JOURNAL OF MATHEMATICAL PHYSICS VOLUME 44, NUMBER 9 SEPTEMBER 2003

Wigner–Yanase information on quantum state space: The geometric approach

Paolo Gibilisco[a)]
Dipartimento di Scienze, Facoltà di Economia, Università di Chieti-Pescara "G. D'Annunzio," Viale Pindaro 42, I–65127 Pescara, Italy

Tommaso Isola[b)]
Dipartimento di Matematica, Università di Roma "Tor Vergata," Via della Ricerca Scientifica, I–00133 Roma, Italy

(Received 10 April 2003; accepted 3 May 2003)

In the search of appropriate Riemannian metrics on quantum state space, the concept of statistical monotonicity, or contraction under coarse graining, has been proposed by Chentsov. The metrics with this property have been classified by Petz. All the elements of this family of geometries can be seen as quantum analogs of Fisher information. Although there exists a number of general theorems shedding light on this subject, many natural questions, also stemming from applications, are still open. In this paper we discuss a particular member of the family, the Wigner–Yanase information. Using a well-known approach that mimics the classical pull-back approach to Fisher information, we are able to give explicit formulas for the geodesic distance, the geodesic path, the sectional and scalar curvatures associated to Wigner–Yanase information. Moreover, we show that this is the only monotone metric for which such an approach is possible. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1598279]

I. INTRODUCTION

The notion of information proposed by Fisher is fundamental in probability and statistics for a number of reasons; here we mention only the Cramer–Rao inequality and the asymptotic behavior of maximum lilkelihood estimators for exponential models (one can see Ref. 5 for unexpected features and applications of Fisher information). In classical statistics Rao was the first to point out that Fisher information can be seen as a Riemannian metric on the space of probability densities. This point of view was nicely complemented by the results of Chentsov, saying that (on the simplex of probability vectors) Fisher information is the unique Riemannian metric contracting under Markov morphisms. This can be rephrased in a more suggestive way. Markov morphisms, or positive mappings, are the mathematical counterpart of the notion of noise. Now suppose that we want to use a distance to distinguish different states (probability densities) in a statistically relevant way. Then the effect of noise must be that of contracting the metric. Chentsov theorem says therefore that in the classical case there is only one choice, the Fisher information (another argument producing Fisher information can be found in Ref. 43).

In the quantum case one deals with density operators instead of density vectors and completely positive mappings play the role of Markov morphisms. As often happens in the quantum counterpart of a classical theory, instead of a uniqueness result, one has a classification theorem, due to Petz. This result states that there is bijection between statistically monotone metrics on quantum state space and the operator monotone functions: we have therefore a rich "garden" of candidates for the role of Fisher information in quantum physics. Among the elements of this family of metrics one can find, in a certain sense, the most relevant Riemannian metrics appeared in the literature.[35,37]

[a)]Electronic mail: gibilisc@sci.unich.it
[b)]Electronic mail: isola@mat.uniroma2.it

Despite the existence of general results for the theory[13,17,19,26,27,28,30,40,41] a number of open problems resists investigation. For example, it does not exist yet a general formula for the geodesic path and the geodesic distance associated to an arbitrary monotone metric. For the use of this kind of distances see, for example Ref. 32. Because of the absence of a general formula, inequalities (giving bounds for the geodesic distance) have been proved.[38]

In this paper we discuss the Wigner–Yanase skew information. To find the formulas for geodesic path and geodesic distance we mimic the classical approach to Fisher information via sphere geometry (one should note the importance of determining the geodesic path in the study of the 2-Wasserstein metric[6]). Indeed Wigner–Yanase information appears as the pull-back of the square root map.[18] Next we prove the formula for the scalar curvature. One proof, due to Dittmann, uses the general formula[13] and requires a long calculation. The second one just uses the pull-back approach. One should emphasize that, since the scalar curvature determines the asymptotic behavior of the volume (for a Riemannian metric) then it has also a statistical meaning in relation to the quantum analog of Jeffrey's rule for determining prior probability distributions (see Ref. 35). Finally we prove, as a corollary of the results in Refs. 25, 26, and 19 that the Wigner–Yanase information is the only monotone metric that can be seen as a pull-back metric.

The paper is organized as follows. In Sec. II we review the geometric approach to Fisher information. Sec. III one finds an introduction to the general theory of statistical monotone metrics. Sec. IV shows how the Wigner–Yanase information can be seen as a monotone Riemannian metric. In Sec. V we show that the Wigner–Yanase geometry can be seen as the sphere geometry transposed on the space of density matrices; moreover, we characterize it as the unique pull-back metric. Section VI contains some comments on the main results and on some open problems.

II. FISHER INFORMATION AND ITS GEOMETRY

The classical definition of Fisher information for an indexed family of densities p_θ is given by the variance of the score. In the case of a family indexed by only one parameter θ it is the number

$$I(\theta)=\mathrm{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log p_\theta\right)^2\right], \tag{2.1}$$

assigned to the parameter θ. For n parameters, say $\theta=(\theta^1,\dots,\theta^n)$, it is a matrix defined on the parameter manifold given by

$$I(\theta)_{ij}=\mathrm{E}_\theta\left[\left(\frac{\partial}{\partial\theta^i}\log p_\theta\right)\left(\frac{\partial}{\partial\theta^j}\log p_\theta\right)\right]. \tag{2.2}$$

Geometrically this means that $I(\theta)$ is a symmetric bilinear form on the tangent spaces of the parameter manifold. In a coordinate free language it reads as

$$I(\theta)(U,V)=\mathrm{E}_\theta[U(\log p_\theta)\,V(\log p_\theta)], \tag{2.3}$$

where U and V are vectors tangent to the parameter manifold and $U(\log p_\theta)$ is the derivative of $\log p_\theta$ along the direction U, which means $U(\log p_\theta)=(\mathrm{d}/\mathrm{d}t)\log p_{\theta+tU}|_{t=0}$.

$I(\theta)$ is a measure for the statistical distinguishabilty of distribution parameters. Under certain regularity conditions for $\theta\mapsto p_\theta$ the image of this mapping is a manifold of distributions. This manifold is the actual object of interest in information geometry rather than the space of distribution parameters and formula (2.3) defines a Riemannian metric g on it (for a general reference see Ref. 1). Indeed, a vector u tangent to this manifold is of the form

$$u=\frac{\mathrm{d}}{\mathrm{d}t}p_{\theta+tU}|_{t=0},$$

and the right hand side of (2.3) now reads as

$$g(u,v) := \mathrm{E}_p\left[\frac{u}{p}\frac{v}{p}\right], \tag{2.4}$$

defining the Fisher metric on the manifold of densities.

We restrict now to $\mathcal{P}_n \subset \mathbb{R}^n$, the simplex of strictly positive probability vectors, that is, $\mathcal{P}_n := \{\rho \in \mathbb{R}^n : \Sigma_{i=1}^n \rho_i = 1, \rho_i > 0, i = 1,\ldots,n\}$. An element $\rho \in \mathcal{P}_n$ is a density on the n-point set $\{1,\ldots,n\}$ with $\rho(i) = \rho_i$. We regard an element u of the tangent space $\mathrm{T}_\rho\mathcal{P}_n \equiv \{u \in \mathbb{R}^n : \Sigma_{i=1}^n u_i = 0\}$ as a function u on $\{1,\ldots,n\}$ with $u(i) = u_i$.

Definition 2.1: The Fisher–Rao Riemannian metric on $\mathrm{T}_\rho\mathcal{P}_n$ is given by

$$\langle u,v \rangle_\rho^F := \sum_{i=1}^n \frac{u_i v_i}{\rho_i}, \tag{2.5}$$

for $u,\ v \in \mathrm{T}_\rho\mathcal{P}_n$.

To see the relation between this metric and the Fisher metric, let $u,v \in \mathrm{T}_\rho\mathcal{P}_n$. We obtain from (2.4),

$$g(u,v) = \sum_{i=1}^n \frac{u(i)}{\rho_i}\frac{v(i)}{\rho_i}\rho_i = \sum_{i=1}^n \frac{u_i v_i}{\rho_i},$$

in accordance with (2.5).

The following result is well known and is a very special case of a far more general situation (see Ref. 15, for example).

Theorem 2.2: *The manifold $\mathcal{P}_n$ equipped with the Fisher–Rao Riemannian metric $\langle \cdot,\cdot \rangle^F$ is isometric with an open subset of the sphere of radius 2 in $\mathbb{R}^n$.*

Proof: We consider the mapping $\varphi : \mathcal{P}_n \to S_2^{n-1} \subset \mathbb{R}^n$,

$$\varphi(\rho) := 2(\sqrt{\rho_1}, \ldots, \sqrt{\rho_n}).$$

Then $D_\rho\varphi(u) = (u_1/\sqrt{\rho_1}, \ldots, u_n/\sqrt{\rho_n})$ and we get

$$D_\rho\varphi(\langle \cdot,\cdot \rangle^F)(u,v) := \langle D_\rho\varphi(u), D_\rho\varphi(v) \rangle^{\mathbb{R}^n} = \sum_{i=1}^n \frac{u_i v_i}{\rho_i} = \langle u,v \rangle_\rho^F.$$

Hence the standard metric on the sphere of radius 2 is pulled back to the Fisher–Rao Riemannian metric. □

This identification of $\mathcal{P}_n$ with an open subset of a radius 2 sphere allows for obtaining differential geometrical quantities of the Riemannian manifold $(\mathcal{P}_n, \langle \cdot,\cdot \rangle^F)$. From the very definition of geodesic distance, geodesic path and scalar curvature, one has for S_r^{n-1}, with $P_1, P_2 \in S_r^{n-1}$, the following:

(1) geodesic distance,

$$d(P_1,P_2) = r \cdot \arccos\left(\frac{\langle P_1, P_2 \rangle}{r^2}\right);$$

(2) geodesic path connecting P_1 and P_2,

$$\gamma^{P_1,P_2}(t) = r\frac{(1-t)P_1 + tP_2}{\|(1-t)P_1 + tP_2\|}$$

(of course, t is not the arc length parameter);

(3) scalar curvature,

$$\mathrm{Scal}(v)=\frac{1}{r^2}(n-1)(n-2),$$

because S_r^{n-1} has constant sectional curvature equal to $1/r^2$.

Let us denote by d_F, γ_F, Scal_F, respectively, the corresponding quantities for the Fisher information. The above considerations give, for $\rho,\sigma\in\mathcal{P}_n$, the following:

(1) Bhattacharya distance,

$$d_{\mathrm{F}}(\rho,\sigma)=2\arccos\left(\sum_i \rho_i^{1/2}\sigma_i^{1/2}\right);$$

(2) geodesic path connecting ρ and σ,

$$\gamma_{\mathrm{F}}^{\rho,\sigma}(t)=2\,\frac{((1-t)\sqrt{\rho}+t\sqrt{\sigma})^2}{\sum_i((1-t)\sqrt{\rho_i}+t\sqrt{\sigma_i})^2};$$

(3) scalar curvature,

$$\mathrm{Scal}_{\mathrm{F}}(\rho)=\frac{1}{4}(n-1)(n-2),\quad \forall\rho\in\mathcal{P}_n.$$

The Levi–Civita connection associated to Fisher metric can be decomposed using the geometry of mixture and exponential models. In the rest of the section we explain how.

Definition 2.3: A dualistic structure on a manifold $\mathcal{M}$ is a triple $(\langle\cdot,\cdot\rangle,\nabla,\tilde{\nabla})$, where $\langle\cdot,\cdot\rangle$ is a Riemannian metric on $\mathcal{M}$ and $\nabla,\tilde{\nabla}$ are affine connections on $\mathcal{M}$ such that

$$X\langle Y,Z\rangle=\langle\nabla_X Y,Z\rangle+\langle Y,\tilde{\nabla}_X Z\rangle,$$

where X,Y,Z are vector fields. If $U^{\nabla},U^{\tilde{\nabla}}$ are the parallel transport associated to $\nabla,\tilde{\nabla}$ then the above equation is equivalent to

$$\langle U^{\nabla}(u),U^{\tilde{\nabla}}(v)\rangle=\langle u,v\rangle.$$

A divergence on a manifold is a smooth non-negative function $D:\mathcal{M}\times\mathcal{M}\to\mathbb{R}$ such that $D(\rho,\sigma)=0$ iff $\rho=\sigma$. To each divergence D one may associate a dualistic structure $(\langle\cdot,\cdot\rangle,\nabla,\tilde{\nabla})$ (see Refs. 1 and 14).

Let ∇^2 be the Levi–Civita connection of Fisher information. The Kullback–Leibler relative entropy $K(\rho,\sigma)=\Sigma_i\rho_i(\log\rho_i-\log\sigma_i)$ gives a dualistic structure $(\langle\cdot,\cdot\rangle^F,\nabla^m,\nabla^e)$ such that

$$\nabla^2=\tfrac{1}{2}(\nabla^m+\nabla^e),$$

where ∇^m,∇^e are the mixture and exponential connections. These connections are torsion free and flat: once the representation by scores is used for the tangent spaces, the associated parallel transports are given by

$$U^m_{\rho\sigma}:T_\rho\mathcal{P}\to T_\sigma\mathcal{P},\quad U^m_{\rho\sigma}(u)=\frac{\rho}{\sigma}u,$$

$$U^e_{\rho\sigma}:T_\rho\mathcal{P}\to T_\sigma\mathcal{P},\quad U^e_{\rho\sigma}(u)=u-\mathrm{E}_\sigma(u).$$

The geodesics of ∇^m,∇^e are, respectively, the mixture and exponential models.

III. METRIC CONTRACTION UNDER COARSE GRAINING

In the commutative case a Markov morphism (or stochastic map) is a stochastic matrix $T:\mathbb{R}^n \to \mathbb{R}^k$. In the noncommutative case a stochastic map is a completely positive and trace preserving operator $T:M_n \to M_k$ where M_n denotes the space of n by n complex matrices. We shall denote by $\mathcal{D}_n$ the manifold of strictly positive elements of M_n and by $\mathcal{D}_n^1 \subset \mathcal{D}_n$ the submanifold of density matrices.

In the commutative case a monotone metric is a family of Riemannian metrics $g=\{g^n\}$ on $\{\mathcal{P}_n\}$, $n\in\mathbb{N}$ such that

$$g^m_{T(\rho)}(TX,TX) \leqslant g^n_\rho(X,X)$$

holds for every stochastic mapping $T:\mathbb{R}^n \to \mathbb{R}^m$ and all $\rho\in\mathcal{P}_n$ and $X\in T_\rho\mathcal{P}_n$.

In perfect analogy, a monotone metric in the noncommutative case is a family of Riemannian metrics $g=\{g^n\}$ on $\{\mathcal{D}_n^1\}$, $n\in\mathbb{N}$ such that

$$g^m_{T(\rho)}(TX,TX) \leqslant g^n_\rho(X,X)$$

holds for every stochastic mapping $T:M_n \to M_m$ and all $\rho\in\mathcal{D}_n^1$ and $X\in T_\rho\mathcal{D}_n^1$.

Let us recall that a function $f:(0,\infty)\to\mathbb{R}$ is called operator monotone if for any $n\in\mathbb{N}$, any A, $B\in M_n$ such that $0\leqslant A\leqslant B$, the inequalities $0\leqslant f(A)\leqslant f(B)$ hold. An operator monotone function is said to be symmetric if $f(x):=xf(x^{-1})$ and normalized if $f(1)=1$. In what follows by operator monotone we mean normalized symmetric operator monotone. With each operator monotone function f one associates also the so-called Chentsov–Morotzova function (see Ref. 8),

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

Define $L_\rho(A):=\rho A$, and $R_\rho(A):=A\rho$. Since L_ρ, R_ρ commute we may define $c(L_\rho,R_\rho)$. Now we can state the fundamental theorems about monotone metrics (uniqueness and classification are up to scalars).

Theorem 3.1: (Ref. 7) *There exists a unique monotone metric on $\mathcal{P}_n$ given by the Fisher information.*

Theorem 3.2: (Ref. 34) *There exists a bijective correspondence between monotone metrics on $\mathcal{D}_n^1$ and operator monotone functions given by the formula*

$$\langle A,B\rangle_{\rho,f}:=\mathrm{Tr}(A\cdot c_f(L_\rho,R_\rho)(B)).$$

The tangent space to $\mathcal{D}_n^1$ at ρ is given by $T_\rho\mathcal{D}_n^1 \equiv \{A\in M_n : A=A^*, \mathrm{Tr}(A)=0\}$, and can be decomposed as $T_\rho\mathcal{D}_n^1=(T_\rho\mathcal{D}_n^1)^c \oplus (T_\rho\mathcal{D}_n^1)^o$, where $(T_\rho\mathcal{D}_n^1)^c:=\{A\in T_\rho\mathcal{D}_n^1 : [A,\rho]=0\}$, and $(T_\rho\mathcal{D}_n^1)^o$ is the orthogonal complement of $(T_\rho\mathcal{D}_n^1)^c$, with respect to the Hilbert–Schmidt scalar product $\langle A,B\rangle:=\mathrm{Tr}(A^*B)$. Each statistically monotone metric has a unique expression (up to a constant) given by $\mathrm{Tr}(\rho^{-1}A^2)$, for $A\in(T_\rho\mathcal{D}_n^1)^c$. The following result will be used in Sec. V.

Proposition 3.3: (see Ref. 3) *Let $A\in T_\rho\mathcal{D}_n^1$ be decomposed as $A=A^c+i[\rho,U]$ where $A^c \in (T_\rho\mathcal{D}_n^1)^c$ and $i[\rho,U]\in(T_\rho\mathcal{D}_n^1)^o$. Suppose $\varphi\in\mathcal{C}^1(0,+\infty)$. Then*

$$(D_\rho\varphi)(A)=\varphi'(\rho)A^c+i[\varphi(\rho),U].$$

As proved by Lesniewski and Ruskai each monotone metric is the Hessian of a suitable relative entropy; to state this result more precisely, we introduce some notation. In what follows g is an operator convex function defined on $(0,+\infty)$ and such that $g(1)=0$. The formula

$$f(x)\equiv f_g(x):=\frac{(x-1)^2}{g(x)+xg(x^{-1})},$$

associates a normalized, symmetric operator monotone function $f=f_g$ to each g. We denote by $\Delta_{\sigma,\rho}=L_\sigma R_\rho^{-1}$ the relative modular operator. The relative g–entropy of ρ and σ is defined as

$$H_g(\rho,\sigma):=\mathrm{Tr}(\rho^{1/2}g(\Delta_{\sigma,\rho})(\rho^{1/2})).$$

H_g is a divergence on $\mathcal{D}_n$ in the sense of Refs. 14, and 1. If ρ,σ are diagonal, H_g reduces to the commutative relative g–entropy (see Ref. 9).

Theorem 3.4: (Ref. 30) *Let g be operator convex, $g(1)=0$, $f=f_g$ and $\rho\in\mathcal{D}_n$. Then*

$$-\frac{\partial}{\partial t}\frac{\partial}{\partial s}H_g(\rho+tA,\rho+sB)\bigg|_{t=s=0}=\mathrm{Tr}(A\cdot c_f(L_\rho,R_\rho)(B)).$$

To state the general formula for the scalar curvature of a monotone metric we need some auxiliary functions. In what follows c',$(\log c)'$ denote derivatives with respect to the first variable, and $c=c_f$:

$$\begin{aligned}
h_1(x,y,z)&:=\frac{c(x,y)-z\,c(x,z)\,c(y,z)}{(x-z)(y-z)c(x,z)c(y,z)},\\
h_2(x,y,z)&:=\frac{(c(x,z)-c(y,z))^2}{(x-y)^2c(x,y)c(x,z)c(y,z)},\\
h_3(x,y,z)&:=z\,\frac{(\ln c)'(z,x)-(\ln c)'(z,y)}{x-y},\\
h_4(x,y,z)&:=z\,(\ln c)'(z,x)\quad(\ln c)'(z,y),\\
h&:=h_1-\tfrac{1}{2}h_2+2h_3-h_4.
\end{aligned}\tag{3.1}$$

The functions h_i have no essential singularities if arguments coincide.

Note that $\langle A,B\rangle_\rho^f:=\mathrm{Tr}(A\cdot c_f(L_\rho,R_\rho)(B))$ defines a Riemannian metric also over $\mathcal{D}_n$ ($\mathcal{D}_n^1$ is a submanifold of codimension 1). Let $\mathrm{Scal}_f(\rho)$ be the scalar curvature of $(\mathcal{D}_n,\langle\cdot,\cdot\rangle_\rho^f)$ at ρ and $\mathrm{Scal}_f^1(\rho)$ be the scalar curvature of $(\mathcal{D}_n^1,\langle\cdot,\cdot\rangle_\rho^f)$.

Theorem 3.5: (Ref. 13) *Let $\sigma(\rho)$ be the spectrum of ρ. Then*

$$\begin{aligned}
\mathrm{Scal}_f(\rho)&=\sum_{x,y,z\in\sigma(\rho)}h(x,y,z)-\sum_{x\in\sigma(\rho)}h(x,x,x),\\
\mathrm{Scal}_f^1(\rho)&=\mathrm{Scal}_f(\rho)+\tfrac{1}{4}(n^2-1)(n^2-2).
\end{aligned}\tag{3.2}$$

IV. WIGNER–YANASE INFORMATION AS A RIEMANNIAN METRIC

Let $\rho\in\mathcal{D}_n^1$ be a density matrix and let A be a self-adjoint matrix. The Wigner–Yanase information (or skew information, information content relative to A) was defined as

$$I(\rho,A):=-\mathrm{Tr}([\rho^{1/2},A]^2),$$

where $[\cdot,\cdot]$ denotes the commutator (see Ref. 42). Consider now $g(x):=g_{wy}(x):=4(1-\sqrt{x})$. In this case

$$H_g(\rho,\sigma)=4(1-\mathrm{Tr}(\rho^{1/2}\sigma^{1/2})).$$

The associated operator monotone and Chentsov–Morotzova functions are

$$f_{wy}(x) := \frac{1}{4}(\sqrt{x}+1)^2, \quad c_{wy}(x,y) := \frac{1}{y f_{wy}\left(\frac{x}{y}\right)} = \frac{4}{(\sqrt{x}+\sqrt{y})^2}.$$

Let us consider the monotone metric,

$$\langle A,B\rangle_\rho^{wy} := \mathrm{Tr}(A\, c_{wy}(L_\rho, R_\rho)(B)).$$

A typical element of $(T_\rho D_n)^o$ has the form $\mathrm{i}[\rho,A]$, where A is self-adjoint. We have

$$\begin{aligned}
\langle \mathrm{i}[\rho,A],\mathrm{i}[\rho,A]\rangle_\rho^{wy} &= \mathrm{Tr}\,(\mathrm{i}[\rho,A]4(L_\rho^{1/2}+R_\rho^{1/2})^{-2}(\mathrm{i}[\rho,A])) \\
&= -4\ \mathrm{Tr}\,((L_\rho^{1/2}+R_\rho^{1/2})^{-1}([\rho,A])\ \ (L_\rho^{1/2}+R_\rho^{1/2})^{-1}([\rho,A])) \\
&= -4\ \mathrm{Tr}\,((L_\rho^{1/2}+R_\rho^{1/2})^{-1}\circ(L_\rho-R_\rho)(A)\ \ (L_\rho^{1/2}+R_\rho^{1/2})^{-1}\circ(L_\rho-R_\rho)(A)) \\
&= -4\ \mathrm{Tr}\,((L_\rho^{1/2}-R_\rho^{1/2})(A)\ \ (L_\rho^{1/2}-R_\rho^{1/2})(A)) \\
&= -4\ \mathrm{Tr}([\rho^{1/2},A]^2) \\
&= 4I(\rho,A),
\end{aligned}$$

and this explains why the monotone metric associated with the function $\frac{1}{4}(\sqrt{x}+1)^2$ is called the Wigner–Yanase monotone metric.

V. THE MAIN RESULT

First of all, we calculate the scalar curvature of Wigner–Yanase information using Theorem 3.5. If $f_{wy}(x) := \frac{1}{4}(\sqrt{x}+1)^2$, we write Scal^1_{wy} for Scal^1_f.

Theorem 5.1:

$$\mathrm{Scal}^1_{wy}(\rho) = \tfrac{1}{4}(n^2-1)(n^2-2).$$

Proof: Let us calculate the auxiliary functions for $c_{wy}(x,y) := 4(\sqrt{x}+\sqrt{y})^{-2}$. We get

$$h_1(x,y,z) = \frac{\sqrt{x}\ \sqrt{y}+3\ \sqrt{x}\ \sqrt{z}+3\ \sqrt{y}\ \sqrt{z}+z}{4\,(\sqrt{x}+\sqrt{y})^2\,(\sqrt{x}+\sqrt{z})\,(\sqrt{y}+\sqrt{z})},$$

$$h_2(x,y,z) = \frac{(\sqrt{x}+\sqrt{y}+2\ \sqrt{z})^2}{4\,(\sqrt{x}+\sqrt{z})^2\,(\sqrt{y}+\sqrt{z})^2},$$

$$h_3(x,y,z) = \frac{\sqrt{z}}{(\sqrt{x}+\sqrt{y})\,(\sqrt{x}+\sqrt{z})\,(\sqrt{y}+\sqrt{z})},$$

$$h_4(x,y,z) = \frac{1}{(\sqrt{x}+\sqrt{z})\,(\sqrt{y}+\sqrt{z}).}$$

Now one can verify by calculation that the symmetrization of $h_1 - \frac{1}{2}h_2$ and the symmetrization of $2\,h_3 - h_4$ vanish. Hence, by (3.1), the symmetrization of h vanishes, too. Since we sum up in formula (3.2) over all triples of eigenvalues we may replace h with its symmetrization without changing the summation result. Therefore

$$\mathrm{Scal}_{wy}(\rho) = 0\,, \quad \mathrm{Scal}^1_{wy}(\rho) = \tfrac{1}{4}(n^2-1)(n^2-2), \quad \forall \rho \in \mathcal{D}^1_n.$$

□

Remark 5.2: The fact that $\mathrm{Scal}_{wy}(\rho) = 0$ can be seen by a different approach (look at the Wigner–Yanase metric over $\mathcal{D}_n$ as the 0-geometry; see Refs. 21 and 27).

In what follows we use the pull-back approach to derive (and explain) the above formula in a direct way. Furthermore we deduce the geodesic distance and geodesic equation.

Let us denote by $\mathcal{S}$ the manifold $\{A \in M_n : \operatorname{Tr} A A^* = 4,\ A = A^*\}$. Clearly, since $\mathcal{S}$ is the intersection of the radius 2 sphere in $\mathbb{C}^{n\times n}$ and the subspace of Hermitian matrices, it is isometric with a radius 2 sphere $S_2^{n^2-1}$.

Now, let $\varphi : \mathcal{D}_n^1 \to \mathcal{S} \subset \mathbb{C}^{n\times n}$, $\varphi(\rho) := 2\sqrt{\rho}$. Then we have the following result (see Refs. 25, 18, 27, and 21).

Theorem 5.3: *The pull-back by the map φ of the natural metric on $\mathcal{S} \equiv S_2^{n^2-1}$ coincides with the Wigner–Yanase monotone metric.*

Proof: Let A and B be vectors tangent to $\mathcal{D}_n^1$ at ρ. Because $\varphi(\rho)\,\varphi(\rho) = 4\rho$ we get from the Leibniz rule $D_\rho\varphi(A)\sqrt{\rho} + \sqrt{\rho}\,D_\rho\varphi(A) = 2A$. Thus, the differential of φ at the point ρ is given by

$$D_\rho\varphi(A) = 2(L_\rho^{1/2} + R_\rho^{1/2})^{-1}(A)\,.$$

Therefore the pull-back of the real part of the Hilbert–Schmidt metric yields

$$\begin{aligned} D_\rho\varphi(\operatorname{Re}\langle\cdot,\cdot\rangle)(A,B) &= \operatorname{Re}\langle D_\rho\varphi(A), D_\rho\varphi(B)\rangle \\ &= 4\operatorname{Re}\langle (L_\rho^{1/2}+R_\rho^{1/2})^{-1}(A), (L_\rho^{1/2}+R_\rho^{1/2})^{-1}(B)\rangle \\ &= 4\,\langle A, (L_\rho^{1/2}+R_\rho^{1/2})^{-2}(B)\rangle \\ &= 4\operatorname{Tr} A\,(L_\rho^{1/2}+R_\rho^{1/2})^{-2}(B) \\ &= \operatorname{Tr} A\, c_{wy}(L_\rho, R_\rho)(B) = \langle A,B\rangle_\rho^{wy}, \end{aligned}$$

which was to be proved. □

From this result one can deduce the following.

Theorem 5.4: *For the geodesic distance, the geodesic path and the scalar curvature of Wigner–Yanase information the following formulas hold:*

(1) geodesic distance,

$$d_{wy}(\rho,\sigma) = 2\arccos(\operatorname{Tr}(\rho^{1/2}\sigma^{1/2})); \tag{5.1}$$

(2) geodesic path,

$$\gamma_{wy}^{\rho,\sigma}(t) = 2\,\frac{((1-t)\sqrt{\rho} + t\sqrt{\sigma})^2}{\operatorname{Tr}(((1-t)\sqrt{\rho} + t\sqrt{\sigma})^2)}; \tag{5.2}$$

(3) scalar curvature

$$\operatorname{Scal}_{wy}^1(\rho) = \tfrac{1}{4}(n^2-1)(n^2-2). \tag{5.3}$$

Proof: The formulas are immediate consequences of the preceding theorem and of sphere geometry. Indeed by the pull-back argument the Wigner–Yanase metric looks locally like a sphere of radius 2 of dimension (n^2-1). But for a sphere of this kind the sectional curvatures are all equal to $\frac{1}{4}$ and therefore the scalar curvature is given by $\frac{1}{4}(n^2-1)(n^2-2)$. □

One may ask if other monotone metrics are the pull-back of some function φ different from the square root. The rest of the section answers this question.

Definition 5.5: A monotone metric $\langle\cdot,\cdot\rangle_{\rho,f}$ is a pull-back metric if there exists a manifold $\mathcal{S} \subset M_n$ and a function $\varphi \in \mathcal{C}^1(0,+\infty)$ such that the pull-back metric of $\varphi : \mathcal{D}_n^1 \to \mathcal{S} \subset M_n$ coincides with $\langle\cdot,\cdot\rangle_{\rho,f}$.

Proposition 5.6: Let $\langle\cdot,\cdot\rangle_{\rho,f}$ be a monotone metric, let $c = c_f$ be the associated CM-function and let $\varphi \in \mathcal{C}^1(0,+\infty)$. We have that $\langle\cdot,\cdot\rangle_{\rho,f}$ is a pull-back metric by φ if and only if

$$\left(\frac{\varphi(x)-\varphi(y)}{x-y}\right)^2 = c(x,y). \tag{5.4}$$

Proof: Apply the proposition (3.3) to tangent vectors in $(T_\rho \mathcal{D}_n^1)^o$. □

Definition 5.7: Let $\varphi,\chi \in \mathcal{C}^1(0,+\infty)$. We say that (φ,χ) is a dual pair if there exists an operator monotone f such that

$$\frac{\varphi(x)-\varphi(y)}{x-y} \cdot \frac{\chi(x)-\chi(y)}{x-y} = c(x,y), \tag{5.5}$$

where $c=c_f$ is the CM-function associated with f.

In such a case we say that f (or c_f) is a dual function. If (φ,φ) is a dual pair with respect to f (or c_f) we say that f (or c_f) is self-dual. Obviously one has the following.

Proposition 5.8: To say that $\langle \cdot,\cdot \rangle_{\rho,f}$ is a pull-back metric by φ it is equivalent to say that f (or c_f) is self-dual with respect to φ.

Definition 5.9: Two dual pairs $(\varphi,\chi),(\tilde{\varphi},\tilde{\chi})$ are equivalent if there exist constants A_1,A_2,B_1,B_2 such that $A_1A_2=1$,

$$\tilde{\varphi} = A_1\varphi + B_1,$$

$$\tilde{\chi} = A_2\chi + B_2.$$

Obviously equivalent pairs define the same CM-function. In what follows we consider dual pairs up to this equivalence relation with the traditional abuse of language. We are ready to state the fundamental result of the theory that classifies dual pairs.

Theorem 5.10: (Refs. 23, 24, 25, 26, 36, and 19) *Let $\varphi,\chi \in \mathcal{C}^1(0,+\infty)$. Then (φ,χ) is a dual pair if and only if one of the following two possibilities hold:*

$$(\varphi(x),\chi(x)) = \left(\frac{x^p}{p},\frac{x^{1-p}}{1-p}\right), \quad p \in [-1,2]\setminus\{0,1\},$$

$$(\varphi(x),\chi(x)) = (x,\log(x)).$$

Corollary 5.11: The function $f(x)=\frac{1}{4}(\sqrt{x}+1)^2$ is the only self-dual operator monotone function, that is: the Wigner–Yanase metric is the only pull-back metric among statistically monotone metrics.

VI. CONCLUSIONS

Remark 6.1: Note that the formula (5.1) implies $d^{wy}(\rho,\sigma) \leqslant 2\pi$. An analogous inequality holds for the Bures metric (see Ref. 10, p. 311), also known as the SLD-metric: this is the monotone metric associated with $f(x)=\frac{1}{2}(1+x)$. Indeed the formula,

$$d_{\text{Bures}}(\rho,\sigma) = \sqrt{2-2\,\text{Tr}(\rho^{1/2}\sigma\rho^{1/2})^{1/2}}, \tag{6.1}$$

seems to be the only other explicit formula for a geodesic distance (in the family of statistically monotone metrics).

Remark 6.2: In general it is difficult to give explicit formulas for geodesic paths of monotone metrics. In the case of the Bures metric these geodesics can be given because they are projections of large circles on a sphere in the purifying space (see Ref. 10, p. 311 and Refs. 12, 4, and 39). For a discussion of geodesics for α-connections see Refs. 27, 28.

Remark 6.3: A classical theorem classifies the spaces of costant curvature.[29] It is not known at the moment if there are other monotone metrics of costant sectional and scalar curvature.

Remark 6.4: We have seen in the commutative case that for the Levi–Civita connection of the pull-back of the square root the decomposition is available,

$$\nabla^2 = \frac{1}{2}(\nabla^m + \nabla^e).$$

In the noncommutative case an analogous decomposition for the pull-back of the square root no longer holds. Indeed, on one hand, the use of Umegaki relative entropy $H(\rho,\sigma) = \mathrm{Tr}(\rho(\log\rho - \log\sigma))$ produces a similar decomposition, but for the Bogoliubov–Kubo–Mori metric.[31,1,22,33] On the other hand, if one uses $H_{wy}(\rho,\sigma) = 4(1 - \mathrm{Tr}(\rho^{1/2}\sigma^{1/2}))$ as a divergence on $\mathcal{D}_n^1$ and constructs the associated dualistic structure $(\langle\cdot,\cdot\rangle^{H_{wy}}, \nabla^{H_{wy}}, \nabla^{H_{wy}})$ (again following the lines of Refs. 14 and 1), then the construction is trivial, namely, the dual connections both coincide with the Levi–Civita connection of the Wigner–Yanase information. This is easily seen on $\mathcal{P}_n$ where $H_g(\rho,\sigma)$ reduces to Csiszar relative g-entropy: it is known that such an entropy induces the α-geometry where α is given by the formula $\alpha = 3 + 2g'''(1)/g''(1)$ (see Ref. 1, p. 57). For $g = 4(1 - \sqrt{x})$ this gives $\alpha = 0$, that is, the Fisher information case (see also Ref. 21).

ACKNOWLEDGMENTS

This research has been supported by the italian MIUR program ''Quantum Probability and Infinite Dimensional Analysis'' 2001–2002. It is a pleasure to thank J. Dittmann for a number of valuable conversations on this subject and for permitting us to reproduce in this paper the proof of Theorem 5.1.

[1] Amari, S. and Nagaoka, S., *Methods of Information Geometry* (American Mathematical Society and Oxford University Press, Oxford, 2000).
[2] Belavkin, V. P., Hirota, O., and Hudson, R. L. *Quantum Communications and Measurement* (Plenum, New York, 1995).
[3] Bhatia, R., *Matrix Analysis* (Springer-Verlag, New York, 1997).
[4] Braunstein, S. L., and Caves, C. M., *Geometry of Quantum States*, pp. 21–30 in Ref. 2.
[5] Carlen, E., ''Superadditivity of Fisher's information and logarithmic Sobolev inequalities,'' J. Funct. Anal. **101**, 194–211 (1991).
[6] Carlen, E., and Gangbo, W., ''Costrained steepest descent in the 2-Wasserstein metric,'' preprint, 2002.
[7] Chentsov, N., *Statistical Decision Rules and Optimal Inference* (American Mathematical Society, Providence, RI, 1982).
[8] Chentsov, N. and Morozova, E., ''Markov invariant geometry on state manifolds,'' J. Sov. Math. **56**, 2648–2669 (1991).
[9] Csiszar, I., ''Information type measures of difference of probability distribution and indirect observation,'' Stud. Sci. Math. Hung. **2**, 299–318 (1967).
[10] Dittmann, J., ''On the Riemannian metric on the space of density matrices,'' Rep. Math. Phys. **36**, 309–315 (1995).
[11] Dittmann, J., ''The scalar curvature of the Bures metric on the space of density matrices,'' J. Geom. Phys. **31**, 16–24 (1999).
[12] Dittmann, J. and Uhlmann, A., ''Connections and metrics respecting standard purification,'' J. Math. Phys. **40**, 3246–3267 (1999).
[13] Dittmann, J., ''On the curvature of monotone metrics and a conjecture concering the Kubo–Mori metric,'' Linear Algebr. Appl. **315**, 83–112 (2000).
[14] Eguchi, S., ''Geometry of minimum contrast,'' Hiroshima Math. J. **22**, 631–647 (1992).
[15] Friedrich, T., ''Die Fisher-Information und symplektische strukturen,'' Math. Nachr. **153**, 273–296 (1991).
[16] Gibilisco, P. and Isola, T., ''Connections on statistical manifolds of density operators by geometry of noncommutative L^p-spaces,'' Infinite Dimen. Anal., Quantum Probab., Relat. Top. **2**, 169–178 (1999).
[17] Gibilisco, P. and Isola, T., ''Monotone metrics on statistical manifolds of density matrices by geometry of noncommutative L^2-spaces,'' in: *Disordered and Complex Systems*, edited by A. C. Coolen, L. Hughston, P. Sollich, and R. F. Streater (AIP, Melville, NY, 2001), pp. 129–140.
[18] Gibilisco, P. and Isola, T., ''A characterisation of Wigner-Yanase skew information among statistically monotone metrics,'' Infinite Dimen. Anal., Quantum Probab., Relat. Top. **4**, 553–557 (2001).
[19] Gibilisco, P. and Isola, T., ''On the characterisation of dual statistically monotone metrics,'' preprint [math. PR/03059], 2003.
[20] Gibilisco, P. and Pistone, G., ''Connections on non-parametric statistical manifolds by Orlicz space geometry,'' Infinite Dimen. Anal., Quantum Probab., Relat. Top. **1**, 325–347 (1998).
[21] Grasselli, M. R., ''Duality, monotonicity and the Wigner–Yanase–Dyson metrics,'' preprint, math-ph/0212022, 2002.
[22] Grasselli, M. R. and Streater, R. F., ''On the uniqueness of Chentsov metric in quantum information geometry,'' Infinite Dimen. Anal., Quantum Probab., Relat. Top. **4**, 173–182 (2001).
[23] Hasegawa, H., ''α-divergence of the non-commutative information geometry,'' Rep. Math. Phys. **33**, 87–93 (1993).
[24] Hasegawa, H., ''Non-commutative extension of the information geometry,'' pp. 327–337 in Ref. 2.
[25] Hasegawa, H. and Petz, D., ''Non-commutative extension of information geometry, II,'' in *Quantum Communication, Computing and Measurement*, edited by D. Hirota *et al.* (Plenum, New York, 1997), pp. 109–118.
[26] Hasegawa, H., ''Dual geometry of the Wigner–Yanase–Dyson information content,'' Preprint, 2003.

[27] Jenčova, A., ''Geometry of quantum states: dual connections and divergence functions,'' Rep. Math. Phys. **47**, 121–138 (2001).
[28] Jenčova, A., ''Quantum information geometry and standard purification,'' J. Math. Phys. **43**, 2187–2201 (2002).
[29] Kobayashi, X. and Nomizu, X., *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. I.
[30] Lesniewski, A. and Ruskai, M. B., ''Monotone Riemannian metrics and relative entropy on noncommutative probability spaces,'' J. Math. Phys. **40**, 5702–5724 (1999).
[31] Nagaoka, H., ''Differential geometric aspects of quantum state estimation and relative entropy,'' pp. 449–452 in Ref. 2.
[32] Nielsen, M. and Chuang, I., *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[33] Petz, D., ''Geometry of canonical correlation on the state space of a quantum system,'' J. Math. Phys. **35**, 780–795 (1994).
[34] Petz, D., ''Monotone metrics on matrix space,'' Linear Algebr. Appl. **244**, 81–96 (1996).
[35] Petz, D., ''Covariance and Fisher information in quantum mechanics,'' J. Phys. A **35**, 929–939 (2002).
[36] Petz, D. and Hasegawa, H., ''On the Riemannian metric of α-entropies of density matrices,'' Lett. Math. Phys. **38**, 221–225 (1996).
[37] Petz, D. and Sudár, C., ''Geometries of quantum states,'' J. Math. Phys. **37**, 2662–2673 (1996).
[38] Ruskai, M. B., ''Contraction of Riemannian metrics and related distance measures on pairs of qubit states,'' preprint, 2002.
[39] Uhlmann, A., ''Parallel transport and quantum holonomy,'' Rep. Math. Phys. **24**, 229–240 (1986).
[40] Uhlmann, A., ''Density operators as an arena for differential geometry,'' Rep. Math. Phys. **33**, 253–263 (1993).
[41] Uhlmann, A., ''Geometric phases and related structures,'' Rep. Math. Phys. **36**, 461–481 (1995).
[42] Wigner, E. and Yanase, Al., ''Information content of distribution,'' Proc. Natl. Acad. Sci. U.S.A. **49**, 910–918 (1963).
[43] Wootters, W. V., ''Statistical distance and Hilbert space,'' Phys. Rev. D **23**, 357–362 (1981).