J. Math. Anal. Appl. 442 (2016) 1–16

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

# An order inequality characterizing invariant barycenters on symmetric cones

Sejong Kim [a], Hosoo Lee [b], Yongdo Lim [b,*]

[a] Department of Mathematics, Chungbuk National University, Cheongju 362-763, Republic of Korea
[b] Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Republic of Korea

ARTICLE INFO

*Article history:*
Received 12 March 2016
Available online 22 April 2016
Submitted by U. Stadtmueller

*Keywords:*
Wasserstein geometry
Contractive barycenter
Cartan barycenter
Hadamard space
Symmetric cone
Ando–Hiai inequality

ABSTRACT

This paper is concerned with invariant and contractive barycenters on the Wasserstein space of probability measures on metric spaces of non-positive curvature, where the center of gravity, also called the Cartan barycenter, is the canonical barycenter on Hadamard spaces. We establish an order inequality of probability measures on partially ordered symmetric spaces of non-compact type, namely symmetric cones (self-dual homogeneous cones), characterizing the Cartan barycenter among other invariant and contractive barycenters. The derived inequality and partially ordered structures on the probability measure space lead also to significant results on (norm) inequalities including the Ando–Hiai inequality for probability measures on symmetric cones.



## 1. Introduction

An isometry invariant barycentric map on the space of probability measures on a metric space with finite first or second moment with respect to the Wasserstein distance (alternatively Kantorovich–Rubinstein distance) plays a fundamental role in the fields of metric geometry, convex analysis, geometric analysis, statistical analysis, probability measure theory, optimal transport theory [12,16], to cite only a few. In [30] K.-T. Sturm develops a theory of barycenters of probability measures for metric spaces of nonpositive curvature, particularly that class of metric spaces known as CAT(0)-spaces or alternatively Hadamard spaces. For these spaces one has available a method for finding the barycenter of a probability measure via an approach stretching back to Cartan by finding the point that minimizes the integral of the square of distances. The canonical barycenter on a Hadamard space is the least squares barycenter. This barycenter has appeared under a variety of other designations: center of gravity, Frechet mean, Riemannian center of mass,

* Corresponding author.
*E-mail addresses:* skim@chungbuk.ac.kr (S. Kim), hosoolee@skku.edu (H. Lee), ylim@skku.edu (Y. Lim).

http://dx.doi.org/10.1016/j.jmaa.2016.04.048

Karcher barycenter, or frequently, Cartan barycenter, the terminology we adopt. The Cartan barycenter on a Hadamard metric space $(M,d)$ is defined as the unique point $\Lambda(\mu)$ that minimizes the variance function $z \mapsto \int_M [d^2(z,x) - d^2(y,x)]d\mu(x)$. It turns out that the Cartan barycentric map contracts the Wasserstein metric, a property that has been called the *fundamental contraction property* (Theorem 6.3, [30]).

A metric space admitting a contractive barycenter, called a *barycentric metric space*, is an important object to study in a variety of pure and applied areas, particularly in random variable theory (expectation and variance) on metric spaces. In fact, a complete, simply connected Riemannian manifold admits a contractive barycenter if and only if it has nonpositive curvature [30]. Another type of contractive barycenter on a Hadamard space, or generally on a Busemann NPC space, is the one by Es-Sahib and Heinich and Navas [10,28,2,19], which plays a fundamental role in geometric ergodic theory, fixed point theory and cocycle theory [28,7,2], and implies in particular that there are infinitely many invariant and contractive barycentric maps on the space of probability measures on Hadamard spaces. However, to the best of our knowledge, no one has found a characterizing property of the Cartan barycenter among other invariant (or contractive) barycenters on a Hadamard space. (Sturm's result gives the less direct characterization that it is the probabilistic limit of the inductive mean [30].) This problem is quite natural and important, and depends heavily and clearly on certain structures of the given Hadamard space.

Our main goal in this paper is to settle this problem on a class of partially ordered symmetric spaces of non-compact type, namely the class of symmetric cones. Symmetric cones, also called domains of positivity, are open convex self-dual cones in Euclidean space which have a transitive group of symmetries. By the Koecher–Vinberg theorem (cf. [11]) these correspond to the cone of squares in finite-dimensional real Euclidean Jordan algebras, originally classified by Jordan, von Neumann and Wigner.

We first consider the symmetric cone of positive definite matrices of fixed size equipped with the trace Riemannian metric, an important example of symmetric cones, and then move to general symmetric cones with extensions of the key ingredients appeared in the cone of positive definite matrices. As positive matrices have gained increased prominence in theoretical, applied, and computational settings, finding appropriate methods for averaging them has become an important task. They appear in a diverse variety of settings: covariance matrices in statistics, in Gaussian measures on Euclidean spaces, elements of the search space in convex and semidefinite programming, kernels in machine learning, density matrices in quantum information, data points in radar imaging, and diffusion tensors in medical imaging [21,27,29].

Let $\mathbb{P}$ be the convex cone of all positive definite matrices of size $m$ equipped with the Löwner order; $A \leq B$ if and only if $B - A$ is positive semidefinite. It is well known that $\mathbb{P}$ is a symmetric space of non-compact type with the trace metric $ds = ||A^{-1/2}dA\, A^{-1/2}||_2 = \left(\mathrm{tr}(A^{-1}dA)^2\right)^{1/2}$, and that congruence transformations $A \mapsto MAM^*$ over non-singular matrices $M$ and inversion $A \mapsto A^{-1}$ act as isometries on $\mathbb{P}$. We introduce and develop the partially ordered structures on the probability measure space equipped with the "upper set ordering" from the Löwner order on $\mathbb{P}$, and prove that the Cartan barycenter satisfies the following property

$$\int_{\mathbb{P}} \log X \; d\mu(X) \leq 0 \qquad \text{implies} \quad \Lambda(\mu) \leq I \tag{1.1}$$

which characterizes the Cartan barycenter on the space of probability measures with finite second moment, where 0 and $I$ stand for zero and identity matrices. As an important consequence of the main result we establish the Ando–Hiai inequality for the Cartan barycenter: for a probability measure $\mu$ with finite second moment,

$$\Lambda(\mu) \leq I \quad \text{implies} \qquad \Lambda(\mu^p) \leq I, \qquad \forall p \geq 1, \tag{1.2}$$

where $A^p$ is the matrix $p$-th power of $A$ and $\mu^p(\mathcal{O}) := \mu(\{A^{\frac{1}{p}} : A \in \mathcal{O}\})$ for $\mathcal{O} \in \mathcal{B}(\mathbb{P})$, the algebra of Borel subsets of $\mathbb{P}$. The Ando–Hiai inequality for two positive definite matrices plays a fundamental and significant role in matrix theory and it has many important applications such as operator means, operator monotone functions, statistical mechanics, quantum information theory, matrix-norm inequalities for unitarily invariant norms related to Araki–Lieb–Thirring, Golden–Thompson trace inequality, and log-majorizations, so our result will be a crucial ingredient in deriving these in the general setting of probability measures on the cone of positive definite matrices and generally on symmetric cones of Euclidean Jordan algebras. For instance, we establish a norm inequality for the Cartan barycenter: for the operator norm $||\cdot||$,

$$||\Lambda(\mu^q)^{\frac{1}{q}}|| \le ||\Lambda(\mu^p)^{\frac{1}{p}}||, \quad 0 < p \le q$$

which is the first result on norm inequalities related to the Cartan barycenter overcoming the limitation of finite discrete probability measures.

The main tools of the paper involve the theory of nonpositively curved metric spaces and techniques from probability and partially ordered Hadamard spaces and the recent combination of the three, particularly by K.-T. Sturm [30] and Lawson and Lim [20]. Not only are these tools crucial for our developments, but also, we believe, significantly enhance the potential usefulness of the Cartan barycenter for geometric analysis and inequalities and (partially ordered) probability measure theory on symmetric cones.

## 2. Invariant and contractive barycenters

Let $\mathbb{H}$ be the Euclidean space of $m \times m$ Hermitian matrices equipped with $\langle X, Y\rangle := \mathrm{Tr}(XY)$. The Frobenius norm $||\cdot||_2$ defined by $||X||_2 = (\mathrm{tr}\, X^2)^{1/2}$ for $X \in \mathbb{H}$ gives rise to the Riemannian structure on the open convex cone $\mathbb{P}$ of positive definite matrices with $\langle X, Y\rangle_A = \mathrm{Tr}(A^{-1}XA^{-1}Y)$, where $A \in \mathbb{P}$ and $X, Y \in T_A(\mathbb{P}) \equiv \mathbb{H}$. Then $\mathbb{P}$ is a Cartan–Hadamard Riemannian manifold, a simply connected complete Riemannian manifold with non-positive sectional curvature (the canonical 2-tensor is non-negative). The Riemannian metric distance between $A$ and $B$ is given by $d(A, B) = ||\log A^{-\frac{1}{2}}BA^{-\frac{1}{2}}||_2$, and the unique (up to parametrization) geodesic line containing $A$ and $B$ is $t \mapsto A\#_tB := A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^tA^{\frac{1}{2}}$. The Riemannian exponential at $A \in \mathbb{P}$ is given by

$$\exp_A(X) = A^{\frac{1}{2}}\exp(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})A^{\frac{1}{2}} \tag{2.3}$$

and its inverse is

$$\log_A(X) = A^{\frac{1}{2}}\log(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})A^{\frac{1}{2}}. \tag{2.4}$$

The following metric inequalities reflect the non-positive curvature of the trace metric [18,3,30]

$$d(A\#_tB, C\#_sD) \le (1-t)d(A, C) + td(B, D) + |t-s|d(C, D), \quad s, t \in [0, 1] \tag{2.5}$$

and

$$d^2(A\#_tB, C\#_tD) \le (1-t)d^2(A, C) + td^2(B, D) - (1-t)t[d(A, B) - d(C, D)]^2.$$

Let $\mathcal{B} := \mathcal{B}(\mathbb{P})$ be the algebra of Borel sets, the smallest $\sigma$-algebra containing the open sets of $\mathbb{P}$. We note that the Euclidean topology on $\mathbb{P}$ coincides with the metric topology of the trace metric $d$. Let $\mathcal{P}$ be the set of all probability measures on $(\mathbb{P}, \mathcal{B})$ and $\mathcal{P}_0$ the set of all $\mu \in \mathcal{P}$ of the form $\mu = (1/n)\sum_{j=1}^{n}\delta_{A_j}$, where $\delta_A$ is the point measure of mass 1 at $A \in \mathbb{P}$. For $p \in [1, \infty)$ let $\mathcal{P}^p$ be the set of probability measures with *finite* *p-moment*: for some (and hence all) $Y \in \mathbb{P}$,

$$\int_{\mathbb{P}} d^p(X,Y)d\mu(X) < \infty.$$

For metric spaces $M$ and $N$, a continuous $f : M \to N$ induces a *push-forward* map $f_* : \mathcal{P}(X) \to \mathcal{P}(Y)$ defined by $f_*(\mu)(B) = \mu(f^{-1}(B))$ for $\mu \in \mathcal{P}(M)$ and $B \in \mathcal{B}(N)$.

We say that $\omega \in \mathcal{P}(\mathbb{P} \times \mathbb{P})$ is a *coupling* for $\mu, \nu \in \mathcal{P}$ if $\mu$, $\nu$ are the marginals for $\omega$, i.e., if for all $B \in \mathcal{B}$, $\omega(B \times \mathbb{P}) = \mu(B)$ and $\omega(\mathbb{P} \times B) = \nu(B)$. Equivalently, $\mu$ and $\nu$ are the push-forwards of $\omega$ under the projection maps $\pi_1$ and $\pi_2$, respectively. We note that one such coupling is the product measure $\mu \times \nu$, and that for any coupling $\omega$ it must be the case that $\operatorname{supp}(\omega) \subseteq \operatorname{supp}(\mu) \times \operatorname{supp}(\nu)$. We denote the set of all couplings for $\mu, \nu \in \mathcal{P}$ by $\Pi(\mu, \nu)$.

The *Wasserstein distance* $d_p^W$ on $\mathcal{P}^p$ is defined by

$$d_p^W(\mu_1, \mu_2) := \left[ \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{\mathbb{P} \times \mathbb{P}} d^p(X,Y)d\pi(X,Y) \right]^{\frac{1}{p}}.$$

It is known that $d_p^W$ is a complete metric on $\mathcal{P}^p$ and $\mathcal{P}_0$ is dense in $\mathcal{P}^p$ [8,30]. Note that $\mathcal{P}_0 \subset \mathcal{P}^q \subset \mathcal{P}^p \subset \mathcal{P}^1$ and $d_p^W \leq d_q^W$ for $1 \leq p \leq q < \infty$.

For the following see the introduction of [31], also [28,7,9].

**Example 2.1.** For $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{A_j}$, $\nu = \frac{1}{n}\sum_{j=1}^n \delta_{B_j}$, and $1 \leq p < \infty$

$$d_p^W(\mu, \nu) = \min_{\sigma \in S_n} \left[ \frac{1}{n} \sum_{j=1}^n d^p(A_j, B_{\sigma(j)}) \right]^{\frac{1}{p}}$$

where $S_n$ denotes the permutation group on $n$-letters.

Explicit calculation of Wasserstein distance is very difficult for most concrete examples, except the previous cases, but the following estimation plays a crucial role for our main results.

**Example 2.2.** ([7]) For $\mu, \nu_1, \nu_2 \in \mathcal{P}^1$ and $t \in [0,1]$,

$$d_1^W((1-t)\mu + t\nu_1, (1-t)\mu + t\nu_2) \leq t \cdot \sup\{d(A_1, A_2) : A_j \in \operatorname{supp} \nu_j\}.$$

In particular, for any $A, B \in \mathbb{P}$

$$d_1^W((1-t)\mu + t\delta_A, (1-t)\mu + t\delta_B) \leq t d(A,B). \tag{2.6}$$

We note that these basic results on probability measure spaces hold in the general setting of complete metric spaces in which cases a separability assumption is necessary. In this paper we restrict our attention to the cases $p = 1$ and $p = 2$, the most important cases in probability measure theory.

Next, we introduce a $\mathrm{GL}_m$-action and $p$-th powers on $\mathcal{P}$. For $M \in \mathrm{GL}_m$, the general linear group, $A \in \mathbb{P}$, $\mu \in \mathcal{P}$, $p \in \mathbb{R} \setminus \{0\}$, and $\mathcal{O} \in \mathcal{B}(\mathbb{P})$, we let

$$M.A := MAM^*, \quad M.\mathcal{O} = \{MAM^* : A \in \mathcal{O}\}, \quad \mathcal{O}^p := \{A^p : A \in \mathcal{O}\}$$

and

$$(M.\mu)(\mathcal{O}) = \mu(M^{-1}.\mathcal{O}), \qquad \mu^p(\mathcal{O}) := \mu(\mathcal{O}^{\frac{1}{p}}). \tag{2.7}$$

For notational convenience, we let $t.\mu := (tI).\mu \in \mathcal{P}^1$ for positive reals $t$; that is,

$$(t.\mu)(\mathcal{O}) = \mu(t^{-2}\mathcal{O}).$$

Note that $M.\mu, \mu^p \in \mathcal{P}^r$ if $\mu \in \mathcal{P}^r$. In terms of push forward measures,

$$M.\mu = f_*\mu \quad \text{and} \quad \mu^p = g_*\mu,$$

where $f(X) = MXM^*$ and $g(X) = X^p$.

**Example 2.3.** The actions in (2.7) are natural and comparable with finitely supported measures: for $\mu = \frac{1}{n}\sum_{j=1}^{n}\delta_{A_j}$ and for $M \in \mathrm{GL}_m$,

$$\mu^p = \frac{1}{n}\sum_{j=1}^{n}\delta_{A_j^p} \quad \text{and} \quad M.\mu = \frac{1}{n}\sum_{j=1}^{n}\delta_{M.A_j} = \frac{1}{n}\sum_{j=1}^{n}\delta_{MA_jM^*}.$$

**Definition 2.4.** A map $\beta : \mathcal{P}^p \to \mathbb{P}$ is said to be a *barycenter* if it is idempotent in the sense that $\beta(\delta_X) = X$ for all $X \in \mathbb{P}$. A barycentric map $\beta$ is said to be *contractive* if

$$d(\beta(\mu_1), \beta(\mu_2)) \leq d_p^W(\mu_1, \mu_2)$$

for all $\mu_1, \mu_2 \in \mathcal{P}^p$, and is said to be *invariant* if for all $M \in \mathrm{GL}_m$ and $\mu \in \mathcal{P}^p$,

(i) $\beta(M.\mu) = M.\beta(\mu)$; and
(ii) $\beta(\mu^{-1}) = \beta(\mu)^{-1}$.

The Cartan barycenter $\Lambda : \mathcal{P}^2 \to \mathbb{P}$ is defined by

$$\Lambda(\mu) = \underset{Z\in\mathbb{P}}{\arg\min} \int_{\mathbb{P}} d^2(Z, X)d\mu(X).$$

The uniqueness and existence of the minimizer is well known in general setting of Hadamard spaces (Proposition 4.3, [30]). The Cartan barycenter $\Lambda$ on $\mathcal{P}^2$ can be extended to $\mathcal{P}^1$: for $\mu \in \mathcal{P}^1$, the unique minimizer of the uniformly convex, continuous function

$$Z \mapsto \int_{\mathbb{P}} [d^2(Z, X) - d^2(Y, X)]d\mu(X).$$

This point is independent of $Y$ and coincides with $\Lambda(\mu)$ for $\mu \in \mathcal{P}^2$. One can directly see that it is invariant under isometries on $\mathbb{P}$. The contractive property of $\Lambda$ follows immediately from Sturm's *fundamental contraction property* on Hadamard spaces.

**Theorem 2.5** *(Fundamental contraction property, [30]). For $\mu, \nu \in \mathcal{P}^2$,*

$$d(\Lambda(\mu), \Lambda(\nu)) \leq d_1^W(\mu, \nu) \leq d_2^W(\mu, \nu).$$

An alternative constructive scheme of contractive and invariant barycenters has been used by Es-Sahib and Heinich in [10]. Quite general conditions for constructing invariant and contractive barycenters in metric spaces including the Es-Sahib and Heinich barycenter are given by Lawson and Lim. From these construction schemes we conclude that there exist infinitely many distinct contractive and invariant barycenters on $\mathbb{P}$.

## 3. The Cartan barycenter and Karcher equation

The logarithm map $\log : \mathbb{P} \to \mathbb{H}$ satisfies $d(X, I) = d_F(\log_I X, 0)$, where the first distance is the trace metric and the second is the metric arising from the Frobenius norm. In the following, we simply denote $\log X = \log_I X$. Then it follows for $r \geq 1$ that $\int_{\mathbb{P}} d^r(X, I)d\mu(X) < \infty$ if and only if $\int_{\mathbb{H}} d_F^r(\log X, 0)d\mu_* < \infty$, where $\mu_* = \log_*(\mu)$, the push-forward of $\mu$. We conclude that the push-forward map $\log_*$ carries $\mathcal{P}^r(\mathbb{P})$ into $\mathcal{P}^r(\mathbb{H})$.

The Cartan barycenter on $\mathcal{P}^2$

$$\Lambda(\mu) = \underset{Z\in\mathbb{P}}{\arg\min} \int_{\mathbb{P}} d^2(Z, X)d\mu(X)$$

arises as the unique point where the gradient of the variance function

$$Z \mapsto \int_{\mathbb{P}} d^2(Z, X)d\mu(X)$$

vanishes:

$$\int_{\mathbb{P}} \log_Z(X)d\mu(X) = 0.$$

For the general setting of Riemannian manifolds with nonpositive curvature, see Karcher [15] and [29, Theorem 2]. Then from (2.4), the Cartan barycenter $\Lambda(\mu)$ is the unique positive definite solution $Z$ of the *Karcher equation*

$$\int_{\mathbb{P}} \log(Z^{-\frac{1}{2}}XZ^{-\frac{1}{2}})d\mu(X) = 0. \tag{3.8}$$

We observe that for a positive probability vector $\omega = (w_1, \ldots, w_n) \in \mathbb{R}^n$ and $\mu = \sum_{j=1}^{n} w_j\delta_{A_j} \in \mathcal{P}^0$, the Cartan barycenter of $\mu$ is determined by

$$\Lambda(\omega; A_1, \ldots, A_n) := \Lambda(\mu) = \underset{X\in\mathbb{P}}{\arg\min} \sum_{j=1}^{n} w_j d^2(X, A_j)$$

and the corresponding Karcher equation is given by

$$\sum_{j=1}^{n} w_j \log(X^{-\frac{1}{2}}A_jX^{-\frac{1}{2}}) = 0.$$

Since the work of Lawson and Lim [20], the Cartan barycenter of finitely supported measures has served as a “geometric center” and as the most attractive averaging among other matrix geometric means. A currently active research topic in matrix analysis is understanding, finding properties of, and computing efficiently the least squares mean [25,22,27].

We establish a characteristic property of the Cartan barycenter via the Karcher equation. For $A \in \mathbb{P}$, $t > 0$ and $\mu \in \mathcal{P}$, we define a probability measure $A\#_t\mu$ as

$$(A\#_t\mu)(\mathcal{O}) := \mu(\{A\#_t X : X \in \mathcal{O}\}).$$

That is,

$$A\#_t\mu = A^{\frac{1}{2}}.(A^{-\frac{1}{2}}.\mu)^{\frac{1}{t}} = f_*\mu,$$

where $f(X) = A\#_{\frac{1}{t}}X$ and the second equality follows from $f^{-1}(X) = A\#_t X$. One can see directly that $f_*\mu \in \mathcal{P}^r$ if $\mu \in \mathcal{P}^r$ and for $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{A_j}$,

$$\frac{1}{n}\sum_{j=1}^{n} \delta_{A\#_t A_j} = A\#_{\frac{1}{t}}\mu.$$

We consider the equation

$$X = \Lambda(X\#_t\mu), \qquad t > 0. \tag{3.9}$$

**Theorem 3.1.** *For $\mu \in \mathcal{P}^2$, the Cartan mean $\Lambda(\mu)$ is the unique solution of (3.9).*

**Proof.** Suppose that $A = \Lambda(A\#_t\mu)$. By the Karcher equation (3.8)

$$\int_{\mathbb{P}} \log(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})d(A\#_t\mu)(X) = 0.$$

Set $\nu := (A^{-\frac{1}{2}}.\mu)^{\frac{1}{t}} = g_*\mu$ and $g(X) = (A^{-\frac{1}{2}}XA^{-\frac{1}{2}})^{\frac{1}{t}}$. Then by change of variable theorem for Bochner integral,

$$0 = \int_{\mathbb{P}} \log X d\nu(X) = \int_{\mathbb{P}} \log(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})^{\frac{1}{t}} d\mu(X) = \frac{1}{t}\int_{\mathbb{P}} \log(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})d\mu(X).$$

This shows that

$$\int_{\mathbb{P}} \log(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})d\mu(X) = 0,$$

that is, $A = \Lambda(\mu)$. Hence, (3.9) has at most one solution, the Cartan barycenter $\Lambda(\mu)$. By the reverse implications in above, $\Lambda(\mu)$ is a solution of (3.9). This completes the proof. □

An important structure on the cone $\mathbb{P}$ is the Löwner ordering; $A \le B$ if $B - A$ is positive semidefinite. We introduce a partial order on $\mathcal{P}$ from the Löwner cone ordering $\le$ on $\mathbb{P}$.

**Definition 3.2** *(Partial ordering).* A subset $\mathcal{U} \subset \mathbb{P}$ is called an *upper set*, if whenever $A \in \mathcal{U}$ and $A \le B$, then $B \in \mathcal{U}$. For $\mu, \nu \in \mathcal{P}$, define $\mu \le \nu$ if $\mu(\mathcal{U}) \le \nu(\mathcal{U})$ for any upper set $\mathcal{U} \in \mathcal{B}(\mathbb{P})$, equivalently $\int_{\mathcal{U}} d\mu(X) \le \int_{\mathcal{U}} d\nu(X)$.

**Remark 3.3.** For $\mu \in \mathcal{P}$, $A \le B$ implies $\delta_A \le \delta_B$. Moreover,

$$A \le B \quad \text{implies} \quad \mu_A \le \mu_B \tag{3.10}$$

for $\mu \in \mathcal{P}$, where $\mu_A := \frac{1}{2}\mu + \frac{1}{2}\delta_A \in \mathcal{P}$.

In [17] Kim and Lee have established the monotonicity of the Cartan barycenter $\Lambda : \mathcal{P}^2 \to \mathbb{P}$,

$$\Lambda(\mu) \le \Lambda(\nu) \ \text{ for } \ \mu \le \nu. \tag{3.11}$$

This extends the same result for finite discrete measures, which is known as the monotonicity conjecture suggested by Bhatia and Holbrook [5] and settled by Lawson and Lim [20].

## 4. Main results

The main theorem of this paper is the following.

**Theorem 4.1.** *Let $\beta : \mathcal{P}^2 \to \mathbb{P}$ be an invariant barycenter on $\mathbb{P}$ satisfying*

$$\int_{\mathbb{P}} \log X \ d\mu(X) \le 0 \quad \textit{implies} \quad \beta(\mu) \le I \tag{4.12}$$

*for all $\mu \in \mathcal{P}^2$. Then $\beta = \Lambda$. Moreover, the Cartan barycenter satisfies (4.12).*

**Proof.** (1) To prove (4.12) for the Cartan barycenter $\Lambda$, we proceed with the following steps. Let $\mu \in \mathcal{P}^2$. Assume that $\int_{\mathbb{P}} \log(X) d\mu(X) \le 0$.

*Step* 1. By definition of Löwner ordering, there exists $A \ge I$ such that

$$\int_{\mathbb{P}} \log(X) d\mu_A(X) = \frac{1}{2}\int_{\mathbb{P}} \log(X) d\mu(X) + \frac{1}{2}\log(A) = 0, \tag{4.13}$$

where $\mu_A := \frac{1}{2}\mu + \frac{1}{2}\delta_A \in \mathcal{P}^2$. By the Karcher equation, $\Lambda(\mu_A) = I$.

*Step* 2. We consider the sequence $G_k$ on $\mathbb{P}$ defined inductively by

$$G_0 = \Lambda(\mu_I) \text{ for } \mu_I = \frac{1}{2}\mu + \frac{1}{2}\delta_I \ \text{ and } \quad G_{k+1} = \Lambda(\mu_{G_k}), \quad k = 0, 1, \ldots$$

It follows from (3.10) and $A \ge I$ that $\mu_A \ge \mu_I$. By monotonicity of the Cartan barycenter (3.11) and Step 1,

$$G_0 = \Lambda(\mu_I) \le \Lambda(\mu_A) = I.$$

This together with (3.10) leads to $\mu_I \ge \mu_{G_0}$ and again by the monotonicity (3.11)

$$G_1 = \Lambda(\mu_{G_0}) \le \Lambda(\mu_I) = G_0.$$

By induction $G_k$ is a decreasing sequence in $\mathbb{P}$ bounded above by $I$:

$$0 < G_k \le \cdots \le G_1 \le G_0 \le I.$$

*Step* 3. We will show that $\lim_{k\to\infty} G_k = \Lambda(\mu)$.

Let $Z = \Lambda(\mu)$. Then by Karcher equation (3.8), we have

$$\begin{aligned}0 &= \frac{1}{2}\int_{\mathbb{P}} \log(Z^{-\frac{1}{2}}XZ^{-1/2})d\mu(X) \\ &= \frac{1}{2}\int_{\mathbb{P}} \log(Z^{-\frac{1}{2}}XZ^{-\frac{1}{2}})d\mu(X) + \frac{1}{2}\log(Z^{-\frac{1}{2}}ZZ^{-\frac{1}{2}}) \\ &= \int_{\mathbb{P}} \log(Z^{-\frac{1}{2}}XZ^{-\frac{1}{2}})d\mu_Z(X)\end{aligned}$$

and therefore by the Karcher equation, $\Lambda(\mu_Z) = Z$. Moreover, by Sturm's fundamental contraction theorem (Theorem 2.5) and (2.6)

$$d(Z, G_k) = d(\Lambda(\mu_Z), \Lambda(\mu_{G_{k-1}})) \le d_1^W(\mu_Z, \mu_{G_{k-1}}) \le \frac{1}{2}d(Z, G_{k-1}).$$

By induction

$$d(Z, G_k) \le \frac{1}{2^k}d(Z, G_0)$$

for all $k$. This implies that $G_k$ converges to $Z = \Lambda(\mu)$ as $k \to \infty$.

*Step* 4. By preceding steps,

$$\Lambda(\mu) = \lim_{k\to\infty} G_k \le I.$$

This completes our first claim.

(2) Let $\beta : \mathcal{P}^2 \to \mathbb{P}$ be an invariant barycenter satisfying (4.12). Let $\mu \in \mathcal{P}^2$ such that $\int_{\mathbb{P}} \log X d\mu(X) \ge 0$. Then $\int_{\mathbb{P}} \log X^{-1} d\mu(X) \le 0$, and alternatively

$$\int_{\mathbb{P}} \log X d\nu(X) \le 0,$$

where $\nu := \mu^{-1}$. Since $\nu \in \mathcal{P}^2$, we have $\beta(\mu^{-1}) = \beta(\nu) \le I$ by (4.12). Or $\beta(\mu) \ge I$ by invariance of $\beta$. This shows that for $\mu \in \mathcal{P}^2$,

$$\int_{\mathbb{P}} \log X d\mu(X) = 0 \quad \text{implies} \quad \beta(\mu) = I. \tag{4.14}$$

Next, let $\mu \in \mathcal{P}^2$ and set $A = \Lambda(\mu)$. Then by the Karcher equation (3.8),

$$\int_{\mathbb{P}} \log(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})d\mu(X) = 0.$$

By change of variables theorem,

$$\int_{\mathbb{P}} \log X d\phi(X) = 0,$$

where $\phi := A^{-\frac{1}{2}}.\mu$. We then have $\beta(\phi) = I$ by (4.14). By the invariance of the Cartan barycenter,

$$\beta(\mu) = \beta(A^{\frac{1}{2}}.\phi) = A^{\frac{1}{2}}\beta(\phi)A^{\frac{1}{2}} = A = \Lambda(\mu).$$

Since $\mu$ varies arbitrary over $\mathcal{P}^2$, we have $\beta = \Lambda$. □

The inequality (4.12) extends the one by Yamazaki on the Cartan mean of finite discrete measures [32]. As applications of our main result, we establish several significant (norm) inequalities associated to the Cartan barycenter on $\mathcal{P}^2$.

We recall Hansen's inequality [13, Theorem 2.1], a direct consequence of Löwner–Heinz inequality: let $X$ be a $m \times m$ matrix with $XX^* \le I$. Then for any $A > 0$,

$$XA^pX^* \le (XAX^*)^p \quad \text{for all} \quad p \in [0, 1]$$

and

$$XA^pX^* \ge (XAX^*)^p \quad \text{for all} \quad p \in [1, 2].$$

The Ando–Hiai inequality [1] of positive definite matrices $A$ and $B$ is given by

$$A\#_t B \le I \quad \text{implies} \quad A^p\#_t B^p \le I$$

for any $t \in [0, 1]$ and for all $p \ge 1$. As a consequence of Theorem 4.1, we establish the Ando–Hiai inequality for the Cartan barycenter on $\mathcal{P}^2$.

**Theorem 4.2.** *Let $\mu \in \mathcal{P}^2$. Then $\Lambda(\mu) \le I$ implies $\Lambda(\mu^p) \le I$ for all $p \ge 1$.*

**Proof.** Let $\mu \in \mathcal{P}^2$. Assume that $Z := \Lambda(\mu) \le I$. Let $p \in [1, 2]$. Then

$$\begin{aligned} 0 &= p \int_{\mathbb{P}} \log(Z^{\frac{1}{2}}X^{-1}Z^{\frac{1}{2}})d\mu(X) \\ &= \int_{\mathbb{P}} \log(Z^{\frac{1}{2}}X^{-1}Z^{\frac{1}{2}})^p d\mu(X) \\ &\le \int_{\mathbb{P}} \log(Z^{\frac{1}{2}}X^{-p}Z^{\frac{1}{2}})d\mu(X), \end{aligned}$$

where the last inequality follows from the facts that log is operator monotone, $Z \le I$, and Hansen's inequality for $p \in [1, 2]$ holds. It is equivalent to

$$0 \ge \int_{\mathbb{P}} \log(Z^{-\frac{1}{2}}X^pZ^{-\frac{1}{2}})d\mu(X) = \int_{\mathbb{P}} \log X \; d\nu(X),$$

where $\nu := Z^{-\frac{1}{2}}.\mu^p$. Indeed, the last equality follows from change of variables theorem and

$$\nu(\mathcal{O}) = (Z^{-\frac{1}{2}}.\mu^p)(\mathcal{O}) = \mu((Z^{\frac{1}{2}}\mathcal{O}Z^{\frac{1}{2}})^{\frac{1}{p}}).$$

It then follows from Theorem 4.1 that

$$Z^{-\frac{1}{2}}\Lambda(\mu^p)Z^{-\frac{1}{2}} = \Lambda(\nu) \leq I.$$

That is,

$$\Lambda(\mu^p) \leq Z \leq I.$$

Repeating inductively this procedure to $\Lambda(\mu^{\frac{p}{2}}) \leq I$ for $p \in [2^k, 2^{k+1}]$ yields that $\Lambda(\mu^p) \leq I$ for all $p \geq 1$, which completes the proof. □

**Remark 4.3.** For $\mu \in \mathcal{P}^2$,

$$\int_{\mathbb{P}} \log X d\mu(X) \leq 0 \quad \text{implies} \quad \Lambda(\mu^p) \leq I, \forall p > 0.$$

The implication follows from the fact that $\int_{\mathbb{P}} \log X d\mu(X) \leq 0$ if and only if for any $p > 0$, $\int_{\mathbb{P}} \log X^p d\mu(X) = \int_{\mathbb{P}} \log X d\mu^p(X) \leq 0$, and from (4.12). It turns out [26] that the reverse implication holds true for finite discrete measures $\mu$;

$$\int_{\mathbb{P}} \log X d\mu(X) \leq 0 \quad \text{if and only if} \quad \Lambda(\mu^p) \leq I, \forall p > 0.$$

A special consequence of Theorem 4.1 is the following operator norm inequality.

**Corollary 4.4.** *Let $\mu \in \mathcal{P}^2$, and let $0 < p \leq q$. Then*

$$\Lambda(\mu^p) \leq I \quad \textit{implies} \quad \Lambda(\mu^q) \leq I. \tag{4.15}$$

*Moreover for the operator norm $||\cdot||$,*

$$||\Lambda(\mu^q)^{\frac{1}{q}}|| \leq ||\Lambda(\mu^p)^{\frac{1}{p}}||. \tag{4.16}$$

**Proof.** Let $\mu \in \mathcal{P}^2$ and let $0 < p \leq q$. Suppose that $\Lambda(\mu^p) \leq I$. Set $\nu := \mu^p$. Applying Theorem 4.2 with $\nu$ and $q/p \geq 1$ we have

$$\Lambda(\mu^q) = \Lambda(\nu^{\frac{q}{p}}) \leq I.$$

Next, we shall prove (4.16). Set $\alpha := ||\Lambda(\mu^p)^{\frac{1}{p}}||^{\frac{p}{2}}$. Then

$$\Lambda(\alpha^{-1}.\mu^p)^{\frac{1}{p}} = \left[\alpha^{-1}.\Lambda(\mu^p)\right]^{\frac{1}{p}} = \frac{1}{\alpha^{\frac{2}{p}}}\Lambda(\mu^p)^{\frac{1}{p}} = \frac{1}{||\Lambda(\mu^p)^{\frac{1}{p}}||}\Lambda(\mu^p)^{\frac{1}{p}} \leq I$$

from invariance of the Cartan barycenter and from the definition of the operator norm and positive definiteness of $\Lambda(\mu^p)$. This implies that $\Lambda(\alpha^{-1}.\mu^p) \leq I$, since $A \leq I$ implies $A^t \leq I$ for any $t \geq 0$ and $A \in \mathbb{P}$. We let $\nu := \alpha^{-1}.\mu^p$. Then $\nu^{\frac{q}{p}} = \alpha^{-\frac{q}{p}}.\mu^q$ because

$$
\begin{aligned}
(\alpha^{-1}.\mu^p)^{\frac{q}{p}}(\mathcal{O}) &= (\alpha^{-1}.\mu^p)(\mathcal{O}^{\frac{p}{q}}) = \mu^p(\alpha^2\mathcal{O}^{\frac{p}{q}}) = \mu([\alpha^2\mathcal{O}^{\frac{p}{q}}]^{\frac{1}{p}}) \\
&= \mu(\alpha^{\frac{2}{p}}\mathcal{O}^{\frac{1}{q}}) = \mu([\alpha^{\frac{2q}{p}}]^{\frac{1}{q}}\mathcal{O}^{\frac{1}{q}}) = \mu^q(\alpha^{\frac{2q}{p}}\mathcal{O}) \\
&= (\alpha^{-\frac{q}{p}}.\mu^q)(\mathcal{O}).
\end{aligned}
$$

Applying Theorem 4.2 with $\nu$ and $q/p \geq 1$ yields $\Lambda(\nu^{\frac{q}{p}}) \leq I$ and thus

$$
\alpha^{-\frac{2q}{p}}\Lambda(\mu^q) = \alpha^{-\frac{q}{p}}.\Lambda(\mu^q) = \Lambda(\alpha^{-\frac{q}{p}}.\mu^q) = \Lambda(\nu^{\frac{q}{p}}) \leq I.
$$

This implies that $\left[\alpha^{-\frac{2q}{p}}\Lambda(\mu^q)\right]^{\frac{1}{q}} \leq I$, and hence,

$$
||\Lambda(\mu^q)^{\frac{1}{q}}|| \leq \alpha^{\frac{2}{p}} = ||\Lambda(\mu^p)^{\frac{1}{p}}||.
$$

This establishes (4.16). $\square$

Let $A$ be an $m\times m$ positive semidefinite matrix with eigenvalues $\lambda_j(A)$ $j = 1, \dots, m$ arranged in decreasing order, i.e., $\lambda_1(A) \geq \cdots \geq \lambda_m(A)$. For $A, B \geq 0$, we define $A \underset{(\log)}{\prec} B$ if

$$
\prod_{i=1}^{k}\lambda_i(A) \leq \prod_{i=1}^{k}\lambda_i(B) \quad \text{for } k = 1, 2, \dots, m-1, \text{ and } \det A = \det B.
$$

This relation is called log-majorization. It is well known that $A \underset{(\log)}{\prec} B$ implies $|||A||| \leq |||B|||$ for all unitarily invariant norms $|||\cdot|||$.

For $1 \leq k \leq m$, let $\Gamma^k$ be the $k$-th asymmetric tensor power (see [1,6] for basic properties of $\Gamma^k$). Then

$$
\begin{aligned}
\Lambda(\omega; \Gamma^k A_1, \dots, \Gamma^k A_n) &= \Gamma^k\Lambda(\omega; A_1, \dots, A_n), \\
\lambda_1(\Lambda^k A) &= \prod_{i=1}^{k}\lambda_i(A), A > 0, \\
\Gamma^k(A^p) &= (\Gamma^k A)^p, \quad p > 0, A > 0.
\end{aligned}
$$

Using these properties, one can see (cf. [14]) that for $\mu = \frac{1}{n}\sum_{j=1}^{n}\delta_{A_j} \in \mathcal{P}^0$,

$$
\Lambda(\mu^q)^{\frac{1}{q}} \underset{(\log)}{\prec} \Lambda(\mu^p)^{\frac{1}{p}}, \qquad 0 < p \leq q, \tag{4.17}
$$

$$
\Lambda(\mu^p)^{\frac{1}{p}} \underset{(\log)}{\prec} \Lambda(\mu) \underset{(\log)}{\prec} \Lambda(\mu^{\frac{1}{p}})^p, \qquad p \geq 1 \tag{4.18}
$$

and therefore

$$
|||\Lambda(\mu^q)^{\frac{1}{q}}||| \leq |||\Lambda(\mu^p)^{\frac{1}{p}}|||, \qquad 0 < p \leq q, \tag{4.19}
$$

$$
|||\Lambda(\mu^p)^{\frac{1}{p}}||| \leq |||\Lambda(\mu)||| \leq |||\Lambda(\mu^{\frac{1}{p}})^p|||, \qquad p \geq 1 \tag{4.20}
$$

for all unitarily invariant norms $|||\cdot|||$. Moreover, using the fact that

$$
\lim_{p\to 0^+}\Lambda(\mu^p)^{\frac{1}{p}} = \exp((1/n)\sum_{j=1}^{n}\log A_j),
$$

we obtain

$$\lim_{p\to 1^-} |||\Lambda(\omega;\mu^{\frac{1}{p}})^p||| = |||\Lambda(\mu)||| \leq \lim_{p\to 0+} |||\Lambda(\mu^p)|||^{\frac{1}{p}} = |||e^{(1/n)\sum_{j=1}^n \log A_j}|||, \tag{4.21}$$

which is a multivariate version of complementary Golden–Thompson inequality and settles a question of Bhatia and Grover [4]. We note that (4.19) extends (4.16) for the case of discrete measures but remains open for general $\mu \in \mathcal{P}^2$. An appropriate version of (4.21) in the setting of $\mu \in \mathbb{P}^2$ is

$$|||\Lambda(\mu)||| \leq \lim_{p\to 0+} |||\Lambda(\mu^p)|||^{\frac{1}{p}} = |||e^{\int_{\mathbb{P}} \log X d\mu(X)}|||. \tag{4.22}$$

We believe that our main tools on $\mu \in \mathcal{P}^2$ involving the theory of nonpositively curved metric spaces and techniques from probability [30] and partially ordered Hadamard manifolds will significantly enhance the potential usefulness of the Cartan barycenter on $\mathcal{P}^2$ in the further work on majorizations and unitarily invariant norm inequalities for probability measures, like (4.22).

**Remark 4.5.** The results in this section still hold in the infinite-dimensional case provided one restricts to Hilbert–Schmidt operators, since the underlying space is still a partially ordered Hadamard space. See [21,22].

## 5. Cartan barycenters on symmetric cones

In this section, we shall see that the techniques and results from the probabilistic treatment of the Cartan barycenter for positive definite matrices carry over, typically with little change, to the case of symmetric cones. We first briefly describe (following mostly [11]) some Jordan-algebraic concepts pertinent to our purpose. A *Jordan algebra* $V$ over $\mathbb{R}$ is a finite-dimensional commutative algebra with identity $e$ satisfying $x^2(xy) = x(x^2y)$ for all $x, y \in V$. For $x \in V$, let $L(x)$ be the linear operator defined by $L(x)y = xy$, and let $P(x) = 2L(x)^2 - L(x^2)$. The map $P$ is called the quadratic representation of $V$. An element $x \in V$ is said to be invertible if there exists an element $x^{-1}$ in the subalgebra generated by $x$ and $e$ such that $xx^{-1} = e$. A useful property of Jordan algebras is power associative, that is, the subalgebra generated by $x$ is associative.

An element $c \in V$ is called an idempotent if $c^2 = c$. We say that $c_1, \ldots, c_k$ is a complete system of orthogonal idempotents if $c_i^2 = c_i$, $c_ic_j = 0$, $i \neq j$, $c_1 + \cdots + c_k = e$. An idempotent is primitive if it is non-zero and cannot be written as the sum of two non-zero idempotents. A Jordan frame is a complete system of primitive idempotents.

A Jordan algebra $V$ is said to be *Euclidean* if there exists an inner product $\langle \cdot, \cdot \rangle$ such that for all $x, y, z \in V$:

$$\langle xy, z\rangle = \langle y, xz\rangle. \tag{5.23}$$

The following spectral theorem for Euclidean Jordan algebras appears in [11].

**Theorem 5.1.** *Any two Jordan frames in an Euclidean Jordan algebra $V$ have the same number of elements (called the rank of $V$, denoted* rank$(V)$*). Given $x \in V$, there exists a Jordan frame $c_1, \ldots, c_r$ and real numbers $\lambda_1, \ldots, \lambda_r$ such that*

$$x = \sum_{i=1}^r \lambda_i c_i.$$

**Definition 5.2.** Let $V$ be a Euclidean Jordan algebra of rank$(V) = r$. The spectral mapping $\lambda : V \to \mathbb{R}^r$ is defined by $\lambda(x) = (\lambda_1(x), \ldots, \lambda_r(x))$, where the $\lambda_i(x)$'s are eigenvalues of $x$ (with multiplicities) as in Theorem 5.1 in non-increasing order $\lambda_{\max}(x) = \lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_r(x) = \lambda_{\min}(x)$. We define $\det(x) = \prod_{i=1}^{r} \lambda_i(x)$ and $\mathrm{tr}(x) = \sum_{i=1}^{r} \lambda_i(x)$. Then tr is a linear form on $V$ and det is a homogeneous polynomial of degree $r$ on $V$.

The trace inner product $\langle x, y\rangle = \mathrm{tr}(xy)$ in a Euclidean Jordan algebra satisfies (5.23). We will assume that $V$ is a Euclidean Jordan algebra of rank $r$ and equipped with the trace inner product $\langle x, y\rangle = \mathrm{tr}(xy)$. Let $Q$ be the set of all square elements of $V$. Then $Q$ is a closed convex cone of $V$ with $Q \cap -Q = \{0\}$, and is the set of elements $x \in V$ such that $L(x)$ is positive semi-definite. It turns out that $Q$ has non-empty interior $\Omega$, and $\Omega$ is a symmetric cone, that is, the group $G(\Omega) = \{g \in \mathrm{GL}(V) | g(\Omega) = \Omega\}$ acts transitively on it and $\Omega$ is a self-dual cone with respect to the inner product $\langle\cdot|\cdot\rangle$. Furthermore, for any $a$ in $\Omega$, $P(a) \in G(\Omega)$ and is positive definite. We note that any symmetric cone (self-dual, homogeneous open convex cone) can be realized as an interior of squares in an appropriate Euclidean Jordan algebra [11].

**Proposition 5.3.** *The symmetric cone $\Omega \subseteq V$ has the following properties:*

$$\Omega = \{x^2 : x \textit{ is invertible}\} = \{x : L(x) \textit{ is positive definite}\} = \{x : \lambda_{\min}(x) > 0\}.$$

We further note that the symmetric cone can be obtained as $\Omega = \exp(V) := \{\exp(x) : x \in V\}$, where $\exp(x) = \sum_{k=1}^{\infty} \frac{x^n}{n!}$, and the exponential map $\exp : V \to \Omega$ is bijective. The logarithm map $\log : \Omega \to V$ is defined as the inverse of the exponential map.

The space $\mathbb{H}_m$ of $m \times m$ Hermitian matrices equipped with the trace inner product $\langle X, Y\rangle = \mathrm{tr}(X^*Y)$ and the Jordan product $X \circ Y = \frac{1}{2}(XY + YX)$ is a typical example of Euclidean Jordan algebras. In this case the corresponding symmetric cone is $\mathbb{P}_m$, the convex cone of $m \times m$ positive definite Hermitian matrices, and the quadratic representation is given by $P(X)Y = XYX$.

It turns out [11] that the symmetric cone $\Omega$ admits a $G(\Omega)$-invariant Riemannian metric defined by $\langle u, v\rangle_a = \langle P(a)^{-1}u, v\rangle$, $a \in \Omega$, $u, v \in V$. The inversion $j(x) = x^{-1}$ is an involutive isometry fixing $e$. It is a symmetric Riemannian space of non-compact type and hence is an NPC space with respect to its distance metric. The unique geodesic curve joining $a$ and $b$ is $t \mapsto a\#_t b := P(a^{1/2})(P(a^{-1/2})b)^t$ and the Riemannian distance $d(a, b)$ is given by $d(a, b) = \left(\sum_{i=1}^{r} \log^2 \lambda_i(P(a^{-1/2})b)\right)^{1/2}$. See [18,23,24] for more details. The geodesic middle (geometric mean) of $a$ and $b$ is given by $a\#b := a\#_{1/2}b = P(a^{1/2})(P(a^{-1/2})b)^{1/2}$. In [24], it is shown that the geometric mean is monotone for the cone ordering, $x \leq y$ if and only if $y - x \in \overline{\Omega}$, and therefore we conclude that every symmetric cone is a Loewner–Heinz NPC space. See below for its definition and the monotonic property of the Cartan barycenter of finite discrete measures [20].

**Definition 5.4.** A *Loewner–Heinz NPC space* is an NPC space equipped with a closed partial order $\leq$ satisfying $x_1\#x_2 \leq y_1\#y_2$ whenever $x_i \leq y_i$ for $i = 1, 2$. Here $x\#y$ denotes the unique midpoint between $x$ and $y$.

**Theorem 5.5.** *Let $(M, d, \leq)$ be a Loewner–Heinz NPC space. Then the Cartan barycenter is monotonic on $\mathcal{P}_0$. That is, for $\mu = \frac{1}{n}\sum_{j=1}^{n} \delta_{x_j}$ and $\nu = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j}$ with $x_j \leq y_j$ for all $j$, $\Lambda(\mu) \leq \Lambda(\nu)$.*

Next, we consider the monotonicity of the Cartan barycenter on $\mathcal{P}^2$ for a symmetric cone $\Omega$. Its monotonic property for the case of positive definite matrices is central for Theorem 4.1.

Let $\mu, \nu \in \mathcal{P}^2$ with $\mu \leq \nu$ and let

$$\mu_n := \mu|_{((1/n)e, ne]} + \mu\left(\Omega \setminus (\frac{1}{n}e, ne]\right) \delta_{\{\frac{1}{n}e\}}, \tag{5.24}$$

$$\nu_n := \nu|_{((1/n)e, ne]} + \nu\left(\Omega \setminus (\frac{1}{n}e, ne]\right) \delta_{\{\frac{1}{n}e\}}, \tag{5.25}$$

where $(x, y] = \{z \in V : x < z \leq y\}$, the Löwner interval determined by $x$ and $y$. Then taking the methods for the cone of positive definite matrices [17], we can show that $\mu_n$ and $\nu_n$ have bounded (and hence compact) supports and satisfy $\mu_n \leq \nu_n$ and $\Lambda(\mu_n) \leq \Lambda(\nu_n)$ for all $n$, and the sequences $\mu_n$ and $\nu_n$ converge to $\mu$ and $\nu$ respectively for the Wasserstein distance $d_2^W$. Then by Sturm's contractive property for $d_2^W$,

$$d(\Lambda(\mu), \Lambda(\mu_n)) \leq d_2^W(\mu, \mu_n) \to 0,$$
$$d(\Lambda(\nu), \Lambda(\nu_n)) \leq d_2^W(\nu, \nu_n) \to 0$$

and therefore

$$\Lambda(\mu) = \lim_{n\to\infty} \Lambda(\mu_n) \leq \lim_{n\to\infty} \Lambda(\nu_n) = \Lambda(\nu).$$

This establishes the monotonicity of the Karcher barycenter on $\mathcal{P}^2$.

The steps in the proof of Theorem 4.1 carry directly over on the setting of symmetric cones:

**Theorem 5.6.** *The Cartan barycenteric map $\Lambda : \mathcal{P}^2(\Omega) \to \Omega$ on the symmetric cone is monotonic and satisfies*

$$\int_{\Omega} \log x \, d\mu(x) \leq 0 \quad \text{implies} \quad \Lambda(\mu) \leq e \tag{5.26}$$

*for all $\mu \in \mathcal{P}^2(\Omega)$. Furthermore, if $\beta$ is an invariant metric for the Riemannian metric $(\Omega, d)$ satisfying (5.26), then $\beta = \Lambda$.*

One can adapt methods in the proof of Theorem 4.2 to derive the Ando–Hiai inequality (and hence Corollary 4.4) on symmetric cones: for $\mu \in \mathcal{P}^2(\Omega)$, $\Lambda(\mu) \leq I$ implies $\Lambda(\mu^p) \leq I$ for all $p \geq 1$.

## 6. Final remarks and acknowledgments

Although the order inequality characterizing invariant barycenters is new and quite attractive, particularly in the theory of matrix analysis, analysis on symmetric cones and probability measures, it depends heavily on the Karcher equation (also, monotonicity and Sturm's fundamental contraction theorem for the Cartan barycenter) and it does not carry over in the $\mathcal{P}^1$-setting. The Cartan barycener on $\mathcal{P}^1$ is also an important object in related research areas and so finding a (order inequality) characterizing property still remains open in our context. We close with the following open problem.

**Problem 1.** Does the Cartan barycenter $\Lambda$ on $\mathcal{P}^1$ satisfy the monotonic property, (5.26) and the Ando–Hiai inequality?

The authors wish to express their gratitude to Professor Jimmie Lawson for helpful insights about the ordered structures and the Löwner–Heinz inequality on $\mathcal{P}^2$. The work of S. Kim was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (NRF-2015R1C1A1A02036407). The work of H. Lee was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (NRF-2015R1D1A1A01059900). The work of Y. Lim was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2015R1A3A2031159).

## References


[1] T. Ando, F. Hiai, Log majorization and complementary Golden–Thompson type inequalities, Linear Algebra Appl. 197 (1994) 113–131.
[2] G. Basso, Fixed point theorems for metric spaces with a conical geodesic bicombing, preprint.
[3] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2007.
[4] R. Bhatia, P. Grover, Norm inequalities related to the matrix geometric mean, Linear Algebra Appl. 437 (2012) 726–733.
[5] R. Bhatia, J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 413 (2006) 594–618.
[6] R. Bhatia, R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. 353 (2012) 1453–1467.
[7] J. Bochi, A. Navas, A geometric path from zero Lyapunov exponents to rotation cocycles, Ergodic Theory Dynam. Systems 35 (2015) 374–402.
[8] F. Bolley, Separability and completeness for the Wasserstein distance, in: Séminaire de Probabilités XLI, in: Lecture Notes in Math., vol. 1934, 2008, pp. 371–377.
[9] T. Champion, L. De Pascale, R. Juutinen, The $\infty$-Wasserstein distance: local solutions and existence of optimal transport maps, SIAM J. Math. Anal. 40 (2008) 1–20.
[10] A. Es-Sahib, H. Heinich, Barycentre canonique pour un espace métrique à courbure négative, in: Séminaire de Probabilités, XXXIII, in: Lecture Notes in Math., vol. 1709, Springer, Berlin, 1999, pp. 355–370.
[11] J. Faraut, A. Korányi, Analysis on Symmetric Cones, The Clarendon Press, Oxford, 1994.
[12] W. Gangbo, R.J. McCann, The geometry of optimal transportation, Acta Math. 177 (1996) 113–161.
[13] F. Hansen, G.K. Pedersen, Jensen's inequality for operators and Löwner's theorem, Math. Ann. 258 (1982) 229–241.
[14] F. Hiai, D. Petz, Riemannian metrics on positive definite matrices related to means II, Linear Algebra Appl. 436 (2012) 2117–2136.
[15] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. 30 (1977) 509–541.
[16] Y. Kim, B. Pass, Multi-marginal optimal transport on Riemannian manifolds, Amer. J. Math. 137 (2015) 1045–1060.
[17] S. Kim, H. Lee, The power mean and the least squares mean of probability measures on the space of positive definite matrices, Linear Algebra Appl. 465 (2015) 325–346.
[18] S. Lang, Fundamentals of Differential Geometry, Grad. Texts in Math., Springer-Verlag, 1999.
[19] J. Lawson, H. Lee, Y. Lim, Weighted geometric means, Forum Math. 24 (2012) 1067–1090.
[20] J. Lawson, Y. Lim, Monotonic properties of the least squares mean, Math. Ann. 351 (2011) 267–279.
[21] J. Lawson, Y. Lim, Weighted means and Karcher equations of positive operators, Proc. Natl. Acad. Sci. USA 110 (2013) 15626–15632.
[22] J. Lawson, Y. Lim, Karcher means and Karcher equations of positive definite operators, Trans. Amer. Math. Soc. Series B 1 (2014) 1–22.
[23] Y. Lim, Finsler metrics on symmetric cones, Math. Ann. 316 (2000) 379–389.
[24] Y. Lim, Applications of geometric means on symmetric cones, Math. Ann. 319 (2001) 457–468.
[25] Y. Lim, M. Pálfia, Matrix power means and the Karcher mean, J. Funct. Anal. 262 (2012) 1498–1514.
[26] Y. Lim, T. Yamazaki, On some inequalities for the matrix power and Karcher means, Linear Algebra Appl. 438 (2013) 1293–1304.
[27] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. 26 (2005) 735–747.
[28] A. Navas, An $L^1$ ergodic theorem with values in a non-positively curved space via a canonical barycenter map, Ergodic Theory Dynam. Systems 33 (2013) 609–623.
[29] X. Pennec, Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements, J. Math. Imaging Vision 25 (2006) 127–154.
[30] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces, Paris, 2002, in: Contemp. Math., vol. 338, Amer. Math. Soc., Providence, RI, 2003, pp. 357–390.
[31] C. Villani, Topics in Optimal Transportation, Grad. Stud. Math., vol. 58, American Mathematical Society, Providence, RI, 2003.
[32] T. Yamazaki, The Riemannian mean and matrix inequalities related to the Ando–Hiai inequality and chaotic order, Oper. Matrices 6 (2012) 577–588.