Lin. Alg. Appl. **430**(2009), 3105–3130.

Riemannian metrics on positive definite matrices related to means

Fumio Hiai[1] and Dénes Petz[2]

[1] Graduate School of Information Sciences, Tohoku University
Aoba-ku, Sendai 980-8579, Japan

[2] Alfréd Rényi Institute of Mathematics,
H-1364 Budapest, POB 127, Hungary

Abstract

The Riemannian metric on the manifold of positive definite matrices is defined by a kernel function ϕ in the form $K_D^\phi(H,K)=\sum_{i,j}\phi(\lambda_i,\lambda_j)^{-1}\mathrm{Tr}\,P_iHP_jK$ when $\sum_i \lambda_i P_i$ is the spectral decomposition of the foot point D and the Hermitian matrices H,K are tangent vectors. For such kernel metrics the tangent space has an orthogonal decomposition. The pull-back of a kernel metric under a mapping $D\mapsto G(D)$ is a kernel metric as well. Several Riemannian geometries of the literature are particular cases, for example, the statistical metric for multivariate Gaussian distributions and the quantum Fisher information. In the paper the case $\phi(x,y)=M(x,y)^\theta$ is mostly studied when $M(x,y)$ is a mean of the positive numbers x and y. There are results about the geodesic curves and geodesic distances. The geometric mean, the logarithmic mean and the root mean are important cases.

AMS classification: 15A45; 15A48; 53B21; 53C22

Keywords: positive definite matrix; Riemannian metric; congruence-invariant metric; quantum Fisher information; quantum skew information; symmetric homogeneous mean; logarithmic mean; geometric mean; geodesic curve; geodesic distance; Fréchet derivative; divided difference

[1]E-mail: hiai@math.is.tohoku.ac.jp
[2]E-mail: petz@math.bme.hu

Introduction

The $n \times n$ positive definite matrices $\mathbb{P}_n$ with complex entries can be parametrized by the real and imaginary parts of the entries, and they form an open subset of the space $\mathbb{H}_n$ of $n \times n$ Hermitian matrices regarded as the Euclidean space $\mathbb{R}^m$, where $m = n^2$. Hence the tangent space of their manifold $\mathbb{P}_n$ at any foot point can be identified with $\mathbb{H}_n$. A Riemannian metric $K_D(H, K)$ is a family of inner products on $\mathbb{H}_n$ depending smoothly on the foot point D. If $\phi(x, y)$ is a positive kernel function on $(0, \infty) \times (0, \infty)$ and D has the spectral decomposition $\sum_{i=1}^k \lambda_i P_i$, then a Riemannian metric can be defined as

$$K_D^\phi(H, K) := \sum_{i,j=1}^k \phi(\lambda_i, \lambda_j)^{-1} \mathrm{Tr}\, P_i H P_j K, \tag{0.1}$$

where Tr is the usual trace on matrices. The goal of the present paper is to study this kind of Riemannian metrics.

As far as the authors know, the first example of (0.1) is historically the case $\phi(x, y) = xy$, which was considered by Skovgaard [41] as a statistical Riemannian metric on positive definite matrices describing multivariate Gaussian distributions. The geodesic properties of this Riemannian metric was formally described in [30, §3]. Another example is related to Fisher information. In the quantum mechanical setting the states correspond to positive semidefinite matrices of trace 1, and in [35, 39] the metric (0.1) was justified in the particular case $\phi(x, y) = y f(x/y)$, where $f : (0, \infty) \to (0, \infty)$ is an operator monotone function. More details on these examples are presented in the rest of this section.

The trivial choice $\phi(x, y) \equiv 1$ gives a flat space where the Riemannian metric is the Hilbert-Schmidt inner product $\langle H, K \rangle_{\mathrm{HS}}$ on $\mathbb{H}_n$. The Hilbert-Schmidt inner product $\langle X, Y \rangle_{\mathrm{HS}} := \mathrm{Tr}\, X^* Y$ and the Hilbert-Schmidt norm $\|X\|_{\mathrm{HS}} := (\mathrm{Tr}\, X^* X)^{1/2}$ are defined on the space $\mathbb{M}_n$ of all $n \times n$ complex matrices, and the space $(\mathbb{H}_n, \langle \cdot, \cdot \rangle_{\mathrm{HS}})$ is a real subspace of the Hilbert space $(\mathbb{M}_n, \langle \cdot, \cdot \rangle_{\mathrm{HS}})$.

The positive definite real matrices might be considered as the variance of multivariate normal distributions and the information geometry of Gaussians yields a natural Riemannian metric. The simplest way to construct an information geometry is to start with an information potential function and to introduce the Riemannian metric by the Hessian of the potential. We want a geometry on the family of non-degenerate multivariate Gaussian distributions with zero mean vector. Those distributions are given by a positive definite real matrix D in the form

$$p_D(x) := \frac{1}{\sqrt{(2\pi)^n \det D}} \exp\left(-\frac{\langle D^{-1} x, x \rangle}{2} \right), \qquad x \in \mathbb{R}^n.$$

We identify the Gaussian p_D with the matrix D, and we can say that the Riemannian geometry is constructed on the space of positive definite real matrices. There are many reasons (originated from statistical mechanics, information theory and mathematical statistics) that the Boltzmann entropy

$$S(p_D) := \frac{1}{2} \log(\det D) + \text{const.}$$

is a candidate for being an information potential.

The $n \times n$ real symmetric matrices can be identified with the Euclidean space of dimension $n(n+1)/2$ and the positive definite real matrices form an open subset. Therefore the set of Gaussians has a simple and natural manifold structure. The tangent space at each foot point is the set of real symmetric matrices. The Riemannian metric is defined as the Hessian

$$g_D(H,K) := \frac{\partial^2}{\partial s \partial t} S(p_{D+sH+tK})\Big|_{s=t=0},$$

where H and K are tangents at D. The differentiation easily gives

$$g_D(H,K) = \mathrm{Tr}\, D^{-1} H D^{-1} K. \tag{0.2}$$

In the statistical model of multivariate Gaussian distributions, (0.2) plays the role of a natural Riemannian metric and the corresponding information geometry of the Gaussians was discussed in [32] in detail. We note here that this geometry has many symmetries. Each congruence transformation of the matrices becomes a symmetry, namely

$$g_{XDX^t}(XHX^t, XKX^t) = g_D(H,K) \tag{0.3}$$

for every real invertible matrix X.

Formula (0.2) determines a Riemannian metric on the manifold $\mathbb{P}_n$ of positive definite complex matrices as well, and below we call this metric extended on $\mathbb{P}_n$ the *congruence-invariant metric* since it is invariant under congruence $X \cdot X^*$ for any invertible $X \in \mathbb{M}_n$. Note that if we want to find the geodesic curve between A and B, then it is sufficient to find the geodesic joining I and $A^{-1/2}BA^{-1/2}$ due to property (0.3). This is essentially easier since they commute. In fact, concerning the geodesic curves in the Riemannian manifold $(\mathbb{P}_n, g)$, it is known [26, 29, 9] that for each $A, B \in \mathbb{P}_n$ there exists a unique geodesic shortest curve joining $A, B \in \mathbb{P}_n$ given by

$$\gamma(t) = A \,\#_t\, B := A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}, \qquad 0 \le t \le 1, \tag{0.4}$$

and the geodesic midpoint $\gamma(1/2)$ is just the *geometric mean* ([40, 1])

$$A \,\#\, B := A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2} A^{1/2}.$$

Furthermore, the geodesic distance is

$$\delta(A,B) = \| \log(A^{-1/2}BA^{-1/2}) \|_{\mathrm{HS}}. \tag{0.5}$$

In this way, the information Riemannian geometry is adequate to treat the geometric mean of positive definite matrices.

For each $A, B \in \mathbb{P}_n$ the mean $C' := A \,\#\, B$ is the midpoint of the geodesic joining A and B, and similarly $A' := B \,\#\, C$ and $B' := C \,\#\, A$ are taken. Since $\delta(B \,\#\, C, C \,\#\, A) \le \frac{1}{2}\delta(A,B)$ by [9, Proposition 6], the diameter of the triangle $A'B'C'$ is at most the half of the diameter of ABC. This result gives a geometric proof of the recursive construction of geometric mean of 3 positive matrices proposed in [3]. Note that another "geometric mean" of $A_1, \dots, A_k \in \mathbb{P}_n$ was introduced in [29, 9] as the unique minimizer of $A \in \mathbb{P}_n \mapsto \sum_{j=1}^k \delta^2(A, A_j)$.

We denote by $\mathcal{D}_n$ the set of all $n \times n$ positive definite matrices of trace 1, which is a smooth differentiable manifold as a submanifold of $\mathbb{P}_n$. The tangent space of the manifold $\mathcal{D}_n$ at each foot point D is the subspace of $\mathbb{H}_n$ consisting of $n \times n$ Hermitian matrices of trace 0, i.e.,

$$T_D\mathcal{D}_n = \mathbb{H}_n \ominus \mathbb{R}I := \{H \in \mathbb{H}_n : \mathrm{Tr}\, H = 0\}.$$

One can define a Riemannian metric on $\mathcal{D}_n$ in the form

$$K_D(H, K) = \langle H, \mathbb{J}_D^{-1}K\rangle_{\mathrm{HS}}, \qquad D \in \mathcal{D}_n,\ H, K \in \mathbb{H}_n \ominus \mathbb{R}I,$$

where $\mathbb{J}_D$ is a positive linear operator on the real Hilbert space $(\mathbb{H}_n \ominus \mathbb{R}I, \langle \cdot, \cdot \rangle_{\mathrm{HS}})$. One can extend $\mathbb{J}_D$ to a positive symmetric operator on $\mathbb{H}_n$ and furthermore to a positive operator on the Hilbert space $(\mathbb{M}_n, \langle \cdot, \cdot \rangle_{\mathrm{HS}})$ by complexification. So we may assume that a Riemannian metric K_D is given on $\mathcal{D}_n$, $n \in \mathbb{N}$, by $K_D(X, Y) = \langle X, \mathbb{J}_D^{-1}Y\rangle_{\mathrm{HS}}$ for $X, Y \in \mathbb{M}_n$. The metric K_D (more precisely, a sequence of metrics K_D on $\mathcal{D}_n$, $n \in \mathbb{N}$) is *monotone* if for any completely positive and trace preserving map (or coarse graining [37, Chapter 9]) $\beta : \mathbb{M}_n \to \mathbb{M}_m$ we have

$$K_{\beta(D)}(\beta(X), \beta(X)) \le K_D(X, X), \qquad D \in \mathcal{D}_n,\ X \in \mathbb{M}_n.$$

Recall that β is completely positive and trace preserving if and only if β^* is completely positive and unital. The Chentsov theorem (see [35]) says that in the commutative case (or when restricted on the diagonal positive matrices) there exists a unique monotone metric (up to a scalar factor) that is the so-called *Fisher-Rao metric*. The situation is quite different in the non-commutative case, and it was proved in Petz [35] that the monotone metrics K_D with normalization $K_D(I, I) = \mathrm{Tr}\,(D^{-1})$ correspond one-to-one to the operator monotone functions $f : (0, \infty) \to (0, \infty)$ with normalization $f(1) = 1$ as follows:

$$K_D^f(X, Y) := \langle X, (\mathbb{J}_D^f)^{-1}Y\rangle_{\mathrm{HS}} \quad \text{and} \quad \mathbb{J}_D^f := f(\mathbb{L}_D\mathbb{R}_D^{-1})\mathbb{R}_D. \tag{0.6}$$

Furthermore, K_D^f is symmetric if and only if f is symmetric, i.e., $xf(x^{-1}) = f(x)$, $x > 0$. We say that an operator monotone function $f \ge 0$ on $(0, \infty)$ is *standard* if $f(1) = 1$ and $xf(x^{-1}) = f(x)$.

On the other hand, the theory of *operator means* due to Kubo and Ando [25] says that there is a one-to-one correspondence between the symmetric operator means (or matrix means) and the standard operator monotone functions f as follows:

$$\sigma_f(A, B) := A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}, \qquad A, B \in \mathbb{P}_n.$$

Thus one may write

$$K_D^f(X, Y) = \langle X, \sigma_f(\mathbb{L}_D, \mathbb{R}_D)^{-1}Y\rangle_{\mathrm{HS}}. \tag{0.7}$$

When $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ is diagonal, one can more explicitly write

$$K_D^f(X, X) = \sum_{i,j=1}^n \frac{1}{\lambda_j f(\lambda_i/\lambda_j)} |X_{ij}|^2, \qquad X = [X_{ij}] \in \mathbb{M}_n.$$

For each standard operator monotone function f, the symmetric monotone metric (or the *quantum Fisher information*) K_D^f originally defined on $\mathcal{D}_n$ by (0.6) or (0.7) can be automatically extended to $\mathbb{P}_n$ by the same formula.

It was also observed in Lesniewski and Ruskai [27] that any of the above metrics K^f can be realized as the Hessian

$$K_D^f(H, K) = -\frac{\partial^2}{\partial s \partial t} S_F(D + sH, D + tK)\Big|_{s=t=0},$$

of the quasi-entropy ([33, 34])

$$S_F(D_1, D_2) := \langle D_1^{1/2}, F(\mathbb{L}_{D_2}\mathbb{R}_{D_1}^{-1}) D_1^{1/2} \rangle_{\mathrm{HS}}$$

defined by a function F on $(0, \infty)$ with the relation $1/f(x) = (F(x) + xF(x^{-1}))/(x-1)^2$.

The *Wigner-Yanase-Dyson skew information* is the quantity

$$I_D^{\mathrm{WYD}}(p, K) := -\frac{1}{2} \mathrm{Tr}\,[D^p, K][D^{1-p}, K], \qquad D \in \mathcal{D}_n,\ K \in \mathbb{H}_n,$$

where $0 < p < 1$. The case $p = 1/2$ is the original Wigner-Yanase skew information. It was observed in [38] that the Wigner-Yanase-Dyson skew information $I_D^{\mathrm{WYD}}(p, K)$ coincides, apart from a constant factor, with a monotone Riemannian metric

$$K_D^{f_p}(\mathrm{i}[D, K], \mathrm{i}[D, K]),$$

where f_p is a standard operator monotone function defined by

$$f_p(x) := p(1-p) \frac{(x-1)^2}{(x^p - 1)(x^{1-p} - 1)}. \tag{0.8}$$

The notion of skew information was recently generalized by Hansen [16] as follows: For each standard operator monotone function f that is regular, i.e., $f(0)$ $(:= \lim_{x \searrow 0} f(x)) > 0$, the *metric adjusted skew information* (or the *quantum skew information*) corresponding to f is

$$I_D^f(K) := \frac{f(0)}{2} K_D^f(\mathrm{i}[D, K], \mathrm{i}[D, K]), \qquad D \in \mathcal{D}_n,\ K \in \mathbb{H}_n, \tag{0.9}$$

which is explicitly written as

$$I_D^f(K) = \frac{f(0)}{2} \sum_{i,j=1}^n \frac{(\lambda_i - \lambda_j)^2}{\lambda_j f(\lambda_i/\lambda_j)} |K_{ij}|^2 \tag{0.10}$$

if $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$.

Via the operator $\mathbb{J}_D^f$ in (0.6), each standard operator monotone function f defines a quantity

$$\varphi_D[K, K] := \langle K, \mathbb{J}_D^f K\rangle_{\mathrm{HS}}, \qquad D \in \mathcal{D}_n,\ K \in \mathbb{H}_n, \tag{0.11}$$

which was called *generalized variance* in [36]. Any such variance has the property $\varphi_D[K, K] = \mathrm{Tr}\, DK^2$ for commuting D and K.

In the present paper we study Riemannian geometry on $\mathbb{P}_n$ with kernel metrics K^ϕ in (0.1) when the kernel function $\phi(x, y)$ is in the form $M(x, y)^\theta$, a degree $\theta \in \mathbb{R}$ power of a certain mean $M(x, y)$ for two positive numbers (as prescribed at the beginning of Section 2). The above quantities (0.2), (0.6) and (0.11) are important special cases where $\theta = 2, 1$ and -1, respectively. The paper is organized as follows. After describing our setting in Section 1 in more detail, in Section 2 we determine Riemannian metrics in our class which are written as a pull-back of the Euclidean metric. For such metrics the geodesic curve and the geodesic distance are explicitly given (Theorem 2.1). Section 3 is concerned with the (non-)completeness of Riemannian metrics in our class (Theorem 3.1) and pull-back metrics from the congruence-invariant metric g (Theorem 3.3). In Section 4 we discuss comparison properties among our Riemannian metrics. The comparison of geodesic distances for two metrics is easily described in terms of the corresponding means and the degrees of power (Theorem 4.1 and Remark 4.4). Finally in Section 5, we treat the generalized situation (of Finsler type metrics rather than Riemannian metrics) where unitarily invariant norms are applied in place of the Hilbert-Schmidt norm.

For basics on Riemannian geometry the reader may refer to texts [21, 28] for example.

1 Riemannian metrics induced by kernel functions

For each $D \in \mathbb{P}_n$ the *left* and *right multiplication* operators $\mathbb{L}_D$ and $\mathbb{R}_D$ are defined as $\mathbb{L}_D X := DX$ and $\mathbb{R}_D X := XD$ for $X \in \mathbb{M}_n$. Note that $\mathbb{L}_D$ and $\mathbb{R}_D$ are commuting positive operators on the Hilbert space $(\mathbb{M}_n, \langle\,\cdot\,,\,\cdot\,\rangle_{\mathrm{HS}})$, i.e., $\mathbb{L}_D\mathbb{R}_D = \mathbb{R}_D\mathbb{L}_D$, $\langle X, \mathbb{L}_D X\rangle_{\mathrm{HS}} \geq 0$ and $\langle X, \mathbb{R}_D X\rangle_{\mathrm{HS}} \geq 0$ for all $X \in \mathbb{M}_n$. For a kernel function $\phi : (0,\infty)\times(0,\infty) \to (0,\infty)$, a positive operator $\phi(\mathbb{L}_D, \mathbb{R}_D)$ on $(\mathbb{M}_n, \langle\,\cdot\,,\,\cdot\,\rangle_{\mathrm{HS}})$ is defined via functional calculus, that is, when $D = \sum_{i=1}^k \lambda_i P_i$ is the spectral decomposition,

$$\phi(\mathbb{L}_D, \mathbb{R}_D)X := \sum_{i,j=1}^k \phi(\lambda_i, \lambda_j) P_i X P_j, \qquad X \in \mathbb{M}_n.$$

In the rest of this section we assume that $\phi(x, y)$ is symmetric, i.e., $\phi(x, y) = \phi(y, x)$, and that $\phi(x, y)$ is smooth in x and y. Then $\phi(\mathbb{L}_D, \mathbb{R}_D)$ maps $\mathbb{H}_n$ into itself, and one can define a *Riemannian metric* K^ϕ on $\mathbb{P}_n$ by

$$K_D^\phi(H, K) := \langle H, \phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} K\rangle_{\mathrm{HS}} = \sum_{i,j=1}^k \phi(\lambda_i, \lambda_j)^{-1} \mathrm{Tr}\, P_i H P_j K \tag{1.1}$$

when $H, K \in \mathbb{H}_n$.

By taking the diagonalization $D = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*$ with a unitary U, one can also write

$$\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H = U\left(\left[\frac{1}{\sqrt{\phi(\lambda_i, \lambda_j)}}\right]_{ij} \circ (U^*HU)\right)U^*, \tag{1.2}$$

where $\circ$ denotes the *Schur* (or *Hadamard*) *product*.

Lemma 1.1. *For each $D \in \mathbb{P}_n$ let*

$$T_D^c := \{H \in \mathbb{H}_n : HD = DH\} \quad \text{and} \quad T_D^q := \{\mathrm{i}[D, K] : K \in \mathbb{H}_n\}.$$

Then

(1) $K_D^\phi(H, K) = \mathrm{Tr}\,\hat{\phi}(D)HK$ *if* $H \in T_D^c$ *and* $K \in \mathbb{H}_n$*, where* $\hat{\phi}(x) := 1/\phi(x, x)$*,* $x > 0$.

(2) $K_D^\phi(H, \mathrm{i}[D, K]) = 0$ *if* $H \in T_D^c$ *and* $K \in \mathbb{H}_n$.

(3) $K_D^\phi(\mathrm{i}[D, K], \mathrm{i}[D, K]) = \langle K, \tilde{\phi}(\mathbb{L}_D, \mathbb{R}_D)K\rangle_{\mathrm{HS}}$ *for all* $K \in \mathbb{H}_n$*, where*

$$\tilde{\phi}(x, y) := \frac{(x - y)^2}{\phi(x, y)}, \qquad x, y > 0.$$

In particular, the tangent space $T_D = \mathbb{H}_n$ has an orthogonal decomposition $T_D = T_D^c \oplus T_D^q$ with respect to K_D^ϕ. Furthermore, the linear map $H \mapsto \mathrm{i}[D, H]$ on the real Hilbert space $(\mathbb{H}_n, K_D^\phi)$ is symmetric and its kernel and range are T_D^c and T_D^q, respectively.

The proof of the lemma is left to the reader, which is easy by using (1.1). See also [37, Example 11.8].

When $\gamma : [0, 1] \to \mathbb{P}_n$ is a C^1 curve (or more generally, a continuous and piecewise C^1 curve), the *length* of γ with respect to the metric K^ϕ is given by

$$L_\phi(\gamma) := \int_0^1 \sqrt{K_{\gamma(t)}^\phi(\gamma'(t), \gamma'(t))}\, dt = \int_0^1 \|\phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t)\|_{\mathrm{HS}}\, dt. \tag{1.3}$$

Note that the length $L_\phi(\gamma)$ is independent of the choice of the parametrization of γ. The *geodesic distance* $\delta_\phi(A, B)$ between $A, B \in \mathbb{P}_n$ is the infimum of $L_\phi(\gamma)$ over all C^1 curves (or equivalently, over all smooth curves) γ from A to B. A *geodesic shortest curve* is a curve from A to B such that $L_\phi(\gamma) = \delta_\phi(A, B)$.

Now let G be a smooth function from an open interval (a, b) into $(0, \infty)$. Assume that $G'(x) \neq 0$ for all $x \in (a, b)$ so that G is a diffeomorphism from (a, b) onto a subinterval of $(0, \infty)$. Let $\mathbb{H}_n(a, b)$ denote the submanifold $\{A \in \mathbb{H}_n : a < A < b\}$ of $\mathbb{H}_n$, where $a < A < b$ means that all the eigenvalues of A are in (a, b). Then the map $A \mapsto G(A)$ defined via functional calculus is a smooth diffeomorphism from $\mathbb{H}_n(a, b)$ into $\mathbb{P}_n$. The *Fréchet derivative* $DG(A) : \mathbb{H}_n \to \mathbb{H}_n$ of G at each $A \in \mathbb{H}_n(a, b)$ is given by

$$DG(A)(H) := \frac{d}{dt}G(A + tH)\Big|_{t=0}, \qquad H \in \mathbb{H}_n.$$

The *divided difference* of G is the function $G^{[1]}(x,y)$ on $(a,b)\times(a,b)$ defined by

$$G^{[1]}(x,y) := \begin{cases} \frac{G(x)-G(y)}{x-y} & \text{if } x\ne y, \\ G'(x) & \text{if } x=y. \end{cases}$$

When $A\in\mathbb{H}_n(a,b)$ has the diagonalization $A=U\mathrm{Diag}(\lambda_1,\dots,\lambda_n)U^*$, the differential formula

$$DG(A)(H) = U\big(\big[G^{[1]}(\lambda_i,\lambda_j)\big]_{ij}\circ(U^*HU)\big)U^* = G^{[1]}(\mathbb{L}_A,\mathbb{R}_A)H \qquad (1.4)$$

holds for all $H\in\mathbb{H}_n$ (see [6, p. 124]). Our next aim is to determine a Riemannian metric K on $\mathbb{H}_n(a,b)$ such that $A\mapsto G(A)$ is an isometry into the Riemannian space $(\mathbb{P}_n,K^\phi)$, that is, for every $A\in\mathbb{P}_n(a,b)$,

$$K_A(H,K) = K^\phi_{G(A)}(DG(A)(H),DG(A)(K)), \qquad H,K\in\mathbb{H}_n.$$

This Riemannian metric on $\mathbb{H}_n(a,b)$ is called the *pull-back* of K^ϕ under the transformation $A\mapsto G(A)$.

Lemma 1.2. *Let K_A, $A\in\mathbb{H}_n(a,b)$, be the pull-back of the Riemannian metric K^ϕ on $\mathbb{P}_n$ under $A\mapsto G(A)$ as mentioned above. Let $A\in\mathbb{H}_n(a,b)$ and $A=\sum_{i=1}^k\lambda_iP_i$ be the spectral decomposition. Furthermore, let $T_A^c:=\{H\in\mathbb{H}_n : HA=AH\}$ as in Lemma 1.1. Then*

(1) *$DG(A)(H)=G'(A)H$ if $H\in T_A^c$.*

(2) *$DG(A)(\mathrm{i}[A,K])=\mathrm{i}[G(A),K]$ for all $K\in\mathbb{H}_n$.*

(3) *For every $H\in T_A^c$,*

$$K_A(H,H) = \sum_{i=1}^k \frac{G'(\lambda_i)^2}{\phi(G(\lambda_i),G(\lambda_i))}\mathrm{Tr}\,P_iH^2.$$

(4) *For every $H\in T_A^c$ and $K\in\mathbb{H}_n$, $K_A(H,\mathrm{i}[D,K])=0$.*

(5) *For every $K\in\mathbb{H}_n$,*

$$K_A(\mathrm{i}[A,K],\mathrm{i}[A,K]) = \sum_{i,j=1}^k \frac{(G(\lambda_i)-G(\lambda_j))^2}{\phi(G(\lambda_i),G(\lambda_j))}\mathrm{Tr}\,P_iKP_jK.$$

Proof. (1) is obvious.

(2) This is found in [37] but a short proof using the differential formula (see [6]) is given here. We may assume without loss of generality that A is diagonal as $A=\mathrm{Diag}(\alpha_1,\dots,\alpha_n)$. It follows from (1.4) that

$$\begin{aligned} DG(A)(\mathrm{i}[A,K]) &= \big[G^{[1]}(\alpha_i,\alpha_j)\big]_{ij}\circ[\mathrm{i}(\alpha_i-\alpha_j)K_{ij}]_{ij} \\ &= \mathrm{i}[(G(\alpha_i)-G(\alpha_j))K_{ij}] = \mathrm{i}[G(A),K]. \end{aligned}$$

(3) By the isometry property together with the above (1) and Lemma 1.1 (1) we have

$$\begin{aligned}K_A(H,H) &= K^{\phi}_{G(A)}(G'(A)H, G'(A)H) = \operatorname{Tr}\hat{\phi}(G(A))G'(A)^2H^2 \\ &= \sum_{i=1}^{k} \frac{G'(\lambda_i)^2}{\phi(G(\lambda_i), G(\lambda_i))} \operatorname{Tr} P_i H^2.\end{aligned}$$

(4) By the isometry property together with the above (1), (2) and Lemma 1.1 (2) we have

$$K_A(H, \mathrm{i}[A,K]) = K^{\phi}_{G(A)}(G'(A)H, \mathrm{i}[G(A),K]) = 0.$$

(5) Similarly, by Lemma 1.1 (3),

$$\begin{aligned}K_A(\mathrm{i}[A,K], \mathrm{i}[A,K]) &= K^{\phi}_{G(A)}(\mathrm{i}[G(A),K], \mathrm{i}[G(A),K]) \\ &= \sum_{i,j=1}^{k} \frac{(G(\lambda_i) - G(\lambda_j))^2}{\phi(G(\lambda_i), G(\lambda_j))} \operatorname{Tr} P_i K P_j K.\end{aligned}$$

□

From Lemmas 1.1 and 1.2 we arrive at the following result.

Theorem 1.3. *Let G be a smooth function from $(0,\infty)$ into $(0,\infty)$ such that $G'(x) \neq 0$ for all $x > 0$. Then the pull-back of the kernel metric K^{ϕ} under the mapping $D \in \mathbb{P}_n \mapsto G(D) \in \mathbb{P}_n$ is a kernel metric K^{ψ} corresponding to the function*

$$\psi(x,y) := \frac{\phi(G(x), G(y))}{G^{[1]}(x,y)^2}, \qquad x, y > 0.$$

2 Pull-back metrics from the Euclidean metric

We are concerned with the Riemannian metric K^{ϕ} related to a kernel function ϕ that is a power of a certain mean for two positive numbers. As in [17] a *symmetric homogeneous mean* is a function $M : (0,\infty) \times (0,\infty) \to (0,\infty)$ such that for every $x, y > 0$,

(1) $M(x,y) = M(y,x)$,

(2) $M(\alpha x, \alpha y) = \alpha M(x,y)$ for all $\alpha > 0$,

(3) $M(x,y)$ is non-decreasing in x, y,

(4) $\min\{x,y\} \le M(x,y) \le \max\{x,y\}$.

In particular, (4) implies that $M(x,x) = x$ for all $x > 0$. The above mean M is determined by a single variable function $M(x,1)$ since $M(x,y) = yM(x/y,1)$. The set of all symmetric homogeneous means was denoted by $\mathfrak{M}$ in [17], so in this paper we denote by $\mathfrak{M}_0$ the set of all smooth symmetric homogeneous means. Here note that a symmetric homogeneous mean M is smooth, i.e., $M(x,y)$ is smooth in $x,y>0$ if so is $M(x,1)$.

In the rest of the paper we always assume $n \ge 2$ since the situation is trivial when $n=1$. We assume that ϕ is a power of an $M \in \mathfrak{M}_0$ with degree $\theta \in \mathbb{R}$, i.e., $\phi(x,y) := M(x,y)^\theta$. The aim of this section is to determine when the Riemannian metric K^ϕ derived from M and θ is a pull-back of the Euclidean metric. We are interested in this problem because the geodesic shortest path in that case is explicitly written as the pull-back of a segment in the Euclidean space.

Theorem 2.1. *Let $M \in \mathfrak{M}_0$, $\theta \in \mathbb{R}$ with $\theta \ne 0$ and $\phi(x,y) := M(x,y)^\theta$. Assume that F is a smooth function from $(0,\infty)$ into $\mathbb{R}$ such that $F'(x) \ne 0$ for all $x>0$. Then the transformation $D \in \mathbb{P}_n \mapsto F(D) \in \mathbb{H}_n$ is isometric from $(\mathbb{P}_n, K^\phi)$ into the Euclidean manifold $(\mathbb{H}_n, \|\cdot\|_{\mathrm{HS}})$ if and only if*

$$F(x) = \begin{cases} \pm\frac{2}{2-\theta}x^{\frac{2-\theta}{2}} + c & \text{if } \theta \ne 2, \\ \pm\log x + c & \text{if } \theta = 2, \end{cases} \tag{2.1}$$

(up to a constant c) and

$$M(x,y) = \begin{cases} \left(\dfrac{2-\theta}{2} \cdot \dfrac{x-y}{x^{\frac{2-\theta}{2}} - y^{\frac{2-\theta}{2}}}\right)^{2/\theta} & \text{if } \theta \ne 2, \\ \dfrac{x-y}{\log x - \log y} & \text{if } \theta = 2. \end{cases} \tag{2.2}$$

Moreover, in this case, for every $A, B \in \mathbb{P}_n$ a unique (up to parametrization) geodesic shortest curve from A to B is given by

$$\gamma(t) = \begin{cases} \left((1-t)A^{\frac{2-\theta}{2}} + tB^{\frac{2-\theta}{2}}\right)^{\frac{2}{2-\theta}}, & 0 \le t \le 1 \quad \text{if } \theta \ne 2, \\ \exp((1-t)\log A + t\log B), & 0 \le t \le 1 \quad \text{if } \theta = 2, \end{cases} \tag{2.3}$$

and the geodesic distance between A and B is

$$\delta_\phi(A,B) = \begin{cases} \frac{2}{|2-\theta|}\|A^{\frac{2-\theta}{2}} - B^{\frac{2-\theta}{2}}\|_{\mathrm{HS}} & \text{if } \theta \ne 2, \\ \|\log A - \log B\|_{\mathrm{HS}} & \text{if } \theta = 2. \end{cases}$$

Proof. Let (a,b) be the range of F (which must be an open interval by assumption) and $G := F^{-1} : (a,b) \to (0,\infty)$ be the inverse of F. The stated property of isometric transformation means that for every $D \in \mathbb{P}_n$,

$$K_D^\phi(H,K) = \langle DF(D)(H), DF(D)(K)\rangle_{\mathrm{HS}}, \qquad H,K \in \mathbb{H}_n. \tag{2.4}$$

For any $A \in \mathbb{H}_n(a,b)$ and $H,K \in \mathbb{H}_n$ let $D := G(A) \in \mathbb{P}_n$ and $\tilde{H} := DG(A)(H)$, $\tilde{K} := DG(A)(K)$, where $DG(A) : \mathbb{H}_n \to \mathbb{H}_n$ is the Fréchet derivative of G at A. Since $F \circ G$ is the identity function on (a,b), it follows (see [6, p. 311]) that $DF(D) \circ DG(A)$

is the identity mapping on $\mathbb{H}_n$ so that $DF(D)(\tilde{H}) = H$ and $DF(D)(\tilde{K}) = K$. Hence it follows from (2.4) that

$$K^{\phi}_{G(A)}(DG(A)(H), DG(A)(K)) = \langle H, K\rangle_{\mathrm{HS}}, \qquad H, K \in \mathbb{H}_n.$$

This means that the pull-back of K^{ϕ} via G is the Euclidean metric on the submanifold $\mathbb{H}_n(a,b)$ of $\mathbb{H}_n$. From (3)–(5) of Lemma 1.2 one can easily see that this property is equivalent to that the following two conditions hold:

$$\frac{G'(t)^2}{G(t)^{\theta}} = 1, \qquad t \in (a,b),$$

$$\frac{(G(s)-G(t))^2}{\phi(G(s),G(t))} = (s-t)^2, \qquad s,t \in (a,b).$$

It is obvious that the above two are respectively equivalent to the following:

$$F'(x)^2 = x^{-\theta}, \qquad x > 0, \tag{2.5}$$

$$\frac{(x-y)^2}{\phi(x,y)} = (F(x)-F(y))^2, \qquad x,y > 0. \tag{2.6}$$

The differential equation (2.5) determines F as (2.1), and this together with (2.6) determines M as (2.2).

The rest of the theorem immediately follows from the isometric transformation via F in (2.1). One may just note that the segment joining $H, K \in \mathbb{H}_n$ is a unique shortest path between H and K in the Euclidean manifold $(\mathbb{H}_n, \|\cdot\|_{\mathrm{HS}})$. □

In the following we present a bit more direct proof of Theorem 2.1. Formula (2.8) below will be also useful in our discussions in the rest of the paper. Let F and $G := F^{-1}$ be as above. For each C^1 curve $\gamma : [0,1] \to \mathbb{P}_n$ we make a change of variable $\xi(t) := F(\gamma(t))$, hence $\gamma(t) = G(\xi(t))$. We then have

$$K^{\phi}_{\gamma(t)}(\gamma'(t), \gamma'(t)) = \|\phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t)\|^2_{\mathrm{HS}}$$

and

$$\gamma'(t) = DG(\xi(t))(\xi'(t)).$$

Under the diagonalization $\xi(t) = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*$ for each fixed $t \in [0,1]$, the differential formula in (1.4) is written as

$$DG(\xi(t))(\xi'(t)) = U\big(\big[G^{[1]}(\lambda_i,\lambda_j)\big]_{ij} \circ (U^*\xi'(t)U)\big)U^* = G^{[1]}(\mathbb{L}_{\xi(t)}, \mathbb{R}_{\xi(t)})\xi'(t). \tag{2.7}$$

From this and (1.2) we obtain

$$\begin{aligned}\phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t) &= \phi(\mathbb{L}_{G(\xi(t))}, \mathbb{R}_{G(\xi(t))})^{-1/2} G^{[1]}(\mathbb{L}_{\xi(t)}, \mathbb{R}_{\xi(t)})\xi'(t) \\ &= U\left(\left[\frac{G^{[1]}(\lambda_i,\lambda_j)}{\sqrt{\phi(G(\lambda_i), G(\lambda_j))}}\right]_{ij} \circ (U^*\xi'(t)U)\right)U^*. \end{aligned} \tag{2.8}$$

Here the eigenvalues of $\xi(t)$ can be arbitrary positive numbers and $\xi'(t)$ can be an arbitrary element of $\mathbb{H}_n$. Hence we see that the metric K^ϕ on $\mathbb{P}_n$ is the pull-back of the Euclidean metric on $\mathbb{H}_n(a,b)$ via F if and only if

$$\frac{G^{[1]}(s,t)}{\sqrt{\phi(G(s),G(t))}} = \pm 1 \tag{2.9}$$

for all $s,t \in (a,b)$, where the right-hand side of (2.9) is 1 or -1 according to G being increasing or decreasing. Since $\phi(x,x) = x^\theta$, (2.9) for $s = t$ yields the differential equation

$$G'(t) = \pm G(t)^{\theta/2}, \qquad t \in (a,b).$$

This is equivalently written as $F'(x) = \pm x^{-\theta/2}$, $x > 0$, which is solved as (2.1). From (2.1) and (2.9) we obtain (2.2). Thus we have proved Theorem 2.1 again. Note that one can even more simply prove the theorem by appealing to

$$F^{[1]}(x,y) = \pm \frac{1}{\sqrt{\phi(x,y)}}.$$

For $\theta \in \mathbb{R}$, $\theta \neq 0$, we write $M_\theta(x,y)$ for $M(x,y)$ given in (2.2) and

$$\phi_\theta(x,y) := M_\theta(x,y)^\theta = \begin{cases} \left(\dfrac{2-\theta}{2} \cdot \dfrac{x-y}{x^{\frac{2-\theta}{2}} - y^{\frac{2-\theta}{2}}}\right)^2 & \text{if } \theta \neq 2, \\[2ex] \left(\dfrac{x-y}{\log x - \log y}\right)^2 & \text{if } \theta = 2. \end{cases} \tag{2.10}$$

The family of means M_θ interpolates the following typical means:

$$M_{-2}(x,y) = M_{\mathrm{A}}(x,y) := \frac{x+y}{2} \quad \text{(arithmetic mean)}, \tag{2.11}$$

$$M_1(x,y) = M_{\sqrt{}}(x,y) := \left(\frac{\sqrt{x}+\sqrt{y}}{2}\right)^2 \quad \text{(root mean)}, \tag{2.12}$$

$$M_2(x,y) = M_{\mathrm{L}}(x,y) := \frac{x-y}{\log x - \log y} \quad \text{(logarithmic mean)}, \tag{2.13}$$

$$M_4(x,y) = M_{\mathrm{G}}(x,y) := \sqrt{xy} \quad \text{(geometric mean)}. \tag{2.14}$$

Furthermore, we may define $M_0(x,y)$ by taking the limit

$$M_0(x,y) := \lim_{\theta \to 0} M_\theta(x,y) = \frac{1}{e}\left(\frac{x^x}{y^y}\right)^{1/(x-y)} \quad \text{(identric mean)}, \tag{2.15}$$

and $\phi_0(x,y) \equiv 1$. Note also that

$$\frac{x-y}{\log x - \log y} = \lim_{\theta \to 2} \frac{2-\theta}{2} \cdot \frac{x-y}{x^{\frac{2-\theta}{2}} - y^{\frac{2-\theta}{2}}}.$$

As mentioned in Introduction, monotone metrics ([35]) are among particularly important class of Riemannian metrics. Those are the kernel metrics K^ϕ in the case where $\theta = 1$ and $M(x,1)$ is operator monotone. In the case $\theta = 1$, the theorem says that the metric corresponding to the root mean $M_{\sqrt{\ }}$ (that is a special case of binomial means [17]), called the *Wigner-Yanase metric*, is a unique monotone metric that is a pull-back of the Euclidean metric. This was in fact proved by Gibilisco and Isola [15] in a slightly different approach. Other famous monotone metrics are the *Bogoliubov metric* (also called the *Kubo-Mori metric*) corresponding to the logarithmic mean M_{L} and the *Bures-Uhlmann metric* corresponding to the arithmetic mean M_{A}.

In this way, we have found a one-parameter family $M_\theta \in \mathfrak{M}_0$, $\theta \in \mathbb{R}$, given in (2.2) and (2.15). It is remarkable that this is a rather familiar family of means introduced in [42] with a different parametrization, which appeared in [14] and was called *Stolarsky means* in [10, §2.6]. A monotonicity property of the family was proved in [42], which we state in the next lemma for the convenience of references.

Lemma 2.2. ([42]) *For every $x, y > 0$ with $x \ne y$, $M_\theta(x,y)$ is strictly decreasing in $\theta \in \mathbb{R}$. Furthermore, $\lim_{\theta\to-\infty} M_\theta(x,y) = \max\{x,y\}$ and $\lim_{\theta\to\infty} M_\theta(x,y) = \min\{x,y\}$.*

Next we are concerned with the relation among the metrics K^ϕ under the reflection map $A \mapsto A^{-1}$.

Proposition 2.3. *Let $M^{(1)}, M^{(2)} \in \mathfrak{M}_0$, $\theta_1, \theta_2 \in \mathbb{R}$ and $\phi^{(k)}(x,y) := M^{(k)}(x,y)^{\theta_k}$, $k = 1,2$. Then the Riemannian manifolds $(\mathbb{P}_n, K^{\phi^{(1)}})$ and $(\mathbb{P}_n, K^{\phi^{(2)}})$ are isometric under the reflection $A \mapsto A^{-1}$ on $\mathbb{P}_n$ if and only if*

$$\theta_1 + \theta_2 = 4 \quad \textit{and} \quad \left(\frac{M^{(1)}(x,y)}{\sqrt{xy}}\right)^{\theta_1} = \left(\frac{M^{(2)}(x,y)}{\sqrt{xy}}\right)^{\theta_2}, \quad x, y > 0.$$

In particular, if $\phi(x,y) = M(x,y)^2$ with an arbitrary $M \in \mathfrak{M}_0$, then $A \mapsto A^{-1}$ is an isometric transformation on $(\mathbb{P}_n, K^\phi)$. Moreover, for every $\theta \in \mathbb{R}$, $(\mathbb{P}_n, K^{\phi_\theta})$ and $(\mathbb{P}_n, K^{\phi_{4-\theta}})$ are isometric under $A \mapsto A^{-1}$, where $\phi_\theta(x,y)$ is given in (2.10).

Proof. Let γ be a C^1 curve in $\mathbb{P}_n$. Since

$$\frac{d}{dt}\big(\gamma(t)^{-1}\big) = -\mathbb{L}_{\gamma(t)}^{-1}\mathbb{R}_{\gamma(t)}^{-1}\gamma'(t),$$

we have

$$\phi^{(2)}(\mathbb{L}_{\gamma(t)}^{-1}, \mathbb{R}_{\gamma(t)}^{-1})^{-1/2}\left(\frac{d}{dt}\big(\gamma(t)^{-1}\big)\right) = -\phi^{(2)}(\mathbb{L}_{\gamma(t)}^{-1}, \mathbb{R}_{\gamma(t)}^{-1})^{-1/2}\mathbb{L}_{\gamma(t)}^{-1}\mathbb{R}_{\gamma(t)}^{-1}\gamma'(t).$$

Hence $A \mapsto A^{-1}$ gives an isometry from $(\mathbb{P}_n, K^{\phi^{(1)}})$ to $(\mathbb{P}_n, K^{\phi^{(2)}})$ if and only if

$$\|\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}} = \|\phi^{(2)}(\mathbb{L}_D^{-1}, \mathbb{R}_D^{-1})^{-1/2}\mathbb{L}_D^{-1}\mathbb{R}_D^{-1}H\|_{\mathrm{HS}}$$

for all $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$. We may assume that D is diagonal. For $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ the above equality is written as

$$\left\| \left[\frac{1}{\sqrt{\phi^{(1)}(\lambda_i, \lambda_j)}} \right]_{ij} \circ H \right\|_{\mathrm{HS}} = \left\| \left[\frac{1}{\sqrt{\phi^{(2)}(\lambda_i^{-1}, \lambda_j^{-1})\, \lambda_i \lambda_j}} \right]_{ij} \circ H \right\|_{\mathrm{HS}}.$$

This hold for all $H \in \mathbb{H}_n$ if and only if

$$\phi^{(1)}(x, y) = \phi^{(2)}(x^{-1}, y^{-1}) x^2 y^2, \qquad x, y > 0,$$

that is,

$$M^{(1)}(x, y)^{\theta_1} = M^{(2)}(x^{-1}, y^{-1})^{\theta_2} x^2 y^2, \qquad x, y > 0.$$

Letting $x = y$ implies that $x^{\theta_1} = x^{4-\theta_2}$ for all $x > 0$. Hence $\theta_1 + \theta_2 = 4$ must hold and the above condition for $M^{(1)}$ and $M^{(2)}$ is rewritten as

$$\left(\frac{M^{(1)}(x, y)}{\sqrt{xy}} \right)^{\theta_1} = \left(\frac{M^{(2)}(x, y)}{\sqrt{xy}} \right)^{\theta_2}, \qquad x, y > 0.$$

Since this is obviously satisfied for $\theta_1 = \theta_2 = 2$ and $M^{(1)} = M^{(2)}$, the second assertion follows. A simple computation with (2.2) gives the last assertion. □

Remark 2.4. The latter assertions of Proposition 2.3 can be extended as follows: For every $\theta, \theta' \in \mathbb{R} \setminus \{2\}$ the Riemannian manifolds $(\mathbb{P}_n, K^{\phi_\theta})$ and $(\mathbb{P}_n, K^{\phi_{\theta'}})$ are isometric under the diffeomorphism $A \mapsto F(A)$ on $\mathbb{P}_n$ given with

$$F(x) := \left| \frac{2 - \theta'}{2 - \theta} \right|^{\frac{2}{2-\theta'}} x^{\frac{2-\theta}{2-\theta'}}, \qquad x > 0,$$

since F satisfies

$$\phi_\theta(x, y) = \frac{\phi_{\theta'}(F(x), F(y))}{F^{[1]}(x, y)^2}, \qquad x, y > 0.$$

Also, for every $\alpha \in \mathbb{R} \setminus \{0\}$, $A \mapsto A^\alpha$ is an isometric transformation on $(\mathbb{P}_n, K^{\phi_2})$.

An interesting problem concerning the family M_θ ($= M$ in (2.2)) is to determine the range of θ for which M_θ is an operator monotone mean, i.e., $M_\theta(x, 1)$ is an operator monotone function on $(0, \infty)$. The cases $\theta = -2, 1, 2$ and 4 are among typical operator monotone functions as listed in (2.11)–(2.14). The problem has been settled by Kosaki [23] in such a way that $M_\theta(x, 1)$ is operator monotone if and only if $-2 \le \theta \le 6$.

We give the next lemma on M_θ for later use.

Lemma 2.5. *Let $M_{\mathrm{H}}(x, y) := 2xy/(x + y)$, the harmonic mean. Then $M_{10}(x, 1) > M_{\mathrm{H}}(x, 1)$ for all $x > 0$ with $x \ne 1$. For every $\theta > 10$, $M_\theta(x, 1) < M_{\mathrm{H}}(x, 1)$ if x $(\ne 1)$ is sufficiently near 1.*

Proof. The first assertion follows from

$$M_{10}(x,1)^5 - \left(\frac{2x}{x+1}\right)^5 = \frac{4x^4(x-1)^4(x+1)}{(x^3+x^2+x+1)(x+1)^5}.$$

To prove the second, let $\theta > 10$ and $\alpha := (\theta-2)/2 > 4$. Direct computations show

$$\begin{aligned} M_\theta(x,1)^{\alpha+1} &= \alpha\frac{x^{\alpha+1}-x^\alpha}{x^\alpha-1} \\ &= 1 + \frac{\alpha+1}{2}(x-1) + \frac{(\alpha+1)(\alpha-1)}{12}(x-1)^2 + o((x-1)^2), \\ M_{\mathrm{H}}(x,1)^{\alpha+1} &= \left(\frac{2x}{x+1}\right)^{\alpha+1} \\ &= 1 + \frac{\alpha+1}{2}(x-1) + \frac{(\alpha+1)(\alpha-2)}{8}(x-1)^2 + o((x-1)^2), \end{aligned}$$

which give the desired assertion. □

3 The degree 2 case

A Riemannian manifold said to be *complete* if the geodesic distance induced from the Riemannian metric is complete. It is a general fact in Riemannian geometry that a geodesic shortest curve joining any two points exists in a complete Riemannian manifold. The next theorem shows that the Riemannian manifold $(\mathbb{P}_n, K^\phi)$ treated in Section 2 is never complete except the case of degree $\theta = 2$.

Theorem 3.1. *Let $M \in \mathfrak{M}_0$, $\theta \in \mathbb{R}$ and $\phi(x,y) := M(x,y)^\theta$. Then the Riemannian manifold $(\mathbb{P}_n, K^\phi)$ is complete if and only if $\theta = 2$. Hence, when $\theta = 2$ (and $M \in \mathfrak{M}_0$ is arbitrary), for any $A, B \in \mathbb{P}_n$ there is a geodesic shortest curve joining A, B in $(\mathbb{P}_n, K^\phi)$.*

Proof. First assume $\theta \neq 2$ and prove the non-completeness of $(\mathbb{P}_n, K^\phi)$. Let $\gamma(t) := tI$ for $t > 0$, where I is the $n \times n$ identity matrix. Since

$$\|\phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t)\|_{\mathrm{HS}} = \|M(t,t)^{-\theta/2}I\|_{\mathrm{HS}} = t^{-\theta/2}\sqrt{n},$$

we have

$$\int_0^1 \sqrt{K^\phi_{\gamma(t)}(\gamma'(t),\gamma'(t))}\,dt < +\infty \quad \text{if } \theta < 2,$$

$$\int_1^\infty \sqrt{K^\phi_{\gamma(t)}(\gamma'(t),\gamma'(t))}\,dt < +\infty \quad \text{if } \theta > 2.$$

Hence, if we define $A_k := k^{-1}I$ if $\theta < 2$ and $A_k := kI$ if $\theta > 2$, then it follows that $\{A_k\}_{k=1}^\infty$ is Cauchy with respect to the geodesic distance δ_ϕ. Now suppose that $\delta_\phi(A_k, A) \to 0$ for some $A \in \mathbb{P}_n$. Since δ_ϕ and $\|\cdot\|_{\mathrm{HS}}$ (hence also the operator norm $\|\cdot\|_\infty$) define the same topology on $\mathbb{P}_n$ (see [21, Chapter IV, Proposition 3.5]), we must have $\|A_k - A\|_\infty \to 0$. Hence there is an $\varepsilon > 0$ such that $\varepsilon I \le A_k \le \varepsilon^{-1} I$ for all k. This contradicts the choice of A_k, so $\{A_k\}$ does not converge in $(\mathbb{P}_n, K^\phi)$.

Next assume $\theta = 2$, and prove that $(\mathbb{P}_n, K^\phi)$ is complete. To do so, we need a lemma.

Lemma 3.2. *If $M \in \mathfrak{M}_0$ and $\phi(x,y) := M(x,y)^2$, then $\delta_\phi(A,I) = \|\log A\|_{\mathrm{HS}}$ for every $A \in \mathbb{P}_n$.*

Proof. We may assume that A is diagonal. Let $\gamma : [0,1] \to \mathbb{P}_n$ be a C^1 curve from A to I, and diagonalize $\gamma(t)$, $0 \le t \le 1$, so that

$$\gamma(t) = U(t)\mathrm{Diag}(\lambda_1(t), \dots, \lambda_n(t))U(t)^*$$

with $\lambda_1(t) \le \cdots \le \lambda_n(t)$ and unitary matrices $U(t)$. Here one can fix $U(t)$, $0 \le t \le 1$, so that $\lambda_1(t), \dots, \lambda_n(t)$ and $U(t)$ are C^1 except branching points of $\lambda_1(t), \dots, \lambda_n(t)$ (see [20] for example). Note that the set of branching points is at most countable. Therefore, for each t except such branching points, we have

$$\begin{aligned}\gamma'(t) = U(t)\mathrm{Diag}(\lambda_1'(t), \dots, \lambda_n'(t))U(t)^* + U'(t)\mathrm{Diag}(\lambda_1(t), \dots, \lambda_n(t))U(t)^* \\ + U(t)\mathrm{Diag}(\lambda_1(t), \dots, \lambda_n(t))U'(t)^*\end{aligned}$$

so that

$$\begin{aligned}U(t)^*\gamma'(t)U(t) = \mathrm{Diag}(\lambda_1'(t), \dots, \lambda_n'(t)) + U(t)^*U'(t)\mathrm{Diag}(\lambda_1(t), \dots, \lambda_n(t)) \\ + \mathrm{Diag}(\lambda_1(t), \dots, \lambda_n(t))U'(t)^*U(t).\end{aligned}$$

Since $U(t)^*U(t) = I$ yields that $U'(t)^*U(t) + U(t)^*U'(t) = O$, the diagonal entries of $U(t)^*\gamma'(t)U(t)$ are $\lambda_1'(t), \dots, \lambda_n'(t)$. Hence it follows that

$$\begin{aligned}&\|\phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t)\|_{\mathrm{HS}} \\ &\quad = \left\| \left[\frac{1}{M(\lambda_i(t), \lambda_j(t))}\right]_{ij} \circ (U(t)^*\gamma'(t)U(t)) \right\|_{\mathrm{HS}} \ge \sqrt{\sum_{i=1}^n \left(\frac{\lambda_i'(t)}{\lambda_i(t)}\right)^2}\end{aligned}$$

for all t except a countable set. Since $\xi(t) := \mathrm{Diag}(\log\lambda_1(t), \dots, \log\lambda_n(t))$ is a curve (continuous in $0 \le t \le 1$ and C^1 except a countable set as mentioned above) from $\log A$ to O, we have

$$L_\phi(\gamma) \ge \int_0^1 \|\xi'(t)\|_{\mathrm{HS}}\,dt \ge \left\|\int_0^1 \xi'(t)\,dt\right\|_{\mathrm{HS}} \ge \|\log A\|_{\mathrm{HS}}.$$

Furthermore, if $A = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ and $\gamma_0(t) := A^{1-t} = \mathrm{Diag}(\lambda_1^{1-t}, \dots, \lambda_n^{1-t})$ for $0 \le t \le 1$, then one can easily compute

$$L_\phi(\gamma_0) = \sqrt{\sum_{i=1}^n (\log\lambda_i)^2} = \|\log A\|_{\mathrm{HS}},$$

implying $\delta_\phi(A,I) = \|\log A\|_{\mathrm{HS}}$. ∎

Proof of Theorem 3.1 (continued). Let $\{A_k\}$ be a δ_ϕ-Cauchy sequence in $\mathbb{P}_n$. Since δ_ϕ and the operator norm $\|\cdot\|_\infty$ define the same topology on $\mathbb{P}_n$ as mentioned in the first part of the proof, it suffices to prove that $\|A_k - A\|_\infty \to 0$ for some $A \in \mathbb{P}_n$. In fact, it is enough to show that there exist a subsequence $\{A_{k_m}\}$ of $\{A_k\}$ and an $A \in \mathbb{P}_n$ such that $\|A_{k_m} - A\|_\infty \to 0$. Since $|\delta_\phi(A_k, I) - \delta_\phi(A_l, I)| \le \delta_\phi(A_k, A_l) \to 0$ as $k, l \to \infty$, it follows from Lemma 3.2 that $\delta_\phi(A_k, I) = \|\log A_k\|_{\mathrm{HS}}$ is a bounded sequence and so $\sup_k \|\log A_k\|_\infty < +\infty$. Hence there is an $\varepsilon > 0$ such that $\varepsilon I \le A_k \le \varepsilon^{-1} I$ for all k. The desired assertion now follows from the $\|\cdot\|_\infty$-compactness of $\{A \in \mathbb{P}_n : \varepsilon I \le A \le \varepsilon^{-1} I\}$. □

Let $\phi_{\mathrm{G}}(x, y)$ denote the degree 2 power of the geometric mean, i.e., $\phi_{\mathrm{G}}(x, y) := M_{\mathrm{G}}(x, y)^2 = xy$. The metric $K^{\phi_{\mathrm{G}}}$ induced from ϕ_{G} is nothing but the congruence-invariant metric g given in (0.2). The completeness of the Riemannian manifold $(\mathbb{P}_n, K^{\phi_{\mathrm{G}}})$ was shown in [9]. Now we define a one-parameter family of kernel functions

$$N_\alpha(x, y) := \alpha(xy)^{\alpha/2} \frac{x - y}{x^\alpha - y^\alpha}, \quad x, y > 0, \quad \alpha \in \mathbb{R}, \tag{3.1}$$

where $N_0(x, y)$ is understood as

$$N_0(x, y) := \lim_{\alpha \to 0} N_\alpha(x, y) = \frac{x - y}{\log x - \log y} \quad \text{(logarithmic mean).}$$

We have $N_1(x, y) = \sqrt{xy}$ (geometric mean) and $N_2(x, y) = 2xy/(x + y)$ (harmonic mean). Note that $N_\alpha(x, y)$ is symmetric and homogeneous in the sense of (1) and (2) at the beginning of Section 2 and $N_{-\alpha}(x, y) = N_\alpha(x, y)$. When $\alpha > 2$, N_α does not belong to $\mathfrak{M}_0$ since $N_\alpha(x, 1) \to 0$ as $x \to \infty$. When $0 < \alpha \le 2$, one can easily see by elementary calculus that $N_\alpha(x, 1)$ is increasing in $x > 0$ and $1 \le N_\alpha(x, 1) \le x$ for all $x \ge 1$. It is also not difficult to see that $N_\alpha(x, y)$ is strictly decreasing in $\alpha > 0$ for each $x, y > 0$ with $x \ne y$. Thus $\{N_\alpha\}_{0 \le \alpha \le 2}$ is a family of means in $\mathfrak{M}_0$ interpolating the logarithmic and the harmonic means. Furthermore, it is seen as in [23] that N_α is an operator monotone mean for each $0 \le \alpha \le 2$.

We determine when our Riemannian metric K^ϕ is a pull-back of $K^{\phi_{\mathrm{G}}}$ up to a multiple constant, and moreover extend the formula for the geodesic curve in (0.4) and that for the geodesic distance in (0.5) for $g = K^{\phi_{\mathrm{G}}}$ to the family of metrics induced from the above N_α.

Theorem 3.3. *Let $M \in \mathfrak{M}_0$, $\theta \in \mathbb{R}$ and $\phi(x, y) := M(x, y)^\theta$. Let $\alpha > 0$. Assume that F is a smooth function from $(0, \infty)$ into itself such that $F'(x) \ne 0$ for all $x > 0$. Then the transformation $D \in \mathbb{P}_n \mapsto F(D) \in \mathbb{P}_n$ is isometric from $(\mathbb{P}_n, \alpha^2 K^\phi)$ into $(\mathbb{P}_n, K^{\phi_{\mathrm{G}}})$ if and only if $\theta = 2$, $\alpha \le 2$, $F(x) = cx^{\pm\alpha}$ (up to a constant $c > 0$) and $M = N_\alpha$, where $\alpha^2 K^\phi$ is the metric $\alpha^2 K^\phi_D(H, K)$ for $D \in \mathbb{P}_n$ and $H, K \in \mathbb{H}_n$.*

In the above case, for every $A, B \in \mathbb{P}_n$ there exists a unique (up to parametrization) geodesic shortest curve in $(\mathbb{P}_n, K^\phi)$ from A to B given by

$$\gamma(t) := (A^\alpha \#_t B^\alpha)^{1/\alpha} \left(= \left(A^{\alpha/2}(A^{-\alpha/2} B^\alpha A^{-\alpha/2})^t A^{\alpha/2}\right)^{1/\alpha} \right)$$

and moreover

$$\delta_\phi(A,B) = \|\log(A^{-\alpha/2}B^\alpha A^{-\alpha/2})^{1/\alpha}\|_{\mathrm{HS}}.$$

Proof. For any C^1 curve γ in $\mathbb{P}_n$ let $\xi(t) := F(\gamma(t))$. Under the diagonalization $\gamma(t) = U\mathrm{Diag}(\lambda_1,\dots,\lambda_n)U^*$ for each fixed $t \in [0,1]$ we have by (1.2) and (2.8)

$$\|\phi(\mathbb{L}_{\gamma(t)},\mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t)\|_{\mathrm{HS}} = \left\|\left[\frac{1}{\sqrt{\phi(\lambda_i,\lambda_j)}}\right]_{ij} \circ (U^*\gamma'(t)U)\right\|_{\mathrm{HS}},$$

$$\|\phi_{\mathrm{G}}(\mathbb{L}_{\xi(t)},\mathbb{R}_{\xi(t)})^{-1/2}\xi'(t)\|_{\mathrm{HS}} = \left\|\left[\frac{F^{[1]}(\lambda_i,\lambda_j)}{\sqrt{F(\lambda_i)F(\lambda_j)}}\right]_{ij} \circ (U^*\gamma'(t)U)\right\|_{\mathrm{HS}}.$$

Hence the isometry property stated in the theorem implies that

$$\frac{F^{[1]}(x,y)}{\sqrt{F(x)F(y)}} = \pm\frac{\alpha}{\sqrt{\phi(x,y)}}, \qquad x,y>0. \tag{3.2}$$

When $x=y$ this yields

$$\frac{F'(x)}{F(x)} = \pm\alpha x^{-\theta/2}, \qquad x>0. \tag{3.3}$$

Suppose $\theta \ne 2$. Then (3.3) is solved as

$$F(x) = c\exp\biggl(\pm\frac{2\alpha}{2-\theta}\,x^{\frac{2-\theta}{2}}\biggr)$$

with a constant $c>0$. By this and (3.2), $\phi(x,1)$ is written as

$$\phi(x,1) = \alpha^2 e^\beta\,\frac{e^{\beta x^r}(x-1)^2}{(e^{\beta x^r}-e^\beta)^2} \quad \text{with } r := \frac{2-\theta}{2},\ \beta := \pm\frac{2\alpha}{2-\theta}.$$

For both $\pm$ signs in the definition of β, all the cases except $\theta = 2$ are excluded as follows: If $0 \le \theta < 2$, then $\phi(x,1) \to 0$ as $x \to \infty$. But this is inconsistent with $\phi(x,1) = M(x,1)^\theta \ge 1$ for $x \ge 1$. If $\theta < 0$, then $x^{-\theta}\phi(x,1) \to 0$ as $x\to\infty$, which is inconsistent with $\phi(x,1) = M(x,1)^\theta \ge x^\theta$ for $x \ge 1$. If $\theta > 2$, then $x^{-\theta}\phi(x,1) \to 0$ as $x \to 0$, which is also inconsistent with $\phi(x,1) \ge x^\theta$ for $0 < x \le 1$. Therefore $\theta = 2$ must hold. When $\theta = 2$, the solution of (3.3) is $F(x) = cx^{\pm\alpha}$ with a constant $c>0$. This and (3.2) determine M as $M = N_\alpha$ since $N_{-\alpha} = N_\alpha$. Then α obeys the restriction $\alpha \le 2$ as shown before the theorem. It is immediate to see that the isometry property actually holds if θ, α, F and M are as stated in the theorem.

When α^2K^ϕ is a pull-back of $K^{\phi_{\mathrm{G}}}$ as above, a geodesic shortest curve in $(\mathbb{P}_n, K^\phi)$ joining each $A,B\in\mathbb{P}_n$ is uniquely determined as the image under $D \mapsto D^{1/\alpha}$ of that in $(\mathbb{P}_n, K^{\phi_{\mathrm{G}}})$ joining A^α, B^α. Thanks to (0.4) its explicit form is

$$\gamma(t) := (A^\alpha \#_t B^\alpha)^{1/\alpha}, \qquad 0 \le t \le 1.$$

Furthermore, thanks to (0.5) it is also immediate to see that

$$\begin{aligned}\delta_\phi(A,B) &= \frac{1}{\alpha}\delta_{\phi_{\mathrm{G}}}(A^\alpha,B^\alpha) = \frac{1}{\alpha}\|\log(A^{-\alpha/2}B^\alpha A^{-\alpha/2})\|_{\mathrm{HS}} \\ &= \|\log(A^{-\alpha/2}B^\alpha A^{-\alpha/2})^{1/\alpha}\|_{\mathrm{HS}},\end{aligned}$$

as required. □

It is desirable to prove the uniqueness of geodesic shortest curves for all metrics treated in Theorem 3.1 in the degree 2 case.

We write ψ_α for ϕ arising in Theorem 3.3, i.e.,

$$\psi_\alpha(x,y) := N_\alpha(x,y)^2 = \alpha^2(xy)^\alpha\left(\frac{x-y}{x^\alpha-y^\alpha}\right)^2, \qquad 0<\alpha\le 2. \tag{3.4}$$

It is worth noting that the geodesic shortest path and its distance in $(\mathbb{P}_n, K^{\psi_\alpha})$ converge as $\alpha \searrow 0$ to those in $(\mathbb{P}_n, K^{\phi_{\mathrm{L}}})$, where $\phi_{\mathrm{L}}(x,y) := M_{\mathrm{L}}(x,y)^2$, the degree 2 power of the logarithmic mean. Namely, we have

$$\lim_{\alpha\searrow 0}(A^\alpha \#_t B^\alpha)^{1/\alpha} = \exp((1-t)\log A + t\log B), \qquad 0\le t\le 1,$$

$$\lim_{\alpha\searrow 0}\|\log(A^{-\alpha/2}B^\alpha A^{-\alpha/2})^{1/\alpha}\|_{\mathrm{HS}} = \|\log A - \log B\|_{\mathrm{HS}}$$

(see the $\theta = 2$ case of Theorem 2.1). In fact, the latter follows from a version of the Lie-Trotter formula

$$\lim_{\alpha\to 0}(A^{-\alpha/2}B^\alpha A^{-\alpha/2})^{1/\alpha} = \exp(-\log A + \log B)$$

and the former is its modification (see [19, Lemma 3.3]). It is also worthwhile to note that $\|\log(A^{-\alpha/2}B^\alpha A^{-\alpha/2})^{1/\alpha}\|_{\mathrm{HS}}$ is increasing in $\alpha > 0$ due to Araki's log-majorization [4] (see also [2]). Hence $\delta_{\psi_\alpha}(A,B)$ decreases to $\delta_{\phi_{\mathrm{L}}}(A,B)$ as $\alpha \searrow 0$ while $\psi_\alpha(x,y)$ increases to $\phi_{\mathrm{L}}(x,y)$ as $\alpha \searrow 0$. In fact, this kind of comparison property is true in general as we will see in the subsequent sections.

When $A_1,\dots,A_k \in \mathbb{P}_n$, since the arithmetic mean $\frac{1}{k}\sum_{j=1}^k A_j$ is the unique minimizer of $A \in \mathbb{P}_n \mapsto \sum_{j=1}^k \|A - A_j\|_{\mathrm{HS}}^2$, it is immediate from Theorem 2.1 that a certain power mean $\left(\frac{1}{k}\sum_{j=1}^k A_j^{\frac{2-\theta}{2}}\right)^{\frac{2}{2-\theta}}$ (understood as $\exp\bigl(\frac{1}{k}\sum_{j=1}^n \log A_j\bigr)$ if $\theta = 2$) is determined as a unique minimizer of $A \in \mathbb{P}_n \mapsto \sum_{j=1}^k \delta_{\phi_\theta}^2(A, A_j)$. Let $G(A_1,\dots,A_k)$ be the "geometric mean" introduced in [9, 8], i.e., the unique minimizer of $A \mapsto \sum_{j=1}^k \delta_{M_{\mathrm{G}}^2}^2(A, A_j)$. It is also immediately seen from Theorem 3.3 that $G(A_1^\alpha,\dots,A_k^\alpha)^{1/\alpha}$ is a unique minimizer of $A \mapsto \sum_{j=1}^k \delta_{\psi_\alpha}^2(A, A_j)$, which is regarded as a k-variable extension of $(A^\alpha \# B^\alpha)^{1/\alpha}$.

4 Comparison property

The aim of this section is to compare the geodesic distances for different Riemannian metrics related to means in $\mathfrak{M}_0$. The following is a result of this kind for general kernel metrics.

Theorem 4.1. *Let $\phi^{(1)}, \phi^{(2)} : (0,\infty) \times (0,\infty) \to (0,\infty)$ be smooth symmetric kernel functions. Then the following conditions are equivalent:*

(i) $\phi^{(1)}(x,y) \le \phi^{(2)}(x,y)$ *for all* $x, y > 0$*;*

(ii) $K_D^{\phi^{(1)}}(H,H) \ge K_D^{\phi^{(2)}}(H,H)$ *for all* $D \in \mathbb{P}_n$ *and* $H \in \mathbb{H}_n$*;*

(iii) $L_{\phi^{(1)}}(\gamma) \ge L_{\phi^{(2)}}(\gamma)$ *for all* C^1 *curve* γ *in* $\mathbb{P}_n$*;*

(iv) $\delta_{\phi^{(1)}}(A,B) \ge \delta_{\phi^{(2)}}(A,B)$ *for all* $A, B \in \mathbb{P}_n$*.*

Hence, the respective conditions with equality in place of inequality in (i)–(iv) *are equivalent.*

The next lemma is useful to prove the theorem while it is meaningful by itself.

Lemma 4.2. *Let $\phi : (0,\infty) \times (0,\infty) \to (0,\infty)$ be a smooth symmetric kernel function. Then for every $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$,*

$$\lim_{\varepsilon \searrow 0} \frac{\delta_\phi(D, D+\varepsilon H)}{\varepsilon} = \|\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H\|_{\mathrm{HS}}.$$

Proof. First recall that if $\mathbb{T}$ is a linear operator on the Hilbert space $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ represented as the Schur multiplication by a matrix $[t_{ij}] \in \mathbb{H}_n$, then $\mathbb{T} \ge 0$ if and only if $t_{ij} \ge 0$ for all $i, j = 1, \dots, n$. We denote by $\mathbb{I}$ the identity operator on $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$, which is represented as the Schur multiplication by the matrix of all entries equal to 1. To prove the lemma, we may assume that $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$. Since

$$\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} H = \left[\phi(\lambda_i, \lambda_j)^{-1}\right]_{ij} \circ H, \qquad H \in \mathbb{H}_n,$$

it follows that

$$\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} \ge \left(\min_{1 \le i,j \le n} \phi(\lambda_i, \lambda_j)^{-1}\right) \mathbb{I}$$

as operators on $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$. For each $\rho > 0$ with $\rho < \min_{i,j} \phi(\lambda_i, \lambda_j)^{-1}$, since $A \in \mathbb{P}_n \mapsto \phi(\mathbb{L}_A, \mathbb{R}_A)$ is continuous, there exists an $r_1 > 0$ such that if $A \in \mathbb{P}_n$ and $\|A - D\|_{\mathrm{HS}} < r_1$ then

$$\|\phi(\mathbb{L}_A, \mathbb{R}_A)^{-1} - \phi(\mathbb{L}_D, \mathbb{R}_D)^{-1}\|_\infty < \rho,$$

where $\|\cdot\|_\infty$ denotes the operator norm for operators on $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$. Furthermore, since δ_ϕ and $\|\cdot\|_{\mathrm{HS}}$ define the same topology on $\mathbb{P}_n$ (see [21, Chapter IV, Proposition 3.5]), there exists an $r_0 > 0$ such that if $A \in \mathbb{P}_n$ and $\delta_\phi(A, D) < r_0$ then $\|A - D\|_{\mathrm{HS}} < r_1$.

Now let $H \in \mathbb{H}_n$ and $\varepsilon > 0$ be sufficiently small so that $\delta_\phi(D, D+\varepsilon H) < r_0$ and $\varepsilon\|H\|_{\mathrm{HS}} < r_1$. Let $\gamma : [0,1] \to \mathbb{P}_n$ be any C^1 curve from D to $D+\varepsilon H$ such that $L_\phi(\gamma) < r_0$. Since $\delta_\phi(\gamma(t), D) < r_0$ and so $\|\gamma(t) - D\|_{\mathrm{HS}} < r_1$ for all $0 \le t \le 1$, we have

$$\begin{aligned}
L_\phi(\gamma) &= \int_0^1 \sqrt{\langle \gamma'(t), \phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1}\gamma'(t)\rangle_{\mathrm{HS}}}\, dt \\
&\ge \int_0^1 \sqrt{\langle \gamma'(t), (\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I})\gamma'(t)\rangle_{\mathrm{HS}}}\, dt \\
&= \int_0^1 \|(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I})^{1/2}\gamma'(t)\|_{\mathrm{HS}}\, dt \\
&\ge \|(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I})^{1/2}(\varepsilon H)\|_{\mathrm{HS}} \\
&= \varepsilon \Big\| \big[(\phi(\lambda_i, \lambda_j)^{-1} - \rho)^{1/2}\big]_{ij} \circ H \Big\|_{\mathrm{HS}}.
\end{aligned}$$

In the above, note that $\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I} \ge 0$ on $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ since $\rho < \min_{i,j}\phi(\lambda_i, \lambda_j)^{-1}$. Also, the second inequality above follows since $\int_0^1 \|(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho I)^{1/2}\gamma'(t)\|_{\mathrm{HS}}\, dt$ is the length of the curve $(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I})^{1/2}\gamma(t)$, $0 \le t \le 1$, from $(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I})^{1/2}D$ to $(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} - \rho\mathbb{I})^{1/2}(D+\varepsilon H)$ in the Euclidean space $(\mathbb{H}_n, \|\cdot\|_{\mathrm{HS}})$ and it is shortest if γ is the segment between D and $D+\varepsilon H$. Taking the infimum of $L_\phi(\gamma)$ gives

$$\delta_\phi(D, D+\varepsilon H) \ge \varepsilon \Big\| \big[(\phi(\lambda_i, \lambda_j)^{-1} - \rho)^{1/2}\big]_{ij} \circ H \Big\|_{\mathrm{HS}}.$$

On the other hand, let $\gamma_0(t) := D + t\varepsilon H$. Since $\|\gamma_0(t) - D\|_{\mathrm{HS}} \le \varepsilon\|H\|_{\mathrm{HS}} < r_1$ for $0 \le t \le 1$, we have

$$\begin{aligned}
\delta_\phi(D, D+\varepsilon H) &\le L_\phi(\gamma_0) \\
&= \int_0^1 \sqrt{\langle \gamma_0'(t), \phi(\mathbb{L}_{\gamma_0(t)}, \mathbb{R}_{\gamma_0(t)})^{-1}\gamma_0'(t)\rangle_{\mathrm{HS}}}\, dt \\
&\le \int_0^1 \sqrt{\langle \gamma_0'(t), (\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} + \rho\mathbb{I})\gamma_0'(t)\rangle_{\mathrm{HS}}}\, dt \\
&= \|(\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1} + \rho\mathbb{I})^{1/2}(\varepsilon H)\|_{\mathrm{HS}} \\
&= \varepsilon \Big\| \big[(\phi(\lambda_i, \lambda_j)^{-1} + \rho)^{1/2}\big]_{ij} \circ H \Big\|_{\mathrm{HS}}.
\end{aligned}$$

Since ρ is arbitrary,

$$\lim_{\varepsilon \searrow 0} \frac{\delta_\phi(D, D+\varepsilon H)}{\varepsilon} = \Big\| \big[\phi(\lambda_i, \lambda_j)^{-1/2}\big]_{ij} \circ H \Big\|_{\mathrm{HS}} = \|\phi(\mathbb{L}_D, \mathbb{L}_D)^{-1/2}H\|_{\mathrm{HS}}.$$

□

Proof of Theorem 4.1. (i) ⇒ (ii) is immediately seen thanks to (1.1) and (1.2). (ii) ⇒ (iii) ⇒ (iv) is obvious. Finally, assume (iv) and apply Lemma 4.2 to obtain

$$\|\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}} \ge \|\phi^{(2)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}} \tag{4.1}$$

for all $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$ (i.e., (ii) holds). When $D := \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \oplus I_{n-2}$ with $x, y > 0$ and $H := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus O_{n-2}$, (4.1) implies (i). □

Corollary 4.3. *Let $\phi^{(k)}$, $k = 1, 2$, be as in Theorem 4.1 and G be as in Theorem 1.3. Then the following conditions are equivalent:*

(i) $\phi^{(1)}(x, y) = \phi^{(2)}(G(x), G(y))/G^{[1]}(x, y)^2$ *for all* $x, y > 0$*;*

(ii) $D \in \mathbb{P}_n \mapsto G(D) \in \mathbb{P}_n$ *gives rise to an isometry between the Riemannian metrics* $K^{\phi^{(1)}}$ *and* $K^{\phi^{(2)}}$*, i.e.,* $K_D^{\phi^{(1)}}(H, K) = K_{G(D)}^{\phi^{(2)}}(DG(D)(H), DG(D)(K))$ *for all* $D \in \mathbb{P}_n$ *and* $H, K \in \mathbb{H}_n$*;*

(iii) $L_{\phi^{(1)}}(\gamma) = L_{\phi^{(2)}}(G(\gamma))$ *for all* C^1 *curve* γ *in* $\mathbb{P}_n$*;*

(iv) $D \in \mathbb{P}_n \mapsto G(D) \in \mathbb{P}_n$ *gives rise to an isometry between the geodesic distances* $\delta_{\phi^{(1)}}$ *and* $\delta_{\phi^{(2)}}$*, i.e.,* $\delta_{\phi^{(1)}}(A, B) = \delta_{\phi^{(2)}}(G(A), G(B))$ *for all* $A, B \in \mathbb{P}_n$*.*

Proof. Theorem 1.3 says that the pull-back of the metric $K^{\phi^{(2)}}$ under $D \mapsto G(D)$ is the kernel metric K^ψ with

$$\psi(x, y) := \frac{\phi^{(2)}(G(x), G(y))}{G^{[1]}(x, y)}, \qquad x, y > 0.$$

Note that

$$K_D^\psi(H, K) = K_{G(D)}^{\phi^{(2)}}(DG(D)(H), DG(D)(K)), \quad D \in \mathbb{P}_n,\ H, K \in \mathbb{H}_n,$$

and hence $L_\psi(\gamma) = L_{\phi^{(2)}}(G(\gamma))$ for every C^1 curve γ in $\mathbb{P}_n$ and $\delta_\psi(A, B) = \delta_{\phi^{(2)}}(G(A), G(B))$ for every $A, B \in \mathbb{P}_n$. Hence the corollary follows from the equivalence of the equality versions of (i)–(iv) of Theorem 4.1 applied to $\phi^{(1)}$ and ψ. □

Remark 4.4. Let $M^{(1)}, M^{(2)} \in \mathfrak{M}_0$, $\theta_1, \theta_2 \in \mathbb{R}$ and $\phi^{(k)}(x, y) := M^{(k)}(x, y)^{\theta_k}$, $k = 1, 2$. Then it is easy to check that condition (i) (hence also (ii)–(iv)) of Theorem 4.1 is equivalent to

(i′) $\theta_1 = \theta_2 = 0$, or $\theta_1 = \theta_2 > 0$ and $M^{(1)}(x, 1) \le M^{(2)}(x, 1)$ for all $x > 0$, or $\theta_1 = \theta_2 < 0$ and $M^{(1)}(x, 1) \ge M^{(2)}(x, 1)$ for all $x > 0$.

Remark 4.5. Let $\mathcal{D}_n := \{D \in \mathbb{P}_n : \operatorname{Tr} D = 1\}$, a submanifold of $\mathbb{P}_n$, and $\phi^{(k)}$, $k = 1, 2$, be as in Remark 4.4. One can replace $(\mathbb{P}_n, \mathbb{H}_n)$ by $(\mathcal{D}_n, \mathbb{H}_n \ominus \mathbb{R}I)$ and slightly modify the proof of Theorem 4.1 to show that conditions (i)–(iv) of Theorem 4.1 are also equivalent to the following conditions reduced on $\mathcal{D}_n$:

(iii′) $L_{\phi^{(1)}}(\gamma) \ge L_{\phi^{(2)}}(\gamma)$ for all C^1 curve γ in $\mathcal{D}_n$;

(iv′) $\delta^{\mathcal{D}}_{\phi^{(1)}}(A, B) \ge \delta^{\mathcal{D}}_{\phi^{(2)}}(A, B)$ for all $A, B \in \mathcal{D}_n$, where $\delta^{\mathcal{D}}_\phi(A, B)$ denotes the geodesic distance in the Riemannian manifold $(\mathcal{D}_n, K^\phi)$.

By Theorem 2.1 and Remark 4.4 we have:

Corollary 4.6. *Let $M \in \mathfrak{M}_0$, $\theta \in \mathbb{R}$ and $\phi(x,y) := M(x,y)^\theta$, and let ϕ_θ be given in* (2.10)*. If $\phi(x,y) \le \phi_\theta(x,y)$ for all $x,y > 0$, then for every $A,B \in \mathbb{P}_n$,*

$$\delta_\phi(A,B) \ge \delta_{\phi_\theta}(A,B) = \begin{cases} \frac{2}{|2-\theta|}\|A^{\frac{2-\theta}{2}} - B^{\frac{2-\theta}{2}}\|_{\mathrm{HS}} & \text{if } \theta \ne 2, \\ \|\log A - \log B\|_{\mathrm{HS}} & \text{if } \theta = 2. \end{cases} \tag{4.2}$$

If $\phi(x,y) \ge \phi_\theta(x,y)$ for all $x,y > 0$, then the reversed inequality holds in (4.2)*.*

The next theorem is a refinement of Corollary 4.6 with strict inequality under additional assumptions.

Theorem 4.7. *Let M, θ, ϕ and ϕ_θ be as in Corollary 4.6. Assume that $A,B \in \mathbb{P}_n$ are not commuting, i.e., $AB \ne BA$. If $\phi(x,y) < \phi_\theta(x,y)$ for all $x,y > 0$ with $x \ne y$, then $\delta_\phi(A,B) > \delta_{\phi_\theta}(A,B)$. Similarly, $\delta_\phi(A,B) < \delta_{\phi_\theta}(A,B)$ if $\phi(x,y) > \phi_\theta(x,y)$ for all $x,y > 0$ with $x \ne y$.*

To prove the theorem, we need a simple lemma.

Lemma 4.8. *Let $\phi^{(k)}$, $k = 1,2$, be as in Remark 4.4, and assume that $\phi^{(1)}(x,y) < \phi^{(2)}(x,y)$ for all $x,y > 0$ with $x \ne y$. If $\gamma : [0,1] \to \mathbb{P}_n$ is a C^1 curve and $\gamma(t)\gamma'(t) \ne \gamma'(t)\gamma(t)$ for some $t \in [0,1]$, then $L_{\phi^{(1)}}(\gamma) > L_{\phi^{(2)}}(\gamma)$.*

Proof. It suffices to show that if $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$ are not commuting, then

$$\|\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}} > \|\phi^{(2)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}}.$$

To prove this, we may assume that $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$. Then $DH \ne HD$ means that $H_{ij} \ne 0$ for some (i,j) with $\lambda_i \ne \lambda_j$, where $H = [H_{ij}]$. Since $\phi^{(1)}(\lambda_i,\lambda_j) < \phi^{(2)}(\lambda_i,\lambda_j)$ for such (i,j), it obviously follows that

$$\begin{aligned} \|\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^n \frac{|H_{ij}|^2}{\phi^{(1)}(\lambda_i,\lambda_j)} > \sum_{i,j=1}^n \frac{|H_{ij}|^2}{\phi^{(2)}(\lambda_i,\lambda_j)} \\ &= \|\phi^{(2)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}}^2, \end{aligned}$$

as required. □

Proof of Theorem 4.7. Assume that $\phi(x,y) < \phi_\theta(x,y)$ for all $x \ne y$ and on the contrary that $\delta_\phi(A,B) = \delta_{\phi_\theta}(A,B)$. Choose a sequence $\{\gamma_k\}$ of C^1 curves from A to B such that $L_\phi(\gamma_k) \to \delta_{\phi_\theta}(A,B)$ as $k \to \infty$. The following proof is given in the case $\theta \ne 2$ but the case $\theta = 2$ is similar with obvious modifications. Let $\xi_k(t) := \gamma_k(t)^{\frac{2-\theta}{2}}$ for $0 \le t \le 1$. Since Remark 4.4 gives

$$\delta_{\phi_\theta}(A,B) \le L_{\phi_\theta}(\gamma_k) \le L_\phi(\gamma_k) \longrightarrow \delta_{\phi_\theta}(A,B)$$

so that by Theorem 2.1

$$\frac{|2-\theta|}{2}L_{\phi_\theta}(\gamma_k)=\int_0^1\|\xi_k'(t)\|_{\mathrm{HS}}\,dt\longrightarrow\|A^{\frac{2-\theta}{2}}-B^{\frac{2-\theta}{2}}\|_{\mathrm{HS}}\quad\text{as } k\to\infty.$$

By reparametrizing $\xi_k(t)$'s (hence $\gamma_k(t)$'s) one may assume that each ξ_k has a constant speed, i.e.,

$$\|\xi_k'(t)\|_{\mathrm{HS}}=\frac{|2-\theta|}{2}L_{\phi_\theta}(\gamma_k),\qquad 0\le t\le 1.$$

Set $\alpha:=\|A^{\frac{2-\theta}{2}}-B^{\frac{2-\theta}{2}}\|_{\mathrm{HS}}$ and $H_0:=\alpha^{-1}\big(B^{\frac{2-\theta}{2}}-A^{\frac{2-\theta}{2}}\big)$, a unit vector in $(\mathbb{H}_n,\langle\cdot,\cdot\rangle_{\mathrm{HS}})$. We notice

$$\begin{aligned}\int_0^1\left(1-\left\langle\frac{\xi_k'(t)}{\|\xi_k'(t)\|_{\mathrm{HS}}},H_0\right\rangle_{\mathrm{HS}}\right)dt&=1-\frac{2}{|2-\theta|L_{\phi_\theta}(\gamma_k)}\big\langle B^{\frac{2-\theta}{2}}-A^{\frac{2-\theta}{2}},H_0\big\rangle_{\mathrm{HS}}\\&=1-\frac{2\alpha}{|2-\theta|L_{\phi_\theta}(\gamma_k)}\longrightarrow 0\quad\text{as } k\to\infty.\end{aligned}$$

Hence, by taking a subsequence, one can assume that

$$\left\|\frac{\xi_k'(t)}{\|\xi_k'(t)\|_{\mathrm{HS}}}-H_0\right\|_{\mathrm{HS}}^2=2\left(1-\left\langle\frac{\xi_k'(t)}{\|\xi_k'(t)\|_{\mathrm{HS}}},H_0\right\rangle_{\mathrm{HS}}\right)\longrightarrow 0\quad\text{a.e. } t\in[0,1].$$

Since $\|\xi_k'(t)\|_{\mathrm{HS}}=2^{-1}|2-\theta|L_{\phi_\theta}(\gamma_k)\to\alpha$, this means that

$$\|\xi_k'(t)-\big(B^{\frac{2-\theta}{2}}-A^{\frac{2-\theta}{2}}\big)\|_{\mathrm{HS}}\longrightarrow 0\quad\text{a.e. } t\in[0,1],\tag{4.3}$$

which implies also that for every $0\le t\le 1$,

$$\begin{aligned}\|\xi_k(t)-\big((1-t)A^{\frac{2-\theta}{2}}+tB^{\frac{2-\theta}{2}}\big)\|_{\mathrm{HS}}&=\left\|\int_0^t\big(\xi_k'(s)-\big(B^{\frac{2-\theta}{2}}-A^{\frac{2-\theta}{2}}\big)\big)\,ds\right\|_{\mathrm{HS}}\\&\le\int_0^t\|\xi_k'(s)-\big(B^{\frac{2-\theta}{2}}-A^{\frac{2-\theta}{2}}\big)\|_{\mathrm{HS}}\,ds\longrightarrow 0.\end{aligned}\tag{4.4}$$

Now define $\xi_0(t):=(1-t)A^{\frac{2-\theta}{2}}+tB^{\frac{2-\theta}{2}}$ and $\gamma_0(t):=\xi_0(t)^{\frac{2}{2-\theta}}$. With $G_\theta(x):=x^{\frac{2}{2-\theta}}$ one can apply (2.8), (4.3) and (4.4) to obtain

$$\begin{aligned}&\|\phi(\mathbb{L}_{\gamma_0(t)},\mathbb{R}_{\gamma_0(t)})^{-1/2}\gamma_0'(t)\|_{\mathrm{HS}}\\&\quad=\|\phi(\mathbb{L}_{G_\theta(\xi_0(t))},\mathbb{R}_{G_\theta(\xi_0(t))})^{-1/2}G_\theta^{[1]}(\mathbb{L}_{\xi_0(t)},\mathbb{R}_{\xi_0(t)})\xi_0'(t)\|_{\mathrm{HS}}\\&\quad=\lim_{k\to\infty}\|\phi(\mathbb{L}_{G_\theta(\xi_k(t))},\mathbb{R}_{G_\theta(\xi_k(t))})^{-1/2}G_\theta^{[1]}(\mathbb{L}_{\xi_k(t)},\mathbb{R}_{\xi_k(t)})\xi_k'(t)\|_{\mathrm{HS}}\\&\quad=\lim_{k\to\infty}\|\phi(\mathbb{L}_{\gamma_k(t)},\mathbb{R}_{\gamma_k(t)})^{-1/2}\gamma_k'(t)\|_{\mathrm{HS}}\end{aligned}$$

for a.e. $t\in[0,1]$. Fatou's lemma yields

$$L_\phi(\gamma_0)\le\liminf_{k\to\infty}L_\phi(\gamma_k)=\delta_{\phi_\theta}(A,B)=L_{\phi_\theta}(\gamma_0)\tag{4.5}$$

thanks to Theorem 2.1. Here it is clear that $\xi_0(t)$ and $\xi_0'(t)$ are not commuting for any $0 \le t \le 1$. Hence $\gamma_0(t)$ and $\gamma_0'(t)$ never commute for $0 \le t \le 1$. In fact, this is seen because $\xi_0'(t)$ can be approximated by polynomials of $\gamma_0(t)$ and $\gamma_0'(t)$ thanks to (2.7) applied to $\xi_0(t) = G_\theta^{-1}(\gamma_0(t))$ so that $\gamma_0(t)\gamma_0'(t) = \gamma_0'(t)\gamma_0(t)$ implies $\xi_0(t)\xi_0'(t) = \xi_0'(t)\xi_0(t)$. Hence (4.5) contradicts the conclusion of Lemma 4.8.

The proof of the second assertion is easy. Assume that $\phi(x,y) > \phi_\theta(x,y)$ for all $x \ne y$, and let $\gamma_0(t)$ be same as in the proof of the first assertion. Since $\gamma_0(t)$ and $\gamma_0'(t)$ never commute for $0 \le t \le 1$ as mentioned above, Lemma 4.8 again implies that

$$\delta_\phi(A,B) \le L_\phi(\gamma_0) < L_{\phi_\theta}(\gamma_0) = \delta_{\phi_\theta}(A,B),$$

as required. □

The above proof of the first assertion is a bit involved. The proof would be much simpler if a geodesic shortest path joining A and B exists in $(\mathbb{P}_n, K^\phi)$, which is not known at the moment.

Example 4.9. The following are examples of the inequality given in Corollary 4.6 in the cases of familiar means. In fact, these are immediate consequences of Corollary 4.6 and Lemma 2.2 together with (2.11)–(2.14) and Lemma 2.5. Furthermore, Theorem 4.7 shows that all inequalities in the following become strict if A, B are not commuting and the respective closed range of θ is replaced by the open range.

(1) For the θ-power $M_{\mathrm{A}}^\theta(x,y) = \left(\frac{x+y}{2}\right)^\theta$ of the arithmetic mean,

$$\delta_{M_{\mathrm{A}}^\theta}(A,B) \begin{cases} \le \delta_{\phi_\theta}(A,B) & \text{if } \theta \le -2,\ \theta \ge 0, \\ \ge \delta_{\phi_\theta}(A,B) & \text{if } -2 \le \theta \le 0. \end{cases}$$

(2) For the θ-power $M_{\sqrt{}}^\theta(x,y) = \left(\frac{\sqrt{x}+\sqrt{y}}{2}\right)^{2\theta}$ of the root mean,

$$\delta_{M_{\sqrt{}}^\theta}(A,B) \begin{cases} \le \delta_{\phi_\theta}(A,B) & \text{if } \theta \le 0,\ \theta \ge 1, \\ \ge \delta_{\phi_\theta}(A,B) & \text{if } 0 \le \theta \le 1. \end{cases}$$

(3) For the θ-power $M_{\mathrm{L}}^\theta(x,y) = \left(\frac{x-y}{\log x - \log y}\right)^\theta$ of the logarithmic mean,

$$\delta_{M_{\mathrm{L}}^\theta}(A,B) \begin{cases} \le \delta_{\phi_\theta}(A,B) & \text{if } \theta \le 0,\ \theta \ge 2, \\ \ge \delta_{\phi_\theta}(A,B) & \text{if } 0 \le \theta \le 2. \end{cases}$$

(4) For the θ-power $M_{\mathrm{G}}^\theta(x,y) = (xy)^{\theta/2}$ of the geometric mean,

$$\delta_{M_{\mathrm{G}}^\theta}(A,B) \begin{cases} \le \delta_{\phi_\theta}(A,B) & \text{if } \theta \le 0,\ \theta \ge 4, \\ \ge \delta_{\phi_\theta}(A,B) & \text{if } 0 \le \theta \le 4. \end{cases}$$

(5) For the θ-power $M_{\rm H}^\theta(x,y)=\left(\frac{2xy}{x+y}\right)^\theta$ of the harmonic mean,

$$\delta_{M_{\rm H}^\theta}(A,B)\begin{cases}\le \delta_{\phi_\theta}(A,B) & \text{if } \theta\le 0,\\ \ge \delta_{\phi_\theta}(A,B) & \text{if } 0\le\theta\le 10.\end{cases}$$

For any $\theta\in\mathbb{R}$, $M_{\rm H}(x,1)<M_\theta(x,1)$ holds for large $x>0$ since $\lim_{x\to\infty}M_\theta(x,1)=+\infty$ while $\lim_{x\to\infty}M_{\rm H}(x,1)=2$. From this and Lemma 2.5 we observe that $\delta_{M_{\rm H}^\theta}(A,B)$ and $\delta_{\phi_\theta}(A,B)$ are not comparable when $\theta>10$.

In the case $\theta=2$ the above example (4) with (0.5) says that

$$\|\log(A^{-1/2}BA^{-1/2})\|_{\rm HS}\ge\|\log A-\log B\|_{\rm HS},\qquad A,B\in\mathbb{P}_n.$$

This is the so-called *exponential metric increasing* (*EMI*) property discussed in [7, 9], whose former description is found in [30, §3]. On the other hand, for instance, (1) says that

$$\delta_{M_{\rm A}^2}(A,B)\le\|\log A-\log B\|_{\rm HS},\qquad A,B\in\mathbb{P}_n,$$

which may be called the "exponential metric decreasing" property. In the case $\theta=1$ the above examples give

$$\delta_{M_{\rm H}}(A,B)\ge\delta_{M_{\rm G}}(A,B)\ge\delta_{M_{\rm L}}(A,B)\ge 2\|A^{1/2}-B^{1/2}\|_{\rm HS}\ge\delta_{M_{\rm A}}(A,B),$$

which may be called the "square metric increasing/decreasing" properties.

In the particular case where $\phi(x,y)=M(x,y)$ (of degree $\theta=1$) is an operator monotone mean, i.e., $M(x,1)$ is a standard operator monotone function and moreover A,B are commuting, the next theorem gives the exact formula for $\delta_M(A,B)$ independently of the choice of M. It seems that this independence of M is reflected by the uniqueness of a monotone Riemannian metric in the classical case (see [35]).

Theorem 4.10. *Let $M\in\mathfrak{M}_0$ and assume that $M(x,1)$ is an operator monotone function. If $A,B\in\mathbb{P}_n$ are commuting, then*

$$\delta_M(A,B)=2\|A^{1/2}-B^{1/2}\|_{\rm HS},$$

and a geodesic shortest curve from A to B is given by

$$\gamma_{A,B}(t):=\big((1-t)A^{1/2}+tB^{1/2}\big)^2,\qquad 0\le t\le 1,\tag{4.6}$$

independently of the choice of M as above. Furthermore, this $\gamma_{A,B}$ is a unique (up to parametrization) geodesic shortest curve from A to B whenever $M\ne M_{\rm A}$, the arithmetic mean.

First we recall that the assumption of the operator monotonicity of $M(x,1)$ implies that K^M is a monotone metric [35] (see also Introduction). Without this assumption we give a small lemma.

Lemma 4.11. *Assume that* $\gamma : [0,1] \to \mathbb{P}_n$ *is a* C^1 *curve and* $\gamma(t)\gamma'(t) = \gamma'(t)\gamma(t)$ *for all* $t \in [0,1]$. *Let* $\xi(t) := \gamma(t)^{1/2}$. *Then* $L_M(\gamma) = 2\int_0^1 \|\xi'(t)\|_{\mathrm{HS}}\,dt$ *for all* $M \in \mathfrak{M}_0$.

Proof. Since $M(x,x) = x$ for all $x > 0$, we note that $\|M(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H\|_{\mathrm{HS}}$ is independent of the choice of M whenever $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$ are commuting. This implies that $L_M(\gamma)$ is independent of $M \in \mathfrak{M}_0$ if γ is as stated in the lemma. Hence the $\theta = 1$ case of Theorem 2.1 implies that

$$L_M(\gamma) = L_{M_{\sqrt{}}}(\gamma) = 2\int_0^1 \|\xi'(t)\|_{\mathrm{HS}}\,dt,$$

where $M_{\sqrt{}}$ is the root mean given in (2.12). □

Proof of Theorem 4.10. Assume that $AB = BA$, and let $\gamma_{A,B}$ be as given in the theorem. By Lemma 4.11 and Theorem 2.1 we have

$$L_M(\gamma_{A,B}) = L_{M_{\sqrt{}}}(\gamma_{A,B}) = 2\|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}$$

so that $\delta_M(A,B) \le 2\|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}$. To prove the converse, let Φ denote the conditional expectation (with respect to Tr) of $\mathbb{M}_n$ onto the commutative subalgebra generated by A, B, and let $\gamma : [0,1] \to \mathbb{P}_n$ be an arbitrary C^1 curve from A to B. Then $\Phi(\gamma)$ is a C^1 curve in $\mathbb{P}_n$ from A to B. Since K^M is a monotone metric as mentioned before Lemma 4.11, we have

$$K^M_{\Phi(\gamma(t))}(\Phi(\gamma'(t)), \Phi(\gamma'(t))) \le K^M_{\gamma(t)}(\gamma'(t), \gamma'(t)), \qquad 0 \le t \le 1,$$

so that $L_\phi(\Phi(\gamma)) \le L_\phi(\gamma)$. Therefore we may assume that $\gamma(t)$'s are in a commutative subalgebra. When $\xi(t) := \gamma(t)^{1/2}$, it then follows from Lemma 4.11 that

$$L_M(\gamma) = 2\int_0^1 \|\xi'(t)\|_{\mathrm{HS}}\,dt \ge 2\|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}. \tag{4.7}$$

Hence $\delta_\phi(A,B) = 2\|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}$ and $\gamma_{A,B}$ is a common geodesic shortest curve from A to B for all metrics K^M with operator monotone M.

Next we show the last assertion on the uniqueness of a geodesic curve. Since M_{A} is the largest standard operator monotone function and $M \ne M_{\mathrm{A}}$, we note that $M(x,1) < M_{\mathrm{A}}(x,1)$ for all $x > 0$ with $x \ne 1$. In fact, if $M(x_0,1) = M_{\mathrm{A}}(x_0,1)$ for some $x_0 > 0$ with $x_0 \ne 1$, then the concavity of $M(x,1)$ implies that $M(x,1) = M_{\mathrm{A}}(x,1)$ for all x between 1 and x_0. Hence the analyticity of $M(x,1)$ gives $M = M_{\mathrm{A}}$. Now let $\gamma_1 : [0,1] \to \mathbb{P}_n$ be a C^1 curve from A to B such that $L_M(\gamma_1) = 2\|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}$. Since $L_M(\gamma_1) \ge L_{M_{\mathrm{A}}}(\gamma_1)$ by Remark 4.4 and $L_{M_{\mathrm{A}}}(\gamma_1) \ge 2\|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}$ by (4.7), we have $L_M(\gamma_1) = L_{M_{\mathrm{A}}}(\gamma_1)$. Hence it follows from Lemma 4.8 that $\gamma_1(t)\gamma_1'(t) = \gamma_1'(t)\gamma_1(t)$ for all $t \in [0,1]$. Lemma 4.11 in turn implies that $\int_0^1 \|\xi_1'(t)\|_{\mathrm{HS}}\,dt = \|A^{1/2} - B^{1/2}\|_{\mathrm{HS}}$, where $\xi_1(t) := \gamma_1(t)^{1/2}$. Therefore we have $\xi_1(t) = (1-t)A^{1/2} + tB^{1/2}$, $0 \le t \le 1$, so that $\gamma_1 = \gamma_{A,B}$. □

When $M = M_{\mathrm{A}}$ and A, B are commuting, it is not known whether $\gamma_{A,B}$ given in (4.6) is a unique geodesic shortest path joining A, B. To prove this, we probably need to examine the equality case in the monotonicity of $K_D^M(H,H)$ under conditional expectation. Another problem for commuting A, B is whether $\gamma_{A,B}$ gives a geodesic shortest path for any metric K^M with $M \in \mathfrak{M}_0$ that is not necessarily operator monotone.

We close the section with a result on comparison of skew informations given in (0.9) though apart from Riemannian metrics.

Proposition 4.12. *Let f and g be two standard operator monotone functions that are regular, i.e., $f(0), g(0) > 0$. Then $I_D^f(K) \ge I_D^g(K)$ for all $D \in \mathbb{P}_n$ and $K \in \mathbb{H}_n$ if and only if $f(0)/f(x) \ge g(0)/g(x)$ for all $x > 0$.*

Proof. Assume that $I_D^f(K) \ge I_D^g(K)$ for all $D \in \mathbb{P}_n$ and $K \in \mathbb{H}_n$. From the expression in (0.10) for diagonal D, it follows that

$$f(0)\,\frac{(x-y)^2}{yf(x/y)} \ge g(0)\,\frac{(x-y)^2}{yg(x/y)}, \qquad x, y > 0,$$

since f, g are standard so that $yf(x/y) = xf(y/x)$ and $yg(x/y) = xg(y/x)$. Hence we have $f(0)/f(x) \ge g(0)/g(x)$ for all $x > 0$. The converse implication is directly seen from definition (0.9). $\square$

For example, as for $f_p = f_{1-p}$, $0 < p \le 1/2$, given in (0.8), $f_p(0)/f_p(x)$ is increasing in $p \in (0, 1/2]$ so that the Wigner-Yanase-Dyson skew information $I_D^{\mathrm{WYD}}(p, K)$ is increasing in $p \in (0, 1/2]$ for fixed D and K. (See [5] for this and a comparison result of different type for quantum skew informations).

5 Unitarily invariant norms

Let $|||\cdot|||$ be a *unitarily invariant norm* on matrices, that is, $|||\cdot|||$ is a norm on $\mathbb{M}_n$, $n \in \mathbb{N}$, such that $|||UXV||| = |||X|||$ for all $X, U, V \in \mathbb{M}_n$ with U, V unitaries. The Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}}$ is a special example of such norms. When a smooth symmetric kernel function $\phi : (0,\infty) \times (0,\infty) \to (0,\infty)$ is given, replacing $\|\cdot\|_{\mathrm{HS}}$ by $|||\cdot|||$ in (1.3) we define the *length*

$$L_{\phi,|||\cdot|||}(\gamma) := \int_0^1 |||\phi(\mathbb{L}_{\gamma(t)}, \mathbb{R}_{\gamma(t)})^{-1/2}\gamma'(t)|||\,dt$$

of a C^1 curve $\gamma : [0,1] \to \mathbb{P}_n$. The *distance* $\delta_{\phi,|||\cdot|||}(A,B)$ between $A, B \in \mathbb{P}_n$ is the infimum of $L_{\phi,|||\cdot|||}(\gamma)$ over all C^1 curves γ from A to B. The manifold $\mathbb{P}_n$ with the distance $\delta_{\phi,|||\cdot|||}$ is no longer a Riemannian manifold but a certain manifold of Finsler type. Such manifolds have been studied by several authors, for example, see [11, 12] for the operator norm (in a C^*-algebra) and [13] for unitarily invariant norms.

In this section we show that many results in the previous sections hold true even when the Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}}$ is replaced by a general unitarily invariant norm $|||\cdot|||$. First, Theorem 2.1 can be extended as follows. We omit the proof that is essentially same as the second proof of Theorem 2.1.

Proposition 5.1. *Let $|||\cdot|||$ be any unitarily invariant norm. Let M, θ, ϕ and F be as in Theorem 2.1. Then the transformation $D \in \mathbb{P}_n \mapsto F(D) \in \mathbb{H}_n$ is isometric from $(\mathbb{P}_n, \delta_{\phi,|||\cdot|||})$ into $(\mathbb{H}_n, |||\cdot|||)$ if and only if F is in the form* (2.1) *and $M = M_\theta$ (so $\phi = \phi_\theta$). Moreover, for every $A, B \in \mathbb{P}_n$,*

$$\delta_{\phi_\theta,|||\cdot|||}(A,B) = \begin{cases} \frac{2}{|2-\theta|}|||A^{\frac{2-\theta}{2}} - B^{\frac{2-\theta}{2}}||| & \text{if } \theta \neq 2, \\ |||\log A - \log B||| & \text{if } \theta = 2, \end{cases} \tag{5.1}$$

and this distance is attained by the curve given in (2.3)*.*

Remark 5.2. Although the distance in $(\mathbb{P}_n, \delta_{\phi_\theta,|||\cdot|||})$ is determined as in (5.1), a curve of shortest length is not necessarily unique (up to parametrization) in Proposition 5.1 unlike Theorem 2.1 in the Riemannian metric case. For example, when $|||\cdot|||$ is the p-norm $\|H\|_p = (\mathrm{Tr}\,|H|^p)^{1/p}$, a curve of shortest length in $(\mathbb{P}_n, \delta_{\phi_\theta,\|\cdot\|_p})$ joining any two points is unique. When $|||\cdot|||$ is the trace-norm $\|\cdot\|_1$, let $A := \mathrm{Diag}(1,2)$ and $B := \mathrm{Diag}(2,1)$. Then the distance $\delta_{\|\cdot\|_1}(A,B)$ in $(\mathbb{P}_n, \|\cdot\|_1)$ is $\|A-B\|_1 = 2$. For each increasing C^1 functions $f, g : [0,1] \to [0,1]$ with $f(0) = g(0) = 0$ and $f(1) = g(1) = 1$, define $\gamma(t) := \mathrm{Diag}(1+f(t), 2-g(t))$ for $0 \le t \le 1$. Then γ is a C^1 curve in $\mathbb{P}_n$ joining A, B, and the length of γ in $(\mathbb{P}_n, \|\cdot\|_1)$ is

$$L_{\|\cdot\|_1}(\gamma) = \int_0^1 \|\gamma'(t)\|_1\,dt = \int_0^1 (f'(t) + g'(t))\,dt = 2.$$

So infinitely many curves of shortest length exist in $(\mathbb{P}_n, \delta_{\|\cdot\|_1})$ and hence in $(\mathbb{P}_n, \delta_{\phi_\theta,\|\cdot\|_1})$ for each θ by Proposition 5.1. The situation is same for the operator norm $\|\cdot\|_\infty$ as well. The existence of such infinitely many curves of shortest length is a typical non-Riemannian phenomenon as was exemplified in [31, Lemma 1.3] in a convex cone of a very general nature.

The next comparison property is a partial extension of Theorem 4.1 (Remark 4.4). An essential point of the proof is similar to that of [17, Theorem 1.1].

Proposition 5.3. *Let $M^{(1)}, M^{(2)} \in \mathfrak{M}_0$, $\theta \in \mathbb{R}$ and $\phi^{(k)}(x,y) := M_k(x,y)^\theta$, $k = 1,2$. Then the following conditions are equivalent:*

(i) *$(M^{(1)}(e^t,1)/M^{(2)}(e^t,1))^{\theta/2}$ is a positive definite function on $\mathbb{R}$;*

(ii) *$L_{\phi^{(1)},|||\cdot|||}(\gamma) \ge L_{\phi^{(2)},|||\cdot|||}(\gamma)$ for all C^1 curve γ in $\mathbb{P}_n$ and for any unitarily invariant norm $|||\cdot|||$;*

(iii) *$L_{\phi^{(1)},\|\cdot\|_\infty}(\gamma) \ge L_{\phi^{(2)},\|\cdot\|_\infty}(\gamma)$ for all C^1 curve γ in $\mathbb{P}_n$ and for the operator norm $\|\cdot\|_\infty$.*

Proof. (i) $\Rightarrow$ (ii). It suffices to show that (i) implies that

$$|||\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H||| \ge |||\phi^{(2)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2}H||| \tag{5.2}$$

for all $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$. To do this, one may assume that $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$. By (1.2) notice that

$$\phi^{(2)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H = \left[\left(\frac{\phi^{(1)}(\lambda_i, \lambda_j)}{\phi^{(2)}(\lambda_i, \lambda_j)} \right)^{1/2} \right]_{ij} \circ (\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H)$$

and

$$\left(\frac{\phi^{(1)}(\lambda_i, \lambda_j)}{\phi^{(2)}(\lambda_i, \lambda_j)} \right)^{1/2} = \left(\frac{M^{(1)}(\lambda_i/\lambda_j, 1)}{M^{(2)}(\lambda_i/\lambda_j, 1)} \right)^{\theta/2} = \left(\frac{M^{(1)}(e^{\log \lambda_i - \log \lambda_j}, 1)}{M^{(2)}(e^{\log \lambda_i - \log \lambda_j}, 1)} \right)^{\theta/2}. \tag{5.3}$$

Since (i) implies that $\left[(\phi^{(1)}(\lambda_i, \lambda_j)/\phi^{(2)}(\lambda_i, \lambda_j)^{1/2} \right]_{ij}$ is a positive definite matrix with all diagonal entries equal to 1, (5.2) is obtained (see [8, 1.4.1] for example).

(ii) $\Rightarrow$ (iii) is trivial.

(iii) $\Rightarrow$ (i). For $k = 1, 2$, since $D \in \mathbb{P}_n \mapsto \phi^{(k)}(\mathbb{L}_D, \mathbb{R}_D)$ is continuous, it is obvious that

$$\lim_{\varepsilon \searrow 0} \frac{L_{\phi^{(k)}, \|\cdot\|_\infty}([D, D + \varepsilon H])}{\varepsilon} = \|\phi^{(k)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H\|_\infty$$

for all $D \in \mathbb{P}_n$ and $H \in \mathbb{H}_n$, where $[D, D + \varepsilon H]$ denotes the straight segment $D + t\varepsilon H$, $0 \le t \le 1$. Hence condition (iii) implies that

$$\|\phi^{(1)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H\|_\infty \ge \|\phi^{(2)}(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H\|_\infty, \qquad D \in \mathbb{P}_n, \ H \in \mathbb{H}_n.$$

When $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$, this means that

$$\|H\|_\infty \ge \left\| \left[\left(\frac{\phi^{(1)}(\lambda_i, \lambda_j)}{\phi^{(2)}(\lambda_i, \lambda_j)} \right)^{1/2} \right]_{ij} \circ H \right\|_\infty, \qquad H \in \mathbb{H}_n.$$

Now the proof of [17, Theorem 1.1] shows that $\left[(\phi^{(1)}(\lambda_i, \lambda_j)/\phi^{(2)}(\lambda_i, \lambda_j))^{1/2} \right]_{ij}$ is positive semidefinite, which means (i) thanks to (5.3). $\square$

Remark 5.4. The geodesic distance versions of the above (ii) and (iii) are

(iv) $\delta_{\phi^{(1)}, |||\cdot|||}(A, B) \ge \delta_{\phi^{(2)}, |||\cdot|||}(A, B)$ for all $A, B \in \mathbb{P}_n$ and for any unitarily invariant norm $|||\cdot|||$;

(v) $\delta_{\phi^{(1)}, \|\cdot\|_\infty}(A, B) \ge \delta_{\phi^{(2)}, \|\cdot\|_\infty}(A, B)$ for all $A, B \in \mathbb{P}_n$.

It is obvious that (ii) $\Rightarrow$ (iv) and (iii) $\Rightarrow$ (v). It may be expected that (iv) and (v) are also equivalent to the conditions of Proposition 5.3. This would be proved as in the proof of (iii) $\Rightarrow$ (i) if we have

$$\lim_{\varepsilon \searrow 0} \frac{\delta_{\phi, \|\cdot\|_\infty}(D, D + \varepsilon H)}{\varepsilon} = \|\phi(\mathbb{L}_D, \mathbb{R}_D)^{-1/2} H\|_\infty$$

for all $D \in \mathbb{P}_n$, $H \in \mathbb{H}_n$ and for $\phi = M^\theta$ with $M \in \mathfrak{M}_0$. Although the above convergence for $\|\cdot\|_{\mathrm{HS}}$ is in Lemma 4.2, we do not know whether it is also true for $\|\cdot\|_\infty$.

For $M^{(1)}, M^{(2)} \in \mathfrak{M}_0$ consider the following conditions:

(a) $M^{(1)}(x,1) \le M^{(2)}(x,1)$ for all $x > 0$;

(b) $M^{(1)}(e^t,1)/M^{(2)}(e^t,1)$ is positive definite on $\mathbb{R}$ (in this case we write $M^{(1)} \preceq M^{(2)}$);

(c) $M^{(1)}(e^t,1)/M^{(2)}(e^t,1)$ is infinitely divisible in the sense that

$$\left(\frac{M^{(1)}(e^t,1)}{M^{(2)}(e^t,1)}\right)^r$$

is positive definite on $\mathbb{R}$ for any $r > 0$ (in this case we write $M^{(1)} \ll M^{(2)}$).

Obviously, (c) $\Rightarrow$ (b) $\Rightarrow$ (a). Condition (a) appeared in Remark 4.4 while (b) is in the case $\theta = 2$ of Proposition 5.3. We also note that (b) played an essential role in [17, 18]. It was recently observed in [10, 22] that the stronger condition (c) is even satisfied for many cases where $M^{(1)}, M^{(2)} \in \mathfrak{M}_0$ satisfy (b). In fact, Kosaki [24] communicated to us that

$$M_{\mathrm{H}} \ll M_{\mathrm{G}} \ll M_{\mathrm{L}} \ll M_{\sqrt{\ }} \ll M_{\mathrm{A}}$$

can be easily shown by applying [22, Corollary 3] and [10, Proposition 4]. Hence by Proposition 5.3 (also Remark 5.4), if $\theta \ge 0$ then

$$\delta_{M_{\mathrm{H}}^\theta,|||\cdot|||}(A,B) \ge \delta_{M_{\mathrm{G}}^\theta,|||\cdot|||}(A,B) \ge \delta_{M_{\mathrm{L}}^\theta,|||\cdot|||}(A,B) \ge \delta_{M_{\sqrt{\ }}^\theta,|||\cdot|||}(A,B) \ge \delta_{M_{\mathrm{A}}^\theta,|||\cdot|||}(A,B),$$

and inequalities are reversed if $\theta \le 0$. For $\{N_\alpha\}_{0\le\alpha\le2}$ given in (3.1), if $0 \le \alpha < \beta \le 2$ then we have $N_\beta \ll N_\alpha$ by [10, Theorem 2] since

$$\frac{N_\beta(e^{2t},1)}{N_\alpha(e^{2t},1)} = \frac{\beta}{\alpha}\cdot\frac{\sinh\alpha t}{\sinh\beta t}.$$

As for $\psi_\alpha = N_\alpha^2$ given in (3.4), similarly to Theorem 3.3 we have

$$\delta_{\psi_\alpha,|||\cdot|||}(A,B) = |||\log(A^{-\alpha/2}B^\alpha A^{-\alpha/2})^{1/\alpha}|||, \qquad 0 < \alpha \le 2,$$

which decreases to $\delta_{M_{\mathrm{L}}^2,|||\cdot|||}(A,B) = |||\log A - \log B|||$ as $\alpha \searrow 0$ (this is also a consequence of Araki's log-majorization [4] as mentioned at the end of Section 3). In particular, the inequality

$$\delta_{M_{\mathrm{G}}^2,|||\cdot|||}(A,B) = |||\log(A^{-1/2}BA^{-1/2})||| \ge |||\log A - \log B|||$$

is the generalized EMI, which was first proved in [11] for the operator norm (in a C^*-algebra) and later in [7] for unitarily invariant norms.

Finally, as for ϕ_θ given in (2.10) we show:

Proposition 5.5. *Let* $|||\cdot|||$ *be any unitarily invariant norm and* $A, B \in \mathbb{P}_n$*. Then* $\delta_{\phi_\theta,|||\cdot|||}(A,B)$ *given in* (5.1) *is decreasing in* $\theta \in (-\infty, 2]$ *and increasing in* $\theta \in [2, \infty)$*. Furthermore,*

$$\delta_{M_G^\theta,|||\cdot|||}(A,B) \begin{cases} \leq \delta_{\phi_\theta,|||\cdot|||}(A,B) & \text{if } \theta \leq 0,\ \theta \geq 4, \\ \geq \delta_{\phi_\theta,|||\cdot|||}(A,B) & \text{if } 0 \leq \theta \leq 4. \end{cases}$$

Proof. Assume that $\theta' < \theta < 2$ or $2 < \theta < \theta'$, and define a kernel function $k : (0,\infty) \times (0,\infty) \to (0,\infty)$ by

$$k(x,y) := \frac{2-\theta'}{2-\theta} \cdot \frac{x^{\frac{2-\theta}{2}} - y^{\frac{2-\theta}{2}}}{x^{\frac{2-\theta'}{2}} - y^{\frac{2-\theta'}{2}}}.$$

The kernel $k(x,y)$ is positive definite (even infinitely divisible) by [10, Theorem 2] and $k(x,x) = 1$ for all $x > 0$. With the diagonalizations $A = U\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)U^*$ and $B = V\mathrm{Diag}(\mu_1, \ldots, \mu_n)V^*$ we write

$$\begin{aligned} \frac{2}{2-\theta}\Big(A^{\frac{2-\theta}{2}} - B^{\frac{2-\theta}{2}}\Big) &= U\left(\frac{2}{2-\theta}\left[\lambda_i^{\frac{2-\theta}{2}} - \mu_j^{\frac{2-\theta}{2}}\right]_{ij} \circ (U^*V)\right)V^* \\ &= U\left(\big[k(\lambda_i,\mu_j)\big]_{ij} \circ \frac{2}{2-\theta'}\left[\lambda_i^{\frac{2-\theta'}{2}} - \mu_j^{\frac{2-\theta'}{2}}\right]_{ij} \circ (U^*V)\right)V^* \\ &= U\left(\big[k(\lambda_i,\mu_j)\big]_{ij} \circ U^*\frac{2}{2-\theta'}\Big(A^{\frac{2-\theta'}{2}} - B^{\frac{2-\theta'}{2}}\Big)V\right)V^*. \end{aligned}$$

Hence [8, 1.4.1] can be applied to obtain $\delta_{\phi_\theta,|||\cdot|||}(A,B) \leq \delta_{\phi_{\theta'},|||\cdot|||}(A,B)$ thanks to Proposition 5.1.

The second assertion (extending (4) of Example 4.9) follows since $M_\theta \ll M_G$ for $\theta \geq 4$ and $M_G \ll M_\theta$ for $\theta \leq 4$ (see [10, §2.6]). □

Acknowledgments

This work is partially supported by the Hungarian Research Grant OTKA T068258 (D.P.) and Grant-in-Aid for Scientific Research (B)17340043 (F.H.) as well as by Hungary-Japan HAS-JSPS Joint Project (D.P. & F.H.). D.P. thanks to Professors Peter Michor and Gabor Toth for communication about Riemannian metrics. F.H. thanks to Professor Hideki Kosaki for communication about operator monotonicity and infinite divisibility for means. The authors are also grateful to Professor Tsuyoshi Ando and two anonymous referees for valuable comments that helped to improve the final version of the paper.

References

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* **26** (1979) 203–241.

[2] T. Ando and F. Hiai, Log majorization and complementary Golden-Thompson type inequalities, *Linear Algebra Appl.* **197/198** (1994), 113–131.

[3] T. Ando, C.-K. Li and R. Mathias, Geometric means, *Linear Algebra Appl.* **385** (2004) 305–334.

[4] H. Araki, On an inequality of Lieb and Thirring, *Lett. Math. Phys.* **19** (1990), 167–170.

[5] K. Audenaert, L. Cai and F. Hansen, Inequalities for quantum skew information, *Lett. Math. Phys.* **85** (2008), 135–146.

[6] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1996.

[7] R. Bhatia, On the exponential metric increasing property, *Linear Algebra Appl.* **375** (2003), 211–220.

[8] R. Bhatia, *Positive Definite Matrices*, Princeton Univ. Press, Princeton, 2007.

[9] R. Bhatia and J. A. R. Holbrook, Riemannian geometry and matrix geometric means, *Linear Algebra. Appl.* **423** (2006), 594–618.

[10] R. Bhatia and H. Kosaki, Mean matrices and infinite divisibility, *Linear Algebra. Appl.* **424** (2007), 36–54.

[11] G. Corach, H. Porta and L. Recht, A geometric interpretation of Segal's inequality $\|e^{X+Y}\| \le \|e^{X/2}e^Y e^{X/2}\|$, *Proc. Amer. Math. Soc.* **115** (1992), 229–231.

[12] G. Corach, H. Porta and L. Recht, Geodesics and operator means in the space of positive operators, *Internat. J. Math.* **4** (1993), 193–202.

[13] J. I. Fujii, Means and metrics for positive operators with unitarily invariant norms, Preprint, 2008.

[14] J. I. Fujii, M. Fujii, T. Miura, H. Takagi and S.-E. Takahasi, Continuously differentiable means, *J. Inequal. Appl.* **2006**, Art. ID 75941, 15 pp.

[15] P. Gibilisco and T. Isola, A characterization of Wigner-Yanase skew information among statistically monotone metrics, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **4** (2001), 553–557.

[16] F. Hansen, Metric adjusted skew information, *Proc. Nat. Acad. Sci. USA* **105** (2008), 9909–9916.

[17] F. Hiai and H. Kosaki, Means for matrices and comparison of their norms, *Indiana Univ. Math. J.* **48** (1999), 899–936

[18] F. Hiai and H. Kosaki, *Means of Hilbert Space Operators*, Lecture Notes in Math. **1820**, Springer-Verlag, 2003.

[19] F. Hiai and D. Petz, The Golden-Thompson trace inequality is complemented, *Linear Algebra Appl.* **181** (1993), 153–185.

[20] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.

[21] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. 1, Wiley Interscience, New York-London, 1963.

[22] H. Kosaki, On infinite divisibility of positive definite functions arising from operator means, *J. Funct. Anal.* **254** (2008) 84–108.

[23] H. Kosaki, Notes on the Stolarsky means, Notes, 2008.

[24] H. Kosaki, Private communication, 2008.

[25] F. Kubo and T. Ando, Means of positive linear operator*Math. Ann.* **246** (1980), 205–224.

[26] J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Amer. Math. Monthly* **108** (2001) 797–812.

[27] A. Lesniewski and M. B. Ruskai, Monotone Riemannian metrics and relative entropy on noncommutative probability spaces, *J. Math. Phys.* **40** (1999), 5702–5724.

[28] P. W. Michor, *Topics in Differential Geometry*, Amer. Math. Soc., Providence, RI, 2008.

[29] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive definite matrices, *SIAM J. Matrix Anal. Appl.* **26** (2005), 735–747.

[30] G. D. Mostow, *Strong Rigidity of Locally Symmetric Spaces*, Annals Math. Studies 78, Princeton University Press, Princeton, 1973.

[31] R. D. Nussbaum, Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations, *Differential Integral Equations* **7** (1994), 1649–1707.

[32] A. Ohara, N. Suda and S. Amari, Dualistic differential geometry of positive definite matrices and its applications to related problems, *Linear Algebra Appl.* **247** (1996), 31–53.

[33] D. Petz, Quasi-entropies for states of a von Neumann algebra, *Publ. Res. Inst. Math. Sci.* **21** (1985), 787–800.

[34] D. Petz, Quasi-entropies for finite quantum systems, *Rep. Math. Phys.* **21** (1986), 57–65.

[35] D. Petz, Monotone metrics on matrix spaces, *Linear Algebra Appl.* **244** (1996), 81–96.

[36] D. Petz, Covariance and Fisher information in quantum mechanics, *J. Phys. A: Math. Gen.* **35** (2002), 1–11.

[37] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer-Verlag, 2008.

[38] D. Petz and H. Hasegawa, On the Riemannian metric of α-entropies of density matrices, *Lett. Math. Phys.* **38** (1996), 221–225.

[39] D. Petz and C. Sudár, Geometries of quantum states, *J. Math. Phys.* **37** (1996), 2662–2673.

[40] W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, *Rep. Math. Phys.* **8** (1975) 159–170.

[41] L. T. Skovgaard, A Riemannian geometry of the multivariate normal model, *Scand. J. Statistics*, **11** (1984), 211–223.

[42] K. B. Stolarsky, Generalizations of the logarithmic mean, *Math. Mag.* **48** (1975), 87–92.