Monotone metrics on statistical manifolds of density matrices by geometry of non-commutative L^2-spaces

Paolo Gibilisco[1], Tommaso Isola[2]

[1]*Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, I–10129 Torino, Italy, and Centro Vito Volterra, Università di Roma "Tor Vergata".*
[2]*Dipartimento di Matematica, Università di Roma "Tor Vergata", via della Ricerca Scientifica, I–00133 Roma, Italy.*

Abstract. Using an integral decomposition of non-commutative monotone metrics we show that each monotone metric described by Petz classification theorem is related to the geometry of a suitable non-commutative L^2-space. This exactly reproduces and generalizes the commutative case where the unique monotone metric (Chentsov theorem about Fisher-Rao metric) is classically related to the commutative L^2-geometry.

INTRODUCTION

The concept of monotone metric for parametric statistical manifolds has been introduced by Chentsov [3,4] and further developed by Petz [14]. There are two fundamental results: i) the Chentsov uniqueness theorem [2,3]; ii) the Petz classification theorem [14]. The first theorem says that in the commutative case there exists a unique monotone metric (up to a scalar factor) and that this metric coincides with the well-known Fisher-Rao metric. In the non-commutative case the situation is, as usual, more complicated and richer. This means that there is no uniqueness and that we have an infinite family of different monotone metrics. The classification theorem by Petz shows that there is a natural bijection between the family of monotone metrics and the family of operator monotone functions.
It is well-known that the Fisher-Rao metric can be related to the geometry of commutative L^2-spaces [5].The purpose of this paper is to answer to the following question: which non-commutative monotone metrics arise by geometry of non-commutative L^2-spaces (following the line of the commutative case)? Such a question is relevant for at least three reasons: i) it is natural to ask which features of the commutative case survive in the non-commutative one; ii) to interpret some norms and scalar products as L^2-norms and scalar products opened the way to the

1) E-mail: gibilisco@polito.it, isola@mat.uniroma2.it

recently established non-parametric version of information geometry [6,16,17]; iii) it has been proved [6,8] in the commutative case that α-geometries for $\alpha \in (-1, 1)$ are related to the geometry of L^p-spaces where $p = \frac{2}{1-\alpha}$. The classic work of Amari shows that the riemannian Fisher-Rao metric induces the 0-connection. It is reasonable therefore to hope that one may associate, using suitable non-commutative L^p-spaces, a family of α-connections to those non-commutative monotone metrics that are associated to L^2-type metrics [8].

The purpose of this note is to show that there is a general positive answer, given by Theorem 12, to the above formulated question. This means the following. Let $\mathcal{P}_n$ be the probability simplex in $\mathbf{R}^n$ and $T_p\mathcal{P}_n$ be the tangent space at $p \in \mathcal{P}_n$, and let us identify $\mathbf{R}^n$ with $L^2(m)$, where m is the counting measure on $\{1, \ldots, n\}$. If one defines $M_p(v)_i := p_i v_i$, $p \in \mathcal{P}_n$, $v \in T_p\mathcal{P}_n$, then Chentsov Theorem can be rephrased by saying that "each" monotone metric turns into an isometry the linear map $v \in T_p\mathcal{P}_n \to M_p^{-1/2}(v) \in L^2(m)$, and that this property characterises monotone metrics.

Now let $\mathcal{D}_n$ be the manifold of invertible density matrices, so that $T_\rho\mathcal{D}_n$, $\rho \in \mathcal{D}_n$, is the space of hermitian, traceless matrices, and denote by $L^2(\tau)$ the Hilbert space of all n by n matrices endowed with the scalar product given by the normalised trace τ. Define $L_\rho(A) := \rho A$, $R_\rho(A) := A\rho$, and for any symmetric measure μ on $[0, 1]$ define $M_{\rho,\mu}^{-1/2} : T_\rho\mathcal{D}_n \to \mathcal{H}_\mu := L^2([0, 1], \mu; L^2(\tau))$, by $M_{\rho,\mu}^{-1/2} := \int_{[0,1]}^{\oplus}((1-s)L_\rho + sR_\rho)^{-1/2} d\mu(s)$. Then Theorem 12 shows that each noncommutative monotone metric turns into an isometry the linear map $A \in T_\rho\mathcal{D}_n \to M_{\rho,\mu}^{-1/2}(A) \in \mathcal{H}_\mu$, by a suitable measure μ, and that this property characterises monotone metrics. In this sense $M_{\rho,\mu}^{-1/2}$ appears to be a noncommutative analogue of the division by the square root, and $L^2([0, 1], \mu; L^2(\tau))$ as an analogue of $L^2(m)$.

In the last section we discuss the possible relevance of this result for the noncommutative theory of α-connections.

THE FISHER-RAO METRIC AND CHENTSOV UNIQUENESS THEOREM

Denote by $\mathcal{P}_n = \{p \in \mathbf{R}^n : \sum_{i=1}^n p_i = 1, p_i > 0, i = 1, \ldots, n\}$ the probability simplex in $\mathbf{R}^n$ and by S the sphere of radius 2 in $\mathbf{R}^n$. Define a function $A : \mathcal{P}_n \to S$ by $A(p)_i = 2p_i^{\frac{1}{2}}$ and consider the riemannian structure that S induces on $\mathcal{P}_n$ by this embedding. Let $p(t)$ be a curve on $\mathcal{P}_n$. We transport this curve on S by the embedding A and determine the induced riemannian metric. As $\frac{d}{dt}A(p(t))_i = \frac{1}{\sqrt{p_i(t)}}\frac{d}{dt}(p_i(t))$, we get

$$\|\frac{d}{dt}A(p(t))\|^2 = \sum_{i=1}^n (\frac{d}{dt}A(p(t))_i)^2 = \sum_{i=1}^n \frac{1}{p_i(t)} \left(\frac{d}{dt}(p_i(t))\right)^2 .$$

So we obtain the well-known Fisher-Rao metric

$$\sum_{i=1}^{n} \frac{1}{p_i} v_i w_i = \sum_{i=1}^{n} \frac{v_i}{\sqrt{p_i}} \frac{w_i}{\sqrt{p_i}}.$$

Now let m be the counting measure on $\{1, \ldots, n\}$, so that $L^2(m)$ can be identified with $\mathbf{R}^n$ with the usual scalar product. Then the Fisher-Rao metric is induced by the linear isomorphism $v \to \frac{v}{\sqrt{p}}$ that identifies the tangent space of $\mathcal{P}_n$ at p, $T_p\mathcal{P}_n = \{v \in \mathbf{R}^n : \sum v_i = 0\}$, with the tangent space of the unit sphere $S^2(m) \subset L^2(m)$ at the point $\sqrt{p}$, $T_{\sqrt{p}}(S^2(m)) = \{u \in L^2(m) : \langle u, \sqrt{p} \rangle = 0\}$, where $\langle u, v \rangle := \sum_{i=1}^{n} u_i v_i$ denotes the scalar product of $\mathbf{R}^n$.

Recall that a monotone metric is a family $\{g_p : p \in \mathcal{P}_n, n \in \mathbf{N}\}$ of inner products on $T_p\mathcal{P}_n$, such that $g_{Tp}(Tu, Tu) \leq g_p(u, u)$, for any stochastic map $T : \mathbf{R}^n \to \mathbf{R}^k$.

Define $(M_p(v))_i = p_i v_i$, $i = 1, \ldots, n$. We may formulate the

Theorem 1. *Chentsov uniqueness Theorem.* There exists a unique (up to a scalar factor) monotone metric on $T_p\mathcal{P}_n$. This metric is the scalar product that turns into an isometry the linear map

$$v \in T_p\mathcal{P}_n \to M_p^{-1/2}(v) \in L^2(m).$$

Proof. Evidently

$$\sum \frac{1}{p_i} v_i w_i = \langle M_p^{-1/2}(v), M_p^{-1/2}(w) \rangle.$$

The proof of monotonicity can be found in [2,3]. ♣

OPERATOR MONOTONE FUNCTIONS AND CHENTSOV-MOROZOVA FUNCTIONS.

Let us recall [1] that a function $f : (0, \infty) \to R$ is called *operator monotone* if for any $n \in N$, any A, $B \in M_n(\mathbf{C})$ such that $0 \leq A \leq B$, the inequalities $0 \leq f(A) \leq f(B)$ hold. By Löwner's results they can be represented in integral form. To make expressions compact, let us introduce the notation

$$\phi(x, t) = \frac{x(1+t)}{x+t}, \quad \text{for} \quad x > 0, \qquad t \geq 0.$$

For fixed $x > 0$ the function $\phi(x, t)$ is bounded and continuous on the extended half–line $[0, \infty]$.

Theorem 2. ([12] p. 208–9) The map $m \to f$, defined by

$$f(x) = \int_{[0,\infty]} \phi(x, t) dm(t), \qquad \text{for} \quad x > 0,$$

establishes an affine isomorphism from the class of positive Radon measures on $[0, \infty]$ onto the class of operator monotone functions.

Remark. In the above representation one has $f(0) = \inf_x f(x) = m(\{0\})$ and $\inf_x \frac{f(x)}{x} = m(\{\infty\})$.

Some other operator monotone functions are associated to a given operator monotone function f ([12] p. 213–4), among them the *transpose* function, $f'(x) := xf(x^{-1})$, and the *dual* function, $f^{\perp}(x) := \frac{x}{f(x)}$. These transformations are involutive, that is $f'' = f$, $f^{\perp\perp} = f$. Moreover f is said *symmetric* if $f = f'$.

We need a different representation for the operator monotone functions. It is based on the following result.

Lemma 3. (see [11]) Define $g : [0,1] \to [0,\infty]$ by $g(s) = \frac{s}{1-s}$, $s \in [0,1)$, and $g(1) = \infty$. Let $A \subset [0,1]$ and $B \subset [0,\infty]$ be measurable sets and let μ be a positive Radon measure on $[0,1]$ and m a positive Radon measure on $[0,\infty]$. The formulae $m_\mu(B) := \mu(g^{-1}(B))$, $\mu_m(A) = m(g(A))$, establish a bijection between the class of positive Radon measures on $[0,1]$ and the class of positive Radon measures on $[0,\infty]$. Moreover, if h is an integrable function w.r.t. m, then

$$\int_{[0,\infty]} h(t)dm(t) = \int_{[0,1]} h(g(s))d\mu_m(s).$$

Proposition 4. The map $\mu \to f$, defined by

$$f(x) = \int_{[0,1]} \frac{x}{(1-s)x+s}\, d\mu(s), \qquad \text{for} \quad x > 0,$$

establishes a bijection between the class of positive Radon measures on $[0,1]$ and the class of operator monotone functions.

Proof.

$$f(x) = \int_{[0,\infty]} \phi(x,t)dm(t) = \int_{[0,1]} \phi(x,g(s))d\mu_m(s) =$$

$$= \int_{[0,1]} \frac{x\left(1+\frac{s}{1-s}\right)}{x+\frac{s}{1-s}}\, d\mu_m(s) = \int_{[0,1]} \frac{x}{(1-s)x+s}\, d\mu_m(s).$$

♣

In the above corrispondence we write $f = f_\mu$ or $\mu = \mu_f$ to indicate that f_μ is the operator monotone function associated to μ or viceversa that μ_f is the measure associated to f.

Corollary 5. ([21] p. 474) The map $\mu \mapsto f$, defined by

$$\frac{1}{f(x)} = \int_{[0,1]} \frac{1}{(1-s)x+s}\, d\mu(s) \qquad \text{for} \quad x > 0$$

establishes a bijection between the class of positive Radon measures on $[0,1]$ and the class of operator monotone functions.
Proof. For each operator monotone function f one has

$$\frac{1}{f(x)} = \frac{1}{x} f^{\perp}(x) = \frac{1}{x} \int_{[0,1]} \frac{x}{(1-s)x+s}\, d\mu_{f^{\perp}}(s) = \int_{[0,1]} \frac{1}{(1-s)x+s}\, d\mu_{f^{\perp}}(s).$$

♣

Definition 6. To each operator monotone function f one associates the so-called *Chentsov–Morozova* function

$$c_f(x,y) := \frac{1}{yf\left(\frac{x}{y}\right)}, \qquad \text{for} \quad x,y>0.$$

Proposition 7. The map $\mu \mapsto c(\cdot,\cdot)$, defined by

$$c(x,y) = \int_{[0,1]} \frac{1}{(1-s)x+sy}\, d\mu(s), \qquad \text{for} \quad x,y>0$$

establishes a bijection between the class of positive Radon measures on $[0,1]$ and the class of Chentsov–Morozova functions.
Proof. By the Corollary 5 we have

$$c_f(x,y) = \frac{1}{yf\left(\frac{x}{y}\right)} = \frac{1}{y} \int_{[0,1]} \frac{1}{(1-s)\frac{x}{y}+s}\, d\mu_{f^{\perp}}(s) = \int_{[0,1]} \frac{1}{(1-s)x+sy}\, d\mu_{f^{\perp}}(s).$$

♣

Note that f is symmetric *iff* c_f (or μ_f) is, *i.e.* it satisfies $c(x,y)=c(y,x)$ (or $d\mu(s)=d\mu(1-s)$).

THE MAIN RESULT

Let τ be the usual trace on $M_n(\mathbf{C})$, and let $\mathcal{D}_n := \{A \in M_n(\mathbf{C}) : A>0, \tau(A)=1\}$ be the set of density matrices. The tangent space to $\mathcal{D}_n$ at $\rho \in \mathcal{D}_n$ can be naturally identified with $\{A \in M_n(\mathbf{C}) : A=A^*, \tau(A)=0\}$. Similarly to the commutative case a symmetric monotone metric is a family $\{g_\rho : \rho \in \mathcal{D}_n, n \in \mathbf{N}\}$ of inner products on $T_\rho\mathcal{D}_n$, such that $\rho \in \mathcal{D}_n \to g_\rho(A,A) \in [0,\infty)$ is continuous, for any $A \in T_\rho\mathcal{D}_n$, and $g_{T\rho}(TA,TA) \leq g_\rho(A,A)$, for any stochastic (i.e. completely positive, trace preserving) map $T : M_n(\mathbf{C}) \to M_k(\mathbf{C})$.

Theorem 8. *Petz classification theorem.* There exists a bijective correspondence between symmetric monotone metrics and symmetric operator monotone functions

$f : (0,\infty) \to (0,\infty)$, which is given by $g_\rho(A,B) = \tau(Ac_f(L_\rho,R_\rho)(B))$, for A, $B \in T_\rho\mathcal{D}_n$, where c_f is the CM-function associated to f.

We want to give a different description of Petz classification theorem. So let us start with some definitions.

Definition 9. Denote by $L^2(\tau)$ the vector space $M_n(\mathbf{C})$ endowed with the inner product $\langle A,B\rangle := \tau(A^*B)$. For any $\rho \in \mathcal{D}_n$, $s \in [0,1]$, define the operators $L_\rho(A) = \rho A$, $R_\rho(A) = A\rho$, and $M_{\rho,s} := (1-s)L_\rho + sR_\rho$. Then L_ρ, R_ρ, $M_{\rho,s}$ are positive invertible linear operators on $L^2(\tau)$.

Definition 10. Set $\mathcal{H}_s := L^2(\tau)$, $s \in [0,1]$, and, for any symmetric positive Radon measure μ on $[0,1]$, set

$$\mathcal{H}_\mu := \int_{[0,1]}^{\oplus} \mathcal{H}_s d\mu(s) \cong L^2([0,1], d\mu; L^2(\tau)) \cong L^2([0,1], d\mu) \otimes L^2(\tau),$$

and $M_{\rho,\mu} := \int_{[0,1]}^{\oplus} M_{\rho,s} d\mu(s)$.

Therefore $\mathcal{H}_\mu$ is endowed with the inner product $\langle A,B\rangle := \int_0^1 \langle A(s), B(s)\rangle d\mu(s)$, if $A : s \in [0,1] \to A(s) \in L^2(\tau)$, and analogously for B.

Definition 11. Let μ be a symmetric positive Radon measure on $[0,1]$. The μ-metric on $T_\rho\mathcal{D}_n$, denoted by $\langle\cdot,\cdot\rangle_{\rho,\mu}$, is the inner product on $T_\rho\mathcal{D}_n$ which turns into an isometry the linear map

$$A \in T_\rho\mathcal{D}_n \to M_{\rho,\mu}^{-1/2}(A) \in \mathcal{H}_\mu,$$

where $M_{\rho,\mu}^{-1/2} = \int_{[0,1]}^{\oplus} M_{\rho,s}^{-1/2} d\mu(s)$.

Theorem 12. The family of μ-metrics coincides with the family of symmetric monotone metrics classified by Petz theorem.

Proof. By Petz theorem each monotone metric has the form $g_\rho(A,B) = \tau(Ac(L_\rho,R_\rho)(B))$. Therefore we get the conclusion by the following calculation

$$\langle A,B\rangle_{\rho,\mu} = \int_{[0,1]} \langle M_{\rho,s}^{-1/2}(A), M_{\rho,s}^{-1/2}(B)\rangle d\mu(s) =$$

$$= \int_{[0,1]} \tau(A \cdot M_{\rho,s}^{-1}(B)) d\mu(s) = \int_{[0,1]} \tau(A((1-s)L_\rho + sR_\rho)^{-1}(B)) d\mu(s)$$

$$= \tau\left(A\left(\int_{[0,1]} ((1-s)L_\rho + sR_\rho)^{-1} d\mu(s)\right)(B)\right) = \tau(Ac_\mu(L_\rho,R_\rho)(B)).$$

♣

The integral decomposition on which Theorem 12 is based, has been sketched by Uhlmann in [21].

A DIFFERENT APPROACH

One could also follow a different approach, choosing as a noncommutative analogue of M_p a different interpolation, namely $\widetilde{M}_{\rho,s} := L_\rho^{1-s} R_\rho^s$, for $s \in [0,1]$, as suggested by one of the forms of the BKM monotone metric. We need some preliminaries.

Proposition 13. Let ν be a symmetric positive Radon measure on $[0,1]$. The formula

$$f^\nu(x) := \int_{[0,1]} x^t \mathrm{d}\nu(t)$$

defines a map from the family of positive Radon measures on [0,1] to the class of operator monotone functions. The map is not surjective.

Proof. It is well-known [11] that $f^\nu(x) = \int_{[0,1]} x^t d\beta(t)$, where $\beta : [0,1] \to [0,\infty)$, is increasing and left-continuous, and $\beta(0) = 0$. Besides, it is easy to prove that there are $\beta_n : [0,1] \to [0,\infty)$ increasing, left-continuous, and piecewise constant functions, such that $\beta_n \to \beta$ uniformly in $[0,1]$. As for any fixed $x \in [0,\infty)$, the function $t \in [0,1] \to x^t \in [0,\infty)$ is continuous and bounded, by means of Helly's theorem we get $f^\nu(x) = \lim_{n\to\infty} f_n(x)$, where $f_n(x) := \int_0^1 x^t d\beta_n(t)$. Moreover, if $k \in \mathbf{N}$, $A = A^* \in M_k(\mathbf{C})$ with spectrum contained in $[0,1]$, then $f^\nu(A) = \lim_{n\to\infty} f_n(A)$. Indeed, if $A = \sum_{i=1}^k \lambda_i e_i$ is its spectral decomposition, then $f^\nu(A) = \sum_{i=1}^k f^\nu(\lambda_i) e_i = \sum_{i=1}^k \lim_{n\to\infty} f_n(\lambda_i) e_i = \lim_{n\to\infty} f_n(A)$. Now it is easy to see that f_n is an operator monotone function, being a linear combination with positive coefficients of functions x^{t_i}, which are operator monotone [1]. Finally if $k \in \mathbf{N}$, A, $B \in M_k(\mathbf{C})$, are such that $0 \le A \le B$ we get $f^\nu(A) = \lim_{n\to\infty} f_n(A) \le \lim_{n\to\infty} f_n(B) = f^\nu(B)$, which proves that f^ν is operator monotone.
As for the last statement, the function $\frac{2x}{x+1}$, which is operator monotone [1] and gives the largest monotone metric, is not in the range of the map, because, if ν is not a multiple of the Dirac measure at 0, any f^ν is such that $\lim_{x\to\infty} f^\nu(x) = \infty$, otherwise f^ν is constant. ♣

Corollary 14. The formula

$$c^\nu(x,y) := \int_{[0,1]} (x^{1-t} y^t)^{-1} \mathrm{d}\nu(t)$$

defines a map from the family of positive Radon measures on [0,1] to the class of CM-functions. The map is not surjective.

Definition 15. Let ν be a symmetric positive Radon measure on $[0,1]$. The ν-metric on $T_\rho \mathcal{D}_n$, denoted by $\langle \cdot, \cdot \rangle_\rho^\nu$, is the inner product on $T_\rho \mathcal{D}_n$ which turns into an isometry the linear map

$$A \in T_\rho \mathcal{D}_n \to \widetilde{M}_{\rho,\nu}^{-1/2}(A) \in \mathcal{H}_\nu,$$

where $\widetilde{M}_{\rho,\nu}^{-1/2} := \int_{[0,1]}^{\oplus} M_{\rho,s}^{-1/2} d\nu(s)$.

Theorem 16. The family of ν-metrics is a proper subset of the family of monotone metrics classified by Petz theorem.
Proof. By definition

$$\langle A, B\rangle_\rho^\nu = \int_{[0,1]} \langle \widetilde{M}_{\rho,s}^{-1/2}(A), \widetilde{M}_{\rho,s}^{-1/2}(B)\rangle d\nu(s) =$$

$$= \int_{[0,1]} \tau(A \cdot \widetilde{M}_{\rho,s}^{-1}(B)) d\nu(s) = \int_{[0,1]} \tau(A(L_\rho^{1-s} R_\rho^s)^{-1}(B)) d\nu(s)$$

$$= \tau\left(A\left(\int_{[0,1]} (L_\rho^{1-s} R_\rho^s)^{-1} d\nu(s)\right)(B)\right) = \tau(Ac^\nu(L_\rho, R_\rho)(B)).$$

Therefore the conclusion follows from the previous results. ♣

UNIFORMLY CONVEX BANACH SPACES

The purpose of this section is to review some results on the geometry of uniformly convex Banach spaces, needed in the sequel. We refer to [8] for full proofs and consider only real Banach spaces. Denote by $\widetilde{X}$ the dual space of X and by S^X the unit sphere of X. If $L \in \widetilde{X}$ and $x \in X$ we write $\langle L, x\rangle = L(x)$.

Definition 17. We say that x is *orthogonal* to y, and denote it by $x \perp y$, if $\|x\| \leq \|x + \lambda y\|$, for any $\lambda \in \mathbf{R}$. Moreover, if $A \subset X$, $x \perp A$ means $x \perp y$, for any $y \in A$.

Definition 18. The *duality mapping* $J : X \to \text{Subsets}(\widetilde{X})$ is defined by $J(x) := \{v \in \widetilde{X} : \langle v, x\rangle = \|x\|^2 = \|v\|^2\}$. We say that X has the *duality map property* if J is single-valued. In this case we set $\widetilde{x} := J(x)$.

Definition 19. We say that X has the *projection property* if for any closed convex $M \subset X$ and any $x \in X$ there is a unique $m \in M$ s.t. $\|x - m\| = \inf\{\|x - z\| : z \in M\} \equiv d(x, M)$. In this case we define $\pi_M(x) := m$.

Definition 20. X is *uniformly convex* if for any $\varepsilon > 0$ there is $\delta > 0$ s.t. $x, y \in S^X$ and $\|\frac{x+y}{2}\| > 1 - \delta$ implies $\|x - y\| < \varepsilon$.

Proposition 21. Let X and $\widetilde{X}$ be uniformly convex Banach spaces. Then
i) X has the projection property.
ii) X has the duality map property.
iii) $x \perp \ker(\widetilde{x})$.
iv) If $M := \ker(\widetilde{x})$, then $\pi_M(v) = v - \frac{\langle \widetilde{x}, v\rangle}{\langle \widetilde{x}, x\rangle} x$.

Now recall that if $\mathcal{M}$ is a Banach manifold and $\mathcal{N} \subset \mathcal{M}$ is a submanifold, then for any $p \in \mathcal{N}$ there is a splitting of the tangent space $T_p\mathcal{M} = T_p\mathcal{N} \oplus V$ and a projection operator $\pi_p : T_p\mathcal{M} \to T_p\mathcal{N}$. Moreover if there is a connection ∇ on $\mathcal{M}$, one gets a connection ∇' on the submanifold $\mathcal{N}$, by setting $\nabla' := \pi \circ \nabla$.

Proposition 22. Let X, $\widetilde{X}$ be uniformly convex Banach spaces. Then
i) S^X is a Banach submanifold of X.
ii) $T_x S^X$, the tangent space to S^X at $x \in S^X$, can be identified with $\ker(\widetilde{x})$.
iii) The projection operator $\pi_x : T_x X \to T_x S^X$ is given by $\pi_x(v) = v - \langle \widetilde{x}, v\rangle x$.

Using this projection, the trivial connection on X induces a connection on S^X, that we call the *natural connection* on S^X.

We may rephrase the content of this section by saying that if X, $\widetilde{X}$ are uniformly convex then X is "almost an Hilbert space". Note that if X is an Hilbert space, then the natural connection on S^X is just Levi-Civita connection on the sphere.

α-CONNECTIONS FOR COMMUTATIVE STATISTICAL MANIFOLDS

In this section we summarise some of the results of [6] in the light of the abstract setting of the previous section. Let $(X, \mathcal{A}, \mu)$ be a measure space. We give the following

Definition 23. If $\alpha \in (-1, 1)$, set $p := \frac{2}{1-\alpha}$. $L^p_{\mathbf{R}} \equiv L^p_{\mathbf{R}}(X, \mathcal{A}, \mu) := \{u : X \to \mathbf{R} : u \text{ is } \mathcal{A}\text{-measurable}, \int_X |u|^p d\mu < \infty\}$, for $p \in [1, \infty)$. The unit sphere is denoted by $S^p := \{f \in L^p_{\mathbf{R}} : \|u\|_p = 1\}$. $\mathcal{P}_\mu := \{\rho \in L^1_{\mathbf{R}} : \rho > 0, \int_X \rho = 1\}$. For any $\rho \in \mathcal{P}_\mu$ we set $\mathcal{F}^\alpha_\rho \equiv L^p_0(\rho) := \{u \in L^p_{\mathbf{R}}(X, \mathcal{A}, \rho\mu) : \int_X u\rho d\mu = 0\}$. If $p > 1$ we define $\widetilde{p}$ by $\frac{1}{p} + \frac{1}{\widetilde{p}} = 1$.

A calculation shows that the duality map is given by $u \in L^p_{\mathbf{R}} \to \widetilde{u} := \|u\|_p^{2-p} \mathrm{sgn} u |u|^{\frac{p}{\widetilde{p}}} \in L^{\widetilde{p}}_{\mathbf{R}}$. Therefore, if $\rho \in \mathcal{P}_\mu$, we have that $\rho^{1/p} \in S^p$ and $\widetilde{\rho^{1/p}} = \rho^{1/\widetilde{p}} \in S^{\widetilde{p}}$. The spaces $L^p_{\mathbf{R}}$ are uniformly convex, so the results of the previous section are applicable. For the tangent space of S^p at $\rho^{1/p}$ we have $T_{\rho^{1/p}} S^p = \{u \in L^p_{\mathbf{R}} : \int u\rho^{1/\widetilde{p}} d\mu = 0\}$. We denote by ∇^p the natural connection on S^p induced by the trivial connection on $L^p_{\mathbf{R}}$. Observe that the isometric isomorphism $I^p_\rho : u \in L^p_{\mathbf{R}}(X, \mathcal{A}, \mu) \to u\rho^{-1/p} \in L^p_{\mathbf{R}}(X, \mathcal{A}, \rho\mu)$ sets up a bijection between $T_{\rho^{1/p}} S^p$ and $L^p_0(\rho)$.

Let $\mathcal{N} \subset \mathcal{P}_\mu$ be a statistical model, equipped with a structure of a differential manifold. Consider the bundle-connection pair on S^p given by the tangent bundle and the natural connection (TS^p, ∇^p). Making use of the Amari embedding $A^\alpha : \rho \in \mathcal{N} \to \rho^{1/p} \in S^p$, we may construct the pull-back $((A^\alpha)^* TS^p, (A^\alpha)^* \nabla^p)$ of the bundle-connection pair (TS^p, ∇^p) to $\mathcal{N}$. This means that the fibre over $\rho \in \mathcal{N}$ of the pull-back bundle is given by $T_{\rho^{1/p}} S^p$. Consider now $\mathcal{F}^\alpha := \cup_{\rho \in \mathcal{N}} \mathcal{F}^\alpha_\rho$. Using the family of isomorphisms I^p_ρ, $\rho \in \mathcal{N}$, it is possible to identify $\mathcal{F}^\alpha$ with the pull-back bundle $(A^\alpha)^*TS^p$. One can also transfer the pull-back connection $(A^\alpha)^*\nabla^p$ using this isomorphism. We denote by ∇^α this last connection on the bundle $\mathcal{F}^\alpha$.

Theorem 24. [6] Consider the bundle-connection pair $(\mathcal{F}^\alpha, \nabla^\alpha)$, $\alpha \in (-1, 1)$, on the statistical manifold $\mathcal{N}$. Then ∇^α coincides with the Amari–Chentsov α-connection.
Proof. One obtains

$$\nabla^\alpha = \frac{1+\alpha}{2}\nabla^e + \frac{1-\alpha}{2}\nabla^m \tag{1}$$

where ∇^m and ∇^e are the usual mixture and exponential connections defined by parallel transport on the mixture and exponential bundles $L_0^{x \log x}$ and $L_0^{\exp}$ (see [6] for details). ♣

It is useful to emphasize the new aspects that this theorem introduces in Information Geometry. First of all, it solves the longstanding problem of an infinite dimensional theory for α-geometries (note that we may discuss orthogonality, projections, etc. also in a non-riemannian, non-hilbertian setting). Moreover α-connections appear as L^p-connections in a disguised form (a new result even in the parametric case). Following this line of thought we want to stress that equality (1) should be seen as a theorem and not as a definition. In this sense the parametric case could be seriously misleading: indeed the α-connections are not defined on the tangent space, in general, but on a suitable α-bundle (this point is still overlooked also in some recent papers). In addition one should note that the problem of different geodesics intersecting at right angles cannot be solved naively. In general these geodesics will be on two different manifolds (the target manifolds of different embeddings of the densities) such that a duality pairing exists between the two tangent bundles. A theory of this type has been outlined in [7] and probably this can be the right approach also in the non-commutative setting (see the work of Streater [18,19] where the use of +1 and -1 geodesics is of great importance in the theory of statistical dynamics). But probably the most important aspect is that one can see all the construction from an abstract point of view (that is for uniformly convex spaces) so that this kind of family of dual geometries should appear whenever one has a family of L^p-type spaces. We have discussed this approach in a previous paper [8] regarding a non-commutative non-parametric generalisation of the α-connections.

NORMS OF L^p-TYPE AND α-CONNECTIONS ASSOCIATED TO MONOTONE METRICS

A general approach to noncommutative α-connections is still missing, even though a a number of different points of view exist [8–10,13]. But now Theorem 12 shows that each monotone metric can be obtained by an L^2 scalar product.

Moreover, motivated by Theorem 24 and by the considerations of the previous section, we suggest that one should try to construct α-geometries associated to an arbitrary monotone metric by the construction of an L^p-norm associated to that monotone metric. What follows is a tentative first step in that direction.

Let $(E, \|\cdot\|)$ be a Banach space and denote by $\mathcal{L}(E)$ the set of continuous linear operators on E, and by $GL(E)$ the subset of the invertible ones. If $T \in \mathcal{L}(E)$, we may define a new Banach space $(E, \|\cdot\|_T)$, where $\|v\|_T := \|Tv\|$. Moreover, if $T \in GL(E)$, then $T^{-1} : (E, \|\cdot\|) \to (E, \|\cdot\|_T)$ is an isometric isomorphism.
Now let $L^p(\tau)$ be the matrix von Neumann-Schatten class, that is $M_n(\mathbf{C})$ endowed with the norm $\|A\|_p := \tau(|A|^p)^{1/p}$, and consider $T := M_{\rho,s}^{1/p} \in \mathcal{L}(L^p(\tau))$. Therefore, we may consider the norm $\|A\|_T := \|T(A)\|_p = (\tau(|M_{\rho,s}^{1/p}(A)|^p))^{1/p}$, for $A \in M_n(\mathbf{C})$. Analogously, if we set $\widetilde{T} := \widetilde{M}_{\rho,s}^{1/p}$, we have the norms $\|A\|_{\widetilde{T}} := (\tau(|\widetilde{M}_{\rho,s}^{1/p}(A)|^p))^{1/p}$. So we may define the Banach spaces $L^p(\rho)_s := (M_n(\mathbf{C}), \|\cdot\|_T)$, and $L^p(\rho)^s := (M_n(\mathbf{C}), \|\cdot\|_{\widetilde{T}})$. The latter spaces are the matrix version of the spaces introduced by Trunov and Zolotarev [20,22] and studied by several authors.
In the construction of commutative α-connections is fundamental the isomorphism $u \in L^p(\rho) \to u\rho^{1/p} \in L^p(\tau)$, that allows to identify $T_{\rho^{1/p}} S^p$ with the space $L_0^p(\rho)$ of p-integrable ρ-centred random variables. The operator $A \in L^p(\rho)_s \to M_{\rho,s}^{1/p}(A) \in L^p(\tau)$ could play the same role. If $p = 2$, we may identify $L^2(\rho)_\mu := \int_{[0,1]}^{\oplus} L^2(\rho)_s d\mu(s)$ with $L^2([0,1], d\mu) \otimes L^2(\tau)$, by means of the operator $\int_{[0,1]}^{\oplus} M_{\rho,s}^{1/2} d\mu(s)$. For example the proof of Theorem 12 can be reformulated using $M_{\rho,\mu}^{-1}$ instead of $M_{\rho,\mu}^{-1/2}$ and $L^2(\rho)_\mu$ instead of $L^2([0,1], d\mu) \otimes L^2(\tau)$. In a similar way one may consider $L^2(\rho)^\nu := \int_{[0,1]}^{\oplus} L^2(\rho)^s d\nu(s)$ (this kind of inner product has been introduced by Petz and Toth [15]) and accordingly give a different proof of Theorem 16.
In view of the above considerations, we conjecture that it could be possible to associate to an arbitrary monotone metric a family of α-connections, using a kind of direct integral of the Banach spaces $L^p(\rho)_s$, $s \in [0,1]$, with respect to a positive Radon measure μ on $[0,1]$.

REFERENCES

1. Bhatia R., *Matrix Analysis*. New York: Springer Verlag, 1997.
2. Campbell L. L., An extended Chentsov characterization of the information metric, *Proc. Amer. Math. Soc.* **98**, 135 (1986).
3. Chentsov N. N, *Statistical decision rules and optimal inference*. Transl. Math. Monographs, **53**, Providence: American Math. Soc., 1982.
4. Chentsov N. N, and Morozova E. A., Markov invariant geometry on state manifolds, (in Russian), *Itogi Nauki i Tekhniki* **36**, 69 (1990).
5. Dawid A. P., Further comments on some comments on a paper by Bradley Efron, *Ann. Stat.* **5**, 1249 (1977).
6. Gibilisco P., and Pistone G., Connections on non-parametric statistical manifolds

by Orlicz space geometry, *Infinite Dimensional Analysis, Quantum Probability & Related Fields* **1**, 325 (1998).

7. Gibilisco P., and Pistone G., Analytical and geometrical properties of statistical connections in information geometry, in *Mathematical Theory of Networks and Systems*, eds. A. Beghi, L. Finesso and G. Ricci, Padova: Il Poligrafo, 1999.
8. Gibilisco P., and Isola T., Connections on statistical manifolds of density operators by geometry of noncommutative L^p-spaces, *Infinite Dimensional Analysis, Quantum Probability & Related Fields* **2**, 169 (1999).
9. Hasegawa H., Non-commutative extension of the Information geometry, Proc. Int. Workshop on *Quantum Communications and Measurement*, Nottingham 1994, eds. V. P. Belavkin, O. Hirota and R. L. Hudson, Plenum Press, 1995.
10. Hasegawa H., and Petz D., Non-commutative extension of Information geometry II, in *Quantum Communication, Computing, and Measurement*, edited by Hirota *et al.*, Plenum Press, New York, 1997.
11. Hewitt E., and Stromberg K., *Real and Abstract Analysis.* New York: Springer Verlag, 1965.
12. Kubo F., and Ando T., Means of positive linear operators, *Math. Ann.* **246**, 205 (1980).
13. Nagaoka H., Differential geometrical aspects of quantum state estimation and relative entropy, Proc. Int. Workshop on *Quantum Communications and Measurement*, Nottingham 1994, eds. V. P. Belavkin, O. Hirota and R. L. Hudson, Plenum Press, 1995.
14. Petz D., Monotone metrics on matrix spaces, *Linear algebra and appl.* **244**, 81 (1996).
15. Petz D., and Toth G., The Bogoliubov inner product in quantum statistics, *Lett. Math. Phys.* **27**, 205 (1993).
16. Pistone G., and Rogantin M. P., The exponential statistical manifold: mean parameters, orthogonality, and space transformation, *Bernoulli* **5**, 721, (1999).
17. Pistone G., and Sempi C., An infinite dimensional geometric structure on the space of all the probability measures equivalent to a given one, *Ann. Stat.* **33**, 1543 (1995).
18. Streater R., *Statistical Dynamics*, London: Imperial College Press, 1995.
19. Streater R., Statistical dynamics and information geometry, in "Geometry and nature", *Contemp. Math.* **203**, p. 117, Amer. Math. Soc., Providence, RI, 1997.
20. Trunov N. V., On a noncommutative analogue of the space L_p, *Iz. VUZ Matematika* (Soviet Mathematics Translation) **23**, 69, (1979).
21. Uhlmann A., Geometric phases and related structures, *Reports on Math. Phys.* **36**, 461, 1995.
22. Zolotarev A. A., L^p spaces with respect to states on a von Neumann algebra and interpolation, *Iz. VUZ Matematika* (Soviet Mathematics Translation) **26**, 36, (1982).