

DYNAMIC CHARACTERIZATION OF BARYCENTRIC OPTIMAL TRANSPORT PROBLEMS AND THEIR MARTINGALE RELAXATION

IVAN GUO, SEVERIN NILSSON, AND JOHANNES WIESEL

ABSTRACT. We extend the Benamou-Brenier formula from classical optimal transport to weak optimal transport and show that the barycentric optimal transport problem studied by Gozlan and Juillet has a dynamic analogue. We also investigate a martingale relaxation of this problem, and relate it to the martingale Benamou-Brenier formula of Backhoff-Veraguas, Beiglböck, Huesmann and Källblad.

1. INTRODUCTION AND MAIN RESULTS

Let μ and ν be two probability measures on \mathbb{R}^d with finite second moments. The optimal transport problem with quadratic cost is given by

$$(OT) \quad \mathcal{T}_2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int |x - y|^2 \pi(dx, dy),$$

where $\Pi(\mu, \nu)$ denotes the set of couplings between μ and ν , i.e.,

$$\pi \in \Pi(\mu, \nu) \iff \pi(A \times \mathbb{R}^d) = \mu(A) \text{ and } \pi(\mathbb{R}^d \times A) = \nu(A) \quad \forall A \subseteq \mathbb{R}^d \text{ Borel};$$

see [Vil21, San15] for an overview. In the seminal work [BB00] it is shown that solving $\mathcal{T}_2(\mu, \nu)$ is equivalent to minimizing the total energy along absolutely continuous curves $(\mu_t)_{t \in [0,1]}$ from μ to ν ; to be precise,

$$(1) \quad \mathcal{T}_2(\mu, \nu) = \inf_{(\mu_t, v_t)} \int_0^1 \int_{\mathbb{R}^d} |v_t|^2 d\mu_t dt,$$

where the infimum is taken over all (μ_t, v_t) such that $\mu_0 = \mu, \mu_1 = \nu$, and (μ_t, v_t) solves

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0$$

in the sense of distributions. Problem (1) is known as the dynamic formulation of optimal transport, or the Benamou-Brenier formula. It has the probabilistic representation

$$(DOT) \quad \mathcal{T}_2(\mu, \nu) = \inf \left\{ \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right] : dX_t = v_t dt \text{ where } X_0 \sim \mu, X_1 \sim \nu \right\}.$$

In this note we extend the Benamou-Brenier formula to the so-called barycentric weak optimal transport problem. Introduced in the series of papers [GRST17, GRS⁺18], this problem is defined as

$$(WOT) \quad \bar{\mathcal{T}}_2(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int |\operatorname{mean}(\pi_x) - x|^2 \mu(dx),$$

Date: November 27, 2025.

where the map $(\pi_x)_{x \in \mathbb{R}^d}$ is the disintegration of π with respect to μ and $\text{mean}(\rho) := \int y \rho(dy)$ for any integrable probability measure ρ . Weak optimal transport covers the settings of martingale optimal transport [BHL13, BJ16], entropic optimal transport [Con19, Nut21] and semi-martingale optimal transport [TT14, GL21, BCH⁺24], among others; see also the related works [Mar96a, Mar96b, Tal95, Tal96, FS18, ABC19, BG18, FS18, Shu20]. It has recently proved to be an extremely versatile tool in OT. Intuitively, $\bar{\mathcal{T}}_2(\mu, \nu)$ measures how far μ and ν are away from being the marginals of a one-step martingale. [GJ20] show that

$$\bar{\mathcal{T}}_2(\mu, \nu) = \inf_{\eta \preceq_c \nu} \mathcal{T}_2(\mu, \eta),$$

where \preceq_c denotes convex order, i.e. $\eta \preceq_c \nu$ if $\int f d\eta \leq \int f d\nu$ for all convex functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$. Our first main result is the following dynamic characterization of $\bar{\mathcal{T}}_2$:

Theorem 1. *We have*

$$\bar{\mathcal{T}}_2(\mu, \nu) = \inf \left\{ \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right] : dX_t = v_t dt + \sigma_t dB_t, X_0 \sim \mu, X_1 \sim \nu \right\},$$

where the infimum is taken over predictable processes v and σ .

Compared to (DOT), the dynamic formulation in Theorem 1 allows for a costless martingale transport via the diffusion term $\sigma_t dB_t$; on the flip side $\bar{\mathcal{T}}_2(\mu, \nu)$ penalizes only the deviation of $x \mapsto \text{mean}(\pi_x)$ from the identity.

We note that the dynamic formulation in Theorem 1 is different from the entropic projection problem, also known as the Schrödinger bridge,

$$\inf \left\{ \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right] : dX_t = v_t dt + dB_t \text{ where } X_0 \sim \mu, X_1 \sim \nu \right\},$$

see [Sch32, Föl06], where the infimum is taken over the drift v only and σ is identically equal to the identity matrix. The Schrödinger bridge minimizes the Kullback-Leibler divergence of the law of X with respect to the Wiener measure, rather than a cost function on the marginals.

As mentioned above, $\bar{\mathcal{T}}_2(\mu, \nu)$ essentially allows for arbitrary martingale transports, as σ does not influence the cost $\mathbb{E}[\int_0^1 |v_t|^2 dt]$. It is thus natural to extend our analysis to the functional

$$\bar{\mathcal{T}}^{\alpha, \beta}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int \alpha |\text{mean}(\pi_x) - x|^2 - \beta \text{MCov}(\pi_x, \gamma_1^d) \mu(dx)$$

for $\alpha, \beta > 0$, see [BPRS25, Section 1.1.6]. In the above, the maximal covariance

$$\text{MCov}(\rho, \varrho) := \sup_{\pi \in \Pi(\rho, \varrho)} \int \langle y, z \rangle \pi(dy, dz), \quad \rho, \varrho \in \mathcal{P}_2(\mathbb{R}^d),$$

measures the 2-Wasserstein distance of the disintegration π_x from the d -dimensional standard normal distribution γ_1^d , up to terms that do not depend on the coupling π .

One of the main results of [BVBHK19] is the representation

$$(2) \quad \begin{aligned} & \sup_{\pi \in \Pi_M(\mu, \nu)} \int \text{MCov}(\pi_x, \gamma_1^d) \mu(dx) \\ &= \sup \left\{ \mathbb{E} \left[\int_0^1 \text{Tr}(\sigma_t) dt \right] : dX_t = \sigma_t dB_t, X_0 \sim \mu, X_1 \sim \nu \right\}, \end{aligned}$$

where

$$(3) \quad \Pi_M(\mu, \nu) = \{\pi \in \Pi(\mu, \nu) : \text{mean}(\pi_x) = x \quad \forall x \in \mathbb{R}^d\}$$

is the set of martingale measures with marginals μ and ν and we recall that $\Pi_M(\mu, \nu) \neq \emptyset$ if and only if $\mu \preceq_c \nu$; see [Str65]. The solution of (2) is given by a so-called stretched Brownian motion. Equation (2) corresponds to $\overline{\mathcal{T}}^{0,1}$ in our notation above. Our second main result gives a similar representation of $\overline{\mathcal{T}}^{\alpha,\beta}$ for the intermediate case $\alpha, \beta > 0$.

Theorem 2. *For $\alpha, \beta > 0$ and $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ we have*

$$\overline{\mathcal{T}}^{\alpha,\beta}(\mu, \nu) = \inf \left\{ \mathbb{E} \left[\int_0^1 \alpha |v_t|^2 - \beta (\langle B_t, v_t \rangle + \text{Tr}(\sigma_t)) dt \right] : dX_t = v_t dt + \sigma_t dB_t, X_0 \sim \mu, X_1 \sim \nu \right\},$$

where the infimum is taken over all predictable processes v and σ . The right hand side is attained by the process

$$dX_t = (\nabla \varphi(X_0) - X_0) dt + \sigma_t dB_t \quad \text{with} \quad X_0 \sim \mu,$$

where the 1-Lipschitz map $\nabla \varphi$ is given in Proposition 4 and σ is given in Proposition 5 below.

Note that Theorem 1 can be formally obtained from Theorem 2 by taking $\alpha = 1, \beta \rightarrow 0$; similarly (2) can be obtained by setting $\alpha \rightarrow \infty, \beta = 1$. Let us also remark that one can actually restrict the minimization in Theorem 2 to drifts v that are independent of B , leading to $\mathbb{E}[\langle B_t, v_t \rangle] = 0$. This follows from the proof of Theorem 2 below. The dynamic formulation in Theorem 2 can also be seen as a version of the semimartingale optimal transport problem.

2. NOTATION

We write $\mathcal{P}_2(\mathbb{R}^d)$ for the set of (Borel) probability measures with finite second moments. We let $\langle \cdot, \cdot \rangle$ denote the standard inner product on \mathbb{R}^d and for $x \in \mathbb{R}^d$ we write $|x|^2 = \langle x, x \rangle$. For a probability measure μ on \mathbb{R}^d and a function $\kappa : \mathbb{R}^d \rightarrow \mathcal{P}(\mathbb{R}^d)$ we define $(\mu \otimes \kappa)(A \times B) := \int_A \kappa_x(B) \mu(dx)$ for all Borel sets $A, B \subseteq \mathbb{R}^d$. Next, we write $(\pi_x)_{x \in \mathbb{R}^d}$ for the disintegration of $\pi \in \Pi(\mu, \nu)$ wrt. μ , i.e. $x \mapsto \pi_x(A)$ is Borel measurable for all Borel sets $A \subseteq \mathbb{R}^d$ and satisfies $\mu \otimes \pi_x = \pi$. Lastly we define the push-forward measure of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}^k$ under μ as $f_\# \mu(A) := \mu(\{x \in \mathbb{R}^d : f(x) \in A\})$ for all Borel sets $A \subseteq \mathbb{R}^k$, $k \in \mathbb{N}$.

We say that a process X is an admissible diffusion process if there exists a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,1]}, \mathbb{P})$ which supports a standard Brownian motion $(B_t)_{t \in [0,1]}$ with $X_0 \perp (B_t)_{t \in [0,1]}$ and predictable processes $v \in L^2(\mathbb{P} \otimes dt; \mathbb{R}^d)$ and $\sigma \in L^2(\mathbb{P} \otimes dt; \mathbb{R}^{d \times d})$ such that

$$dX_t = v_t dt + \sigma_t dB_t.$$

For $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, we denote by $\mathcal{D}(\mu, \nu)$ the set of all admissible diffusion processes X with $X_0 \sim \mu$ and $X_1 \sim \nu$. We set $\gamma_t^d := \text{Law}(B_t)$. We also define

$$\mathcal{BB}^{\alpha,\beta}(\mu, \nu) := \inf_{X \in \mathcal{D}(\mu, \nu)} \mathbb{E} \left[\int_0^1 \alpha |v_t|^2 - \beta (\langle B_t, v_t \rangle + \text{Tr}(\sigma_t)) dt \right].$$

Using this more compact notation, Theorem 1 reads $\overline{\mathcal{T}}_2 = \mathcal{BB}^{1,0}$, while Theorem 2 reads $\overline{\mathcal{T}}^{\alpha,\beta} = \mathcal{BB}^{\alpha,\beta}$ for $\alpha, \beta > 0$.

3. PRELIMINARY RESULTS

Before we turn to the proofs of Theorems 1 and 2, we need to investigate the relation between two results, which were mentioned in the introduction.

Proposition 3 ([BVBHK19, Theorem 2.2.]). *Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mu \preceq_c \nu$. Then (2) holds and the problem*

$$\sup \left\{ \mathbb{E} \left[\int_0^1 \text{Tr}(\sigma_t) dt \right] : dX_t = \sigma_t dB_t, X_0 \sim \mu, X_1 \sim \nu \right\}.$$

admits a unique (in law) maximizer \widehat{M} .

The authors call the maximizer \widehat{M} a *stretched Brownian motion*; \widehat{M} is the martingale M whose trajectories are as close as possible to Brownian motion in the adapted Wasserstein distance, while satisfying the marginal conditions $M_0 \sim \mu$ and $M_1 \sim \nu$ (see [BVBHK19, Section 6]).

In the follow-up paper [BVBST25] it is shown that under an irreducibility condition¹ on μ and ν , \widehat{M} is a *Bass martingale* between μ and ν . Bass martingales, which go back to [Bas83] as a solution to the Skorokhod embedding problem, are martingales M of the form

$$M_t = \mathbb{E}[\nabla\phi(W_1)|W_t],$$

where the Brownian motion W is started at some $W_0 \sim \alpha$, $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ is a convex function and $\nabla\phi(W_1)$ is square integrable. Bass' construction can be viewed as a natural analogue of Brenier's Theorem [Bre91], which states that for regular enough measures μ and ν , the minimizing vector field v_t appearing in the dynamic formulation on $\mathcal{T}_2(\mu, \nu)$ is of the form $v_t = \nabla\phi - \text{Id}$ for some convex function ϕ .

Next we recall the following result of [GJ20], which was later refined in [BPRS25] and [BVBST25].

Proposition 4 ([GJ20, Theorem 1.2]). *There exists a unique $\bar{\mu} \preceq_c \nu$ such*

$$\overline{\mathcal{T}}_2(\mu, \nu) = \mathcal{T}_2(\mu, \bar{\mu}) = \inf_{\eta \preceq_c \nu} \mathcal{T}_2(\mu, \eta).$$

In particular, $\bar{\mu}$ is given by

$$\bar{\mu} = \nabla\varphi\#\mu$$

where $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ is a convex $C^1(\mathbb{R}^d)$ -function and $\nabla\varphi$ is 1-Lipschitz. Furthermore, the optimizers of $\overline{\mathcal{T}}_2(\mu, \nu)$ and $\mathcal{T}_2(\mu, \bar{\mu})$ are connected via the relation

$$\begin{aligned} \pi \in \Pi(\mu, \nu) \text{ is optimal for } \overline{\mathcal{T}}_2(\mu, \nu) \\ \iff \pi_x = \kappa_{\nabla\varphi(x)} \mu\text{-a.e for some } \kappa \in \Pi_M(\nabla\varphi\#\mu, \nu), \end{aligned}$$

where Π_M was defined in (3).

We can now make a connection between Propositions 3 and 4: indeed, an admissible choice in Proposition 4 is $\kappa = \text{Law}(\widehat{M}_0, \widehat{M}_1)$ where \widehat{M} is a stretched Brownian motion between $\nabla\varphi\#\mu$ and ν from Proposition 3. In fact, the following holds:

¹Two measures μ and ν are irreducible if for any martingale M with $M_0 \sim \mu$ and $M_1 \sim \nu$ we have the implication $\mu(A), \nu(B) > 0 \implies \mathbb{P}(M_0 \in A, M_1 \in B) > 0$ for any $A, B \subseteq \mathbb{R}^d$ Borel.

Proposition 5 ([BPRS25, Theorem 5.4]). *Let $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ be as in Proposition 4 and let $\kappa = \text{Law}(\widehat{M}_0, \widehat{M}_1)$, where \widehat{M} is a stretched Brownian motion between $\nabla\varphi_{\#}\mu$ and ν . Then the coupling $\pi = \mu \otimes \kappa_{\nabla\varphi(x)} \in \Pi(\mu, \nu)$ is optimal for $\overline{\mathcal{T}}^{\alpha, \beta}(\mu, \nu)$, for all $\alpha, \beta > 0$.*

4. PROOFS

We start with the following lemma.

Lemma 6. *We have*

$$\mathcal{BB}^{1,0}(\mu, \nu) = \inf_{\eta \preceq_c \nu} \mathcal{T}_2(\mu, \eta).$$

Proof. We begin by proving the inequality $\mathcal{T}_2(\mu, \eta) \geq \mathcal{BB}^{1,0}(\mu, \nu)$ for any $\eta \preceq_c \nu$. Take any vector field $v \in L^2(\mathbb{P} \otimes dt; \mathbb{R}^d)$ that pushes μ onto η , i.e.

$$dX_t = v_t dt \quad \text{with} \quad X_0 \sim \mu, X_1 \sim \eta.$$

Since $\eta \preceq_c \nu$, by the martingale representation theorem there exists $\sigma \in L^2(\mathbb{P} \otimes dt; \mathbb{R}^{d \times d})$, $M_0 \perp (B_t)_{t \in [0,1]}$ such that

$$(4) \quad dM_t = \sigma_t dB_t \quad \text{with} \quad M_0 \sim \eta, M_1 \sim \nu.$$

For any $\varepsilon \in (0, 1)$ define the process X^ε via

$$(5) \quad dX_t^\varepsilon = \frac{v_{\frac{t}{1-\varepsilon}}}{1-\varepsilon} \mathbf{1}_{\{0 \leq t \leq 1-\varepsilon\}} dt + \frac{\sigma_{\frac{t+\varepsilon-1}{\varepsilon}}}{\sqrt{\varepsilon}} \mathbf{1}_{\{1-\varepsilon < t \leq 1\}} dB_t \quad \text{with} \quad X_0^\varepsilon = X_0.$$

Then X^ε is an element of $\mathcal{D}(\mu, \nu)$ and we have

$$(6) \quad \mathcal{BB}^{1,0}(\mu, \nu) \leq \frac{1}{(1-\varepsilon)^2} \mathbb{E} \left[\int_0^1 |v_{\frac{t}{1-\varepsilon}}|^2 \mathbf{1}_{\{0 \leq t \leq 1-\varepsilon\}} dt \right] = \frac{1}{1-\varepsilon} \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right].$$

Minimizing over all such vector fields v , appealing to the Benamou-Brenier formula (DOT), and taking $\varepsilon \downarrow 0$, we get the desired inequality $\mathcal{BB}^{1,0}(\mu, \nu) \leq \mathcal{T}_2(\mu, \eta)$.

We now turn to proving the inequality $\inf_{\eta \preceq_c \nu} \mathcal{T}_2(\mu, \nu) \leq \mathcal{BB}^{1,0}(\mu, \nu)$. Suppose that $X \in \mathcal{D}(\mu, \nu)$, i.e.

$$dX_t = v_t dt + \sigma_t dB_t \quad \text{with} \quad X_0 \sim \mu, X_1 \sim \nu.$$

Let Y be given by

$$dY_t = \mathbb{E}[v_t | X_0] dt \quad \text{with} \quad Y_0 = X_0$$

and set $\widehat{\mu} := \text{Law}(Y_1)$. Then $\widehat{\mu} \preceq_c \nu$ as

$$\begin{aligned} Y_1 &= X_0 + \int_0^1 \mathbb{E}[v_t | X_0] dt = \mathbb{E} \left[X_0 + \int_0^1 v_t dt \middle| X_0 \right] \\ &= \mathbb{E} \left[X_0 + \int_0^1 v_t dt + \int_0^1 \sigma_t dB_t \middle| X_0 \right] = \mathbb{E}[X_1 | X_0]. \end{aligned}$$

Thus, (DOT), Jensen's inequality and Tonelli's theorem yield

$$\begin{aligned} \inf_{\eta \preceq_c \nu} \mathcal{T}_2(\mu, \eta) &\leq \mathcal{T}_2(\mu, \widehat{\mu}) \leq \mathbb{E} \left[\int_0^1 |\mathbb{E}[v_t | X_0]|^2 dt \right] \\ &\leq \mathbb{E} \left[\int_0^1 \mathbb{E}[|v_t|^2 | X_0] dt \right] = \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right]. \end{aligned}$$

As $X \in \mathcal{D}(\mu, \nu)$ was arbitrary, this concludes the proof. \square

We now give the proof of Theorem 1.

Proof of Theorem 1. We first show $\overline{\mathcal{T}}_2(\mu, \nu) \leq \mathcal{BB}^{1,0}(\mu, \nu)$. Take a process $X \in \mathcal{D}(\mu, \nu)$, i.e.

$$dX_t = v_t dt + \sigma_t dB_t \quad \text{with} \quad X_0 \sim \mu, X_1 \sim \nu.$$

By definition, $\text{Law}(X_0, X_1) \in \Pi(\mu, \nu)$. Applying Jensen's inequality,

$$\overline{\mathcal{T}}_2(\mu, \nu) \leq \mathbb{E} \left[|\mathbb{E}[X_1|X_0] - X_0|^2 \right] = \mathbb{E} \left[\left| \mathbb{E} \left[\int_0^1 v_t dt \middle| X_0 \right] \right|^2 \right] \leq \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right].$$

Minimizing over X yields the inequality $\overline{\mathcal{T}}_2(\mu, \nu) \leq \mathcal{BB}^{1,0}(\mu, \nu)$.

For the opposite inequality, let $(X_0, Y) \sim \pi \in \Pi(\mu, \nu)$. We set $v_t := \mathbb{E}[Y|X_0] - X_0$ and let X solve $dX_t = v_t dt$. Note that here v_t only depends on X_0 and is constant in t . Then

$$\eta := \text{Law}(X_1) = \text{Law}(\mathbb{E}[Y|X_0]) \preceq_c \text{Law}(Y) = \nu.$$

We now define (4) and (5) as in the proof of Lemma 6 above to obtain

$$\mathcal{BB}^{1,0}(\mu, \nu) \leq \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right] = \mathbb{E} \left[|\mathbb{E}[Y|X_0] - X_0|^2 \right]$$

as in (6). Minimizing over $(X_0, Y) \sim \pi \in \Pi(\mu, \nu)$ concludes the proof. \square

Combining Lemma 6 and the proof of Theorem 1 actually gives an independent proof of Proposition 4.

Corollary 7. *We have*

$$\overline{\mathcal{T}}_2(\mu, \nu) = \mathcal{BB}^{1,0}(\mu, \nu) = \inf_{\eta \preceq_c \nu} \mathcal{T}_2(\mu, \eta).$$

We now turn to the proof of Theorem 2.

Proof of Theorem 2. Suppose that $X \in \mathcal{D}(\mu, \nu)$, i.e.

$$dX_t = v_t dt + \sigma_t dB_t \quad \text{with} \quad X_0 \sim \mu, X_1 \sim \nu,$$

and define $\pi := \text{Law}(X_0, X_1) \in \Pi(\mu, \nu)$. Then

$$\begin{aligned} \int |\text{mean}(\pi_x) - x|^2 \mu(dx) &= \mathbb{E} \left[|\mathbb{E}[X_1|X_0] - X_0|^2 \right] \\ (7) \quad &= \mathbb{E} \left[\left| \mathbb{E} \left[\int_0^1 v_t dt + \int_0^1 \sigma_t dB_t \middle| X_0 \right] \right|^2 \right] \\ &= \mathbb{E} \left[\left| \mathbb{E} \left[\int_0^1 v_t dt \middle| X_0 \right] \right|^2 \right] \leq \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right], \end{aligned}$$

where the last inequality follows by two applications of Jensen's inequality. Similarly, recalling that $X_0 \perp (B_t)_{t \in [0,1]}$ and taking the possibly sub-optimal candidate $\varrho_x := \text{Law}(X_1, B_1 | X_0 = x) \in \Pi(\pi_x, \gamma_1^d)$ yields

$$\begin{aligned} (8) \quad \int_{\mathbb{R}^d} \text{MCov}(\pi_x, \gamma_1^d) \mu(dx) &\geq \mathbb{E} [\mathbb{E}[\langle X_1, B_1 \rangle | X_0]] \\ &= \mathbb{E} [\langle X_1, B_1 \rangle] = \mathbb{E} \left[\int_0^1 \langle v_t, B_t \rangle + \text{Tr}(\sigma_t) dt \right]. \end{aligned}$$

Combining (7) and (8) we deduce the inequality

$$\begin{aligned} \int_{\mathbb{R}^d} \alpha |\text{mean}(\pi_x) - x| - \beta \text{MCov}(\pi_x, \gamma_1^d) \mu(dx) \\ \leq \mathbb{E} \left[\int_0^1 \alpha |v_t|^2 - \beta (\langle v_t, B_t \rangle + \text{Tr}(\sigma_t)) dt \right], \end{aligned}$$

showing $\bar{\mathcal{T}}^{\alpha, \beta}(\mu, \nu) \leq \mathcal{BB}^{\alpha, \beta}(\mu, \nu)$.

For the inequality $\bar{\mathcal{T}}^{\alpha, \beta}(\mu, \nu) \geq \mathcal{BB}^{\alpha, \beta}(\mu, \nu)$, let κ and $\nabla\varphi$ be as in Proposition 3 and 4, i.e. $\kappa = \text{Law}(\widehat{M}_0, \widehat{M}_1)$ where \widehat{M} denotes the stretched Brownian motion from $\nabla\varphi_{\#}\mu$ to ν . Let us take $X_0 \sim \mu$ and apply the martingale representation theorem to write

$$\widehat{M}_t = \nabla\varphi(X_0) + \int_0^t \sigma_s dB_s$$

for some $\sigma \in L^2(\mathbb{P} \otimes dt; \mathbb{R}^{d \times d})$ and $X_0 \perp (B_t)_{t \in [0, 1]}$. Next, we set $v_t = \nabla\varphi(X_0) - X_0$ and define the process X via

$$dX_t = v_t dt + \sigma_t dB_t.$$

By definition, $\pi := \text{Law}(X_0, X_1)$ is an element of $\Pi(\mu, \nu)$ and $\pi_x = \kappa_{\nabla\varphi(x)}$. By Proposition 5 we conclude that π is the minimizer of $\bar{\mathcal{T}}^{\alpha, \beta}(\mu, \nu)$. Furthermore,

$$(9) \quad \mathbb{E} \left[\int_0^1 |v_t|^2 dt \right] = \mathbb{E} \left[|\nabla\varphi(X_0) - X_0|^2 \right] = \int |\text{mean}(\kappa_{\nabla\varphi(x)}) - x|^2 \mu(dx).$$

Next we observe that by Proposition 3,

$$(10) \quad \int \text{MCov}(\kappa_{\nabla\varphi(x)}, \gamma_1^d) \mu(dx) = \int \text{MCov}(\pi_x, \gamma_1^d) \mu(dx) = \mathbb{E} \left[\int_0^1 \text{Tr}(\sigma_t) dt \right].$$

Lastly, by Fubini's theorem and $X_0 \perp (B_t)_{t \in [0, 1]}$, we have

$$\begin{aligned} (11) \quad \mathbb{E} \left[\int_0^1 \langle v_t, B_t \rangle dt \right] &= \int_0^1 \mathbb{E} [\langle \nabla\varphi(X_0) - X_0, B_t \rangle] dt \\ &= \int_0^1 \langle \mathbb{E} [\nabla\varphi(X_0) - X_0], \mathbb{E} [B_t] \rangle dt = 0. \end{aligned}$$

Combining (9)-(11) and using optimality of π we obtain

$$\begin{aligned} \bar{\mathcal{T}}^{\alpha, \beta}(\mu, \nu) &= \int_{\mathbb{R}^d} \alpha |x - \text{mean}(\kappa_{\nabla\varphi(x)})| - \beta \text{MCov}(\kappa_{\nabla\varphi(x)}, \gamma_1^d) \mu(dx) \\ &= \mathbb{E} \left[\int_0^1 \alpha |v_t|^2 - \beta (\langle v_t, B_t \rangle + \text{Tr}(\sigma_t)) dt \right] \geq \mathcal{BB}^{\alpha, \beta}(\mu, \nu). \end{aligned}$$

This concludes the proof. \square

Remark 8. In Theorems 1 and 2, the quadratic cost function can be generalized to any convex cost function using the same argument, noting that [BPRS25, Theorem 5.4] also holds for general convex cost functions. This is analogous to the extension of the Benamou-Brenier formula to convex cost functions [Bre04, PS25].

REFERENCES

- [ABC19] J-J Alibert, Guy Bouchitté, and Thierry Champion, *A new class of costs for optimal transport planning*, European Journal of Applied Mathematics **30** (2019), no. 6, 1229–1263.
- [Bas83] Richard F Bass, *Skorokhod imbedding via stochastic integrals*, Séminaire de Probabilités **17** (1983), 221–224.
- [BB00] Jean-David Benamou and Yann Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. (Heidelb.) **84** (2000), no. 3, 375–393.
- [BCH⁺24] Jean-David Benamou, Guillaume Chazareix, Marc Hoffmann, Gregoire Loeper, and François-Xavier Vialard, *Entropic semi-martingale optimal transport*, arXiv preprint arXiv:2408.09361 (2024).
- [BG18] Malcolm Bowles and Nassif Ghoussoub, *A theory of transfers: Duality and convolution*, arXiv preprint arXiv:1804.08563 (2018).
- [BHL13] Mathias Beiglböck, Pierre Henry-Labordere, and Friedrich Penkner, *Model-independent bounds for option prices—a mass transport approach*, Finance and Stochastics **17** (2013), no. 3, 477–501.
- [BJ16] Mathias Beiglböck and Nicolas Juillet, *On a problem of optimal transport under marginal martingale constraints*.
- [BPRS25] Mathias Beiglböck, Gudmund Pammer, Lorenz Riess, and Stefan Schrott, *The fundamental theorem of weak optimal transport*, 2025.
- [Bre91] Yann Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Communications on pure and applied mathematics **44** (1991), no. 4, 375–417.
- [Bre04] ———, *Extended monge-kantorovich theory*, Optimal Transportation and Applications: Lectures given at the CIME Summer School, held in Martina Franca, Italy, September 2-8, 2001, Springer, 2004, pp. 91–121.
- [BVBHK19] Julio Backhoff-Veraguas, Mathias Beiglböck, Martin Huesmann, and Sigrid Källblad, *Martingale benamou–brenier: a probabilistic perspective*, 2019.
- [BVBST25] Julio Backhoff-Veraguas, Mathias Beiglböck, Walter Schachermayer, and Bertram Tschiderer, *Existence of bass martingales and the martingale benamou–brenier problem in \mathbb{R}^d* , 2025.
- [Con19] Giovanni Conforti, *A second order equation for schrödinger bridges with applications to the hot gas experiment and entropic transportation cost*, Probability Theory and Related Fields **174** (2019), no. 1, 1–47.
- [Föl06] Hans Föllmer, *Random fields and diffusion processes*, École d’Été de Probabilités de Saint-Flour XV–XVII, 1985–87, Springer, 2006, pp. 101–203.
- [FS18] Max Fathi and Yan Shu, *Curvature and transport inequalities for markov chains in discrete spaces*.
- [GJ20] Nathael Gozlan and Nicolas Juillet, *On a mixture of brenier and strassen theorems*, Proceedings of the London Mathematical Society **120** (2020), no. 3, 434–463.
- [GL21] Ivan Guo and Grégoire Loeper, *Path dependent optimal transport and model calibration on exotic derivatives*, The Annals of Applied Probability **31** (2021), no. 3, 1232–1263.
- [GRS⁺18] Nathael Gozlan, Cyril Roberto, Paul-Marie Samson, Yan Shu, and Prasad Tetali, *Characterization of a class of weak transport-entropy inequalities on the line*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 3, 1667–1693.
- [GRST17] Nathael Gozlan, Cyril Roberto, Paul-Marie Samson, and Prasad Tetali, *Kantorovich duality for general transport costs and applications*, J. Funct. Anal. **273** (2017), no. 11, 3327–3405 (en).
- [Mar96a] Katalin Marton, *Bounding \bar{d} -distance by informational divergence: a method to prove measure concentration*, The Annals of Probability **24** (1996), no. 2, 857–866.
- [Mar96b] ———, *A measure concentration inequality for contracting markov chains*, Geometric & Functional Analysis GAFA **6** (1996), no. 3, 556–571.
- [Nut21] Marcel Nutz, *Introduction to entropic optimal transport*, Lecture notes, Columbia University (2021).

- [PS25] Brendan Pass and Yair Shenfeld, *A dynamical formulation of multi-marginal optimal transport*, arXiv preprint arXiv:2509.22494 (2025).
- [San15] Filippo Santambrogio, *Optimal transport for applied mathematicians*.
- [Sch32] Erwin Schrödinger, *Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique*, Annales de l'institut Henri Poincaré, vol. 2, 1932, pp. 269–310.
- [Shu20] Yan Shu, *From hopf-lax formula to optimal weak transfer plan*, SIAM Journal on Mathematical Analysis **52** (2020), no. 3, 3052–3072.
- [Str65] Volker Strassen, *The existence of probability measures with given marginals*, The Annals of Mathematical Statistics **36** (1965), no. 2, 423–439.
- [Tal95] Michel Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques **81** (1995), no. 1, 73–205.
- [Tal96] ———, *New concentration inequalities in product spaces*, Inventiones mathematicae **126** (1996), no. 3, 505–563.
- [TT14] Xiaolu Tan and Nizar Touzi, *Optimal transportation under controlled stochastic dynamics*, Annals of Probability **41** (2014), no. 5, 3201–3240.
- [Vil21] Cédric Villani, *Topics in optimal transportation*, vol. 58, American Mathematical Soc., 2021.

IVAN GUO
 SCHOOL OF MATHEMATICS
 MONASH UNIVERSITY
Email address: `Ivan.Guo@monash.edu`

SEVERIN NILSSON
 DEPARTMENT OF MATHEMATICS
 CARNEGIE MELLON UNIVERSITY
Email address: `snilsson@andrew.cmu.edu`

JOHANNES WIESEL
 DEPARTMENT OF MATHEMATICS
 UNIVERSITY OF COPENHAGEN
Email address: `wiesel@math.ku.dk`