Wasserstein Barycenter Convexity Detects Hilbertian Geometry

Bang-Xian Han[*] Deng-Yu Liu [†]

July 5, 2026

Abstract

A striking feature of the Lott–Sturm–Villani curvature-dimension condition is its Finsler flatness phenomenon: every finite-dimensional normed space equipped with Lebesgue measure satisfies the weak $\mathrm{CD}(0, n)$ condition, independently of whether the norm is Euclidean. This phenomenon, emphasized by Villani as both natural and surprising, shows that entropy convexity along Wasserstein geodesics alone does not distinguish Hilbertian flat geometry from genuinely non-Hilbertian Finsler geometry.

We show that Wasserstein barycentric convexity provides precisely such a distinction. We prove that if a finite-dimensional normed vector space, equipped with Lebesgue measure, satisfies the Wasserstein Jensen's inequality for the Boltzmann entropy at barycenters, then its norm must be induced by an inner product. Thus Wasserstein barycenters provide an intrinsic optimal-transport test for Hilbertian structures. As a consequence, smooth reversible Finsler manifolds satisfying the corresponding barycentric curvature-dimension condition have Riemannian tangent norms.

The proof does not assume smoothness or strict convexity of the norm. Its two main ingredients, a rank-one polarization argument and a maximal-face trapping argument, are also of independent interest for the optimal transport theory.

Keywords. Wasserstein barycenter; curvature-dimension condition; displacement convexity; Finsler geometry; optimal transport.

MSC 2020. Primary 53C23, 49Q22; Secondary 46B20, 53C60, 28A75.

Contents

1 Introduction 1

2 Preliminaries 4

3 The strictly convex case 9

4 Strict convexity from BCD 13

1 Introduction

The displacement convexity of entropy functionals along W_2 geodesics, introduced by McCann in [15], is one of the most important notions in modern optimal transport theory. On Riemannian

[*]School of Mathematics, Shandong University, Jinan, China. Email: hanbx@sdu.edu.cn.

[†]Corresponding Author, School of Mathematical Sciences, University of Science and Technology of China, Hefei, China. Email: yzldy@mail.ustc.edu.cn

manifolds, this notion relates Ricci curvature bounds, optimal transport, and Prékopa–Leindler type inequalities via Riemannian interpolation inequalities of Cordero-Erausquin–McCann–Schmuckenschläger [6, 7]. This point of view led Lott–Villani and Sturm to the synthetic curvature-dimension theory on metric-measure spaces [14, 22, 23]. In Lott–Sturm–Villani's theory, curvature-dimension conditions are expressed through convexity of entropy along L^2-Wasserstein geodesics. This provides a formulation of lower Ricci bounds outside the smooth Riemannian setting.

The flatness paradox of the CD condition. As Villani notes in *the discussion around the last theorem in his book "Old and New"* [24, p. 926], every finite-dimensional normed vector space equipped with Lebesgue measure satisfies the standard synthetic non-negative Ricci condition in the weak Lott–Villani–Sturm sense. This conclusion is compatible with the Finsler interpretation of curvature-dimension bounds, but it is geometrically surprising from the Riemannian viewpoint: when the norm is not Euclidean, the space is not infinitesimally Hilbertian and fails the splitting behaviour expected for Ricci-limit spaces; if the norm is not strictly convex, branching of geodesics also occurs. In [24, p. 926-931], Villani explicitly raised the question of what additional regularity principles could rule out such non-Euclidean normed models. In Ambrosio–Gigli–Savaré's RCD theory [2, 8], Finsler structures are excluded by extra Riemannian assumptions: infinitesimal Hilbertianity, linearity of the heat flow, or equivalently the quadraticity of the Cheeger energy. For the Finsler side of the curvature-dimension theory, including weighted Ricci curvature and nonlinear analytic aspects, see also [16, 17, 18, 20].

The basic question of this paper is whether one can detect the Riemannian, rather than merely Finsler, nature of a space using only entropy inequalities in optimal transport. For ordinary displacement convexity, the answer is negative: in the weak Lott–Villani–Sturm theory, finite-dimensional Banach spaces already satisfy the nonnegative curvature-dimension condition. Our result shows that the answer becomes positive once geodesic interpolation is replaced by Wasserstein barycenters.

Motivated by Wasserstein barycenter geometry [1, 13], an abstract barycenter curvature-dimension condition BCD was introduced in [9, 10]. In this condition displacement interpolation in the Lott–Sturm–Villani theory is replaced by Wasserstein barycenters, and the entropy functionals are required to satisfy the Wasserstein Jensen's inequality at Wasserstein barycenters of finite weighted families of measures.

Definition 1.1 (BCD condition)**.** *Let $(X, \mathsf{d}, \mathfrak{m})$ be a geodesic metric measure space. We say that it satisfies $\mathrm{BCD}(K, \infty)$ if for every finite family $\mu_i \in \mathcal{P}_2(X)$ and every set of weights $\lambda_i > 0$, $\sum_i \lambda_i = 1$, there exists $\bar{\mu} \in \mathrm{argmin}_{\eta \in \mathcal{P}_2(X)} \sum_i \lambda_i W_2^2(\eta, \mu_i)$ (called a Wasserstein barycenter) such that*

$$\mathrm{Ent}_{\mathfrak{m}}(\bar{\mu}) \leq \sum_i \lambda_i \,\mathrm{Ent}_{\mathfrak{m}}(\mu_i) - \frac{K}{2} \sum_i \lambda_i W_2^2(\bar{\mu}, \mu_i), \tag{1.1}$$

where the Boltzmann entropy $\mathrm{Ent}_{\mathfrak{m}}(\cdot)$ is defined by

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) := \begin{cases} \int \rho \ln \rho \, \mathrm{d}\mathfrak{m} & \text{if } \mu \ll \mathfrak{m},\ \mu = \rho\, \mathfrak{m}, \\ +\infty & \text{otherwise.} \end{cases}$$

and $W_2(\cdot, \cdot)$ is the 2-Wasserstein distance associated with d, with quadratic cost d^2.

The condition is stated in the usual existential form: for each finite family, it suffices that at least one barycenter satisfies Jensen's inequality.

For two marginals, weighted Wasserstein barycenters are the corresponding points on W_2-geodesics. Thus the above condition gives

$$\mathrm{BCD}(0, \infty) \Longrightarrow \mathrm{CD}(0, \infty)$$

in the weak Lott–Villani sense, with the same entropy convention.

On smooth Riemannian manifolds, this stronger condition has the expected meaning. Barycentric Jensen inequalities are closely related to the multi-marginal extension of the Cordero-Erausquin–McCann–Schmuckenschläger theory developed by Kim–Pass [13]; in particular, the non-negative Ricci case is covered under the regularity hypotheses considered there. Conversely, applying the two-marginal case recovers the usual displacement convexity of entropy, and the theorem of von Renesse–Sturm [25] gives the Ricci lower bound.

The main theorem below states that, outside the Riemannian category, BCD has a rigidity that CD does not have. In particular, the barycenter curvature-dimension condition excludes the Finsler examples. The point is that a single Wasserstein geodesic tests one transport direction, whereas a barycenter tests the compatibility of several calibrated directions. This multi-directional compatibility is where the Hilbertian structure is forced.

Theorem 1.2 (Barycenter convexity rigidity)**.** *If $(E, \|\cdot\|, \mathcal{L}^n)$ is a finite-dimensional normed vector space with Lebesgue measure and satisfies* $\mathrm{BCD}(0,\infty)$*, then $\|\cdot\|$ is induced by an inner product.*

Thus the passage from geodesics to barycenters is not a cosmetic strengthening of convexity: it changes the class of infinitesimal models. Together with the stability of the barycenter curvature-dimension condition under measured blow-up, the following corollary rules out non-Hilbertian reversible Finsler tangent norms. Its proof is given in Section 4.

Corollary 1.3 (Exclusion of non-Riemannian reversible Finsler tangents)**.** *Let $(M, F, \mathfrak{m})$ be a smooth reversible Finsler manifold with a smooth positive reference measure. Assume that it satisfies* $\mathrm{BCD}(K,\infty)$*. Then F_x is induced by an inner product on T_xM for every $x \in M$. In particular, the Finsler structure is Riemannian.*

In finite-dimensional normed spaces and in the corresponding Finsler setting, BCD detects the Riemannian case. Thus the barycenter curvature-dimension condition gives a new synthetic Riemannian curvature-dimension condition: it implies the usual CD condition, but excludes the Finsler examples allowed by CD.

According to [9], RCD spaces satisfy the corresponding Wasserstein Jensen's inequality. The comparison with the RCD theory would nevertheless be understood as a motivation rather than as an equivalence statement. In the RCD framework, the Riemannian character of the space is imposed through infinitesimal Hilbertianity. The present theorem points to another route: the Wasserstein Jensen's inequality already carries a Riemannianity constraint. This suggests studying a curvature-dimension theory in which Wasserstein barycenters, rather than only Wasserstein geodesics, are the testing objects. In that direction, one may ask which analytic and geometric consequences follow directly from the Wasserstein Jensen's inequality, and whether it yields stability, localization, rigidity, or regularity results that are not visible from two-point displacement convexity alone.

The theorem also gives a Hilbertian characterization. Kakutani's theorem [12] detects the inner product by requiring every two-dimensional section of the unit ball to be an ellipse; the Jordan–von Neumann theorem [11] does so through the parallelogram law. For background on Hilbertian characterizations, see [3]. The characterization here comes from the Wasserstein Jensen's inequality.

Proof strategy

The proof has two main ingredients. In the strictly convex case, the argument works directly with calibrated rank-one transport branches; in the non-strictly convex case, it exploits the convex geometry of the unit ball.

The rank-one polarization method. For every covector r, the square potential $\alpha r(x)^2/2$ generates the affine rank-one map

$$x \longmapsto x - \alpha r(x) J(r).$$

For small α this potential is c-concave by a one-dimensional dual-norm estimate. The four-branch identity

$$\Phi_{p+q} + \Phi_{p-q} - 2\Phi_p - 2\Phi_q = 0, \qquad \Phi_r(x) = \frac{1}{2} r(x)^2,$$

produces a zero-sum calibrated family with a unique Wasserstein barycenter. Strict convexity gives uniqueness, and the BCD condition can then be applied directly. Applying BCD to the family gives the parallelogram equality.

The maximal-face method. To prove strict convexity, we argue by contradiction. If the unit ball has a non-trivial flat face, then a supporting covector exposes a positive-dimensional face. In this face we choose two extremal translation anchors and a third face-valued perturbation. The anchors force the uniqueness of the Wasserstein barycenter, while the third branch gives a strict first-order entropy decrease through the Jacobian formula. This contradicts BCD and rules out flat faces.

Organization.

Section 2 collects preliminaries on normed spaces, c-concavity, barycenters, calibrated maps, and entropy under changes of variables. Section 3 proves the Hilbertianity under the additional assumption of strict convexity. Section 4 proves that BCD rules out flat faces of the unit ball. Combining these two parts, we prove Theorem 1.2 and then prove Corollary 1.3.

2 Preliminaries

2.1 Norms, duality, and the Legendre map

Let $(E, \|\cdot\|)$ be a normed space and $(E^*, \|\cdot\|_*)$ be its dual space. Throughout the paper we set

$$h(v) = \frac{1}{2}\|v\|^2, \qquad c(x, y) = h(y - x).$$

This rescaling does not affect barycenters. For the case $K = 0$, which is the only case used in the proof of the main theorem, it also leaves Jensen's inequality unchanged; for general K it only changes the normalization of the curvature parameter. For probability measures $\mu, \nu \in \mathcal{P}_2(E)$, we write

$$\mathsf{C}(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \int_{E \times E} c(x, y) \, \mathrm{d}\pi(x, y).$$

Thus $\mathsf{C} = \frac{1}{2} W_2^2$ for the 2-Wasserstein distance induced by $\|\cdot\|$.

The convex conjugate of h is

$$h^*(p) = \frac{1}{2}\|p\|_*^2.$$

If $\|\cdot\|$ is strictly convex, by standard Fenchel duality and the differentiability criterion for finite convex functions (Šmulian duality); see Rockafellar [21, Theorems 23.5 and 25.1], h^* is differentiable on E^*, and we write

$$J(p) := \mathrm{D}h^*(p) \in E.$$

The map J is continuous, odd, and one-homogeneous; equivalently,

$$J(\lambda p) = \lambda J(p) \qquad \forall \lambda \in \mathbb{R}.$$

Moreover,

$$J(p) \in \partial h^*(p) = \left\{ v \in E : p(v) = \|p\|_*^2,\ \|v\| = \|p\|_* \right\}, \qquad p \in \partial h(J(p)).$$

Conversely, if $p \in \partial h(v)$ and h^* is differentiable at p, then $v = J(p)$.

2.2 c-concavity and optimal maps

For a function $\psi : E \to \mathbb{R} \cup \{-\infty\}$, define the c-transform by

$$\psi^c(y) := \inf_{x \in E} \{c(x, y) - \psi(x)\}.$$

A function ψ is c-concave if $\psi = \chi^c$ for some χ. Equivalently, ψ admits global c-supports: for every x there is y such that

$$\psi(z) \le c(z, y) - c(x, y) + \psi(x) \qquad \forall z \in E.$$

The c-subdifferential $\partial^c \psi(x)$ is the set of such y.

If ψ is differentiable at x and $y \in \partial^c \psi(x)$, then

$$\mathrm{D}\psi(x) \in -\partial h(y - x).$$

If $\|\cdot\|$ is strictly convex, equivalently h^* is differentiable, this gives

$$y - x = J(-\mathrm{D}\psi(x)).$$

Lemma 2.1 (Negative convex functions are c-concave)**.** *Let $F : E \to \mathbb{R}$ be convex. Then $-F$ is c-concave.*

Proof. Fix x_0 and choose $p \in \partial F(x_0)$. Choose $v \in \partial h^*(-p)$, equivalently $-p \in \partial h(v)$, and set $y = x_0 - v$. Convexity of F gives

$$F(z) - F(x_0) \ge p(z - x_0),$$

and the subgradient inequality for h gives

$$h(z - y) - h(x_0 - y) \ge -p(z - x_0).$$

Combining these inequalities yields

$$-F(z) \le h(z - y) - h(x_0 - y) - F(x_0) = c(z, y) - c(x_0, y) - F(x_0).$$

Thus $-F$ admits a global c-support at every x_0 and is c-concave. $\square$

We shall use the following dual calibration principle, whose idea goes back to [1, 5].

Lemma 2.2 (Zero-sum calibration/balance condition)**.** *Let $E = \mathbb{R}^n$, let $c(x, y) = h(y - x)$, and let $\lambda_1, \ldots, \lambda_m > 0$ satisfy $\sum_i \lambda_i = 1$. Let $\psi_i : E \to \mathbb{R}$ be c-concave functions such that*

$$\sum_{i=1}^{m} \lambda_i \psi_i = 0.$$

Set $\chi_i := \psi_i^c$. Assume that all integrals appearing below are well-defined; this will be the case in all applications, where the potentials are affine or quadratic and the measures have finite second moments. Let $\nu \in \mathcal{P}_2(E)$, and suppose that there are maps $T_i : E \to E$ such that

$$T_i(x) \in \partial^c \psi_i(x) \qquad \text{for } \nu\text{-a.e. } x.$$

Define

$$\mu_i := (T_i)_{\#}\nu.$$

Then ν is a barycenter of $(\mu_i)_{i=1}^m$ with weights $(\lambda_i)_{i=1}^m$, namely

$$\nu \in \operatorname*{argmin}_{\eta\in\mathcal{P}_2(E)} \sum_{i=1}^m \lambda_i W_2^2(\eta, \mu_i).$$

Here W_2 is the 2-Wasserstein distance induced by $\|\cdot\|$.

Proof. Since $\chi_i = \psi_i^c$, we have

$$\psi_i(x) + \chi_i(y) \le c(x,y) \qquad \forall x, y \in E.$$

Therefore, for every $\eta \in \mathcal{P}_2(E)$ and every $\pi \in \Pi(\eta, \mu_i)$,

$$\int c(x,y)\,\mathrm{d}\pi(x,y) \ge \int \psi_i\,\mathrm{d}\eta + \int \chi_i\,\mathrm{d}\mu_i.$$

Taking the infimum over π gives

$$\frac{1}{2}W_2^2(\eta,\mu_i) \ge \int \psi_i\,\mathrm{d}\eta + \int \chi_i\,\mathrm{d}\mu_i.$$

After multiplying by λ_i and summing in i, we obtain

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^m \lambda_i W_2^2(\eta,\mu_i) &\ge \int \sum_{i=1}^m \lambda_i\psi_i\,\mathrm{d}\eta + \sum_{i=1}^m \lambda_i \int \chi_i\,\mathrm{d}\mu_i \\ &= \sum_{i=1}^m \lambda_i \int \chi_i\,\mathrm{d}\mu_i.\end{aligned}$$

On the other hand, since

$$T_i(x) \in \partial^c\psi_i(x),$$

we have

$$\psi_i(x) + \chi_i(T_i(x)) = c(x, T_i(x)) \qquad \text{for } \nu\text{-a.e. } x.$$

Thus $(\mathrm{id}, T_i)_{\#}\nu$ is optimal from ν to μ_i, and

$$\frac{1}{2}W_2^2(\nu,\mu_i) = \int c(x,T_i(x))\,\mathrm{d}\nu(x) = \int \psi_i\,\mathrm{d}\nu + \int \chi_i\,\mathrm{d}\mu_i.$$

Averaging again and using $\sum_i \lambda_i\psi_i = 0$, we get

$$\frac{1}{2}\sum_{i=1}^m \lambda_i W_2^2(\nu,\mu_i) = \sum_{i=1}^m \lambda_i \int \chi_i\,\mathrm{d}\mu_i.$$

Hence ν attains the lower bound and is a barycenter. ☐

2.3 A lemma in convex geometry

In Section 4, we use this lemma to prove the uniqueness of the Wasserstein barycenter of a particular triple of measures $\{\mu_1, \mu_2, \mu_3\}$ in the non-strictly convex setting. The lemma and its proof are well known to experts, but we include them for completeness.

Lemma 2.3. *Let $F \subset \mathbb{R}^n$ be a nonempty compact convex set. Then there exists a nonzero linear functional $\ell \in (\mathbb{R}^n)^*$, such that ℓ attains both its maximum and its minimum on F at unique points.*

Proof. Define the support function of F by

$$h_F(\ell) := \max_{x\in F} \ell(x), \qquad \ell \in (\mathbb{R}^n)^*.$$

Since F is compact, h_F is finite everywhere. Moreover, since F is bounded, h_F is a convex Lipschitz function.

We first observe that

$$\partial h_F(\ell) = \operatorname*{argmax}_{x\in F} \ell(x),$$

where $\mathbb{R}^n$ is identified with its bidual.

Indeed, let $x \in F$ satisfy

$$\ell(x) = h_F(\ell).$$

Then, for every $m \in (\mathbb{R}^n)^*$,

$$h_F(m) \geq m(x) = \ell(x) + (m - \ell)(x) = h_F(\ell) + (m - \ell)(x).$$

Hence $x \in \partial h_F(\ell)$.

Conversely, suppose that $x \in \partial h_F(\ell)$. Since h_F is the Fenchel conjugate of the indicator function of F,

$$h_F^* = \iota_F,$$

the Fenchel equality implies that $x \in F$. Taking $m = 0$ in the subgradient inequality gives

$$0 = h_F(0) \geq h_F(\ell) - \ell(x),$$

and therefore

$$\ell(x) \geq h_F(\ell).$$

Since $x \in F$, the reverse inequality is automatic, and thus

$$\ell(x) = h_F(\ell).$$

This proves the subdifferential identity.

For a finite convex function on a finite-dimensional space, differentiability at a point is equivalent to the subdifferential at that point being a singleton. Therefore,

$$h_F \text{ is differentiable at } \ell \quad \Longleftrightarrow \quad \operatorname*{argmax}_{x\in F} \ell(x) \text{ is a singleton.}$$

By the Rademacher theorem for Lipschitz functions, the complement of the set

$$D_F := \{\ell \in (\mathbb{R}^n)^* : h_F \text{ is differentiable at } \ell\}$$

is a Lebesgue-null set. The same is true for $-F$.

Thus,

$$D_F \cap D_{-F}$$

contains a nonzero functional ℓ. For such an ℓ, the functional ℓ has a unique maximizer on F and a unique minimizer on F. □

2.4 Jacobian formulas for entropy

The entropy computations used below are based on finite-dimensional change-of-variables formulas. We state them here to avoid ambiguity with Wasserstein first variation formulas (which ask for more regularity).

Lemma 2.4 (Entropy under a change of variables)**.** *Let $\nu = \rho\mathcal{L}^n$ be an absolutely continuous probability measure with finite entropy. Let S be a C^1-diffeomorphism from an open neighbourhood of* $\operatorname{spt}\nu$ *onto its image, and assume*

$$\det \mathrm{D}S(x) > 0 \qquad \text{for all } x \in \operatorname{spt}\nu.$$

Set $\nu_S := S_\#\nu$. Then

$$\operatorname{Ent}(\nu_S) = \operatorname{Ent}(\nu) - \int_E \log\det \mathrm{D}S(x)\,\mathrm{d}\nu(x).$$

Proof. Let ρ_S be the density of ν_S. The Jacobian identity gives

$$\rho_S(S(x))\det \mathrm{D}S(x) = \rho(x) \qquad \text{for } \nu\text{-a.e. } x.$$

Changing variables in the entropy integral yields

$$\begin{aligned}\operatorname{Ent}(\nu_S) &= \int_E \rho_S(S(x))\log\rho_S(S(x))\det \mathrm{D}S(x)\,\mathrm{d}x\\ &= \int_E \rho(x)\big(\log\rho(x) - \log\det \mathrm{D}S(x)\big)\,\mathrm{d}x,\end{aligned}$$

which is the claim. □

Corollary 2.5 (Smooth perturbations)**.** *Let $\nu = \rho\mathcal{L}^n$ be as in Lemma 2.4. Let*

$$V = v_0 + Z, \qquad v_0 \in E, \quad Z \in C^1_c(E;E).$$

For $|t|$ sufficiently small define $S_t(x) := x + tV(x)$. Then

$$\operatorname{Ent}((S_t)_\#\nu) = \operatorname{Ent}(\nu) - \int_E \log\det(I + t\mathrm{D}V(x))\,\mathrm{d}\nu(x),$$

and consequently

$$\operatorname{Ent}((S_t)_\#\nu) = \operatorname{Ent}(\nu) - t\int_E \operatorname{div} V\,\mathrm{d}\nu + O(t^2).$$

Equivalently, if $\mathrm{d}\nu = \rho\,\mathrm{d}x$, then

$$\operatorname{Ent}((S_t)_\#\nu) = \operatorname{Ent}(\nu) - t\int_E \rho\,\operatorname{div} V\,\mathrm{d}x + O(t^2).$$

Proof. For $|t|$ small, S_t is a C^1-diffeomorphism on a neighbourhood of $\operatorname{spt}\nu$. Lemma 2.4 gives the exact logarithmic determinant formula. The expansion follows from

$$\log\det(I + tA) = t\operatorname{tr} A + O(t^2\|A\|^2),$$

uniformly for $A = \mathrm{D}V(x)$, since $\mathrm{D}V$ is bounded. □

Corollary 2.6 (Affine maps)**.** *Let $A : E \to E$ be an invertible linear map with* $\det A > 0$, *and let $b \in E$. Then, for every absolutely continuous probability measure ν with finite entropy,*

$$\operatorname{Ent}((x \mapsto Ax + b)_\#\nu) = \operatorname{Ent}(\nu) - \log\det A.$$

3 The strictly convex case

In this section, assume that a finite-dimensional normed space $(E, \|\cdot\|, \mathcal{L}^n)$ satisfies $\mathrm{BCD}(0,\infty)$ and that $\|\cdot\|$ is strictly convex. Strict convexity gives uniqueness for the Wasserstein barycenters used below, so the BCD condition can be applied directly. Strict convexity of the primal norm implies that $h^*(p) = \frac{1}{2}\|p\|_*^2$ is differentiable on E^*. We write

$$J := \mathrm{D}h^* : E^* \to E.$$

The purpose of this section is to prove that $\|\cdot\|_*$ satisfies the parallelogram identity. This will imply that the dual norm, and hence also the original norm, is Hilbertian.

The perturbations used below are rank-one affine perturbations generated by square functions

$$x \longmapsto \frac{1}{2}r(x)^2, \qquad r \in E^*.$$

Each branch is an extendable Wasserstein geodesic. The entropy computations are affine Jacobian computations covered by Corollary 2.6.

3.1 Rank-one square potentials

For $r \in E^*$, set

$$m_r^2 := r(J(r)) = \|r\|_*^2.$$

The equality follows from the two-homogeneity of h^*. If $r = 0$, then $m_r = 0$ and all formulas below are interpreted in the trivial sense.

For $\alpha \in \mathbb{R}$, define

$$\Phi_{\alpha,r}(x) := \frac{\alpha}{2}\, r(x)^2,$$

and define the affine rank-one map

$$T_{\alpha,r}(x) := x - \alpha r(x) J(r).$$

Thus $T_{\alpha,r} = I - \alpha J(r) \otimes r$. Throughout this subsection we use the strict convexity assumption, so that h^* is differentiable on E^*.

Lemma 3.1 (Rank-one square potentials are calibrated)**.** *Let $r \in E^*$ and $\alpha \in \mathbb{R}$. Assume*

$$\alpha m_r^2 < 1. \tag{3.1}$$

Then $\Phi_{\alpha,r}$ is c-concave and

$$T_{\alpha,r}(x) \in \partial^c \Phi_{\alpha,r}(x) \qquad \forall x \in E.$$

Moreover the equality set is single-valued:

$$\partial^c \Phi_{\alpha,r}(x) = \{T_{\alpha,r}(x)\} \qquad \forall x \in E.$$

Proof. If $r = 0$, then $\Phi_{\alpha,r} \equiv 0$ and $T_{\alpha,r} = \mathrm{id}$. Since $h \geq 0$ and $h(v) = 0$ only for $v = 0$, the assertion is immediate. We henceforth assume $r \neq 0$.

Fix $x \in E$ and put

$$y := T_{\alpha,r}(x) = x - \alpha r(x) J(r).$$

We prove the global c-support inequality

$$\Phi_{\alpha,r}(z) - \Phi_{\alpha,r}(x) \leq h(y - z) - h(y - x) \qquad \forall z \in E. \tag{3.2}$$

Write

$$s := r(x), \qquad u := z - x, \qquad t := r(u) = r(z - x).$$

Then

$$\Phi_{\alpha,r}(z) - \Phi_{\alpha,r}(x) = \frac{\alpha}{2}\big((s+t)^2 - s^2\big) = \alpha st + \frac{\alpha}{2}t^2. \tag{3.3}$$

On the other hand, by the dual norm inequality,

$$h(w) = \frac{1}{2}\|w\|^2 \ge \frac{r(w)^2}{2\|r\|_*^2} = \frac{r(w)^2}{2m_r^2} \qquad \forall w \in E. \tag{3.4}$$

Since $J(r) \in \partial h^*(r)$, Fenchel equality gives

$$h(J(r)) + h^*(r) = r(J(r)) = m_r^2, \qquad h^*(r) = \frac{1}{2}m_r^2,$$

hence

$$h(J(r)) = \frac{1}{2}m_r^2. \tag{3.5}$$

Using (3.4) for $w = y - z = -\alpha s J(r) - u$, and using (3.5) for $y - x = -\alpha s J(r)$, we obtain

$$\begin{aligned} h(y-z) - h(y-x) &\ge \frac{r(-\alpha s J(r) - u)^2}{2m_r^2} - \frac{\alpha^2 s^2 m_r^2}{2} \\ &= \frac{(-\alpha s m_r^2 - t)^2}{2m_r^2} - \frac{\alpha^2 s^2 m_r^2}{2} \\ &= \alpha st + \frac{t^2}{2m_r^2}. \end{aligned}$$

The assumption gives $\alpha < m_r^{-2}$, hence in particular $\alpha \le m_r^{-2}$. Therefore

$$\alpha st + \frac{t^2}{2m_r^2} \ge \alpha st + \frac{\alpha}{2}t^2.$$

Together with (3.3), this proves (3.2). Hence $T_{\alpha,r}(x) \in \partial^c \Phi_{\alpha,r}(x)$ for every x, and $\Phi_{\alpha,r}$ is c-concave.

It remains to prove uniqueness of the c-subgradient. Let $z \in \partial^c \Phi_{\alpha,r}(x)$. Since $\Phi_{\alpha,r}$ is differentiable, the first-order necessary condition for a c-support gives

$$\mathrm{D}\Phi_{\alpha,r}(x) = \alpha r(x) r \in -\partial h(z - x).$$

Equivalently,

$$z - x \in \partial h^*(-\alpha r(x) r).$$

Because h^* is differentiable, the right-hand side is the singleton

$$\{J(-\alpha r(x) r)\} = \{-\alpha r(x) J(r)\}.$$

Thus $z = x - \alpha r(x) J(r) = T_{\alpha,r}(x)$. This proves the single-valued equality set. □

Lemma 3.2. *Let $r \in E^*$, $\alpha \in \mathbb{R}$, and assume $\alpha m_r^2 < 1$. Then $T_{\alpha,r}$ is invertible and*

$$\det T_{\alpha,r} = 1 - \alpha m_r^2 > 0.$$

Proof. The determinant formula is the rank-one determinant identity

$$\det(I - u \otimes r) = 1 - r(u),$$

applied to $u = \alpha J(r)$. The positivity follows from the assumption $\alpha m_r^2 < 1$, so $T_{\alpha,r}$ is invertible. □

3.2 The polarization perturbation

For $r \in E^*$, write

$$\Phi_r(x) := \frac{1}{2} r(x)^2.$$

We use the polarization identity

$$\Phi_{p+q} + \Phi_{p-q} - 2\Phi_p - 2\Phi_q = 0 \qquad \forall p, q \in E^*. \tag{3.6}$$

Indeed, it is just $(p+q)^2 + (p-q)^2 = 2p^2 + 2q^2$.

Fix $p, q \in E^*$. Choose $\varepsilon > 0$ so small that

$$\varepsilon M(p,q) < 1, \qquad M(p,q) := \max\{\|p+q\|_*^2, \|p-q\|_*^2, 2\|p\|_*^2, 2\|q\|_*^2\}, \tag{3.7}$$

with the convention that there is no restriction if $M(p,q) = 0$. Define four potentials

$$\psi_1 := \varepsilon\Phi_{p+q}, \qquad \psi_2 := \varepsilon\Phi_{p-q}, \qquad \psi_3 := -2\varepsilon\Phi_p, \qquad \psi_4 := -2\varepsilon\Phi_q. \tag{3.8}$$

By (3.6),

$$\psi_1 + \psi_2 + \psi_3 + \psi_4 = 0. \tag{3.9}$$

By Lemma 2.1 and Lemma 3.1, each ψ_i is c-concave. The corresponding calibrated maps are

$$\begin{aligned} T_1(x) &= x - \varepsilon(p+q)(x)J(p+q), \\ T_2(x) &= x - \varepsilon(p-q)(x)J(p-q), \\ T_3(x) &= x + 2\varepsilon p(x)J(p), \\ T_4(x) &= x + 2\varepsilon q(x)J(q). \end{aligned} \tag{3.10}$$

Their determinants are

$$\begin{aligned} d_1 &:= \det T_1 = 1 - \varepsilon\|p+q\|_*^2, \\ d_2 &:= \det T_2 = 1 - \varepsilon\|p-q\|_*^2, \\ d_3 &:= \det T_3 = 1 + 2\varepsilon\|p\|_*^2, \\ d_4 &:= \det T_4 = 1 + 2\varepsilon\|q\|_*^2. \end{aligned} \tag{3.11}$$

All four determinants are positive by (3.7).

Proposition 3.3 (Unique barycenter for the polarization family)**.** *Let ν be a compactly supported absolutely continuous probability measure with finite entropy, and set*

$$\mu_i := (T_i)_\# \nu, \qquad i = 1, 2, 3, 4,$$

where the T_i are the four maps defined in (3.10). Then ν is the unique Wasserstein barycenter of $(\mu_1, \mu_2, \mu_3, \mu_4)$ with equal weights $1/4$.

Proof. By Lemma 2.2, we know that ν is a Wasserstein barycenter.

We now prove uniqueness. Let η be any Wasserstein barycenter of $(\mu_1, \mu_2, \mu_3, \mu_4)$ with equal weights. Hence

$$\frac{1}{4}\sum_{i=1}^{4} W_2^2(\eta, \mu_i) = \frac{1}{4}\sum_{i=1}^{4} W_2^2(\nu, \mu_i).$$

By Lemma 2.2 and (3.9), the value at ν is

$$\frac{1}{8}\sum_{i=1}^{4} W_2^2(\eta, \mu_i) = \frac{1}{8}\sum_{i=1}^{4} W_2^2(\nu, \mu_i) = \frac{1}{4}\sum_{i=1}^{4} \int_E \psi_i^c \, \mathrm{d}\mu_i. \tag{3.12}$$

For each i, define the nonnegative Kantorovich gap for the normalized cost c by

$$G_i(x,y) := c(x,y) - \psi_i(x) - \psi_i^c(y) \geq 0.$$

Let π_i be an optimal plan from η to μ_i for the cost c. Equality in (3.12) implies

$$\frac{1}{4}\sum_{i=1}^{4}\int_{E\times E} G_i(x,y)\,\mathrm{d}\pi_i(x,y) = 0.$$

Because each term is nonnegative, we have

$$G_i(x,y) = 0 \qquad \pi_i\text{-a.e.}$$

for every i. Hence

$$y \in \partial^c\psi_i(x) \qquad \pi_i\text{-a.e.}$$

By Lemma 3.1, the equality set of each ψ_i is single-valued:

$$\partial^c\psi_i(x) = \{T_i(x)\} \qquad \forall x \in E.$$

Therefore

$$\pi_i = (\mathrm{id}, T_i)_\# \eta, \qquad \text{and hence} \qquad \mu_i = (T_i)_\# \eta.$$

In particular, for $i = 1$,

$$(T_1)_\#\eta = \mu_1 = (T_1)_\#\nu.$$

By Lemma 3.2, T_1 is an invertible affine map. Pushing forward by T_1^{-1}, we obtain

$$\eta = \nu.$$

Thus ν is the unique barycenter. □

3.3 The entropy inequality gives the parallelogram law

Proposition 3.4 (BCD forces the dual parallelogram identity)**.** *Assume* $(E, \|\cdot\|, \mathcal{L}^n)$ *satisfies* $\mathrm{BCD}(0,\infty)$*, and assume the norm is strictly convex. Then*

$$\|p+q\|_*^2 + \|p-q\|_*^2 = 2\|p\|_*^2 + 2\|q\|_*^2 \qquad \forall p, q \in E^*. \tag{3.13}$$

Proof. Fix $p, q \in E^*$, and choose $\varepsilon > 0$ satisfying (3.7). Let ν be any compactly supported absolutely continuous probability measure with finite entropy, and construct $\mu_1, \ldots, \mu_4$ from (3.10). By Proposition 3.3, the unique barycenter is ν. Hence BCD gives

$$\mathrm{Ent}(\nu) \leq \frac{1}{4}\sum_{i=1}^{4}\mathrm{Ent}(\mu_i). \tag{3.14}$$

Each T_i is affine, so Corollary 2.6 gives

$$\mathrm{Ent}(\mu_i) = \mathrm{Ent}(\nu) - \log d_i,$$

where the determinants d_i are listed in (3.11). Substituting this into (3.14) yields

$$d_1 d_2 d_3 d_4 \leq 1. \tag{3.15}$$

In other words,

$$(1 - \varepsilon\|p+q\|_*^2)(1 - \varepsilon\|p-q\|_*^2)(1 + 2\varepsilon\|p\|_*^2)(1 + 2\varepsilon\|q\|_*^2) \leq 1. \tag{3.16}$$

Expanding (3.16) at first order as $\varepsilon \downarrow 0$ gives the inequality

$$\|p+q\|_*^2 + \|p-q\|_*^2 \geq 2\|p\|_*^2 + 2\|q\|_*^2. \tag{3.17}$$

Since (3.17) holds for every pair of covectors, we may apply it to the pair $p+q, p-q$. This gives

$$\|p+q\|_*^2 + \|p-q\|_*^2 \leq 2\|p\|_*^2 + 2\|q\|_*^2. \tag{3.18}$$

Combining (3.17) and (3.18) proves (3.13). ☐

Corollary 3.5 (Dual parallelogram identity implies Hilbertianity)**.** *If $\|\cdot\|_*$ satisfies* (3.13)*, then $\|\cdot\|$ is induced by an inner product on E.*

Proof. By the Jordan–von Neumann theorem, the dual norm $\|\cdot\|_*$ is induced by an inner product. Hence its dual norm $\|\cdot\|_{**}$ is also induced by an inner product. Since $\|\cdot\| = \|\cdot\|_{**}$, the result follows. ☐

4 Strict convexity from BCD

It remains to consider the case where the norm is not strictly convex. In this section, we show that BCD$(0,\infty)$ in fact forces strict convexity. The main difficulty in the non-strictly convex case is that Wasserstein barycenters of general distributions $\sum_{i=1}^m \lambda_i \delta_{\mu_i}$ are typically non-unique, so the BCD condition cannot be applied directly. We overcome this difficulty by constructing a special distribution below whose Wasserstein barycenter is unique, using the flat faces of the unit ball. Recall that the presence of flat faces in the unit ball is precisely the obstruction to strict convexity.

Before entering the proof, let us isolate the mechanism. A non-trivial boundary segment produces a positive-dimensional exposed face F. Two extremal translations inside this face will trap any barycenter by a one-dimensional separating functional, while a third face-valued branch is chosen to have negative averaged divergence. The trapping gives uniqueness of the barycenter, and the negative divergence contradicts the Jensen entropy inequality.

Proposition 4.1 (Maximal-face trapping: barycenter convexity rules out flat faces)**.** *If $(E, \|\cdot\|, \mathcal{L}^n)$ satisfies* BCD$(0,\infty)$*, then the unit ball of $\|\cdot\|$ has no non-trivial line segment. Hence $\|\cdot\|$ is strictly convex.*

Proof. Step 1: Characterization of the exposed face.

Assume that the unit ball $B = \{v : \|v\| \leq 1\}$ is not strictly convex. Then there exist distinct $u_0, u_1 \in \partial B$ with $[u_0, u_1] \subset \partial B$. Let $m = (u_0+u_1)/2$ and choose a supporting covector $p \in E^*$ such that

$$\|p\|_* = 1, \qquad p(m) = 1, \qquad p(v) \leq 1 \quad \forall v \in B.$$

Then $p(u_0) = p(u_1) = 1$, and

$$F := \{v \in B : p(v) = 1\}$$

is a positive-dimensional exposed face.

We first prove the face identity used throughout the proof. Since $h^*(p) = \frac{1}{2}\|p\|_*^2 = \frac{1}{2}$, Fenchel equality gives

$$v \in \partial h^*(p) \quad \Longleftrightarrow \quad h(v) + h^*(p) = p(v).$$

Writing $r = \|v\|$, the right-hand side implies

$$\frac{1}{2}r^2 + \frac{1}{2} = p(v) \leq \|p\|_* \|v\| = r.$$

Thus $(r-1)^2 \le 0$, so $r = 1$ and $p(v) = 1$, i.e. $v \in F$. Conversely, if $v \in F$, then $\|v\| = 1$ and $p(v) = 1$, so the Fenchel equality holds. Hence

$$\partial h^*(p) = F.$$

By the two-homogeneity and evenness of h^*, for every $\alpha \ge 0$,

$$\partial h^*(\alpha p) = \alpha F, \qquad \partial h^*(-\alpha p) = -\alpha F,$$

with the convention $0F = \{0\}$.

Step 2: A face-valued perturbation with negative averaged divergence.

Let $T = \operatorname{span}(F - F)$. Choose a linear functional ℓ on E whose restriction to F has a unique maximizer a and a unique minimizer b:

$$a = \operatorname*{argmax}_{f \in F} \ell(f), \qquad b = \operatorname*{argmin}_{f \in F} \ell(f).$$

Such an ℓ exists by Lemma 2.3.

All directional derivatives and divergences below are computed with respect to a fixed linear coordinate system inducing the Lebesgue measure $\mathcal{L}^n$. Pick $e_0 \in \operatorname{ri} F$ and $0 \ne \tau \in T$. Choose $\rho \in C_c^\infty(E)$, $\rho \ge 0$, $\int_E \rho \,\mathrm{d}x = 1$, such that $D_\tau \rho$ is not identically zero, where D_τ denotes the directional derivative in the direction τ. Set

$$\varphi := D_\tau \rho.$$

Then $\varphi \in C_c^\infty(E)$, and integration by parts gives

$$\int_E \rho \, D_\tau \varphi \,\mathrm{d}x = -\int_E (D_\tau \rho)^2 \,\mathrm{d}x < 0.$$

For sufficiently small $\delta > 0$, define

$$e(x) := e_0 + \delta \varphi(x) \tau$$

so that $e(x) \in F$ for every x. Indeed, $e_0 \in \operatorname{ri} F$ and $\tau \in \operatorname{span}(F - F)$, so $e_0 + s\tau \in F$ for all sufficiently small $|s|$. Since $\operatorname{div} e = \delta D_\tau \varphi$, we have

$$\int_E \rho \operatorname{div} e \,\mathrm{d}x < 0. \tag{4.1}$$

Step 3: Three calibrated branches and uniqueness of the barycenter.

Let $\nu = \rho \mathcal{L}^n$. For small $t > 0$, define

$$T_1^t(x) = x + ta, \qquad T_2^t(x) = x + tb, \qquad T_3^t(x) = x - 2te(x),$$

and set $\mu_i^t = (T_i^t)_\# \nu$. Since $\mathrm{D}e$ is bounded, for sufficiently small t the map $T_3^t = \mathrm{id} - 2te$ is a C^1-diffeomorphism on an open neighbourhood of $\operatorname{spt} \nu$. Thus all μ_i^t have finite entropy and compact support.

Set

$$q_1 = p, \qquad q_2 = p, \qquad q_3 = -2p,$$

and

$$\psi_i^t(y) := -t\, q_i(y).$$

Then $\psi_1^t + \psi_2^t + \psi_3^t = 0$. For a general covector $q \in E^*$,

$$\begin{aligned}(\psi_q^t)^c(z) &= \inf_y \{h(z-y) + tq(y)\} \\ &= tq(z) - h^*(tq).\end{aligned}$$

Equality holds if and only if

$$z - y \in \partial h^*(tq),$$

again by Fenchel equality. Therefore

$$T_1^t(x) - x = ta \in tF = t\partial h^*(p),$$

$$T_2^t(x) - x = tb \in tF = t\partial h^*(p),$$

and

$$T_3^t(x) - x = -2te(x) \in -2tF = t\partial h^*(-2p).$$

Thus the plans $(\mathrm{id}, T_i^t)_{\#}\nu$ are calibrated by ψ_i^t and their c-transforms. Lemma 2.2 implies that ν is a Wasserstein barycenter of $\mu_1^t, \mu_2^t, \mu_3^t$ with equal weights.

We now prove uniqueness. Let η be any Wasserstein barycenter, and denote by $\chi_i^t = (\psi_i^t)^c$ the dual partners. For each i, let π_i be an optimal plan from η to μ_i^t, and set

$$G_i^t(y, z) := c(y, z) - \psi_i^t(y) - \chi_i^t(z) \geq 0.$$

Since $\psi_1^t + \psi_2^t + \psi_3^t = 0$, we have

$$\frac{1}{3}\sum_{i=1}^{3}\int G_i^t \, \mathrm{d}\pi_i = \frac{1}{6}\sum_{i=1}^{3} W_2^2(\eta, \mu_i^t) - \frac{1}{3}\sum_{i=1}^{3}\int \chi_i^t \, \mathrm{d}\mu_i^t.$$

The calibrated plans from ν to μ_i^t give, by Lemma 2.2,

$$\frac{1}{6}\sum_{i=1}^{3} W_2^2(\nu, \mu_i^t) = \frac{1}{3}\sum_{i=1}^{3}\int \chi_i^t \, \mathrm{d}\mu_i^t.$$

Since η is also a Wasserstein barycenter, it has the same barycenter value as ν:

$$\frac{1}{3}\sum_{i=1}^{3} W_2^2(\eta, \mu_i^t) = \frac{1}{3}\sum_{i=1}^{3} W_2^2(\nu, \mu_i^t).$$

Therefore

$$\frac{1}{3}\sum_{i=1}^{3}\int G_i^t \, \mathrm{d}\pi_i = 0.$$

Because every $G_i^t \geq 0$, it follows that $G_i^t = 0$ for π_i-almost every point and every i. Consequently, for $i = 1, 2$ the optimal plan $\pi_i \in \Pi(\eta, \mu_i^t)$ is supported on the corresponding equality set,

$$z - y \in tF.$$

Consider $i = 1$. Let $S_a(x) = x + ta$. Since $\mu_1^t = (S_a)_{\#}\nu$, define

$$\widehat{\pi}_1 := \big((y, z) \mapsto (y, S_a^{-1}(z))\big)_{\#}\pi_1.$$

Writing $x = S_a^{-1}(z) = z - ta$, the plan $\widehat{\pi}_1$ has marginals η and ν. On its support, $z - y = tf$ for some $f \in F$, and $z = x + ta$; hence

$$y = x + t(a - f).$$

Therefore

$$\ell(y) - \ell(x) = t\big(\ell(a) - \ell(f)\big) \geq 0.$$

Integrating with respect to $\widehat{\pi}_1$ gives

$$\int \ell \,\mathrm{d}\eta - \int \ell \,\mathrm{d}\nu = \int t\big(\ell(a) - \ell(f)\big) \,\mathrm{d}\widehat{\pi}_1 \geq 0. \tag{4.2}$$

For $i = 2$, with $S_b(x) = x + tb$, define $\widehat{\pi}_2 := ((y,z) \mapsto (y, S_b^{-1}(z)))_{\#}\pi_2$. Writing $x = S_b^{-1}(z) = z - tb$, the same argument gives a plan with marginals η and ν and

$$y = x + t(b - f), \qquad f \in F.$$

Thus

$$\ell(y) - \ell(x) = t\big(\ell(b) - \ell(f)\big) \leq 0,$$

and hence

$$\int \ell \,\mathrm{d}\eta \leq \int \ell \,\mathrm{d}\nu. \tag{4.3}$$

Combining (4.2) and (4.3), equality holds in (4.2). The integrand $t(\ell(a) - \ell(f))$ is nonnegative, so it vanishes $\widehat{\pi}_1$-almost everywhere. Since a is the unique maximizer of ℓ on F, we have $f = a$ $\widehat{\pi}_1$-almost everywhere. Hence $y = x$ $\widehat{\pi}_1$-almost everywhere. The two marginals of $\widehat{\pi}_1$ are therefore equal, and $\eta = \nu$. Thus the barycenter is unique.

Step 4: The entropy contradiction.
Applying $\mathrm{BCD}(0,\infty)$ gives

$$\mathrm{Ent}(\nu) \leq \frac{1}{3}\,\mathrm{Ent}(\mu_1^t) + \frac{1}{3}\,\mathrm{Ent}(\mu_2^t) + \frac{1}{3}\,\mathrm{Ent}(\mu_3^t).$$

The first two measures are translations of ν, hence have the same entropy. Therefore

$$\mathrm{Ent}(\nu) \leq \mathrm{Ent}(\mu_3^t). \tag{4.4}$$

But $\mu_3^t = (\mathrm{id} - 2te)_{\#}\nu$. Applying Corollary 2.5 with $V = -2e$ gives

$$\mathrm{Ent}(\mu_3^t) = \mathrm{Ent}(\nu) + 2t \int_E \rho \,\mathrm{div}\, e \,\mathrm{d}x + O(t^2).$$

By (4.1), this is strictly smaller than $\mathrm{Ent}(\nu)$ for small $t > 0$, contradicting (4.4). Since every non-trivial boundary segment produces the positive-dimensional exposed face constructed in Step 1, this contradiction rules out all non-trivial boundary segments. Hence the unit ball is strictly convex. ☐

Proofs of the main theorem and the Finsler corollary

Proof of Theorem 1.2. Proposition 4.1 implies that the norm is strictly convex. Consequently h^* is differentiable and the rank-one square perturbations constructed above are available.

Proposition 3.4 gives the parallelogram identity. Corollary 3.5 then implies that the original norm is induced by an inner product. This proves the theorem. ☐

Proof of Corollary 1.3. Fix $x \in M$. For $r > 0$, consider the rescaled pointed metric-measure space

$$(M, r^{-1}\mathsf{d}_F, c_r\mathfrak{m}, x),$$

where $c_r > 0$ is a normalizing constant chosen for the measured blow-up. For the rescaled distance $r^{-1}\mathsf{d}_F$, the squared Wasserstein distances are multiplied by r^{-2}. Hence the original $\mathrm{BCD}(K,\infty)$ inequality becomes a $\mathrm{BCD}(Kr^2,\infty)$ inequality for the rescaled space. Multiplying the reference measure by $c_r > 0$ only adds the same constant to all entropy terms and therefore does not affect Jensen's inequality.

By the standard first-order blow-up of a smooth reversible Finsler manifold, the pointed measured spaces above converge, as $r \downarrow 0$, to

$$(T_x M, F_x, a_x \mathcal{L}_x, 0),$$

where $a_x > 0$ and $\mathcal{L}_x$ is a Lebesgue measure on $T_x M$; see [4, Chapter 1] for the Finsler tangent norm and [19] for the weighted Finsler metric-measure setting. Since $Kr^2 \to 0$, the stability theorem for the barycenter curvature-dimension condition under measured Gromov–Hausdorff convergence [9, Theorem 6.6] gives the limiting Wasserstein Jensen's inequality on the tangent space for compactly supported test families with finite entropy. In fact, such stability is a direct, almost verbatim adaptation of the proof of Lott–Villani [14, Theorem 4.15], since only the same Γ-convergence mechanism for the entropy and quadratic transport costs is involved.

It follows that F_x is induced by an inner product. Since x is arbitrary, F is Riemannian. □

Funding

This work was supported in part by the National Key R&D Programs of China (grants 2021YFA1000900 and 2021YFA1002200), the National Natural Science Foundation of China (12201596), the Shandong Provincial Natural Science Foundation (ZR2025QB05), and the Taishan Scholars Program of Shandong Province (tsqn202408059).

Acknowledgements

The authors would like to thank Emanuel Milman for valuable discussions concerning the relation between BCD and the classical curvature-dimension theory.

Declarations

The authors declare that they have no conflict of interest and that the manuscript has no associated data.

References

[1] M. Agueh and G. Carlier, *Barycenters in the Wasserstein space*, SIAM J. Math. Anal. 43 (2011), no. 2, 904–924.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J. 163 (2014), no. 7, 1405–1490.

[3] D. Amir, *Characterizations of Inner Product Spaces*, Operator Theory: Advances and Applications 20, Birkhäuser, Basel, 1986.

[4] D. Bao, S.-S. Chern, and Z. Shen, *An Introduction to Riemann–Finsler Geometry*, Graduate Texts in Mathematics 200, Springer, New York, 2000.

[5] G. Carlier and I. Ekeland, *Matching for teams*, Econom. Theory 42 (2010), no. 2, 397–418.

[6] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math. 146 (2001), no. 2, 219–257.

[7] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *Prékopa–Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport*, Ann. Fac. Sci. Toulouse Math. (6) 15 (2006), no. 4, 613–635.

[8] N. Gigli, *On the differential structure of metric measure spaces and applications*, Mem. Amer. Math. Soc. 236 (2015), no. 1113.

[9] B.-X. Han, D.-Y. Liu, and Z.-N. Zhu, *On the geometry of Wasserstein barycenter I*, arXiv:2412.01190v2 [math.MG], 2025.

[10] B.-X. Han, D.-Y. Liu, and Z.-N. Zhu, *Barycenter curvature-dimension condition for extended metric measure spaces*, Indag. Math. (N.S.) 37 (2026), no. 3, 892–902.

[11] P. Jordan and J. von Neumann, *On inner products in linear metric spaces*, Ann. of Math. (2) 36 (1935), no. 3, 719–723.

[12] S. Kakutani, *Some characterizations of Euclidean space*, Japan. J. Math. 16 (1939), 93–97.

[13] Y.-H. Kim and B. Pass, *Wasserstein barycenters over Riemannian manifolds*, Adv. Math. 307 (2017), 640–683.

[14] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) 169 (2009), no. 3, 903–991.

[15] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. 128 (1997), no. 1, 153–179.

[16] S. Ohta, *Finsler interpolation inequalities*, Calc. Var. Partial Differential Equations 36 (2009), no. 2, 211–249.

[17] S. Ohta, *Ricci curvature, entropy, and optimal transport*, in *Optimal Transportation*, London Math. Soc. Lecture Note Ser. 413, Cambridge University Press, Cambridge, 2014, 145–199.

[18] S. Ohta, *Nonlinear geometric analysis on Finsler manifolds*, Eur. J. Math. 3 (2017), no. 4, 916–952.

[19] S. Ohta and K.-T. Sturm, *Heat flow on Finsler manifolds*, Comm. Pure Appl. Math. 62 (2009), no. 10, 1386–1433.

[20] S. Ohta and K.-T. Sturm, *Bochner–Weitzenböck formula and Li–Yau estimates on Finsler manifolds*, Adv. Math. 252 (2014), 429–448.

[21] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series 28, Princeton University Press, Princeton, 1970.

[22] K.-T. Sturm, *On the geometry of metric measure spaces. I*, Acta Math. 196 (2006), no. 1, 65–131.

[23] K.-T. Sturm, *On the geometry of metric measure spaces. II*, Acta Math. 196 (2006), no. 1, 133–177.

[24] C. Villani, *Optimal transport: old and new*, Grundlehren der mathematischen Wissenschaften 338, Springer, Berlin, 2009.

[25] M.-K. von Renesse and K.-T. Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math. 58 (2005), no. 7, 923–940.