Commun. Math. Phys. 334, 1617–1648 (2015)
Digital Object Identifier (DOI) 10.1007/s00220-014-2248-x

Communications in
Mathematical Physics

Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Rényi Relative Entropies

Milán Mosonyi[1,2], **Tomohiro Ogawa**[3]

[1] Física Teòrica: Informació i Fenomens Quàntics, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. E-mail: milan.mosonyi@gmail.com
[2] Mathematical Institute, Budapest University of Technology and Economics, Egry József u 1., Budapest 1111, Hungary
[3] Graduate School of Information Systems, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan. E-mail: ogawa@is.uec.ac.jp

Received: 19 December 2013 / Accepted: 5 August 2014
Published online: 4 December 2014 –

Abstract: We show that the new quantum extension of Rényi's α-relative entropies, introduced recently by Müller-Lennert et al. (J Math Phys 54:122203, 2013) and Wilde et al. (Commun Math Phys 331(2):593–622, 2014), have an operational interpretation in the strong converse problem of quantum hypothesis testing. Together with related results for the direct part of quantum hypothesis testing, known as the quantum Hoeffding bound, our result suggests that the operationally relevant definition of the quantum Rényi relative entropies depends on the parameter α: for $\alpha < 1$, the right choice seems to be the traditional definition $D_\alpha^{(\text{old})}(\rho \,\|\, \sigma) := \frac{1}{\alpha-1}\log \operatorname{Tr} \rho^\alpha \sigma^{1-\alpha}$, whereas for $\alpha > 1$ the right choice is the newly introduced version $D_\alpha^{(\text{new})}(\rho \,\|\, \sigma) := \frac{1}{\alpha-1}\log \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha$.

On the way to proving our main result, we show that the new Rényi α-relative entropies are asymptotically attainable by measurements for $\alpha > 1$. From this, we obtain a new simple proof for their monotonicity under completely positive trace-preserving maps.

1. Introduction

Rényi, in his seminal paper [45], introduced a generalization of the Kullback–Leibler divergence (relative entropy). According to his definition, the α-divergence of two probability distributions (more generally, two positive functions) p and q on a finite set $\mathcal{X}$ for a parameter $\alpha \in [0, +\infty) \setminus \{1\}$ is given by

$$D_\alpha(p \,\|\, q) := \begin{cases} \frac{1}{\alpha-1}\log \sum_{x\in\mathcal{X}} p(x)^\alpha q(x)^{1-\alpha} - \frac{1}{\alpha-1}\log \sum_{x\in\mathcal{X}} p(x), & \operatorname{supp} p \subseteq \operatorname{supp} q \text{ or } \alpha \in [0,1), \\ +\infty, & \text{otherwise.} \end{cases} \tag{1}$$

The limit $\alpha \to 1$ yields the standard relative entropy. These quantities turned out to play a central role in information theory and statistics; indeed, the Rényi relative entropies and derived quantities quantify the trade-off between the exponents of the relevant quantities in many information-theoretic tasks, including hypothesis testing, source coding and noisy channel coding; see, e.g. [10] for an overview of these results. It was also shown in [10] that the Rényi relative entropies, and other related quantities, like the Rényi entropies and the Rényi capacities, have direct operational interpretations as so-called generalized cutoff rates in the corresponding information-theoretic tasks.

In quantum theory, the state of a system is described by a density operator instead of a probability distribution, and the definition (1) can be extended for pairs of density operators (more generally, positive operators) in various inequivalent ways, due to the non-commutativity of operators. There are some basic requirements any such extension should satisfy; most importantly, positivity and monotonicity under CPTP (completely positive and trace-preserving) maps. That is, if D_α is an extension of (1) to pairs of positive semidefinite operators, then it should satisfy

$$D_\alpha(\rho \,\|\, \sigma) \geq 0 \quad \text{and} \quad D_\alpha(\rho \,\|\, \sigma) = 0 \Longleftrightarrow \rho = \sigma \quad \text{(positivity)}$$

for any density operators ρ, σ and $\alpha > 0$, and if $\mathcal{F}$ is a CPTP map then

$$D_\alpha(\mathcal{F}(\rho) \,\|\, \mathcal{F}(\sigma)) \leq D_\alpha(\rho \,\|\, \sigma) \quad \text{(monotonicity)} \tag{2}$$

should hold.

One formal extension has been known in the literature for a long time, defined as

$$D_\alpha^{(\mathrm{old})}(\rho \,\|\, \sigma) := \begin{cases} \frac{1}{\alpha-1}\log \operatorname{Tr} \rho^\alpha \sigma^{1-\alpha} - \frac{1}{\alpha-1}\log \operatorname{Tr} \rho, & \operatorname{supp}\rho \subseteq \operatorname{supp}\sigma \text{ or } \alpha \in [0,1), \\ +\infty, & \text{otherwise.} \end{cases} \tag{3}$$

Hölder's inequality ensures positivity of $D_\alpha^{(\mathrm{old})}$ for every $\alpha > 0$. Monotonicity has been proved for $\alpha \in [0,2]\setminus\{1\}$ with various methods [29,40,48], but it doesn't hold for $\alpha > 2$ in general, as it was noted, e.g., in [34]. Monotonicity under measurements, however, is still true for $\alpha > 2$ [17]. In the limit $\alpha \to 1$, these divergences yield Umegaki's relative entropy [49]

$$\begin{aligned} D_1(\rho \,\|\, \sigma) &:= \lim_{\alpha \to 1} D_\alpha^{(\mathrm{old})}(\rho \,\|\, \sigma) \\ &= D(\rho \,\|\, \sigma) := \begin{cases} \frac{1}{\operatorname{Tr}\rho} \operatorname{Tr} \rho(\log\rho - \log\sigma), & \operatorname{supp}\rho \subseteq \operatorname{supp}\sigma, \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned} \tag{4}$$

The quantum Stein's lemma [22,39] gives an operational interpretation to Umegaki's relative entropy (which we will call simply relative entropy for the rest) in a state discrimination problem, as the optimal decay rate of the type II error under the assumption that the type I error goes to 0 (see Sect. 4.1 for details). This shows that Umegaki's relative entropy is the right non-commutative extension of the Kullback–Leibler divergence from an information-theoretic point of view.

It has been shown in [31] that, similarly to the classical case, the Rényi α-relative entropies $D_\alpha^{(\mathrm{old})}$ with $\alpha \in (0,1)$ have a direct operational interpretation as generalized cutoff rates in binary state discrimination. This in turn is based on the so-called quantum Hoeffding bound theorem, which quantifies the trade-off between the optimal exponential decay rates of the two error probabilities in binary state discrimination [2,18,24,36].

In more detail, it says that if the type II error is required to vanish asymptotically as $\sim e^{-nr}$ for some $r > 0$ (n is the number of the copies of the system, all prepared in state ρ or all prepared in state σ) then the optimal type I error goes to 0 exponentially fast with the exponent given by the Hoeffding divergence

$$H_r(\rho\|\sigma) := \sup_{0<\alpha<1} \frac{\alpha-1}{\alpha}\left[r - D_\alpha^{(\mathrm{old})}(\rho\,\|\,\sigma)\right], \tag{5}$$

as long as $r < D(\rho\,\|\,\sigma)$. The transformation rule defining $H_r(\rho\|\sigma)$ from the α-relative entropies can be inverted, and $D_\alpha^{(\mathrm{old})}(\rho\,\|\,\sigma)$ can be expressed in terms of the Hoeffding divergences for any $\alpha \in (0,1)$. These results suggest that $D_\alpha^{(\mathrm{old})}$ gives the right quantum extension of the Rényi α-relative entropies for the parameter range $\alpha \in (0,1)$.

Recently, a new quantum extension of the Rényi α-relative entropies has been proposed in [34,50], defined as

$$D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma) := \begin{cases} \frac{1}{\alpha-1}\log \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^{\alpha} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho, & \operatorname{supp}\rho\subseteq\operatorname{supp}\sigma \text{ or } \alpha\in[0,1),\\ +\infty, & \text{otherwise.}\end{cases} \tag{6}$$

These new Rényi divergences also yield Umegaki's relative entropy in the limit $\alpha \to 1$. Monotonicity for the range $\alpha \in (1,2]$ has been shown in [34,50] and extended to $\alpha \in (1,+\infty)$ in [4] and, independently and with a different proof method, for the range $\alpha \in [\frac{1}{2},1)\cup(1,+\infty)$ in [13]. It is claimed in [34] that these new Rényi relative entropies are not monotone for $\alpha \in [0,\frac{1}{2})$. Positivity follows immediately from the monotonicity for $\alpha \in [\frac{1}{2},1)\cup(1,+\infty)$. The Araki–Lieb–Thirring inequality [1,30] (see also [6, Theorem IX.2.10]) implies that

$$D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma) \le D_\alpha^{(\mathrm{old})}(\rho\,\|\,\sigma) \tag{7}$$

for every ρ,σ and $\alpha \in (0,+\infty)\setminus\{1\}$. Moreover, the results of [23] yield that for non-commuting operators the above inequality is strict for all $\alpha \in (0,+\infty)\setminus\{1\}$. The converse Araki–Lieb–Thirring inequality of [3] implies lower bounds on $D_\alpha^{(\mathrm{new})}$ in terms of $D_\alpha^{(\mathrm{old})}$ [32].

In this paper we show that the new Rényi relative entropies with $\alpha > 1$ play the same role in the converse part of binary state discrimination as the old Rényi relative entropies with $\alpha \in (0,1)$ play in the direct part. Namely, we show (in Theorem 4.10) that if the type II error is required to vanish asymptotically as $\sim e^{-nr}$ with some $r > D(\rho\,\|\,\sigma)$ then the optimal type I error goes to 1 exponentially fast, with the exponent given by the converse Hoeffding divergence

$$H_r^*(\rho\|\sigma) := \sup_{1<\alpha} \frac{\alpha-1}{\alpha}\left[r - D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma)\right]. \tag{8}$$

From this, we derive (in Theorem 4.18) a representation of the new Rényi relative entropies as generalized cutoff rates in the strong converse domain, thus providing a direct operational interpretation of the new Rényi relative entropies for $\alpha > 1$. These results are direct quantum counterparts of the well-known classical results by Han and Kobayashi [14] and Csiszár [10]. In the quantum case, Hayashi [17] obtained a limiting formula for the strong converse exponent using the classical Rényi relative entropies;

see Remarks 3.4 and 4.14. Our Eq. (8) can be seen as a single-letterization of Hayashi's exponent.

In the proof we only use the monotonicity of the new Rényi relative entropies under pinching [34, Proposition 13], and show (in Theorem 3.7) that the new Rényi relative entropies can be asymptotically attained by measurements, similarly to the relative entropy [22]. Based on this, we provide a simple new proof for the monotonicity of $D_\alpha^{(\text{new})}$ under CPTP maps for $\alpha > 1$. We give an overview of the monotonicity and attainability properties of the old and the new Rényi relative entropies in Appendix A.

Our results suggest that, somewhat surprisingly, the right formula to define the Rényi α-relative entropies for quantum states depends on whether the parameter α is below or above 1; it seems that for $\alpha < 1$, one should use the old Rényi relative entropies, while for $\alpha > 1$, the new Rényi relative entropies are the right choice. Hence, we suggest to define the Rényi relative entropies for quantum states (more generally, for positive operators) ρ, σ as

$$D_\alpha(\rho\,\|\,\sigma) := \begin{cases} \frac{1}{\alpha-1}\log\operatorname{Tr}\rho^\alpha\sigma^{1-\alpha} - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho, & \alpha\in[0,1),\\ \frac{1}{\alpha-1}\log\operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha - \frac{1}{\alpha-1}\log\operatorname{Tr}\rho, & \alpha>1 \text{ and } \operatorname{supp}\rho\subseteq\operatorname{supp}\sigma,\\ +\infty, & \text{otherwise.}\end{cases}$$

2. Preliminaries

For a finite-dimensional Hilbert space $\mathcal{H}$, let $\mathcal{L}(\mathcal{H})$ denote the set of linear operators on $\mathcal{H}$, let $\mathcal{L}(\mathcal{H})_+$ denote the set of positive semidefinite operators, and $\mathcal{S}(\mathcal{H})$ be the set of density operators (states) on $\mathcal{H}$ (i.e., positive semidefinite operators with trace 1). A finite-valued POVM (positive operator valued measure) on $\mathcal{H}$ is a map $M:\mathcal{I}\to\mathcal{L}(\mathcal{H})$, where $\mathcal{I}$ is some finite set, $0\le M_i$, $i\in\mathcal{I}$, and $\sum_{i\in\mathcal{I}} M_i = I$. We denote the set of POVMs on $\mathcal{H}$ by $\mathcal{M}(\mathcal{H})$.

Any Hermitian operator $A\in\mathcal{L}(\mathcal{H})$ admits a spectral decomposition $A=\sum_i a_i P_i$, where $a_i\in\mathbb{R}$ and the P_i are orthogonal projections. We introduce the notation $\{A>0\} := \sum_{i:\,a_i>0} P_i$ for the spectral projection of A corresponding to the positive half-line $(0,+\infty)$. The spectral projections $\{A\ge 0\}$, $\{A<0\}$ and $\{A\le 0\}$ are defined similarly. The positive part of A is defined as

$$A_+ := A\{A>0\}, \tag{9}$$

and it is easy to see that

$$\operatorname{Tr} A_+ = \operatorname{Tr} A\{A>0\} = \max_{0\le T\le I}\operatorname{Tr} AT \ge 0. \tag{10}$$

In particular, if ρ_n and σ_n are self-adjoint operators then for any $a\in\mathbb{R}$ the application of (10) to $A=\rho_n - e^{na}\sigma_n$ yields

$$\operatorname{Tr}\rho_n\{\rho_n - e^{na}\sigma_n>0\} \ge e^{na}\operatorname{Tr}\sigma_n\{\rho_n - e^{na}\sigma_n>0\}. \tag{11}$$

If $\mathcal{F}$ is a positive trace-preserving map then

$$\operatorname{Tr}\mathcal{F}(A)_+ = \max_{0\le T\le I}\operatorname{Tr}\mathcal{F}(A)T = \max_{0\le T\le I}\operatorname{Tr} A\mathcal{F}^*(T) \le \max_{0\le S\le I}\operatorname{Tr} AS = \operatorname{Tr} A_+.$$

In particular, we have the following lemma.

Lemma 2.1. *Let ρ_n and σ_n be self-adjoint operators and $\mathcal{F}$ be a positive trace-preserving map. Then for any $a \in \mathbb{R}$,*

$$\operatorname{Tr}(\rho_n - e^{na}\sigma_n)_+ \geq \operatorname{Tr}(\mathcal{F}(\rho_n) - e^{na}\mathcal{F}(\sigma_n))_+. \tag{12}$$

Let A be a Hermitian operator on $\mathcal{H}$ with spectral decomposition $A = \sum_i a_i E_i$. The *pinching operation* $\mathcal{E}_A$ corresponding to A is defined as

$$\mathcal{E}_A(B) := \sum_i E_i B E_i, \quad B \in \mathcal{L}(\mathcal{H}). \tag{13}$$

It is also denoted by $\mathcal{E}_E(B)$ in terms of the PVM (projection-valued measure) $E = \{E_i\}_i$. Note that $\mathcal{E}_A(B)$ is the unique operator in the commutant $\{A\}'$ of $\{A\}$ satisfying

$$\forall C \in \{A\}', \quad \operatorname{Tr} BC = \operatorname{Tr} \mathcal{E}_A(B)C. \tag{14}$$

The following lemma is from [16,17]:

Lemma 2.2 (Pinching inequality). *Let A be self-adjoint and B be a positive semidefinite operator on $\mathcal{H}$. Then*

$$B \leq v(A)\mathcal{E}_A(B),$$

where $v(A)$ denotes the number of different eigenvalues of A.

All through the paper, ρ and σ will denote positive semidefinite operators on some finite-dimensional Hilbert space $\mathcal{H}$, and we use the notation

$$\rho_n := \rho^{\otimes n}, \quad \sigma_n := \sigma^{\otimes n}, \quad \widehat{\rho}_n := \mathcal{E}_{\sigma_n}(\rho_n), \quad v_n := v(\sigma_n), \tag{15}$$

where $\mathcal{E}_{\sigma_n}$ is the pinching operation corresponding to σ_n, and v_n denotes the number of different eigenvalues of σ_n. Note that $v_n \leq (n+1)^{\dim \mathcal{H}}$, and Lemma 2.2 yields

$$\rho_n \leq v_n \widehat{\rho}_n \leq (n+1)^{\dim \mathcal{H}} \widehat{\rho}_n. \tag{16}$$

The power of the pinching inequality for asymptotic analysis comes from the fact that

$$\lim_{n\to+\infty} \frac{1}{n} \log v_n = 0,$$

which we will use repeatedly and without further explanation in the paper.

We will use the convention that powers of a positive semidefinite operator are only taken on its support and defined to be 0 on the orthocomplement of its support. That is, if $a_1, \ldots, a_r$ are the eigenvalues of $A \geq 0$, with corresponding eigenprojections $P_1, \ldots, P_r$, then $A^p := \sum_{i:\, a_i > 0} a_i^p P_i$ for any $p \in \mathbb{R}$. In particular, A^0 is the projection onto the support of A.

We will also use the convention $\log 0 := -\infty$.

3. Properties of the New Rényi Relative Entropies

For positive semidefinite operators ρ and σ, and $\alpha \in \mathbb{R}$, let

$$F_\alpha(\rho\|\sigma) := \log \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha. \tag{17}$$

For a POVM $M = \{M_x\}_x$, we can consider the corresponding classical quantity as

$$F_\alpha^M(\rho\|\sigma) := \log\left(\sum_x \{\operatorname{Tr}\rho M_x\}^\alpha \{\operatorname{Tr}\sigma M_x\}^{1-\alpha}\right). \tag{18}$$

Note that for states ρ and σ such that $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, $\frac{1}{\alpha-1}F_\alpha(\rho\|\sigma)$ is the new Rényi α-relative entropy defined in (6), and $\frac{1}{\alpha-1}F_\alpha^M(\rho\|\sigma)$ is the post-measurement Rényi α-relative entropy.

In this section we show that for every $\alpha > 1$, the new Rényi α-relative entropies are asymptotically attainable by measurements in the limit of infinitely many copies of ρ and σ; for this we only use that the new Rényi α-relative entropies are monotonic under pinching by the reference state, which is very simple to show. From this we derive a new simple proof for the monotonicity of the new Rényi α-relative entropies.

Monotonicity in the classical case is well-known and easy to prove; we state it explicitly here for completeness:

Lemma 3.1 (Classical monotonicity). *Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})_+$ be commuting operators such that* $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$*, and let $\mathcal{F} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a positive trace-preserving map such that $\mathcal{F}(\rho)$ commutes with $\mathcal{F}(\sigma)$. For every $\alpha > 1$, $F_\alpha(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) \le F_\alpha(\rho\|\sigma)$.*

Proof. The proof is an elementary argument based on the convexity of the function $x \mapsto x^\alpha$ on $[0, +\infty)$ for $\alpha > 1$; details can bee found e.g. in [25, Proposition A.3]. □

The following has been shown in [34, Proposition 13]. We reproduce the proof here for readers' convenience.

Lemma 3.2 (Monotonicity under pinching). *Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ and $\alpha \ge 1$. Then*

$$F_\alpha(\mathcal{E}_\sigma(\rho)\|\sigma) \le F_\alpha(\rho\|\sigma). \tag{19}$$

Proof. It is easy to see that $\sigma^{\frac{1-\alpha}{2\alpha}}\mathcal{E}_\sigma(\rho)\sigma^{\frac{1-\alpha}{2\alpha}} = \mathcal{E}_\sigma\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)$, and Problem II.5.5 with Theorem II.3.1 in [6], applied to the convex function $f(t) = t^\alpha$, yields the assertion. □

Using the above two lemmas, we can prove monotonicity under measurements.

Lemma 3.3 (Monotonicity under measurements). *Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ be such that* $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$. *For any POVM $M = \{M_x\}_x \in \mathcal{M}(\mathcal{H})$, we have*

$$F_\alpha^M(\rho\|\sigma) \le F_\alpha(\rho\|\sigma), \quad \alpha \ge 1. \tag{20}$$

Proof. For any POVM $M_n = \{M_n(x)\}_x$ on $\mathcal{H}^{\otimes n}$ and any $\alpha \geq 1$,

$$\sum_x (\operatorname{Tr} \rho_n M_n(x))^\alpha (\operatorname{Tr} \sigma_n M_n(x))^{1-\alpha} \leq v_n^\alpha \sum_x (\operatorname{Tr} \widehat{\rho}_n M_n(x))^\alpha (\operatorname{Tr} \sigma_n M_n(x))^{1-\alpha} \tag{21}$$

$$\leq v_n^\alpha \operatorname{Tr} \widehat{\rho}_n^\alpha \sigma_n^{1-\alpha} \tag{22}$$

$$\leq v_n^\alpha \operatorname{Tr} \left(\sigma_n^{\frac{1-\alpha}{2\alpha}} \rho_n \sigma_n^{\frac{1-\alpha}{2\alpha}} \right)^\alpha, \tag{23}$$

where the first inequality is due to (16), the second inequality follows from Lemma 3.1, and the third one from Lemma 3.2.

Now let $M = \{M_x\}_{x \in \mathcal{X}} \in \mathcal{M}(\mathcal{H})$ be a POVM on a single copy, and M_n be its nth i.i.d. extension, i.e.,

$$M_n(\underline{x}) := M_{x_1} \otimes \cdots \otimes M_{x_n}, \quad \underline{x} \in \mathcal{X}^n. \tag{24}$$

Then we obtain

$$\left(\sum_x (\operatorname{Tr} \rho M_x)^\alpha (\operatorname{Tr} \sigma M_x)^{1-\alpha} \right)^n = \sum_{\underline{x}} \left(\operatorname{Tr} \rho_n M_n(\underline{x})\right)^\alpha \left(\operatorname{Tr} \sigma_n M_n(\underline{x})\right)^{1-\alpha}$$

$$\leq v_n^\alpha \operatorname{Tr} \left(\sigma_n^{\frac{1-\alpha}{2\alpha}} \rho_n \sigma_n^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$$

$$= v_n^\alpha \left(\operatorname{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right)^n. \tag{25}$$

Taking the logarithm and dividing by n yields

$$F_\alpha^M(\rho\|\sigma) \leq F_\alpha(\rho\|\sigma) + \frac{\alpha}{n} \log v_n, \tag{26}$$

which proves the lemma by taking the limit $n \to \infty$. □

Remark 3.4. The technique used in the proof of the above lemma is essentially due to [17] (see around page 88), where the inequalities (21) and (22) have been shown.

Remark 3.5. Note that the assumption $\operatorname{supp} \rho \subseteq \operatorname{supp} \sigma$ was necessary to apply classical monotonicity in (22). In fact, the statement of Lemma 3.3 need not hold without this assumption. Indeed, in the extreme case where ρ and σ have orthogonal supports, we have $F_\alpha(\rho\|\sigma) = -\infty$, and the trivial POVM $M = \{I\}$ yields $F_\alpha^M(\rho\|\sigma) = \log(\operatorname{Tr} \rho)^\alpha (\operatorname{Tr} \sigma)^{1-\alpha}$, which is a finite number unless ρ or σ is equal to 0.

The following lemma is standard:

Lemma 3.6. *Let A and B be Hermitian operators on $\mathcal{H}$ with their spectrum in some interval I, and let $f : I \to \mathbb{R}$ be a monotone increasing function. If $A \leq B$ then* $\operatorname{Tr} f(A) \leq \operatorname{Tr} f(B)$. *In particular,*

$$0 \leq A \leq B \Longrightarrow \operatorname{Tr} A^\alpha \leq \operatorname{Tr} B^\alpha \quad \alpha > 0.$$

Proof. Let $\{\lambda_i^\downarrow(A)\}_{i=1}^{\dim \mathcal{H}}$ denote the sequence of decreasingly ordered eigenvalues of A. By the Courant–Fischer–Weyl minimax principle [6, Corollary III.1.2], $\lambda_i^\downarrow(A) \leq \lambda_i^\downarrow(B)$, $1 \leq i \leq \dim \mathcal{H}$, from which the assertion follows. □

Theorem 3.7 (Asymptotic attainability). *Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ be such that* $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$*. For any $\alpha \geq 1$, we have*

$$F_\alpha(\rho\|\sigma) = \lim_{n\to\infty} \frac{1}{n} F_\alpha(\widehat{\rho}_n\|\sigma_n) \\ = \lim_{n\to\infty} \frac{1}{n} \max_{M_n\in\mathcal{M}(\mathcal{H}^{\otimes n})} F_\alpha^{M_n}(\rho_n\|\sigma_n), \tag{27}$$

where the maximization in the second line is over all POVMs on $\mathcal{H}^{\otimes n}$.

Proof. Since σ_n and $\widehat{\rho}_n$ commute, they have a common eigenbasis $\{e_n(i)\}_{i=1}^{d_n}$, $d_n = (\dim\mathcal{H})^n$. Let $E_n = \{E_n(i) = |e_n(i)\rangle\langle e_n(i)|\}_{i=1}^{d_n}$ be the corresponding projection-valued measure. Then

$$\frac{1}{n}F_\alpha(\widehat{\rho}_n\|\sigma_n) = \frac{1}{n}F_\alpha^{E_n}(\rho_n\|\sigma_n) \leq \frac{1}{n}\max_{M_n} F_\alpha^{M_n}(\rho_n\|\sigma_n) \leq \frac{1}{n}F_\alpha(\rho_n\|\sigma_n) = F_\alpha(\rho\|\sigma), \tag{28}$$

where the last inequality is due to Lemma 3.3. By Lemma 2.2,

$$0 \leq \sigma_n^{\frac{1-\alpha}{2\alpha}}\rho_n\sigma_n^{\frac{1-\alpha}{2\alpha}} \leq v_n\sigma_n^{\frac{1-\alpha}{2\alpha}}\widehat{\rho}_n\sigma_n^{\frac{1-\alpha}{2\alpha}} = v_n\sum_{i=1}^{d_n}(\operatorname{Tr}\rho_n E_n(i))\,(\operatorname{Tr}\sigma_n E_n(i))^{\frac{1-\alpha}{\alpha}}\,E_n(i), \tag{29}$$

and Lemma 3.6 yields

$$\operatorname{Tr}\left(\sigma_n^{\frac{1-\alpha}{2\alpha}}\rho_n\sigma_n^{\frac{1-\alpha}{2\alpha}}\right)^\alpha \leq v_n^\alpha \operatorname{Tr}\left(\sigma_n^{\frac{1-\alpha}{2\alpha}}\widehat{\rho}_n\sigma_n^{\frac{1-\alpha}{2\alpha}}\right)^\alpha = v_n^\alpha\sum_{i=1}^{d_n}(\operatorname{Tr}\rho_n E_n(i))^\alpha\,(\operatorname{Tr}\sigma_n E_n(i))^{1-\alpha}. \tag{30}$$

Taking the logarithm, we obtain

$$F_\alpha(\rho\|\sigma) \leq \frac{1}{n}F_\alpha(\widehat{\rho}_n\|\sigma_n) + \frac{\alpha}{n}\log v_n = \frac{1}{n}F_\alpha^{E_n}(\rho_n\|\sigma_n) + \frac{\alpha}{n}\log v_n \\ \leq \frac{1}{n}\max_{M_n} F_\alpha^{M_n}(\rho_n\|\sigma_n) + \frac{\alpha}{n}\log v_n. \tag{31}$$

Combining this with (28), and taking the limit $n \to +\infty$, the assertion follows. □

Theorem 3.7 implies the asymptotic attainability for the Rényi relative entropies:

Corollary 3.8. *For any $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ and $\alpha > 1$, we have*

$$D_\alpha^{(\mathrm{new})}(\rho\|\sigma) = \lim_{n\to\infty}\frac{1}{n}D_\alpha^{(\mathrm{new})}(\widehat{\rho}_n\|\sigma_n) \\ = \lim_{n\to\infty}\frac{1}{n}\max_{M_n\in\mathcal{M}(\mathcal{H}^{\otimes n})} D_\alpha^{(\mathrm{new})}\left(\{\operatorname{Tr}\rho_n M_n(x)\}_{x\in\mathcal{X}}\|\{\operatorname{Tr}\sigma_n M_n(x)\}_{x\in\mathcal{X}}\right), \tag{32}$$

where the maximization in the second line is over all POVMs on $\mathcal{H}^{\otimes n}$.

Proof. The case where $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$ is immediate from Theorem 3.7. On the other hand, if $\operatorname{supp}\rho \nsubseteq \operatorname{supp}\sigma$ then also $\operatorname{supp}\widehat{\rho}_n \nsubseteq \operatorname{supp}\sigma_n$, and hence, by the definition (6), $D_\alpha^{(\text{new})}(\rho\|\sigma) = D_\alpha^{(\text{new})}(\widehat{\rho}_n\|\sigma_n) = \max_{M_n\in\mathcal{M}(\mathcal{H}^{\otimes n})} D_\alpha^{(\text{new})}(\{\operatorname{Tr}\rho_n M_n(x)\}_{x\in\mathcal{X}} \|\{\operatorname{Tr}\sigma_n M_n(x)\}_{x\in\mathcal{X}}) = +\infty$ for every $n \in \mathbb{N}$, making the assertion trivial. □

Remark 3.9. The same statement for the relative entropy has been shown in [22].

Remark 3.10. The maximum over all measurements in (32) can be replaced by a concrete binary POVM given by a Neyman–Pearson test; see Corollary 4.6.

Theorem 3.7 has a number of important further corollaries:

Corollary 3.11 (Convexity). *For any fixed* $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ *such that* $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, $F_\alpha(\rho\|\sigma)$ *is a convex function of* α *for* $\alpha \geq 1$.

Proof. It is easy to see (by computing its second derivative) that $F_\alpha(\widehat{\rho}_n\|\sigma_n)$ is a convex function of α. Thus by Theorem 3.7, $F_\alpha(\rho\|\sigma)$ is a pointwise limit of convex functions, and hence it is convex. □

Corollary 3.12. *For any fixed* $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$, *the function* $\alpha \mapsto D_\alpha^{(\text{new})}(\rho \,\|\, \sigma)$ *is monotone increasing for* $\alpha > 1$.

Proof. We can assume that $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, since otherwise $D_\alpha^{(\text{new})}(\rho \,\|\, \sigma) = +\infty$ for every $\alpha > 1$, and the assertion holds trivially. Note that $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$ implies that $F_1(\rho\|\sigma) = \log\operatorname{Tr}\rho$, and hence $D_\alpha^{(\text{new})}(\rho \,\|\, \sigma) = \frac{F_\alpha(\rho\|\sigma)-F_1(\rho\|\sigma)}{\alpha-1}$. The assertion then follows from Corollary 3.11. □

Corollary 3.13 (Monotonicity). *Let* $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ *be such that* $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, *and let* $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ *be a CPTP map. Then*

$$F_\alpha(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) \leq F_\alpha(\rho\|\sigma), \quad \alpha > 1.$$

Proof. By complete positivity, $\mathcal{F}_n := \mathcal{F}^{\otimes n}$ is positive for every $n \in \mathbb{N}$. Let $\mathcal{F}_n^* : \mathcal{L}(\mathcal{K}^{\otimes n}) \to \mathcal{L}(\mathcal{H}^{\otimes n})$ be the dual (adjoint) of $\mathcal{F}_n$, defined by

$$\forall \omega \in \mathcal{S}(\mathcal{H}^{\otimes n}), \quad \forall A \in \mathcal{L}(\mathcal{K}^{\otimes n}), \quad \operatorname{Tr}\mathcal{F}_n(\omega)A = \operatorname{Tr}\omega\mathcal{F}_n^*(A). \tag{33}$$

Then $\mathcal{F}_n^*$ is a unital positive map. Thus, if $\{M(x)\}_{x\in\mathcal{X}} \in \mathcal{M}(\mathcal{K}^{\otimes n})$ is a POVM on $\mathcal{K}^{\otimes n}$ then $\mathcal{F}_n^*(M) := \{\mathcal{F}_n^*(M(x))\}_{x\in\mathcal{X}}$ is a POVM on $\mathcal{H}^{\otimes n}$. Hence,

$$\begin{aligned} \max_{M\in\mathcal{M}(\mathcal{K}^{\otimes n})} F_\alpha^M(\mathcal{F}_n(\rho_n)\|\mathcal{F}_n(\sigma_n)) &= \max_{M\in\mathcal{M}(\mathcal{K}^{\otimes n})} F_\alpha^{\mathcal{F}_n^*(M)}(\rho_n\|\sigma_n) \\ &\leq \max_{M\in\mathcal{M}(\mathcal{H}^{\otimes n})} F_\alpha^M(\rho_n\|\sigma_n) \end{aligned} \tag{34}$$

for any n. Now (34) and Theorem 3.7 yield the assertion. □

Corollary 3.13 immediately implies the following:

Corollary 3.14. *The new Rényi relative entropies are monotone under CPTP maps for* $\alpha > 1$. *That is, if* $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ *and* $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ *is a CPTP map then*

$$D_\alpha(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) \leq D_\alpha(\rho\|\sigma), \quad \alpha > 1, \tag{35}$$

and the limit $\alpha \searrow 1$ *yields the same monotonicity property for the relative entropy.*

For $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$, let

$$Q_\alpha^{(\mathrm{new})}(\rho\|\sigma) := \mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha, \quad \alpha \in \mathbb{R}_+.$$

This is an analogy of the quasi-entropy [40] (or quantum f-divergence [25]) corresponding to the function $x \mapsto x^\alpha$. However, $Q_\alpha^{(\mathrm{new})}$ cannot be written in the form of an f-divergence [25, Corollary 2.10]. Corollary 3.13 is equivalent to the monotonicity of Q:

Corollary 3.15 (Monotonicity of Q). *Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ be such that* $\mathrm{supp}\,\rho \subseteq \mathrm{supp}\,\sigma$*, and let $\mathcal{F}\colon \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ be a CPTP map. Then*

$$Q_\alpha^{(\mathrm{new})}(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) \leq Q_\alpha^{(\mathrm{new})}(\rho\|\sigma), \quad \alpha > 1.$$

Following the argument of [40], we immediately obtain the joint convexity of Q:

Corollary 3.16 (Joint convexity). *Let $\rho_i, \sigma_i \in \mathcal{L}(\mathcal{H})_+$ be such that* $\mathrm{supp}\,\rho_i \subseteq \mathrm{supp}\,\sigma_i$*, $i = 1, \ldots, r$, and let $p_1, \ldots, p_r$ be a probability distribution. Then*

$$Q_\alpha^{(\mathrm{new})}\left(\sum_{i=1}^r p_i\rho_i \,\middle\|\, \sum_{i=1}^r p_i\sigma_i\right) \leq \sum_{i=1}^r p_i Q_\alpha^{(\mathrm{new})}(\rho_i\|\sigma_i).$$

Proof. Let $\delta_1, \ldots, \delta_r$ be orthogonal rank 1 projections on $\mathcal{K} := \mathbb{C}^r$, and define $\rho := \sum_{i=1}^r p_i\delta_i \otimes \rho_i$, $\sigma := \sum_{i=1}^r p_i\delta_i \otimes \sigma_i$. Taking $\mathcal{F} := \mathrm{Tr}_{\mathcal{K}}$ to be the partial trace over $\mathcal{K}$ in Corollary 3.15, the assertion follows. □

Remark 3.17. In Corollary 3.16, we obtained the joint convexity from the monotonicity of $Q_\alpha^{(\mathrm{new})}$. In [13] (and also in [34,50] for $\alpha \in (1, 2]$) the authors followed the opposite approach: they first established joint convexity of $Q_\alpha^{(\mathrm{new})}$, and from that they obtained its monotonicity under CPTP maps by a standard argument using the Stinespring representation and decomposing the trace as a convex combination of unitary conjugations.

Remark 3.18. Note that the monotonicity properties in Corollaries 3.13, 3.14 and 3.15 hold for any trace-preserving linear map $\mathcal{F}$ such that $\mathcal{F}^{\otimes n}$ is positive for every $n \in \mathbb{N}$. This is a weaker condition than complete positivity.

We give an overview of the various monotonicity and attainability properties of the old and the new Rényi relative entropies in Appendix A.

4. Strong Converse Exponent in Quantum Hypothesis Testing

4.1. Simple quantum hypothesis testing. We study the simple hypothesis testing problem for the null hypothesis H_0: ρ_n versus the alternative hypothesis H_1: σ_n, where $\rho_n = \rho^{\otimes n}$ and $\sigma_n = \sigma^{\otimes n}$ are the n-fold tensor products of arbitrarily given density operators ρ and σ in $\mathcal{S}(\mathcal{H})$. The problem is to decide which hypothesis is true based on the outcome drawn from a quantum measurement, which is described by a POVM on $\mathcal{H}_n = \mathcal{H}^{\otimes n}$. In the hypothesis testing problem, it is sufficient to treat a two-valued POVM $\{T_n(0), T_n(1)\} \in \mathcal{M}(\mathcal{H}^{\otimes n})$, where 0 and 1 indicate the acceptance of H_0 and H_1, respectively. Since $T_n(1) = I - T_n(0)$, the POVM is uniquely determined by $T_n = T_n(0)$, and the only constraint on T_n is that $0 \leq T_n \leq I_n$. We will call such operators tests. For a test T_n, the error probabilities of the first and the second kind are, respectively, defined by

$$\alpha_n(T_n) := \operatorname{Tr} \rho_n(I_n - T_n), \tag{36}$$

$$\beta_n(T_n) := \operatorname{Tr} \sigma_n T_n. \tag{37}$$

In general there is a trade-off between these error probabilities, and we can not make these probabilities unconditionally small, as described below. First, we consider the optimal value for $\beta_n(T_n)$ under the constant constraint on $\alpha_n(T_n)$, that is,

$$\beta_n^*(\epsilon) := \min\{\beta_n(T_n) | T_n\text{: test}, \quad \alpha_n(T_n) \leq \epsilon\}. \tag{38}$$

The quantum Stein's lemma [22,39] states that for all $\varepsilon \in (0, 1)$,

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n^*(\epsilon) = -D(\rho\|\sigma), \tag{39}$$

where $D(\rho\|\sigma)$ is the quantum relative entropy given in (4).

For the study of the trade-off between the error probabilities, it is natural to ask what happens if we require the type II error probabilities to vanish with an exponent below or above the relative entropy, i.e., we want to study the asymptotic behavior of $\alpha_n(T_n)$ under the exponential constraint $\beta_n(T_n) \leq e^{-nr}$, $r > 0$. Specifically, let us define

$$\begin{aligned} B_e(r) &:= \sup\left\{ -\limsup_{n\to\infty} \frac{1}{n} \log \alpha_n(T_n) \,\middle|\, \limsup_{n\to\infty} \frac{1}{n} \log \beta_n(T_n) \leq -r \right\} \\ &= \sup\Big\{ R \,\Big|\, \exists\{T_n\}_{n=1}^{\infty},\ 0 \leq T_n \leq I_n,\ \text{s.t.} \\ &\qquad \limsup_{n\to\infty} \frac{1}{n} \log \beta_n(T_n) \leq -r,\ \limsup_{n\to\infty} \frac{1}{n} \log \alpha_n(T_n) \leq -R \Big\}, \end{aligned} \tag{40}$$

where the supremum in the first line is taken over all sequences of tests $\{T_n\}_{n\in\mathbb{N}}$ satisfying the condition. It was shown in [18,36] that

$$B_e(r) = \sup_{0\leq s<1} \frac{-sr - \log \operatorname{Tr} \rho^{1-s}\sigma^s}{1-s} = \sup_{0<\alpha<1} \frac{\alpha-1}{\alpha} \left[r - D_\alpha^{(\text{old})}(\rho \,\|\, \sigma) \right] = H_r(\rho\|\sigma), \tag{41}$$

where $D_\alpha^{(\text{old})}$ is the traditional definition of the quantum Rényi relative entropy, given in (3), and $H_r(\rho\|\sigma)$ is the Hoeffding divergence defined in (5). [Note that the roles of the type I and the type II errors are reversed here as compared to some previous work on the Hoeffding bound, and hence our $H_r(\rho\|\sigma)$ corresponds to $H_r(\sigma\|\rho)$ in those works.] It can be shown that $B_e(r) > 0$ when $0 < r < D(\rho\|\sigma)$, and $\alpha_n(T_n)$ goes to zero exponentially with the rate $B_e(r)$ for an optimal sequence of tests $\{T_n\}_{n=1}^{\infty}$.

On the other hand, if $\operatorname{supp} \rho \subseteq \operatorname{supp} \sigma$ and $\beta_n(T_n) \leq e^{-nr}$ with $r > D(\rho\|\sigma)$ then $\alpha_n(T_n)$ inevitably goes to 1 exponentially fast [39]; this is called the strong converse property. In this case, we are interested in determining the exponent with which the success probabilities $1 - \alpha_n(T_n) = \operatorname{Tr} \rho_n T_n$ go to zero. The optimal such exponent is the strong converse exponent $B_e^*(r)$; formally,

$$B_e^*(r) := \inf\left\{-\liminf_{n\to+\infty}\frac{1}{n}\log\operatorname{Tr}\rho_n T_n \;\middle|\; \limsup_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n T_n \le -r\right\}, \tag{42}$$

where the infimum is taken over all possible sequences of tests $\{T_n\}_{n\in\mathbb{N}}$ satisfying the condition. Note that one's aim is to make the success probabilities decay as slow as possible, and hence optimality means taking the smallest possible exponent along all sequences of tests with a fixed decay rate of the type II errors. It is easy to see that $B_e^*(r)$ can be alternatively written as

$$\begin{aligned}B_e^*(r) = \sup\Big\{R \;\Big|\; &\forall\{T_n\}_{n=1}^{\infty},\ 0\le T_n\le I_n,\\ &\limsup_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n T_n \le -r \;\Rightarrow\; \liminf_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\rho_n T_n \le -R\Big\}\\ = \inf\Big\{R \;\Big|\; &\exists\{T_n\}_{n=1}^{\infty},\ 0\le T_n\le I_n,\\ &\limsup_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n T_n \le -r,\ \liminf_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\rho_n T_n \ge -R\Big\}.\end{aligned} \tag{43}$$

The main result of Sect. 4 is Theorem 4.10, where we show that, in complete analogy with (41),

$$B_e^*(r) = \sup_{1<\alpha}\frac{\alpha-1}{\alpha}\left[r - D_\alpha^{(\text{new})}(\rho\,\|\,\sigma)\right] = H_r^*(\rho\|\sigma), \tag{44}$$

where $H_r^*(\rho\|\sigma)$ is the converse Hoeffding divergence (8). The inequality $B_e^*(r) \ge H_r^*(\rho\|\sigma)$ follows easily from the monotonicity of the Rényi divergences, as we show in Lemma 4.7. We show that this is in fact an equality by determining the asymptotics of the error probabilities for the Neyman–Pearson tests. This is interesting in itself, as these quantities play a central role in the information spectrum method [15,37]. We start with this problem in Sect. 4.2.

Remark 4.1. Note that if $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$ is not satisfied then the strong converse property doesn't hold; indeed, the choice $T_n := I - \sigma_n^0$, $n\in\mathbb{N}$, yields a sequence of tests for which $\beta_n(T_n) = 0 \le e^{-nr}$, $r>0$, and $\alpha_n(T_n) = (\operatorname{Tr}\rho\sigma^0)^n$, $n\in\mathbb{N}$, which converges to zero exponentially fast with an exponent $-\log\operatorname{Tr}\rho\sigma^0 > 0$. Hence, for the rest we will assume that $\operatorname{supp}\rho\subseteq\operatorname{supp}\sigma$.

4.2. Exponents for the Neyman–Pearson tests. Let ρ and σ be quantum states such that

$$\operatorname{supp}\rho\subseteq\operatorname{supp}\sigma, \tag{45}$$

and let ρ_n, σ_n, etc. be defined as in (15). To exclude a trivial case, we assume that $\rho\ne\sigma$. Let us define the quantum Neyman–Pearson tests by

$$S_n(a) := \{\rho_n - e^{na}\sigma_n > 0\}, \tag{46}$$

where $a\in\mathbb{R}$ is a trade-off parameter. Our goal in this section is to determine the asymptotics of the corresponding type I success probabilities $\operatorname{Tr}\rho_n S_{n,a}$ and the type II error probabilities $\operatorname{Tr}\sigma_n S_{n,a}$. Note that

$$S_n(a) = 0 \iff a \ge D_{\max}(\rho\,\|\,\sigma) := \inf\{\gamma\colon \rho\le e^{\gamma}\sigma\}. \tag{47}$$

Here $D_{\max}(\rho \| \sigma)$ is the *max-relative entropy* [11,44], and it was shown in [34, Theorem 4] that

$$D_{+\infty}^{(\text{new})}(\rho\|\sigma) := \lim_{\alpha\to+\infty} D_{\alpha}^{(\text{new})}(\rho \| \sigma) = D_{\max}(\rho \| \sigma).$$

Thus,

$$\operatorname{Tr} \rho_n S_{n,a} = \operatorname{Tr} \sigma_n S_{n,a} = 0, \quad a \geq D_{\max}(\rho \| \sigma),$$

and, with the convention $\log 0 := -\infty$,

$$\lim_{n\to+\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n S_{n,a} = \lim_{n\to+\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n S_{n,a} = -\infty, \quad a \geq D_{\max}(\rho \| \sigma).$$

Hence, for the rest we can restrict our attention to $a < D_{\max}(\rho \| \sigma)$.

For every $s \in \mathbb{R}$, let

$$\psi(s) := F_{s+1}(\rho\|\sigma) = \log \operatorname{Tr} \left(\sigma^{\frac{-s}{2(s+1)}} \rho \sigma^{\frac{-s}{2(s+1)}}\right)^{s+1}, \tag{48}$$

and

$$\phi(a) := \sup_{s\geq 0}\{as - \psi(s)\} \tag{49}$$

be its Legendre–Fenchel transform on the interval $[0, +\infty)$.

Lemma 4.2. *We have*

$$\psi(0) = 0, \tag{50}$$

$$\psi'(0) = D(\rho \| \sigma), \tag{51}$$

$$\lim_{s\to+\infty} \psi'(s) = D_{\max}(\rho \| \sigma), \tag{52}$$

and

$$\phi(a) \begin{cases} = 0, & a \leq D(\rho \| \sigma) \\ > 0, & D(\rho \| \sigma) < a \leq D_{\max}(\rho \| \sigma), \\ = +\infty, & D_{\max}(\rho \| \sigma) < a. \end{cases} \tag{53}$$

Proof. The identity in (50) is immediate from the definition of ψ. $\psi(0) = 0$ yields $\psi'(0) = \lim_{s\to 0} \frac{1}{s}\psi(s) = \lim_{\alpha\to 1} D_\alpha(\rho \| \sigma) = D(\rho \| \sigma)$, where the last identity is due to [34, Theorem 4]. Using again [34, Theorem 4] and the L'Hospital rule, $\lim_{s\to+\infty} \psi'(s) = \lim_{s\to+\infty} \frac{1}{s}\psi(s) = \lim_{\alpha\to+\infty} D_\alpha(\rho \| \sigma) = D_{\max}(\rho \| \sigma)$. By Corollary 3.11, $s \mapsto \psi(s)$ is convex, and hence (53) follows immediately from (50)–(52). □

Lemma 4.3. *For any $a \in \mathbb{R}$ and $n \in \mathbb{N}$, we have*

$$\frac{1}{n} \log \operatorname{Tr} \rho_n S_n(a) \leq -\phi(a), \tag{54}$$

$$\frac{1}{n} \log \operatorname{Tr} \sigma_n S_n(a) \leq -\{a + \phi(a)\}. \tag{55}$$

Proof. For any $a \in \mathbb{R}$ and $s \geq 0$, we have

$$\begin{aligned}
\operatorname{Tr} \rho_n S_n(a) &= \{\operatorname{Tr} \rho_n S_n(a)\}^{s+1} \{\operatorname{Tr} \rho_n S_n(a)\}^{-s} \\
&\leq e^{-nas} \{\operatorname{Tr} \rho_n S_n(a)\}^{s+1} \{\operatorname{Tr} \sigma_n S_n(a)\}^{-s} \\
&\leq e^{-nas} \Big[\{\operatorname{Tr} \rho_n S_n(a)\}^{s+1} \{\operatorname{Tr} \sigma_n S_n(a)\}^{-s} \\
&\quad + \{\operatorname{Tr} \rho_n (I_n - S_n(a))\}^{s+1} \{\operatorname{Tr} \sigma_n (I_n - S_n(a))\}^{-s} \Big] \\
&\leq e^{-nas} \operatorname{Tr} \left(\sigma_n^{\frac{-s}{2(s+1)}} \rho_n \sigma_n^{\frac{-s}{2(s+1)}} \right)^{s+1} \\
&= e^{-nas} e^{n\psi(s)}, \qquad (56)
\end{aligned}$$

where in the first inequality we used (11), the second inequality is trivial, and the last inequality follows from Lemma 3.3. Taking the logarithm and the infimum in s yields the inequality in (54).

Using (11) and (56), we get

$$\operatorname{Tr} \sigma_n S_n(a) \leq e^{-na} \operatorname{Tr} \rho_n S_n(a) \leq e^{-na(s+1)} e^{n\psi(s)}, \qquad (57)$$

which yields (55). □

Note that the bounds in (54) and (55) are trivial for $a \geq D_{\max}(\rho \,\|\, \sigma)$, due to (47). For $a \leq D(\rho \,\|\, \sigma)$ we have $\phi(a) = 0$ [cf. (53)], and hence the upper bound in (54) is trivial in this range. More detailed information about the values of $\operatorname{Tr} \sigma_n S_n(a)$ in this range is given in the setting of the Hoeffding bound; Corollary 4.5 in [24] states that

$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n S_n(a) = - \sup_{0\leq t\leq 1} \{at - \log \operatorname{Tr} \rho^t \sigma^{1-t}\} \leq -a = -\{\phi(a) + a\}$$

for every $a < D(\rho \,\|\, \sigma)$. Theorems 4.4 and 4.5 below show that the inequalities in (54) and (55) hold asymptotically as equalities in the non-trivial range $D(\rho \,\|\, \sigma) < a < D_{\max}(\rho \,\|\, \sigma)$.

Theorem 4.4. *For any $a \in (D(\rho\|\sigma), D_{\max}(\rho \,\|\, \sigma))$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n S_n(a) = \lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr}(\rho_n - e^{na}\sigma_n)_+ = -\phi(a). \qquad (58)$$

Proof. For a fixed $m \in \mathbb{N}$, let $\widehat{\rho}_m := \mathcal{E}_{\sigma_m}(\rho_m)$, and define

$$\widehat{S}_{m,k}(a) := \{\widehat{\rho}_m^{\otimes k} - e^{kma}\sigma_m^{\otimes k} > 0\}. \qquad (59)$$

Write $n \in \mathbb{N}$ in the form $n = km + r$, $k, r \in \mathbb{N}$, $0 \leq r < m$. For any $a, b \in \mathbb{R}$, we have

$$\begin{aligned}
\operatorname{Tr} \rho_n S_n(a) &= \operatorname{Tr}(\rho_n - e^{na}\sigma_n) S_n(a) + e^{na} \operatorname{Tr} \sigma_n S_n(a) \\
&\geq \operatorname{Tr}(\rho_n - e^{na}\sigma_n)_+ \\
&\geq \operatorname{Tr}(\widehat{\rho}_m^{\otimes k} - e^{na}\sigma_m^{\otimes k})_+ \qquad (60) \\
&\geq \operatorname{Tr}(\widehat{\rho}_m^{\otimes k} - e^{na}\sigma_m^{\otimes k}) \widehat{S}_{m,k}(b) \qquad (61) \\
&\geq \operatorname{Tr} \widehat{\rho}_m^{\otimes k} \widehat{S}_{m,k}(b) - e^{na} e^{-kmb} \operatorname{Tr} \widehat{\rho}_m^{\otimes k} \widehat{S}_{m,k}(b) \qquad (62) \\
&= \{1 - e^{ra} e^{-km(b-a)}\} \operatorname{Tr} \widehat{\rho}_m^{\otimes k} \widehat{S}_{m,k}(b), \qquad (63)
\end{aligned}$$

where (60) follows from Lemma 2.1 (with the choice $\mathcal{F} := \mathcal{E}_{\sigma_m}^{\otimes k} \otimes \mathrm{Tr}_{[km+1,r]}$), (61) follows from (10), and we used (11) in (62). Hence, by choosing $b > a$, we get

$$-\phi(a) \geq \limsup_{n\to+\infty} \frac{1}{n} \log \mathrm{Tr}\, \rho_n S_n(a) \geq \liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{Tr}\, \rho_n S_n(a)$$
$$\geq \liminf_{n\to\infty} \frac{1}{n} \log \mathrm{Tr}(\rho_n - e^{na}\sigma_n)_+ \geq \frac{1}{m} \liminf_{k\to\infty} \frac{1}{k} \mathrm{Tr}\, \widehat{\rho}_m^{\otimes k} \widehat{S}_{m,k}(b), \tag{64}$$

where the first inequality is due to (54).

Note that $\widehat{\rho}_m$ and σ_m are commuting density operators, and hence they can be represented as probability density functions on some finite set $\mathcal{X}$, which is the interpretation we will be using in the following. Then $Y := \log \frac{\widehat{\rho}_m}{\sigma_m}$ is a random variable on $\mathcal{X}$, and its logarithmic moment generating function w.r.t. $\widehat{\rho}_m$ is

$$m\psi_m(s) := \Psi_m(s) := \log \mathbb{E}_{\widehat{\rho}_m}\, e^{s \log \frac{\widehat{\rho}_m}{\sigma_m}} = \log \mathrm{Tr}\, \widehat{\rho}_m e^{s \log \frac{\widehat{\rho}_m}{\sigma_m}} = \log \mathrm{Tr}\, \widehat{\rho}_m^{1+s} \sigma_m^{-s}. \tag{65}$$

Note that $\log \frac{\widehat{\rho}_m^{\otimes k}}{\sigma_m^{\otimes k}}$ can naturally be identified with $Y_1 + \cdots + Y_k$, where Y_i is the ith translate of Y on $\times_{j=1}^{+\infty} \mathcal{X}$. Obviously, these translates form a sequence of i.i.d. random variables under the product law $\widehat{\rho}_m^{\otimes\infty}$, and hence, by Cramér's theorem [12, Theorem 2.1.24], we have

$$\liminf_{k\to\infty} \frac{1}{k} \log \mathrm{Tr}\, \widehat{\rho}_m^{\otimes k} \widehat{S}_{m,k}(b) = \liminf_{k\to\infty} \frac{1}{k} \log \mathrm{Tr}\, \widehat{\rho}_m^{\otimes k} \left\{ \frac{1}{k} \log \frac{\widehat{\rho}_m^{\otimes k}}{\sigma_m^{\otimes k}} > mb \right\}$$
$$\geq -\inf_{\kappa > mb} \sup_{s\in\mathbb{R}} \{\kappa s - \Psi_m(s)\}.$$

Assume now that $D(\rho \,\|\, \sigma) < a < b < D_{\max}(\rho \,\|\, \sigma)$. Then we have

$$mb > mD(\rho\|\sigma) = D(\rho_m \,\|\, \sigma_m) \geq D(\widehat{\rho}_m \|\sigma_m) = \mathbb{E}_{\widehat{\rho}_m} \log \frac{\widehat{\rho}_m}{\sigma_m} = \Psi_m'(0),$$

where the second inequality is due to the monotonicity of the quantum relative entropy. Since Ψ_m is convex, it follows that

$$\inf_{\kappa > mb} \sup_{s\in\mathbb{R}} \{\kappa s - \Psi_m(s)\} = \sup_{s\in\mathbb{R}} \{mbs - \Psi_m(s)\} = \sup_{s\geq 0} \{mbs - \Psi_m(s)\}$$
$$= m \sup_{s\geq 0} \{bs - \psi_m(s)\}.$$

Let $\delta_m := \frac{\log v_m}{m}$. From (31), we obtain

$$\psi(s) \leq \psi_m(s) + (1+s)\delta_m, \tag{66}$$

and hence,

$$\sup_{s\geq 0} \{bs - \psi_m(s)\} \leq \sup_{s\geq 0} \{bs - \psi(s) + (1+s)\delta_m\}$$
$$= \sup_{s\geq 0} \{(b+\delta_m)\, s - \psi(s)\} + \delta_m$$
$$= \phi(b+\delta_m) + \delta_m.$$

Putting it all together, we get

$$\frac{1}{m}\liminf_{k\to\infty}\frac{1}{k}\log\operatorname{Tr}\widehat{\rho}_m^{\otimes k}\widehat{S}_{m,k}(b)\geq-\{\phi(b+\delta_m)+\delta_m\}. \tag{67}$$

Substituting it back to (64), taking the limit $m\to+\infty$ and using that $\lim_{m\to+\infty}\delta_m=0$, and that ϕ is continuous on $(D(\rho\,\|\,\sigma), D_{\max}(\rho\,\|\,\sigma))$, we obtain the assertion. □

Theorem 4.5. *For any* $a\in(D(\rho\|\sigma), D_{\max}(\rho\,\|\,\sigma))$, *we have*

$$\lim_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n S_n(a)=-\{\phi(a)+a\}. \tag{68}$$

Proof. By (10), we have

$$\operatorname{Tr}(\rho_n-e^{nb}\sigma_n)_+\geq\operatorname{Tr}(\rho_n-e^{nb}\sigma_n)S_n(a) \tag{69}$$

for any $b\in\mathbb{R}$, and hence,

$$\operatorname{Tr}(\rho_n-e^{nb}\sigma_n)_++e^{nb}\operatorname{Tr}\sigma_n S_n(a)\geq\operatorname{Tr}\rho_n S_n(a). \tag{70}$$

Assume now that $D(\rho\,\|\,\sigma)<a<b<D_{\max}(\rho\,\|\,\sigma)$. Applying Theorem 4.4 to (70), we get

$$-\phi(a)=\liminf_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\rho_n S_n(a)\leq\max\left\{-\phi(b), b+\liminf_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n S_n(a)\right\}.$$

Note that $D(\rho\,\|\,\sigma)<a<b<D_{\max}(\rho\,\|\,\sigma)$ implies $\phi(a)<\phi(b)$, and hence we have

$$-\phi(a)\leq b+\liminf_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n S_n(a). \tag{71}$$

Taking $b\searrow a$, we obtain

$$-\{\phi(a)+a\}\leq\liminf_{n\to\infty}\frac{1}{n}\log\operatorname{Tr}\sigma_n S_n(a). \tag{72}$$

Now combining (55) and (72) yields the assertion. □

Theorems 4.4 and 4.5 yield the following refinement of Corollary 3.8. Note that $\phi(a)$ can also be written as $\phi(a)=\sup_{\alpha>1}\{a(\alpha-1)-F_\alpha(\rho\|\sigma)\}$, where $F_\alpha(\rho\|\sigma)$ is defined in (17). For simplicity, we will use the notation $F(\alpha):=F_\alpha(\rho\|\sigma)$. By Corollary 3.11, $\alpha\mapsto F(\alpha)$ is convex on $(1,+\infty)$, and Lemma 4.2 yields that for every $\alpha\in(1,+\infty)$ there exists an $a_\alpha\in(D(\rho\|\sigma), D_{\max}(\rho\,\|\,\sigma))$ such that

$$\phi(a_\alpha)=a_\alpha(\alpha-1)-F(\alpha). \tag{73}$$

Corollary 4.6. *For every* $\alpha>1$, *let* a_α *be as above, and let* $p_{n,\alpha}:=\{\operatorname{Tr}\rho_n S_n(a_\alpha),\operatorname{Tr}\rho_n(I_n-S_n(a_\alpha))\}$, $q_{n,\alpha}:=\{\operatorname{Tr}\sigma_n S_n(a_\alpha),\operatorname{Tr}\sigma_n(I_n-S_n(a_\alpha))\}$ *be the post-measurement states corresponding to the Neyman–Pearson test* $S_n(a_\alpha)$. *Then*

$$\lim_{n\to+\infty}\frac{1}{n}D_\alpha\left(p_{n,\alpha}\,\|\,q_{n,\alpha}\right)=D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma).$$

Proof. Omitting a standard $\varepsilon-\delta$ argument, we can write Theorems 4.4 and 4.5 as $\operatorname{Tr}\rho_n S_n(a_\alpha)\sim e^{-n\phi(a_\alpha)}$ and $\operatorname{Tr}\sigma_n S_n(a_\alpha)\sim e^{-n(\phi(a_\alpha)+a_\alpha)}$, which then yields

$$\begin{aligned}(\operatorname{Tr}\rho_n S_n(a_\alpha))^\alpha(\operatorname{Tr}\sigma_n S_n(a_\alpha))^{1-\alpha}&\sim\exp\left(-n\left[\alpha\phi(a_\alpha)+(1-\alpha)(\phi(a_\alpha)+a_\alpha)\right]\right)\\&=\exp(nF(\alpha)),\end{aligned}$$

where the last identity is due to (73). Note that $F(\alpha) > 0$ for $\alpha > 1$, and $\lim_{n\to+\infty} \operatorname{Tr} \rho_n (I_n - S_n(a_\alpha)) = \lim_{n\to+\infty} \operatorname{Tr} \sigma_n(I_n - S_n(a_\alpha)) = 1$. Hence, $Q_\alpha^{(\mathrm{new})}(p_{n,\alpha}\|q_{n,\alpha}) \sim \exp(nF(\alpha))$, from which the assertion follows. □

4.3. The strong converse exponent. Consider the hypothesis testing problem from Sect. 4.1. Our aim here is to prove the identity (44), i.e., that the strong converse exponent $B_e^*(r)$, defined in (42), is equal to the converse Hoeffding bound $H_r^*(\rho\|\sigma)$ defined in (8). We will assume that $\rho \neq \sigma$ to avoid a trivial case, and that $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$ so that we actually have a strong converse (cf. Remark 4.1).

We start with the following lemma, which is a direct analogue of Nagaoka's proof of the strong converse to the quantum Stein's lemma [35], except that we use the new Rényi divergences instead of the old ones.

Lemma 4.7. *For any $r \geq 0$, we have*

$$B_e^*(r) \geq H_r^*(\rho\|\sigma). \tag{74}$$

Proof. Let $T_n \in \mathcal{L}(\mathcal{H}_n)$ be a test and let $p_n := (\operatorname{Tr}\rho_n T_n, \operatorname{Tr}\rho_n(I - T_n))$ and $q_n := (\operatorname{Tr}\sigma_n T_n, \operatorname{Tr}\sigma_n(I - T_n))$ be the post-measurement states. By the monotonicity of the Rényi relative entropies under measurements (Lemma 3.3), we have, for any $\alpha > 1$,

$$\begin{aligned} D_\alpha^{(\mathrm{new})}(\rho_n \,\|\, \sigma_n) \geq D_\alpha^{(\mathrm{new})}(p_n \,\|\, q_n) &\geq \frac{1}{\alpha - 1} \log\left[(\operatorname{Tr}\rho_n T_n)^\alpha (\operatorname{Tr}\sigma_n T_n)^{1-\alpha}\right] \\ &= \frac{\alpha}{\alpha - 1} \log(1 - \alpha_n(T_n)) - \log\beta_n(T_n), \end{aligned}$$

or equivalently,

$$\frac{1}{n}\log(1 - \alpha_n(T_n)) \leq \frac{\alpha - 1}{\alpha}\left[D_\alpha^{(\mathrm{new})}(\rho \,\|\, \sigma) + \frac{1}{n}\log\beta_n(T_n)\right]. \tag{75}$$

If $\limsup_{n\to\infty} \frac{1}{n}\log\operatorname{Tr}\sigma_n T_n \leq -r$ then

$$\limsup_{n\to\infty} \frac{1}{n}\log(1 - \alpha_n(T_n)) \leq \frac{\alpha - 1}{\alpha}\left[D_\alpha^{(\mathrm{new})}(\rho \,\|\, \sigma) - r\right], \quad \alpha > 1.$$

Taking the infimum in $\alpha > 1$, the statement follows. □

Remark 4.8. Using that the old Rényi relative entropies are also monotonic under measurements [17], exactly the same argument as above yields that

$$B_e^*(r) \geq \sup_{1<\alpha} \frac{\alpha - 1}{\alpha}\left[r - D_\alpha^{(\mathrm{old})}(\rho \,\|\, \sigma)\right]. \tag{76}$$

This was already pointed out in [39] with a restricted optimization over $\alpha \in (1, 2]$, and later extended by Hayashi to the above form [17].

Our goal in the rest of the section is to show that (74) holds as an equality. To start with, we give some alternative expressions for $H_r^*(\rho\|\sigma)$. Let

$$a_{\max} := D_{\max}(\rho\,\|\,\sigma), \quad \text{and} \quad r_{\max} := \phi(a_{\max}) + a_{\max}. \tag{77}$$

Note that

$$H_r^*(\rho\|\sigma) = \sup_{s\geq 0} \frac{rs - \psi(s)}{s+1} = \sup_{0\leq u<1} \{ur - \tilde{\psi}(u)\}, \tag{78}$$

where

$$\tilde{\psi}(u) := (1-u)\psi\left(\frac{u}{1-u}\right), \quad u \in [0,1).$$

It is easy to see that $\tilde{\psi}'(u) = -\psi(s) + (1+s)\psi'(s)$ with the notational convention $u = s/(s+1)$, and hence

$$\tilde{\psi}(0) = \psi(0) = 0, \quad \tilde{\psi}'(0) = \psi'(0) = D(\rho\,\|\,\sigma), \tag{79}$$

and

$$\begin{aligned}\lim_{u\nearrow 1} \tilde{\psi}'(u) &= \lim_{s\to+\infty} \left(s\psi'(s) - \psi(s)\right) + \lim_{s\to+\infty} \psi'(s) \\ &= \lim_{s\to+\infty} \phi\left(\psi'(s)\right) + D_{\max}(\rho\,\|\,\sigma) = \phi(a_{\max}) + a_{\max} \\ &= r_{\max}.\end{aligned}$$

It is also easy to see, by computing the second derivative, that $\tilde{\psi}$ is convex for commuting ρ and σ; convexity in the general case then follows the same way as in Corollary 3.11. Convexity and (79) yield

$$H_r^*(\rho\|\sigma) = 0, \quad r \leq D(\rho\,\|\,\sigma). \tag{80}$$

Lemma 4.9. *For any $r \geq 0$, we have*

$$H_r^*(\rho\|\sigma) = \begin{cases} r - a_r = \phi(a_r), & r < \phi(a_{\max}) + a_{\max}, \\ r - D_{\max}(\rho\,\|\,\sigma), & r \geq \phi(a_{\max}) + a_{\max}, \end{cases} \tag{81}$$

where $a_{\max}$ and $r_{\max}$ are defined in (77), and a_r is the unique solution of $r - a_r = \phi(a_r)$.

Proof. First, we consider the case $0 \leq r < r_{\max}$. Note that $a \mapsto \phi(a) + a$ is strictly increasing and continuous on $(-\infty, a_{\max})$, and hence for every $r < r_{\max}$ there exists a unique a_r such that $r = \phi(a_r) + a_r$. By definition,

$$\phi(a_r) \geq a_r s - \psi(s) = s(r - \phi(a_r)) - \psi(s), \quad s \geq 0,$$

and equality holds in the above inequality for some $s_r \in [0, +\infty)$. Rearranging, we get

$$\phi(a_r) \geq \frac{sr - \psi(s)}{1+s}, \quad s \geq 0,$$

with equality for s_r, and hence

$$\phi(a_r) = \max_{s\geq 0} \frac{sr - \psi(s)}{1+s}.$$

Taking into account (78), this proves the assertion.

Next, assume that $r \geq r_{\max}$. Note that

$$\lim_{s\to+\infty} \frac{rs-\psi(s)}{s+1} = r - \lim_{s\to+\infty} \frac{\psi(s)}{s+1} = r - D_{\max}(\rho \,\|\, \sigma), \tag{82}$$

due to [34, Theorem 4]. Hence it is enough to show that

$$\frac{rs-\psi(s)}{s+1} \leq r - D_{\max}(\rho \,\|\, \sigma) \tag{83}$$

for every $s \geq 0$. Note that $r \geq r_{\max} = \phi(a_{\max}) + a_{\max}$ implies

$$r - a_{\max} \geq \phi(a_{\max}) \geq a_{\max} s - \psi(s) \tag{84}$$

for every $s \geq 0$, from which we obtain

$$\frac{r+\psi(s)}{s+1} \geq a_{\max}. \tag{85}$$

Thus we have

$$r - a_{\max} \geq r - \frac{r+\psi(s)}{s+1} = \frac{rs-\psi(s)}{s+1}, \tag{86}$$

and hence $H_r^*(\rho\|\sigma) = r - D_{\max}(\rho \,\|\, \sigma)$, as required. □

Now we are ready to prove the identity (44) for the strong converse exponent.

Theorem 4.10. *For any $r \geq 0$, we have*

$$B_e^*(r) = H_r^*(\rho\|\sigma). \tag{87}$$

Proof. Since we have already shown $B_e^*(r) \geq H_r^*(\rho\|\sigma)$ in Lemma 4.7, we only have to show the converse inequality $B_e^*(r) \leq H_r^*(\rho\|\sigma)$. Due to the definition (43) of $B_e^*(r)$ as an infimum of rates, this is equivalent to showing that for any rate $R > H_r^*(\rho\|\sigma)$ there exists a sequence of tests $\{T_n\}_{n=1}^{\infty}$ satisfying

$$\limsup_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n T_n \leq -r \qquad \text{and} \qquad \liminf_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n T_n \geq -R. \tag{88}$$

We prove the claim by considering three different regions of r.

(i) In the case $D(\rho\|\sigma) < r < r_{\max}$, there exists a unique $a_r \in (D(\rho \,\|\, \sigma), D_{\max}(\rho \,\|\, \sigma))$ satisfying $r - a_r = \phi(a_r)$, and Theorems 4.4 and 4.5 yield

$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n S_n(a_r) = -(\phi(a_r) + a_r) = -r,$$
$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n S_n(a_r) = -\phi(a_r) = -H_r^*(\rho\|\sigma),$$

where the last identity is due to Lemma 4.9.

(ii) In the case $0 \le r \le D(\rho\|\sigma)$, we have $H_r^*(\rho\|\sigma) = 0$, according to (80). For any $R > 0$, we can find an $a \in (D(\rho\,\|\,\sigma), D_{\max}(\rho\,\|\,\sigma))$ such that $0 < \phi(a) < R$. Note that $\phi(a) + a > D(\rho\|\sigma) \ge r$, and Theorems 4.4 and 4.5 yield

$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n S_n(a) = -(\phi(a) + a) < -r,$$
$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n S_n(a) = -\phi(a) > -R.$$

(iii) In the case $r \ge r_{\max}$, we use a modification of the Neyman–Pearson tests, following the method of the proof of Theorem 4 in [37]. For every $a, r \in \mathbb{R}$, let

$$T_n(r, a) := e^{-n\{r-a-\phi(a)\}} S_n(a).$$

Note that for $r \ge r_{\max}$ we have $H_r^*(\rho\|\sigma) = r - D_{\max}(\rho\,\|\,\sigma)$ due to Lemma 4.9. Assume now that $a \in (D(\rho\,\|\,\sigma), D_{\max}(\rho\,\|\,\sigma))$. Then $r > \phi(a) + a$, and hence $0 \le T_n(r, a) \le I$, i.e., $T_n(r, a)$ is a test, and

$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n T_n(r, a) = -r + a + \phi(a) - (a + \phi(a)) = -r,$$
$$\lim_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n T_n(r, a) = -r + a + \phi(a) - \phi(a) = -(r - a),$$

by Theorems 4.4 and 4.5. Now for a given $R > H_r^*(\rho\|\sigma) = r - D_{\max}(\rho\,\|\,\sigma)$, we can find an $a \in (D(\rho\,\|\,\sigma), D_{\max}(\rho\,\|\,\sigma))$ such that $r - D_{\max}(\rho\,\|\,\sigma) < r - a < R$, and the assertion follows. □

Remark 4.11. It is easy to see, by applying a standard diagonal argument, that there exists a sequence of tests $\{T_n\}_{n\in\mathbb{N}}$ such that (88) holds with $H_r^*(\rho\|\sigma)$ in place of R, and the proof of Theorem 4.10 yields that for this sequence, we actually have

$$\limsup_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n T_n \le -r \qquad \text{and} \qquad \liminf_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n T_n = -H_r^*(\rho\|\sigma).$$

Moreover, it is also possible to have $\limsup_{n\to\infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n T_n = -r$ above; this is obvious in cases (i) and (iii) in the proof of Theorem 4.10, and in case (ii) this follows from the Hoeffding bound theorem [18,36].

Remark 4.12. The direct region $(0 \le r < D(\rho\|\sigma))$ and the strong converse region $(r > D(\rho\|\sigma))$ in quantum hypothesis testing are considered to be dual, and the theory of both regions can be developed logically independently of the other, which is the approach that we followed here.

Following a different approach, one could prove $B_e^*(r) \le H_r^*(\rho\|\sigma)$ in the case $0 \le r < D(\rho\|\sigma)$ [case (ii) of the above proof] based on Stein's lemma rather than our argument. Indeed, the Stein's lemma (39) implies the existence of a sequence of tests $\{T_n\}_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n(T_n) = -D(\rho\|\sigma) < -r \quad \text{and} \lim_{n\to\infty} \alpha_n(T_n) = 0.$$

By the latter,

$$\lim_{n\to+\infty} \frac{1}{n} \log \operatorname{Tr} \rho_n T_n = 0 = H_r^*(\rho\|\sigma).$$

This proves (88) for every $R > 0$.

Remark 4.13. By Theorem 4.10 and (78), we have

$$B_e^*(r) = H_r^*(\rho\|\sigma) = \sup_{0\le u<1}\{ru - \tilde{\psi}(u)\},$$

where $\tilde{\psi}(u)$ is a continuous convex function on $[0, 1)$. Hence, $B_e^*(r)$ is the Legendre–Fenchel transform (polar function) of $\tilde{\psi}$, and the bipolar theorem says that

$$\sup_{r\ge 0}\{ur - B_e^*(r)\} = \tilde{\psi}(u) = \frac{\alpha-1}{\alpha} D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma)\,, \quad \alpha > 1, \tag{89}$$

where in the last formula we set $\alpha := 1/(1-u)$ and used the definition (48) of ψ. That is, the new Rényi relative entropies can be expressed essentially as the Legendre–Fenchel transform of the operational quantities $B_e^*(r),\ r \ge 0$. A more direct operational interpretation is provided in the next section.

Remark 4.14. A proof for the following representation of the strong converse exponent:

$$B_e^*(r) = \max_{s\ge 0} \frac{rs - \lim_{m\to\infty}\psi_m(s)}{s+1}, \tag{90}$$

where ψ_m is defined in (65), has been outlined in Hayashi's book [17]. More precisely, he proves that the RHS of (90) is a lower bound on $B_e^*(r)$ (optimality); the achievability can be proved by applying the corresponding classical result to the commuting states $\hat{\rho}_m$ and σ_m for every m [19], at least in the region $r < r_{\max}$. Apart from identifying the limit $\lim_{m\to\infty}\psi_m(s)$ as $sD_{1+s}^{(\mathrm{new})}(\rho\,\|\,\sigma)$, our approach here differs from Hayashi's also in that we prove the achievability part by computing explicitly the asymptotic error rates of the Neyman–Pearson tests, providing yet another operational interpretation for the new Rényi divergences.

We note that Theorem 4.10 yields an operational proof of the Lieb–Thirring inequality. Indeed, combining (76) with (89), we get that

$$D_\alpha^{(\mathrm{old})}(\rho\,\|\,\sigma) \ge D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma)\,, \quad \alpha > 1,$$

or equivalently,

$$\operatorname{Tr}\rho^\alpha\sigma^{1-\alpha} \ge \operatorname{Tr}\left(\rho^{\frac{1}{2}}\sigma^{\frac{1-\alpha}{\alpha}}\rho^{\frac{1}{2}}\right)^\alpha, \quad \alpha > 1.$$

Introducing $A := \rho^{\frac{1}{2}}$ and $B := \sigma^{\frac{1-\alpha}{\alpha}}$, the above can be rewritten as

$$\operatorname{Tr} A^\alpha B^\alpha A^\alpha \ge \operatorname{Tr}(ABA)^\alpha\,, \quad \alpha > 1. \tag{91}$$

Since we were interested in hypothesis testing, we only derived Theorem 4.10 for density operators; however, it is easy to see that it also holds, with obvious modifications, for arbitrary positive semidefinite operators. Hence we arrive at the following:

Corollary 4.15 (Lieb–Thirring inequality). *For any positive semidefinite operators A and B,* (91) *holds.*

To close the section, we give one more representation of $H_r^*(\rho\|\sigma)$. This is closely related to the information spectrum approach [37], and although we didn't need it in our proof for the strong converse exponent, an alternative proof could be given based on this representation.

Lemma 4.16. *For any $r \geq 0$, we have*

$$H_r^*(\rho\|\sigma) = \inf_{a\in\mathbb{R}} \max\{\phi(a), r-a\} \tag{92}$$

$$= \inf\{\max\{\phi(a), r-a\} | D(\rho\,\|\,\sigma) < a < D_{\max}(\rho\,\|\,\sigma)\}. \tag{93}$$

Proof. Let $a_{\max}$ and $r_{\max}$ be as in (77). First, we consider the case $0 \leq r < r_{\max}$. Let a_r be the unique solution of $r = \phi(a_r) + a_r$, as in the proof of Lemma 4.9. Then

$$\max\{\phi(a_r), r-a_r\} = \phi(a_r) = r - a_r.$$

Now if $a < a_r$ then $r-a > r-a_r$ and $\phi(a) \leq \phi(a_r)$, which implies $\max\{\phi(a), r-a\} = r-a > r-a_r$. On the other hand, if $a > a_r$ then $r-a < r-a_r$, while $\phi(a) \geq \phi(a_r)$, and hence $\max\{\phi(a), r-a\} = \phi(a) \geq \phi(a_r)$. Thus

$$R(r) := \inf_{a\in\mathbb{R}} \max\{\phi(a), r-a\} = \max\{\phi(a_r), r-a_r\} = \phi(a_r) = r-a_r, \tag{94}$$

and (92) follows by taking into account (81).

Note that when $D(\rho\,\|\,\sigma) < r < r_{\max}$ then $D(\rho\,\|\,\sigma) < a_r < D_{\max}(\rho\,\|\,\sigma)$, and (93) is immediate from (94). In the case $0 \leq r \leq D(\rho\|\sigma)$, we have $r = a_r$ and $R(r) = \phi(a_r) = r - a_r = 0$. On the other hand, for every $D(\rho\,\|\,\sigma) < a < D_{\max}(\rho\,\|\,\sigma)$ we have $\phi(a) > 0 > r - a$, and thus

$$\begin{aligned}&\inf\{\max\{\phi(a), r-a\} | D(\rho\,\|\,\sigma) < a < D_{\max}(\rho\,\|\,\sigma)\}\\ &\quad = \inf\{\phi(a) | D(\rho\,\|\,\sigma) < a < D_{\max}(\rho\,\|\,\sigma)\}\\ &\quad = 0 = R(r),\end{aligned}$$

proving (93).

Next, assume that $r \geq r_{\max}$. Then $r \geq \phi(a) + a$, or equivalently, $r - a \geq \phi(a)$ for every $a \leq a_{\max}$, and hence $\max\{\phi(a), r-a\} = r-a$ for $a \leq a_{\max}$, while for $a > a_{\max}$ we have $\max\{\phi(a), r-a\} = \phi(a) = +\infty$. Hence,

$$\begin{aligned}R(r) &= \inf_{a\in\mathbb{R}} \max\{\phi(a), r-a\} = \inf\{\max\{\phi(a), r-a\} | D(\rho\,\|\,\sigma) < a < D_{\max}(\rho\,\|\,\sigma)\}\\ &= \inf_{a\leq a_{\max}} \{r-a\} = r - a_{\max} = r - D_{\max}(\rho\,\|\,\sigma).\end{aligned}$$

Taking into account (81), we get (92) and (93). □

4.4. Representation as cutoff rates. In the setting of Sect. 4.1, let

$$\alpha_{n,r} := \alpha_{e^{-nr}}(\rho^{\otimes n}\|\sigma^{\otimes n}) := \min\{\mathrm{Tr}\,\rho_n(I-T) \,:\, 0 \leq T \leq I,\ \mathrm{Tr}\,\sigma_n T \leq e^{-nr}\}.$$

Following [10], we define the *generalized* κ*-cutoff rate* $C_\kappa(\rho\|\sigma)$ for any $\kappa > 0$ as the smallest r_0 such that

$$\limsup_{n\to\infty} \frac{1}{n} \log(1-\alpha_{n,r}) \le -\kappa(r - r_0), \quad r > 0. \tag{95}$$

As before, we assume that $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$ and $\rho \neq \sigma$.

Lemma 4.17. *For every* $r > 0$*,*

$$\lim_{n\to+\infty} \frac{1}{n} \log(1-\alpha_{n,r}) = -H_r^*(\rho\|\sigma).$$

Proof. Consider the inequality (75). Taking the supremum over all test T_n such that $\operatorname{Tr}\sigma_n T_n \le e^{-nr}$, we get

$$\frac{1}{n}\log(1-\alpha_{n,r}) \le \frac{\alpha-1}{\alpha}\left[D_\alpha^{(\mathrm{new})}(\rho\,\|\,\sigma) - r\right].$$

Taking now the limsup in n and the infimum in α, we obtain

$$\limsup_{n\to+\infty} \frac{1}{n} \log(1-\alpha_{n,r}) \le -H_r^*(\rho\|\sigma). \tag{96}$$

According to Remark 4.11, for every $r' > 0$, there exists a sequence of tests $T_{n,r'}$, $n \ge 1$, such that

$$\limsup_{n\to+\infty} \frac{1}{n} \log \operatorname{Tr}\sigma_n T_{n,r'} \le -r' \quad \text{and} \quad \liminf_{n\to+\infty} \frac{1}{n} \log \operatorname{Tr}\rho_n T_{n,r'} \ge -H_{r'}(\rho\|\sigma). \tag{97}$$

Hence, for any $r' > r$, there exists an $N_{r'}$ such that for all $n > N_{r'}$, $\operatorname{Tr}\sigma_n T_{n,r'} \le e^{-nr}$, and thus $\operatorname{Tr}\rho_n T_{n,r'} \le 1-\alpha_{n,r}$. By the second inequality in (97),

$$\liminf_{n\to+\infty} \frac{1}{n} \log(1-\alpha_{n,r}) \ge \liminf_{n\to+\infty} \frac{1}{n} \log \operatorname{Tr}\rho_n T_{n,r'} \ge -H_{r'}(\rho\|\sigma). \tag{98}$$

From the definition (8) of the converse Hoeffding divergence, it is clear that $r \mapsto H_r^*(\rho\|\sigma)$ is a monotone increasing convex function on $(0,+\infty)$. Moreover, Lemma 4.9 implies that $H_r^*(\rho\|\sigma)$ is finite for every $r > 0$. Thus, $r \mapsto H_r^*(\rho\|\sigma)$ is continuous on $(0,+\infty)$, and (98) yields

$$\liminf_{n\to+\infty} \frac{1}{n} \log(1-\alpha_{n,r}) \ge \sup_{r'>r} -H_{r'}(\rho\|\sigma) = -H_r(\rho\|\sigma). \tag{99}$$

Finally, (96) and (99) yield the assertion. □

Theorem 4.18. *For every* $\kappa \in (0,1)$*,*

$$C_\kappa(\rho\|\sigma) = D_{\frac{1}{1-\kappa}}^{(\mathrm{new})}(\rho\,\|\,\sigma)\,.$$

Proof. By Lemma 4.17 and (78), we have

$$\lim_{n\to\infty}\frac{1}{n}\log(1-\alpha_{n,r})=-H_r^*(\rho\|\sigma)=-\sup_{0\le u<1}\{ru-\tilde\psi(u)\}.$$

By definition, we have

$$H_r^*(\rho\|\sigma)\ge r\kappa-\tilde\psi(\kappa)=\kappa\left(r-\frac{1}{\kappa}\tilde\psi(\kappa)\right),$$

and the above inequality holds with equality for $r_\kappa:=\tilde\psi'(\kappa)$, and hence

$$\frac{1}{\kappa}\tilde\psi(\kappa)=\frac{1}{\kappa}(1-\kappa)\psi\left(\frac{\kappa}{1-\kappa}\right)=D^{(\text{new})}_{\frac{1}{1-\kappa}}(\rho\,\|\,\sigma)$$

is the smallest r_0 for which (95) holds. □

The above theorem immediately yields the following operational interpretation of the new Rényi relative entropies:

Corollary 4.19. *For every $\alpha>1$,*

$$D_\alpha^{(\text{new})}(\rho\,\|\,\sigma)=C_{\frac{\alpha-1}{\alpha}}(\rho\|\sigma).$$

The above operational interpretation yields as an immediate consequence an alternative proof for the monotonicity of the new Rényi divergences, Corollary 3.14 and Remark 3.18:

Corollary 4.20. *Let $\rho,\sigma\in\mathcal{B}(\mathcal{H})_+$ and $\mathcal{F}:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{K})$ be a trace-preserving linear map such that $\mathcal{F}^{\otimes n}$ is positive for every $n\in\mathbb{N}$. Then*

$$D_\alpha^{(\text{new})}(\mathcal{F}(\rho)\|\mathcal{F}(\sigma))\le D_\alpha^{(\text{new})}(\rho\|\sigma),\quad \alpha>1.$$

In particular, $D_\alpha^{(\text{new})}$ is monotone non-increasing under CPTP maps for every $\alpha>1$.

Proof. By assumption, the Hilbert–Schmidt dual $(\mathcal{F}^{\otimes n})^*$ is positive and unital for every $n\in\mathbb{N}$, and hence

$$\begin{aligned}
&\alpha_{e^{-nr}}(\mathcal{F}(\rho)^{\otimes n}\|\mathcal{F}(\sigma)^{\otimes n})\\
&\quad=\min\{\operatorname{Tr}\mathcal{F}^{\otimes n}(\rho^{\otimes n})(I-T)\colon 0\le T\le I,\ \operatorname{Tr}\mathcal{F}^{\otimes n}(\sigma^{\otimes n})T\le e^{-nr}\}\\
&\quad=\min\{\operatorname{Tr}\rho^{\otimes n}(I-(\mathcal{F}^{\otimes n})^*(T))\colon 0\le T\le I,\ \operatorname{Tr}\sigma^{\otimes n}(\mathcal{F}^{\otimes n})^*(T)\le e^{-nr}\}\\
&\quad\ge\min\{\operatorname{Tr}\rho^{\otimes n}(I-T)\colon 0\le T\le I,\ \operatorname{Tr}\sigma^{\otimes n}T\le e^{-nr}\}\\
&\quad=\alpha_{e^{-nr}}(\rho^{\otimes n}\|\sigma^{\otimes n}).
\end{aligned}$$

Thus for every $\kappa\in(0,1)$, and every $r>0$,

$$\begin{aligned}
&\limsup_{n\to+\infty}\frac{1}{n}\log(1-\alpha_{e^{-nr}}(\mathcal{F}(\rho)^{\otimes n}\|\mathcal{F}(\sigma)^{\otimes n}))\\
&\quad\le\limsup_{n\to+\infty}\frac{1}{n}\log(1-\alpha_{e^{-nr}}(\rho^{\otimes n}\|\sigma^{\otimes n}))\le-\kappa r+\kappa D^{(\text{new})}_{\frac{1}{1-\kappa}}(\rho\,\|\,\sigma),
\end{aligned}$$

where in the last inequality we used Theorem 4.18. By the definition of the κ-cutoff rate and Theorem 4.18, we get

$$D^{(\mathrm{new})}_{\frac{1}{1-\kappa}}(\mathcal{F}(\rho)\,\|\,\mathcal{F}(\sigma)) = C_\kappa(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) \leq D^{(\mathrm{new})}_{\frac{1}{1-\kappa}}(\rho\,\|\,\sigma)\,,$$

proving the assertion. □

5. Conclusion

In this paper we have determined the exact strong converse exponent for binary quantum hypothesis testing, and showed that it can be expressed in terms of the recently introduced version of quantum Rényi α-relative entropies $D_\alpha^{(\mathrm{new})}$ [34,50] with parameters $\alpha > 1$. Following then Csiszár's approach, we gave a direct operational interpretation of these Rényi relative entropies as generalized cutoff rates. Our results show that, at least in the context of hypothesis testing, the operationally relevant quantum generalization of Rényi's α-relative entropies for $\alpha > 1$ are given by $D_\alpha^{(\mathrm{new})}$. On the other hand, previous results [2,18,31,36] show that for $\alpha < 1$, the operationally relevant quantum generalization is the traditional notion $D_\alpha^{(\mathrm{old})}$.

Our proof for the optimality of the converse Hoeffding divergence for the strong converse rate follows immediately from the monotonicity of $D_\alpha^{(\mathrm{new})}$, $\alpha > 1$, under measurements; this proof technique goes back to Nagaoka's proof for the strong converse [35]. We proved the achievability of the converse Hoeffding divergence for the strong converse rate by showing that the quantum Neyman–Pearson tests (or suitable modifications for large r) achieve it for a suitably chosen trade-off parameter. The proof uses the pinching technique developed by Hayashi [16,17], classical large deviation theory, and, for (66), the asymptotic attainability of the new Rényi relative entropies by pinching. An alternative proof for the achievability of the converse Hoeffding divergence can be obtained by combining the pinching technique with the Gärtner–Ellis theorem; this approach can be used also for the hypothesis testing problem of various non-i.i.d. states [33].

Acknowledgements. The authors would like to thank Dr. Gen Kimura for his hospitality, and MM would further like to thank Prof. Fumio Hiai, Dr. Hiromichi Ohno and Prof. Takashi Sano for their hospitality during his visit in Japan. MM acknowledges support by the European Commission (Marie Curie Fellowship "QUANTSTAT") and by the European Research Council (Advanced Grant "IRQUAT"). Part of this work was done when MM was a Marie Curie research fellow at the School of Mathematics, University of Bristol. TO was partially supported by MEXT Grant-in-Aid (A) No. 20686026 "Project on Multi-user Quantum Network". The authors are grateful to Nilanjana Datta, Masahito Hayashi, Ke Li, Marco Tomamichel and Andreas Winter for comments and discussions, and to an anonymous referee for his/her comments. The authors are also grateful to Masahito Hayashi and Marco Tomamichel for sharing with them the manuscript [21] before publication.

Appendix A: Monotonicity and Attainability Properties of the Rényi Divergences

For a general quantum divergence D (i.e., a function on pairs of density operators), one can consider various monotonicity and attainability properties. By a monotonicity property we mean that for every $\rho, \sigma \in \mathcal{B}(\mathcal{H})_+$ and every $\mathcal{F} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ belonging to a certain class of maps,

$$D(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) \leq D(\rho\|\sigma). \tag{A1}$$

Here we will consider the monotonicity properties MON, SMON, EPPMON, MMON and PMON, where in each case, the map $\mathcal{F}$ in (A1) is a trace-preserving positive linear map, with the following additional properties:

MON	$\mathcal{F}$ is completely positive.
SMON	$\mathcal{F}$ is a stochastic map in the sense of [25], i.e., it is the convex combination of two trace-preserving maps $\mathcal{F}_1$ and $\mathcal{F}_2$, such that the adjoint (w.r.t. the Hilbert–Schmidt inner product) of $\mathcal{F}_1$ is a Schwarz map, and the adjoint of $\mathcal{F}_2$ is a Schwarz map composed with the transposition in some basis.
EPPMON	$\mathcal{F}$ is such that every tensor power $\mathcal{F}^{\otimes n}$ is positive, $n \in \mathbb{N}$.
MMON	$\mathcal{F}$ is a measurement, i.e., all operators in $\mathcal{F}(\mathcal{B}(\mathcal{H}))$ commute with each other.
PMON	$\mathcal{F}$ is the pinching with respect to the reference state σ.

The following implications are obvious:

$$\begin{array}{ccccc} \text{SMON} & & & & \\ \Downarrow & & & & \\ \text{MON} & \Longrightarrow & \text{MMON} & \Longrightarrow & \text{PMON}. \\ \Uparrow & & & & \\ \text{EPPMON} & & & & \end{array}$$

By an asymptotic attainability property we mean that for every $\rho, \sigma \in \mathcal{B}(\mathcal{H})_+$, there exists a sequence of maps $\mathcal{F}_n\colon \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{K}_n)$, $n \in \mathbb{N}$, with each $\mathcal{F}_n$ belonging to some class further specified below, such that

$$D(\rho\|\sigma) = \lim_{n\to+\infty} \frac{1}{n} D(\mathcal{F}_n(\rho^{\otimes n})\|\mathcal{F}_n(\sigma^{\otimes n})).$$

Here we will consider

AAM	(asymptotic attainability by measurements) Every $\mathcal{F}_n$ is a measurement.
AAP	(asymptotic attainability by pinching) Every $\mathcal{F}_n$ is the pinching with respect to the reference state $\sigma^{\otimes n}$.

The following implication is obvious:

$$\text{AAP} \Longrightarrow \text{AAM}. \tag{A2}$$

Furthermore, we say that D satisfies AAMmax if

$$D(\rho\|\sigma) = \lim_{n\to+\infty} \frac{1}{n} \max_{\mathcal{F}_n \text{measurement}} D(\mathcal{F}_n(\rho^{\otimes n})\|\mathcal{F}_n(\sigma^{\otimes n})).$$

We have

$$\text{MMON} + \text{AAM} \Longrightarrow \text{AAMmax} \Longrightarrow \text{EPPMON}, \tag{A3}$$

where the first implication is straightforward to verify, and the second one follows the same way as in Corollary 3.13.

The following table summarizes the monotonicity and attainability properties of the old and the new Rényi relative entropies (NK stands for "Not Known"):

<table>
<tr><th></th><th></th><th>$(0, 1/2)$</th><th>$[1/2, 1)$</th><th>$(1, 2]$</th><th>$(2, +\infty)$</th></tr>
<tr><td rowspan="2">SMON</td><td>$D_\alpha^{(\mathrm{old})}$</td><td colspan="3">YES[1]</td><td>NO[2]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td>NO[3]</td><td colspan="3">NK</td></tr>
<tr><td rowspan="2">EPPMON</td><td>$D_\alpha^{(\mathrm{old})}$</td><td colspan="2">YES[1]</td><td>NK</td><td>NO[2]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td>NO[3]</td><td colspan="3">YES[4]</td></tr>
<tr><td rowspan="2">MON</td><td>$D_\alpha^{(\mathrm{old})}$</td><td colspan="3">YES[1]</td><td>NO[2]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td>NO[3]</td><td colspan="3">YES[4]</td></tr>
<tr><td rowspan="2">MMON</td><td>$D_\alpha^{(\mathrm{old})}$</td><td colspan="4">YES[1]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td>NK</td><td colspan="3">YES[4]</td></tr>
<tr><td rowspan="2">PMON</td><td>$D_\alpha^{(\mathrm{old})}$</td><td colspan="4">YES[1]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td colspan="4">YES[4]</td></tr>
<tr><td rowspan="2">AAP</td><td>$D_\alpha^{(\mathrm{old})}$</td><td colspan="4">NO[5]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td colspan="4">YES[4]</td></tr>
<tr><td rowspan="2">AAM</td><td>$D_\alpha^{(\mathrm{old})}$</td><td>NK</td><td colspan="3">NO[5]</td></tr>
<tr><td>$D_\alpha^{(\mathrm{new})}$</td><td colspan="4">YES[4]</td></tr>
</table>

[1] Monotonicity of $D_\alpha^{(\mathrm{old})}$ for $\alpha \in [0, 2]$ under 2-positive maps has been proved in [40], and has been extended to stochastic maps in [25]. MMON and PMON for $\alpha \in [0, 2]$ are immediate consequences, and for $\alpha > 2$ they have been proved by a different method in [17, Section 3.7]. EPPMON follows from the operational interpretation of $D_\alpha^{(\mathrm{old})}$ for $\alpha \in (0, 1)$ in the context of the Hoeffding bound; see, e.g., [36].

[2] Failure of MON for $\alpha > 2$ was pointed out in [34, page 7]. One can easily see that MON is equivalent to joint convexity for the core quantities of the old Rényi divergences, $Q_\alpha(\rho\|\sigma) := \operatorname{Tr} \rho^\alpha \sigma^{1-\alpha}$; see, e.g., [40]. An easy argument [20], omitted in [34], shows that even convexity of Q_α in its first argument implies the operator convexity of the power function $\mathbb{R}_+ \ni x \mapsto x^\alpha$. Since the latter is not true for $\alpha > 2$ (see, e.g., [6, Exercise V.2.11]), MON cannot hold for $D_\alpha^{(\mathrm{old})}$, $\alpha > 2$, from which the failure of SMON and EPPMON for the same range of α are obvious.

[3] MON for $D_\alpha^{(\mathrm{new})}$ is also equivalent to joint convexity, the failure of which for $\alpha < 1/2$ has been confirmed by numerical examples according to [34]. Failure of MON obviously yields failure of SMON and EPPMON.

[4] MON for $D_\alpha^{(\mathrm{new})}$ have been proved by various methods, applicable to different parameter ranges, in [4,13,34,50]. These approaches either prove monotonicity directly, or through joint convexity, and rely on techniques from matrix analysis or functional analysis.

In this paper we followed a different approach, starting from PMON, that has been proved for all parameter values $\alpha \geq 0$ in [34]. We then proved, for $\alpha > 1$, MMON in Lemma 3.3 and AAP in Theorem 3.7, which in turn yield AAM and the stronger monotonicity property EPPMON, according to (A2) and (A3); see also Corollary 3.14 and Remark 3.18.

AAP for $\alpha \in [0, 1)$ has been proved very recently in [21]. It is not clear whether MMON and thus EEPMON for $\alpha \in [1/2, 1)$ can be obtained from it the same way as for $\alpha > 1$ in the present paper. However, when combined with MON for $\alpha \in [1/2, 1)$, derived by other methods as mentioned above, it implies AAM and thus EPPMON for $\alpha \in [1/2, 1)$, according to (A2) and (A3).

[5] For commuting states the old and the new Rényi relative entropies coincide, whereas for non-commuting states the inequality in (7) is strict according to [23]. Thus, AAP for $D_\alpha^{(\text{new})}$ implies that AAP cannot hold for $D_\alpha^{(\text{old})}$, for any fixed value $\alpha \in (0, +\infty)\setminus\{1\}$. For $\alpha \geq 1/2$, AAM+MMON yields AAMmax according to (A3), and hence

$$\lim_{n\to+\infty} \frac{1}{n} \max_{\mathcal{F}_n \text{measurement}} D_\alpha^{(\text{old})}(\mathcal{F}_n(\rho^{\otimes n}) \| \mathcal{F}_n(\sigma^{\otimes n})) = D_\alpha^{(\text{new})} < D_\alpha^{(\text{old})}$$

whenever ρ and σ don't commute, showing that AAM fails for $D_\alpha^{(\text{old})}$, $\alpha \geq 1/2$.

Remark A.1. In Corollary 4.20 we presented an approach to obtain EPPMON from the operational representation in Corollary 4.19. However, to obtain Corollary 4.19, we used MMON (to prove Lemma 4.7) and AAP [for (66)], from which properties EEPMON is immediate, as we have seen above. It is an interesting open question whether the cutoff rate representation, or Theorem 4.10, can be obtained without the use of monotonicity and achievability properties, thus providing a fully operational proof for the monotonicity of the new Rényi divergences for $\alpha > 1$. We remark that such a fully operational proof for $D_\alpha^{(\text{old})}$, $\alpha \in (0, 1)$, follows from the Hoeffding bound theorem, as it was pointed out in [36].

Remark A.2. For $\alpha = 1$, the old and the new Rényi relative entropies yield the same limit D_1, Umegaki's relative entropy. SMON and EPPMON for $D_\alpha^{(\text{old})}$ yield immediately the same properties for D_1 by taking the limit $\alpha \to 1$. AAP has been shown in [22], and it was the key technical tool to prove the direct part of the quantum Stein's lemma [22], and various generalizations of it [7–9]. From these, the rest of the properties, MON, MMON, PMON, AAM and AMMmax, follow immediately, as we have seen before.

The above properties show that the new Rényi relative entropies provide the smallest possible quantum extension of the classical Rényi relative entropies, under very mild conditions.

Proposition A.3. *For a fixed $\alpha \geq 0$, let $\widehat{D}_\alpha$ be a function on pairs of quantum states on the same Hilbert space, with the following properties:*

1. *$\widehat{D}_\alpha$ coincides with the classical Rényi relative entropy D_α on commuting states;*
2. *$\widehat{D}_\alpha$ is additive, i.e., for every ρ, σ and every $n \in \mathbb{N}$, $\widehat{D}_\alpha(\rho^{\otimes n} \| \sigma^{\otimes n}) = n\widehat{D}_\alpha(\rho\|\sigma)$;*
3. *$\widehat{D}_\alpha$ satisfies PMON.*

Then $D_\alpha^{(\text{new})} \leq \widehat{D}_\alpha$. In particular, $D_\alpha^{(\text{new})} \leq D_\alpha^{(\text{old})}$ for every $\alpha \in [0, +\infty]\setminus\{1\}$.

Proof. Let ρ and σ be fixed. By assumption, we have

$$D_\alpha(\mathcal{E}_{\sigma^{\otimes n}}(\rho^{\otimes n}) \| \sigma^{\otimes n}) = \widehat{D}_\alpha(\mathcal{E}_{\sigma^{\otimes n}}(\rho^{\otimes n}) \| \sigma^{\otimes n}) \leq \widehat{D}_\alpha(\rho^{\otimes n} \| \sigma^{\otimes n}) = n\widehat{D}_\alpha(\rho\|\sigma).$$

Using that $D_\alpha^{(\mathrm{new})}$ satisfies AAP, we get

$$D_\alpha^{(\mathrm{new})}(\rho\|\sigma) = \lim_{n\to+\infty} \frac{1}{n} D_\alpha(\mathcal{E}_{\sigma^{\otimes n}}(\rho^{\otimes n})\|\sigma^{\otimes n}) \le \widehat{D}_\alpha(\rho\|\sigma).$$

□

Sufficiency and single-shot attainability

Instead of the asymptotic attainability properties studied above, one can also consider single-shot attainability. Here we will be interested in attainability by measurements (AM), which is satisfied by a quantum divergence D if for every pair of states ρ, σ, there exists a measurement $\mathcal{F}$ such that $D(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) = D(\rho\|\sigma)$. It is easy to see that

$$\text{AM} + \text{MMON} \Longrightarrow \text{monotonicity under trace-preserving positive maps}, \tag{A4}$$

a very strong monotonicity property. It is clear that $D_\alpha^{(\mathrm{old})}$ cannot satisfy AM for any $\alpha \in (0, +\infty)\setminus\{1\}$, due to the strict inequality in (7) for non-commuting states. It is an open question whether AAM for $D_\alpha^{(\mathrm{new})}$ can be strengthened to AM in general. However, we have the following special cases:

Lemma A.4. *$D_{1/2}^{(\mathrm{new})}$ and $D_{+\infty}^{(\mathrm{new})} = D_{\max}$ satisfy AM.*

Proof. Note that $D_{1/2}^{(\mathrm{new})} = -2\log F$, where F is Uhlmann's fidelity [47]. Since the fidelity is known to be attainable by measurements (see, e.g., [38, Chapter 9]), the assertion follows for $D_{1/2}^{(\mathrm{new})}$.

If $\rho, \sigma \in \mathcal{B}(\mathcal{H})_+$ are such that $\operatorname{supp}\rho \le \operatorname{supp}\sigma$ then one can use the duality of linear programming to write the max-relative entropy of ρ and σ as [5,46,51]

$$\begin{aligned} D_{\max}(\rho\|\sigma) &= \max\{\log \operatorname{Tr} M\rho \colon 0 \le M,\ \operatorname{Tr} M\sigma = 1\} \\ &= \max\left\{\log \frac{\operatorname{Tr} M\rho}{\operatorname{Tr} M\sigma} \colon 0 \le M \le I\right\} \\ &= \max\left\{\max_{x\in\mathcal{X}}\left\{\log \frac{\operatorname{Tr} M_x\rho}{\operatorname{Tr} M_x\sigma}\right\} : \{M_x\}_{x\in\mathcal{X}} \text{ POVM}\right\} \\ &= \max\left\{D_{\max}(\{\operatorname{Tr}\rho M_x\}_{x\in\mathcal{X}}\|\{\operatorname{Tr}\sigma M_x\}_{x\in\mathcal{X}}) \colon \{M_x\}_{x\in\mathcal{X}} \text{ POVM}\right\}. \end{aligned}$$

The equality between the first and the last expression above holds trivially when $\operatorname{supp}\rho \le \operatorname{supp}\sigma$ is not satisfied. □

It is well-known that the fidelity is monotone non-decreasing, or equivalently, $D_{1/2}^{(\mathrm{new})}$ is monotone non-increasing, under CPTP maps. Combining this with Lemma A.4, we get the following stronger monotonicity property:

Corollary A.5. *The fidelity is monotone non-decreasing, or equivalently, $D_{1/2}^{(\mathrm{new})}$ is monotone non-increasing, under trace-preserving positive maps.*

Proof. Monotonicity under CPTP maps implies MMON, and thus the assertion is immediate from Lemma A.4 and (A4). □

Remark A.6. Monotonicity of $D_{\max}$ under trace-preserving positive maps is trivial from its definition (47).

Remark A.7. It is easy to see that for fixed states, the classical Rényi relative entropies are monotone increasing in the parameter α. Lemma A.4 thus yields that

$$D_{\max}(\rho\|\sigma) = \max_{\alpha\in[0,+\infty]} \max\left\{D_\alpha\left(\{\mathrm{Tr}\, M_i\rho\}\|\{\mathrm{Tr}\, M_i\sigma\}\right) : \{M_i\}\ \mathrm{POVM}\right\},$$

i.e., the max-relative entropy of ρ and σ is the largest Rényi α-relative entropy of the classical distributions that can be obtained from ρ and σ after performing a measurement.

We say that a quantum divergence D satisfies the sufficiency property (S) if the following holds: for every states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, and CPTP map $\mathcal{F}$: $B(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$,

$$D(\mathcal{F}(\rho)\|\mathcal{F}(\sigma)) = D(\rho\|\sigma) \tag{A5}$$

implies the existence of a CPTP map $\mathcal{F}'$: $\mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ such that

$$\mathcal{F}'(\mathcal{F}(\rho)) = \rho \quad \text{and} \quad \mathcal{F}'(\mathcal{F}(\sigma)) = \sigma. \tag{A6}$$

Obviously, if D is monotone under CPTP maps then (A6) implies (A5). Thus, for a monotone divergence, sufficiency means that the monotonicity inequality is strict in the sense that it can only be saturated in a trivial way.

The old Rényi relative entropies $D_\alpha^{(\mathrm{old})}$ satisfy MON for every $\alpha \in [0, 2]$, and they are known to have the sufficiency property for every parameter value in this interval, except for its endpoints 0 and 2; see [25–27,41,42]. Failure of (S) for $\alpha = 0$ is trivial to see, and for $\alpha = 2$ it follows from a counterexample given in [28, Example 2.2] and [25, Section 5].

Sufficiency for the new Rényi relative entropies is an open question for every parameter value, except at the endpoints of the monotonicity interval $[1/2, +\infty]$. Below we show that, similarly to the case of the old Rényi relative entropies, sufficiency fails at these points.

The following lemma is due to Petz [43, Lemma 4.1].

Lemma A.8. *Let ρ, σ be states and $\{M_x\}_{x\in\mathcal{X}}$ be a measurement such that*

$$D_{1/2}^{(\mathrm{old})}\left(\{\mathrm{Tr}\,\rho M_x\}_{x\in\mathcal{X}}\|\{\mathrm{Tr}\,\sigma M_x\}_{x\in\mathcal{X}}\right) = D_{1/2}^{(\mathrm{old})}(\rho\|\sigma). \tag{A7}$$

Then ρ and σ commute.

Corollary A.9. *No quantum divergence can satisfy (A)+(S). In particular, $D_{1/2}^{(\mathrm{new})}$ and $D_\infty^{(\mathrm{new})}$ do not satisfy (S).*

Proof. Assume that D satisfies (A) and (S), and let ρ, σ be non-commuting states. By (A), there exists a POVM $\{M_x\}_{x\in\mathcal{X}}$ such that $D(\rho\|\sigma) = D\left(\{\mathrm{Tr}\,\rho M_x\}_{x\in\mathcal{X}}\|\{\mathrm{Tr}\,\sigma M_x\}_{x\in\mathcal{X}}\right)$. By (S), there exists a CPTP map Ψ such that $\Psi(\{\mathrm{Tr}\,\rho M_x\}_{x\in\mathcal{X}}) = \rho$ and $\Psi(\{\mathrm{Tr}\,\sigma M_x\}_{x\in\mathcal{X}}) = \sigma$. By the monotonicity of $D_{1/2}^{(\mathrm{old})}$, we have (A7), and by Lemma A.8, ρ and σ commute, which is a contradiction.

The assertion about $D_{1/2}^{(\mathrm{new})}$ and $D_\infty^{(\mathrm{new})}$ follows as a special case, due to Lemma A.4. □

References

1. Araki, H.: On an inequality of Lieb and Thirring. Lett. Math. Phys. **19**(2), 167–170 (1990)
2. Audenaert, K.M.R., Nussbaum, M., Szkoła, A., Verstraete, F.: Asymptotic error rates in quantum hypothesis testing. Commun. Math. Phys. **279**, 251–283 (2008)
3. Audenaert, K.M.R.: On the Araki–Lieb–Thirring inequality. Int. J. Inf. Syst. Sci. **4**, 78–83 (2008)
4. Beigi, S.: Quantum Rényi divergence satisfies data processing inequality. J. Math. Phys. **54**, 122202 (2013)
5. Berta, M., Furrer, F., Scholz, V.B.: The smooth entropy formalism on von Neumann algebras (2011). arXiv:1107.5460
6. Bhatia, R.: Matrix Analysis. Springer, New York (1997)
7. Bjelakovic, I., Krüger, T., Siegmund-Schultze, R., Szkoła, A.: The Shannon–McMillan theorem for ergodic quantum lattice systems. Invent. Math. **155**(1), 203–222 (2004)
8. Bjelakovic, I., Siegmund-Schultze, R.: An ergodic theorem for the quantum relative entropy. Commun. Math. Phys. **247**, 697–712 (2004)
9. Bjelakovic, I., Deuschel, J.-D., Krüger, T., Seiler, R., Siegmund-Schultze, R., Szkoła, A.: A quantum version of Sanov's theorem. Commun. Math. Phys. **260**(3), 659–671 (2005)
10. Csiszár, I.: Generalized cutoff rates and Rényi's information measures. IEEE Trans. Inf. Theory **41**, 26–34 (1995)
11. Datta, N.: Min- and max-relative entropies and a new entanglement monotone. IEEE Trans. Inf. Theory **55**(6), 2816–2826 (2009)
12. Dembo, A., Zeitouni, O.: Large deviations techniques and applications. In: Application of Mathematics, 2nd edn, vol. 38. Springer, New York (1998)
13. Frank, R.L., Lieb, E.H.: Monotonicity of a relative Rényi entropy. J. Math. Phys **54**, 122201 (2013)
14. Han, T.S., Kobayashi, K.: The strong converse theorem for hypothesis testing. IEEE Trans. Inf. Theory **35**, 178–180 (1989)
15. Han, T.S.: Information-Spectrum Methods in Information Theory. Springer, New York (2003)
16. Hayashi, M.: Optimal sequence of POVM's in the sense of Stein's lemma in quantum hypothesis testing. J. Phys. A Math. Gen. **35**, 10759–10773 (2002)
17. Hayashi, M.: Quantum Information Theory: An Introduction. Springer, New York (2006)
18. Hayashi, M.: Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. Phys. Rev. A **76**, 062301 (2007)
19. Hayashi, M.: Private communication (2014)
20. Hayashi, M., Tomamichel, M.: Private communication (2013)
21. Hayashi, M., Tomamichel, M.: Correlation detection and an operational interpretation of the Renyi mutual information. arXiv:1408.6894
22. Hiai, F., Petz, D.: The proper formula for relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. **143**, 99–114 (1991)
23. Hiai, F.: Equality cases in matrix norm inequalities of Golden–Thompson type. Linear Multilinear Algebra **36**, 239–249 (1994)
24. Hiai, F., Mosonyi, M., Ogawa, T.: Error exponents in hypothesis testing for correlated states on a spin chain. J. Math. Phys. **49**, 032112 (2008)
25. Hiai, F., Mosonyi, M., Petz, D., Bény, C.: Quantum f-divergences and error correction. Rev. Math. Phys. **23**(7), 691–747 (2011)
26. Jenčová, A., Petz, D.: Sufficiency in quantum statistical inference. Commun. Math. Phys. **263**, 259–276 (2006)
27. Jenčová, A., Petz, D.: Sufficiency in quantum statistical inference. A survey with examples. Infin. Dimens. Anal. Quantum Probab. Relat. Top. **9**, 331–351 (2006)
28. Jenčová, A., Petz, D., Pitrik, J.: Markov triplets on CCR algebras. Acta Sci. Math. (Szeged) **76**, 27–50 (2010)
29. Lieb, E.H.: Convex trace functions and the Wigner–Yanase–Dyson conjecture. Adv. Math. **11**, 267–288 (1973)
30. Lieb, E.H., Thirring, W.: Studies in Mathematical Physics, pp. 269–297. Princeton University Press, Princeton (1976)
31. Mosonyi, M., Hiai, F.: On the quantum Rényi relative entropies and related capacity formulas. IEEE Trans. Inf. Theory **57**, 2474–2487 (2011)
32. Mosonyi, M.: Inequalities for the quantum Rényi divergences with applications to compound coding problems (2013). arxiv:1310.7525
33. Mosonyi, M., Ogawa, T.: The strong converse rate of quantum hypothesis testing for correlated quantum states. arXiv:1407.3567
34. Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., Tomamichel, M.: On quantum Renyi entropies: a new definition and some properties. J. Math. Phys. **54**, 122203 (2013)

35. Nagaoka, H.: Strong converse theorems in quantum information theory. In: Proceedings of ERATO Workshop on Quantum Information Science, Tokyo, p. 33 (2001). (Hayashi, M. (ed.) Asymptotic Theory of Quantum Statistical Inference, pp. 64–65. World Scientific (2005)
36. Nagaoka, H.: The converse part of the theorem for quantum Hoeffding bound. quant-ph/0611289
37. Nagaoka, H., Hayashi, M.: An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses. IEEE Trans. Inf. Theory **53**, 534–549 (2007)
38. Nielsen, M.A., Chuang, I.L.: Quantum Information and Quantum Computation. Cambridge University Press, Cambridge (2000)
39. Ogawa, T., Nagaoka, H.: Strong converse and Stein's lemma in quantum hypothesis testing. IEEE Trans. Inf. Theory **47**, 2428–2433 (2000)
40. Petz, D.: Quasi-entropies for finite quantum systems. Rep. Math. Phys. **23**, 57–65 (1986)
41. Petz, D.: Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. Commun. Math. Phys. **105**, 123–131 (1986)
42. Petz, D.: Sufficiency of channels over von Neumann algebras. Q. J. Math. Oxf. Ser. (2) **39**(153), 97–108 (1988)
43. Petz, D.: Monotonicity of quantum relative entropy revisited. Rev. Math. Phys. **15**(1), 79–91 (2003)
44. Renner, R.: Security of quantum key distribution. Ph.D. dissertation, Swiss Federal Institute of Technology Zurich, Diss. ETH No. 16242 (2005)
45. Rényi, A.: On measures of entropy and information. In: Proceedings of 4th Berkeley Sympos. Math. Statist. and Prob., vol. I, pp. 547–561. Univ. California Press, Berkeley (1961)
46. Tomamichel, M.: A framework for non-asymptotic quantum information theory. Ph.D. thesis, Department of Physics, ETH Zurich. arXiv:1203.2142
47. Uhlmann, A.: The "transition probability" in the state space of a *-algebra. Rep. Math. Phys. **9**, 273–278 (1976)
48. Uhlmann, A.: Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory. Commun. Math. Phys. **54**, 21–32 (1977)
49. Umegaki, H.: Conditional expectation in an operator algebra. Kodai Math. Sem. Rep. **14**, 59–85 (1962)
50. Wilde, M.M., Winter, A., Yang, D.: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels. Commun. Math. Phys. **331**(2), 593–622 (2014)
51. Winter, A.: Private communication (2012)

Communicated by M. M. Wolf