Advances in Mathematics 380 (2021) 107595

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

The metric property of the quantum Jensen-Shannon divergence

Dániel Virosztek [1]

Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria

A R T I C L E I N F O

Article history:
Received 7 November 2019
Received in revised form 8 July 2020
Accepted 8 January 2021
Available online 21 January 2021
Communicated by Dan Voiculescu

To Róza Virosztek

MSC:
primary 46N50
secondary 47N50, 81Q10

Keywords:
Quantum Jensen-Shannon divergence
Metric property

A B S T R A C T

In this short note, we prove that the square root of the quantum Jensen-Shannon divergence is a true metric on the cone of positive matrices, and hence in particular on the quantum state space.

E-mail address: daniel.virosztek@ist.ac.at.
URL: http://pub.ist.ac.at/~dviroszt.

[1] D. Virosztek was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 846294, and partially supported by the Hungarian National Research, Development and Innovation Office (NKFIH) via grants no. K124152, and no. KH129601.

https://doi.org/10.1016/j.aim.2021.107595
0001-8708/

1. Introduction

1.1. Motivation, goals

Quantum Jensen-Shannon divergence (QJSD) is among the most popular dissimilarity measures in information theory as it has a wide range of applications in the theory of complex networks [6,7], pattern recognition [2], graph theory [16], and chemical physics [1], beyond its applications in quantum theory [5,14,15]. It is a symmetrized version of the quantum relative entropy, and hence several desirable properties of a divergence measure on quantum states (e.g., monotonicity under quantum maps) can be justified for the QJSD directly by the fact that they hold for the relative entropy. Although QJSD is a symmetric divergence measure, it is not a metric as it does not satisfy the triangle inequality. However, its square root is known to be a metric for commuting states and for pure states [9].

Based on numerical investigations, it was conjectured more than ten years ago that the square root of the QJSD is a metric on the quantum state space in any finite dimension [4,9]. However, this conjecture has not been proved yet, as mentioned, for example, in [5,7,12,14,15]. Now we prove more: we prove that the square root of the quantum Jensen-Shannon divergence is a true metric on the cone of positive matrices in any dimension.

Furthermore, we will interpret a result of Carlen, Lieb, and Seiringer which tells us that on the qubit state space, the metric induced by the QJSD is a Hilbert space metric, and this is the case for a wide family of quantum Jensen divergences indexed by operator convex functions, as well.

1.2. Notation, definition and basic properties of the QJSD

Throughout this note, $M_n^+(\mathbb{C})$ stands for the cone of all positive semidefinite complex $n \times n$ matrices, I denotes the identity matrix, and $\eta(x) = x \log x$ is the standard entropy function on $[0, \infty)$. Density matrices are positive semidefinite matrices with unit trace.

The most attractive quantum version of the classical Kullback-Leibler divergence is the Umegaki relative entropy, which is denoted by $S(.,.)$ and is defined for densities ρ and σ by

$$S(\rho, \sigma) = \operatorname{Tr} \rho (\log \rho - \log \sigma).$$

The Umegaki relative entropy is jointly convex [10,11] and hence monotone under quantum maps, i.e., completely positive and trace preserving transformations of the state space [11,20].

The quantum Jensen-Shannon divergence (QJSD) is denoted by $J(.,.)$ and is defined by

$$J(\rho, \sigma) = \frac{1}{2} S\left(\rho, \frac{1}{2}(\rho + \sigma)\right) + \frac{1}{2} S\left(\sigma, \frac{1}{2}(\rho + \sigma)\right).$$

Note that the QJSD can be written in the form

$$J(\rho,\sigma)=\frac{1}{2}\operatorname{Tr}\eta(\rho)+\frac{1}{2}\operatorname{Tr}\eta(\sigma)-\operatorname{Tr}\eta\left(\frac{\rho+\sigma}{2}\right)\qquad\left(\rho,\sigma\in M_n^+(\mathbb{C})\right),$$

and we will prefer this latter form in the sequel. The QJSD is jointly convex, monotone under quantum maps, and the symmetry transformations of the state space and the positive cone with respect to the QJSD has been described in [12,13,21]. It was conjectured more than ten years ago that the square root of the QJSD is a metric on the quantum state space [4,9], but this has not been proved yet. In the next section, we give a proof of this metric property.

2. Main result

Theorem 1. *The square root of the quantum Jensen-Shannon divergence given by*

$$J(A,B)=\frac{1}{2}\operatorname{Tr}\eta(A)+\frac{1}{2}\operatorname{Tr}\eta(B)-\operatorname{Tr}\eta\left(\frac{A+B}{2}\right)\qquad\left(A,B\in M_n^+(\mathbb{C})\right)$$

is a true metric.

Proof. Symmetry, positivity, and definiteness is clear (the latter by the strict convexity of η), we have to show the triangle inequality.

Step 1. The quantum Jensen-Shannon divergence admits the integral representation

$$J(A,B)=\int_0^\infty d_S^2(A+tI,B+tI)\mathrm{d}t\qquad\left(A,B\in M_n^+(\mathbb{C})\right)\tag{1}$$

where d_S is the square root of the S-divergence defined by

$$d_S^2(A,B)=-\frac{1}{2}\operatorname{Tr}\log A-\frac{1}{2}\operatorname{Tr}\log B+\operatorname{Tr}\log\left(\frac{A+B}{2}\right)$$

for positive definite matrices [19]. On one hand,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}J\left(A+tI,B+tI\right)&=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\operatorname{Tr}\eta(A+tI)+\frac{1}{2}\operatorname{Tr}\eta(B+tI)-\operatorname{Tr}\eta\left(\frac{A+B}{2}+tI\right)\right)\\&=\frac{1}{2}\operatorname{Tr}\left(\log(A+tI)+I\right)+\frac{1}{2}\operatorname{Tr}\left(\log(B+tI)+I\right)-\operatorname{Tr}\left(\log\left(\frac{A+B}{2}+tI\right)+I\right)\\&=-d_S^2\left(A+tI,B+tI\right).\end{aligned}$$

On the other hand,

$$\lim_{t\to\infty} J\left(A+tI, B+tI\right) = 0$$

for every $A, B \in M_n^+(\mathbb{C})$, because the Jensen-Shannon divergence is homogeneous, $\eta(1+r) = r + \mathcal{O}\left(r^2\right)$, and hence

$$J\left(A+tI, B+tI\right) = tJ\left(I+\frac{A}{t}, I+\frac{B}{t}\right)$$

$$= t\left(\frac{1}{2}\operatorname{Tr}\eta\left(I+\frac{A}{t}\right)+\frac{1}{2}\operatorname{Tr}\eta\left(I+\frac{B}{t}\right)-\operatorname{Tr}\eta\left(I+\frac{A+B}{2t}\right)\right)$$

$$= t\left(\frac{1}{2}\operatorname{Tr}\frac{A}{t}+\mathcal{O}\left(t^{-2}\right)+\frac{1}{2}\operatorname{Tr}\frac{B}{t}+\mathcal{O}\left(t^{-2}\right)-\operatorname{Tr}\frac{A+B}{2t}+\mathcal{O}\left(t^{-2}\right)\right)=\mathcal{O}\left(t^{-1}\right).$$

So

$$J(A,B) = \lim_{t\to\infty}\left(J(A,B) - J(A+tI,B+tI)\right) = -\lim_{t\to\infty}\int_0^t \frac{\mathrm{d}}{\mathrm{d}r} J(A+rI,B+rI)\mathrm{d}r$$

$$= \int_0^\infty d_S^2(A+rI,B+rI)\mathrm{d}r$$

as claimed.

Step 2. The main result of [19] is that d_S is a metric on the positive definite cone. Therefore, for every $A, B, C \in M_n^+(\mathbb{C})$ we have

$$J(A,C) = \int_0^\infty d_S^2(A+tI,C+tI)\mathrm{d}t$$

$$\leq \int_0^\infty d_S(A+tI,C+tI)\left(d_S(A+tI,B+tI)+d_S(B+tI,C+tI)\right)\mathrm{d}t$$

$$\leq \sqrt{\int_0^\infty d_S^2(A+tI,C+tI)\mathrm{d}t}\left(\sqrt{\int_0^\infty d_S^2(A+tI,B+tI)\mathrm{d}t}+\sqrt{\int_0^\infty d_S^2(B+tI,C+tI)\mathrm{d}t}\right)$$

$$= \sqrt{J(A,C)}\left(\sqrt{J(A,B)}+\sqrt{J(B,C)}\right)$$

where the first inequality relies on Sra's result, and the second one follows from the Cauchy-Schwartz and the Minkowski inequalities. Dividing by $\sqrt{J(A,C)}$ completes the proof. $\square$

3. The 2 × 2 case

In this section we discuss the case of 2×2 matrices, because in this special case, one can prove more than the metric property of the square root of the QJSD in the following sense: on one hand, it can be shown that this metric is a Hilbert space metric, on the other hand, this is true not only for the QJSD, but for a large family of quantum Jensen divergences indexed by operator convex functions.

In the sequel, I interpret the argument of Eric Carlen, Elliott Lieb, and Robert Seiringer that I learned from Robert Seiringer in 2017. The symbol $\mathcal{S}\left(\mathbb{C}^2\right)$ stands for the space of 2×2 density matrices.

Theorem 2 *(Carlen-Lieb-Seiringer). For an operator convex function $f : [0, \infty) \to \mathbb{R}$, the square root of the symmetric quantum Jensen f-divergence defined by*

$$J_f(\rho, \sigma) := \frac{1}{2}(\operatorname{Tr} f(\rho) + \operatorname{Tr} f(\sigma)) - \operatorname{Tr} f\left(\frac{\rho+\sigma}{2}\right) \qquad \left(\rho, \sigma \in \mathcal{S}\left(\mathbb{C}^2\right)\right)$$

is a true metric on $\mathcal{S}\left(\mathbb{C}^2\right)$, moreover, it admits a Hilbert space embedding.

Proof. Operator convex functions on $[0, \infty)$ admit the integral representation

$$\begin{aligned} f(x) &= a + bx + cx^2 + \int_0^\infty \left(\frac{x}{1+t} - \frac{x}{x+t}\right) \mathrm{d}\mu(t) \\ &= a + bx + cx^2 + \int_0^\infty \left(\frac{x}{1+t} - 1 + \frac{t}{x+t}\right) \mathrm{d}\mu(t), \end{aligned}$$

where $a, b \in \mathbb{R}, c \geq 0$, and μ is a positive measure on $(0, \infty)$ with $\int_{(0,\infty)} (1+t)^{-2} \mathrm{d}\mu(t) < +\infty$ (see, e.g., [8, 8.1]).

Jensen divergences are linear in the sense that $J_{\alpha f+g}(\cdot, \cdot) = \alpha J_f(\cdot, \cdot) + J_g(\cdot, \cdot)$. Affine functions do not matter, $J_a(\cdot, \cdot) \equiv 0$ for $a(x) = \alpha x + \beta$. Note that the quadratic function $q(x) = x^2$ gives the Hilbert-Schmidt norm square, $J_q(\rho, \sigma) = \frac{1}{4} ||\rho - \sigma||_{HS}^2$.

By Schoenberg's theorem on the Hilbert space representation of negative definite kernels (see the original works [17,18] or [3, Chapter 3, 3.2.]), it suffices to prove that $J_f(\cdot, \cdot)$ is negative definite on $\mathcal{S}\left(\mathbb{C}^2\right)$, that is,

$$\sum_{j,k=1}^{m} c_j c_k J_f(\rho_j, \rho_k) \leq 0 \tag{2}$$

for all $c_1, \ldots, c_m \in \mathbb{R}$ with $\sum_{j=1}^m c_j = 0$ and for all $\rho_1, \ldots, \rho_m \in \mathcal{S}\left(\mathbb{C}^2\right)$. This clearly holds for $J_q(\cdot, \cdot)$, so we need to prove it only for $J_{f_t}(\cdot, \cdot)$, where $f_t(x) = \frac{t}{t+x}$ and $t > 0$.

Let $\mathbf{B}$ denote the closed unit ball in $\mathbb{R}^3$, and let $r_j \in \mathbf{B}$ denote the Bloch vector of ρ_j, that is, $\rho_j = \frac{1}{2}\left(I + r_j \cdot \sigma\right)$, where $\sigma = \{\sigma_x, \sigma_y, \sigma_z\}$ is the vector containing the Pauli matrices. Now

$$\sum_{j,k=1}^{m} c_j c_k J_{f_t}\left(\rho_j, \rho_k\right) = -\sum_{j,k=1}^{m} c_j c_k \operatorname{Tr} \frac{t}{t + \frac{\rho_j+\rho_k}{2}}$$

$$= -\sum_{j,k=1}^{m} c_j c_k \left(\frac{t}{t + \frac{1}{2}\left(1 + \left|\left|\frac{r_j+r_k}{2}\right|\right|\right)} + \frac{t}{t + \frac{1}{2}\left(1 - \left|\left|\frac{r_j+r_k}{2}\right|\right|\right)} \right)$$

$$= -\sum_{j,k=1}^{m} c_j c_k \left(\frac{2t^2 + t}{\left(t + \frac{1}{2}\right)^2 - \left|\left|\frac{r_j+r_k}{4}\right|\right|^2} \right) = -\sum_{j,k=1}^{m} c_j c_k \left(\frac{2t^2+t}{\left(t+\frac{1}{2}\right)^2} \right) \left(\frac{1}{1 - \left|\left|\frac{r_j+r_k}{4t+2}\right|\right|^2} \right)$$

where $||.||$ is the Euclidean norm in $\mathbb{R}^3$. Moreover,

$$\frac{1}{1 - \left|\left|\frac{r_j+r_k}{4t+2}\right|\right|^2} = \int_0^\infty \exp\left(-s\left(1 - \left|\left|\frac{r_j+r_k}{4t+2}\right|\right|^2\right)\right) \mathrm{d}s$$

$$= \int_0^\infty e^{-s} \exp\left(s\left|\left|\frac{r_j}{4t+2}\right|\right|^2\right) \exp\left(s\left|\left|\frac{r_k}{4t+2}\right|\right|^2\right) \exp\left(s\frac{2\left\langle r_j, r_k\right\rangle}{(4t+2)^2}\right) \mathrm{d}s.$$

Therefore,

$$\sum_{j,k=1}^{m} c_j c_k J_{f_t}\left(\rho_j, \rho_k\right) = -\frac{2t^2+t}{\left(t+\frac{1}{2}\right)^2} \int_0^\infty e^{-s} \sum_{j,k=1}^{m} c_j(s) c_k(s) \exp\left(\frac{2s\left\langle r_j, r_k\right\rangle}{(4t+2)^2}\right) \mathrm{d}s, \tag{3}$$

where $c_j(s) := c_j \exp\left(s\left|\left|\frac{r_j}{4t+2}\right|\right|^2\right)$.

The right hand side of (3) is non-positive for any $c_1, \dots, c_m \in \mathbb{R}$ and $r_1, \dots r_m \in \mathbf{B}$ as the product of positive definite functions is positive definite, and all the coefficients of the power series of the exponential function are positive. The proof is done. □

We note that Briët and Harremoës gave a proof of the Hilbert space embedding property for 2×2 densities and for a special class of operator convex functions containing the standard entropy function [4]. However, their proof contains a mistake, namely, in the proof of [4, Lemma 4], which is the key step in their argument, it is erroneously claimed that the function

$$\mathcal{S}\left(\mathbb{C}^2\right) \times \mathcal{S}\left(\mathbb{C}^2\right) \to \mathbb{R}; \qquad (\rho, \sigma) \mapsto 2 \operatorname{Tr}\left(\frac{1}{2}(\rho+\sigma)\right)^2 - 1$$

is positive definite. For example, if

$$\rho = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \sigma = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix},$$

then the matrix

$$\begin{bmatrix} 2\operatorname{Tr}\rho^2 - 1 & 2\operatorname{Tr}\left(\frac{1}{2}(\rho+\sigma)\right)^2 - 1 \\ 2\operatorname{Tr}\left(\frac{1}{2}(\rho+\sigma)\right)^2 - 1 & 2\operatorname{Tr}\sigma^2 - 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & 0 \end{bmatrix}$$

is indefinite.

We also note that it is known that the cone of 2×2 positive definite matrices equipped with the Jensen divergence corresponding to the operator convex function $x \mapsto -\log x$ does not admit any Hilbert space embedding [19]. This fact shows that the result of Carlen, Lieb, and Seiringer is optimal in the sense that one can not go beyond the qubit state space in proving the Hilbert space embedding property in its full generality, that is, for all operator convex functions. (To be precise, $x \mapsto -\log x$ is operator convex only on $(0, \infty)$ and not on $[0, \infty)$, but the integral representation $\log x = \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{x+t}\right) \mathrm{d}t$ shows that from the viewpoint of negative definite kernels, it behaves similarly to operator convex functions on $[0, \infty)$.) However, for the standard entropy function $\eta(x) = x \log x$, we do not have such counterexamples. So it is still an open question, whether the metric induced by the QJSD is a Hilbert space metric on the cone of positive semidefinite $n \times n$ matrices for $n \geq 2$, or on the space of $n \times n$ density matrices for $n \geq 3$.

Acknowledgments

I would like to thank Robert Seiringer for great discussions on the topic and for showing me their argument on the case of 2×2 matrices. I am also grateful to Lajos Molnár for proposing this problem to me around 2015, and to the anonymous referee for his/her valuable comments and suggestions.

References

[1] J. Antolín, J.C. Angulo, S. López-Rosa, Fisher and Jensen–Shannon divergences: quantitative comparisons among distributions. Application to position and momentum atomic densities, J. Chem. Phys. 130 (2009) 074110.
[2] L. Bai, L. Rossi, A. Torsello, E.R. Hancock, A quantum Jensen–Shannon graph kernel for unattributed graphs, Pattern Recognit. 48 (2015) 344–355.
[3] C. Berg, J.P.R. Christensen, P. Ressel, Harmonic Analysis on Semigroups, Springer-Verlag, New York, 1984.
[4] J. Briët, P. Harremoës, Properties of classical and quantum Jensen-Shannon divergence, Phys. Rev. A 79 (2009) 052311.
[5] J. Dajka, J. Łuczka, P. Hänggi, Distance between quantum states in the presence of initial qubit-environment correlations: a comparative study, Phys. Rev. A 84 (2011) 032120.

[6] M. De Domenico, M.A. Porter, A. Arenas, MuxViz: a tool for multilayer analysis and visualization of networks, J. Complex Netw. 3 (2015) 159–176.
[7] M. De Domenico, V. Nicosia, A. Arenas, V. Latora, Structural reducibility of multilayer networks, Nat. Commun. 6 (2015) 6864.
[8] F. Hiai, M. Mosonyi, D. Petz, C. Bény, Quantum f-divergences and error correction, Rev. Math. Phys. 23 (2011) 691–747.
[9] P.W. Lamberti, A.P. Majtey, A. Borras, M. Casas, A. Plastino, Metric character of the quantum Jensen-Shannon divergence, Phys. Rev. A 77 (2008) 052311.
[10] E.H. Lieb, Convex trace functions and the Wigner–Yanase–Dyson conjecture, Adv. Math. 11 (1973) 267–288.
[11] G. Lindblad, Expectations and entropy inequalities for finite quantum systems, Commun. Math. Phys. 39 (1974) 111–119.
[12] L. Molnár, G. Nagy, Transformations on density operators that leave the Holevo bound invariant, Int. J. Theor. Phys. 53 (2014) 3273–3278.
[13] L. Molnár, J. Pitrik, D. Virosztek, Maps on positive definite matrices preserving Bregman and Jensen divergences, Linear Algebra Appl. 495 (2016) 174–189.
[14] C. Radhakrishnan, M. Parthasarathy, S. Jambulingam, T. Byrnes, Distribution of quantum coherence in multipartite systems, Phys. Rev. Lett. 116 (2016) 150504.
[15] W. Roga, M. Fannes, K. Życzkowski, Universal bounds for the Holevo quantity, coherent information, and the Jensen-Shannon divergence, Phys. Rev. Lett. 105 (2010) 040505.
[16] L. Rossi, A. Torsello, E.R. Hancock, R.C. Wilson, Characterizing graph symmetries through quantum Jensen-Shannon divergence, Phys. Rev. E 88 (2013) 032806.
[17] I.J. Schoenberg, Metric spaces and completely monotone functions, Ann. Math. 39 (1938) 811–841.
[18] I.J. Schoenberg, Metric spaces and positive definite functions, Trans. Am. Math. Soc. 44 (1938) 522–536.
[19] S. Sra, Positive definite matrices and the S-divergence, Proc. Am. Math. Soc. 144 (2016) 2787–2797.
[20] A. Uhlmann, Endlich dimensionale dichtmatrizen, II, Wiss. Z., Karl-Marx-Univ. Leipzig, Math.-Nat.wiss. Reihe 22 (1973) 139.
[21] D. Virosztek, Maps on quantum states preserving Bregman and Jensen divergences, Lett. Math. Phys. 106 (2016) 1217–1234.