数理解析研究所講究録
第 1695 巻 2010 年 35-55

Variational characterization of the Knothe-Rosenblatt type rearrangements and its stochastic version

広島大学・工学研究科　三上　敏夫 (Toshio Mikami)

Department of Applied Mathematics

Hiroshima University

1 Introduction.

The Knothe-Rosenblatt rearrangement plays a crucial role in many fields, e.g., the Brunn-Minkowski inequality and statistics (see [12], [13], [22] and the references therein).

Let $d\geq 1$ and let $\mathcal{M}_1(\mathbf{R}^d)$ denote the set of all Borel probability measures on $\mathbf{R}^d$ with a weak topology. For a distribution function F on $\mathbf{R}$, let

$$F^{-1}(u):=\inf\{x\in\mathbf{R}|u\leq F(x)\}\quad(0\leq u\leq 1). \tag{1.1}$$

For P_0, $P_1\in\mathcal{M}_1(\mathbf{R}^d)$, $x\in\mathbf{R}$, and $i=0,1$, let

$$\begin{aligned} F_{i,k}(x|\mathbf{x}_{k-1}) &:= \begin{cases} P_i((-\infty,x]\times\mathbf{R}^{d-1}) & (k=1),\\ P_i((-\infty,x]\times\mathbf{R}^{d-k}|\mathbf{x}_{k-1}) & (1<k<d),\\ P_i((-\infty,x]|\mathbf{x}_{d-1}) & (k=d),\end{cases}\\ \varphi_k(\mathbf{x}_k) &:= F_{1,k}(\cdot|\varphi_1(x_1),\cdots,\varphi_{k-1}(\mathbf{x}_{k-1}))^{-1}(F_{0,k}(x_k|\mathbf{x}_{k-1}))(1\leq k\leq d), \end{aligned} \tag{1.2}$$

where $\mathbf{x}_k:=(x_i)_{1\leq i\leq k}\in\mathbf{R}^k$ for $x=(x_i)_{1\leq i\leq d}\in\mathbf{R}^d$ and $P_i(\cdot|\mathbf{x}_{k-1})$ denotes the regular conditional probability of P_i given $\mathbf{x}_{k-1}$.

Suppose that $F_{0,k}(\cdot|\mathbf{x}_{k-1})$ is continuous for all $k=1,\cdots,d$. Then P_1 is the image measure of P_0 by

$$T_{KR}(\mathbf{x}_d):=(\varphi_1(x_1),\cdots,\varphi_d(\mathbf{x}_d)).$$

T_{KR} is called the **Knothe-Rosenblatt rearrangement**. Suppose, in addition, that $F_{1,k}(\cdot|\mathbf{x}_{k-1})$ is continuous for all $k=1,\cdots,d$. Then T_{KR} is invertible and the minimizer of the following weakly converges to $P_0(dx)\delta_{T_{KR}(x)}(dy)$ as $\varepsilon\to 0$: for $p>1$,

$$\inf\Big\{\int_{\mathbf{R}^d\times\mathbf{R}^d}\sum_{k=1}^{d}\varepsilon^{2(k-1)}|y_k-x_k|^p\mu(dxdy)\Big|\mu(dx\times\mathbf{R}^d)=P_0(dx),$$
$$\mu(\mathbf{R}^d\times dy)=P_1(dy)\Big\}, \tag{1.3}$$

provided $\int_{\mathbf{R}^d}|x|^p(P_0(dx)+P_1(dx))$ is finite (see [2]). Here $\delta_x(dy)$ denotes the delta measure on $\{x\}$.

For $1\le k\le d$, $\mathbf{x}_{k-1}\in\mathbf{R}^{k-1}$, $dF_{0,k}(x|\mathbf{x}_{k-1})\delta_{\varphi_k(\mathbf{x}_k)}(dy)$ is the unique minimizer of

$$\inf\Big\{\int_{\mathbf{R}\times\mathbf{R}}|y-x|^p\mu(dxdy)\Big|\mu(dx\times\mathbf{R})=dF_{0,k}(x|\mathbf{x}_{k-1}),$$
$$\mu(\mathbf{R}\times dy)=dF_{1,k}(y|\varphi_1(x_1),\cdots,\varphi_{k-1}(\mathbf{x}_{k-1}))\Big\} \tag{1.4}$$

(see e.g. [21], [24]). (1.4) also implies that $P_0(d\mathbf{x}_k\times\mathbf{R}^{d-k})\delta_{(\varphi_1(x_1),\cdots,\varphi_k(\mathbf{x}_k))}(d\mathbf{y}_k)$ is the unique minimizer of

$$\inf\Big\{\int_{\mathbf{R}^k\times\mathbf{R}^k}|y_k-x_k|^p\mu(dxdy)\Big|\mu(dx\times\mathbf{R}^k)=P_0(dx\times\mathbf{R}^{d-k}),$$
$$\mu(\mathbf{R}^k\times dy)=P_1(dy\times\mathbf{R}^{d-k}),$$
$$y_i=\varphi_{i-1}(\mathbf{x}_{i-1})(i=1,\cdots,k-1),\mu-a.s.\Big\}. \tag{1.5}$$

We generalize (1.5) and call the minimizer the **Knothe-Rosenblatt type rearrangement**. We also prove the duality theorem, give the convergence result which generalizes (1.3) by the idea of [2] and consider the similar problems in the stochastic control setting.

2 Knothe-Rosenblatt type rearrangement.

Let $d\ge 2$, $1\le d_1<d$, $c(x,y):\mathbf{R}^{d-d_1}\times\mathbf{R}^{d-d_1}\mapsto[0,\infty)$ be Borel measurable and $\nu\in\mathcal{M}_1(\mathbf{R}^{2d_1})$. For P_0, $P_1\in\mathcal{M}_1(\mathbf{R}^d)$, let

$$T(P_0,P_1|\nu) := \inf\Big\{\int_{\mathbf{R}^d\times\mathbf{R}^d}c(\mathbf{x}_{d_1,d},\mathbf{y}_{d_1,d})\mu(dxdy)\Big|$$
$$\mu(d\mathbf{x}_{d_1}\times\mathbf{R}^{d-d_1}\times d\mathbf{y}_{d_1}\times\mathbf{R}^{d-d_1})=\nu(d\mathbf{x}_{d_1}d\mathbf{y}_{d_1}),$$
$$\mu(dx\times\mathbf{R}^d)=P_0(dx),\mu(\mathbf{R}^d\times dy)=P_1(dy)\Big\}, \tag{2.1}$$

where $\mathbf{x}_{i,j} := (x_k)_{i+1\le k\le j} \in \mathbf{R}^{j-i}$ for $x = (x_k)_{1\le k\le d} \in \mathbf{R}^d$. If the set over which the infimum is taken is empty, then we consider the infimum is equal to infinity. If there exists a Borel measurable function $\varphi : \mathbf{R}^{d_1} \mapsto \mathbf{R}^{d_1}$ such that $\mathbf{y}_{d_1} = \varphi(\mathbf{x}_{d_1})$, ν-a.s., then we write, for simplicity,

$$T(P_0, P_1|\varphi) := T(P_0, P_1|\nu).$$

We first show the existence of the Knothe-Rosenblatt type rearrangement.

Proposition 2.1 *Suppose that c is lower semi-continuous. Then, for any $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$, $T(P_0, P_1|\nu)$ has a minimizer, provided it is finite.*

(Proof) Let $\{\mu_n\}_{n\ge 1}$ be a minimizing sequence of $T(P_0, P_1|\nu)$. Since $\mu_n(dx \times \mathbf{R}^d) = P_0(dx)$ and $\mu_n(\mathbf{R}^d \times dy) = P_1(dy)$, it has a weakly convergent subsequence which we denote by $\{\mu_{n(k)}\}_{k\ge 1}$. Let μ denote the limit. Then by Skorohod's representation theorem, Fatou's lemma and the lower semicontinuity of c,

$$\begin{aligned} T(P_0, P_1|\nu) &= \lim_{k\to\infty} \int_{\mathbf{R}^d\times\mathbf{R}^d} c(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu_{n(k)}(dxdy) \\ &\ge \int_{\mathbf{R}^d\times\mathbf{R}^d} c(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu(dxdy). \end{aligned} \tag{2.2}$$

For any $f \in C(\mathbf{R}^{d_1} \times \mathbf{R}^{d_1})$,

$$\begin{aligned} \int_{\mathbf{R}^d\times\mathbf{R}^d} f(\mathbf{x}_{d_1}, \mathbf{y}_{d_1})\mu(dxdy) &= \lim_{k\to\infty} \int_{\mathbf{R}^d\times\mathbf{R}^d} f(\mathbf{x}_{d_1}, \mathbf{y}_{d_1})\mu_{n(k)}(dxdy) \\ &= \int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_1}} f(\mathbf{x}_{d_1}, \mathbf{y}_{d_1})\nu(d\mathbf{x}_{d_1} d\mathbf{y}_{d_1}). \end{aligned} \tag{2.3}$$

In the same way, one can show that $\mu(dx \times \mathbf{R}^d) = P_0(dx)$ and $\mu(\mathbf{R}^d \times dy) = P_1(dy)$.□

2.1 Duality Theorem

It is easy to see that the following holds:

$$\begin{aligned} T(P_0, P_1|\varphi) &= \inf\Big\{\int_{\mathbf{R}^d\times\mathbf{R}^d} \frac{c(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})}{1_{\{\varphi(\mathbf{x}_{d_1})\}}(\mathbf{y}_{d_1})}\mu(dxdy)\Big| \\ &\qquad \mu(dx \times \mathbf{R}^d) = P_0(dx), \mu(\mathbf{R}^d \times dy) = P_1(dy)\Big\}, \end{aligned} \tag{2.4}$$

where $1_A(x) := 1$ if $x \in A$ and $:= 0$ if $x \notin A$ for the set A. This leads us to the duality theorem for $T(P_0, P_1|\varphi)$ which can be obtained from [11] (see also p. 76 in Vol. 1 of [21]).

Theorem 2.1 *For any $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$T(P_0,P_1|\varphi) = \sup\Big\{\int_{\mathbf{R}^d} f_1(y)P_1(dy) - \int_{\mathbf{R}^d} f_0(x)P_0(dx)\Big| f_0, f_1 \in C_b(\mathbf{R}^d), \quad f_1(y) - f_0(x) \le \frac{c(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})}{1_{\{\varphi(\mathbf{x}_{d_1})\}}(\mathbf{y}_{d_1})}\Big\}. \tag{2.5}$$

For $f \in C_b(\mathbf{R}^d)$ and $x = (\mathbf{x}_{d_1}, \mathbf{x}_{d_1,d}) \in \mathbf{R}^d$,

$$v(x; f|\varphi) := \sup\{f(\varphi(\mathbf{x}_{d_1}), y) - c(\mathbf{x}_{d_1,d}, y)| y \in \mathbf{R}^{d-d_1}\}. \tag{2.6}$$

Then, from (2.5),

$$T(P_0,P_1|\varphi) = \sup\Big\{\int_{\mathbf{R}^d} f(y)P_1(dy) - \int_{\mathbf{R}^d} v(x; f|\varphi)P_0(dx)\Big| f \in C_b(\mathbf{R}^d)\Big\}. \tag{2.7}$$

We easily obtain the following (see e.g. (2.8)-(2.9) in [16]).

Proposition 2.2 *Suppose that $\varphi \in C(\mathbf{R}^{d_1} : \mathbf{R}^{d_1})$, $c(x,y) \in C(\mathbf{R}^{d-d_1} \times \mathbf{R}^{d-d_1} : [0,\infty))$ and $\lim_{|y-x|\to\infty} c(x,y) = \infty$. Then for any $f \in C_b(\mathbf{R}^d)$, $v(\cdot; f|\varphi)$ is continuous.*

We formally derive the Hamilton-Jacobi Equation (HJ Eqn for short) for $v(x; f|\varphi)$. Let

$$\begin{aligned} \Phi(t,x) &:= x + t(\varphi(x) - x), \\ b(t,x) &:= \varphi(\Phi(t,\cdot)^{-1}(x)) - \Phi(t,\cdot)^{-1}(x) \quad ((t,x) \in [0,1] \times \mathbf{R}^{d_1}), \end{aligned} \tag{2.8}$$

provided it exists. Then

$$\frac{d\Phi(t,x)}{dt} = \varphi(x) - x = b(t, \Phi(t,x)). \tag{2.9}$$

In case $c(x,y) = \ell(y-x)$ for a convex ℓ, we consider the following HJ Eqn:

$$\frac{\partial v(t,x)}{\partial t} + < \nabla_{d_1} v(t,x), b(t, \mathbf{x}_{d_1}) > + h(\nabla_{d_1,d} v(t,x)) = 0 \quad ((t,x) \in (0,1) \times \mathbf{R}^d), \tag{2.10}$$

where $\nabla_{d_1} := (\partial/\partial x_i)_{i=1}^{d_1}$, $\nabla_{d_1,d} := (\partial/\partial x_i)_{i=d_1+1}^{d}$ and

$$h(z) := \sup\{< u, z > -\ell(u) | u \in \mathbf{R}^{d-d_1}\} \quad (z \in \mathbf{R}^{d-d_1}).$$

Then we have

Proposition 2.3 *Suppose that $c(x,y)=\ell(y-x)$ for a convex ℓ, that $\Phi(t,\cdot)$ is injective for all $t\in[0,1]$, that the HJ Eqn (2.10) has a classical solution v and that the following ODE has an absolutely continuous solution: for any $\phi_2(0)=\mathbf{x}_{d_1,d}$*

$$\frac{d\phi_2(t)}{dt}=\nabla h(\nabla_{\mathbf{x}_{d_1,d}}v(t,\Phi(t,\mathbf{x}_{d_1}),\phi_2(t))). \tag{2.11}$$

Then $v(0,x)=v(x;v(1,\cdot)|\varphi)$.

(Proof) For any $\phi_2\in AC(\mathbf{R}^{d-d_1})$, from (2.9), we have

$$\begin{aligned}
& v(1,\Phi(1,\mathbf{x}_{d_1}),\phi_2(1))-v(0,\Phi(0,\mathbf{x}_{d_1}),\phi_2(0)) \\
= & \int_0^1\Big\{\frac{\partial v(t,\Phi(t,\mathbf{x}_{d_1}),\phi_2(t))}{\partial t}+<\nabla_{d_1}v(t,\Phi(t,\mathbf{x}_{d_1}),\phi_2(t)),b(t,\Phi(t,\mathbf{x}_{d_1}))> \\
& \quad +<\nabla_{d_1,d}v(t,\Phi(t,\mathbf{x}_{d_1}),\phi_2(t)),\frac{d\phi_2(t)}{dt}>\Big\}dt \\
= & \int_0^1\Big\{-h(\nabla_{d_1,d}v(t,\Phi(t,\mathbf{x}_{d_1}),\phi_2(t))) \\
& \quad +<\nabla_{d_1,d}v(t,\Phi(t,\mathbf{x}_{d_1}),\phi_2(t)),\frac{d\phi_2(t)}{dt}>\Big\}dt \\
\leq & \int_0^1\ell\Big(\frac{d\phi_2(t)}{t}\Big)dt,
\end{aligned} \tag{2.12}$$

where the equality holds if (2.11) holds. By Jensen's inequality,

$$\begin{aligned}
v(0,x) &= \sup\Big\{v(1,\varphi(\mathbf{x}_{d_1}),\phi_2(1))-\int_0^1\ell\Big(\frac{d\phi_2(t)}{t}\Big)dt\Big|\phi_2(0)=\mathbf{x}_{d_1,d}\Big\} \\
&= \sup\Big\{v(1,\varphi(\mathbf{x}_{d_1}),\phi_2(1))-\ell(\phi_2(1)-\phi_2(0))\Big|\phi_2(0)=\mathbf{x}_{d_1,d}\Big\} \\
&= v(x;v(1,\cdot)|\varphi).\square
\end{aligned} \tag{2.13}$$

Before we formulate the duality theorem in the framework of the theory of viscosity solutions, we give assumptions.

(A.1). $b(t,x)$ is bounded and there exists $K>0$ such that

$$|b(t,x)-b(t,y)|\leq K|x-y|\quad (t\in[0,1],x,y\in\mathbf{R}^{d_1}).$$

(A.2). There exists $m\in C([0,1]\times\mathbf{R}^{d_1}\times[0,1]\times\mathbf{R}^{d_1}\times[0,\infty))$ such that $m(t,x,s,y,0)=0$ and that

$$|b(t,x)-b(s,y)|\leq m(t,x,s,y,|t-s|+|x-y|)\quad (t,s\in[0,1],x,y\in\mathbf{R}^{d_1}).$$

(A.3). $\ell : \mathbf{R}^{d-d_1} \mapsto [0,\infty)$ is convex and $\liminf_{|v|\to\infty} \frac{\ell(v)}{|v|} = \infty$.

Example 2.1 *Suppose that $d_1 = 1$. Then (A.1)-(A.2) holds if $1 < d\varphi(x)/dx \le K+1$.*

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

$$\begin{aligned} & v(t,x;f|\varphi) \\ := \ & \sup\Big\{ f(\Phi(1,\mathbf{y}_{d_1}),\phi_2(1)) - \int_t^1 \ell\Big(\frac{d\phi_2(s)}{ds}\Big) ds \Big| (\Phi(t,\mathbf{y}_{d_1}),\phi_2(t)) = x \Big\}. \end{aligned} \quad (2.14)$$

Then it is easy to see that the following holds:

$$\begin{aligned} & v(t,x;f|\varphi) \\ = \ & \sup\Big\{ f(\mathbf{x}_{d_1} + (1-t)b(t,\mathbf{x}_{d_1}), y) - (1-t)\ell\Big(\frac{y - \mathbf{x}_{d_1,d}}{1-t}\Big) \Big| y \in \mathbf{R}^{d-d_1} \Big\}. \end{aligned} \quad (2.15)$$

(see (2.8)). We also have

Corollary 2.1 *Suppose that $c(x,y) = \ell(y-x)$ and that (A.1)-(A.3) hold. Then for any Lipschitz continuous $f : \mathbf{R}^d \mapsto \mathbf{R}$, $v(t,x;f|\varphi)$ is a Lipschitz continuous viscosity solution of (2.10). In particular, for any $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$T(P_0,P_1|\varphi) = \sup\Big\{ \int_{\mathbf{R}^d} v(1,y)P_1(dy) - \int_{\mathbf{R}^d} v(0,x)P_0(dx) \Big| v(1,\cdot) \in C_b^\infty(\mathbf{R}^d) \Big\}, \quad (2.16)$$

where $v(t,x)$ denotes a bounded uniformly continuous viscosity solution of (2.10).

(Proof) In the same way as in p. 127 in [4], by (A.1) and (A.3), one can prove that $v(\cdot,\cdot;f|\varphi)$ is Lipschitz continuous for Lipschitz continuous $f : \mathbf{R}^d \mapsto \mathbf{R}$. In addition, from Chap. II.16 of [7], under (A.1)-(A.3), $v(t,x;f|\varphi)$ is a bounded, uniformly continuous viscosity solution of (2.10). It is easy to see that the supremum in (2.7) can be taken only over all $f \in C_b^\infty(\mathbf{R}^d)$. For $n \ge 1$, $f \in C_b^\infty(\mathbf{R}^d)$ and $(t,x) \in [0,1] \times \mathbf{R}^d$,

$$\begin{aligned} v_n(t,x;f) \ := \ & \sup\Big\{ f(\mathbf{x}_{d_1} + (1-t)b(t,\mathbf{x}_{d_1}), \phi_2(1)) - \int_t^1 \ell\Big(\frac{d\phi_2(s)}{ds}\Big) ds \Big| \\ & \phi_2(t) = \mathbf{x}_{d_1,d}, \Big|\frac{d\phi_2(s)}{ds}\Big| \le n \Big\}. \end{aligned} \quad (2.17)$$

Then, from Theorem 10.1 in p. 95 of [7], under (A.1), $v_n(t,x;f)$ is the unique bounded uniformly continuous viscosity solution of the following HJ Eqn: for $(t,x) \in (0,1) \times \mathbf{R}^d$,

$$\begin{aligned}\frac{\partial v(t,x)}{\partial t} + < \nabla_{d_1} v(t,x), b(t, \mathbf{x}_{d_1}) > + h_n(\nabla_{d_1,d} v(t,x)) &= 0,\\ v(1,x) &= f(x), \qquad (2.18)\end{aligned}$$

where

$$h_n(z) := \sup\{< u, z > -\ell(u) | u \in \mathbf{R}^{d-d_1}, |u| \le n\}.$$

Let $\overline{v}$ be a bounded uniformly continuous viscosity solution of (2.10) with $\overline{v}(1,x) = f(x)$. Then it is a bounded uniformly continuous viscosity supersolution of (2.18) with $\overline{v}(1,x) = f(x)$ and

$$v_n(t,x;f) \le \overline{v}(t,x) \qquad (2.19)$$

from Theorem 9.1 in p. 86 of [7]. Let $n \to \infty$ in (2.19). Then we obtain $v(t,x;f|\varphi) \le \overline{v}(t,x)\Box$.

2.2 Convergence Theorem

Let $2 \le k \le d$, $0 = d_0 < d_1 < \cdots < d_k = d$ and
(A.4) $c_i \in LSC(\mathbf{R}^{d_i - d_{i-1}} \times \mathbf{R}^{d_i - d_{i-1}} : [0,\infty))$ $(i = 1, \cdots, k)$.
For $\varepsilon \ge 0$, P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned}T^{\varepsilon}(P_0, P_1) := \inf\Big\{\int_{\mathbf{R}^d \times \mathbf{R}^d} \sum_{i=1}^{k} \varepsilon^{i-1} c_i(\mathbf{x}_{d_{i-1},d_i}, \mathbf{y}_{d_{i-1},d_i}) \mu(dxdy) \Big| \\ \mu(dx \times \mathbf{R}^d) = P_0(dx), \mu(\mathbf{R}^d \times dy) = P_1(dy)\Big\}. \qquad (2.20)\end{aligned}$$

It is known that if $c_i(x,y) = \ell_i(y-x)$ and ℓ_i is strictly convex and superlinear $(i = 1, \cdots, k)$ and if $P_0(dx)$ is absolutely continuous with respect to the Lebesgue measure dx, then $T^{\varepsilon}(P_0, P_1)$ has the unique minimizer, provided that it is finite (see e.g. [21], [24], [25]).

$$\begin{aligned}T_1(P_{0,1}, P_{1,1}) := \inf\Big\{\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_1}} c_1(x,y) \mu(dxdy) \Big| \\ \mu(dx \times \mathbf{R}^{d_1}) = P_{0,1}(dx), \mu(\mathbf{R}^{d_1} \times dy) = P_{1,1}(dy)\Big\}, \qquad (2.21)\end{aligned}$$

where $P_{t,i}(d\mathbf{x}_{d_i}) := P_t(d\mathbf{x}_{d_i} \times \mathbf{R}^{d-d_i})$ $(t = 0, 1)$. For $i \ge 2$ and $\nu_{i-1} \in \mathcal{M}_1(\mathbf{R}^{2d_{i-1}})$,

$$\begin{aligned} T_i(P_{0,i}, P_{1,i}|\nu_{i-1}) \quad := \quad & \inf\Big\{ \int_{\mathbf{R}^{d_i}\times\mathbf{R}^{d_i}} c_i(\mathbf{x}_{d_{i-1},d_i}, \mathbf{y}_{d_{i-1},d_i})\mu(dxdy) \Big| \\ & \mu(d\mathbf{x}_{d_{i-1}} \times \mathbf{R}^{d_i-d_{i-1}} \times d\mathbf{y}_{d_{i-1}} \times \mathbf{R}^{d_i-d_{i-1}}) = \nu_{i-1}(d\mathbf{x}_{d_{i-1}}d\mathbf{y}_{d_{i-1}}), \\ & \mu(dx \times \mathbf{R}^{d_i}) = P_{0,i}(dx), \mu(\mathbf{R}^{d_i} \times dy) = P_{1,i}(dy)\Big\}. \qquad (2.22) \end{aligned}$$

The following theorem can be proved in the same way as [2] (see also section 1) and is proved for the readers' convenience.

Theorem 2.2 *Let P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$. Suppose that $k = 2$ and (A.4) holds and that $T_1(P_{0,1}, P_{1,1})$ and $T_2(P_0, P_1|\nu_1)$ have the unique minimizers ν_1 and ν_2, respectively. Then a minimizer of $T^\varepsilon(P_0, P_1)$ exists and weakly converges to ν_2 as $\varepsilon \to 0$ and the following holds:*

$$\lim_{\varepsilon\to 0}\int_{\mathbf{R}^d\times\mathbf{R}^d} c_1(\mathbf{x}_{d_1}, \mathbf{y}_{d_1})\mu^\varepsilon(dxdy) = T_1(P_{0,1}, P_{1,1}), \qquad (2.23)$$

$$\lim_{\varepsilon\to 0}\int_{\mathbf{R}^d\times\mathbf{R}^d} c_2(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu^\varepsilon(dxdy) = T_2(P_0, P_1|\nu_1). \qquad (2.24)$$

(Proof). In the same way as in the proof of Proposition 2.1, by a standard method, one can show that $T^\varepsilon(P_0, P_1)$ has a minimizer μ^ε, since

$$T^\varepsilon(P_0, P_1) \le T_1(P_{0,1}, P_{1,1}) + \varepsilon T_2(P_0, P_1|\nu_1) < +\infty. \qquad (2.25)$$

Since the set of μ for which $\mu(dx \times \mathbf{R}^d) = P_0(dx)$ and $\mu(\mathbf{R}^d \times dy) = P_1(dy)$ is compact, any sequence $\{\mu^{\varepsilon_n}\}_{n\ge 1}$ ($\varepsilon_n \to 0$ as $n \to \infty$) has a weakly convergent subsequence $\{\mu^{\varepsilon_{n(\ell)}}\}_{\ell\ge 1}$ and for the limit μ,

$$\mu_1(d\mathbf{x}_{d_1}d\mathbf{y}_{d_1}) := \mu(d\mathbf{x}_{d_1} \times \mathbf{R}^{d-d_1} \times d\mathbf{y}_{d_1} \times \mathbf{R}^{d-d_1})$$

is the minimizer of $T_1(P_{0,1}, P_{1,1})$ by the uniqueness of the minimizer and (2.23) holds. Indeed, from (2.25),

$$\begin{aligned} T_1(P_{0,1}, P_{1,1}) \le & \int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_1}} c_1(x,y)\mu_1(dxdy) = \int_{\mathbf{R}^d\times\mathbf{R}^d} c_1(\mathbf{x}_{d_1}, \mathbf{y}_{d_1})\mu(dxdy) \\ \le & \liminf_{\ell\to\infty} T^{\varepsilon_{n(\ell)}}(P_0, P_1) \le \limsup_{\ell\to\infty} T^{\varepsilon_{n(\ell)}}(P_0, P_1) \\ \le & T_1(P_{0,1}, P_{1,1}). \qquad (2.26) \end{aligned}$$

Since

$$T_1(P_{0,1}, P_{1,1}) + \varepsilon \int_{\mathbf{R}^d\times\mathbf{R}^d} c_2(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu^\varepsilon(dxdy) \le T^\varepsilon(P_0, P_1), \tag{2.27}$$

we also have, from (2.25) and (2.27),

$$\begin{aligned} T_2(P_0, P_1|\nu_1) &\le \int_{\mathbf{R}^d\times\mathbf{R}^d} c_2(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu(dxdy) \\ &\le \liminf_{\ell\to\infty} \int_{\mathbf{R}^d\times\mathbf{R}^d} c_2(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu^{\varepsilon_{n(\ell)}}(dxdy) \\ &\le \limsup_{\ell\to\infty} \int_{\mathbf{R}^d\times\mathbf{R}^d} c_2(\mathbf{x}_{d_1,d}, \mathbf{y}_{d_1,d})\mu^{\varepsilon_{n(\ell)}}(dxdy) \\ &\le T_2(P_0, P_1|\nu_1). \end{aligned} \tag{2.28}$$

The uniqueness of the minimizer of $T_2(P_0, P_1|\nu_1)$ completes the proof.□

Theorem 2.3 *Let P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$. Suppose that (A.4) holds, that $T_1(P_{0,1}, P_{1,1})$ and $T_i(P_{0,i}, P_{1,i}|\nu_{i-1})$ have the unique minimizers ν_1 and ν_i $(i = 2, \cdots, k)$, respectively and that $\nu \mapsto T_i(P_{0,i}, P_{1,i}|\nu)$ is continuous $(i = 3, \cdots, k)$. Then a minimizer of $T^\varepsilon(P_0, P_1)$ exists and weakly converges to ν_k as $\varepsilon \to 0$ and the following holds:*

$$\lim_{\varepsilon\to 0} \int_{\mathbf{R}^d\times\mathbf{R}^d} c_1(\mathbf{x}_{d_1}, \mathbf{y}_{d_1})\mu^\varepsilon(dxdy) = T_1(P_{0,1}, P_{1,1}), \tag{2.29}$$

$$\lim_{\varepsilon\to 0} \int_{\mathbf{R}^d\times\mathbf{R}^d} c_i(\mathbf{x}_{d_{i-1},d_i}, \mathbf{y}_{d_{i-1},d_i})\mu^\varepsilon(dxdy) = T_i(P_{0,i}, P_{1,i}|\nu_{i-1})(i = 2, \cdots, k). \tag{2.30}$$

(Proof). In the same way as in (2.25), one can show that $T^\varepsilon(P_0, P_1)$ has a minimizer μ^ε and that any subsequence $\{\mu^{\varepsilon_n}\}_{n\ge 1}$ $(\varepsilon_n \to 0$ as $n \to \infty)$ has a weakly convergent subsequence $\{\mu^{\varepsilon_{n(\ell)}}\}_{\ell\ge 1}$. Let μ denote the weak limit of $\mu^{\varepsilon_{n(\ell)}}$ as $\ell \to \infty$. We prove the theorem by induction. For $i = 2, \cdots, k$,

$$\begin{aligned} &T^\varepsilon_{i-1}(P_{0,i-1}, P_{1,i-1}) \\ :=\ &\inf\Big\{\int_{\mathbf{R}^{d_{i-1}}\times\mathbf{R}^{d_{i-1}}} \sum_{j=1}^{i-1} \varepsilon^{j-1} c_j(\mathbf{x}_{d_{j-1},d_j}, \mathbf{y}_{d_{j-1},d_j})\nu(dxdy)\Big| \\ &\qquad \nu(dx\times\mathbf{R}^{d_{i-1}}) = P_{0,i-1}(dx), \nu(\mathbf{R}^{d_{i-1}}\times dy) = P_{1,i-1}(dy)\Big\}. \end{aligned} \tag{2.31}$$

Let μ^ε_{i-1} and $\nu^\varepsilon_{i,j}$ denote a minimizer of $T^\varepsilon_{i-1}(P_{0,i-1}, P_{1,i-1})$ and $T_j(P_{0,j}, P_{1,j}|\nu^\varepsilon_{i,j-1})$ $(j = i, \cdots, k)$, respectively, where $\nu^\varepsilon_{i,i-1} := \mu^\varepsilon_{i-1}$. Then

$$\begin{aligned}
&T^{\varepsilon}_{i-1}(P_{0,i-1}, P_{1,i-1}) + \int_{\mathbf{R}^d\times\mathbf{R}^d} \sum_{j=i}^{k} \varepsilon^{j-1} c_j(\mathbf{x}_{d_{j-1},d_j}, \mathbf{y}_{d_{j-1},d_j})\mu^{\varepsilon}(dxdy)\\
\le\ & T^{\varepsilon}(P_0, P_1)\\
\le\ & T^{\varepsilon}_{i-1}(P_{0,i-1}, P_{1,i-1}) + \sum_{j=i}^{k} \varepsilon^{j-1} T_j(P_{0,j}, P_{1,j}|\nu^{\varepsilon}_{i,j-1}). \qquad (2.32)
\end{aligned}$$

From Theorem 2.2, $\mu^{\varepsilon}_2 \longrightarrow \nu_2$ as $\varepsilon \longrightarrow 0$ and (2.23)-(2.24) holds. Suppose that $\mu^{\varepsilon}_i \longrightarrow \nu_i$ as $\varepsilon \longrightarrow 0$ for $i \le k-1$. In the same way as in Theorem 2.2, one can show that for $j = 1, 2$,

$$\mu(d\mathbf{x}_{d_j} \times \mathbf{R}^{d-d_j} \times d\mathbf{y}_{d_j} \times \mathbf{R}^{d-d_j}) = \nu_j(d\mathbf{x}_{d_j} d\mathbf{y}_{d_j}). \qquad (2.33)$$

Suppose that (2.33) holds for $j = i-1$. Then, from (2.32) and the assumption of induction,

$$\begin{aligned}
T_i(P_{0,i}, P_{1,i}|\nu_{i-1}) \ &\le\ \int_{\mathbf{R}^d\times\mathbf{R}^d} c_i(\mathbf{x}_{d_{i-1},d_i}, \mathbf{y}_{d_{i-1},d_i})\mu(dxdy)\\
&\le\ \liminf_{\ell\to\infty} \int_{\mathbf{R}^d\times\mathbf{R}^d} c_i(\mathbf{x}_{d_{i-1},d_i}, \mathbf{y}_{d_{i-1},d_i})\mu^{\varepsilon_{n(\ell)}}(dxdy)\\
&\le\ \limsup_{\ell\to\infty} \int_{\mathbf{R}^d\times\mathbf{R}^d} c_i(\mathbf{x}_{d_{i-1},d_i}, \mathbf{y}_{d_{i-1},d_i})\mu^{\varepsilon_{n(\ell)}}(dxdy)\\
&\le\ \lim_{\ell\to\infty} T_i(P_{0,i}, P_{1,i}|\mu^{\varepsilon_{n(\ell)}}_{i-1}) = T_i(P_0, P_1|\nu_{i-1}). \qquad (2.34)
\end{aligned}$$

(2.34) implies (2.30) and the uniqueness of the minimizer ν_i of $T_i(P_{0,i}, P_{1,i}|\nu_{i-1})$ implies that (2.33) holds for $j = i$.□

From (2.32), we also have

Proposition 2.4 *Suppose that the assumption in Theorem 2.3 holds. Then, for* $i = 1, \cdots, k-1$,

$$0 \le \frac{\int_{\mathbf{R}^d\times\mathbf{R}^d} \sum_{j=1}^{i} \varepsilon^{j-1} c_j(\mathbf{x}_{d_{j-1},d_j}, \mathbf{y}_{d_{j-1},d_j})\mu^{\varepsilon}(dxdy) - T^{\varepsilon}_i(P_{0,i}, P_{1,i})}{\varepsilon^i} \longrightarrow 0 \quad (\varepsilon \longrightarrow 0). \qquad (2.35)$$

We don't know the real convergence rate of (2.35).

Example 2.2 *Let* P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$. *Suppose that*
(i) $d_{i+1} = d_i + 1$ *(*$i = 1, \cdots, k-1$*),*

(ii) $c_i(x,y) = \ell_i(y-x)$ *and* $\ell_i : \mathbf{R}^{d_i} \mapsto [0,\infty)$ *is strictly convex and superlinear* $(i = 1,\cdots,k)$*,*
(iii) P_0 *is absolutely continuous with respect to the Lebesgue measure* dx*,*
(iv) $T_1(P_{0,1}, P_{1,1})$ *is finite.*
Then $T_1(P_{0,1}, P_{1,1})$ *has the unique minimizer* ν_1 *which can be written as follows:*

$$\nu_1(d\mathbf{x}_{d_1} d\mathbf{y}_{d_1}) = P_{0,1}(d\mathbf{x}_{d_1})\delta_{\phi_1(\mathbf{x}_{d_1})}(d\mathbf{y}_{d_1}), \tag{2.36}$$

where ϕ_1 *is a Borel measurable function (see e.g. [21], [24]).*
Suppose, in addition, that
(v) $T_i(P_{0,i}, P_{1,i}|\nu_{i-1})$ *is finite for* $i = 2,\cdots,k$*. (If* $T_i(P_{0,i}, P_{1,i}|\nu_{i-1})$ *is finite, then it has a minimizer (see the proof of Prop. 2.1).)*
Then the following holds:

$$\nu_i(d\mathbf{x}_{d_i} d\mathbf{y}_{d_i}) = P_{0,i}(d\mathbf{x}_{d_i})\delta_{\Phi_{\nu_0,\cdots,\nu_{i-1}}(\mathbf{x}_{d_i})}(d\mathbf{y}_{d_i}), \tag{2.37}$$

where $\Phi_{\nu_0,\cdots,\nu_{i-1}}(\mathbf{x}_{d_i}) := (\phi_{\nu_0}(\mathbf{x}_{d_1}),\cdots,\phi_{\nu_{i-1}}(\mathbf{x}_{d_i}))$*,* $\phi_{\nu_0} := \phi_1$ *and*

$$\begin{aligned}
\phi_{\nu_{i-1}}(\mathbf{x}_{d_i}) &:= (F_{\nu_{i-1},1}(\cdot|\mathbf{x}_{d_{i-1}}, \Phi_{\nu_0,\cdots,\nu_{i-2}}(\mathbf{x}_{d_{i-1}})))^{-1}(F_{\nu_{i-1},0}(x_{d_i}|\mathbf{x}_{d_{i-1}}, \Phi_{\nu_0,\cdots,\nu_{i-2}}(\mathbf{x}_{d_{i-1}}))), \\
& \quad F_{\nu_{i-1},1}(x|\mathbf{x}_{d_{i-1}}, \Phi_{\nu_0,\cdots,\nu_{i-2}}(\mathbf{x}_{d_{i-1}})) \\
&:= \nu_{i-1}(\mathbf{R}\times(-\infty,x]|(\mathbf{x}_{d_{i-1}}, \Phi_{\nu_0,\cdots,\nu_{i-2}}(\mathbf{x}_{d_{i-1}}))), \\
& \quad F_{\nu_{i-1},0}(x_{d_i}|\mathbf{x}_{d_{i-1}}, \Phi_{\nu_0,\cdots,\nu_{i-2}}(\mathbf{x}_{d_{i-1}})) \\
&:= \nu_{i-1}((-\infty,x]\times\mathbf{R}|(\mathbf{x}_{d_{i-1}}, \Phi_{\nu_0,\cdots,\nu_{i-2}}(\mathbf{x}_{d_{i-1}}))).
\end{aligned}$$

In particular, $\phi_{\nu_{i-1}}$ *is a minimizer of the following:*

$$\min\left\{\int_{\mathbf{R}^{d_i}} \ell_i(\phi(x)-x_{d_i})P_{0,i}(dx)\middle| P_{0,i}(\Phi_{\nu_0,\cdots,\nu_{i-2}},\phi)^{-1} = P_{1,i}\right\} = T_i(P_{0,i}, P_{1,i}|\nu_{i-1}). \tag{2.38}$$

3 Stochastic version of Knothe-Rosenblatt type rearrangement.

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) $X(t) = X(0) + \int_0^t \beta_X(s, X)ds + W_X(t)$ $(0 \le t \le 1)$.
Here $\mathcal{B}(C([0,t]))_+ := \cap_{s>t}\mathcal{B}(C([0,s]))$, $\mathcal{B}(C([0,t]))$ and W_X denote the Borel σ-field of $C([0,t])$ and an $(\mathcal{F}_t^X)$-Brownian motion, respectively, and $\mathcal{F}_t^X := \sigma[X(s) : 0 \le s \le t]$ (see e.g. [14]). Let $d \ge 2$ and $1 \le d_1 < d$, and let $b_1 : [0,1] \times \mathbf{R}^{d_1} \mapsto \mathbf{R}^{d_1}$ be a Borel measurable function such that the following SDE has a weak solution for a given initial distribution:

$$dX_1(t) = b_1(t, X_1(t))dt + dW_{X_1}(t). \tag{3.1}$$

Let $L(t, x; u) : [0,1] \times \mathbf{R}^d \times \mathbf{R}^{d-d_1} \mapsto [0, \infty)$.

A minimizer of the following can be considered as the stochastic optimal control (SOC for short) version of the Knothe Rosenblatt type rearrangement: for P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned} V(P_0, P_1|b_1) \;:=\; & \inf\Big\{E\Big[\int_0^1 L(t, Y(t); \beta_{Y,2}(t, Y))dt\Big]\Big|Y \in \mathcal{A}, \beta_{Y,1}(t, Y) = b_1(t, Y_1(t)), \\ & PY(0)^{-1} = P_0, PY(1)^{-1} = P_1\Big\}, \end{aligned} \tag{3.2}$$

where we write $\beta_Y(t, Y) = (\beta_{Y,1}(t, Y), \beta_{Y,2}(t, Y)) \in \mathbf{R}^{d_1} \times \mathbf{R}^{d-d_1}$.

Example 3.1 *For P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$, take T_{KR} in section 1 and, on a complete filtered probability space, consider*

$$Z(t) = Z(0) + \int_0^t \frac{T_{KR}(Z(0)) - Z(s)}{1-s}ds + W_Z(t). \tag{3.3}$$

Then $Z(1) = T_{KR}(Z(0))$. In particular, $PZ(1)^{-1} = P_1$, provided $PZ(0)^{-1} = P_0$. Besides, $\beta_{Z,i}(t, Z) = \beta_{\mathbf{Z}_i}(t, \mathbf{Z}_i)$ for all $i = 1, \cdots, d$. Suppose that $p \in [1, 2)$ and that $\int_{\mathbf{R}^d} |x|^p(P_0(dx) + P_1(dx))$ is finite. Then

$$E\Big[\int_0^1 \Big|\frac{T_{KR}(Z(0)) - Z(s)}{1-s}\Big|^p ds\Big] < \infty. \tag{3.4}$$

Indeed, $W_o(t) := Z(t) - Z(0) - (T_{KR}(Z(0)) - Z(0))t$ is a tided down brownian motion starting and arriving at o, and

$$\frac{T_{KR}(Z(0)) - Z(s)}{1-s} = T_{KR}(Z(0)) - Z(0) - \frac{W_o(s)}{1-s}.$$

We describe our assumption in this section to show the existence of the stochastic analogue of the Knothe Rosenblatt type rearrangement.

(H.1). (i) $L \in C([0,1] \times \mathbf{R}^d \times \mathbf{R}^{d-d_1} : [0,\infty))$, (ii) $u \mapsto L(t,x;u)$ is strictly convex.
(H.2). There exists $\gamma > 1$ such that

$$\liminf_{|u|\to\infty} \frac{\inf\{L(t,x;u) : (t,x) \in [0,1] \times \mathbf{R}^d\}}{|u|^\gamma} > 0. \tag{3.5}$$

(H.3).

$$\Delta L(\varepsilon_1, \varepsilon_2) := \sup \frac{L(t,x;u) - L(s,y;u)}{1 + L(s,y;u)} \to 0 \quad \text{as } \varepsilon_1,\ \varepsilon_2 \to 0, \tag{3.6}$$

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 $u \in \mathbf{R}^d$.

The following can be proved in the same way as Prop. 2.1 in [19], and the proof is omitted.

Proposition 3.1 *Suppose that (H.1)-(H.3) hold. Then for any $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$, $V(P_0, P_1|b_1)$ has a minimizer, provided it is finite.*

3.1 Duality Theorem

We consider the following HJB Equation:

$$\begin{aligned} \frac{\partial v(t,x)}{\partial t} + \frac{1}{2}\Delta v(t,x) + &< \nabla_{\mathbf{x}_{d_1}} v(t,x), b_1(t, \mathbf{x}_{d_1}) > \\ &+ H(t,x; \nabla_{\mathbf{x}_{d_1,d}} v(t,x)) \;=\; 0, \end{aligned} \tag{3.7}$$

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

$$H(t,x;z) := \sup\{< u, z > - L(t,x;u) | u \in \mathbf{R}^{d-d_1}\} \quad (z \in \mathbf{R}^{d-d_1}).$$

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

$$\begin{aligned} u(t,x; f|b_1)(x) \;:=\; \sup\Big\{ & E\Big[f(Y(1)) - \int_t^1 L(s, Y(s); \beta_{Y,2}(s,Y))ds\Big]\Big| \\ & Y(t) = x, \beta_{Y,1}(s,Y) = b_1(s, Y_1(s)), Y \in \mathcal{A}\Big\}. \end{aligned} \tag{3.8}$$

(H.4). (i) $L(t,x;o)$ is bounded; (ii) $\Delta L(0,\infty)$ is finite;(iii) $b_1 \in C^{1,2}([0,1] \times \mathbf{R}^d) \cap C_b^{0,1}([0,1] \times \mathbf{R}^d)$, $|D_x L(t,x;u)|/(1 + L(t,x;u))$ is bounded on $[0,1] \times \mathbf{R}^d \times \mathbf{R}^{d_1}$ and $D_u L(t,x;u)$ is bounded on $[0,1] \times \mathbf{R}^d \times B_R$ for all $R > 0$, where $B_R := \{x \in \mathbf{R}^{d_1} | |x| \le R\}$.

The following can be proved in the same way as Theorem 11.1 in IV.11 of [7], and the proof is omitted.

Proposition 3.2 *Suppose that (H.1)-(H.2) and (H.4,i,iii) hold. Then, for any* $f \in C^5(\mathbf{R}^d) \cap C_b^3(\mathbf{R}^d)$, $u(t,x;f|b_1) \in C^{1,2}([0,1] \times \mathbf{R}^d) \cap C_b^{0,1}([0,1] \times \mathbf{R}^d)$ *and is the unique classical solution of the HJB Equation (3.7) with* $v(1,x) = f(x)$.

It is easy to see that the following holds:

$$\begin{aligned} V(P_0, P_1|b_1) \quad &:= \quad \inf\Big\{ E\Big[\int_0^1 \frac{L(t,Y(t);\beta_{Y,2}(t,Y))}{1_{\{b_1(t,Y_1(t))\}}(\beta_{Y,1}(t,Y))} dt\Big] \Big| Y \in \mathcal{A}, \\ &\qquad PY(0)^{-1} = P_0, PY(1)^{-1} = P_1 \Big\}, \end{aligned} \tag{3.9}$$

which implies the duality theorem for $V(P_0, P_1|b_1)$.

Theorem 3.1 *Suppose that (H.1)-(H.4) hold. Then for any* $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$V(P_0, P_1|b_1) = \sup\Big\{ \int_{\mathbf{R}^d} v(1,y) P_1(dy) - \int_{\mathbf{R}^d} v(0,x) P_0(dx) \Big\}, \tag{3.10}$$

where the supremum is taken over all classical solutions v *of (3.7) with* $v(1,y) \in C_b^\infty(\mathbf{R}^d)$.

(Proof). Under (H.1)-(H.3) and (H.4,i,ii), (3.9) implies that $V(P_0, \cdot|b_1)$ is convex and lower-semicontinuous, which can be proved in the same way as in [19] and is not identically equal to infinity by considering the case where $\beta_{Y,2}(s,Y) = o$ from (H.4,i). Hence, from Theorem 2.2.15 and Lemma 3.2.3 in [3],

$$V(P_0, P_1|b_1) = \sup\Big\{ \int_{\mathbf{R}^d} f(y) P_1(dy) - V(P_0, \cdot|b_1)^*(f) \Big| f \in C_b(\mathbf{R}^d) \Big\}, \tag{3.11}$$

where

$$V(P_0, \cdot|b_1)^*(f) := \sup\Big\{ \int_{\mathbf{R}^d} f(y) P(dy) - V(P_0, P|b_1) \Big| P \in \mathcal{M}_1(\mathbf{R}^d) \Big\}. \tag{3.12}$$

One can replace $C_b(\mathbf{R}^d)$ by $C_b^\infty(\mathbf{R}^d)$ in (3.11) in the same way as in the proof of Theorem 2.1 in [19]. For $f \in C_b^\infty(\mathbf{R}^d)$, from Proposition 3.2,

$$
\begin{aligned}
& V(P_0,\cdot|b_1)^*(f) \\
= \ & \sup\Big\{E\Big[f(Y(1)) - \int_0^1 \frac{L(t,Y(t);\beta_{Y,2}(t,Y))}{1_{b_1(t,Y_1(t))}(\beta_{Y,1}(t,Y))}dt\Big]\Big|Y \in \mathcal{A}, P(Y(0))^{-1} = P_0\Big\} \\
= \ & \int_{\mathbf{R}^d} u(0,x;f|b_1)P_0(dx),
\end{aligned} \tag{3.13}
$$

where the optimal control is $\beta_{Y,2}(t,Y) = \nabla_{d_1,d}u(t,Y(t);f|b_1)$.□

As a corollary to Theorem 3.1, in the same way as [19], we easily obtain

Corollary 3.1 *Suppose that (H.1)-(H.4) hold. Then for any $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$ for which $V(P_0, P_1|b_1)$ is finite, there exists a Borel measurable function $b_2^o : [0,1] \times \mathbf{R}^d \mapsto \mathbf{R}^{d-d_1}$ such that for a minimizer $\{Y(t)\}_{0\le t\le 1}$, $\beta_{Y,2}(t,Y) = b_2^o(t,Y(t))$.*

We consider the following marginal problem:

$$
\mathrm{v}(P_0,P_1|b_1) := \inf \int_0^1 dt \int_{\mathbf{R}^d} L(t,x;B_2(t,x))Q_t(dx), \tag{3.14}
$$

where the infimum is taken over all $\{Q_t(dx)\}_{0\le t\le 1} \subset \mathcal{M}_1(\mathbf{R}^d)$ for which $B_1 = b_1$, $Q_t = P_t$ $(t = 0,1)$ and

$$
\frac{\partial Q_t(dx)}{\partial t} = \frac{1}{2}\triangle Q_t(dx) - div(B(t,x)Q_t(dx)),
$$

in a weak sense. Here we write $B(t,x) = (B_1(t,x), B_2(t,x)) \in \mathbf{R}^{d_1} \times \mathbf{R}^{d-d_1}$.

In the same way as [17], we have

Theorem 3.2 *Suppose that (H.1)-(H.4) hold. Then for any $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,*

$$
\mathrm{v}(P_0,P_1|b_1) = \sup\Big\{\int_{\mathbf{R}^d} v(1,y)P_1(dy) - \int_{\mathbf{R}^d} v(0,x)P_0(dx)\Big\}, \tag{3.15}
$$

where the supremum is taken over all classical solutions v of (3.7) with $v(1,y) \in C_b^\infty(\mathbf{R}^d)$. In particular, $V(P_0,P_1|b_1) = \mathrm{v}(P_0,P_1|b_1)(\in [0,\infty))$.

We introduce an additional assumption to formulate the duality theorem in the framework of the theory of viscosity solutions.

(H.4)'. (i) $\partial L(t,x;u)/\partial t$ and $D_xL(t,x;u)$ is bounded on $[0,1] \times \mathbf{R}^d \times B_R$ for all $R > 0$; (ii) $\Delta L(0,\infty)$ is finite;(iii) $b_1 \in C_b^1([0,1] \times \mathbf{R}^d)$.

In the same way as in Lemma 4.5 in [17], one can prove

Proposition 3.3 *Suppose that (H.1)-(H.3) and (H.4)' hold. Then for any* $f \in UC_b(\mathbf{R}^d)$, $u(t,x;f|b_1)$ *is a bounded continuous viscosity solution of (3.7) with* $v(1,x) = f(x)$ *and for any* $Q \in \mathcal{M}_1(\mathbf{R}^d)$ *and* $t \in [0,1]$,

$$\begin{aligned}\int_{\mathbf{R}^d} u(t,x;f|b_1)Q(dx) &= \sup\Big\{E\Big[f(Y(1)) - \int_t^1 L(s,Y(s);\beta_{Y,2}(s,Y))ds\Big]\Big| \\ &\qquad PY^{-1}(t) = Q, \beta_{Y,1}(s,Y) = b_1(s,Y_1(s)), Y \in \mathcal{A}\Big\}.\end{aligned} \quad (3.16)$$

In addition, for any bounded continuous viscosity solution u *of (3.7) with* $u(1,x) = f(x)$, $u(t,x) \geq u(t,x;f|b_1)$, *that is,* $u(t,x;f|b_1)$ *is minimal.*

In the same was as in Theorem 3.1, from Prop. 3.3, we have

Theorem 3.3 *Suppose that (H.1)-(H.3) and (H.4)' hold. Then for any* $P_0, P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$V(P_0,P_1|b_1) = \sup\Big\{\int_{\mathbf{R}^d} v(1,y)P_1(dy) - \int_{\mathbf{R}^d} v(0,x)P_0(dx)\Big\}, \quad (3.17)$$

where the supremum is taken over all bounded continuous viscosity solutions $v(t,x;f)$ *of (3.7) with* $v(1,x) \in C_b^\infty(\mathbf{R}^d)$.

Remark 3.1 *(H.3) and (i) in (H.4)' implies (i) in (H.4).*

3.2 Convergence Theorem

Let $L_1 : [0,1] \times \mathbf{R}^{d_1} \times \mathbf{R}^{d_1} \mapsto [0,\infty)$ and $L_2 : [0,1] \times \mathbf{R}^d \times \mathbf{R}^{d-d_1} \mapsto [0,\infty)$. For $\varepsilon > 0$, P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$,

$$\begin{aligned}V^\varepsilon(P_0,P_1) &:= \inf\Big\{E\Big[\sum_{i=1}^2 \varepsilon^{i-1}\int_0^1 L_i(t,\mathbf{Y}_i(t);\beta_{Y,i}(t,Y))dt\Big]\Big| \\ &\qquad PY(0)^{-1} = P_0, PY(1)^{-1} = P_1, Y \in \mathcal{A}\Big\},\end{aligned} \quad (3.18)$$

where $\mathbf{Y}_1(t) := Y_1(t)$ and $\mathbf{Y}_2(t) := Y(t)$ for $Y(t) = (Y_1(t), Y_2(t)) \in \mathbf{R}^{d_1} \times \mathbf{R}^{d-d_1}$.

If (H.1)-(H.3) holds for $L = L_i$ for all $i = 1,2$, then $V^\varepsilon(P_0,P_1)$ has a minimizer, provided it is finite (see Prop. 2.1 in [19]).

$$\begin{aligned}V_1(P_{0,1},P_{1,1}) &:= \inf\Big\{E\Big[\int_0^1 L_1(t,Y(t);\beta_Y(t,Y))dt\Big]\Big|Y \in \mathcal{A}_1, \\ &\qquad PY(0)^{-1} = P_{0,1}, PY(1)^{-1} = P_{1,1}\Big\},\end{aligned} \quad (3.19)$$

where $\mathcal{A}_1$ denotes $\mathcal{A}$ with $d = d_1$.

Remark 3.2 *If (H.1)-(H.4) with $L = L_1$ holds and that $V_1(P_{0,1}, P_{1,1})$ is finite. Then there exists a Borel measurable function $b : [0,1] \times \mathbf{R}^{d_1} \mapsto \mathbf{R}^{d_1}$ such that for any minimizer $\{Y(t)\}_{0\le t\le 1}$ of $V_1(P_{0,1}, P_{1,1})$, $\beta_Y(t, Y) = b(t, Y(t))$ (see [19]).*

Let b_1 denote the drift vector of the minimizer of $V_1(P_{0,1}, P_{1,1})$, provided it exists and let $V_2(P_0, P_1|b_1)$ denote $V(P_0, P_1|b_1)$ with $L = L_2$. Then

Theorem 3.4 *Let P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$. Suppose that (H.1)-(H.3) with $L = L_i$ holds ($i = 1, 2$) and that $V_1(P_{0,1}, P_{1,1})$ and $V_2(P_0, P_1|b_1)$ is finite and have the unique minimizers $\{X_1(t)\}_{0\le t\le 1}$ and $\{X(t)\}_{0\le t\le 1}$, respectively. Then a minimizer $\{Y^\varepsilon(t)\}_{0\le t\le 1}$ of $V^\varepsilon(P_0, P_1)$ exists and weakly converges to $\{X(t)\}_{0\le t\le 1}$ as $\varepsilon \to 0$. In particular,*

$$\lim_{\varepsilon\to 0} E\Big[\int_0^1 L_1(t, Y_1^\varepsilon(t); \beta_{Y^\varepsilon,1}(t, Y^\varepsilon))dt\Big] = V_1(P_{0,1}, P_{1,1}), \tag{3.20}$$

$$\lim_{\varepsilon\to 0} E\Big[\int_0^1 L_2(t, Y^\varepsilon(t); \beta_{Y^\varepsilon,2}(t, Y^\varepsilon))dt\Big] = V_2(P_0, P_1|b_1). \tag{3.21}$$

(Proof) In the same way as Prop. 2.1 in [19], one can show that there exists a minimizer $Y^\varepsilon(t)$ of $V^\varepsilon(P_0, P_1)$ since

$$V^\varepsilon(P_0, P_1) \le V_1(P_{0,1}, P_{1,1}) + \varepsilon V_2(P_0, P_1|b_1). \tag{3.22}$$

In the same way as in Lemma 3.1 in [19], from (H.2), one can show that any sequence $\{Y^{\varepsilon_n}(\cdot)\}_{n\ge 1}$ in $\mathcal{A}$ ($\varepsilon_n \to 0$ as $n \to \infty$) has a weakly convergent subsequence $\{Y^{\varepsilon_{n(k)}}(\cdot)\}_{k\ge 1}$. Indeed,

$$E\Big[\int_0^1 L_1(t, Y_1^\varepsilon(t); \beta_{Y^\varepsilon,1}(t, Y^\varepsilon))dt\Big] \le V^\varepsilon(P_0, P_1), \tag{3.23}$$

$$E\Big[\int_0^1 L_2(t, Y^\varepsilon(t); \beta_{Y^\varepsilon,2}(t, Y^\varepsilon))dt\Big] \le V_2(P_0, P_1|b_1). \tag{3.24}$$

We prove (3.24). In the same way as in Lemma 3.1 in [19], from (H.1,ii), by Jensen's inequality,

$$\begin{aligned} V_1(P_{0,1}, P_{1,1}) &\le E\Big[\int_0^1 L_1(t, Y_1^\varepsilon(t); \beta_{Y_1^\varepsilon}(t, Y_1^\varepsilon))dt\Big] \\ &\le E\Big[\int_0^1 L_1(t, Y_1^\varepsilon(t); \beta_{Y^\varepsilon,1}(t, Y^\varepsilon))dt\Big]. \end{aligned} \tag{3.25}$$

Indeed, $Y_1^\varepsilon \in \mathcal{A}_1$ with

$$\beta_{Y_1^\varepsilon}(t, Y_1^\varepsilon) = E[\beta_{1,Y^\varepsilon}(t, Y^\varepsilon))|Y_1^\varepsilon(s), 0 \le s \le t]$$

(see e.g., p. 258 of [14]). (3.25) and (3.22) implies (3.24).

Let $Y^0(t)$ denote the weak limit of $\{Y^{\varepsilon_{n(k)}}(\cdot)\}_{k\ge 1}$ as $n \to \infty$. Then, again in the same way as in Lemma 3.1 in [19] and (3.25), from (H.1,ii) and (3.22)-(3.23), by Jensen's inequality,

$$\begin{aligned} V_1(P_{0,1}, P_{1,1}) &\le E\Big[\int_0^1 L_1(t, Y_1^0(t); \beta_{Y_1^0}(t, Y_1^0))dt\Big] \\ &\le E\Big[\int_0^1 L_1(t, Y_1^0(t); \beta_{Y^0,1}(t, Y^0))dt\Big] \\ &\le \liminf_{k\to\infty} V^{\varepsilon_{n(k)}}(P_0, P_1) \le \limsup_{k\to\infty} V^{\varepsilon_{n(k)}}(P_0, P_1) \\ &\le V_1(P_{0,1}, P_{1,1}). \end{aligned} \tag{3.26}$$

$\beta_{Y^0,1}(t, Y^0) = \beta_{Y_1^0}(t, Y_1^0)$ from the strict convexity of L_1 in u, and Y_1^0 is equal to the minimizer X_1 of $V_1(P_{0,1}, P_{1,1})$ by the uniqueness of the minimizer of $V_1(P_{0,1}, P_{1,1})$ and we obtain (3.20). From (3.24), we also have

$$\begin{aligned} V_2(P_0, P_1|b_1) &\le E\Big[\int_0^1 L_2(t, Y^0(t); \beta_{Y^0,2}(t, Y^0))dt\Big] \\ &\le \liminf_{k\to\infty} E\Big[\int_0^1 L_2(t, Y^{\varepsilon_{n(k)}}(t); \beta_{Y^{\varepsilon_n},2}(t, Y^{\varepsilon_{n(k)}}))dt\Big] \\ &\le \limsup_{k\to\infty} E\Big[\int_0^1 L_2(t, Y^{\varepsilon_{n(k)}}(t); \beta_{Y^{\varepsilon_n},2}(t, Y^{\varepsilon_{n(k)}}))dt\Big] \\ &\le V_2(P_0, P_1|b_1). \end{aligned} \tag{3.27}$$

The uniqueness of the minimizer of $V_2(P_0, P_1|b_1)$ completes the proof.□

One can easily prove

Corollary 3.2 *Let P_0, $P_1 \in \mathcal{M}_1(\mathbf{R}^d)$. Suppose that (H.1)-(H.3) with $L = L_i$ holds ($i = 1, 2$), that $\gamma = 2$ in (H.2), and that $V_1(P_{0,1}, P_{1,1})$ and $V_2(P_0, P_1|b_1)$ is finite. Then the minimizers $\{X_1(t)\}_{0\le t\le 1}$, $\{X(t)\}_{0\le t\le 1}$ and $\{Y^\varepsilon(t)\}_{0\le t\le 1}$ of $V_1(P_{0,1}, P_{1,1})$, $V_2(P_0, P_1|b_1)$ and $V^\varepsilon(P_0, P_1)$ exist uniquely, respectively. In addition, $\{Y^\varepsilon(t)\}_{0\le t\le 1}$ weakly converges to $\{X(t)\}_{0\le t\le 1}$ as $\varepsilon \to 0$ and (3.20)-(3.21) holds.*

From (3.21)-(3.22) and (3.25), we easily have

Proposition 3.4 *Supoose that the assumption in Theorem 3.3 holds. Then for any minimizer* $\{Y^\varepsilon\}_{0\le t\le 1}$ *of* $V^\varepsilon(P_0,P_1)$,

$$0 \le \frac{E\left[\int_0^1 L_1(t, Y_1^\varepsilon(t); \beta_{Y^\varepsilon,1}(t, Y^\varepsilon))dt\right] - V_1(P_{0,1}, P_{1,1})}{\varepsilon} \to 0 \quad (\varepsilon \to 0). \tag{3.28}$$

We don't know the real convergence rate!

4 Discussion

In section 2, Theorem 2.3, we assumed that $\nu \mapsto T_i(P_{0,i}, P_{1,i}|\nu)$ is continuous ($i = 3, \cdots, k$). This continuity is known only in the case of the Knothe-Rosenblatt rearrangement where the representation of the minimizer is known. It is difficult to prove that $\nu \mapsto T(P_0, P_1|\nu)$ is continuous, which is our future problem.
In section 3.2, we only considered the case where $k = 2$ because of the similar reason to above. The point is that we do not even know any example such as the Knothe-Rosenblatt rearrangement. This is also our future problem.
The Knothe-Rosenblatt rearrangement implies the Brunn-Minkowskii inequality. We would like to find, in future, the inequality which can be obtained by the result in section 3.

References

[1] Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math., **44**, no. 4, 375–417 (1991).

[2] Carlier, G., Galichon, F. and Santambrogio, F.: From Knothe's transport to Brenier's map and a continuation method for optimal transport, preprint.

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

[4] L.C. Evans, Partial differential equations. AMS, 1998.

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

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

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

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

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

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

[11] Kellerer, H. G.: Duality theorems for marginal problems. Z. Wahrsch. Verw. Gebiete 67, no. 4, 399–432 (1984).

[12] Knothe, H.: Contributions to the theory of convex bodies. Michigan Mathematical J. **4**, 39–52 (1957).

[13] Lévy, P.: Théorie de l'Addition des Variables Aléatoires. pp. 71-73, 121-123, Paris: Gauthier-Villars 1937.

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

[15] 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).

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

[17] Mikami, T.: Marginal problem for semimartingales via duality. International Conference for the 25th Anniversary of Viscosity Solutions, Gakuto International Series. Mathematical Sciences and Applications, Vol. 30, 133-152 (2008).

[18] Mikami, T.: Optimal transportation problem as stochastic mechanics. To appear in Selected Papers on Probability and Statistics. (AMS Translations Series 2, Vol. 227) Providence: Amer. Math. Soc. 2009.

[19] Mikami, T., Thieullen, M.: Duality Theorem for Stochastic Optimal Control Problem. Stoc. Proc. Appl. **116**, 1815-1835 (2006).

[20] Mikami, T., Thieullen, M.: Optimal Transportation Problem by Stochastic Optimal Control. SICON. **47**, 1127-1139 (2008).

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

[22] Rosenblatt, M.: Remarks on a multivariate transformation. Annals of Mathematical Statistics **23**, 470-472 (1952).

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

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

[25] Villani, C.: Optimal Transport: Old and New. Berlin Heidelberg New York Tokyo: Springer-Verlag 2009.