BANACH JOURNAL of MATHEMATICAL ANALYSIS

Banach J. Math. Anal. 11 (2017), no. 4, 744–763
http://dx.doi.org/10.1215/17358787-2017-0015
ISSN: 1735-8787 (electronic)
http://projecteuclid.org/bjma

MAPS PRESERVING A NEW VERSION OF QUANTUM f-DIVERGENCE

MARCELL GAÁL

Dedicated to Letícia Tari

Communicated by M. Frank

Abstract. For an arbitrary nonaffine operator convex function defined on the nonnegative real line and satisfying $f(0) = 0$, we characterize the bijective maps on the set of all positive definite operators preserving a new version of quantum f-divergence. We also determine the structure of all transformations leaving this quantity invariant on quantum states for any strictly convex functions with the properties $f(0) = 0$ and $\lim_{x\to\infty} f(x)/x = \infty$. Finally, we derive the corresponding result concerning those transformations on the set of positive semidefinite operators. We emphasize that all the results are obtained for finite-dimensional Hilbert spaces.

1. Introduction

Nonlinear preserver problems appear in many parts of mathematics, and the number of those which are related to quantum structures has recently been significantly increasing. These structures are, among others, the set of rank 1 projections, the set of density operators and the set of self-adjoint operators which are the mathematical representatives of the pure states, the quantum states, and the bounded observables, respectively. The transformations leaving a certain quantity or a relation between operators invariant can be viewed as symmetries of

Received Jul. 6, 2016; Accepted Oct. 23, 2016.
First published online Jun. 22, 2017.
2010 *Mathematics Subject Classification.* Primary, 47B49; Secondary 47N50.
Keywords. preservers, positive definite operators, density operators, quantum states, relative entropy.

the underlying quantum system. The most famous result is Wigner's fundamental theorem concerning the "traditional" symmetry transformations, which states that every transformation preserving the transition probability (the trace of the product) between pure states is necessarily implemented by either a unitary or an antiunitary similarity transformation. This result motivated the exploration of the structure of symmetry transformations with respect to some particular quantum relative entropies and quantum divergences, which quantities are used as measures of dissimilarity between quantum states. (Several proofs of Wigner's theorem under different assumptions can be found in [7], [8], and [13]; a general comprehensive survey article about this theorem is [3]; for recent Wigner-type results on density operators, see [6], [14], [17]–[19] and [24].)

In 1985, Petz introduced the general notion of quasientropy for the state space of a von Neumann algebra, and he considered the quantum f-divergence as a special case of the large class of quasientropies (for further details, see [22], [23] and the references therein). In 2014, a different concept of quantum f-divergence was introduced by Dragomir in [5] as

$$D_f(A \parallel B) = \operatorname{Tr} B f(B^{-1/2} A B^{-1/2}) \tag{1.1}$$

for any invertible density operators A and B. The properties of this quantity were studied by Matsumoto [12] in detail. In the rest of this article, we refer to this latter quantity as "the divergence $D_f(\cdot \parallel \cdot)$" to avoid any confusion with the traditional notion of quantum f-divergence. In [16], the structure of bijective transformations on the set of positive definite operators leaving the divergence $D_f(\cdot \parallel \cdot)$ invariant was determined in the case where f is a nonconstant operator monotone decreasing function on the positive real line satisfying $\lim_{x\to\infty} f(x)/x = 0$.

In this article we are interested in questions of similar kind. First, we point out the fact that in a special choice of the function f the divergence $D_f(\cdot \parallel \cdot)$ is a quantum f-divergence in the traditional sense, but this is not the case in general. Second, for an arbitrary nonaffine operator convex function defined on the nonnegative real line and satisfying $f(0) = 0$, we characterize all the bijective transformations on the cone of positive definite operators which preserve the divergence $D_f(\cdot \parallel \cdot)$. We remark that the set of positive definite operators is a rich mathematical object which may be equipped with differential manifold structure. Therefore, in some investigations in quantum theory (especially where differential geometric tools are applied) it is more natural to consider the positive definite operators instead of density operators. Next we describe all the symmetries (not necessarily bijective) on the set of density operators in the case where the function f is strictly convex and the conditions $f(0) = 0$ and $\lim_{x\to\infty} f(x)/x = \infty$ hold. Finally, under the same conditions on f, we characterize the corresponding symmetry transformations on the set of positive semidefinite operators. We obtain the interesting fact that in each case all those symmetries are exactly the unitary and the antiunitary similarity transformations on the underlying Hilbert space. We emphasize that all the statements are obtained for finite-dimensional Hilbert spaces. It means that, while the problems and the statements could be formulated in the context of matrix theory, we nevertheless prefer the operator theoretical point of view and hope one can extend the results to other settings, such as certain classes of linear operators acting on an infinite-dimensional Hilbert space.

Remark 1.1. This question naturally arises: where do the above conditions concerning the function f come from? If operator convexity is required, then according to [21] there are only a few types of numerical functions which are particular important in quantum information theory. These functions are the following:

(1) $x \mapsto -\log x$,
(2) $x \mapsto -x^p$ with some exponent $0 < p < 1$,
(3) $x \mapsto x^q$ with some exponent $-1 \leq q < 0$ or $1 < q \leq 2$,
(4) the so-called standard entropy function, i.e.,

$$\eta(x) = \begin{cases} x \log x, & x > 0, \\ 0, & x = 0, \end{cases}$$

(5) $x \mapsto x - x^r$ with some exponent $0 < r < 1$.

It is mentioned in [21, p. 113] that other numerical functions are physically irrelevant. It follows from the celebrated Löwner–Heinz theorem (see, e.g., [2, Theorem 2.6]) that the numerical functions $x \mapsto -\log x$, $x \mapsto -x^p$ (with $0 < p < 1$) and $x \mapsto x^q$ (with $-1 \leq q < 0$) are nonconstant operator monotone decreasing functions. Furthermore, the standard entropy function $\eta(x)$, the functions $x \mapsto x^q$ (with some $1 < q \leq 2$) and $x \mapsto x - x^r$ (with some $0 < r < 1$) are operator convex functions. Therefore, the cases (1), (2), and (3) under the assumption $-1 \leq q < 0$ are discussed in [16] on the cone of positive definite operators. Our first result covers the remaining cases. In these cases the corresponding results on the set of density operators and positive semidefinite operators are also presented here.

Remark 1.2. We note that if f is affine, then due to the condition $f(0) = 0$ there exists an $a \in \mathbb{R}$ such that $f(x) = ax$. Then we have

$$D_f(A \parallel B) = a \operatorname{Tr} A \quad \big(A, B \in \mathcal{L}(\mathcal{H})_+^{-1}\big),$$

and thus this quantity cannot be used as a measure of dissimilarity between quantum states because it depends only on the trace of the operator A. So, it is not necessary to deal with the case when f is an affine function.

1.1. **Notation.** Next we give a short description of some notation and basic concepts that we use throughout this paper. Assume that $\mathcal{H}$ is at least 2-dimensional. The set $\mathcal{L}(\mathcal{H})$ consists of all linear operators mapping $\mathcal{H}$ to itself, and I stands for the identity operator on $\mathcal{H}$. We recall that $\mathcal{L}(\mathcal{H})$ is a complex Hilbert space with the Hilbert–Schmidt (HS) inner product $\langle \cdot, \cdot \rangle_{\mathrm{HS}} \colon \mathcal{L}(\mathcal{H}) \times \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ defined by

$$\langle A, B \rangle_{\mathrm{HS}} = \operatorname{Tr} AB^* \quad \big(A, B \in \mathcal{L}(\mathcal{H})\big),$$

where Tr stands for the usual trace functional on $\mathcal{L}(\mathcal{H})$. The collections of positive semidefinite operators and positive definite operators are denoted by $\mathcal{L}(\mathcal{H})_+$ and $\mathcal{L}(\mathcal{H})_+^{-1}$, respectively. Finally, the set of density operators (i.e. positive semidefinite operators with unit trace) is denoted by $\mathcal{S}(\mathcal{H})$ and the symbol $\mathcal{P}_1(\mathcal{H})$ stands for the set of all rank 1 (orthogonal) projections.

1.2. **The quantum f-divergence.** Here we recall the traditional definition of quantum f-divergence. In order to do so, for any $A, B \in \mathcal{L}(\mathcal{H})$ we introduce the left and right multiplication operators, respectively, acting on $\mathcal{L}(\mathcal{H})$ by

$$L_A \colon X \mapsto AX, \qquad R_A \colon X \mapsto XA$$

and the relative modular operator $\Delta_{A,B}$ as

$$\Delta_{A,B} = L_A R_{B^{-1}}.$$

Following [10, 2.1 Definition], for $A \in \mathcal{L}(\mathcal{H})_+$ and $B \in \mathcal{L}(\mathcal{H})_+^{-1}$ the quantum f-divergence between the operators A and B is defined by

$$S_f(A \parallel B) = \left\langle B^{1/2}, f(\Delta_{A,B}) B^{1/2} \right\rangle_{\mathrm{HS}},$$

and as for the general case, we set

$$S_f(A \parallel B) = \lim_{\varepsilon \searrow 0} S_f(A \parallel B + \varepsilon I). \tag{1.2}$$

Let $A, B \in \mathcal{L}(\mathcal{H})_+$, and for any $\lambda \geq 0$ denote by P_λ and Q_λ the projections onto the kernels of $A - \lambda I$ and $B - \lambda I$, respectively. If f is any strictly convex function, then due to [10, Proposition 2.2] the limit in (1.2) exists on the extended real line and according to [10, Corollary 2.3] it can be computed as

$$S_f(A \parallel B) = \sum_{a \in \sigma(A)} \Bigg(\sum_{b \in \sigma(B) \setminus \{0\}} b f\left(\frac{a}{b}\right) \operatorname{Tr} P_a Q_b + \omega_f a \operatorname{Tr} P_a Q_0 \Bigg), \tag{1.3}$$

where $\omega_f = \lim_{x \to \infty} f(x)/x \in [-\infty, \infty]$ and where $\sigma(\cdot)$ stands for the spectrum of the operators. (For some examples and important results related to quantum f-divergence, the reader can consult [10], [18] and the references therein.)

1.3. **A new version of quantum f-divergence.** Referring to [5] in (1.1), we presented the definition of the divergence $D_f(\cdot \parallel \cdot)$ between invertible density operators. Obviously, we can define this quantity between positive definite operators as in (1.1). Furthermore, and similarly to (1.2), in the general case we would like to define it as

$$D_f\big(A \parallel B) = \lim_{\varepsilon \searrow 0} D_f(A \parallel B + \varepsilon I). \tag{1.4}$$

In the next proposition, we see this really can be done under certain conditions on f because the limit in (1.4) exists in that given case.

Proposition 1.3. *Assume that $f : [0, \infty[\to \mathbb{R}$ is a convex function satisfying $f(0) = 0$ and $\lim_{x\to\infty} f(x)/x = \infty$. Select $A, B \in \mathcal{L}(\mathcal{H})_+$, and let P_B denote the orthogonal projection on $\mathcal{H}$ onto the support of B. Then the limit in (1.4) exists and we have*

$$D_f(A \parallel B) = \operatorname{Tr} B|_{\operatorname{supp} B} f\big((B|_{\operatorname{supp} B})^{-1/2} P_B A P_B (B|_{\operatorname{supp} B})^{-1/2}\big) \tag{1.5}$$

if $\operatorname{supp} A \subseteq \operatorname{supp} B$, *and* $D_f(A \parallel B) = +\infty$ *otherwise.*

Next we give some important examples.

(1) If we consider the standard entropy function

$$\eta(x) = \begin{cases} x \log x, & x > 0, \\ 0, & x = 0, \end{cases}$$

then

$$D_\eta(A \parallel B) = \operatorname{Tr} B^{1/2} A B^{-1/2} \log(B^{-1/2} A B^{-1/2}). \tag{1.6}$$

(2) If $g\colon [0, \infty[\to \mathbb{R}, x \mapsto x^2 - 1$, then

$$D_g(A \parallel B) = \operatorname{Tr} A^2 B^{-1} - 1,$$

which is called the χ^2-*distance*.

(3) If we take the function $h\colon\ ]0, \infty[\to \mathbb{R},\ x \mapsto -\log x$, then

$$D_h(A \parallel B) = -\operatorname{Tr} B \log B^{-1/2} A B^{-1/2}$$

which is just the usual Belavkin–Staszewski relative entropy (see [21, p. 125]).

For further examples, see [5].

In [18], the authors determined the structure of the transformations on the set of density operators which preserve the quantum f-divergence. Obviously, if the quantity defined in (1.1) were quantum f-divergence for some strictly convex functions, then our problem would have no sense because [18, Theorem] could be applied. By elementary computation, we can verify that, for example, the χ^2-distance is a quantum f-divergence on $\mathcal{S}(\mathcal{H})$, but the content of the next proposition tells us that is not the case in general.

Proposition 1.4. *If we consider the function*

$$\eta(x) = \begin{cases} x \log x, & x > 0, \\ 0, & x = 0, \end{cases}$$

then the divergence $D_\eta(\cdot \parallel \cdot)$ is not a quantum f-divergence on $\mathcal{S}(\mathcal{H})$, meaning that there is no strictly convex function f such that

$$S_f(A \parallel B) = \operatorname{Tr} B^{1/2} A B^{-1/2} \log(B^{-1/2} A B^{-1/2}) = D_\eta(A \parallel B)$$

holds for every $A, B \in \mathcal{S}(\mathcal{H})$.

Remark 1.5. By the same argument as in Proposition 1.4, we can check that if we consider the function $f_p(x) = x^p$, then the divergence $D_{f_p}(\cdot \parallel \cdot)$ is not a quantum f-divergence when $p \neq 2$. It is an interesting open question which functions gives the divergence $D_f(\cdot \parallel \cdot)$ a quantum f-divergence in the traditional sense. Our examples show that, in general, the two concepts may lead to different quantities.

2. Main results

Now we are in a position to formulate the main results of this paper. It is easy to observe that if we take any unitary or antiunitary similarity transformation ϕ (i.e., when ϕ is of the form $\phi(\cdot) = U(\cdot)U^*$, where U is an arbitrary unitary or antiunitary operator) on the set of positive definite operators, then it preserves the divergence $D_f(\cdot \parallel \cdot)$; that is,

$$D_f\big(\phi(A) \,\big\|\, \phi(B)\big) = D_f(A \parallel B)$$

holds for every $A, B \in \mathcal{L}(\mathcal{H})_+^{-1}$. Our first theorem states that the converse is also true under certain conditions on f.

Theorem 2.1. *If $f\colon [0,\infty[\to \mathbb{R}$ is a nonaffine operator convex function satisfying $f(0) = 0$ and if $\phi\colon \mathcal{L}(\mathcal{H})_+^{-1} \to \mathcal{L}(\mathcal{H})_+^{-1}$ is a bijective transformation fulfilling*

$$D_f\big(\phi(A) \,\big\|\, \phi(B)\big) = D_f(A \parallel B) \quad \big(A, B \in \mathcal{L}(\mathcal{H})_+^{-1}\big),$$

then there is either a unitary or an antiunitary operator U on $\mathcal{H}$ such that ϕ is of the form

$$\phi(A) = UAU^* \quad \big(A \in \mathcal{L}(\mathcal{H})_+^{-1}\big).$$

Under different conditions on f the same conclusion can be obtained concerning the density operators, given below.

Theorem 2.2. *Let $f\colon [0,\infty[\to \mathbb{R}$ be a strictly convex function satisfying $f(0) = 0$ and let $\lim_{x\to\infty} f(x)/x = \infty$. If $\phi\colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is a transformation with the property that*

$$D_f\big(\phi(A) \,\big\|\, \phi(B)\big) = D_f(A \parallel B) \quad \big(A, B \in \mathcal{S}(\mathcal{H})\big)$$

holds, then there exists either a unitary or an antiunitary operator U on $\mathcal{H}$ such that ϕ is of the form

$$\phi(A) = UAU^* \quad \big(A \in \mathcal{S}(\mathcal{H})\big).$$

As for the positive semidefinite operators, we have the following structural result.

Theorem 2.3. *Let $f\colon [0,\infty[\to \mathbb{R}$ be a strictly convex function for which $f(0) = 0$ and $\lim_{x\to\infty} f(x)/x = \infty$ holds. If $\phi\colon \mathcal{L}(\mathcal{H})_+ \to \mathcal{L}(\mathcal{H})_+$ is a bijective map satisfying*

$$D_f\big(\phi(A) \,\big\|\, \phi(B)\big) = D_f(A \parallel B) \quad \big(A, B \in \mathcal{L}(\mathcal{H})_+^{-1}\big),$$

then there exists either a unitary or an antiunitary operator U on $\mathcal{H}$ such that ϕ is of the form

$$\phi(A) = UAU^* \quad \big(A \in \mathcal{L}(\mathcal{H})_+\big).$$

We emphasize that we do not require any sort of linearity or bijectivity from the transformation ϕ in Theorem 2.2, but in the present article we could not remove the bijectivity condition from the assumptions of Theorem 2.1 or Theorem 2.3.

3. Proofs

Before turning to the proofs, we recall the notion of strength along a rank 1 projection and collect some of its important properties. This concept was introduced by Bush and Gudder in [1].

Let $A \in \mathcal{L}(\mathcal{H})_+$ be a positive operator, and consider any rank 1 projection $P \in \mathcal{P}_1(\mathcal{H})$. Then the strength of A along P is defined by

$$\lambda(A,P) = \sup\{t \geq 0 \,:\, tP \leq A\},$$

and the function $P \mapsto \lambda(A,P)$ is called the *strength function* of A. Due to [1, Theorem 1], we have the following characterization of the usual Löwner order on the cone $\mathcal{L}(\mathcal{H})_+$:

$$A \leq B \quad \Longleftrightarrow \quad \lambda(A,P) \leq \lambda(B,P) \quad \big(P \in \mathcal{P}_1(\mathcal{H})\big). \tag{3.1}$$

Naturally, this implies that the operators $A, B \in \mathcal{L}(\mathcal{H})_+$ are equal if and only if they have the same strength function. According to [1, Theorem 4], this quantity can be computed as

$$\lambda(A,P) = \begin{cases} \frac{1}{\operatorname{Tr} P(A|_{\operatorname{supp} A})^{-1}}, & \operatorname{supp} P \subseteq \operatorname{supp} A, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

In the case where $\operatorname{supp} P \subseteq \operatorname{supp} A$ holds and the operator A is given by its spectral decomposition as $A = \sum_{a \in \sigma(A)} aQ_a$, (3.2) yields the following explicit and very useful formula:

$$\frac{1}{\lambda(A,P)} = \sum_{a \in \sigma(A)\setminus\{0\}} \frac{1}{a} \operatorname{Tr} PQ_a. \tag{3.3}$$

Now we are in a position to present the proof of our first result. In the proof, we apply the main ideas of [20, Lemma 13] and [6, Proposition 2].

Proof of Proposition 1.3. Select any $B \in \mathcal{L}(\mathcal{H})_+$. With respect to the orthogonal decomposition $\mathcal{H} = \operatorname{supp} B \oplus \ker B$, we can write

$$B = \begin{pmatrix} B_0 & 0 \\ 0 & 0 \end{pmatrix},$$

where $B_0 = B|_{\operatorname{supp} B}$. We select an arbitrary $A \in \mathcal{L}(\mathcal{H})_+$. With respect to the same orthogonal decomposition, we have

$$A = \begin{pmatrix} A_0 & C \\ C^* & A_1 \end{pmatrix},$$

where A_0, C and A_1 are appropriate operators. Direct computation shows that

$$\begin{aligned} &(B+\varepsilon I)^{-1/2} A (B+\varepsilon I)^{-1/2} \\ &\quad = \begin{pmatrix} (B_0+\varepsilon I)^{-1/2} A_0 (B_0+\varepsilon I)^{-1/2} & \frac{1}{\sqrt{\varepsilon}}(B_0+\varepsilon I)^{-1/2} C \\ \frac{1}{\sqrt{\varepsilon}} C^* (B_0+\varepsilon I)^{-1/2} & \frac{1}{\varepsilon} A_1 \end{pmatrix}, \end{aligned}$$

where I denotes the identity operator on an appropriate subspace of $\mathcal{H}$. Under the assumption $\operatorname{supp} A \subseteq \operatorname{supp} B$, we obtain $C = 0, A_1 = 0$ and thus

$$\begin{aligned}
&D_f(A \parallel B + \varepsilon I) \\
&\quad = \operatorname{Tr}(B + \varepsilon I) f\big((B + \varepsilon I)^{-1/2} A (B + \varepsilon I)^{-1/2}\big) \\
&\quad = \operatorname{Tr} B_0 f\big((B_0 + \varepsilon I)^{-1/2} A_0 (B_0 + \varepsilon I)^{-1/2}\big) \\
&\qquad + \varepsilon \operatorname{Tr} f\big((B_0 + \varepsilon I)^{-1/2} A_0 (B_0 + \varepsilon I)^{-1/2}\big).
\end{aligned}$$

In the above displayed formula, the second term tends to zero whenever $\varepsilon \searrow 0$ and this yields

$$D_f(A \parallel B) = \operatorname{Tr} B|_{\operatorname{supp} B} f\big((B|_{\operatorname{supp} B})^{-1/2} P_B A P_B (B|_{\operatorname{supp} B})^{-1/2}\big).$$

Assume now that $\operatorname{supp} A \nsubseteq \operatorname{supp} B$ holds. Then there exists a unit vector $v \in \mathcal{H}$ such that $v \in \ker B$ and $v \notin \ker A$. With respect to the decomposition $\operatorname{supp} B \oplus \ker B$ of $\mathcal{H}$, the vector v is of the form

$$v = \begin{pmatrix} 0 \\ z \end{pmatrix}$$

and thus

$$Av = \begin{pmatrix} A_0 & C \\ C^* & A_1 \end{pmatrix} \begin{pmatrix} 0 \\ z \end{pmatrix} = \begin{pmatrix} Cz \\ A_1 z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3.4}$$

holds. We claim that $A_1 z \neq 0$. Assume on the contrary that $A_1 z = 0$. Since $A \in \mathcal{L}(\mathcal{H})_+$, for arbitrary $w \in \operatorname{supp} B$ we have

$$0 \leq \left\langle \begin{pmatrix} A_0 & C \\ C^* & A_1 \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix}, \begin{pmatrix} w \\ z \end{pmatrix} \right\rangle = \langle A_0 w, w\rangle + 2\Re\langle Cz, w\rangle.$$

Hence for every $t \in \mathbb{R}$ and for an arbitrary $w \in \operatorname{supp} B$, we have

$$0 \leq t^2 \langle A_0 w, w\rangle + 2t\Re\langle Cz, w\rangle,$$

implying that, for every $w \in \operatorname{supp} B$, the equality

$$2\Re\langle Cz, w\rangle = 0$$

holds. From this we deduce that $Cz = 0$, which contradicts (3.4). Therefore, we have $A_1 z \neq 0$. Now we compute

$$\begin{aligned}
&\big\langle (B + \varepsilon I)^{-1/2} A (B + \varepsilon I)^{-1/2} v, v\big\rangle \\
&\quad = \left\langle \begin{pmatrix} (B_0 + \varepsilon I)^{-1/2} & 0 \\ 0 & \frac{1}{\sqrt{\varepsilon}} I \end{pmatrix} \begin{pmatrix} A_0 & C \\ C^* & A_1 \end{pmatrix} \begin{pmatrix} (B_0 + \varepsilon I)^{-1/2} & 0 \\ 0 & \frac{1}{\sqrt{\varepsilon}} I \end{pmatrix} \begin{pmatrix} 0 \\ z \end{pmatrix}, \begin{pmatrix} 0 \\ z \end{pmatrix} \right\rangle \\
&\quad = \frac{1}{\varepsilon} \langle A_1 z, z\rangle.
\end{aligned}$$

Then an application of Peierls inequality (see [2, Theorem 2.9]) gives us

$$\begin{aligned}
\operatorname{Tr} f\big((B + \varepsilon I)^{-1/2} A (B + \varepsilon I)^{-1/2}\big) &\geq f\big(\big\langle (B + \varepsilon I)^{-1/2} A (B + \varepsilon I)^{-1/2} v, v\big\rangle\big) \\
&= f\Big(\frac{1}{\varepsilon} \langle A_1 z, z\rangle\Big)
\end{aligned}$$

and this yields

$$D_f(A \parallel B+\varepsilon I) \geq \operatorname{Tr} Bf\big((B+\varepsilon I)^{-1/2}A(B+\varepsilon I)^{-1/2}\big)+\varepsilon f\Big(\frac{1}{\varepsilon}\langle A_1 z, z\rangle\Big).$$

Since the kernel of A_1 is the same as the kernel of its square root, we have $\langle A_1 z, z\rangle \neq 0$. Therefore, if we take the limit $\varepsilon \searrow 0$ in the above displayed formula, then on the right-hand side the first term is bounded from below and, under the assumption $\lim_{x\to\infty} f(x)/x = \infty$, the second term tends to infinity. The proof is complete. $\square$

In what follows we derive a formula which will be employed several times in the rest of this article. For the proof, we recall that for a fixed $w, z \in \mathcal{H}$ the operation $\otimes$ is defined by $(w \otimes z)v = \langle v, z\rangle w$ for every $v \in \mathcal{H}$. Plainly, for any $T \in \mathcal{L}(\mathcal{H})$, $v, w \in \mathcal{H}$, the computational rules

$$T \cdot w \otimes z = (Tw) \otimes z, \qquad w \otimes z \cdot T \quad = w \otimes (T^* z), \qquad \operatorname{Tr}(w \otimes z) = \langle w, z\rangle$$

concerning the operation $\otimes$ are simple consequences of the definition.

Corollary 3.1. *For any $t > 0, A \in \mathcal{L}(\mathcal{H})_+$ and $P \in \mathcal{P}_1(\mathcal{H})$ with* $\operatorname{supp} P \subseteq \operatorname{supp} A$*, we have*

$$D_f(tP \parallel A) = \lambda(A, P) f\Big(\frac{t}{\lambda(A, P)}\Big).$$

Proof. Consider a unit vector z from the range of P. Then we can write $P = z \otimes z$. Hence we compute

$$\begin{aligned}
&D_f(tP \parallel A)\\
&\quad= \operatorname{Tr} A^{1/2} f(tA^{-1/2} z \otimes z A^{-1/2}) A^{1/2}\\
&\quad= \operatorname{Tr} f\big(t\|A^{-1/2}z\|^2\big) A^{1/2}\big(A^{-1/2}z/\|A^{-1/2}z\|\big) \otimes \big(A^{-1/2}z/\|A^{-1/2}z\|\big) A^{1/2}\\
&\quad= \frac{f(t\|A^{-1/2}z\|^2)}{\|A^{-1/2}z\|^2} = \frac{f(t \operatorname{Tr} PA^{-1})}{\operatorname{Tr} PA^{-1}},
\end{aligned}$$

which, comparing it with (3.2), yields the desired formula. $\square$

We are now in a position to present the proof of Proposition 1.4.

Proof of Proposition 1.4. For temporary use, let us denote $n = \dim \mathcal{H}$. Assume that $D_\eta(\cdot \parallel \cdot)$ is a quantum f-divergence, where $f : [0, \infty[\to \mathbb{R}$ is a strictly convex function such that the limit $\omega_f = \lim_{x\to\infty} f(x)/x$ exists in the extended sense. Choose an orthonormal basis in $\mathcal{H}$, and consider an operator $A \in \mathcal{S}(\mathcal{H})$ and an invertible operator $B \in \mathcal{S}(\mathcal{H})$ whose matrices with respect to that basis are $A = \operatorname{diag}[a_1, a_2, \ldots, a_n]$ and $B = \operatorname{diag}[b_1, b_2, \ldots, b_n]$, respectively. Plainly, the numbers b_i $(i = 1, 2, \ldots, n)$ are strictly positive. By plugging the matrices of A and B into the formula (1.6), we obtain

$$\begin{aligned}
S_f(A \parallel B) = D_\eta(A \parallel B) &= \operatorname{Tr} A \log AB^{-1}\\
&= \sum_{i=1}^n a_i \log\Big(\frac{a_i}{b_i}\Big) = S_\eta(A \parallel B),
\end{aligned}$$

and it means that $S_f(\cdot \parallel \cdot)$ and $S_\eta(\cdot \parallel \cdot)$ are equal on classical probability distributions. It follows from [10, Corollary 2.10] that $S_f(\cdot \parallel \cdot) = D_\eta(\cdot \parallel \cdot)$ coincides with $S_\eta(\cdot \parallel \cdot)$ on quantum states, as well. Therefore, we obtain

$$S_f(A \parallel B) = \begin{cases} \operatorname{Tr} A(\log A - \log B), & \operatorname{supp} A \subseteq \operatorname{supp} B, \\ \infty, & \text{otherwise} \end{cases}$$

for any $A, B \in \mathcal{S}(\mathcal{H})$.

It follows that for any invertible density operator $B \in \mathcal{S}(\mathcal{H})$ and arbitrary density operator $A \in \mathcal{S}(\mathcal{H})$, we have

$$\operatorname{Tr} A(\log A - \log B) = \operatorname{Tr} B^{1/2} A B^{-1/2} \log(B^{-1/2} A B^{-1/2}).$$

Substitute any rank 1 projection $P \in \mathcal{P}_1(\mathcal{H})$ into the place of A. By applying (3.2), Corollary 3.1, and the above displayed equality, we arrive at the equation

$$-\operatorname{Tr} P \log B = \log \operatorname{Tr} P B^{-1}, \tag{3.5}$$

which holds for every invertible density operator $B \in \mathcal{S}(\mathcal{H})$ and rank 1 projection $P \in \mathcal{P}_1(\mathcal{H})$.

Let us consider the spectral decomposition $\sum_{b \in \sigma(B)} b R_b$ of the operator B having at least two different eigenvalues. Here the numbers $\operatorname{Tr} P R_b$ ($b \in \sigma(B)$) are nonnegative and their sum is 1. Hence while on the one hand applying Jensen's inequality to the strictly concave function $x \mapsto \log x$ with the numbers $1/b$ ($b \in \sigma(B)$) in its domain yields

$$\sum_{b \in \sigma(B)} \log \frac{1}{b} \operatorname{Tr} P R_b < \log\Big(\sum_{b \in \sigma(B)} \frac{1}{b} \operatorname{Tr} P R_b \Big),$$

on the other hand applying (3.5) leads us to

$$\sum_{b \in \sigma(B)} \log \frac{1}{b} \operatorname{Tr} P R_b = \log\Big(\sum_{b \in \sigma(B)} \frac{1}{b} \operatorname{Tr} P R_b \Big).$$

This is a contradiction because the operator B has at least two eigenvalues. Therefore, the proof of the proposition is complete. □

Now we turn to the presentation of the proofs of the main results of the paper. In the first one, our strategy is the following. We derive a characterization of the usual order on the set of positive definite operators. Next we show that if a bijective transformation preserves the divergence $D_f(\cdot \parallel \cdot)$, then it is necessarily an order automorphisms. Then we will apply a former structural result concerning order automorphisms on the set of positive definite operators, which helps us to finish the proof.

Proof of Theorem 2.1. We first prove that

$$A \leq B \iff D_f(X \parallel A) \geq D_f(X \parallel B) \quad \big(X \in \mathcal{L}(\mathcal{H})_+^{-1}\big) \tag{3.6}$$

holds for every $A, B \in \mathcal{L}(\mathcal{H})_+^{-1}$. The first part of this statement is in fact [12, Theorem 6] considering only positive definite operators instead of positive semidefinite operators. To see this, assume that $f\colon [0, \infty[\to \mathbb{R}$ is a nonaffine operator convex function satisfying $f(0) = 0$ and consider invertible operators $A, B \in \mathcal{L}(\mathcal{H})_+^{-1}$ such that $A \leq B$ holds. The assumption $A \leq B$ implies that

$$(A^{1/2}B^{-1/2})^* A^{1/2}B^{-1/2} = B^{-1/2}A^{1/2}A^{1/2}B^{-1/2} \leq I,$$

and thus $\|A^{1/2}B^{-1/2}\| \leq 1$.

Therefore, similarly to the last part of the proof of [12, Theorem 6], by the Hansen–Pedersen characterization of operator convexity (see [9]) we infer that

$$\begin{aligned}
&(A^{1/2}B^{-1/2})^* f(A^{-1/2}XA^{-1/2})A^{1/2}B^{-1/2} \\
&\quad \geq f\big((A^{1/2}B^{-1/2})^* A^{-1/2}XA^{-1/2}(A^{1/2}B^{-1/2})\big) \\
&\quad = f(B^{-1/2}XB^{-1/2})
\end{aligned}$$

and thus

$$A^{1/2}f(A^{-1/2}XA^{-1/2})A^{1/2} \geq B^{1/2}f(B^{-1/2}XB^{-1/2})B^{1/2}.$$

By taking the trace, we obtain

$$\operatorname{Tr} A^{1/2}f(A^{-1/2}XA^{-1/2})A^{1/2} \geq \operatorname{Tr} B^{1/2}f(B^{-1/2}XB^{-1/2})B^{1/2}$$

and it verifies the necessity.

As for the sufficiency, assume that $D_f(X \parallel A) \geq D_f(X \parallel B)$ holds for every $X \in \mathcal{L}(\mathcal{H})_+^{-1}$. Let $t > 0$ and X converge to a rank 1 operator tP such that $P \in \mathcal{P}_1(\mathcal{H})$. Then we obtain from Corollary 3.1

$$\lambda(A, P) f\Big(\frac{t}{\lambda(A, P)}\Big) \geq \lambda(B, P) f\Big(\frac{t}{\lambda(B, P)}\Big).$$

Introducing the new variable $x = t/\lambda(B, P)$ and denoting $\gamma = \lambda(B, P)/\lambda(A, P)$ yields

$$f(\gamma x) \geq \gamma f(x) \quad (x > 0).$$

We intend to show that f is strictly convex from which the desired characterization of the order follows rather easily. If f is a nonaffine operator convex function on the interval $]0, +\infty[$, then according to Kraus's theorem (see [11, pp. 40–41]) we obtain that, for every fixed positive number x_0, the first-order divided-difference function

$$R(x, x_0) := \frac{f(x) - f(x_0)}{x - x_0}$$

is nonconstant and operator monotone. Since every nonconstant operator monotone function is strictly monotone-increasing (as follows from the celebrated Löwner's theorem, which claims that every operator monotone function defined on some open interval has a holomorphic extension onto the upper complex half-plane), the function $R(x, x_0)$ is strictly monotone-increasing, as well. Recalling that a function defined on some interval is strictly convex if and only if, for any fixed number from the interval, the corresponding first-order divided-difference function is strictly monotone-increasing gives us that the function f is strictly convex. Since $f(0) = 0$ is satisfied, we have $\gamma \geq 1$, implying that $\lambda(B, P) \geq \lambda(A, P)$ holds for every $P \in \mathcal{P}_1(\mathcal{H})$. This yields $A \leq B$, which verifies (3.6).

As a consequence of (3.6), by the preservation of $D_f(\cdot \parallel \cdot)$ under the transformation ϕ, we infer that the following equivalences hold:

$$\begin{aligned} A \leq B \quad &\iff \quad D_f(X \parallel A) \geq D_f(X \parallel B) \quad \big(X \in \mathcal{L}(\mathcal{H})_+^{-1}\big) \\ &\iff \quad D_f\big(\phi(X) \,\big\|\, \phi(A)\big) \geq D_fl\big(\phi(X) \,\big\|\, \phi(B)\big) \quad \big(X \in \mathcal{L}(\mathcal{H})_+^{-1}\big) \\ &\iff \quad \phi(A) \leq \phi(B). \end{aligned}$$

We conclude that the bijective map ϕ is an order automorphism of $\mathcal{L}(\mathcal{H})_+^{-1}$. The structure of such transformations is described in [15] and, according to [15, Theorem 1], ϕ is of the form

$$\phi(X) = TXT^* \quad \big(X \in \mathcal{L}(\mathcal{H})_+^{-1}\big),$$

where T is an invertible linear or conjugate-linear operator on $\mathcal{H}$. The proof can be completed as in [16, Theorem 1].

The preserver property of ϕ implies that

$$\operatorname{Tr} TAT^* f\big((TAT^*)^{-1/2}(TBT^*)(TAT^*)^{-1/2}\big) = \operatorname{Tr} Af(A^{-1/2}BA^{-1/2})$$

holds for every $A, B \in \mathcal{L}(\mathcal{H})_+^{-1}$. Consider a $t > 0$ such that $f(t) \neq 0$ and plug $B = tA$. This yields $\operatorname{Tr} TAT^* f(tI) = \operatorname{Tr} Af(tI)$ and thus $\operatorname{Tr} AT^*T = \operatorname{Tr} A$ holds for all $A \in \mathcal{L}(\mathcal{H})_+^{-1}$). We deduce that $T^*T = I$, which means that T is either a unitary or an antiunitary operator on $\mathcal{H}$. □

Now we present the proof of Theorem 2.2. The proof technique used in our argument was developed by Molnár and his coauthors in [18, Theorem], and then was employed in [6, Theorem 4], where the author and Molnár applied it to determine the structure of quantum Rényi divergence preservers on the set of density operators. Consequently, our proof is very similar to what the authors have presented in [18, Theorem] and [6, Theorem 4]. In the present paper we focus on the differences, meaning that we present all the computations and core ideas, but if at some point the exact same argument will be used, then we would rather omit the details and refer to [18].

Proof of Theorem 2.2. We first show that ϕ preserves the rank of operators. According to Proposition 1.3, the quantity $D_f(A \parallel B)$ is finite if and only if $\operatorname{supp} A \subseteq \operatorname{supp} B$. We infer from this that ϕ preserves the strict inclusion between density operators on $\mathcal{H}$. If $\mathcal{H}$ is an n-dimensional Hilbert space, then the rank of an operator A is k if and only if there is a strictly increasing (with respect to inclusion) sequence of supports of n density operators on $\mathcal{H}$ such that its kth element is $\operatorname{supp} A$. Therefore, ϕ is rank-preserving.

Next we verify that ϕ is injective. Let $B, B' \in \mathcal{S}(\mathcal{H})$, and suppose that $\phi(B) = \phi(B')$ holds. For every $P \in \mathcal{P}_1(\mathcal{H})$ we have $D_f(\phi(P) \parallel \phi(B)) = D_f(\phi(P) \parallel \phi(B'))$, and by the preserver property of ϕ this implies that $D_f(P \parallel B) = D_f(P \parallel B')$. Therefore, for any $P \in \mathcal{P}_1(\mathcal{H})$ we have $D_f(P \parallel B') < \infty$ if and only if $D_f(P \parallel B) < \infty$, and hence we obtain $\operatorname{supp} B = \operatorname{supp} B'$.

Let us define the function $g : \,]0, \infty[\, \to \mathbb{R}, x \mapsto f(x)/x$. Select any $P \in \mathcal{P}_1(\mathcal{H})$ with $\operatorname{supp} P \subseteq \operatorname{supp} B$, and apply Corollary (3.1) and the preserver property of ϕ.

We deduce that

$$g\Big(\frac{1}{\lambda(B,P)}\Big) = g\Big(\frac{1}{\lambda(B',P)}\Big)$$

holds for every $P \in \mathcal{P}_1(\mathcal{H})$ with $\operatorname{supp} P \subseteq \operatorname{supp} B$. Due to the facts that f is strictly convex and the condition $f(0) = 0$ is satisfied, the function g is strictly monotone-increasing, which implies that

$$\lambda(B,P) = \lambda(B',P)$$

is valid for every $P \in \mathcal{P}_1(\mathcal{H})$ satisfying $\operatorname{supp} P \subseteq \operatorname{supp} B$. Since both the operators B and B' vanish on $\ker B$, we conclude that ϕ is injective.

In the following, $\mathcal{H}$ is assumed to be 2-dimensional. We intend to show that for any $B \in \mathcal{S}(\mathcal{H})$,

$$\big[\min \sigma(B), \max \sigma(B)\big] \subseteq \big[\min \sigma\big(\phi(B)\big), \max \sigma\big(\phi(B)\big)\big]$$

holds, meaning that ϕ can only enlarge the convex hull of the spectrum of the elements of $\mathcal{S}(\mathcal{H})$. To verify this property first observe that the inclusion above holds for all $B \in \mathcal{P}_1(\mathcal{H})$ because ϕ preserves the rank and hence the rank 1 operators, as well. Now select a rank 2 operator $B \in \mathcal{S}(\mathcal{H})$ and set $\mu = \max \sigma(B) \in [1/2, 1[$. Then there are mutually orthogonal projections $Q, R \in \mathcal{P}_1(\mathcal{H})$ such that $B = \mu Q + (1-\mu) R$.

Applying Corollary 3.1 and (3.3), for any $P \in \mathcal{P}_1(\mathcal{H})$ we obtain

$$D_f(P \parallel B) = g\Big(\frac{1}{\mu} \operatorname{Tr} PQ + \frac{1}{1-\mu} \operatorname{Tr} PR\Big). \tag{3.7}$$

Letting P run through the set $\mathcal{P}_1(\mathcal{H})$, the numbers $\operatorname{Tr} PQ, \operatorname{Tr} PR \in [0,1]$ provide all pairs of nonnegative reals such that $\operatorname{Tr} PQ + \operatorname{Tr} PR = 1$. It follows by the continuity of g that the quantity $D_f(P \parallel B)$ runs through the interval $[g(1/\mu), g(1/(1-\mu))]$. Moreover, $D_f(P \parallel B) = g(1/\mu)$ if and only if $\operatorname{Tr} PQ = 1$ which holds exactly when $P = Q$. Similarly, we infer that for any $P \in \mathcal{P}_1(\mathcal{H})$ the number $D_f(\phi(P) \parallel \phi(B))$ belongs to $[g(1/\mu'), g(1/(1-\mu'))]$, where $\mu' = \max \sigma(\phi(B))$. By the preserver property of ϕ, we obtain

$$g(1/\mu') \leq g(1/\mu) \leq g\big(1/(1-\mu)\big) \leq g\big((1-\mu')\big),$$

and this yields

$$\min \sigma\big(\phi(B)\big) \leq \min \sigma(B) \leq \max \sigma(B) \leq \max \sigma\big(\phi(B)\big),$$

which verifies our claim.

The key step in the proof is to show that $\phi(I/2) = I/2$. Assume on the contrary that there is a number $\mu_1 \in]1/2, 1[$ and mutually orthogonal projections $Q_1, R_1 \in \mathcal{P}_1(\mathcal{H})$ such that

$$\phi\Big(\frac{1}{2} I\Big) = \mu_1 Q_1 + (1-\mu_1) R_1. \tag{3.8}$$

We observe that $D_f(P \parallel I/2) = g(2)$ holds for every $P \in \mathcal{P}_1(\mathcal{H})$. By applying this, (3.8), and the preserver property of ϕ, we infer that

$$g\Big(\frac{1}{\mu_1} \operatorname{Tr} \phi(P) Q_1 + \frac{1}{1-\mu_1} \operatorname{Tr} \phi(P) R_1\Big) = g(2),$$

and due to the fact that g is strictly monotone this implies that

$$\frac{1}{\mu_1} \operatorname{Tr} \phi(P) Q_1 + \frac{1}{1-\mu_1} \operatorname{Tr} \phi(P) R_1 = 2. \tag{3.9}$$

As $\operatorname{Tr} \phi(P) Q_1 + \operatorname{Tr} \phi(P) R_1 = 1$, we get that the number 2 is a convex combination of $1/\mu_1$ and $1/(1-\mu_1)$. Due to the strict monotonicity of the function $x \mapsto 1/x$, these latter numbers are different. Since every element of a nondegenerate compact interval is a unique convex combination of the endpoints, this implies that $\operatorname{Tr} \phi(P) Q_1$ has the same value for any $P \in \mathcal{P}_1(\mathcal{H})$ and the same holds for $\operatorname{Tr} \phi(P) R_1$, as well. We intend to prove that

$$\operatorname{Tr} \phi(P) Q_1 > \operatorname{Tr} \phi(P) R_1. \tag{3.10}$$

To this end, we note that if $x \in]1/2, 1[$, then in any representation of 2 as a convex combination of $1/x$ and $1/(1-x)$ the coefficient of the former term is necessarily greater than the coefficient of the latter one, because the function $x \mapsto 1/x$ is strictly monotone decreasing. Referring to (3.9), it follows that $\operatorname{Tr} \phi(P) Q_1 > 1/2$, which verifies (3.10).

As we have learned before, the quantity $\operatorname{Tr} \phi(P) Q_1$ remains constant when P runs through the set $\mathcal{P}_1(\mathcal{H})$, and the same holds for $\operatorname{Tr} \phi(P) R_1$, as well. Consequently, we can rewrite (3.9) in the form

$$\frac{\kappa}{\mu_1} + \frac{1-\kappa}{1-\mu_1} = 2, \tag{3.11}$$

where $\kappa = \operatorname{Tr} \phi(P) Q_1$, and this number is plainly different from $0, 1$ (see the discussion above).

Next let us consider $\phi(\phi(I/2))$. We have

$$\phi\Big(\phi\Big(\frac{1}{2} I\Big)\Big) = \mu_2 Q_2 + (1-\mu_2) R_2,$$

for some $1/2 \le \mu_2 < 1$ and mutually orthogonal projections $Q_2, R_2 \in \mathcal{P}_1(\mathcal{H})$. Since ϕ can only enlarge the convex hull of the spectrum and $\mu_1 > 1/2$, it follows that $\mu_2 \ge \mu_1 > 1/2$. Select an arbitrary rank 1 projection $P \in \mathcal{P}_1(\mathcal{H})$ and set $P_2 = \phi(\phi(P))$. Due to the preserver property of ϕ similarly to (3.9), we infer that

$$\begin{aligned} g(2) &= D_f\Big(\phi\big(\phi(P)\big) \,\Big\|\, \phi\Big(\phi\Big(\frac{1}{2} I\Big)\Big)\Big) \\ &= D_f\big(P_2 \,\big\|\, \mu_2 Q_2 + (1-\mu_2) R_2\big) \\ &= g\Big(\frac{1}{\mu_2} \operatorname{Tr} P_2 Q_2 + \frac{1}{1-\mu_2} \operatorname{Tr} P_2 R_2\Big) \end{aligned}$$

and this yields

$$\frac{1}{\mu_1} \operatorname{Tr} P_2 Q_2 + \frac{1}{1-\mu_1} \operatorname{Tr} P_2 R_2 = 2, \tag{3.12}$$

where $\mu_2 > 1/2$ is fixed. Since the pair $\operatorname{Tr} P_2 Q_2, \operatorname{Tr} P_2 R_2$ of nonnegative real numbers has sum 1, it follows just as above that the numbers $\operatorname{Tr} P_2 Q_2$ and $\operatorname{Tr} P_2 R_2$

are also fixed (i.e., they do not change when P varies). Moreover, by the strict convexity of the function $x \mapsto 1/x$, we also necessarily have

$$\operatorname{Tr} P_2 Q_2 > \operatorname{Tr} P_2 R_2. \tag{3.13}$$

Now choose unit vectors u and v from the ranges of Q_1 and R_1, respectively. Consider a unit vector from the range of Q_2. Let ξ, η be its coordinates with respect to the basis $\{u, v\}$. It is easy to see that the representing matrix of Q_2 is

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \begin{pmatrix} \overline{\xi} \\ \overline{\eta} \end{pmatrix}^t,$$

where ^{t} denotes the transposition. Moreover, since P_2 is a rank 1 projection which is the image (under ϕ) of a rank 1 projection, its matrix representation is of the form

$$\begin{pmatrix} \kappa & \varepsilon\sqrt{\kappa(1-\kappa)} \\ \overline{\varepsilon}\sqrt{\kappa(1-\kappa)} & 1-\kappa \end{pmatrix},$$

where κ is the same as in (3.11), and $\varepsilon \in \mathbb{C}$ with $|\varepsilon| = 1$ varies as P varies. Then similarly to the corresponding part in [18, Theorem], we can verify that the column vector

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

is a scalar multiple of

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Obviously, this can happen only when $Q_2 = Q_1$ or $Q_2 = R_1$. Since P_2 is the image of a rank 1 projection under ϕ, it follows from (3.10) that

$$\operatorname{Tr} P_2 Q_1 > \operatorname{Tr} P_2 R_1. \tag{3.14}$$

If $Q_2 = R_1$, then $Q_1 = R_2$ and due to (3.13) we have

$$\operatorname{Tr} P_2 R_1 > \operatorname{Tr} P_2 Q_1$$

which contradicts (3.14). Therefore, the possibility $Q_2 = R_1$ is ruled out and, consequently, $Q_2 = Q_1$ and $R_2 = R_1$. Thus we obtain

$$\phi\Big(\phi\Big(\frac{1}{2}I\Big)\Big) = \mu_2 Q_1 + (1-\mu_2) R_1. \tag{3.15}$$

By (3.12) we have

$$\frac{1}{\mu_2} \operatorname{Tr} P_2 Q_1 + \frac{1}{1-\mu_2} \operatorname{Tr} P_2 R_1 = 2.$$

On the other hand, referring to the sentence preceding (3.11), we see that

$$\operatorname{Tr} P_2 Q_1 = \kappa$$

and $\operatorname{Tr} P_2 R_1 = 1 - \kappa$, and thus it follows that

$$\frac{\kappa}{\mu_2} + \frac{1-\kappa}{1-\mu_2} = 2. \tag{3.16}$$

Plainly, the solutions of the equation

$$\frac{\kappa}{x} + \frac{1-\kappa}{1-x} = 2 \tag{3.17}$$

are only $x = \kappa$ and $x = 1/2$, but as we have already noticed, both the numbers μ_2, μ_1 are also solutions. Due to the relations $\mu_2 \geq \mu_1 > 1/2$ it follows that $\mu_1 = \mu_2$, and then referring to (3.8) and (3.15) we see that $\phi(\phi(I/2)) = \phi(I/2)$. Since ϕ is injective, this gives $\phi(I/2) = I/2$, which verifies our claim.

We now prove that ϕ is spectrum-preserving; that is, $\sigma(\phi(A)) = \sigma(A)$. In order to do so, let $I/2 \neq A \in \mathcal{S}(\mathcal{H})$ be a rank 2 operator and denote by $\mu \in \,]1/2, 1[$ its maximal eigenvalue. Consider the function

$$h\colon \ ]0,1[\,\to \mathbb{R}, \qquad x \mapsto \frac{f(2x) + f(2(1-x))}{2}.$$

We note that h is strictly convex and symmetric with respect to the middle point $1/2$ of its domain, which implies that h is strictly monotonic on the interval $[1/2, 1[$. Moreover, we infer that $D_f(A \parallel I/2) = h(\mu)$ holds. Since ϕ sends $I/2$ to itself, we similarly have $D_f(\phi(A) \parallel I/2) = h(\mu')$, where μ' denotes the maximal eigenvalue of $\phi(A)$. By virtue of the preserver property of ϕ, we conclude that $h(\mu) = h(\mu')$ and thus we obtain $\mu = \mu'$. This verifies our claim.

Our aim is now to show that ϕ preserves the nonzero transition probability between rank 1 projections; that is,

$$\operatorname{Tr} PQ = \operatorname{Tr} \phi(P)\phi(Q)$$

holds for every $P, Q \in \mathcal{P}_1(\mathcal{H})$ whenever the supports of the operators P, Q are not orthogonal to each other. Let Q, R be mutually orthogonal projections in $\mathcal{P}_1(\mathcal{H})$ and pick a number $\mu \in \,]1/2, 1[$. If we consider the operator $B = \mu Q + (1-\mu)R$, then by the spectrum preserver property of ϕ we can choose other $Q', R' \in \mathcal{P}_1(\mathcal{H})$ such that $\phi(B) = \mu Q' + (1-\mu)R'$. Referring to the sentences following (3.7) we have seen that the quantity $D_f(P \parallel B)$ runs through the interval $[g(1/\mu), g(1/(1-\mu))]$ and the minimum is taken only in the case where $P = Q$. Therefore, the following equivalences hold

$$\begin{aligned} P = Q \quad &\iff \quad D_f(P \parallel B) = g(1/\mu) \\ &\iff \quad D_f\big(\phi(P) \parallel \phi(B)\big) = g(1/\mu) \\ &\iff \quad \phi(P) = Q' \end{aligned}$$

implying that $\phi(Q) = Q'$ and, similarly, $\phi(R) = R'$. It means that ϕ preserves the orthogonality between rank 1 projections. In addition, we have

$$\phi(B) = \mu\phi(Q) + (1-\mu)\phi(R).$$

Now let us consider two different projections P, Q which are not orthogonal to each other. Choose another orthogonal projection R such that $Q + R = I$ and pick a number $\mu \in \,]1/2, 1[$. Compute

$$D_f\big(P \parallel \mu Q + (1-\mu)R\big) = g\Big(\frac{1}{\mu}\operatorname{Tr} PQ + \frac{1}{1-\mu}\operatorname{Tr} PR\Big)$$

and

$$D_f\big(\phi(P) \parallel \phi\big(\mu Q + (1-\mu)R\big)\big) = D_f\big(\phi(P) \parallel \mu\phi(Q) + (1-\mu)\phi(R)\big)$$
$$= g\Big(\frac{1}{\mu}\operatorname{Tr}\phi(P)\phi(Q) + \frac{1}{1-\mu}\operatorname{Tr}\phi(P)\phi(R)\Big).$$

Due to the preserver property of ϕ and the strict monotonicity of g, we infer that

$$\frac{1}{\mu}\operatorname{Tr} PQ + \frac{1}{1-\mu}\operatorname{Tr} PR = \frac{1}{\mu}\operatorname{Tr}\phi(P)\phi(Q) + \frac{1}{1-\mu}\operatorname{Tr}\phi(P)\phi(R),$$

and this yields

$$\operatorname{Tr} PQ = \operatorname{Tr}\phi(P)\phi(Q). \tag{3.18}$$

Therefore, ϕ preserves the nonzero transition probability between rank 1 projections.

Above, $\mathcal{H}$ is supposed to be 2 dimensional. Using the same argument as in the corresponding part of [18, Theorem] it can be shown that (3.18) holds for every $P, Q \in \mathcal{P}_1(\mathcal{H})$ in an arbitrary finite-dimensional Hilbert space $\mathcal{H}$. By the nonbijective version of Wigner's theorem (see, e.g., [7]), we infer that there exists either a unitary or an antiunitary operator U on $\mathcal{H}$ such that

$$\phi(P) = UPU^* \quad \big(P \in \mathcal{P}_1(\mathcal{H})\big).$$

Next consider the transformation

$$\psi\colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}), \qquad A \mapsto U^*\phi(A)U.$$

Since both the unitary and antiunitary similarity transformations leave the quantity $D_f(\cdot \parallel \cdot)$ invariant, the map ψ also preserves this quantity. Moreover, it has the additional property that it acts as the identity on $\mathcal{P}_1(\mathcal{H})$. Similar to the first part of this proof, it follows that ψ preserves the inclusion between the supports of elements of $\mathcal{S}(\mathcal{H})$. This yields

$$\operatorname{supp} P \subseteq \operatorname{supp} A \quad \iff \quad \operatorname{supp} P \subseteq \operatorname{supp}\psi(A),$$

and we obtain that $\operatorname{supp} A = \operatorname{supp}\psi(A)$ holds. Moreover, we have

$$g\Big(\frac{1}{\lambda(A,P)}\Big) = g\Big(\frac{1}{\lambda(\psi(A),P)}\Big)$$

for every $P \in \mathcal{P}_1(\mathcal{H})$, satisfying $\operatorname{supp} P \subseteq \operatorname{supp} A$. Since g is strictly increasing, we infer from this that

$$\lambda(A,P) = \lambda\big(\psi(A),P\big) \quad \big(P \in \mathcal{P}_1(\mathcal{H})\big),$$

implying that $A = \psi(A)$. This means that $\phi(A) = UAU^*$ holds for every $A \in \mathcal{S}(\mathcal{H})$, thus completing the proof of the theorem. $\square$

In the remaining part of the paper we present the proof of our last result. In the proof we employ the core ideas of [25].

Proof of Theorem 2.3. The most crucial part of the proof is to prove that if ϕ is a bijective transformation preserving the divergence $D_f(\cdot \parallel \cdot)$ on the set of positive semidefinite operators, then ϕ preserves the trace of operators. In the case where $f(1) \neq 0$, this property follows from the observation

$$f(1)\operatorname{Tr} A = D_f(A \parallel A) = D_f\big(\phi(A) \,\big\|\, \phi(A)\big) = f(1)\operatorname{Tr}\phi(A) \quad \big(A \in \mathcal{L}(\mathcal{H})_+\big).$$

So, in what follows we assume that the condition $f(1) = 0$ is satisfied. The continuity of f and the limit property $\lim_{x\to\infty} f(x)/x = \infty$ implies immediately that f is bounded from below and thus there is a real number c such that $c = \inf_{x\in[0,\infty[} f(x)$. Plainly, $c < 0$ under the assumption $f(0) = f(1) = 0$ because of the strict convexity of the function f. We assert that

$$\inf_{X\in\mathcal{L}(\mathcal{H})_+} D_f(X \parallel A) = c\operatorname{Tr} A \quad \big(A \in \mathcal{L}(\mathcal{H})_+\big). \tag{3.19}$$

Indeed, since

$$D_f(cA \parallel A) = c\operatorname{Tr} A \quad \big(A \in \mathcal{L}(\mathcal{H})_+\big)$$

we obtain that $\inf_{X\in\mathcal{L}(\mathcal{H})_+} D_f(X \parallel A) \le c\operatorname{Tr} A$.

To see the reverse inequality, if $\operatorname{supp} X \nsubseteq \operatorname{supp} A$, we conclude that $D_f(X \parallel A) = \infty > c\operatorname{Tr} A$. Otherwise, if $A = 0$ and $\operatorname{supp} X \subseteq \operatorname{supp} A$, then we necessarily have $X = 0$ and thus $D_f(X \parallel A) = 0 = c\operatorname{Tr} A$ since $\operatorname{Tr} A = 0$. If $A \neq 0$, then according to (1.5) we calculate

$$\begin{aligned}\frac{D_f(X \parallel A)}{\operatorname{Tr} A} &= \operatorname{Tr}\Big(\frac{A|_{\operatorname{supp} A}}{\operatorname{Tr} A} f\big((A|_{\operatorname{supp} A})^{-1/2} P_A X P_A (A|_{\operatorname{supp} A})^{-1/2}\big)\Big)\\ &\ge f\Big(\frac{\operatorname{Tr}((A|_{\operatorname{supp} A})^{1/2} P_A X P_A (A|_{\operatorname{supp} A})^{-1/2})}{\operatorname{Tr} A}\Big) \ge c,\end{aligned}$$

where the first inequality follows from [10, Lemma A.2], which claims that for every self-adjoint operator H and density operator D acting on a finite-dimensional Hilbert space and for any convex function f, we necessarily have $\operatorname{Tr} Df(H) \ge f(\operatorname{Tr} DH)$. Therefore, $\inf_{X\in\mathcal{L}(\mathcal{H})_+} D_f(X \parallel A) \ge c\operatorname{Tr} A$ holds for every $A \in \mathcal{L}(\mathcal{H})_+$. Then the preserver property and the bijectivity of ϕ results in

$$\begin{aligned}c\operatorname{Tr} A &= \inf_{X\in\mathcal{L}(\mathcal{H})_+} D_f(X \parallel A) = \inf_{X\in\mathcal{L}(\mathcal{H})_+} D_f\big(\phi(X) \,\big\|\, \phi(A)\big)\\ &= \inf_{X\in\mathcal{L}(\mathcal{H})_+} D_f\big(X \,\big\|\, \phi(A)\big) = c\operatorname{Tr}\phi(A).\end{aligned}$$

Since $c \neq 0$, the above displayed formula implies that ϕ is trace-preserving.

Now for any fixed positive number q let us define the operator

$$\psi_q\colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}), \qquad A \mapsto \frac{1}{q}\phi(qA).$$

Since ϕ preserves the trace, ψ_q is a bijective map on the set of density operators, and one can verify easily that it preserves the divergence $D_f(\cdot \parallel \cdot)$, as well. Therefore, applying Theorem 2.2 gives us that for every fixed $q \in\,]0,\infty[$ there exists a unitary or antiunitary operator U_q on $\mathcal{H}$ such that

$$\psi_q(A) = U_q A U_q^* \quad \big(A \in \mathcal{S}(\mathcal{H})\big).$$

Finally, we show that the operator U_q is independent of the choice of the number q. Let $P \in \mathcal{P}_1(\mathcal{H})$ and $q, r \in]0, \infty[$ be arbitrary. Due to Proposition 1.3 and the preserver property of ϕ, the following equivalences hold:

$$\begin{aligned} \operatorname{supp} qP \subseteq \operatorname{supp} rP \quad &\iff \quad D_f(qP \,\|\, rP) < \infty \\ &\iff \quad D_f\big(\phi(qP) \,\big\|\, \phi(rP)\big) < \infty \\ &\iff \quad \operatorname{supp} \phi(qP) \subseteq \operatorname{supp} \phi(rP). \end{aligned}$$

Similarly to the argument at the beginning of the proof of Theorem 2.2, we can verify that ϕ preserves the rank of operators. Hence the operators $\phi(qP)$ and $\phi(rP)$ are rank 1's with the same support. Altogether we have

$$\psi_q(P) = \frac{1}{q}\phi(qP) = \frac{\phi(qP)}{\operatorname{Tr} qP} = \frac{\phi(rP)}{\operatorname{Tr} rP} = \frac{1}{r}\phi(rP) = \psi_r(P)$$

for every $P \in \mathcal{P}_1(\mathcal{H})$ and $q, r \in]0, \infty[$. Since both the maps ψ_q and ψ_r are affine and every element of $\mathcal{S}(\mathcal{H})$ is a convex combination of rank 1 projections, we obtain that $\psi_q = \psi_r$ holds for every positive numbers q, r. The proof of the theorem is complete. $\square$

Acknowledgments. The author wishes to express his deep gratitude to Lajos Molnár for professional and fruitful discussions on the topic of relative entropy preserver problems and for mentioning some open problems discussed here.

The author was partially supported by the "Lendület" Program (LP2012-46/2012) of the Hungarian Academy of Sciences and the National Research, Development and Innovation Office—NKFIH Reg. No. K115383.

References

1. P. Busch and S. P. Gudder, *Effects as functions on projective Hilbert space*, Lett. Math. Phys. **47** (1999), no. 4, 329–337. Zbl 0932.47055. MR1693743. DOI 10.1023/A:1007573216122. 750
2. E. Carlen, "Trace inequalities and quantum entropy: An introductory course" in *Entropy and the Quantum (Tucson, 2009)*, Contemp. Math. **529**, Amer. Math. Soc., Providence, 2010, 73–140. Zbl 1218.81023. MR2681769. DOI 10.1090/conm/529/10428. 746, 751
3. G. Chevalier, "Wigner's theorem and its generalizations" in *Handbook of Quantum Logic and Quantum Structures*, Elsevier, Amsterdam, 2007, 429–475. Zbl 1126.81031. MR2423613. DOI 10.1016/B978-044452870-4/50032-7. 745
4. I. Csiszár, *Information-type measures of difference of probability distributions and indirect observations*, Studia. Sci. Math. Hungar. **2** (1967), 299–318. Zbl 0157.25802. MR0219345.
5. S. S. Dragomir, *A new quantum f-divergence for trace class operators in Hilbert spaces*, Entropy **16** (2014), no. 11, 5853–5875. Zbl 1338.94035. MR3285443. DOI 10.3390/e16115853. 745, 747, 748
6. M. Gaál and L. Molnár, *Transformations on density operators and on positive definite operators preserving the quantum Rényi divergence*, Period. Math. Hungar. **74** (2017), no. 1, 88–107. Zbl 06707857. MR3604111. DOI 10.1007/s10998-016-0174-8. 745, 750, 755
7. Gy. P. Gehér, *An elementary proof for the non-bijective version of Wigner's theorem*, Phys. Lett. A **378** (2014), no. 30–31, 2054–2057. Zbl 1331.46066. MR3226075. DOI 10.1016/j.physleta.2014.05.039. 745, 760
8. M. Győry, *A new proof of Wigner's theorem*, Rep. Math. Phys. **54** (2004), no. 2, 159–167. Zbl 1161.81381. MR2107864. DOI 10.1016/S0034-4877(04)80012-0. 745

9. F. Hansen and G. K. Pedersen, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann. **258** (1982), no. 3, 229–241. Zbl 0473.47011. MR1513286. DOI 10.1007/BF01450679. 754
10. F. Hiai, M. Mosonyi, D. Petz, and C. Bény, *Quantum f-divergences and error correction*, Rev. Math. Phys. **23** (2011), no. 7, 691–747. Zbl 1230.81007. MR2826462. DOI 10.1142/S0129055X11004412. 747, 753, 761
11. F. Kraus, *Über konvexe matrixfuntionen*, Math. Z. **41** (1936), no. 1, 18–42. Zbl 0013.39701. MR1545602. DOI 10.1007/BF01180403. 754
12. K. Matsumoto, *A new quantum version of f-divergence*, preprint, arXiv:1311.4722v3 [quant-ph]. 745, 753, 754
13. L. Molnár, *Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces*, Lecture Notes in Math. **1895**, Berlin, Springer, 2007. Zbl 1119.47001. MR2267033. 745
14. L. Molnár, *Maps on states preserving the relative entropy*, J. Math. Phys. **49** (2008), no. 3, art. ID 032114. Zbl 1153.81407. MR2406782. DOI 10.1063/1.2898693. 745
15. L. Molnár, *Order automorphisms on positive definite operators and a few applications*, Linear Algebra Appl. **434** (2011), no. 10, 2158–2169. Zbl 1220.47051. MR2781684. DOI 10.1016/j.laa.2010.12.007. 755
16. L. Molnár, *Two characterizations of unitary-antiunitary similarity transformations of positive definite operators on a finite-dimensional Hilbert space*, Ann. Univ. Sci. Budapest Eötvös Sect. Math. **58** (2015), 83–93. Zbl 06656760. MR3467456. 745, 746, 755
17. L. Molnár and G. Nagy, *Isometries and relative entropy preserving maps on density operators*, Linear Multilinear Algebra **60** (2012), no. 1, 93–108. Zbl 1241.47034. MR2869675. DOI 10.1080/03081087.2011.570267. 745
18. L. Molnár, G. Nagy, and P. Szokol, *Maps on density operators preserving quantum f-divergences*, Quantum Inf. Process. **12** (2013), no. 7, 2309–2323. Zbl 1270.81021. MR3065488. DOI 10.1007/s11128-013-0528-6. 745, 747, 748, 755, 758, 760
19. L. Molnár and P. Szokol, *Maps on states preserving the relative entropy, II*, Linear Algebra Appl. **432** (2010), no. 12, 3343–3350. Zbl 1187.47030. MR2639288. DOI 10.1016/j.laa.2010.01.025. 745
20. M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, *On quantum Rényi entropies: A new generalization and some properties*, J. Math. Phys. **54** (2013), no. 12, art. ID 122203. Zbl 1290.81016. MR3156110. DOI 10.1063/1.4838856. 750
21. M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer, Berlin, 1993. Zbl 0891.94008. MR1230389. DOI 10.1007/978-3-642-57997-4. 746, 748
22. D. Petz, *Quasientropies for states of a von Neumann algebra*, Publ. Res. Inst. Math. Sci. **21**(1985), no. 4, 787–800. Zbl 0606.46039. MR0817164. DOI 10.2977/prims/1195178929. 745
23. D. Petz, *Quasi-entropies for finite quantum systems*, Rep. Math. Phys. **23** (1986), no. 1, 57–65. Zbl 0629.46061. MR0868631. DOI 10.1016/0034-4877(86)90067-4. 745
24. D. Virosztek, *Maps on quantum states preserving Bregman and Jensen divergences*, Lett. Math. Phys. **106** (2016), no. 9, 1217–1234. Zbl 06628823. MR3533566. DOI 10.1007/s11005-016-0868-0. 745
25. D. Virosztek, *Quantum f-divergence preserving maps on positive semidefinite operators acting on finite dimensional Hilbert spaces*, Linear Algebra Appl. **501** (2016), 242–253. Zbl 1341.47044. MR3485066. DOI 10.1016/j.laa.2016.03.031. 760

Department of Analysis, Bolyai Institute, University of Szeged, H-6720 Szeged, Aradi vértanúk tere 1., Hungary.

E-mail address: gaal@math.u-szeged.hu; marcell.gaal.91@gmail.com