $m \geq 1$,

$$\begin{aligned} & E\Big[\sup_{0\leq t\leq \min(1,\tau_m)} |X_n(t) - X(t)|^2\Big] \\ \leq\ & 2E\Big[\Big|\int_0^{\min(1,\tau_m)} 1_{\{|\beta_X(s,X)|>n\}} |\beta_X(s,X)|ds\Big|^2\Big] \exp(8C^2) \to 0 \end{aligned} \tag{3.14}$$

as $n \to \infty$ by the bounded convergence theorem. Here C denotes the Lipschitz constant of $\sigma(t,x)$ and we used the following:

$$\int_0^{\min(1,\tau_m)} 1_{\{|\beta_X(s,X)|>n\}} |\beta_X(s,X)|ds \leq m. \tag{3.15}$$

Since an L^2-convergent sequence of randam variables has an a.s. convergent subsequence, one can take from (3.12), by a diagonal method, a subsequence $\{n(k)\}_{k\geq 1}$ so that

$$\lim_{k\to\infty} \sup_{0\le t\le \min(1,\tau_m)} |X_{n(k)}(t) - X(t)| \to 0 \quad \text{for each } m \ge 1, \text{ a.s..} \tag{3.16}$$

Since $P(\cup_{m\ge1}\{\tau_m \ge 1\}) = 1$, the proof is complete$\Box$.

For $f \in UC_b(\mathbf{R}^d)$ and $(t,x) \in [0,1]\times\mathbf{R}^d$, since $n \mapsto \varphi_n(t,x;f)$ is nondecreasing (see (2.17)), we can define

$$\varphi(t,x;f) := \lim_{n\to\infty} \varphi_n(t,x;f). \tag{3.17}$$

We also have

Lemma 3.6 *Suppose that (A.0)-(A.4) hold. Then for any $f \in UC_b(\mathbf{R}^d)$, $Q \in \mathcal{M}_1(\mathbf{R}^d)$ and $t \in [0,1]$,*

$$\begin{aligned}\int_{\mathbf{R}^d} \varphi(t,x;f)Q(dx) \;&=\; \sup\Big\{E\Big[f(X(1)) - \int_t^1 L(s,X(s);\beta_X(s,X))ds\Big]\Big| \\ &\qquad X \in \mathcal{A}_t, PX^{-1}(t) = Q\Big\}.\end{aligned} \tag{3.18}$$

$\varphi(t,x;f)$ is a bounded continuous viscosity solution of (2.19). In addition, for any bounded continuous viscosity solution u of (2.19) with $u(1,x) = f(x)$, $u \ge \varphi$, that is, φ is minimal.

(Proof) We write $\varphi(t,x;f) = \varphi(t,x)$ and $\varphi_n(t,x;f) = \varphi_n(t,x)$ for the sake of simplicity. We first prove (3.18). It is easy to see that the left hand side is less than or equal to the right hand side in (3.18) from Lemma 3.4. Indeed,

$$\int_{\mathbf{R}^d} \varphi(t,x)Q(dx) = \lim_{n\to\infty} \int_{\mathbf{R}^d} \varphi_n(t,x)Q(dx) \tag{3.19}$$

by the bounded convergence theorem since, for any $n \ge 1$,

$$\sup_{x\in\mathbf{R}^d} |f(x)| \ge \varphi_n(t,x) \ge -\sup_{x\in\mathbf{R}^d} |f(x)| - \sup_{(t,x)\in[0,1]\times\mathbf{R}^d} L(t,x;0) > -\infty \tag{3.20}$$

(see Remark 2.2, (i)). We prove the opposite inequality. For $X \in \mathcal{A}_t$ for which $PX^{-1}(t) = Q$, take $\{X_{n(k)}\}_{k\ge1}$ in Lemma 3.5. Then for $k \ge 1$, from Lemma 3.4,

$$
\begin{aligned}
&\int_{\mathbf{R}^d} \varphi_{n(k)}(t,x)Q(dx) \\
\geq\ & E\Big[f(X_{n(k)}(1)) - \int_t^1 L(s, X_{n(k)}(s); \beta_{X_{n(k)}}(s, X_{n(k)}))ds\Big] \\
\geq\ & E\Big[f(X_{n(k)}(1)) - \int_t^1 L(s, X_{n(k)}(s); 1_{\{|\beta_X(s,X)|\leq n(k)\}}\beta_X(s,X))ds\Big] \\
\to\ & E\Big[f(X(1)) - \int_t^1 L(s, X(s); \beta_X(s,X))ds\Big] \quad \text{as } k\to\infty, \qquad (3.21)
\end{aligned}
$$

by Jensen's inequality and the dominated convergence theorem. Indeed, from (A.4,ii),

$$
\begin{aligned}
&L(s, X_n(s); 1_{\{|\beta_X(s,X)|\leq n\}}\beta_X(s,X)) \\
\leq\ & (1+\Delta L(0,\infty))(1 + L(s, X(s); 1_{\{|\beta_X(s,X)|\leq n\}}\beta_X(s,X))) \\
\leq\ & (1+\Delta L(0,\infty))(1 + L(s, X(s); \beta_X(s,X)) + L(s, X(s); 0)). \qquad (3.22)
\end{aligned}
$$

(3.19) and (3.21) complete the proof of (3.18).

(3.20) implies that φ is bounded.

We prove the upper semicontinuity of φ. For $(t,x)\in\mathbf{R}^d$ and any $n\geq 1$, take $X_{n,t,x}\in\mathcal{A}_t$ for which $X_{n,t,x}(t)=x$ and

$$
\begin{aligned}
&\varphi(t,x) - \frac{1}{n} \\
<\ & E\Big[f(X_{n,t,x}(1)) - \int_t^1 L(s, X_{n,t,x}(s); \beta_{X_{n,t,x}}(s, X_{n,t,x}))ds\Big]. \qquad (3.23)
\end{aligned}
$$

Put $X_{n,t,x}(u) := x$ for $u<t$. Then $\{X_{n,t,x}\}_{(t,x)\in[0,1]\times\mathbf{R}^d}$ is tight by (A.2) (see [17]) since, from (3.17), (3.20) and (3.23),

$$
\begin{aligned}
&E\Big[\int_t^1 L(s, X_{n,t,x}(s); \beta_{X_{n,t,x}}(s, X_{n,t,x}))ds] \\
\leq\ & 2\sup_{\overline{x}\in\mathbf{R}^d}|f(\overline{x})| + \sup_{(\overline{t},\overline{x})\in[0,1]\times\mathbf{R}^d} L(\overline{t},\overline{x};0) + 1 < \infty. \qquad (3.24)
\end{aligned}
$$

Hence, in the same way as in Lemma 3.2, from (A.3),

$$\limsup_{(t,x)\to(s,y)} \varphi(t,x) \le \varphi(s,y). \tag{3.25}$$

We prove the lower semicontinuity of φ. For (t,x), $(s,y) \in \mathbf{R}^d$ and $X_{s,y} \in \mathcal{A}_s$ for which $X_{s,y}(s) = y$, define $Y_{t,x} \in \mathcal{A}_t$ by the following: for $\alpha \in [t,1]$,

$$\begin{aligned} Y_{t,x}(\alpha) &= x + \int_t^\alpha 1_{[s,1]}(\gamma)\beta_{X_{s,y}}(\gamma, X_{s,y})d\gamma \\ &\quad + \int_t^\alpha 1_{[s,1]}(\gamma)\sigma(\gamma, Y_{s,y}(\gamma))dW_{X_{s,y}}(\gamma) \\ &\quad + \int_t^\alpha 1_{[0,s]}(\gamma)\sigma(\gamma, Y_{s,y}(\gamma))d\tilde{W}_(\gamma), \end{aligned} \tag{3.26}$$

where $\{\tilde{W}(\gamma)\}_{0\le\gamma\le s}$ is a Brownian motion which is independent of $\{W_{X_{s,y}}(\gamma) - W_{X_{s,y}}(s)\}_{s\le\gamma\le 1}$ This is possible from (A.0,ii). In the same way as in (3.21), taking a subsequence if necessary,

$$\begin{aligned} &\liminf_{(t,x)\to(s,y)} \varphi(t,x) \\ \ge\ &\liminf_{(t,x)\to(s,y)} E\Big[f(Y_{t,x}(1)) - \int_t^1 L(\alpha, Y_{t,x}(\alpha); \beta_{Y_{t,x}}(\alpha, Y_{t,x}))d\alpha\Big] \\ \ge\ &E\Big[f(X_{s,y}(1)) - \int_s^1 L(\alpha, X_{s,y}(\alpha); \beta_{X_{s,y}}(\alpha, X_{s,y}))d\alpha\Big]. \end{aligned} \tag{3.27}$$

In particular,

$$\liminf_{(t,x)\to(s,y)} \varphi(t,x) \ge \varphi(s,y). \tag{3.28}$$

We prove that φ is a viscosity subsolution of (2.19). Suppose that $h \in C^{1,2}([0,1)\times\mathbf{R}^d)$ and $\varphi - h$ takes its maximum, say, 0 at $(s,y) \in [0,1)\times\mathbf{R}^d$. Then $\varphi(t,x) - h(t,x) - (|t-s|^2 + |x-y|^2)/(2m)$ takes its strict maximum at $(s,y) \in [0,1)\times\mathbf{R}^d$ for any $m \ge 1$. It is easy to see that $\varphi_n(t,x) - h(t,x) - (|t-s|^2 + |x-y|^2)/(2m)$ takes its maximum at some point $(s_n, y_n) \in [0,1)\times\mathbf{R}^d$ since

$$\begin{aligned} &\varphi_n(t,x) - h(t,x) = \varphi_n(t,x) - \varphi(t,x) + \varphi(t,x) - h(t,x) \\ \le\ &\varphi_n(t,x) - \varphi(t,x) \le 2\sup_{\overline{x}\in\mathbf{R}^d} |f(\overline{x})| + \sup_{(\overline{t},\overline{x})\in[0,1]\times\mathbf{R}^d} L(\overline{t},\overline{x};0) < \infty \end{aligned} \tag{3.29}$$

from (3.17) and (3.20). Since φ_n is a bounded continuous viscosity solution of the HJB Eqn (2.11),

$$\frac{\partial h(s_n,y_n)}{\partial s}+\frac{s_n-s}{m}+\frac{1}{2}\sum_{i,j=1}^{d}a_{ij}(s_n,y_n)\frac{\partial^2 h(s_n,y_n)}{\partial x_i\partial x_j}$$
$$+\frac{1}{2m}\sum_{i=1}^{d}a_{ii}(s_n,y_n)+H_n(s_n,y_n;D_xh(s_n,y_n)+\frac{y_n-y}{m})\geq 0. \quad (3.30)$$

Since $H_n\uparrow H$ (resp. $\varphi_n\uparrow\varphi$) as $n\to\infty$ and H_n and H (resp. φ_n and φ) are continuous, $H_n\uparrow H$ (resp. $\varphi_n\uparrow\varphi$) uniformly on every compact subset of $[0,1]\times\mathbf{R}^d\times\mathbf{R}^d$ (resp. $[0,1]\times\mathbf{R}^d$) as $n\to\infty$ by Dini's Theorem. $(s_n,y_n)\to(s,y)$ as $n\to\infty$. Indeed,

$$\varphi_n(s_n,y_n)-h(s_n,y_n)-\frac{|s_n-s|^2+|y_n-y|^2}{2m}\geq\varphi_n(s,y)-h(s,y), \quad (3.31)$$

which together with (3.17) and (3.29) implies the boundedness of $\{(s_n,y_n)\}_{n\geq1}$. For any convergent subsequence of $\{(s_n,y_n)\}_{n\geq1}$ and its limit $(\overline{s},\overline{y})$, taking the limit in (3.31),

$$\varphi(\overline{s},\overline{y})-h(\overline{s},\overline{y})-\frac{|\overline{s}-s|^2+|\overline{y}-y|^2}{2m}\geq\varphi(s,y)-h(s,y), \quad (3.32)$$

which implies $(\overline{s},\overline{y})=(s,y)$. Let $n\to\infty$ and then $m\to\infty$ in (3.30). Then

$$\frac{\partial h(s,y)}{\partial s}+\frac{1}{2}\sum_{i,j=1}^{d}a_{ij}(s,y)\frac{\partial^2 h(s,y)}{\partial x_i\partial x_j}+H(s,y;D_xh(s,y))\geq 0. \quad (3.33)$$

We can prove that φ is a viscosity supersolution of (2.19) in the same way as in (3.29)-(3.33), by considering a function $\varphi(t,x)-h(t,x)+(|t-s|^2+|x-y|^2)/(2m)$ instead of $\varphi(t,x)-h(t,x)-(|t-s|^2+|x-y|^2)/(2m)$.

We prove that φ is a minimal bounded continuous viscosity solution of (2.19) with $\varphi(1,x)=f(x)$. Let u be a bounded continuous viscosity solution of (2.19) with $u(1,x)=f(x)$. Then u is a bounded continuous viscosity supersolution of (2.11) with $u(1,x)=f(x)$ for all $n\geq1$ since $H_n\leq H$. Since

φ_n is a bounded continuous viscosity subsolution of (2.11) with $\varphi_n(1,x) = f(x)$, by the comparison principle (see [13, p. 249, Theorem 9.1]), $\varphi_n \le u$. Letting $n \to \infty$, the proof is over from (3.17).□

To prove Theorem 2.1, we have to improve the idea in [30, Theorem 2.1]. For the sake of completeness, we write the whole proof.

(Proof of Theorem 2.1). $V(P_0,\cdot) \not\equiv \infty$. Indeed, for $P_1 = P(X^0(1))^{-1}$ (see (2.4) for notation), from (A.4,ii) and Remark 2.2, (i),

$$V(P_0,P_1) \le \sup\{L(t,x;0) | (t,x) \in [0,1]\times \mathbf{R}^d\} < \infty. \tag{3.34}$$

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 3.2 and 3.3 and [9, Theorem 2.2.15 and Lemma 3.2.3],

$$V(P_0,P_1) = \sup\Big\{\int_{\mathbf{R}^d} f(x)P_1(dx) - V^*_{P_0}(f)\Big| f \in C_b(\mathbf{R}^d)\Big\}, \tag{3.35}$$

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

$$V^*_{P_0}(f) := \sup\Big\{\int_{\mathbf{R}^d} f(x)P(dx) - V(P_0,P)\Big| P \in \mathcal{M}_1(\mathbf{R}^d)\Big\}. \tag{3.36}$$

Denote by $\mathcal{V}(P_0,P_1)$ the right hand side of (2.18). Then, from Lemma 3.6 and (3.35), $V(P_0,P_1) \ge \mathcal{V}(P_0,P_1)$.

We prove the opposite inequality. Take $\rho \in C_0^\infty([-1,1]^d : [0,\infty))$ for which $\int_{\mathbf{R}^d}\rho(x)dx = 1$. For $\varepsilon > 0$ and $f \in C_b(\mathbf{R}^d)$, put

$$\rho_\varepsilon(x) := \varepsilon^{-d}\rho(x/\varepsilon), \quad f_\varepsilon(x) := \int_{\mathbf{R}^d} f(y)\rho_\varepsilon(y-x)dy. \tag{3.37}$$

Then $f_\varepsilon \in C_b^\infty(\mathbf{R}^d)$ and, from Lemma 3.6,

$$\mathcal{V}(P_0,P_1) \ge \int_{\mathbf{R}^d} f_\varepsilon(x)P_1(dx) - V^*_{P_0}(f_\varepsilon). \tag{3.38}$$

Take $X_\varepsilon \in \mathcal{A}$ ($\varepsilon > 0$) for which $PX_\varepsilon(0)^{-1} = P_0$ and

$$V^*_{P_0}(f_\varepsilon) - \varepsilon < E[f_\varepsilon(X_\varepsilon(1))] - E\Big[\int_0^1 L(t,X_\varepsilon(t);\beta_{X_\varepsilon}(t,X_\varepsilon))dt\Big]. \tag{3.39}$$

Then $\{X_\varepsilon\}_{\varepsilon\in(0,1)}$ is tight from (A.2) and (3.39) (see [17]) since

$$V^*_{P_0}(f_\varepsilon) \geq -\sup_{x\in\mathbf{R}^d}|f(x)| - \sup_{(t,x)\in[0,1]\times\mathbf{R}^d} L(t,x;0) > -\infty \tag{3.40}$$

from Remark 2.2, (i). From (3.39), in the same way as in Lemma 3.2, there exists a weak limit point $\overline{X}$ of X_ε as $\varepsilon\to 0$ such that

$$\begin{aligned}\limsup_{\varepsilon\to 0} V^*_{P_0}(f_\varepsilon) &\leq E[f(\overline{X}(1))] - E\Big[\int_0^1 L(t,\overline{X}(t);\beta_{\overline{X}}(t,\overline{X}))dt\Big] \\ &\leq V^*_{P_0}(f) \end{aligned}\tag{3.41}$$

since $E[f_\varepsilon(X_\varepsilon(1))] = \int_{\mathbf{R}^d}\rho(z)dzE[f(X_\varepsilon(1)+\varepsilon z)]$. (3.35), (3.38) and (3.41) imply that $\mathcal{V}(P_0,P_1)\geq V(P_0,P_1)$.□

4 Nelson's Problem.

In this section we consider Nelson's Problem under the generalized finite energy condition. We describe Nelson's Problem. Let $b:[0,1]\times\mathbf{R}^d\mapsto\mathbf{R}^d$ be measurable and $\{P_t\}_{0\leq t\leq 1}\subset\mathcal{M}_1(\mathbf{R}^d)$ satisfy the following Fokker-Planck equation: for any $f\in C_b^{1,2}([0,1]\times\mathbf{R}^d)$ and $t\in[0,1]$,

$$\begin{aligned}&\int_{\mathbf{R}^d} f(t,x)P_t(dx) - \int_{\mathbf{R}^d} f(0,x)P_0(dx) \\ =\ & \int_0^t ds\int_{\mathbf{R}^d}\Big(\frac{\partial f(s,x)}{\partial s} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(s,x)\frac{\partial^2 f(s,x)}{\partial x_i\partial x_j} \\ &\qquad + <b(s,x), D_x f(s,x)>\Big)P_s(dx).\end{aligned}\tag{4.1}$$

Inspired by Born's probabilistic interpretation of a solution to Schrödinger's equation, in the case where $a(t,x)=(a_{ij}(t,x))_{i,j=1}^d$ is an identity matrix, Nelson proposed the problem of the construction of a diffusion process $\{X(t)\}_{0\leq t\leq 1}$ for which the following holds (see [34, 35]): for $t\in[0,1]$,

$$X(t) = X(0) + \int_0^t b(s,X(s))ds + \int_0^t \sigma(s,X(s))dW_X(s), \tag{4.2}$$

$$PX(t)^{-1} = P_t. \tag{4.3}$$

The first result was given by Carlen [2] (see also [42]) where σ is an identity matrix. It was generalized, by Mikami [22], to the case where σ is a matrix function. The further generalization and almost complete resolution was made by Cattiaux and Léonard [3-6] (see also [1, 24-25] for the related topics). In these papers, they assumed that

$$\int_0^1 dt \int_{\mathbf{R}^d} |b(t,x)|^2 P_t(dx) < \infty \tag{4.4}$$

for some b for which (4.1) holds. This is called the **finite energy condition** for $\{P_t\}_{0\le t\le 1}$.

Remark 4.1 *It is known that b is not unique for $\{P_t\}_{0\le t\le 1}$ in (4.1) (see [22] or [3-6]). If (4.1) holds, then we will write $b \in \mathbf{A}(\{P_t\}_{0\le t\le 1})$.*

In [28] where σ is an identity matrix, we considered Nelson's Problem under a weaker assumption than (4.4): there exists $\gamma > 1$ such that

$$\int_0^1 dt \int_{\mathbf{R}^d} |b(t,x)|^\gamma P_t(dx) < \infty \tag{4.5}$$

for some $b \in \mathbf{A}(\{P_t\}_{0\le t\le 1})$. We call (4.5) the **generalized finite energy condition** (GFEC for short) for $\{P_t\}_{0\le t\le 1}$.

In Sect. 4.1 we study Nelson's Problem under the GFEC when σ is not an identity matrix as an application of Theorem 2.2. In Sect. 4.2 we study the Duality Theorem for Nelson's Problem which gives a direct approach to Nelson's Problem. In Sects. 4.3-4.4 we study the existence and the uniqueness of a minimizer of and the Duality Theorem for deterministic variational problems for $\{P_t\}_{0\le t\le 1}$ in (4.1) (see (4.6) and (4.9)). Assumptions (A.0)-(A.5) can be found in Sect. 2.2.

4.1 Nelson's Problem under the GFEC.

For P_0 and $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned} v(P_0,P_1) \;:=\; \inf\Big\{ \int_0^1 dt \int_{\mathbf{R}^d} L(t,x;b(t,x)) Q_t(dx) \Big| \{Q_t\}_{0\le t\le 1} \subset \mathcal{M}_1(\mathbf{R}^d), \\ Q_t = P_t (t=0,1), b \in \mathbf{A}(\{Q_t\}_{0\le t\le 1}) \Big\}. \end{aligned} \tag{4.6}$$

From Theorem 2.2, in the same way as in [28, Theorem 2.1], we obtain the following of which the proof is omitted.

Corollary 4.1 *Suppose that (A.0)-(A.5) hold. Then for any* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$V(P_0, P_1) = v(P_0, P_1)(\in [0, \infty]). \tag{4.7}$$

For $\mathbf{P} := \{P_t\}_{0 \le t \le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned} \mathbf{V}(\mathbf{P}) \;:=\; \inf\Big\{ E\Big[\int_0^1 L(t, X(t); \beta_X(t, X))dt\Big] \Big| X \in \mathcal{A}, \\ PX(t)^{-1} = P_t (0 \le t \le 1) \Big\}, \end{aligned} \tag{4.8}$$

$$\mathbf{v}(\mathbf{P}) := \inf\Big\{ \int_0^1 dt \int_{\mathbf{R}^d} L(t, x; b(t, x)) P_t(dx) \Big| b \in \mathbf{A}(\mathbf{P}) \Big\}. \tag{4.9}$$

Using a similar result to (4.7) on small time intervals $\subset [0, 1]$ and then taking the continuum limit, in the same way as in [28, Theorem 2.2], we have the following which is omitted the proof.

Theorem 4.1 *Suppose that (A.0)-(A.5) hold. Then*
(i) for any $\mathbf{P} := \{P_t\}_{0 \le t \le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$,

$$\mathbf{V}(\mathbf{P}) = \mathbf{v}(\mathbf{P})(\in [0, \infty]). \tag{4.10}$$

(ii) For any $\mathbf{P} := \{P_t\}_{0 \le t \le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$ *for which* $\mathbf{v}(\mathbf{P})$ *is finite, there exist a unique minimizer* b_o *of* $\mathbf{v}(\mathbf{P})$ *and a minimizer* X *of* $\mathbf{V}(\mathbf{P})$. *In particular, for any minimizer* X *of* $\mathbf{V}(\mathbf{P})$,

$$\beta_X(t, X) = b_o(t, X(t)) \tag{4.11}$$

and (4.2)-(4.3) with $b = b_o$ *hold.*

Remark 4.2 *(i) If* $\mathbf{v}(\mathbf{P})$ *is finite, then the GFEC (4.5) holds from (A.2). (ii) When (4.4) holds, the semimartingale in (4.2) is Markovian. But we do not know if it is also true even when* $\gamma < 2$. *This is our future problem.*

4.2 Duality Theorem for Nelson's Problem.

Theorem 2.2 and Corollary 4.1 implies the Duality Theorem for $v(P_0, P_1)$.

Corollary 4.2 *Suppose that (A.0)-(A.5) hold. Then for any* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$v(P_0, P_1) = \sup\Big\{\int_{\mathbf{R}^d} \varphi(1,x)P_1(dx) - \int_{\mathbf{R}^d} \varphi(0,x)P_0(dx)\Big\} (\in [0,\infty]), \quad (4.12)$$

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

$$\frac{\partial \varphi(t,x)}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 \varphi(t,x)}{\partial x_i \partial x_j} + H(t,x; D_x\varphi(t,x)) = 0 \qquad (4.13)$$

$((t,x) \in [0,1) \times \mathbf{R}^d)$.

In this section, for $\mathbf{P} := \{P_t\}_{0\le t\le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$, we first prove the Duality Theorem for $\mathbf{V}(\mathbf{P})$ which gives a new direct proof to Theorem 4.1. As a corollary, we obtain the Duality Theorem for $\mathbf{v}(\mathbf{P})$ in the same way as in Corollary 4.2.

We extend $\mathbf{V}(\mathbf{P})$ to a variational problem on $\mathcal{M}_1([0,1] \times \mathbf{R}^d)$. For $\lambda \in \mathcal{M}_1([0,1] \times \mathbf{R}^d)$, if $t \mapsto \lambda(t, dx) := \lambda(dtdx)/dt$ exists and has a weakly continuous version, we denote it by $\lambda_t(dx)$.

Fix $P_0 \in \mathcal{M}_1(\mathbf{R}^d)$. For $\lambda \in \mathcal{M}_1([0,1] \times \mathbf{R}^d)$,

$$\overline{\mathbf{V}}_{P_0}(\lambda) := \begin{cases} \mathbf{V}(\{\lambda_t\}_{0\le t\le 1}) & \text{if } \{\lambda_t\}_{0\le t\le 1} \text{ exists and } \lambda_0 = P_0, \\ \infty & \text{otherwise.} \end{cases} \qquad (4.14)$$

Then it is easy to see that $\overline{\mathbf{V}}_{P_0}(\cdot)$ is convex, lower semi-continuous and not identically equal to ∞ on $\mathcal{M}_1([0,1] \times \mathbf{R}^d)$ under (A.0)-(A.3) and (A.4,ii). Indeed, the convexity can be proved in the same way as in Lemma 3.2. To prove the lower semicontinuity, suppose that $\lambda^n(dtdx) \to \lambda(dtdx)$ as $n \to \infty$ weakly and that $\liminf_{n\to\infty} \overline{\mathbf{V}}_{P_0}(\lambda^n)$ is finite. Then there exist $\{n(k)\}_{k\ge 1}$, $\{X_{n(k)}\}_{k\ge 1}$ and $X \in \mathcal{A}$ for which $PX_{n(k)}(0)^{-1} = P_0$, $PX_{n(k)}(t)^{-1} = \lambda_t^{n(k)}$ for all $t \in [0,1]$ $(k \ge 1)$, $X_{n(k)} \to X$ as $k \to \infty$ weakly and

$$\begin{aligned}\liminf_{n\to\infty}\overline{\mathbf{V}}_{P_0}(\lambda^n) &= \lim_{k\to\infty}E\Big[\int_0^1 L(t,X_{n(k)}(t);\beta_{X_{n(k)}}(t,X_{n(k)}))dt\Big]\\ &\geq E\Big[\int_0^1 L(t,X(t);\beta_X(t,X))dt\Big] \qquad (4.15)\end{aligned}$$

in the same way as in Proposition 2.1. Since $dtPX_n(t)^{-1}(dx)=\lambda^n(dtdx)$, it is easy to see that $\{\lambda_t\}_{0\leq t\leq 1}$ exists, $\lambda_0=P_0$, $PX(t)^{-1}=\lambda_t$ $(0\leq t\leq 1)$ and

$$\liminf_{n\to\infty}\overline{\mathbf{V}}_{P_0}(\lambda^n)\geq\overline{\mathbf{V}}_{P_0}(\lambda). \qquad (4.16)$$

For $f\in C_b^\infty([0,1]\times\mathbf{R}^d)$, let $\phi(t,x;f)$ denote a function $\varphi(t,x;0)$ in (3.17) with $H(t,x,;z)$ replaced by $H(t,x,;z)+f(t,x)$. Then under (A.0)-(A.4), from Lemma 3.6, it is the minimal bounded continuous viscosity solution of the following HJB Eqn:

$$\begin{aligned}&\frac{\partial\phi(t,x;f)}{\partial t}+\frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2\phi(t,x;f)}{\partial x_i\partial x_j}\\ &\qquad +H(t,x;D_x\phi(t,x;f))+f(t,x) = 0 \quad ((t,x)\in[0,1)\times\mathbf{R}^d) \qquad (4.17)\\ &\qquad \phi(1,x;f) = 0 \quad (x\in\mathbf{R}^d). \qquad (4.18)\end{aligned}$$

In the same way as in Theorem 2.1, we obtain the Duality Theorem for $\mathbf{V}(\mathbf{P})$.

Theorem 4.2 *Suppose that (A.0)-(A.4) hold. Then for any* $\mathbf{P}:=\{P_t\}_{0\leq t\leq 1}\subset\mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned}\mathbf{V}(\mathbf{P}) = \sup\Big\{&\Big|\int_{[0,1]\times\mathbf{R}^d}f(t,x)dtP_t(dx)-\int_{\mathbf{R}^d}\phi(0,x;f)P_0(dx)\Big|\\ &f\in C_b^\infty([0,1]\times\mathbf{R}^d)\Big\}(\in[0,\infty]). \qquad (4.19)\end{aligned}$$

If (A.0) and (A.5) hold, then for $f\in C_b^\infty([0,1]\times\mathbf{R}^d)$, the HJB Eqn (4.17)-(4.18) has a unique classical solution $\phi(t,x;f)$ in $C_b^{1,2}([0,1]\times\mathbf{R}^d)$. This can be proved almost in the same way as in [13, p. 208, Lemma 11.3] (see also [13, pp. 169-170, Theorems 4.2 and 4.4]). Hence Theorem 4.2 implies Theorem 4.1 without Corollary 4.1 in the same way as in Theorem 2.2. In the same way as in Corollaries 4.1 and 4.2, we obtain the Duality Theorem for $\mathbf{v}(\mathbf{P})$.

Corollary 4.3 *Suppose that (A.0)-(A.5) hold. Then for any* $\mathbf{P} := \{P_t\}_{0\le t\le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned}\mathbf{v}(\mathbf{P}) &= \mathbf{V}(\mathbf{P}) \\ &= \sup\Big\{\Big|\int_{[0,1]\times\mathbf{R}^d} f(t,x)dtP_t(dx) - \int_{\mathbf{R}^d}\phi(0,x;f)P_0(dx)\Big| \\ &\qquad\qquad f\in C_b^\infty([0,1]\times\mathbf{R}^d)\Big\}(\in[0,\infty]), \end{aligned} \tag{4.20}$$

where for $f \in C_b^\infty([0,1]\times\mathbf{R}^d)$, $\phi(t,x;f)$ *is the unique bounded classical solution to (4.17)-(4.18).*

4.3 Minimizers of $v(P_0,P_1)$ and $\mathbf{v}(\mathbf{P})$.

In this section, inspired by Kantorovich's approach, we introduce a new idea to study minimizers of $v(P_0,P_1)$ and $\mathbf{v}(\mathbf{P})$ defined in Sect. 4.1.

We use (A.1) and the following, instead of (A.0) and (A.2).

(A.0)'. $a(t,x) = (a_{ij}(t,x))_{i,j=1}^d$ in (4.1) is bounded, uniformly Lipschitz continuous in x uniformly in t and nonnegative definite.

(A.2)'.

$$\liminf_{|u|\to\infty} \frac{\inf\{L(t,x;u) : (t,x)\in[0,1]\times\mathbf{R}^d\}}{|u|} > 0. \tag{4.21}$$

In particular, we can consider the case where $a(t,x)$ is a zero matrix and the PDE (4.1) becomes the Liouville equation.

For $\mu(dxdv)\in\mathcal{M}_1(\mathbf{R}^d\times\mathbf{R}^d)$, $\mu_1(dx) := \mu(dx\times\mathbf{R}^d)$, $\mu_2(dv) := \mu(\mathbf{R}^d\times dv)$. We write $\nu(dtdxdv)\in\tilde{\mathcal{A}}$ if the following holds: $\nu(dtdxdv)\in\mathcal{M}_1([0,1]\times\mathbf{R}^d\times\mathbf{R}^d)$, $\nu(t,dxdv) := \nu(dtdxdv)/dt$ exists and $t\mapsto\nu_1(t,dx)$ has a weakly continuous version $\nu_{1,t}(dx)$ for which for any $t\in[0,1]$ and $f\in C_b^{1,2}([0,1]\times\mathbf{R}^d)$,

$$\begin{aligned}&\int_{\mathbf{R}^d} f(t,x)\nu_{1,t}(dx) - \int_{\mathbf{R}^d} f(0,x)\nu_{1,0}(dx) \\ =\ &\int_{[0,t]\times\mathbf{R}^d\times\mathbf{R}^d} \mathcal{L}_{s,x,v}f(s,x)\nu(dsdxdv). \end{aligned} \tag{4.22}$$

Here

$$\mathcal{L}_{s,x,v}f(s,x) := \frac{\partial f(s,x)}{\partial s} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(s,x)\frac{\partial^2 f(s,x)}{\partial x_i \partial x_j} + < v, D_x f(s,x) > . \tag{4.23}$$

Remark 4.3 *If* $\{P_t\}_{0\le t\le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$ *and* $b \in \mathbf{A}(\{P_t\}_{0\le t\le 1})$, *then* $dtP_t(dx)\delta_{b(t,x)}(dv) \in \tilde{\mathcal{A}}$ *(see Remark 4.1 for notation).*

$v(P_0, P_1)$ and $\mathbf{v}(\mathbf{P})$ are the minimization problems of nonlinear functionals of b in (4.1) and the existence of minimizers can be proved from that of $V(P_0, P_1)$ and $\mathbf{V}(\mathbf{P})$ respectively, under (A.0)-(A.5) (see Proposition 2.1, Corollary 4.1 and Theorem 4.1). The following implies that $v(P_0, P_1)$ and $\mathbf{v}(\mathbf{P})$ are the minimization problems of linear functionals of ν in (4.22), from which the existence of minimizers can be proved directly, under (A.0)', (A.1) and (A.2)' only (see Theorem 4.3 given later). It also implies the convexities of $(P_0, P_1) \mapsto v(P_0, P_1)$ and of $\mathbf{P} \mapsto \mathbf{v}(\mathbf{P})$.

Proposition 4.1 *Suppose that (A.1) holds. Then: (i) for* P_0 *and* $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned} v(P_0, P_1) &= \inf\Big\{\Big|\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv)\Big| \\ &\qquad \nu \in \tilde{\mathcal{A}}, \nu_{1,t} = P_t(t = 0, 1)\Big\}, \end{aligned} \tag{4.24}$$

(ii) for $\mathbf{P} := \{P_t\}_{0\le t\le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned} \mathbf{v}(\mathbf{P}) &= \inf\Big\{\Big|\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv)\Big| \\ &\qquad \nu \in \tilde{\mathcal{A}}, \nu_{1,t} = P_t(0 \le t \le 1)\Big\}. \end{aligned} \tag{4.25}$$

(Proof) We only prove (i) since (ii) can be proved similarly. It is easy to see that the left hand side of (4.24) is greater than or equal to the right hand side (see Remark 4.3). We prove the opposite inequality. For $\nu \in \tilde{\mathcal{A}}$, put $b_\nu(t,x) := \int_{\mathbf{R}^d} v\nu(t,x,dv)$, where $\nu(t,x,dv)$ is a regular conditional probability of ν given (t,x). Then by Jensen's inequality,

$$\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv) \geq \int_0^1 dt \int_{\mathbf{R}^d} L(t,x;b_\nu(t,x))\nu_{1,t}(dx). \quad (4.26)$$

$b_\nu \in \mathbf{A}(\{\nu_{1,t}\}_{0\leq t\leq 1})$ from (4.22)-(4.23) since, for any $t \in [0,1]$ and $f \in C_b^{1,2}([0,1]\times\mathbf{R}^d)$,

$$\begin{aligned} & \int_{[0,t]\times\mathbf{R}^d\times\mathbf{R}^d} < v, D_x f(s,x) > \nu(dsdxdv). \\ = & \int_0^t ds \int_{\mathbf{R}^d} < b_\nu(s,x), D_x f(s,x) > \nu_{1,s}(dx). \square \end{aligned} \quad (4.27)$$

The following partially generalizes Theorem 4.1, (ii).

Theorem 4.3 *Suppose that (A.0)', (A.1) and (A.2)' hold. Then: (i) 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, (ii) for any $\mathbf{P} := \{P_t\}_{0\leq t\leq 1} \subset \mathcal{M}_1(\mathbf{R}^d)$ for which $\mathbf{v}(\mathbf{P})$ is finite, $\mathbf{v}(\mathbf{P})$ has a minimizer. Suppose in addition that L is strictly convex in u. Then the minimizer of $\mathbf{v}(\mathbf{P})$ is unique.*

Remark 4.4 *Under (A.0)-(A.5) with $\gamma = 2$, from Proposition 2.2 and Corollary 4.1, the minimizer of $v(P_0,P_1)$ is unique. Indeed, for any $\{Q_t\}_{0\leq t\leq 1} \subset \mathcal{M}_1(\mathbf{R}^d)$ and $b \in \mathbf{A}(\{Q_t\}_{0\leq t\leq 1})$ for which $\int_{[0,1]\times\mathbf{R}^d} L(t,x;b(t,x))dtQ_t(dx)$ is finite, there exists $X \in \mathcal{A}$ such that (4.2)-(4.3) with $P_t = Q_t$ holds under (A.0) and (A.2) with $\gamma = 2$ (see [5]).*

For any $s \geq 0$ and $P \in \mathcal{M}_1(\mathbf{R}^d)$,

$$\mathbf{\Psi}_P(s) := \left\{ \nu \in \tilde{\mathcal{A}} \middle| \nu_{1,0} = P, \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv) \leq s \right\}. \quad (4.28)$$

As a preparation of the proof of Theorem 4.3, we prove the following.

Lemma 4.1 *Suppose that (A.0)', (A.1,i) and (A.2)' hold. Then for any $s \geq 0$ and compact $K \subset \mathcal{M}_1(\mathbf{R}^d)$, the set $\cup_{P\in K}\mathbf{\Psi}_P(s)$ is compact in $\mathcal{M}_1([0,1]\times\mathbf{R}^d\times\mathbf{R}^d)$.*

(Proof) We only have to consider the case where $\cup_{P\in K}\mathbf{\Psi}_P(s)\neq\emptyset$. We first prove that $\cup_{P\in K}\mathbf{\Psi}_P(s)$ is tight. For $\nu\in\cup_{P\in K}\mathbf{\Psi}_P(s)$, from (A.2)', there exists $C_1>0$ such that

$$\int_{[0,1]\times\mathbf{R}^d\times B_R^c}\nu(dtdxdv)\leq\frac{C_1}{R}\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d}L(t,x;v)\nu(dtdxdv)\leq\frac{C_1 s}{R}\quad(4.29)$$

for sufficiently large $R>0$. Take $\psi\in C_b^\infty(\mathbf{R}^d:[0,1])$ for which $\psi(x)=0$ if $|x|\leq 1$ and $=1$ if $|x|\geq 2$, and put $\psi_r(x):=\psi(x/r)$. Then, from (4.22), (A.0)' and (A.2)', there exists $C_2>0$ such that for any $t\in[0,1]$,

$$\begin{aligned}
&\nu_{1,t}(B_{2r}^c)\leq\int_{\mathbf{R}^d}\psi(x/r)\nu_{1,t}(dx)\\
=\;&\int_{\mathbf{R}^d}\psi(x/r)\nu_{1,0}(dx)+\int_{[0,t]\times\mathbf{R}^d\times\mathbf{R}^d}\Big(\frac{1}{2r^2}\sum_{i,j=1}^d a_{ij}(s,x)\frac{\partial^2\psi(x/r)}{\partial x_i\partial x_j}\\
&\qquad\qquad+\frac{1}{r}<v,D_x\psi(x/r)>\Big)\nu(dsdxdv)\\
\leq\;&\nu_{1,0}(B_r^c)+\frac{C_2}{r}\Big(1+\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d}L(t,x;v)\nu(dtdxdv)\Big)\\
\leq\;&\nu_{1,0}(B_r^c)+\frac{C_2(1+s)}{r}.\qquad(4.30)
\end{aligned}$$

Since $\nu_{1,0}\in K$, (4.29)-(4.30) implies the tightness of $\cup_{P\in K}\mathbf{\Psi}_P(s)$.

Next we prove that $\cup_{P\in K}\mathbf{\Psi}_P(s)$ is closed. Suppose that $\nu^n\in\cup_{P\in K}\mathbf{\Psi}_P(s)$ and that $\nu^n\to\nu$ as $n\to\infty$ weakly. Then it is easy to see that

$$\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d}L(t,x;v)\nu(dtdxdv)\leq\liminf_{n\to\infty}\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d}L(t,x;v)\nu^n(dtdxdv)\leq s\quad(4.31)$$

from (A.1,i). We can also prove that $\nu_{1,0}^n$ is convergent. Indeed, integrating the both sides of (4.22) with ν replaced by ν^n, in t, since $\nu^n\in\cup_{P\in K}\mathbf{\Psi}_P(s)$, we have, for any $f\in C_b^2(\mathbf{R}^d)$,

$$\int_{\mathbf{R}^d}f(x)\nu_{1,0}^n(dx)$$

$$
\begin{aligned}
&= \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} (f(x) - (1-t)\mathcal{L}_{t,x,v}f(x))\nu^n(dtdxdv) \\
&\to \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} (f(x) - (1-t)\mathcal{L}_{t,x,v}f(x))\nu(dtdxdv) \quad (n\to\infty). \qquad (4.32)
\end{aligned}
$$

$P_0 := \lim_{n\to\infty} \nu^n_{1,0}$. Then $P_0 \in K$ since K is compact. Besides, $\nu \in \tilde{\mathcal{A}}$ and $\nu_{1,0} = P_0$. Indeed, for any open $A \subset [0,1]$,

$$
\nu(A\times\mathbf{R}^d\times\mathbf{R}^d) \le \liminf_{n\to\infty} \nu^n(A\times\mathbf{R}^d\times\mathbf{R}^d) = \int_A dt. \qquad (4.33)
$$

For any $g \in C_b([0,1])$ and $f \in C_b^{1,2}([0,1]\times\mathbf{R}^d)$,

$$
\begin{aligned}
&\int_0^1 g(t)\Big(\int_{\mathbf{R}^d} f(t,x)\nu_1(t,dx)\Big)dt \\
&= \int_0^1 g(t)\Big(\int_{\mathbf{R}^d} f(0,x)P_0(dx) + \int_{[0,t]\times\mathbf{R}^d\times\mathbf{R}^d} \mathcal{L}_{s,x,v}f(s,x)ds\nu(s,dxdv)\Big)dt
\end{aligned} \qquad (4.34)
$$

since from (4.22),

$$
\begin{aligned}
&\int_0^1 g(t)\Big(\int_{\mathbf{R}^d} f(t,x)\nu^n_{1,t}(dx)\Big)dt \\
&= \int_0^1 g(t)\Big(\int_{\mathbf{R}^d} f(0,x)\nu^n_{1,0}(dx) + \int_{[0,t]\times\mathbf{R}^d\times\mathbf{R}^d} \mathcal{L}_{s,x,v}f(s,x)\nu^n(dsdxdv)\Big)dt,
\end{aligned} \qquad (4.35)
$$

$$
\int_0^1 g(t)\Big(\int_{\mathbf{R}^d} f(t,x)\nu^n_{1,t}(dx)\Big)dt = \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} g(t)f(t,x)\nu^n(dtdxdv), \qquad (4.36)
$$

$$
\begin{aligned}
&\int_0^1 g(t)\Big(\int_{[0,t]\times\mathbf{R}^d\times\mathbf{R}^d} \mathcal{L}_{s,x,v}f(s,x)\nu^n(dsdxdv)\Big)dt \\
&= \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} \Big(\int_t^1 g(s)ds\Big)\mathcal{L}_{t,x,v}f(t,x)\nu^n(dtdxdv).\square
\end{aligned} \qquad (4.37)
$$

(Proof of Theorem 4.3) We first prove (i). A minimizing sequence of the right hand side of (4.24) is a subset of $\mathbf{\Psi}_{P_0}(s)$ for some $s > 0$. From Lemma 4.1, it

has a convergent subsequence and its limit $\nu \in \tilde{\mathcal{A}}$ is a minimizer of the right hand side of (4.24). Indeed, $\nu_{1,1} = P_1$ since, for any $f \in C_b^{1,2}([0,1] \times \mathbf{R}^d)$ and $t \in [0,1]$,

$$\begin{aligned} & \int_{\mathbf{R}^d} f(t,x)\nu_{1,t}(dx) \\ = & \int_{\mathbf{R}^d} f(1,x)P_1(dx) - \int_{[t,1]\times\mathbf{R}^d\times\mathbf{R}^d} \mathcal{L}_{s,x,v} f(s,x) ds\nu(s,dxdv) \end{aligned} \tag{4.38}$$

in the same way as in (4.34)-(4.37). From (4.26), $(b_\nu, \{\nu_{1,t}\}_{0\le t\le 1})$ is a minimizer of $v(P_0, P_1)$ (see (4.6) for the definition of $v(P_0, P_1)$).

The existence of a minimizer of $\mathbf{v}(\mathbf{P})$ can be proved in the same way as in (i). We prove the uniqueness of a minimizer of $\mathbf{v}(\mathbf{P})$. Take minimizers $b_i \in \mathbf{A}(\mathbf{P})$ $(i = 1, 2)$ of $\mathbf{v}(\mathbf{P})$. Then for any $\lambda \in [0,1]$, $\lambda b_1 + (1-\lambda)b_2 \in \mathbf{A}(\mathbf{P})$ and from (A.1),

$$\begin{aligned} \mathbf{v}(\mathbf{P}) &\le \int_{[0,1]\times\mathbf{R}^d} L(t,x;\lambda b_1(t,x) + (1-\lambda)b_2(t,x))dtP_t(dx) \\ &\le \lambda \int_{[0,1]\times\mathbf{R}^d} L(t,x;b_1(t,x))dtP_t(dx) \\ &\qquad +(1-\lambda)\int_{[0,1]\times\mathbf{R}^d} L(t,x;b_2(t,x))dtP_t(dx) \\ &= \mathbf{v}(\mathbf{P}). \end{aligned} \tag{4.39}$$

The strict convexity of L in u implies that $b_1(t,x) = b_2(t,x)$, $dtP_t(dx)$-a.e.□

Remark 4.5 *When $a(t,x)$ is a zero matrix and $L = L(u)$ and $c(x,y) = L(y-x)$, $T_K(P_0,P_1) \ge v(P_0,P_1)$ from (2.1)-(2.2). Indeed, for any absolutely continuous stochastic process $\varphi(t)$ for which $P\varphi(t)^{-1} = P_t(t = 0,1)$, putting $\mu_\varphi(dtdxdv) := dtP(\varphi(t) \in dx, \dot{\varphi}(t) \in dv)$, $\mu_\varphi \in \tilde{\mathcal{A}}$, $(\mu_\varphi)_{1,t} = P_t(t = 0,1)$ and*

$$E\Big[\int_0^1 L(\dot{\varphi}(t))dt\Big] = \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(v)\mu_\varphi(dtdxdv). \tag{4.40}$$

4.4 Duality Theorems for $v(P_0, P_1)$ and $\mathbf{v}(\{P_t\}_{0\le t\le 1})$.

We extend $v(P_0,P_1)$ and $\mathbf{v}(\{P_t\}_{0\le t\le 1})$ to variational problems on $\mathcal{M}_1([0,1]\times\mathbf{R}^d)$ and generalize Corollaries 4.2 and 4.3 in Sect. 4.2.

For $\lambda \in \mathcal{M}_1([0,1] \times \mathbf{R}^d)$,

$$\begin{aligned} \overline{v}(\lambda) &:= \begin{cases} v(\lambda_0, \lambda_1) & \text{if } \lambda(dtdx) = \frac{1}{2}\delta_0(dt)\lambda_0(dx) + \frac{1}{2}\delta_1(dt)\lambda_1(dx), \\ \infty & \text{otherwise,} \end{cases} \\ \overline{\mathbf{v}}(\lambda) &:= \begin{cases} \mathbf{v}(\{\lambda_t\}_{0\le t\le 1}) & \text{if } \{\lambda_t\}_{0\le t\le 1} \text{ exists,} \\ \infty & \text{otherwise} \end{cases} \end{aligned} \tag{4.41}$$

(see (4.14) for notation). For $f \in C_b([0,1] \times \mathbf{R}^d)$,

$$\begin{aligned} \overline{v}^*(f) &:= \sup\Big\{ \int_{[0,1]\times\mathbf{R}^d} f(t,x)\lambda(dtdx) - \overline{v}(\lambda) \Big| \lambda \in \mathcal{M}_1([0,1] \times \mathbf{R}^d) \Big\}, \\ \overline{\mathbf{v}}^*(f) &:= \sup\Big\{ \int_{[0,1]\times\mathbf{R}^d} f(t,x)\lambda(dtdx) - \overline{\mathbf{v}}(\lambda) \Big| \lambda \in \mathcal{M}_1([0,1] \times \mathbf{R}^d) \Big\}. \end{aligned} \tag{4.42}$$

Since $\overline{v}^*(f)$ only depends on $f(0,\cdot)$ and $f(1,\cdot)$, we write $\overline{v}^*(f) = \overline{v}^*(f(0,\cdot), f(1,\cdot))$ for the sake of simplicity.

(A.1,i) and (A.4,ii) imply that $L(t,x;0)$ is bounded (see Remark2.2, (i)). We introduce the following which is stronger than (A.1).

(A.1)'. (i) $L \in C([0,1] \times \mathbf{R}^d \times \mathbf{R}^d : [0,\infty))$, (ii) $u \mapsto L(t,x;u)$ is convex, and (iii) $L(t,x;0)$ is bounded.

We state the following to point out the convexity and the lower semicontinuity of $\lambda \mapsto \overline{v}(\lambda)$ and $\lambda \mapsto \overline{\mathbf{v}}(\lambda)$.

Proposition 4.2 *Suppose that (A.0)', (A.1)' and (A.2)' hold. Then: (i) for any $\lambda \in \mathcal{M}_1([0,1] \times \mathbf{R}^d)$,*

$$\begin{aligned} \overline{v}(\lambda) &= \sup\Big\{ \int_{[0,1]\times\mathbf{R}^d} f(t,x)\lambda(dtdx) - \overline{v}^*(f) \\ &\qquad \Big| f \in C_b([0,1] \times \mathbf{R}^d) \Big\} (\in [0,\infty]). \end{aligned} \tag{4.43}$$

In particular, for any $\lambda(dtdx) = \frac{1}{2}\delta_0(dt)\lambda_0(dx) + \frac{1}{2}\delta_1(dt)\lambda_1(dx) \in \mathcal{M}_1([0,1] \times \mathbf{R}^d)$,

$$\overline{v}(\lambda) = \sup\Big\{\int_{\mathbf{R}^d}(f(0,x)\lambda_0(dx)+f(1,x)\lambda_1(dx)) - \overline{v}^*(2f(0,\cdot),2f(1,\cdot)) \Big| f(0,\cdot),f(1,\cdot)\in C_b(\mathbf{R}^d)\Big\}. \tag{4.44}$$

(ii) for any $\lambda \in \mathcal{M}_1([0,1]\times\mathbf{R}^d)$,

$$\overline{\mathbf{v}}(\lambda) = \sup\Big\{\int_{[0,1]\times\mathbf{R}^d} f(t,x)\lambda(dtdx) - \overline{\mathbf{v}}^*(f)\Big| f\in C_b([0,1]\times\mathbf{R}^d)\Big\}(\in[0,\infty]). \tag{4.45}$$

(Proof) We first prove (i). $\lambda \mapsto \overline{v}(\lambda)$ is convex from Proposition 4.1 and is not indentically equal to ∞ from (A.1)'. We prove that it is lower semicontinuous. Suppose that $\lambda^n(dtdx)\to\lambda(dtdx)$ as $n\to\infty$ weakly and that $s_0 := \liminf_{n\to\infty}\overline{v}(\lambda^n)$ is finite. Take $\nu^n\in\tilde{\mathcal{A}}$ for which $\nu^n_{1,t}=\lambda^n_t$ $(t=0,1)$ and

$$\overline{v}(\lambda^n)+\frac{1}{n} \geq \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu^n(dtdxdv). \tag{4.46}$$

Since $\lambda^n(dtdx)=\frac{1}{2}\delta_0(dt)\lambda^n_0(dx)+\frac{1}{2}\delta_1(dt)\lambda^n_1(dx)$, $\{\nu^n_{1,0}\}_{n\geq1}$ is tight. In particular, there exist $s>0$ and a compact $K\subset\mathcal{M}_1(\mathbf{R}^d)$ such that $\nu^n\in\cup_{P\in K}\mathbf{\Phi}_P(s)$. From Lemma 4.1, one can take a weak limit point ν of $\{\nu^n\}_{n\geq1}$ so that $\nu\in\cup_{P\in K}\mathbf{\Phi}_P(s_0)$, $\lambda(dtdx)=\frac{1}{2}\delta_0(dt)\lambda_0(dx)+\frac{1}{2}\delta_1(dt)\lambda_1(dx)$ and $\nu_{1,t}=\lambda_t$ $(t=0,1)$ (see (4.34) and (4.38)). Putting $\overline{v}(\lambda)=\infty$ for $\lambda\notin\mathcal{M}_1([0,1]\times\mathbf{R}^d)$, from [9, Theorem 2.2.15 and Lemma 3.2.3], we obtain (i).

(ii) can be proved in the same way as in (i). Indeed, from (4.30),

$$\begin{aligned}
&\nu_{1,0}(B^c_{2r}) \leq \int_{\mathbf{R}^d}\psi(x/r)\nu_{1,0}(dx)\\
=\ & \int_0^1 dt\int_{\mathbf{R}^d}\psi(x/r)\nu_{1,t}(dx) - \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d}(1-t)\\
&\quad\times\Big(\frac{1}{2r^2}\sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2\psi(x/r)}{\partial x_i\partial x_j}+\frac{1}{r}<v,D_x\psi(x/r)>\Big)\nu(dtdxdv)\\
\leq\ & \int_{[0,1]\times B^c_r\times\mathbf{R}^d}\nu(dtdxdv)\\
&\quad+\frac{C_2}{r}\Big(1+\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d}L(t,x;v)\nu(dtdxdv)\Big).
\end{aligned}\tag{4.47}$$

This implies, from Lemma 4.1, the tightness of a class of ν for which $\lambda(dtdx) = \nu(dtdx \times \mathbf{R}^d)$ is tight and for which $\int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv)$ is bounded.$\square$

For $f \in C_b([0,1]\times\mathbf{R}^d)$, measurable $I \subset [0,1]$ and $\{P_t\}_{t\in I} \subset \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned}
\overline{v}^*_{I,\{P_t\}_{t\in I}}(f) &:= \sup\Big\{\int_{(\{0,1\}\setminus I)\times\mathbf{R}^d} f(t,x)\lambda(dtdx) - \overline{v}(\lambda) \\
&\qquad \Big|\lambda \in \mathcal{M}_1([0,1]\times\mathbf{R}^d), \lambda_t(dx) = P_t(t \in \{0,1\}\cap I)\Big\}, \\
\overline{\mathbf{v}}^*_{I,\{P_t\}_{t\in I}}(f) &:= \sup\Big\{\int_{([0,1]\setminus I)\times\mathbf{R}^d} f(t,x)\lambda(dtdx) - \overline{\mathbf{v}}(\lambda)\Big| \\
&\qquad \lambda \in \mathcal{M}_1([0,1]\times\mathbf{R}^d), \lambda_t(dx) = P_t(t\in I)\Big\}. \qquad (4.48)
\end{aligned}$$

For $I = \emptyset$, $\overline{v}^*_{I,\{P_t\}_{t\in I}} = \overline{v}^*$ and $\overline{\mathbf{v}}^*_{I,\{P_t\}_{t\in I}} = \overline{\mathbf{v}}^*$. Since $\overline{v}^*_{I,\{P_t\}_{t\in I}}(f)$ only depends on $f(0,\cdot)$ and $f(1,\cdot)$, we write $\overline{v}^*_{I,\{P_t\}_{t\in I}}(f) = \overline{v}^*_{I,\{P_t\}_{t\in I}}(f(0,\cdot), f(1,\cdot))$ for the sake of simplicity.

From Proposition 4.1, we have,

$$\begin{aligned}
&\overline{v}^*_{I,\{P_t\}_{t\in I}}(2f) \\
=\ &\sup\Big\{\sum_{t\in\{0,1\}\setminus I}\int_{\mathbf{R}^d} f(t,x)\nu_{1,t}(dx) - \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv)\Big| \\
&\qquad \nu\in\tilde{\mathcal{A}}, \nu_{1,t} = P_t(t\in\{0,1\}\cap I)\Big\} \\
=\ &\sup\Big\{\sum_{t\in\{0,1\}\setminus I}\int_{\mathbf{R}^d} f(t,x)Q_t(dx) - \int_{[0,1]\times\mathbf{R}^d} L(t,x;b(t,x))dtQ_t(dx)\Big| \\
&\qquad \{Q_t\}_{0\le t\le 1}\subset\mathcal{M}_1(\mathbf{R}^d), b\in\mathbf{A}(\{Q_t\}_{0\le t\le 1}), Q_t = P_t(t\in\{0,1\}\cap I)\Big\}, \qquad (4.49) \\
&\overline{\mathbf{v}}^*_{I,\{P_t\}_{t\in I}}(f) \\
=\ &\sup\Big\{\int_{([0,1]\setminus I)\times\mathbf{R}^d} f(t,x)dt\nu_{1,t}(dx) - \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv)\Big| \\
&\qquad \nu\in\tilde{\mathcal{A}}, \nu_{1,t} = P_t(t\in I)\Big\} \\
=\ &\sup\Big\{\int_{([0,1]\setminus I)\times\mathbf{R}^d} f(t,x)dtQ_t(dx) - \int_{[0,1]\times\mathbf{R}^d} L(t,x;b(t,x))dtQ_t(dx)
\end{aligned}$$

$$\Big|\{Q_t\}_{0\le t\le 1}\subset \mathcal{M}_1(\mathbf{R}^d), b\in \mathbf{A}(\{Q_t\}_{0\le t\le 1}), Q_t = P_t(t\in I)\Big\}. \quad (4.50)$$

From Proposition 4.2, we easily obtain

Proposition 4.3 *Suppose that (A.0)', (A.1)' and (A.2)' hold. Then: (i) for any P_0 and $P_1\in\mathcal{M}_1(\mathbf{R}^d)$ and $I\subset\{0,1\}$,*

$$\begin{aligned} v(P_0,P_1) &= \sup\Big\{\sum_{t\in\{0,1\}\setminus I}\int_{\mathbf{R}^d} f(t,x)P_t(dx) - v^*_{I,\{P_t\}_{t\in I}}(2f) \\ &\qquad \Big| f(0,\cdot), f(1,\cdot)\in C_b(\mathbf{R}^d)\Big\}(\in[0,\infty]), \end{aligned} \quad (4.51)$$

(ii) for any $\mathbf{P}:=\{P_t\}_{0\le t\le 1}\subset\mathcal{M}_1(\mathbf{R}^d)$ and measurable $I\subset[0,1]$,

$$\begin{aligned} \mathbf{v}(\mathbf{P}) &= \sup\Big\{\int_{([0,1]\setminus I)\times\mathbf{R}^d} f(t,x)dtP_t(dx) - \mathbf{v}^*_{I,\{P_t\}_{t\in I}}(f) \\ &\qquad \Big| f\in C_b([0,1]\times\mathbf{R}^d)\Big\}(\in[0,\infty]). \end{aligned} \quad (4.52)$$

We introduce a new assumption which is stronger than (A.2)' but is weaker than (A.2).
(A.2)".

$$\liminf_{|u|\to\infty}\frac{\inf\{L(t,x;u):(t,x)\in[0,1]\times\mathbf{R}^d\}}{|u|}=\infty. \quad (4.53)$$

If (A.0) and (A.5) hold, then the HJB Eqn (2.19) with a terminal function $f\in C^3_b(\mathbf{R}^d)$ has a unique classical solution $\varphi(t,x;f)$ in $C^{1,2}_b([0,1]\times\mathbf{R}^d)$ as we pointed out before Theorem 2.2. In particular, under (A.2)", $D_zH(t,x;D_x\varphi(t,x;f))$ is bounded (see the proof of the following corollary). Hence Proposition 4.3 implies the following which generalizes Corollaries 4.2 and 4.3 (see (4.17)-(4.18) for notation).

Corollary 4.4 *Suppose that (A.0), (A.1)', (A.2)" and (A.5) hold. Then: (i) for any P_0 and $P_1\in\mathcal{M}_1(\mathbf{R}^d)$,*

$$v(P_0,P_1) = \sup\Big\{\int_{\mathbf{R}^d} f(x)P_1(dx) - \int_{\mathbf{R}^d}\varphi(0,x;f)P_0(dx) \Big| f\in C_b^\infty(\mathbf{R}^d)\Big\}, \tag{4.54}$$

(ii) for any $\mathbf{P} := \{P_t\}_{0\le t\le 1}\subset \mathcal{M}_1(\mathbf{R}^d)$,

$$\mathbf{v}(\mathbf{P}) = \sup\Big\{\int_{[0,1]\times\mathbf{R}^d} f(t,x)dtP_t(dx) - \int_{\mathbf{R}^d}\phi(0,x;f)P_0(dx) \Big| f\in C_b^\infty([0,1]\times\mathbf{R}^d)\Big\}. \tag{4.55}$$

(Proof). We only prove (i) since (ii) can be proved similarly. From (4.51),

$$v(P_0,P_1) = \sup\Big\{\int_{\mathbf{R}^d} f(x)P_1(dx) - v^*_{\{0\},\{P_0\}}(0,2f)\Big| f\in C_b(\mathbf{R}^d)\Big\}. \tag{4.56}$$

We prove that $C_b(\mathbf{R}^d)$ can be replaced by $C_b^\infty(\mathbf{R}^d)$ in (4.56). Take $\rho\in C_0^\infty([-1,1]^d:[0,\infty))$ for which $\int_{\mathbf{R}^d}\rho(x)dx=1$. For $\varepsilon>0$ and $f\in C_b(\mathbf{R}^d)$, put

$$\rho_\varepsilon(x) := \varepsilon^{-d}\rho(x/\varepsilon),\quad f_\varepsilon(x) := \int_{\mathbf{R}^d} f(y)\rho_\varepsilon(y-x)dy. \tag{4.57}$$

Then $f_\varepsilon\in C_b^\infty(\mathbf{R}^d)$ and, from (4.56),

$$v(P_0,P_1)\ge \int_{\mathbf{R}^d} f_\varepsilon(x)P_1(dx) - v^*_{\{0\},\{P_0\}}(0,2f_\varepsilon). \tag{4.58}$$

Take $\nu^\varepsilon\in\tilde{\mathcal{A}}$ for which $\nu^\varepsilon_{1,0}=P_0$ and

$$v^*_{\{0\},\{P_0\}}(0,2f_\varepsilon) - \varepsilon < \int_{\mathbf{R}^d} f_\varepsilon(x)\nu^\varepsilon_{1,1}(dx) - \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu^\varepsilon(dtdxdv). \tag{4.59}$$

Then $\{\nu^\varepsilon\}_{\varepsilon\in(0,1)}$ is tight from Lemma 4.1 since

$$v^*_{\{0\},\{P_0\}}(0,2f_\varepsilon)\ge -\sup_{x\in\mathbf{R}^d}|f(x)| - \sup_{(t,x)\in[0,1]\times\mathbf{R}^d} L(t,x;0) > -\infty \tag{4.60}$$

from (A.1)'. Take a weak limit point $\overline{\nu}$ of ν_ε as $\varepsilon \to 0$ such that

$$\begin{aligned} \limsup_{\varepsilon\to 0} v^*_{\{0\},\{P_0\}}(0, 2f_\varepsilon) &\leq \int_{\mathbf{R}^d} f(x)\overline{\nu}_{1,1}(dx) - \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\overline{\nu}(dtdxdv) \\ &\leq v^*_{\{0\},\{P_0\}}(0, 2f) \end{aligned} \tag{4.61}$$

(see (4.38)). (4.58) and (4.61) imply that $C_b(\mathbf{R}^d)$ can be replaced by $C_b^\infty(\mathbf{R}^d)$ in (4.56). For $f \in C_b^\infty(\mathbf{R}^d)$,

$$v^*_{\{0\},\{P_0\}}(0, 2f) = \int_{\mathbf{R}^d} \varphi(0, x; f)P_0(dx) \tag{4.62}$$

since for any $\nu \in \tilde{\mathcal{A}}$ for which $\nu_{1,0} = P_0$, from (4.22),

$$\begin{aligned} & \int_{\mathbf{R}^d} f(x)\nu_{1,1}(dx) - \int_{\mathbf{R}^d} \varphi(0,x;f)P_0(dx) \\ = & \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} \mathcal{L}_{t,x;v}\varphi(t,x;f)\nu(dtdxdv) \\ = & \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} (< v, D_x\varphi(t,x;f) > -H(t,x;D_x\varphi(t,x;f)))\nu(dtdxdv) \\ \leq & \int_{[0,1]\times\mathbf{R}^d\times\mathbf{R}^d} L(t,x;v)\nu(dtdxdv), \end{aligned} \tag{4.63}$$

where the equality holds if and only if $\nu(dtdxdv) = dt\nu_{1,t}(dx)\delta_{D_zH(t,x;D_x\varphi(t,x;f))}(dv)$. Notiec that (4.2) has a unique strong solution if $b(t,x) = D_zH(t,x;D_x\varphi(t,x;f))$. Indeed, for any $r > 0$,

$$\sup\{|D_zH(t,x;z)| : (t,x,z) \in [0,1]\times\mathbf{R}^d\times\mathbf{R}^d, |z| < r\} < \infty. \tag{4.64}$$

We prove (4.64). For any $r > 0$, there exists $R(r) > 0$ such that

$$\inf\{|D_uL(t,x;u)| : (t,x,u) \in [0,1]\times\mathbf{R}^d\times\mathbf{R}^d, |u| > R(r)\} \geq r \tag{4.65}$$

since from (A.5,ii), for any $(t,x,u) \in [0,1]\times\mathbf{R}^d\times\mathbf{R}^d$,

$$L(t,x;0) \geq L(t,x;u) + < D_uL(t,x;u), -u >, \tag{4.66}$$

from which

$$
\begin{aligned}
& \inf\{|D_u L(t,x;u)| : (t,x) \in [0,1] \times \mathbf{R}^d\} \\
\geq\ & \frac{1}{|u|}\{\inf\{L(t,x;u) : (t,x) \in [0,1] \times \mathbf{R}^d\} \\
& \qquad - \sup\{L(t,x;0) : (t,x) \in [0,1] \times \mathbf{R}^d\}\} \\
\to\ & \infty \quad (\text{as } |u| \to \infty \text{ from (A.2)"}). \qquad (4.67)
\end{aligned}
$$

The supremum in (4.64) 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 $|z| < r$ and $|D_z H(t,x;z)| > R(r)$. The second inequality implies that $|z| \geq r$ since $z = D_u L(t,x; D_z H(t,x;z))$ from (A.2)" and (A.5,ii). This contradicts to the fact that $|z| < r$.□

In Theorem 4.3 we proved the existence of a minimizer $(b, \{Q_t\}_{0 \leq t \leq 1})$ of $v(P_0, P_1)$ and the existence and the uniqueness of a minimizer $b \in \mathbf{A}(\{Q_t\}_{0 \leq t \leq 1})$ of $\mathbf{v}(\mathbf{P})$. As a corollary to Corollary 4.4, we can prove the uniqueness of $b \in \mathbf{A}(\{Q_t\}_{0 \leq t \leq 1})$ for which $(b, \{Q_t\}_{0 \leq t \leq 1})$ is a minimizer of $v(P_0, P_1)$, under a stronger assumption than Theorem 4.3.

Corollary 4.5 *Suppose that (A.0), (A.1)', (A.2)" 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, $b \in \mathbf{A}(\{Q_t\}_{0 \leq t \leq 1})$ for which $(b, \{Q_t\}_{0 \leq t \leq 1})$ is a minimizer of $v(P_0, P_1)$ is unique.*

(Proof). Take a maximizing sequence $\{\varphi(\cdot, \cdot; f_n)\}_{n \geq 1}$ in (4.54) and a minimizer ν of the right hand side of (4.24). Then

$$
\begin{aligned}
& \int_{[0,1] \times \mathbf{R}^d \times \mathbf{R}^d} |L(t,x;v) - (< v, D_x \varphi(t,x;f_n) > \\
& \qquad\qquad - H(t,x; D_x \varphi(t,x;f_n)))| \nu(dtdxdv) \\
=\ & \int_{[0,1] \times \mathbf{R}^d \times \mathbf{R}^d} L(t,x;v) \nu(dtdxdv) - \int_{\mathbf{R}^d} (f_n(x) P_1(dx) - \varphi(0,x;f_n) P_0(dx)) \\
\to\ & 0 \quad (\text{as } n \to \infty). \qquad (4.68)
\end{aligned}
$$

Taking a subsequnce if necessary, since an L^1-convergent sequence of random variables has an a.s. convergent subsequence, from (A.5,ii),

$$
\nu(dtdxdv) = dt \nu_{1,t}(dx) \delta_{\lim_{n \to \infty} D_z H(t,x; D_x \varphi(t,x;f_n))}(dv). \qquad (4.69)
$$

Since a subsequence of a maximizing sequence in (4.54) is also a maximizing sequence in (4.54), the proof is over.□

References

[1] Blumenthal, R.M., Corson, H.H.: On continuous collections of measures. In: Le Cam, L. etal. (eds.): Proc. 6th Berkeley Sympos. Math. Statist. Probab. **2**, Berkeley 1970/1971. Berkeley: Univ. California Press 1972, pp. 33-40

[2] Carlen, E.A.: Conservative diffusions. Commun Math Phys **94**, 293-315 (1984)

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

[4] Cattiaux, P., Léonard, C.: Correction to: "Minimization of the Kullback information of diffusion processes" [Ann. Inst. H. Poincar Probab. Statist. 30 (1994), no. 1, 83–132]. Ann Inst H Poincar Probab Statist **31**, 705–707 (1995)

[5] Cattiaux, P., Léonard, C.: Large deviations and Nelson processes. Forum Math **7**, 95–115 (1995)

[6] Cattiaux, P., Léonard, C.: Minimization of the Kullback information for some Markov processes. In: Azema, J. etal. (eds.): Séminaire de Probabilités, XXX, Lecture Notes in Math., 1626. Berlin Heidelberg New York Tokyo: Springer-Verlag 1996, pp. 288–311

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

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

[9] Deuschel, J.D., Stroock, D.W.: Large deviations (Pure and Applied Mathematics 137) Boston: Academic Press, Inc. 1989

[10] Evans, L.C.: Partial differential equations (Graduate Studies in Mathematics 19) Providence: AMS 1998

[11] Evans, L.C.: Partial differential equations and Monge-Kantorovich mass transfer. In: Current developments in mathematics, 1997 (Cambridge, MA). Boston: Int. Press 1999, pp. 65–126

[12] Evans, L.C., Gangbo, W.: Differential equations methods for the Monge-Kantorovich mass transfer problem. Mem. Amer. Math. Soc. **137**, No. 653 (1999)

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

[14] Föllmer, H.: Random fields and diffusion processes. In: Hennequin, PL (ed.): École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87, Lecture Notes in Math. 1362. Berlin Heidelberg New York Tokyo: Springer-Verlag 1988, pp. 101–203

[15] Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**, 113-161 (1996)

[16] Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes. Amsterdam New York Oxford Tokyo: North-Holland/Kodansha 1981

[17] Jacod, J., Shiryaev, A.N.: Limit theorems for stochastic processes. Berlin Heidelberg New York Tokyo: Springer-Verlag 1987

[18] Jamison, B.: Reciprocal processes. Z. Wahrsch. Verw. Gebiete **30**, 65–86 (1974)

[19] Jamison, B.: The Markov process of Schrödinger. Z. Wahrsch. Verw. Gebiete **32**, 323–331 (1975)

[20] Kellerer, H.G.: Duality theorems for marginal problems. Z. Wahrsch. Verw. Gebiete **67**, 399–432 (1984)

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

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

[23] Mikami, T.: Markov marginal problems and their applications to Markov optimal control. In: McEneaney, W. M. etal. (eds.): Stochastic Analysis, Control, Optimization and Applications, A Volume in Honor of W. H. Fleming. Boston: Birkhäuser 1999, pp. 457-476.

[24] Mikami, T.: Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric. Appl Math Optim **42**, 203-227 (2000)

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

[26] Mikami, T.: Monge's problem with a quadratic cost by the zero-noise limit of h-path processes. Probab. Theory Related Fields **129**, 245-260 (2004)

[27] Mikami, T.: A Simple Proof of Duality Theorem for Monge-Kantorovich Problem. Kodai Math. J. **29**, 1-4 (2006)

[28] Mikami, T.: Semimartingales from the Fokker-Planck equation. Appl. Math. Optim. **53**, 209-219 (2006)

[29] Mikami, T.: Optimal transportation problem as stochastic mechanics. To appear in Sugaku (in Japanese), Math. Soc. Japan. (English translation: Sugaku Exposions, AMS).

[30] Mikami, T., Thieullen, M.: Duality Theorem for Stochastic Optimal Control Problem. Stoc. Proc. Appl. online first (2006).

[31] Mikami, T., Thieullen, M.: Optimal Transportation Problem by Stochastic Optimal Control. Submitted to SICON. revision is in progress, 2005.

[32] Monge, G.: Mémoire sur la théorie des déblais et des remblais. In: Histoire de l'Académie Royale des Sciences de Paris, 666-704 (1781)

[33] Nagasawa, M.: Stochastic processes in quantum physics (Monographs in Mathematics 94) Basel: Birkhauser 2000

[34] Nelson, E.: Dynamical theories of Brownian motion. Princeton: Princeton University Press 1967

[35] Nelson, E.: Quantum fluctuations. Princeton: Princeton University Press 1984

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

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

[38] Smith, C., Knott, M.: On Hoeffding-Fréchet bounds and cyclic monotone relations. J. Multivariate Anal. **40**, 328–334 (1992)

[39] Trudinger, N.S., Wang, Xu-Jia.: On the Monge mass transfer problem. Calc. Var. **13**, 19–31 (2001)

[40] Villani, C.: Topics in optimal transportation (Graduate Studies in Mathematics 58) Providence: Amer. Math. Soc. 2003

[41] Zambrini, J.C.: Variational processes. In: Albeverio, S. etal. (eds.): Stochastic processes in classical and quantum systems, Ascona 1985, Lecture Notes in Phys. 262. Berlin Heidelberg New York Tokyo: Springer-Verlag 1986, pp. 517-529

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