On quantum Fisher information

Petz, D. and Jenčová, A

December 2, 2021

Abstract

A parametrized family of classical probability distributions can naturaly be viewed as a differentiable manifold. The Riemannian structure on this manifold is best introduced by means of the Fisher information matrix. In the case of a (finite) quantum system, there seems to be no unique generalization of the metric given by Fisher information. In this paper, we investigate the family of monotone Riemannian metrics for other important properties of Fisher information, namely the generalizations of the Cramér-Rao inequality and additivity.

Keywords: Fisher information, monotone metric, generalized covariance, Cramér-Rao inequality, α-family, α-representation, duality

Introduction

The differential-geometrical aspects of a smooth family $\mathcal{S} = \{p(x,\theta) \mid \theta \in \Theta \subseteq R^m\}$ of classical probability distributions on a sample space $\mathcal{X}$ were studied by many authors (Chentsov [3], Amari [1] and others). The Riemannian structure on the manifold $\mathcal{S}$, first introduced by Rao in 1945, is given by the Fisher information metric tensor

$$g_{ij}(\theta) = E_\theta[\frac{\partial}{\partial\theta_i}\log p(x,\theta)\frac{\partial}{\partial\theta_j}\log p(x,\theta)]$$

From a statistical point of view, such metric has some important properties. First, it is *invariant* with respect to transformations on the sample space (as well as the parameter space). Moreover, it is the only metric induced by a large class of invariant divergence measures for two infinitesimaly distant points. The statistical meaning of the Fisher information is in its relation to the variance of unbiased estimators of θ, given by the well-known Cramér-Rao inequality

$$V_\theta(A) \geq J(\theta)^{-1}$$

for every A such that $E_\theta[A] = \theta$, $\forall \theta \in \Theta$. Here $V_\theta(A)_{ij} = E_\theta[(A_i - \theta_i)(A_j - \theta_j)]$, $J(\theta)_{ij} = g_{ij}(\theta)$ and the inequality is in the sense of the order on positive definite matrices. Another important property is additivity, i.e. the Fisher information given by two independent random variables is the sum of the corresponding Fisher informations.

For a classical system with n alternatives, it was proved by Chentsov [3] that the Fisher information metric is the only invariant metric up to a constant factor. The quantum analogy of such metric seems not to be unique ([4]). The aim of the present paper is to investigate the family of monotone metrics and find the corresponding generalizations of the Cramér-Rao inequality in a neighborhood of a given density matrix D. Moreover, additivity (under tensor product) is proved for all such metrics. The last part of the paper deals with the problem of choice of a proper monotone Riemannian metric for some parametrized families, which correspond to the α-families of Amari [1], extended by Hasegawa [6] to the noncommutative framework.

1 The quantum state space and monotone metrics

The state space of an n-level quantum system may be identified with the manifold $\mathcal{D}_n = \{D \in \mathcal{M}_n(C)^+ \ : \ \mathrm{Tr}\, D = 1\}$, where $\mathcal{M}_n(C)^+$ is the space of all nonnegative definite n-dimensional complex matrices. The tangent space at $D \in \mathcal{D}_n$ is the (n^2-1)-dimensional real vector space of all self-adjoint traceless matrices $T_D = \{H \in \mathcal{M}_n(C)^{sa}, \mathrm{Tr}\, H = 0\}$. An inner product in T_D is given by means of a positive self-adjoint superoperator J_D on $\mathcal{M}_n(C)$ as $\lambda_D(H,K) = \mathrm{Tr}\, H J_D(K)$.

An important condition on such metric is *monotonicity*: if T is a stochastic (i.e. trace-preserving, completely positive) map on $\mathcal{D}_n$, then

$$\lambda_{T(D)}(T(K), T(K)) \leq \lambda_D(K,K) \quad \forall D \in \mathcal{D}_n, \quad \forall K \in T_D$$

It means that a coarse graining T of a state does not increase the metric ([12]). Clearly, monotonicity implies invariancy with respect to bijective stochastic maps, which corresponds to the invariancy property of the classical Fisher information metric. Unlike the classical case, a monotone Riemannian metric on $\mathcal{D}_n$ is far from being unique. Chentsov and Morozova ([4]) found that for a monotone Riemannian metric, there exists a symmetric function $c(x,y)$, satisfying $c(\lambda x, \lambda y) = \lambda^{-1} c(x,y)$, and a constant C such that the squared length of a tangent vector A at a diagonal point $D = \mathrm{diag}(d_1, \ldots, d_n)$ is equal to

$$\|A\|^2 = C \sum_{k=1}^{n} d_k^{-1} A_{kk}^2 + 2 \sum_{j<k} c(d_j, d_k) |A_{jk}|^2 \tag{1}$$

Later Petz [12] gave a characterization of monotone metrics in terms of operator monotone functions. Recall that a function $f : \mathbb{R}^+ \to \mathbb{R}$ is operator monotone if for any matrices K and H, $0 \leq K \leq H$ implies $0 \leq f(K) \leq f(H)$ (see [2] for more details). The above (1) gives a monotone metric if and only if $1/c(x,1)$ is an operator monotone function. For an alternative formulation we set $L_D(A) = DA$ and $R_D(A) = AD$, $A \in \mathcal{M}_n(C)$.

Theorem 1.1 *[12] Let $f : \mathbb{R}^+ \to \mathbb{R}$ be an operator monotone function such that $f(t) = t f(t^{-1})$ for every $t > 0$ and set a superoperator on matrices*

$$K_D = R_D^{\frac{1}{2}} f(L_D R_D^{-1}) R_D^{\frac{1}{2}}$$

Then the relation $\lambda_D(A,B) = \mathrm{Tr}\, K_D^{-1}(A)B$ determines a monotone Riemannian metric on $\mathcal{M}_n(C)$. Conversely, every monotone Riemannian metric is obtained in this way.

Let us now return to the geometric structure of $\mathcal{D}_n$. The cotangent space T_D^* of $\mathcal{D}_n$ at D is, by definition, the (n^2-1)-dimensional vector space of linear functionals on T_D. Using the representation theorem in the Hilbert space $\mathcal{M}_n(C)^{sa}$ with $\langle A, B \rangle = \mathrm{Tr}\, AB$, by $\omega(H) = \mathrm{Tr}\, AH = \mathrm{Tr}\,(A - \mathrm{Tr}\, DA)H$ the space T_D^* can be identified with the space of all observables with zero mean at D. Let the inner product $\lambda_D(H,K) = \mathrm{Tr}\, H J_D(K)$ in T_D satisfy the monotonicity condition, then from Theorem 1.1,

$$T_D^* = \{A \in \mathcal{M}_n(C)^{sa} \ : \ \mathrm{Tr}\, DA = 0\} = \{J_D(H) \ : \ H \in T_D\}$$

and the corresponding inner product in T_D^* is defined by $\varphi_D(A,B) = \mathrm{Tr}\, A J_D^{-1}(B)$. In this case, we will say that φ_D is a *generalized covariance*. It is also easy to see that $\varphi_D(A,B) = \mathrm{Tr}\, DAB$ if A and B commute with D.

Let $H_1, \dots, H_N$ form a basis of T_D. Then the dual (biorthogonal) basis of T_D^* is a basis $G_1, \dots, G_N$ satisfying $\operatorname{Tr} H_i G_j = \delta_{ij}$. It is easy to see that it is uniquely given by

$$G_i = \sum_j \lambda_D^{ij} J_D(H_j)$$

where $\lambda_{Dij} = \lambda_D(H_i, H_j)$ and $(\lambda_D^{ij}) = (\lambda_{Dij})^{-1}$. The coordinates of the metric tensor in T_D^* with respect to this basis are then $\varphi_D(G_i, G_j) = \lambda_D^{ij}$.

2 The Cramér-Rao inequalities

Let $D \in \mathcal{D}_n$ and let $\mathcal{D}' \subseteq \mathcal{D}_n$ be a smooth m- dimensional submanifold, parametrized by $\mathcal{D}' = \{D_t \ : \ t \in I\}$, such that $0 \in I$ and $D_0 = D$. In a neighborhood of D, we will consider an affine parametrization $D_t = D + \sum_i t_i H_i$, where $H_i \in T_D$ and t_i is in a small interval containing 0 for all $i = 1, \dots, m$. A (locally) unbiased estimator of t at $t = 0$ is a collection $A = (A_1, \dots, A_m)$ of self-adjoint matrices, such that

(i) $\operatorname{Tr} D_0 A_i = 0$, $\forall i$, i.e. $A \subset T_D^*$

(ii) $\frac{\partial}{\partial t_i} \operatorname{Tr} D_t A_j|_{t=0} = \operatorname{Tr} H_i A_j = \delta_{ij}$, for $i, j = 1, \dots, m$.

Let T_t' be the tangent space of $\mathcal{D}'$ at t, then a natural basis of T_0' is given by $H_1, \dots, H_m$. If λ_D is a monotone metric in T_D, then the coordinates of the metric tensor in T_0' with respect to this basis are $\lambda_{ij} = \lambda_D(H_i, H_j)$. Let φ_D be the corresponding generalized covariance. Then the generalized variance matrix of the estimator A is a positive definite matrix, defined by $\varphi_D(A)_{ij} = \varphi_D(A_i, A_j)$. The following theorem is the Cramer-Rao inequality for generalized variances.

Theorem 2.1 *Let λ_D and φ_D be as above and let $A = (A_1, \dots, A_m)$ be a locally (at $t = 0$) unbiased estimator of t. Then*

$$\varphi_D(A) \geq (\lambda_{ij})^{-1}$$

in the sense of the order on positive definite matrices. Moreover, equality is attained iff A is the biorthogonal basis of $T_0'^*$.

Proof. Let $G = (G_1, \dots, G_m)$ be a locally unbiased estimator, then any other locally unbiased estimator $A = (A_1, \dots, A_m)$ can be written in the form $A = (G_1 + s_1 X_1, \dots, G_m + s_m X_m)$, where $X_i \in \mathcal{S} = \{X \in \mathcal{M}_n(C)^{sa} : \mathrm{Tr}\, DX_i = 0, \mathrm{Tr}\, H_i X_j = 0, \forall i, j = 1, \dots, m\}$. Suppose that G minimizes $\varphi_D(A)$ over all localy unbiased estimators, then for each $v_1, \dots, v_m \in R$, $G_v = \sum_i v_i G_i$ is a locally unbiased estimator of the linear combination $\sum_i v_i t_i$ such that $\varphi_D(G_v) \leq \varphi_D(A_v)$, for each locally unbiased estimator A_v of $\sum_i v_i t_i$. Thus G is the minimizer iff $\forall k$, $\forall v_1, \dots, v_m \in R$ and for $X_i \in \mathcal{S}$, we have

$$0 = \frac{\partial}{\partial s_k} \mathrm{Tr} \sum_i v_i (G_i + s_i X_i) K_D(\sum_j (G_j + s_j X_j))|_{s_k=0} = 2 \sum_i v_i v_k \mathrm{Tr}\, X_k K_D(G_i)$$

which is equivalent to

$$\mathrm{Tr}\, K_D(G_i) X = 0, \quad \forall X \in \mathcal{S}$$

It follows that G is the solution of $K_D(G_i) = c^{i0} D + \sum_j c^{ij} H_j$. From $K_D(I) = D$ we obtain $\mathrm{Tr}\, K_D(G_i) = \mathrm{Tr}\,(DG_i) = 0$. Hence $c^{i0} = 0$ for each i, so that $G_i = J_D(\sum_j c^{ij} H_j)$. Further, $\delta_{ik} = \mathrm{Tr}\, G_i H_k = \sum_j c^{ij} \lambda_{jk}$. It follows that the minimizer is of the form

$$G_i = \sum_j \lambda^{ij} J_D(H_j)$$

where $(\lambda^{ij}) = (\lambda_{ij})^{-1}$.

Example 2.1 Let $\varphi_D(A, B) = \mathrm{Tr}\, DAB$. Then the corresponding monotone metric is determined by $J_D(H) = G$, where $GD + DG = 2H$, which is the definition of the metric of the symmetric logarithmic derivative, see [9], [12].

Example 2.2 Another important example of a generalized covariance is $\varphi_D(A, B) = \int_0^1 D^{1-s} A D^s B ds$, which is sometimes called the canonical correlation or Bogoljubov inner product. We use the expansion (c.f. [10])

$$\exp(\log(D + tH)) = D + t \int_0^1 D^{1-s} \frac{d}{dt} \log(D + tH)|_{t=0} D^s ds + \dots$$

to obtain $J_D(H) = \frac{d}{dt} \log(D + tH)|_{t=0}$. It follows that the monotone metric is given by

$$\lambda_D(H, K) = \mathrm{Tr}\,(H J_D(K)) = \frac{\partial^2}{\partial t \partial s} \mathrm{Tr}\,(D + tH) \log(D + sK)|_{t,s=0}$$

which is the well-known Kubo-Mori metric. Note that

$$J_D(H) = \int_0^\infty (D+s)^{-1} H (D+s)^{-1} ds$$

which is obtained from the integral representation [10]

$$\log B = \int_0^\infty (1+t)^{-1} - (t+B)^{-1} dt.$$

Example 2.3 The Kubo-Mori metric from previous example arises from relative entropy $D(\rho,\sigma) = \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$ when the density matrices ρ and σ are infinitesimaly distant from each other. In [8], a class of important examples of monotone metrics is obtained if α-divergences $D_\alpha(\rho,\sigma) = \frac{1-\alpha^2}{4}\mathrm{Tr}\,(\rho - \sigma^{\frac{1+\alpha}{2}}\rho^{\frac{1-\alpha}{2}})$, introduced by Hasegawa [5], are used, with $-3 < \alpha < 3$. We also refer to [13] for further examples and use of monotone metrics.

3 Additivity of quantum Fisher information

Let us consider an m-dimensional submanifold $\mathcal{D}' \subset \mathcal{D}_{n_1} \otimes \mathcal{D}_{n_2}$, such that $\mathcal{D}' = \{D_t = D_{1t} \otimes D_{2t} \ : \ t \in I\}$ Let $D_{1t} = D^1 + \sum_i t_i H_i^1$ and $D_{2t} = D^2 + \sum_i t_i H_i^2$, where $D^i \in \mathcal{D}_{n_i}$ and H_j^i is a traceless self-adjoint matrix for $j = 1,\dots,m$, $i = 1,2$. It is easy to see that $A = (A_1,\dots,A_m) \subset \mathcal{M}_{n_1 n_2}^{sa}$ is a locally unbiased estimator iff $\mathrm{Tr}\,(D^1 \otimes D^2)A = 0$ and $\mathrm{Tr}\,H_i A_j = \delta_{ij}$, where $H_i = (H_i^1 \otimes D^2 + D^1 \otimes H_i^2)$. The next theorem states the additivity property for all monotone metrics.

Theorem 3.1 *Let D^i, H_j^i, H_i, $i = 1,2$, $j = 1,\dots,m$ be as above and let the superoperator J_D determine a monotone metric. Then*

$$\mathrm{Tr}\,H_i J_{D^1 \otimes D^2}(H_j) = \mathrm{Tr} H_i^1 J_{D^1}(H_j^1) + \mathrm{Tr}\,H_i^2 J_{D^2}(H_j^2)$$

for each $i = 1,2$, $j = 1,\dots,m$.

Proof. Let $K_D = J_D^{-1}$. Suppose that $G^1 = (G_1^1,\dots,G_m^1)$ and $G^2 = (G_1^2,\dots,G_m^2)$ are such that $K_{D^i}(G_j^i) = H_j^i$ for $i = 1,2$, $j = 1,\dots,m$. We will prove that for $j = 1,\dots,m$, $K_{D^1 \otimes D^2}(G_j) = H_j$, where $G_j = G_j^1 \otimes I + I \otimes G_j^2$; from this, it follows by elementary computation that

$$\mathrm{Tr}\,H_i J_{D^1 \otimes D^2}(H_j) = \mathrm{Tr}\,H_i G_j = \mathrm{Tr}\,H_i^1 J_{D^1}(H_j^1) + \mathrm{Tr}\,H_i^2 J_{D^2}(H_j^2)$$

It is clear that we may assume that D^1 and D^2 are diagonal matrices, $D^1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_1})$, $D^2 = \mathrm{diag}(\mu_1, \ldots, \mu_{n_2})$. According to Theorem 1.1, there is an operator-monotone function f such that the superoperator K_D is of the form

$$K_D(X)_{ij} = f\Big(\frac{d_i}{d_j}\Big) d_j X_{ij} \tag{2}$$

for diagonal $D = \mathrm{diag}(d_1, \ldots, d_n)$ (of any order n). Put $D = D^1 \otimes D^2$. From (2) we see that for $l = 1, \ldots, m$,

$$K_D(G_l^1 \otimes I)_{\alpha\beta} = \begin{cases} f(\frac{\lambda_j}{\lambda_k})\lambda_k \mu_i (G_l^1)_{jk} & \text{if } \alpha = (j-1)n_2 + i, \beta = (k-1)n_2 + i, \\ & \quad i = 1, \ldots, n_2,\ j, k = 1, \ldots, n_1 \\ 0 & \quad \text{otherwise} \end{cases}$$

But $f(\frac{\lambda_j}{\lambda_k})\lambda_k (G_l^1)_{jk} = K_{D_1}(G_l^1)_{jk} = (H_l^1)_{jk}$, and we see that $K_D(G_l^1 \otimes I) = H_l^1 \otimes D^2$. Similarly, we obtain $K_D(I \otimes G_l^2) = D^1 \otimes H_l^2$.

4 α-families and α-representations

In the previous paragraphs, we considered affine parametrizations of $\mathcal{D}'$ in a neighborhood of the point $D = D_0$ and the basis vector ∂_i of the tangent space T_t was identified with $H_i = \frac{\partial}{\partial t_i} D_t$, $i = 1, \ldots, m$. In the classical case, Amari ([1]) introduced α-families of density functions and α-representations of the tangent space, where ∂_i is represented by means of a one-parameter family of differentiable functions as $\frac{\partial}{\partial t_i} F_\alpha(p(x,t))$. His results were, to certain extend, generalized to the noncommutative case by Hasegawa ([6]). We will use these ideas to show that for some families of density matrices there is a "natural" choice of monotone metric.

Let $\mathcal{M}$ denote the differentiable manifold $\mathcal{M}_n(C)^{sa}$. Let $T_D(\mathcal{M})$ be its tangent space at D and let the inner product be given by $\langle A, B\rangle = \mathrm{Tr}\, AB$. It is sometimes convenient to use the orthogonal decomposition of the tangent space introduced by Hasegava in [6], $T_D(\mathcal{M}) = C(D) \oplus C(D)^\perp$, where $C(D) = \{X \in \mathcal{M}_n(C)^{sa} \ :\ XD = DX\}$ and $C(D)^\perp = \{i[D, X] \ :\ X \in \mathcal{M}_n(C)^{sa}\}$.

Let now $g : \mathbb{R} \to \mathbb{R}$ be a smooth function and let the mapping $L_g(D) : T_D(\mathcal{M}) \to T_{g(D)}(\mathcal{M})$ be defined by $L_g(D)(H) = \frac{d}{ds} g(D + sH)|_{s=0}$.

Lemma 4.1 *(i) $L_g(D)$ is a linear map.*

(ii) $L_{f\circ g}(D) = L_f(g(D))L_g(D)$. *In particular, if g is invertible then $L_g(D)$ is invertible and $L_g(D)^{-1} = L_{g^{-1}}(g(D))$.*

(iii) $L_g(D)$ *is self-adjoint.*

Proof. Let $H \in T_D(\mathcal{M})$ be decomposed as $H = H^c + i[D, X]$, then, according to [2] (p. 124), $\frac{d}{ds}g(D + sH)|_{s=0} = g'(D)H^c + i[g(D), X]$. The statements (i) and (ii) follow easily from this equality. Let $K \in T_D(\mathcal{M})$, $K = K^c + i[D, Y]$, then

$$\begin{aligned}\mathrm{Tr}\, KL_g(D)(H) &= \mathrm{Tr}\, K^c g'(D)H^c - \mathrm{Tr}\,[D, Y][g(D), X] = \\ &= \mathrm{Tr}\, g'(D)K^c H^c - \mathrm{Tr}\,[g(D), Y][D, X] = \mathrm{Tr}\, L_g(D)(K)H\end{aligned}$$

which proves (iii).

Let $\tilde{\mathcal{D}} = \{D_t, t \in I\}$ be a smooth submanifold in $\mathcal{D}_n$, such that $D_0 = D$, I being a small neighborhood around 0. We will say that $\tilde{\mathcal{D}}$ is a g-family if

$$D_t = \frac{1}{\mathrm{Tr}\, g^{-1}(g(D) + \sum_i t_i G_i)} g^{-1}(g(D) + \sum_i t_i G_i) \tag{3}$$

where G_i is such that $\frac{\partial}{\partial t_i}\mathrm{Tr}\, g^{-1}(g(D) + \sum_j t_j G_j) = \mathrm{Tr}\,(g'(D))^{-1}G_i = 0$ for each i. In this case, it seems natural to represent the basis vector ∂_i of $\tilde{T}_0$ by G_i. Let $H_i = \frac{\partial}{\partial t_i}D_t|_{t=0}$, then it is easy to see that $G_i = L_g(D)(H_i)$. The vector space $\tilde{T}_0^g = \{L_g(D)(H) \ : \ H \in \tilde{T}_0\}$ will be called the g-representation of the tangent space $\tilde{T}_0$.

Let the inner product in $\tilde{T}_0$ be given by $\lambda(H_i, H_j) = \mathrm{Tr}\, H_i J(H_j)$, then the corresponding inner product in $\tilde{T}_0^g$ is

$$\lambda^g(G_i, G_j) = \mathrm{Tr}\, L_g^{-1}(G_i) J L_g^{-1}(G_j) = \mathrm{Tr}\, G_i L_{g^{-1}} J L_{g^{-1}}(G_j) = \mathrm{Tr}\, G_i J_g(G_j)$$

The g-representation of the cotangent space $\tilde{T}_0^*$ can now be defined by

$$\tilde{T}_0^{*g} = \{J_g(G) \ : \ G \in \tilde{T}_0^g\} = \{L_{g^{-1}}J(H) \ : \ H \in \tilde{T}_0\}$$

Let us further suppose that there is another such function g^*, such that $\tilde{T}_0^{*g}$ corresponds to a g^*-representation of $\tilde{T}_0$, in this case we will say that the functions g and g^* are dual. It is easy to see that in this case $J = L_{g^*}L_g$ and

$$\lambda(H, K) = \mathrm{Tr}\, L_{g^*}(H)L_g(K) = \frac{\partial^2}{\partial t \partial s}\mathrm{Tr}\, g^*(D + sH)g(D + tK)|_{t,s=0}$$

As it was proved in [8], the only possible choice of the functions g and g^*, satisfying the above conditions, is $g_\alpha(x) = \frac{2}{1-\alpha} x^{\frac{1-\alpha}{2}}$ and $g^* = g_{-\alpha}$, $-3 \leq \alpha \leq 3$. This corresponds to Amari's α-families. We will use the expression α-family rather than g_α-family. In [8], it was also shown that the monotone metric, defined by the operator $L_{-\alpha} L_\alpha$ is given by Example 2.3.

We see that as the Fisher information metric in the classical case, the metric given by $L_{-\alpha} L_\alpha$ reflects the dualistic structure of α-families (which has not yet been investigated in details). From this point of view, it seems that this is the metric to be chosen in such families of density matrices. An important example of dual α-families is given by $g_\alpha(x) = x$ and $g_{-\alpha}(x) = \log x$, which is obtained by a limiting procedure for $\alpha = \pm 1$. These are the families $D_t = D + \sum_i t_i H_i$, $\operatorname{Tr} H_i = 0$ and $D_t = (\operatorname{Tr} \exp(\log(D) + \sum_i t_i G_i))^{-1} \exp(\log(D) + \sum_i t_i G_i)$, called the mixture and exponential family, respectively. These families were shown to be important in the classical case, regarding some optimality and asymptotical optimality properties of the estimators of parameters and were studied also in the quantum case in [7]. The corresponding dual metric is the Kubo-Mori metric, see Example 2.2.

References

[1] Amari, S. *Differential-geometrical methods in statistic*, Lecture Notes in Statistics 28(1985)

[2] Bhatia, R. *Matrix analysis*, Springer, New York, 1997

[3] Chentsov, N.N. *Statistical decision rules and optimal inferences*, Translation of Math. Monog. 53 (Amer. Math. Society, Providence 1982)

[4] Chentsov, N.N. and Morozova, E.A. Markov invariant geometry on state manifolds (in Russian), Itogi Nauki i Tekhniki 36 (1990), 69–102

[5] Hasegawa, H. α-divergence of the non-commutative information geometry, Rep. Math. Phys. 33 (1993) 87–93

[6] Hasegawa, H., Non-commutative extension of information geometry, in *Quantum Communication, Computing and Measurement*, eds. Hirota et al., Plenum Press, New York (1994)

[7] Hasegawa, H. Exponential and mixture families in quantum statistics: Dual structure and unbiased parameter estimation, Rep. Math. Phys. 39 (1997) 49–68

[8] Hasegawa, H. and Petz, D. Non-commutative extension of information geometry II, Quantum Communication, Computing and Measurement, Eds. Hirota et al., Plenum Press, New York, (1997)

[9] Holevo, A.S., *Probabilistic and statistical aspects of quantum theory*, North-Holland, Amsterdam, 1982

[10] Ohya, M. and Petz, D. *Quantum Entropy and Its Use*, Springer-Verlag, Heidelberg, 1993

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

[12] Petz, D. Information geometry of quantum states, in *Quantum Probability Communications*, vol. 10, ed. R.L. Hudson et al, World Scientific, 135–158

[13] Slater, P.B. Comment on "Volume of the set of separable states", quant-ph/9904014