On the Uniqueness of the Chentsov Metric in Quantum Information Geometry

M. R. Grasselli[*]   R. F. Streater

Dept. of Mathematics,
King's College,
Strand,
London, WC2R 2LS

11/06/2000

Abstract

We show that, in finite dimensions, the only monotone metrics on the space of invertible density matrices for which the $(+1)$ and (-1) affine connections are mutually dual are constant multiples of the Bogoliubov-Kubo-Mori metric.

1   Introduction

The first task of quantum information geometry is to extend to non-commutative probability spaces the uniqueness and rigidity of the geometrical structures used in the classical version of the theory. This includes, for instance, the uniqueness result of Chentsov concerning the Fisher metric [4] and the two equivalent definitions of α-connections given by Amari [1]: either using an α-embedding of the form $p \mapsto \frac{2}{1-\alpha} p^{\frac{1-\alpha}{2}}$ or as the convex mixture of the exponential and mixture connections.

As for the first problem, Chentsov himself made the first attempt to find the possible Riemannian metrics on a quantum information manifold with the property of having its line element reduced under stochastic maps [13] (cf [18]). This monotonicity property, which is the quantum analogue of being reduced under Markov morphisms, was later investigated by Petz [18]. Unlike the classical case, he found that there are infinitely many Riemannian metrics satisfying it. Concerning the second problem, several definitions of α-connections have been proposed [15, 7, 5], both for finite and infinite dimensional quantum systems. Since some of these definitions involve finding dual connections with respect to some chosen Fisher metric, it is clear that the multitude of possible candidates for the metric encourages the appearance of non-equivalent definitions for the α-connections.

We take the position that quantum information manifolds are equipped with two natural flat connections: the mixture connection, obtained from the linear structure of trace class operators themselves, and the exponential connection, obtained when combinations of states are performed by adding their logarithms [6, 21]. In §2 we present our preferred definition of the ± 1-connections in finite dimensional quantum information manifolds.

Following Amari [1, 2], we consider duality as a fundamental structure to be explored. Thus, given these two connections, we should ask what are the Riemannian metrics that make them dual. We give our answer to this question in §3. In §4, we combine this result with Petz's characterisation of monotone metrics in finite dimensional quantum systems to find that the *BKM* metric [3, 11, 12] is, up to a factor, the unique monotone Riemannian metric with respect to which the exponential and mixture connections are dual.

2 The Exponential and Mixture Connections

Let $\mathcal{H}^N$ be a finite dimensional complex Hilbert space, $\mathcal{A}$ the subspace of self-adjoint operators and $\mathcal{M}$ the set of all invertible density operators on $\mathcal{H}^N$. Then $\mathcal{A}$ is an N^2-dimensional real vector space and $\mathcal{M}$ is an n-dimensional submanifold with $n = N^2 - 1$. Defining the 1-embedding of $\mathcal{M}$ into $\mathcal{A}$ as

$$\ell_1 \;:\; \mathcal{M} \to \mathcal{A}$$
$$\rho \mapsto \log \rho,$$

we can use the linear space structure of $\mathcal{A}$ to obtain a representation of the tangent bundle of $\mathcal{M}$ in terms of operators in $\mathcal{A}$ even though $\mathcal{M}$ is not a vector space itself. At each point $\rho \in \mathcal{M}$, consider the subspace $\mathcal{A}_\rho = \{A \in \mathcal{A} : \mathrm{Tr}(\rho A) = 0\}$ of $\mathcal{A}$, called the space of 'scores'; we define the isomorphism

$$(\ell_1)_{*(\rho)} \;:\; T_\rho \mathcal{M} \to \mathcal{A}_\rho$$
$$v \mapsto (\ell_1 \circ \gamma)'(0),$$

where $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ is a curve in the equivalence class of the tangent vector v. We call this isomorphism the 1-representation of the tangent space $T_\rho \mathcal{M}$. If $(\theta^1, \ldots, \theta^n)$ is a coordinate system for $\mathcal{M}$, put $\partial_i := \partial/\partial\theta^i$, $i = 1, \ldots, n$, evaluated at ρ. Then the 1-representation of the basis $\{\partial_1, \ldots, \partial_n\}$ of $T_\rho \mathcal{M}$ is $\left\{ \frac{\partial \log \rho}{\partial \theta^1}, \ldots, \frac{\partial \log \rho}{\partial \theta^n} \right\}$. The 1-representation of a vector field X on $\mathcal{M}$ is therefore the $\mathcal{A}$-valued function $(X)^{(1)}$ given by $(X)^{(1)}(\rho) = (\ell_1)_{*(\rho)} X_\rho$.

The exponential or 1-connection is the connection obtained from the 1-embedding through the following parallel transport [21]

$$\tau^{(1)}_{\rho_0, \rho_1} \;:\; T_{\rho_0} \mathcal{M} \to T_{\rho_1} \mathcal{M}$$
$$v \mapsto (\ell_1)^{-1}_{*(\rho_1)} \left((\ell_1)_{*(\rho_0)} v - \mathrm{Tr}[\rho_1 (\ell_1)_{*(\rho_0)} v] \right).$$

Giving the parallel transport in a neighbourhood of ρ is equivalent to specifying the covariant derivative. It is readily verified that the 1-representation of the 1-covariant derivative, applied to the vector field $\partial_j := \partial/\partial\theta^j$, is

$$\left(\nabla^{(1)}_{\partial_i} \frac{\partial}{\partial \theta^j} \right)^{(1)} = \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j} - \mathrm{Tr}\left(\rho \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j} \right) \tag{1}$$

At a more abstract level, the construction above corresponds to making $\mathcal{M}$ into an affine space and endowing it with the *natural* flat connection [14]. Rather than exploring this concept further, we benefit from dealing with finite dimensional spaces and prove that the 1-connection is flat by exhibiting an affine coordinate system for it. Let $\{\mathbf{1}, X_1, \ldots, X_n\}$ be a basis for $\mathcal{A}$. Since $\log \rho \in \mathcal{A}$, there exist real numbers $\{\theta^1, \ldots, \theta^n, \Psi\}$ such that

$$\log \rho = \theta^1 X_1 + \cdots + \theta^n X_n - \Psi \mathbf{1},$$

that is,
$$\rho = \exp\left(\theta^1 X_1 + \cdots + \theta^n X_n - \Psi \mathbf{1}\right). \tag{2}$$

The normalisation condition $\mathrm{Tr}\rho = 1$ means, however, that only n among these numbers are independent, so we choose $\Psi \equiv \Psi(\theta)$ to be the one determined by the others. Then $\theta = (\theta^1, \ldots, \theta^n)$ form a 1-affine coordinate system, as can be seen from the following calculation

$$
\begin{aligned}
\left(\nabla^{(1)}_{\partial_i} \frac{\partial}{\partial \theta^j}\right)^{(1)} &= \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j} - \mathrm{Tr}\left(\rho \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j}\right) \\
&= -\frac{\partial^2 \Psi}{\partial \theta^i \partial \theta^j}(\theta) + \mathrm{Tr}\left(\rho \frac{\partial^2 \Psi}{\partial \theta^i \partial \theta^j}(\theta)\right) \\
&= 0.
\end{aligned}
$$

Now let $\mathcal{A}_0$ be the subspace of traceless operators in $\mathcal{A}$. Consider the -1-embedding

$$
\begin{aligned}
\ell_{-1} &: \mathcal{M} \to \mathcal{A} \\
&\quad \rho \mapsto \rho,
\end{aligned}
$$

and define, at each $\rho \in \mathcal{M}$, the -1-representation of tangent vectors as

$$
\begin{aligned}
(\ell_{-1})_{*(\rho)} &: T_\rho \mathcal{M} \to \mathcal{A}_0 \\
&\quad v \mapsto (\ell_{-1} \circ \gamma)'(0),
\end{aligned}
$$

where $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ is again a curve in the equivalence class of the tangent vector v. In coordinates, the -1-representation of the basis $\{\partial_1, \ldots, \partial_n\}$ of $T_\rho \mathcal{M}$ is $\left\{\frac{\partial \rho}{\partial \theta^1}, \ldots, \frac{\partial \rho}{\partial \theta^n}\right\}$. As before, the -1-representation of a vector field X on $\mathcal{M}$ is an $\mathcal{A}_0$-valued function denoted by $(X)^{(-1)}$ or X^-.

We obtain the mixture or -1-connection by defining the parallel transport

$$
\begin{aligned}
\tau^{(-1)}_{\rho_0, \rho_1} &: T_{\rho_0} \mathcal{M} \to T_{\rho_1} \mathcal{M} \\
&\quad v \mapsto (\ell_{-1})^{-1}_{*(\rho_1)}\left((\ell_{-1})_{*(\rho_0)} v\right),
\end{aligned}
$$

and we find that its covariant derivative in the direction ∂_i is

$$\left(\nabla^{(-1)}_{\partial_i} \frac{\partial}{\partial \theta^j}\right)^{(-1)} = \frac{\partial^2 \rho}{\partial \theta^i \partial \theta^j}. \tag{3}$$

If we equip $\mathcal{A}$ with the trace norm, then the -1-embedding maps $\mathcal{M}$ into the unit sphere $\mathcal{S}$ of $\mathcal{A}$, and the -1-connection given here is nothing but the projection onto $\mathcal{S}$ of the *natural* flat connection in this space. It turns out that the unit sphere with respect to the trace norm is flat in $\mathcal{A}$, hence the -1-connection is flat on $\mathcal{M}$. Again, it is a convenience of the finite dimensional case that we can prove flatness of the -1-connection by direct construction of a -1-affine coordinate system. Suppose $\{X_1, \ldots, X_{n+1}\}$ is a normalised basis for $\mathcal{A}$, then there exist real numbers $(\xi_1, \ldots, \xi_{n+1})$ such that

$$\rho = \xi_1 X_1 + \cdots + \xi_{n+1} X_{n+1}. \tag{4}$$

Since $\mathrm{Tr}\rho = 1$, we can take $\xi_{n+1} = 1 - (\xi_1 + \cdots + \xi_n)$ as a function of the first n independent parameters $\xi_1, \ldots, \xi_n$. Then $\xi = (\xi_1, \ldots, \xi_n)$ is a -1-affine coordinate system for $\mathcal{M}$, because

$$
\begin{aligned}
\left(\nabla^{(-1)}_{\frac{\partial}{\partial \xi_i}} \frac{\partial}{\partial \xi_j} \right)^{(-1)} &= \frac{\partial^2 \rho}{\partial \xi_i \partial \xi_j} \\
&= \frac{\partial}{\partial \xi_i}(X_j - X_{n+1}) \\
&= 0.
\end{aligned}
$$

3 Duality and the BKM Metric

Two connections ∇ and ∇^* on a Riemannian manifold $(\mathcal{M}, g)$ are dual with respect to g if and only if

$$
Xg(Y,Z) = g\left(\nabla_X Y, Z\right) + g\left(Y, \nabla^*_X Z\right), \tag{5}
$$

for any vector fields X, Y, Z on $\mathcal{M}$ [1, 14]. Equivalently, if $\tau_{\gamma(t)}$ and $\tau^*_{\gamma(t)}$ are the respective parallel transports along a curve $\{\gamma(t)\}_{0 \leq t \leq 1}$ on $\mathcal{M}$, with $\gamma(0) = \rho$, then ∇ and ∇^* are dual with respect to g if and only if for all $t \in [0, 1]$,

$$
g_\rho(Y, Z) = g_{\gamma(t)}\left(\tau_{\gamma(t)} Y, \tau^*_{\gamma(t)} Z\right). \tag{6}
$$

Given any connection ∇ on $(\mathcal{M}, g)$, we can always find a unique connection ∇^* such that ∇ and ∇^* are dual with respect to g. On the other hand, given two connections ∇ and ∇^*, we can ask what are the possible Riemannian metrics g with respect to which they are dual. In particular, we want to explore this question for the case of the exponential and mixture connections on a manifold of density matrices.

A different concept of duality also used by Amari [1] is that of dual coordinate systems, regardless of any connection. Two coordinate systems $\theta = (\theta^i)$ and $\eta = (\eta_i)$ on a Riemannian manifold $(\mathcal{M}, g)$ are dual with respect to g if and only if their natural bases for $T_\rho \mathcal{M}$ are *biorthogonal* at every point $\rho \in \mathcal{M}$, that is,

$$
g\left(\frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \eta_j} \right) = \delta^i_j.
$$

Equivalently, $\theta = (\theta^i)$ and $\eta = (\eta_i)$ are dual with respect to g if and only if

$$
g_{ij} = \frac{\partial \eta_i}{\partial \theta^j} \quad \text{and} \quad g^{ij} = \frac{\partial \theta_i}{\partial \eta^j}
$$

at every point $\rho \in \mathcal{M}$, where, as usual, $g^{ij} = (g_{ij})^{-1}$.

The following theorem gives a characterisation of dual coordinate systems in terms of potential functions, thus bringing convexity theory and the related duality with respect to Legendre transforms into the discussion.

Theorem 7 (Amari, 1985) *When a Riemannian manifold $(\mathcal{M}, g)$ has a pair of dual coordinate systems (θ, η), there exist potential functions $\Psi(\theta)$ and $\Phi(\eta)$ such that*

$$
g_{ij}(\theta) = \frac{\partial^2 \Psi(\theta)}{\partial \theta^i \partial \theta^j} \quad \text{and} \quad g^{ij} = \frac{\partial^2 \Phi(\eta)}{\partial \eta_i \partial \eta_j}.
$$

Conversely, when either potential function Ψ or Φ exists from which the metric is derived by differentiating it twice, there exist a pair of dual coordinate systems. The dual coordinate systems and the potential functions are related by the following Legendre transforms

$$\theta^i = \frac{\partial\Phi(\eta)}{\partial\eta_i}, \quad \eta_i = \frac{\partial\Psi(\theta)}{\partial\theta^i}$$

and

$$\Psi(\theta) + \Phi(\eta) - \theta^i\eta_i = 0$$

In contrast to the case of dual connections, dual coordinate systems do not necessarily exist on every Riemannian manifold [1]. When the additional property of flatness is required, the following theorem provides a link between the two concepts of duality. In the sense used in this paper, a connection ∇ on manifold $\mathcal{M}$ is said to be flat if $\mathcal{M}$ admits a global ∇-affine coordinate system. This is equivalent to its curvature and torsion both being zero.

Theorem 8 (Amari, 1985) *Suppose that ∇ and ∇^* are two flat connections on a manifold $\mathcal{M}$. If they are dual with respect to a Riemannian metric g on $\mathcal{M}$, then there exists a pair (θ, η) of dual coordinate systems such that θ is ∇-affine and η is a ∇^*-affine.*

We now return to our manifold $\mathcal{M}$ of density matrices and consider the problem of finding a unique Riemannian metric for it. Using either the 1 or the -1 representation of the tangent bundle $T\mathcal{M}$, we define a Riemannian metric on $\mathcal{M}$ by a smooth assignment of an inner product $\langle \cdot, \cdot \rangle_\rho$ in $\mathcal{A} \subset B(\mathcal{H}^N)$ for each point $\rho \in \mathcal{M}$. The *BKM* (Bogoliubov-Kubo-Mori) metric is the Riemannian metric on $\mathcal{M}$ obtained from the *BKM* inner product [17]. If $A^{(1)}, B^{(1)}$ and $A^{(-1)}, B^{(-1)}$ are, respectively, the 1 and -1 representations of $A, B \in T_\rho\mathcal{M}$, then the *BKM* metric has the following equivalent expressions:

$$
\begin{aligned}
g_\rho^B(A, B) &= \text{Tr}\left(A^{(-1)}B^{(1)}\right) \\
&= \int_0^1 \text{Tr}\left(\rho^\lambda A^{(1)}\rho^{1-\lambda}B^{(1)}\right) d\lambda \\
&= \int_0^\infty \text{Tr}\left(\frac{1}{t+\rho}A^{(-1)}\frac{1}{t+\rho}B^{(-1)}\right) dt
\end{aligned}
\tag{9}
$$

It is well known that the exponential and mixture connections are dual with respect to the *BKM* metric [15, 8]. A natural question is whether it is the unique metric with this property. The next theorem tells us how much uniqueness can be achieved from duality alone.

Theorem 10 *If the connections $\nabla^{(1)}$ and $\nabla^{(-1)}$ are dual with respect to a Riemannian metric g on $\mathcal{M}$, then there exist a constant (independent of ρ) $n \times n$ matrix M, such that $(g_\rho)_{ij} = \sum_{k=1}^n M_i^k (g_\rho^B)_{kj}$, in some 1-affine coordinate system.*

Proof: Since the two connections are flat, by theorem 8, there exist dual coordinate systems (θ, η) such that θ is $\nabla^{(1)}$-affine and η is $\nabla^{(-1)}$-affine. Thus, applying theorem 7, there exist a potential function $\Psi(\theta)$ such that

$$g_{ij}(\theta) = \frac{\partial^2\Psi(\theta)}{\partial\theta^i\partial\theta^j}$$

5

and
$$\eta_i = \frac{\partial \Psi(\theta)}{\partial \theta^i}.$$

On the other hand, since θ is $\nabla^{(1)}$-affine, there exist linearly independent operators $\{\mathbf{1}, X_1, \ldots, X_n\}$ such that

$$\rho = \exp\left(\theta^1 X_1 + \cdots + \theta^n X_n - \tilde{\Psi}(\theta)\mathbf{1}\right), \tag{11}$$

where

$$\tilde{\Psi}(\theta) = \log \mathrm{Tr}\left[\exp\left(\theta^1 X_1 + \cdots + \theta^n X_n\right)\right]. \tag{12}$$

Without loss of generality, we can assume that the operators $X_1, \ldots, X_n$ are traceless, since if we add multiples of the identity to any X_j in (11), we can still have the same parameters θ as coordinates for the same point ρ just by modifying the function $\tilde{\Psi}$. But any such set of operators define a $\nabla^{(-1)}$-affine coordinate system through the formula

$$\tilde{\eta}_i = \mathrm{Tr}(\rho X_i),$$

because the latter are affinely related to the ξ coordinates defined in section 2 (with $X_{n+1} = \mathbf{1}/n$). Differentiating $\tilde{\Psi}$ with respect to θ^i we obtain

$$\frac{\partial \tilde{\Psi}(\theta)}{\partial \theta^i} = \mathrm{Tr}(\rho X_i) = \tilde{\eta}_i.$$

Thus $\tilde{\eta}_i = \frac{\partial \tilde{\Psi}(\theta)}{\partial \theta^i}$ and $\eta_i = \frac{\partial \Psi(\theta)}{\partial \theta^i}$ are two $\nabla^{(-1)}$-affine coordinate systems, so they must be related by an affine transformation [14]. So there exist an $n \times n$ matrix M and numbers $(a_1, \ldots, a_n)$ such that

$$\eta_i = \sum_{k=1}^{n} M_i^k \tilde{\eta}_k + a_i,$$

that is,

$$\frac{\partial \Psi(\theta)}{\partial \theta^i} = \sum_{k=1}^{n} M_i^k \frac{\partial \Psi(\theta)}{\partial \theta^k} + a_i,$$

and differentiating this equation with respect to θ^j gives

$$g_{ij}(\theta) = \frac{\partial^2 \Psi(\theta)}{\partial \theta^i \partial \theta^j} = \sum_{k=1}^{n} M_i^k \frac{\partial^2 \tilde{\Psi}(\theta)}{\partial \theta^j \partial \theta^k}. \tag{13}$$

But we can calculate the second derivative of $\tilde{\Psi}$ directly from (12), obtaining

$$\begin{aligned}
\frac{\partial^2 \tilde{\Psi}(\theta)}{\partial \theta^j \partial \theta^k} &= \int_0^1 \mathrm{Tr}\left(\rho^\lambda \frac{\partial \log \rho}{\partial \theta^j} \rho^{1-\lambda} \frac{\partial \log \rho}{\partial \theta^k}\right) d\lambda \\
&= g_\rho^B\left(\frac{\partial}{\partial \theta^j}, \frac{\partial}{\partial \theta^k}\right) \\
&= g_{jk}^B(\theta).
\end{aligned}$$

Inserting this in (13) completes the proof. $\qquad\square$

4 The Condition of Monotonicity

We have seen in the previous section that requiring duality between the exponential and mixture connections reduces the set of possible Riemannian metrics on $\mathcal{M}$ to matrix multiples of the BKM metric. We now investigate the effect of imposing a monotonicity property on this set.

If we use the -1-representation to define a Riemannian metric g on $\mathcal{M}$ by means of the inner product $\langle \cdot, \cdot \rangle_\rho$ in $\mathcal{A} \subset B(\mathcal{H}^N)$, then we say that g is *monotone* if and only if

$$\left\langle S(A^{(-1)}), S(A^{(-1)}) \right\rangle_{S(\rho)} \leq \left\langle A^{(-1)}, A^{(-1)} \right\rangle_\rho, \tag{14}$$

for every $\rho \in \mathcal{M}$, $A \in T_\rho \mathcal{M}$, and every completely positive, trace preserving map $S : \mathcal{A} \to \mathcal{A}$.

In a series of papers [16, 18, 20], Petz has given a complete characterisation of monotone metrics on full matrix spaces in terms of operator monotone functions. The Chentsov condition, however, is defined for metrics on the space $\mathcal{M}$ of faithful *states*, and must first be extended to $\mathcal{A}$ before we can use Petz's results. Let $\widehat{\mathcal{M}}$ be the manifold of faithful weights (the positive definite matrices). Let g be a metric on $T\mathcal{M}$. We can extend g to $\mathcal{A} \simeq T\widehat{\mathcal{M}}$ as follows. At $\rho \in \mathcal{M}$ and $\hat{A}, \hat{B} \in T\widehat{\mathcal{M}}_\rho$, put $\hat{A}^{(-1)} = A_0\rho + A^-$, where $A_0 = \mathrm{Tr}\,\hat{A}^{(-1)} \in \mathbf{R}$ and $\mathrm{Tr}\,A^- = 0$, and similarly for B. Then put

$$\hat{g}_\rho(\hat{A}, \hat{B}) := A_0 B_0 + g_\rho(A^-, B^-). \tag{15}$$

For g^B this extension coincides with that given by eq. (9). More generally, if g is monotone on $T\mathcal{M}$, then $\hat{g}$ is monotone on $T\widehat{\mathcal{M}}$. For, let S be a trace-preserving completely positive map on $T\widehat{\mathcal{M}}$, and $\rho \in \mathcal{M}$. Then S maps $T\mathcal{M}$ to itself, and

$$\hat{g}_{S\rho}(S\hat{A}, S\hat{A}) = A_0^2 + g_{S\rho}(SA^-, SA^-) \leq A_0^2 + g_\rho(A^-, A^-) = \hat{g}_\rho(A, A).$$

For any metric $\hat{g}$ on $T\widehat{\mathcal{M}}$, and putting $\hat{A}^{(-1)} = A_0\rho + A^-$, we define the positive (super) operator K_ρ on $\mathcal{A}$ by

$$\hat{g}_\rho(\hat{A}, \hat{B}) = \left\langle \hat{A}^{(-1)}, K_\rho\left(\hat{B}^{(-1)}\right) \right\rangle_{HS} = \mathrm{Tr}\left(\hat{A}^{(-1)} K_\rho\left(\hat{B}^{(-1)}\right)\right). \tag{16}$$

Note that our K is denoted K^{-1} by Petz. He also defines the (super) operators, $L_\rho X := \rho X$ and $R_\rho X := X\rho$, for $X \in \mathcal{A}$, which are also positive. Then he proved [18]

Theorem 17 (Petz) *A Riemannian metric g on $\mathcal{A}$ is monotone if and only if*

$$K_\rho = \left(R_\rho^{1/2} f(L_\rho R_\rho^{-1}) R_\rho^{1/2}\right)^{-1},$$

where K_ρ is defined in (16) and $f : R^+ \to R^+$ is an operator monotone function satisfying $f(t) = t f(t^{-1})$.

In particular, the BKM metric is monotone and its corresponding operator monotone function is

$$f^B(t) = \frac{t-1}{\log t}.$$

Combining this characterisation with our theorem (10), we obtain the following improved uniqueness result.

Theorem 18 *If the connections $\nabla^{(1)}$ and $\nabla^{(-1)}$ are dual with respect to a monotone Riemannian metric g on $\mathcal{M}$, then g is a scalar multiple of the BKM metric.*

Proof: We first extend the $(N^2 - 1) \times (N^2 - 1)$ matrix M of theorem (10) by one row and column, to a matrix $\hat{M}$, where $\hat{M}_{ij} = M_{ij}$, $(1 \leq i, j \leq N^2 - 1)$, and

$$\hat{M}_{i,N^2} = \hat{M}_{N^2,i} = \delta_{i,N^2}, (i = 1, \ldots, N^2).$$

Now, in the coordinates $\hat{A}^{(-1)} = A_0 \rho + A^-$ of eq. (15), the extension of g to $\hat{g}$ in the orthogonal direction ρ is the same for all g. In particular,

$$\hat{g}(A_0\rho, B^-) = \hat{g}^B(A_0\rho, B^-) = 0; \qquad \hat{g}(A_0\rho, B_0\rho) = \hat{g}^B(A_0\rho, B_0\rho) = A_0 B_0.$$

It follows that the equation proved in theorem (10), when expressed in the -1-affine coordinates η (by taking inverses and using theorem 7), gives the following relation between the kernels of $\hat{g}$ and $\hat{g}^B$, say K^g and K^B respectively,

$$K_\rho^g = \hat{M}^t K_\rho^B. \tag{19}$$

Therefore, if f^g and f^B are the operator monotone functions corresponding respectively to g and g^B, from theorem 17, we have

$$\begin{aligned}
\left(R_\rho^{1/2} f^g (L_\rho R_\rho^{-1}) R_\rho^{1/2}\right)^{-1} &= \hat{M}^t \left(R_\rho^{1/2} f^{BKM} (L_\rho R_\rho^{-1}) R_\rho^{1/2}\right)^{-1} \\
\left(R_\rho^{1/2} f^g (L_\rho R_\rho^{-1}) R_\rho^{1/2}\right) \hat{M}^t &= \left(R_\rho^{1/2} f^{BKM} (L_\rho R_\rho^{-1}) R_\rho^{1/2}\right) \\
\hat{M}^t &= f^g (L_\rho R_\rho^{-1})^{-1} f^{BKM} (L_\rho R_\rho^{-1}),
\end{aligned}$$

as everything commutes. Thus, the operator $\hat{M}$ is given as a function of the operator $L_\rho R_\rho^{-1}$, but it is itself independent of the point ρ, so we conclude that it must be a scalar multiple of the identity operator. $\qquad\square$

5 Discussion

We have proved that, in a finite dimensional quantum system, the *BKM* metric is, up to a scalar factor, the unique monotone Riemannian metric with respect to which the exponential and mixture connections are dual. It should be emphasised that we have considered the manifold of *all* invertible density operators. The uniqueness result might not be true for a submanifold of $\mathcal{M}$, such as occurs in parametric estimation theory.

   We should compare our uniqueness theorem with the results of Hasegawa [9] and Petz and Hasegawa [10, 19], who obtained a family of metrics $\{g^\alpha\}$, each corresponding to a divergence based on the corresponding Wigner-Yanase information, as generalised by Dyson. They show that these metrics are monotone, and that the α-divergence 'admits dual connections', in a sense they specify. This does not contradict our result; the ± 1-connections are not dual relative to g^α unless $\alpha = \pm 1$, which is the *BKM* case.

   Having fixed the metric, the next step in the development of the theory is to define the α-connections, for $\alpha \in (-1, 1)$. We look for a definition that makes $\nabla^{(\alpha)}$ and $\nabla^{(-\alpha)}$ dual and such that the extended manifold $\widehat{\mathcal{M}}$ is α-flat. Only then we can try to find the quantum analogue of the α-divergence. Duality is easily achieved if we define the α-connections as the convex mixture $\nabla^{(\alpha)} = \frac{1+\alpha}{2}\nabla^{(1)} + \frac{1-\alpha}{2}\nabla^{(-1)}$. Flatness, on the other hand, is more apparent if we use the α-embedding. We have not yet been able to prove that the two definitions are equivalent. All these questions are deferred to a subsequent paper.

References

[1] S.-I. Amari, **Differential Geometric Methods in Statistics**, *Lecture Notes in Statistics*, **28**, Springer-Verlag, New York, 1985.

[2] S.-I. Amari, *Information Geometry*, Contemp. Math., **203**, 1985, pp. 81-95.

[3] N. N. Bogoliubov, Phys. Abh. Sov. Union, **1**, 229, 1962.

[4] N. N. Čencov, **Statistical Decision Rules and Optimal Inferences**, *Translations of Mathematical Monographs*, American Mathematical Society, Providence, 1982.

[5] P. Gibilisco and T. Isola, *Connections on Statistical Manifolds of Density Operators by Geometry of Noncommutative L^p-spaces*, Inf. Dim. Analysis, Quant. Prob. and Rel. Top., **2**, 169-178, 1999.

[6] M. R. Grasselli and R. F. Streater, *The Quantum Information Manifold for ε-bounded Forms*, to appear in Rep. Math. Phys., math-hp/9910031.

[7] H. Hasegawa, *Noncommutative Extension of the Information Geometry*, in **Quantum Communication and Measurement**, eds. V. P. Balavkin, O. Hirota and R. L. Hudson, Plenum Press, 1995; pp. 327-337.

[8] H. Hasegawa, *Exponential and mixture families in quantum statistics: dual structures and unbiased parameter estimation.*, Rep. Math. Phys., **39**, 49-68, 1997.

[9] H. Hasegawa, *α-divergence of the non-commutative information geometry*, Rep. Math. Phys., **33**, 87-93, 1993.

[10] H. Hasegawa and D. Petz, *Non-commutative extension of information geometry II*, in **Quantum Communication and Measurement**, eds. V. P. Belavkin, O. Hirota and R. L. Hudson, Plenum Press, pp 109-118, 1997.

[11] R. Kubo, *Statistical Mechanical Theory of Irreversible Processes I*, J. Phys. Soc. Japan, **12**, 570, 1957, also *The Fluctuation-Dissipation Theorem*, Rep. Prog. Phys, **29**, 255, 1966.

[12] H. Mori, *Transport, Collective Motion and Brownian Motion*, Prog. Theor. Phys., **33**, 423, 1965.

[13] E. A. Morozova and N. N. Chentsov, *Markov Invariant Geometry on State Manifolds* (Russian), Ito. Nauki i Tek., **36**, 69-102, 1990.

[14] M. K. Murray and J. W. Rice, **Differential Geometry and Statistics**, *Monographs on Statistics and Applied Probability*, **48**, Chapman & Hall, 1993.

[15] H. Nagaoka, *Differential Geometric Aspects of Quantum State Estimation and Relative Entropy*, in **Quantum Communication and Measurement**, eds. V. P. Balavkin, O. Hirota and R. L. Hudson, Plenum Press, 1995.

[16] D. Petz, *Quasi-entropies for Finite Quantum Systems*, Rep. Math. Phys., **23**, 57-65, 1986.

[17] D. Petz, *Geometry of Canonical Correlation on the State Space of a Quantum System*, J. Math. Phys., **35**, 780-795, 1994.

[18] D. Petz, *Monotone Metrics on Matrix Spaces*, Lin. Alg. Appl., **244**, 81-96, 1996.

[19] D. Petz and H. Hasegawa, *On the Riemannian metric of α-entropies of density matrices*, Lett. Math. Phys., **38**, 221-241, 1996.

[20] D. Petz and C.Sudar, *Geometries of Quantum States*, J. Math. Phys., **37**, 2662-2673, 1996.

[21] R. F. Streater, *The Information Manifold for Relatively Bounded Potentials*, Proc. Stecklov Institute Maths. **228**, 205-223, 2000, math-ph/9910035.