Pure and Applied Functional Analysis
Volume 9, Number 3, 2024, 811–824

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

A REMARK ON THE LAGRANGIAN FORMULATION OF OPTIMAL TRANSPORT WITH A NON–CONVEX COST

TOSHIO MIKAMI AND HARUKA YAMAMOTO

ABSTRACT. We study the Lagrangian formulation of a class of the Monge–Kantorovich optimal transportation problem. It can be considered a stochastic optimal transportation problem for absolutely continuous stochastic processes. A cost function and stochastic processes under consideration is not convex and have essentially bounded time derivatives almost surely, respectively. This paper is a continuation of the second author's master thesis.

1. Introduction

For $d \geq 1$, let $\mathcal{P}(\mathbb{R}^d)$ denote the space of all Borel probability measures on $\mathbb{R}^d$ endowed with weak topology. For $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$, let $\Pi(P_0, P_1)$ denote the set of $\mathbb{R}^d \times \mathbb{R}^d$–valued random variables (X_0, X_1) defined on a possibly different probability space such that $P^{X_i} = P_i, i = 0, 1$. Here P^X denotes the probability distribution of a random variable X. Throughout this paper, the probability space under consideration is not fixed.

Let $L : \mathbb{R}^d \to [0, \infty)$ be Borel measurable. The following is a typical Monge–Kantorovich optimal transportation problem: for $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$,

$$T(P_0, P_1) := \inf\{E[L(X_1 - X_0)] : (X_0, X_1) \in \Pi(P_0, P_1)\} \tag{1.1}$$

(see, e.g. [18, 21] and the references therein).

In the case where $L(u) = |u|^p$ for $p > 0$, we denote (1.1) by $T_p(P_0, P_1)$. In the case where $d = 1$, the minimizer of $T_p, p \geq 1$ was obtained in [6]. In the case where $d \geq 2$, the minimizer of T_2 was obtained in [2, 3] and the generalization to more general costs, including concave ones, was given in [8]. The probabilistic proof of the existence and the uniqueness of the minimizer of T_2 via the stochastic control approach was given in [13] by the zero–noise limit of Schrödinger's problem (see [1, 19, 20] for Schrödinger's problem and [9, 10, 11, 15, 23] for related topics). Schrödinger's problem is also called the entropic regularized optimal transport these days and plays a crucial role in data science (see [4, 5, 17] and the reference therein). Notice that $\mathbb{R}^d \ni u \mapsto |u|^p$ is convex for $p \in [1, \infty)$. Notice also that for $p \in (0, 1)$,

---

2010 *Mathematics Subject Classification.* 49Q22, 93E20, 49K45.

Key words and phrases. Lagrangian formulation, non–convex cost, Monge–Kantorovich problem, stochastic optimal transport.

This work was partially supported by JSPS KAKENHI Grant Number 19K03548.

$[0,\infty) \ni r \mapsto r^p$ is concave but $\mathbb{R}^d \ni u \mapsto |u|^p$ is not. Indeed, for $p > 0$,

$$0 = \left|\frac{1}{2}u + \frac{1}{2}(-u)\right|^p < |u|^p = \frac{1}{2}|u|^p + \frac{1}{2}|-u|^p, \quad u \in \mathbb{R}^d\backslash\{0\}.$$

Let $\mathcal{A}(P_0, P_1)$ denote the set of stochastic processes $X(\cdot)$ such that

$$X(\cdot) \in AC([0,1]), \text{a.s.}, \quad (X(0), X(1)) \in \Pi(P_0, P_1).$$

Here, for $T > 0$, $AC([0,T])$ denotes the space of all absolutely continuous functions from $[0,T]$ to $\mathbb{R}^d$. We consider the following stochastic optimal transport for absolutely continuous stochastic processes:

$$V(P_0, P_1) := \inf\left\{E\left[\int_0^1 L(X'(t))dt\right] : X(\cdot) \in \mathcal{A}(P_0, P_1)\right\}, \tag{1.2}$$

where $X'(t) := \frac{d}{dt}X(t)$.

We briefly describe the relation between T and V. The following holds without any assumption:

$$V(P_0, P_1) \le T(P_0, P_1). \tag{1.3}$$

If $L : \mathbb{R}^d \to [0,\infty)$ is convex, then the equality holds in (1.3) and $V(P_0, P_1)$ can be considered the Lagrangian formulation of $T(P_0, P_1)$ (see, e.g. [15] for the proof of (1.3) and also [12] for related topics).
If (i) $L(ru) \ge rL(u), 0 < r < 1, u \in \mathbb{R}^d$; (ii) $L(u)/|u| \to 0, |u| \to \infty$, and (iii) $T(P_0, P_1)$ is finite, then

$$V(P_0, P_1) = 0 \tag{1.4}$$

(see Appendix for the proof and also Theorem 2.8 and Corollary 2.11 in section 2). A typical example of such $L(u)$ is $|u|^p, p \in (0,1)$ (see Remark 2.1 in section 2 for more examples).

(1.3)–(1.4) imply that, to study the Lagrangian formulation of $T(P_0, P_1)$ when L is not convex, we have to modify a cost function or restrict a class of stochastic processes in (1.2).

We first modify a cost function and give two Lagrangian formulations for $T(P_0, P_1)$. For $t > 0, \varphi \in L^\infty([0,t])$,

$$1 \le N_1(\varphi)_t := \begin{cases} \dfrac{t||\varphi||_{\infty,t}}{|\int_0^t \varphi(s)ds|}, & \text{if } \int_0^t \varphi(s)ds \ne 0, \\ 1, & \text{otherwise}, \end{cases} \tag{1.5}$$

$$1 \le N_2(\varphi)_t := \begin{cases} \dfrac{t||\varphi||_{\infty,t}}{||\varphi||_{1,t}}, & \text{if } ||\varphi||_{1,t} > 0, \\ 1, & \text{otherwise}, \end{cases} \tag{1.6}$$

where

$$||\varphi||_{\infty,t} := ess.sup\left\{|\varphi(s)| : 0 \le s \le t\right\}, \quad ||\varphi||_{1,t} := \int_0^t |\varphi(s)|ds.$$

For simplicity, $||\varphi||_p := ||\varphi||_{p,1}$ for $p = 1, \infty$ and $N_i(\varphi) := N_i(\varphi)_1$, $i = 1, 2$.

Let $\ell : [0,\infty) \to [0,\infty)$ and $\ell(0) = 0$ (see (A1) in section 2). For $i = 1, 2$,

$$L_i(t,\varphi) := N_i(\varphi)\ell\left(\frac{|\varphi(t)|}{N_i(\varphi)}\right), \quad (t,\varphi) \in [0,1] \times L^\infty([0,1]), \tag{1.7}$$

$$\tilde{V}_i(P_0,P_1) := \inf\left\{E\left[\int_0^1 L_i(t,X')dt\right] : X(\cdot) \in \mathcal{A}_\infty(P_0,P_1)\right\}, \tag{1.8}$$

where

$$\mathcal{A}_\infty(P_0,P_1) := \{X(\cdot) \in \mathcal{A}(P_0,P_1) : ||X'||_\infty < \infty \text{ a.s.}\}.$$

We show that

$$T(P_0,P_1) = \tilde{V}_i(P_0,P_1), \quad i = 1, 2,$$

under different assumptions (see Theorems 2.2–2.4 in section 2).

Remark 1.1. If $L : \mathbb{R}^d \to [0,\infty)$ is convex, $L(0) = 0$ and $V(P_0,P_1)$ is finite, then the following holds (see Appendix for the proof): for $i = 1, 2$,

$$\begin{aligned} V(P_0,P_1) &= T(P_0,P_1) \\ &= \inf\left\{E\left[\int_0^1 \frac{1}{N_i(X')}L\left(N_i(X')X'(t)\right)dt\right] : X(\cdot) \in \mathcal{A}_\infty(P_0,P_1)\right\}. \end{aligned} \tag{1.9}$$

For $X(\cdot) \in \mathcal{A}_\infty(P_0,P_1)$ and $t \geq 0$, $L_i(t,X') \geq 0$ and $= 0$ if $X'(t) = 0$. When we consider minimizers of $\tilde{V}_i$, we assume that $\ell(u) > 0, u > 0$ so that we only have to consider $X(\cdot)$ such that $X(t) = X(0)$ if and only if $||X'||_{1,t} = 0$. In particular, we can assume that the following holds:

$$N_1(X') \geq N_2(X') \geq 1, \quad \text{a.s.}, \tag{1.10}$$

which implies the following:

$$|X'(t)| \geq \frac{|X'(t)|}{N_2(X')} \geq \frac{|X'(t)|}{N_1(X')}, \quad L_1(t,X') \geq L_2(t,X'), \tag{1.11}$$

provided $\ell(ru) \geq r\ell(u)$ for $(r,u) \in (0,1) \times (0,\infty)$ (see (3.13) and also Theorems 2.2–2.4 and Proposition 2.5 in section 2). The following also holds (see Appendix for the proof):

$$\begin{aligned} &\tilde{V}_i(P_0,P_1) \\ &= \inf\left\{E\left[\int_0^\tau \ell\left(|X'(t)|\right)dt\right] : \tau = \tau(\omega) \geq 1, X(\cdot) \in AC([0,\tau]),\right. \\ &\qquad \left. N_i(X')_\tau = \tau, \text{a.s.}, (X(0),X(\tau)) \in \Pi(P_0,P_1)\right\}. \end{aligned} \tag{1.12}$$

Next, we consider a restricted class of absolutely continuous stochastic processes with almost surely essentially bounded time derivatives. For $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$, and $B \subset \mathcal{P}([0,\infty))$,

$$\begin{aligned} &\mathcal{A}_\infty(P_0,P_1;B) \\ :=\ &\{(X(\cdot),M) : X(\cdot) \in \mathcal{A}_\infty(P_0,P_1), P^M \in B, ||X'||_\infty \leq M, \text{a.s.}\}, \\ &\Pi_\infty(P_0,P_1;B) \\ :=\ &\{(X_0,X_1,M) : (X_0,X_1) \in \Pi(P_0,P_1), P^M \in B, |X_1 - X_0| \leq M, \text{a.s.}\}. \end{aligned}$$

$$
\begin{aligned}
(1.13)\qquad & V(P_0, P_1; B) \\
:= \ & \inf\left\{ E\left[\int_0^1 \ell(|X'(t)|)dt\right] : (X(\cdot), M) \in \mathcal{A}_\infty(P_0, P_1; B)\right\}, \\
(1.14)\qquad & T^V(P_0, P_1; B) \\
:= \ & \inf\{E[\ell(M)M^{-1}|X_1 - X_0|; M > 0] : (X_0, X_1, M) \in \Pi_\infty(P_0, P_1; B)\}.
\end{aligned}
$$

We show that the following holds (see Theorem 2.8 in section 2):

$$V(P_0, P_1; B) = T^V(P_0, P_1; B).$$

It is a continuation of the second author's master thesis [22] in which she only considered the case where $L(u) = |u|^p, p \in (0, 1)$ and B is a set of a delta measure.

A generalization of our result to the case where stochastic processes under consideration are semimartingales is the first step to the theory of stochastic optimal transport with a non–convex cost and is our future project.

We state our results in section 2 and prove them in section 3. In Appendix, we give the proofs for (1.4), (1.9), and (1.12) for the sake of completeness.

2. Main result

In this section, we state our results. We first state the assumptions.

(A1). (i) $\ell : [0, \infty) \to [0, \infty)$, $\ell(0) = 0$,

$$\ell(ru) \geq r\ell(u), \quad (r, u) \in (0, 1) \times (0, \infty). \tag{2.1}$$

(ii) In (2.1), the equality does not hold for any $(r, u) \in (0, 1) \times (0, \infty)$. (iii) $\ell(u) > 0, u > 0$.

(A2). (i) $\ell : [0, \infty) \to [0, \infty)$ is non–decreasing. (ii) $\ell : [0, \infty) \to [0, \infty)$ is strictly increasing. (iii) $\ell \in C([0, \infty))$ and $\ell(u) \to \infty$, as $|u| \to \infty$.

We state remarks on (A1)–(A2).

Remark 2.1. (i) (2.1) and (A1,ii) mean that $(0, \infty) \ni u \mapsto \ell(u)/u$ is non–increasing and is strictly decreasing, respectively. In particular, (A1,ii) implies (A1,iii), provided $\ell(u) \geq 0$.

(ii) If ℓ is concave and $\ell(0) = 0$, then (2.1) holds. If ℓ is strictly convex and $\ell(0) = 0$, then (2.1) does not hold. $\ell(u) = u$ satisfies (A1,i), but not (A1,ii).

(iii)

$$\ell(u) = \begin{cases} 2u\exp(-u), & 0 \leq u < 1, \\ u\exp(-u), & u \geq 1 \end{cases}$$

is concave on $[0, 1)$ and $[1, 2]$ and is convex on $[2, \infty)$. It is strictly increasing and strictly decreasing on $[0, 1)$ and $[1, \infty)$, respectively. It is not continuous at $u = 1$ and satisfies (A1).

(iv) (2.1) and (A2,i) imply that $\ell \in C((0, \infty))$ since, if $0 < h < u$,

$$\frac{\ell(u)}{u+h} \leq \frac{\ell(u+h)}{u+h} \leq \frac{\ell(u)}{u} \leq \frac{\ell(u-h)}{u-h} \leq \frac{\ell(u)}{u-h}.$$

We describe a list of notations of the sets of minimizers.

$\Pi_{T,opt}(P_0, P_1)$:=the set of minimizers of $T(P_0, P_1)$.

$\Pi_{T^V,opt}(P_0, P_1; m)$:=the set of minimizers of $T^V(P_0, P_1; m)$.

$\mathcal{A}_{i,opt}(P_0, P_1)$:=the set of minimizers of $\tilde{V}_i(P_0, P_1), i = 1, 2$.
$\mathcal{A}_{opt}(P_0, P_1; m)$:=the set of minimizers of $V(P_0, P_1; m)$.

We say that $A \subset [0, 1]$ is a random measurable set if and only if there exists a $\{0, 1\}$–valued stochastic process $\{\eta(t, \omega)\}_{0 \le t \le 1}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ such that

$$[0, 1] \times \Omega \ni (t, \omega) \mapsto \eta(t, \omega) \in \{0, 1\}$$

is jointly measurable and $A = A(\omega) = \eta(\cdot, \omega)^{-1}(1)$, i.e. $\eta(t, \omega) = I_{A(\omega)}(t)$, where $I_B(x) = 1, x \in B; = 0, x \notin B$. It is easy to see that the Lebesgue measure $|A(\omega)| = \int_0^1 I_{\{1\}}(\eta(t, \omega))dt$ is a random variable.

For $x, y \in \mathbb{R}^d$, a Lebesgue measurable set $A \subset [0, 1]$, and $t \in [0, 1]$,

$$X(t; x, y, A) := \begin{cases} x + \dfrac{|A \cap [0, t]|}{|A|}(y - x), & \text{if } x \neq y, |A| > 0, \\ x, & \text{otherwise.} \end{cases} \tag{2.2}$$

The following gives the relation between $T(P_0, P_1)$ and $\tilde{V}_1(P_0, P_1)$ (see (1.1) and (1.8) for notation).

Theorem 2.2. *Suppose that (A1,i) holds. Then for any $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$, the following holds.*
(i)

$$T(P_0, P_1) = \tilde{V}_1(P_0, P_1). \tag{2.3}$$

(ii) If $(X_0, X_1) \in \Pi_{T,opt}(P_0, P_1)$ and a random measurable set $A \subset [0, 1]$ are defined on the same probability space and if

$$P(|A| > 0 | X_0 \neq X_1) = 1,$$

then $X(\cdot; X_0, X_1, A) \in \mathcal{A}_{1,opt}(P_0, P_1)$.
(iii) If $X(\cdot) \in \mathcal{A}_{1,opt}(P_0, P_1)$, then $(X(0), X(1)) \in \Pi_{T,opt}(P_0, P_1)$. Suppose, in addition, that (A1,ii) holds. Then $X(\cdot) = X(\cdot; X(0), X(1), (X')^{-1}(\mathbb{R}^d \setminus \{0\}))$, where

$$(X')^{-1}(\mathbb{R}^d \setminus \{0\}) := \{t \in [0, 1] : X'(t) \neq 0\}.$$

Remark 2.3. In the case where L is strictly convex, for an optimal path $X(\cdot)$ of $V(P_0, P_1)$, $X(\cdot) = X(\cdot; X(0), X(1), [0, 1])$ by Jensen's inequality (see, e.g. [15]). In particular, it moves at constant velocity. Theorem 2.2 implies that under (A1), an optimal path $X(\cdot) \in \mathcal{A}_{1,opt}(P_0, P_1)$ can stop even randomly. But when it moves, the velocity is constant in t and can be random.

Under (A1,i,iii), $\tilde{V}_1 \ge \tilde{V}_2$ (see (3.13)). The following implies that equality holds under an additional assumption (A2,i).

Theorem 2.4. *Suppose that (A1,i,iii) and (A2,i) hold. Then for any $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$, the following holds.*
(i)

$$\tilde{V}_1(P_0, P_1) = \tilde{V}_2(P_0, P_1). \tag{2.4}$$

(ii) Suppose, in addition, that (A2,ii) holds. Then

$$\mathcal{A}_{1,opt}(P_0, P_1) = \mathcal{A}_{2,opt}(P_0, P_1).$$

In particular, for any $X(\cdot) \in \mathcal{A}_{1,opt}(P_0, P_1)$,

$$N_1(X') = N_2(X'). \tag{2.5}$$

The following implies that Theorem 2.4 does not necessarily hold without (A2,i) (see Remark 2.1, (iii) for an example and also Theorem 2.2).

Proposition 2.5. *Suppose that there exists* $r_0 > 0$ *such that* ℓ *is strictly decreasing on* $[r_0, \infty)$. *Then for any* $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$ *for which* $T(P_0, P_1)$ *has a minimizer* (X_0, X_1) *such that* $P(|X_1 - X_0| \geq r_0) > 0$, *the following holds:*

$$T(P_0, P_1) > \tilde{V}_2(P_0, P_1). \tag{2.6}$$

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

$$f^{\ell}(x) := \inf\{\ell(|y - x|) + f(y) | y \in \mathbb{R}^d\}, \quad x \in \mathbb{R}^d. \tag{2.7}$$

From (i) in Theorems 2.2–2.4, we easily obtain the following and omit the proof (see the proof of Theorem 2.1 in [14]).

Corollary 2.6. *Suppose that (A2,iii) holds. Suppose also that "(A1,i)" or "(A1,i,iii) and (A2,i)" hold. Then for* $i = 1$ *or* 2, *the following holds, respectively: for any* $P_0 \in \mathcal{P}(\mathbb{R}^d)$ *such that* $P_0(dx) \ll dx$ *and any* $f \in C_b(\mathbb{R}^d)$,

$$\begin{aligned} &\inf\left\{E\left[\int_0^1 L_i(t, X)dt + f(X(1))\right] : X \in \mathcal{A}_\infty(P_0, P^{X(1)})\right\} \\ &= \int_{\mathbb{R}^d} f^{\ell}(x)P_0(dx). \end{aligned} \tag{2.8}$$

Remark 2.7. (2.8) is a finite–time horizon optimal control problem for absolutely continuous stochastic processes (see [7] for stochastic control theory) and the l. h. s. can be also written as follows:

$$\inf\left\{\tilde{V}_i(P_0, P) + \int_{\mathbb{R}^d} f(x)P(dx) : P \in \mathcal{P}(\mathbb{R}^d)\right\}.$$

The following gives the relation between $V(P_0, P_1; B)$ and $T^V(P_0, P_1; B)$.

Theorem 2.8. *Suppose that (A1,i) holds. Then for any* $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$ *and* $B \subset \mathcal{P}([0, \infty))$, *the following holds.*
(i)

$$V(P_0, P_1; B) = T^V(P_0, P_1; B). \tag{2.9}$$

(ii) If $(X_0, X_1, M) \in \Pi_{T^V,opt}(P_0, P_1; B)$ *and a random measurable set* $A \subset [0, 1]$ *are defined on the same probability space and if*

$$P\left(|A| = \frac{|X_1 - X_0|}{M}\middle| M > 0\right) = 1,$$

then $(X(\cdot; X_0, X_1, A), M) \in \mathcal{A}_{opt}(P_0, P_1; B)$.

(iii) If $(X(\cdot), M) \in \mathcal{A}_{opt}(P_0, P_1; B)$*, then* $(X(0), X(1), M) \in \Pi_{TV,opt}(P_0, P_1; B)$*. Suppose, in addition, that (A1,ii) holds. Then*

$$P\left(|(X')^{-1}(\mathbb{R}^d\backslash\{0\})| = \frac{|X(1) - X(0)|}{M}\middle| M > 0\right) = 1, \tag{2.10}$$

and $X(\cdot) = X(\cdot; X(0), X(1), (X')^{-1}(\mathbb{R}^d\backslash\{0\}))$*.*

Remark 2.9. Even if ℓ is lower semicontinuous,

$$AC([0,1]) \ni \varphi \mapsto \int_0^1 \ell(|\varphi'(t)|)dt$$

is not necessarily lower semicontinuous in the supnorm. In particular, it is not trivial if $\mathcal{A}_{opt}(P_0, P_1; B)$ is not empty.

(A1,i) implies that $\ell(u)/u$ is convergent as $u \to \infty$ (see Remark 2.1, (i)):

$$C_\ell := \lim_{u\to\infty} \frac{\ell(u)}{u}.$$

In particular, the following holds from Theorem 2.8, (i).

Corollary 2.10. *Suppose that (A1,i) holds. Then for any* $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$*, the following holds:*

$$\inf\left\{E\left[\int_0^1 \ell(|X'(t)|)dt\right] : X(\cdot) \in \mathcal{A}_\infty(P_0, P_1)\right\} = C_\ell \cdot T_1(P_0, P_1). \tag{2.11}$$

In particular, if $P_0 \neq P_1$ *and the l. h. s. of (2.11) has a minimizer, then*

$$\inf\left\{u > 0 : \frac{\ell(u)}{u} = C_\ell\right\} < \infty. \tag{2.12}$$

$r \mapsto V(P_0, P_1; \{\delta_r\})$ is non–increasing and converges to $V(P_0, P_1; \{\delta_M\}_{M>0})$, as $r \to \infty$, where δ_r denotes the delta measure on $\{r\}$. In particular, we easily obtain the following from Theorem 2.8, (i) and we omit the proof.

Corollary 2.11. *Suppose that (A1,i) holds. Then for any* $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$ *with bounded supports, the following holds: for any* $r \geq \sup\{|x_0 - x_1|; x_i \in supp(P_i), i = 0, 1\}$*,*

$$V(P_0, P_1; \{\delta_r\}) = \frac{\ell(r)}{r} T_1(P_0, P_1). \tag{2.13}$$

In particular,

$$V(P_0, P_1; \{\delta_M\}_{M>0}) = C_\ell \cdot T_1(P_0, P_1) \tag{2.14}$$

and the left–hand sides of (2.11) and (2.14) coincide.

Remark 2.12. For $a \geq 0$, $\ell(u) = au + 1 - \exp(-u)$ is concave, satisfies (A1,i), and $C_\ell = a$.

3. Proofs of results in section 2

In this section, we prove our results. When it is not confusing, we omit "a.s." for the sake of simplicity.

We first prove Theorem 2.2.

Proof. (Theorem 2.2) We first prove (i). We prove

$$T(P_0,P_1) \le \tilde{V}_1(P_0,P_1). \tag{3.1}$$

Suppose that $X(\cdot) \in \mathcal{A}_\infty(P_0,P_1)$. If $X(1) \ne X(0)$, then

$$N_1 \int_0^1 \ell\left(\frac{|X'(t)|}{N_1}\right) dt \ge \ell\left(|X(1)-X(0)|\right), \tag{3.2}$$

where $N_1 = N_1(X')$ (see (1.5) for notation). Indeed, from (A1,i),

$$\begin{aligned} \ell\left(\frac{|X'(t)|}{N_1}\right) &= \ell\left(\frac{|X'(t)|}{N_1|X(1)-X(0)|}|X(1)-X(0)|\right) \\ &\ge \frac{|X'(t)|}{N_1|X(1)-X(0)|}\ell\left(|X(1)-X(0)|\right), \quad dt-\text{a.e.} \end{aligned} \tag{3.3}$$

since

$$\frac{|X'(t)|}{N_1|X(1)-X(0)|} = \frac{|X'(t)|}{||X'||_\infty} \le 1, \quad dt-\text{a.e.}.$$

Besides,

$$||X'||_1 \ge |X(1)-X(0)|. \tag{3.4}$$

If $X(1)=X(0)$, then (3.2) holds trivially. (3.2) implies (3.1) immediately.

We prove

$$\tilde{V}_1(P_0,P_1) \le T(P_0,P_1). \tag{3.5}$$

Suppose that $(X_0,X_1) \in \Pi(P_0,P_1)$.

$$X(t) := X_0 + t(X_1-X_0), \quad 0 \le t \le 1. \tag{3.6}$$

Then $X(\cdot) \in \mathcal{A}_\infty(P_0,P_1), N_1 = N_1(X') = 1$, and

$$\ell(|X_1-X_0|) = N_1 \int_0^1 \ell\left(\frac{|X'(t)|}{N_1}\right) dt, \tag{3.7}$$

which implies (3.5).

We prove (ii). We write $X(\cdot;X_0,X_1,A) = X(\cdot;A)$ for simplicity. $X(\cdot;A) \in \mathcal{A}_\infty(P_0,P_1)$ and (3.7) with $X(\cdot)=X(\cdot;A)$ holds since

$$|X'(t;A)| = I_A(t)\frac{|X_1-X_0|}{|A|}, \quad dt-\text{a.e.}, \quad N_1(X'(\cdot;A)) = \frac{1}{|A|},$$

provided $X_1 \ne X_0, |A|>0$. (2.3) and (3.7) with $X(\cdot)=X(\cdot;A)$ imply (ii).

(2.3) and (3.2) imply the first part of (iii). We prove that $X(\cdot) = X(\cdot;X(0),X(1),(X')^{-1}(\mathbb{R}^d\setminus\{0\}))$. For $X(\cdot) \in \mathcal{A}_{1,opt}(P_0,P_1)$, if $X(1) \ne X(0)$, then the equality holds in (3.2)–(3.4). In particular, the following holds:

$$\begin{aligned} ||X'||_1 &= |X(1)-X(0)|, \quad \text{a.s.}, \\ |X'(t)| &= 0, N_1|X(1)-X(0)|, \quad dtdP-\text{a.e.} \end{aligned} \tag{3.8}$$

from (A1,ii), where $N_1 := N_1(X')$. Notice that the equalities in (3.8) hold if $X(1) = X(0)$ (see an explanation above (1.10) and Remark 2.1, (i)). The following completes the proof:

$$X'(t) = \frac{X(1) - X(0)}{|(X')^{-1}(\mathbb{R}^d \backslash \{0\})|}, \quad \text{on } (X')^{-1}(\mathbb{R}^d \backslash \{0\}), \quad dtdP\text{−a.e.}. \tag{3.9}$$

We prove (3.9). (3.8) implies that for $P-$ almost all ω, there exists $Z = Z(\omega)$ such that $|Z| = 1$ and

$$X'(t) = N_1 |X(1) - X(0)| Z, \quad \text{on } (X')^{-1}(\mathbb{R}^d \backslash \{0\}), \quad dt\text{−a.e.}. \tag{3.10}$$

In particular,

$$X(1) - X(0) = N_1 |X(1) - X(0)| Z \times |(X')^{-1}(\mathbb{R}^d \backslash \{0\})|. \tag{3.11}$$

(3.10)–(3.11) imply (3.9). □

We prove Theorem 2.4.

Proof. (Theorem 2.4) We first prove (i). We prove

$$\tilde{V}_2(P_0, P_1) \le \tilde{V}_1(P_0, P_1). \tag{3.12}$$

For $X(\cdot) \in \mathcal{A}_\infty(P_0, P_1)$ such that $||X'||_\infty = 0$ if $X(1) = X(0)$ and hence $N_2 \le N_1$, from (A1,i),

$$N_1 \ell\left(\frac{|X'(t)|}{N_1}\right) = N_1 \ell\left(\frac{|X'(t)|}{N_2}\frac{N_2}{N_1}\right) \ge N_2 \ell\left(\frac{|X'(t)|}{N_2}\right). \tag{3.13}$$

(A1,iii) implies (3.12) (see (1.10)).

We prove

$$T(P_0, P_1) \le \tilde{V}_2(P_0, P_1), \tag{3.14}$$

which completes the proof of (i) from Theorem 2.2. The following implies (3.14): for $X(\cdot) \in \mathcal{A}_\infty(P_0, P_1)$, from (A2,i),

$$N_2 \int_0^1 \ell\left(\frac{|X'(t)|}{N_2}\right) dt \ge \ell\left(|X(1) - X(0)|\right), \tag{3.15}$$

in the same way as (3.3), where $N_2 = N_2(X')$. Indeed, if $||X'||_1 > 0$, then

$$\ell\left(\frac{|X'(t)|}{N_2}\right) = \ell\left(\frac{|X'(t)|}{N_2 ||X'||_1}||X'||_1\right) \ge \frac{|X'(t)|}{N_2 ||X'||_1} \ell\left(||X'||_1\right). \tag{3.16}$$

We prove (ii). For $X(\cdot) \in \mathcal{A}_{1,opt}(P_0, P_1)$, the equality holds in (3.13) from (2.4), which implies that $X(\cdot) \in \mathcal{A}_{2,opt}(P_0, P_1)$.

For $X(\cdot) \in \mathcal{A}_{2,opt}(P_0, P_1)$, the equalities hold in (3.15)–(3.16), since from (2.3) and (2.4), $\tilde{V}_2(P_0, P_1) = T(P_0, P_1)$. This implies that $N_2(X') = N_1(X')$ from (A2,ii). In particular, $X(\cdot) \in \mathcal{A}_{1,opt}(P_0, P_1)$ from (2.4). □

We prove Proposition 2.5.

Proof. (Proposition 2.5) If $|X_1 - X_0| \geq r_0$, then take a random variable Y such that the following holds:

$$C := |Y - X_1| = |Y - X_0| = 1 + \frac{|X_1 - X_0|}{2}.$$

$$Y(t) := \begin{cases} X_0 + 2t(Y - X_0), & 0 \leq t \leq \frac{1}{2}, \\ Y + (2t-1)(X_1 - Y), & \frac{1}{2} \leq t \leq 1. \end{cases} \tag{3.17}$$

Then $Y(t) = X_t, t = 0, 1$ and the following holds under our assumption:

$$N_2(Y') \int_0^1 \ell\left(\frac{|Y'(t)|}{N_2(Y')}\right) dt = \ell(2C) < \ell(|X_1 - X_0|), \tag{3.18}$$

since $|Y'(t)| = ||Y'||_\infty = ||Y'||_1 = 2C$.

If $|X_1 - X_0| < r_0$, then

$$Y(t) := X_0 + t(X_1 - X_0), \quad 0 \leq t \leq 1. \tag{3.19}$$

Then $Y(t) = X_t, t = 0, 1$ and the following holds:

$$N_2(Y') \int_0^1 \ell\left(\frac{|Y'(t)|}{N_2(Y')}\right) dt = \ell(|X_1 - X_0|). \tag{3.20}$$

since $|Y'(t)| = ||Y'||_\infty = ||Y'||_1 = |X_1 - X_0|$.

From (3.18) and (3.20), under our assumption, the following holds:

$$\begin{aligned} \tilde{V}_2(P_0, P_1) &\leq E\left[N_2(Y') \int_0^1 \ell\left(\frac{|Y'(t)|}{N_2(Y')}\right) dt\right] \\ &< E[\ell(|X_1 - X_0|)] = T(P_0, P_1). \end{aligned} \tag{3.21}$$

□

We prove Theorem 2.8.

Proof. (Theorem 2.8) We first prove (i). We prove

$$T^V(P_0, P_1; B) \leq V(P_0, P_1; B). \tag{3.22}$$

If $(X(\cdot), M) \in \mathcal{A}_\infty(P_0, P_1; B)$, then $P^M \in B$ and

$$|X(1) - X(0)| \leq \int_0^1 |X'(t)| dt \leq M, \tag{3.23}$$

$$\begin{aligned} \int_0^1 \ell\left(\frac{|X'(t)|}{M} M\right) dt &\geq \int_0^1 \frac{|X'(t)|}{M} \ell(M) dt \\ &\geq \frac{\ell(M)}{M} |X(1) - X(0)|, \end{aligned} \tag{3.24}$$

from (A1,i), provided $M > 0$, which implies (3.22).

We prove

$$V(P_0, P_1; B) \leq T^V(P_0, P_1; B). \tag{3.25}$$

If $(X_0, X_1, M) \in \Pi_\infty(P_0, P_1; B)$, then $(X_M(\cdot) := X(\cdot; X_0, X_1, A_M), M) \in \mathcal{A}_\infty(P_0, P_1; B)$, where $A_M := [0, |X_1 - X_0|/M]$ if $M > 0$ and $= \{0\}$ if $M = 0$. In the case where $M \neq 0$,

$$|X'_M(t)| = \begin{cases} M, & 0 < t < \dfrac{|X_1 - X_0|}{M}, \\ 0, & \dfrac{|X_1 - X_0|}{M} < t < 1, \end{cases}$$

$$\int_0^1 \ell(|X'_M(t)|)dt = \int_0^{\frac{|X_1 - X_0|}{M}} \ell(M)dt = \frac{\ell(M)}{M}|X_1 - X_0|, \tag{3.26}$$

which implies (3.25). (3.22) and (3.25) imply (2.9).

We prove (ii). Since $(X_0, X_1, M) \in \Pi_{T^V,opt}(P_0, P_1; B)$, the following holds: from (2.9),

$$\begin{aligned} V(P_0, P_1; B) &\leq E\left[\int_0^1 \ell(|X'(t; X_0, X_1, A)|)dt\right] \\ &= E\left[\frac{\ell(M)}{M}|X_1 - X_0|; M > 0\right] \\ &= T^V(P_0, P_1; B) = V(P_0, P_1; B). \end{aligned} \tag{3.27}$$

Indeed, if $M \geq |X_1 - X_0| > 0$, then $|A| > 0$ and

$$|X'(t; X_0, X_1, A)| = \frac{|X_1 - X_0|}{|A|} = M \quad \text{on } A, \quad dt\text{–a.e..}$$

We prove the first part of (iii). Since $(X(\cdot), M) \in \mathcal{A}_{opt}(P_0, P_1; B)$, the following holds: from (2.9) and (3.23)–(3.24),

$$\begin{aligned} T^V(P_0, P_1; B) &= V(P_0, P_1; B) \\ &= E\left[\int_0^1 \ell(|X'(t)|)dt\right] \\ &\geq E\left[\frac{\ell(M)}{M}|X(1) - X(0)|; M > 0\right] \\ &\geq T^V(P_0, P_1; B). \end{aligned} \tag{3.28}$$

We prove the second part of (iii). For $(X(\cdot), M) \in \mathcal{A}_{opt}(P_0, P_1; B)$, from (A1,ii),

$$|X'(t)| = 0 \text{ or } M, \quad dtdP\text{–a.e.,} \tag{3.29}$$

$$|X(1) - X(0)| = ||X'||_1, \quad \text{a.s.,} \tag{3.30}$$

since the equality holds in (3.24) from (2.9). The following can be proved in the same way as (3.9):

$$X'(t) = \frac{X(1) - X(0)}{|(X')^{-1}(\mathbb{R}^d \backslash \{0\})|}, \quad \text{on } (X')^{-1}(\mathbb{R}^d \backslash \{0\}), \quad dtdP\text{–a.e..} \tag{3.31}$$

Indeed, replace $N_1|X(1) - X(0)|$ by M in (3.8). (3.11) also implies the following:

$$|X(1) - X(0)| = M \times |(X')^{-1}(\mathbb{R}^d \backslash \{0\})|, \quad \text{a.s.,} \tag{3.32}$$

which completes the proof. □

We prove Corollary 2.10.

Proof. (Corollary 2.10) From Theorem 2.8, (2.11) can be obtained by the following:

$$T^V(P_0, P_1; \mathcal{P}([0,\infty))) = C_\ell \cdot T_1(P_0, P_1), \tag{3.33}$$

$$V(P_0, P_1; \mathcal{P}([0,\infty))) = \inf \left\{ E\left[\int_0^1 \ell(|X'(t)|)dt\right] : X(\cdot) \in \mathcal{A}_\infty(P_0, P_1) \right\}. \tag{3.34}$$

(3.33) can be proved by the following. If $(X_0, X_1, M) \in \Pi_\infty(P_0, P_1; \mathcal{P}([0,\infty)))$, then $(X_0, X_1) \in \Pi(P_0, P_1)$ and

$$E[\ell(M)M^{-1}|X_1 - X_0|; M > 0] \geq C_\ell E[|X_1 - X_0|; M > 0] = C_\ell E[|X_1 - X_0|].$$

If $(X_0, X_1) \in \Pi(P_0, P_1)$, then for $R > 0$, $(X_0, X_1, \max(|X_1 - X_0|, R)) \in \Pi_\infty(P_0, P_1; \mathcal{P}([0,\infty)))$, and by the dominated convergence theorem,

$$E\left[\frac{\ell(\max(|X_1 - X_0|, R))}{\max(|X_1 - X_0|, R)}|X_1 - X_0|\right] \to C_\ell E[|X_1 - X_0|], \quad R \to \infty.$$

(3.34) can be proved by the following. If $(X(\cdot), M) \in \mathcal{A}_\infty(P_0, P_1; \mathcal{P}([0,\infty)))$, then $X(\cdot) \in \mathcal{A}_\infty(P_0, P_1)$. If $X(\cdot) \in \mathcal{A}_\infty(P_0, P_1)$, then $(X(\cdot), ||X'||_\infty) \in \mathcal{A}_\infty(P_0, P_1; \mathcal{P}([0,\infty)))$.

For $X(\cdot) \in \mathcal{A}_\infty(P_0, P_1)$, from (A1,i),

$$\int_0^1 \ell(|X'(t)|)dt \geq C_\ell \int_0^1 |X'(t)|dt, \tag{3.35}$$

where the equality holds if and only if

$$\ell(|X'(t)|) = C_\ell |X'(t)|, \quad dtdP - a.e..$$

If $P_0 \neq P_1$, then $P(||X'||_\infty = 0) < 1$, which implies (2.12). □

4. Appendix

In this section, we state the proofs of (1.4), (1.9), and (1.12).

We prove (1.4). For $(X_0, X_1) \in \Pi(P_0, P_1)$ such that $E[L(X_1 - X_0)]$ is finite,

$$Y_n(t) := \begin{cases} X_0 + n(X_1 - X_0)t, & 0 \leq t \leq \dfrac{1}{n}, \\ X_1, & \dfrac{1}{n} \leq t \leq 1. \end{cases}$$

Then $Y_n \in \mathcal{A}(P_0, P_1)$, and

$$\begin{aligned} 0 \leq V(P_0, P_1) \leq E\left[\int_0^1 L(Y_n'(t))dt\right] &= E[n^{-1}L(n(X_1 - X_0))] \\ &\to 0, \quad n \to \infty \end{aligned}$$

by Lebesgue's dominated convergence theorem since

$$n^{-1}L(n(X_1 - X_0)) \leq L(X_1 - X_0).$$

We prove (1.9). For $(X_0, X_1) \in \Pi_{T,opt}(P_0, P_1)$, $X(\cdot)$ defined by (3.6) is a minimizer of $V(P_0, P_1)$ by Jensen's inequality and $N_1(X') = 1$, provided $V(P_0, P_1)$ is finite. The following implies that (1.9) holds: for $N \geq 1$,

$$L(x) = L\left(\frac{1}{N}Nx + (1 - \frac{1}{N})0\right) \leq \frac{1}{N}L(Nx) + (1 - \frac{1}{N})L(0) = \frac{1}{N}L(Nx)$$

since $L : \mathbb{R}^d \to [0, \infty)$ is convex and $L(0) = 0$.

We prove (1.12). For $T \geq 1, \varphi \in AC([0,T])$ such that $N_i(\varphi')_T = T$,

$$\varphi(T\cdot) \in AC([0,1]), \quad N_i(\varphi'(T\cdot)) = N_i(\varphi')_T = T,$$

$$\begin{aligned}\int_0^T \ell\left(|\varphi'(t)|\right) dt &= \int_0^1 \ell\left(|\varphi'(Ts)|\right) T ds \\ &= \int_0^1 N_i(\varphi'(T\cdot))\ell\left(\frac{1}{N_i(\varphi'(T\cdot))}\left|\frac{d}{ds}\varphi(Ts)\right|\right) ds,\end{aligned}$$

which implies that (l. h. s.) $\leq$ (r. h. s.) in (1.12).

For $\varphi \in AC([0,1]), T \geq 1$,

$$\varphi\left(\frac{\cdot}{T}\right) \in AC([0,T]), \quad N_i\left(\varphi'\left(\frac{\cdot}{T}\right)\right)_T = N_i(\varphi'),$$

$$\begin{aligned}\int_0^1 N_i(\varphi')\ell\left(\frac{|\varphi'(t)|}{N_i(\varphi')}\right) dt &= \int_0^T N_i(\varphi')\ell\left(\frac{1}{N_i(\varphi')}\left|\varphi'\left(\frac{s}{T}\right)\right|\right)\frac{1}{T} ds \\ &= \int_0^T \frac{N_i\left(\varphi'\left(\frac{\cdot}{T}\right)\right)_T}{T}\ell\left(\frac{T}{N_i\left(\varphi'\left(\frac{\cdot}{T}\right)\right)_T}\left|\frac{d}{ds}\varphi\left(\frac{s}{T}\right)\right|\right) ds \\ &= \int_0^T \ell\left(\left|\frac{d}{ds}\varphi\left(\frac{s}{T}\right)\right|\right) ds,\end{aligned}$$

provided $T = N_i(\varphi')$. This implies that (l. h. s.) $\geq$ (r. h. s.) in (1.12).

References

[1] S. Bernstein, *Sur les liaisons entre les grandeurs alétoires*, Verh. des intern. Mathematikerkongr. Zurich 1932 Band **1** (1932), 288–309.

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

[3] Y. Brenier, *Polar factorization and monotone rearrangement of vector–valued functions*, Comm. Pure Appl. Math. **44** (1991), 375–417.

[4] Y. Chen, T. T. Georgiou and M. Pavon, *Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge*, SIAM Review **63** (2021), 249–313.

[5] M. Cuturi, *Sinkhorn distances: Lightspeed computation of optimal transportation*, in: Neural Information Processing Systems 2013, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K. Q. Weinberger (eds.) , Advances in Neural Information Processing Systems, vol. 26, pp. 2292–2300.

[6] G. Dall'Aglio, *Sugli estremi dei momenti delle funzioni di ripartizione doppia*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **10**, (1956), 35–74.

[7] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions, 2nd ed.*, Springer, New York, 2006.

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

[9] B. Jamison, *The Markov process of Schrödinger*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **32** (1975), 323–331.

[10] C. Léonard, *From the Schrödinger problem to the Monge–Kantorovich problem*, J. Funct. Anal. **262** (2012) , 1879–1920.

[11] C. Léonard, *A survey of the Schrödinger problem and some of its connections with optimal transport*, Special Issue on Optimal Transport and Applications. Discrete Contin. Dyn. Syst. **34** (2014), 1533–1574.

[12] T. Mikami, *Optimal control for absolutely continuous stochastic processes and the mass transportation problem*, Elect. Comm. in Probab. **7** (2002), 199–213.
[13] T. Mikami, *Monge's problem with a quadratic cost by the zero–noise limit of h–path processes*, Probab. Theory Related Fields **129** (2004), 245–260.
[14] T. Mikami, *A simple proof of duality theorem for Monge–Kantorovich problem*, Kodai Math. J. **29**, (2006), 1–4.
[15] T. Mikami, *Stochastic Optimal Transportation: Stochastic Control with Fixed Marginals*, Springer Briefs in Mathematics, Springer, Singapore, 2021.
[16] T. Mikami, *Stochastic optimal transport revisited*, SN Partial Differ. Equ. Appl. **2** (2021): article 5, 26 pages.
[17] G. Peyré and M. Cuturi, *Computational Optimal Transport: With Applications to Data Science*, Now Publishers, Boston, 2019.
[18] S. T. Rachev and L. Rüschendorf, *Mass transportation problems, Vol. I: Theory, Vol. II: Application*, Springer, Heidelberg, 1998.
[19] E. Schrödinger, *Ueber die Umkehrung der Naturgesetze*, Sitz. Ber. der Preuss. Akad. Wissen. Berlin, Phys. Math. (1931), 144–153.
[20] E. Schrödinger, *Théorie relativiste de l'electron et l'interprétation de la mécanique quantique*, Ann. Inst. H. Poincaré **2** (1932), 269–310.
[21] C. Villani, *Optimal Transport: Old and New*, Springer, Heidelberg, 2008.
[22] H. Yamamoto, *Continuity of a one–dimensional optimal transport with respect to the cost function*, in Japanese, Master Thesis, Tsuda University, Tokyo, 2022.
[23] J. C. Zambrini, *Variational processes*, in: Stochastic Processes in Classical and Quantum Systems, Ascona 1985, Albeverio, S. et al. (eds.) Lecture Notes in Phys., vol. 262, pp. 517–529, Springer, Heidelberg, 1986.

Manuscript received February 1 2023
revised May 7 2023

T. Mikami
Department of Mathematics, Tsuda University, Kodaira 187-8577, Japan
E-mail address: t.mikami@tsuda.ac.jp

H. Yamamoto
c/o Prof. T. Mikami, Department of Mathematics, Tsuda University, Kodaira 187-8577, Japan