Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies

Milán Mosonyi\textsuperscript{1,2,3,} and Tomohiro Ogawa\textsuperscript{4,}\textsuperscript{†}

\textsuperscript{1}School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.
\textsuperscript{2}Física Teórica: Informació i Fenomens Quàntics, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain.
\textsuperscript{3}Mathematical Institute, Budapest University of Technology and Economics, Egyház József u 1., Budapest, 1111 Hungary.
\textsuperscript{4}Graduate School of Information Systems, University of Electro-Communications, 1-5-1 Chofugoka, Chofu-shi, Tokyo, 182-8585, Japan.

We show that the new quantum extension of Rényi’s $\alpha$-relative entropies, introduced recently by Müller-Lennert, Dupuis, Szehr, Fehr and Tomamichel, J. Math. Phys. \textbf{54}, 122203, (2013), and Wilde, Winter, Yang, arXiv:1306.1586, 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 $\alpha$: for $\alpha < 1$, the right choice seems to be the traditional definition $D_{\alpha}^{(\text{old})}(\rho \parallel \sigma) := \frac{1}{\alpha-1} \log \text{Tr} \rho^\alpha \sigma^{1-\alpha}$, whereas for $\alpha > 1$ the right choice is the newly introduced version $D_{\alpha}^{(\text{new})}(\rho \parallel \sigma) := \frac{1}{\alpha-1} \log \text{Tr} (\sigma^{\frac{1-\alpha}{2\alpha}} \rho^{\frac{1-\alpha}{2\alpha}})$.

As a side result, we show that the new Rényi $\alpha$-relative entropies are asymptotically attainable by measurements for $\alpha > 1$, and give a new simple proof for their monotonicity under completely positive trace-preserving maps.

I. INTRODUCTION

Rényi in his seminal paper \cite{Renyi1961} introduced a generalization of the Kullback-Leibler divergence (relative entropy). According to his definition, the $\alpha$-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

\begin{equation}
D_{\alpha}(p \parallel 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), & \text{supp } p \subseteq \text{supp } q \ or \ \alpha \in [0,1), \\
+\infty, & \text{otherwise}.
\end{cases}
\end{equation}

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. \cite{Dupuis2009} for an overview of these results. It was also shown in \cite{Dupuis2009} 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 $D_{\alpha}$ is an extension of (1) to pairs of positive semidefinite operators, then it should satisfy

\begin{equation}
D_{\alpha}(\rho \parallel \sigma) \geq 0 \quad \text{and} \quad \lim_{\alpha \to 1} D_{\alpha}(\rho \parallel \sigma) = D(\rho \parallel \sigma) = 0 \iff \rho = \sigma \quad \text{(positivity)}
\end{equation}

for any density operators $\rho, \sigma$ and $\alpha > 0$, and if $\Phi$ is a CPTP map then

\begin{equation}
D_{\alpha}(\Phi(\rho) \parallel \Phi(\sigma)) \leq D_{\alpha}(\rho \parallel \sigma) \quad \text{(monotonicity)}
\end{equation}
should hold.

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

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

Hölder’s inequality ensures positivity of $D_{\alpha}^{(\text{old})}$ for every $\alpha > 0$. Monotonicity has been proved for $\alpha \in [0, 2] \setminus \{1\}$ with various methods [20, 38, 14], but it doesn’t hold for $\alpha > 2$ in general, as it was noted, e.g., in [30]. Monotonicity under measurements, however, is still true for $\alpha > 2$ [17]. In the limit $\alpha \to 1$, these divergences yield Umegaki’s relative entropy [45] in a state discrimination problem, as the optimal exponential decay rates of the two error probabilities in binary state discrimination. This in turn is based on the so-called quantum Hoeffding bound theorem, that quantifies bounds on the type II error under the assumption that the type I error goes to 0 (see section IV A (which we will call simply relative entropy for the rest) in [19] 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 [28] that, similarly to the classical case, the Rényi $\alpha$-relative entropies $D_{\alpha}^{(\text{old})}$ 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, that quantifies the trade-off between the optimal exponential decay rates of the two error probabilities in binary state discrimination [3, 18, 22, 32]. 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 $\rho$ or all prepared in state $\sigma$) then the optimal type I error goes to 0 exponentially fast with the exponent given by the Hoeffding divergence

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

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

Recently, a new quantum extension of the Rényi $\alpha$-relative entropies have been proposed in [30, 46], defined as

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

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 [30, 46] and extended to $\alpha \in (1, +\infty)$ in [3] 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 [30] that these new Rényi relative entropies are not monotone for $\alpha \in (0, \frac{1}{2})$. Monotonicity has been proved for $\alpha \in (\frac{1}{2}, 1) \cup (1, +\infty)$. The Araki-Lieb-Thirring inequality [1, 27] (see also [6, Theorem I.2.10]) implies that

$$D_{\alpha}^{(\text{new})} (\rho \| \sigma) \leq D_{\alpha}^{(\text{old})} (\rho \| \sigma)$$

for every $\rho, \sigma$ and $\alpha \in (0, +\infty) \setminus \{1\}$. The converse Araki-Lieb-Thirring inequality of [4] implies lower bounds on $D_{\alpha}^{(\text{new})}$ in terms of $D_{\alpha}^{(\text{old})}$ [29].

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 IV.9) 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 0 exponentially fast, with the exponent given by the converse Hoeffding divergence

$$H^*_r (\rho |\sigma) := \sup_{1 < \alpha} \frac{\alpha - 1}{\alpha} \left[ r - D_{\alpha}^{(\text{new})} (\rho \| \sigma) \right]. \quad (7)$$
From this, we derive (in Theorem [[14.13]]) 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 proof we only use the monotonicity of the new Rényi relative entropies under pinching [30, Proposition 13], and show (in Theorem [[11.7]]) that the new Rényi relative entropies can be asymptotically attained by measurements, similarly to the relative entropy [19]. Based on this, we provide a simple new proof for the monotonicity of $D_\alpha$ under CPTP maps for $\alpha > 1$ as a side-result.

Our results suggest that, somewhat surprisingly, the right formula to define the Rényi $\alpha$-relative entropies for quantum states depends on whether the parameter $\alpha$ 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) $\rho, \sigma$ as

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

II. 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, $0 \leq M_i$, $i \in \mathcal{I}$, 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 $\{ A \geq 0 \}$, $\{ A < 0 \}$, and $\{ A \leq 0 \}$ are defined similarly. The positive part of $A$ is defined as

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

and it is easy to see that

$$\Tr A_+ = \Tr \{ A > 0 \} = \max_{0 \leq T \leq I} \Tr AT \geq 0.$$  \hfill (9)

In particular, if $\rho_n$ and $\sigma_n$ are self-adjoint operators then for any $a \in \mathbb{R}$ the application of (9) to $A = \rho_n - e^{na} \sigma_n$ yields

$$\Tr \rho_n \{ \rho_n - e^{na} \sigma_n > 0 \} \geq e^{na} \Tr \sigma_n \{ \rho_n - e^{na} \sigma_n > 0 \}.  \hfill (10)$$

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

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

In particular, we have the following lemma.

**Lemma II.1** Let $\rho_n$ and $\sigma_n$ be self-adjoint operators and $\mathcal{F}$ be a positive trace-preserving map. Then for any $a \in \mathbb{R}$,

$$\Tr (\rho_n - e^{na} \sigma_n)_+ \geq \Tr (\mathcal{F}(\rho_n) - e^{na} \mathcal{F}(\sigma_n))_+. \hfill (11)$$

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

$$\mathcal{E}_A(B) := \sum_i E_i B E_i, \quad B \in \mathcal{L}(\mathcal{H}).$$
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
\[ \forall C \in \{A\}', \quad \text{Tr } BC = \text{Tr } \mathcal{E}_A(B)C. \tag{13} \]

The following lemma is from [16, 17]:

**Lemma II.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, $\rho$ and $\sigma$ will denote positive semidefinite operators on some finite-dimensional Hilbert space $\mathcal{H}$, and we use the notation
\[ \rho_n := \rho \otimes^n, \quad \sigma_n := \sigma \otimes^n, \quad \hat{\rho}_n := \mathcal{E}_{\sigma_n}(\rho_n), \quad v_n := v(\sigma_n), \tag{14} \]
where $\mathcal{E}_{\sigma_n}$ is the pinching operation corresponding to $\sigma_n$, and $v_n$ denotes the number of different eigenvalues of $\sigma_n$. Note that $v_n \leq (n + 1)^{\dim \mathcal{H}}$, and lemma II.2 yields
\[ \rho_n \leq v_n \hat{\rho}_n \leq (n + 1)^{\dim \mathcal{H}} \hat{\rho}_n. \tag{15} \]
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 $A^p := \sum_i a_i^p P_i$ for any $p \in \mathbb{R}$. In particular, $A^0$ is the projection onto the support of $A$. We will also use the convention $\log 0 := -\infty$.

**III. PROPERTIES OF THE NEW RÉNYI RELATIVE ENTROPIES**

For positive semidefinite operators $\rho$ and $\sigma$ and $\alpha \in \mathbb{R}$, let
\[ F_{\alpha}(\rho||\sigma) := \log \text{Tr} \left( \sigma^{\frac{1-\alpha}{\alpha}} \rho \sigma^{\frac{\alpha}{1-\alpha}} \right)^{\alpha}. \tag{16} \]
For a POVM $M = \{M_x\}_x$, we can consider the corresponding classical quantity as
\[ F^M_{\alpha}(\rho||\sigma) := \log \left( \sum_x (\text{Tr } \rho M_x)^{\alpha} (\text{Tr } \sigma M_x)^{1-\alpha} \right). \tag{17} \]

Note that for states $\rho$ and $\sigma$ such that $\text{supp } \rho \subseteq \text{supp } \sigma$, the $\frac{1}{\alpha-1} F_{\alpha}(\rho||\sigma)$ is the new Rényi $\alpha$-relative entropy defined in [11], and the $\frac{1}{\alpha-1} F^M_{\alpha}(\rho||\sigma)$ is the post-measurement Rényi $\alpha$-relative entropy.

In this section we show that for every $\alpha > 1$, the new Rényi $\alpha$-relative entropies are asymptotically attainable by measurements in the limit of infinitely many copies of $\rho$ and $\sigma$; for this we only use that the new Rényi $\alpha$-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 $\alpha$-relative entropies.

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

**Lemma III.1 (classical monotonicity)** Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})_+$ be commuting operators such that $\text{supp } \rho \subseteq \text{supp } \sigma$, 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 $\alpha > 1$, $F_{\alpha}(\mathcal{F}(\rho)||\mathcal{F}(\sigma)) \leq F_{\alpha}(\rho||\sigma)$.
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 be found e.g. in \[23\] Proposition A.3.

The following has been shown in \[30\] Proposition 13. We reproduce the proof here for readers' convenience.

**Lemma III.2 (monotonicity under pinching)** Let \( \rho, \sigma \in \mathcal{L}(\mathcal{H})_+ \) and \( \alpha \geq 1 \). Then

\[
\text{Tr} \left( \sigma \frac{1}{\alpha} \mathcal{E}_\sigma(\rho)\sigma \frac{1}{\alpha} \right) \leq \text{Tr} \left( \sigma \frac{1}{\alpha} \rho \sigma \frac{1}{\alpha} \right) \alpha.
\]

Proof: It is easy to see that \( \frac{1}{\alpha} \mathcal{E}_\sigma(\rho)\sigma \frac{1}{\alpha} = \mathcal{E}_\sigma \left( \sigma \frac{1}{\alpha} \rho \sigma \frac{1}{\alpha} \right) \), and Problem II.5.5 with Theorem II.3.1 in \[8\], applied to the convex function \( f(t) = t^\alpha \), yields the assertion.

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

**Lemma III.3 (monotonicity under measurements)** Let \( \rho, \sigma \in \mathcal{L}(\mathcal{H})_+ \) be such that \( \text{supp} \rho \subseteq \text{supp} \sigma \). For any POVM \( M = \{M_x\}_x \in \mathcal{M}(\mathcal{H}) \), we have

\[
F^M_\alpha (\rho\|\sigma) \leq F_\alpha (\rho\|\sigma), \quad \alpha \geq 1.
\]

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

\[
\sum_x \left( \text{Tr} \rho_n M_n(x) \right)^\alpha \left( \text{Tr} \sigma_n M_n(x) \right)^{1-\alpha} \leq v_n^\alpha \sum_x \left( \text{Tr} \rho_n M_n(x) \right)^\alpha \left( \text{Tr} \sigma_n M_n(x) \right)^{1-\alpha}
\]

\[
\leq v_n^\alpha \text{Tr} \rho_n^\alpha \sigma_n^{1-\alpha}
\]

\[
\leq v_n^\alpha \text{Tr} \left( \sigma_n^\alpha \rho_n \sigma_n^{1-\alpha} \right) \alpha,
\]

where the first inequality is due to \(15\), the second inequality follows from Lemma III.2, and the third one from Lemma III.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 \( n \)th i.i.d. extension, i.e.,

\[
M_n(x) := M_{x_1} \otimes \ldots \otimes M_{x_n}, \quad x \in \mathcal{X}^n.
\]

Then we obtain

\[
\left( \sum_x \left( \text{Tr} \rho M_x \right)^\alpha \left( \text{Tr} \sigma M_x \right)^{1-\alpha} \right)^n = \sum_x \left( \text{Tr} \rho M_n(x) \right)^\alpha \left( \text{Tr} \sigma M_n(x) \right)^{1-\alpha}
\]

\[
\leq v_n^\alpha \text{Tr} \left( \sigma_n^\alpha \rho_n \sigma_n^{1-\alpha} \right) \alpha
\]

\[
= v_n^\alpha \left( \text{Tr} \left( \sigma_n^\alpha \rho_n \sigma_n^{1-\alpha} \right) \right)^n.
\]

Taking the logarithm yields

\[
F^M_\alpha (\rho\|\sigma) \leq F_\alpha (\rho\|\sigma) + \frac{\alpha}{n} \log v_n,
\]

which proves the lemma by taking the limit \( n \to \infty \).

Remark III.4 The technique used in the proof of the above lemma is essentially due to \[17\] (see around page 88), where the inequalities \(20\) and \(21\) have been shown.

Remark III.5 Note that the assumption \( \text{supp} \rho \subseteq \text{supp} \sigma \) was necessary to apply classical monotonicity in \(21\). In fact, the statement of Lemma III.2 need not hold without this assumption. Indeed, in the extreme case where \( \rho \) and \( \sigma \) have orthogonal supports, we have \( F_\alpha (\rho\|\sigma) = -\infty \), and the trivial POVM \( M = \{I\} \) yields \( F^M_\alpha (\rho\|\sigma) = \log(\text{Tr} \rho^\alpha (\text{Tr} \sigma)^{1-\alpha}) \), which is a finite number unless \( \rho \) or \( \sigma \) is equal to 0.
The following lemma is standard:

**Lemma III.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 $\text{Tr} f(A) \leq \text{Tr} f(B)$. In particular,

$$0 \leq A \leq B \implies \text{Tr} A^\alpha \leq \text{Tr} B^\alpha \quad \alpha > 0.$$ 

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

**Theorem III.7** (asymptotic attainability) Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ be such that $\text{supp} \rho \subseteq \text{supp} \sigma$. For any $\alpha \geq 1$, we have

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

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

**Proof:** Since $\sigma_n$ and $\hat{\rho}_n$ commute, there is a projection-valued measure $E_n = \{E_{n}(1), \ldots, E_{n}(\dim \mathcal{H})\}$, with all projections of rank 1, that jointly diagonalizes both operators, and hence

$$\frac{1}{n} F_\alpha(\hat{\rho}_n||\sigma_n) = \frac{1}{n} F_{\alpha}^E(\rho_n||\sigma_n) \leq \frac{1}{n} \max_{M_n} F_{\alpha}^M(\rho_n||\sigma_n) \leq \frac{1}{n} F_\alpha(\rho_n||\sigma_n) = F_\alpha(\rho||\sigma),$$

where the last inequality is due to Lemma III.3. By Lemma II.2

$$0 \leq \frac{\text{Tr} \rho_n \text{Tr} \sigma_n}{\text{Tr} \rho_n \text{Tr} \sigma_n} \leq \frac{1}{n} \frac{1}{n} \frac{1}{n} \ldots \frac{1}{n} \frac{1}{n} \leq \frac{1}{n} \frac{1}{n} \frac{1}{n} \ldots \frac{1}{n} \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{\dim \mathcal{H}} (\text{Tr} \rho_n E_n(i)) \frac{1}{n} \frac{1}{n} \frac{1}{n} \ldots \frac{1}{n} \frac{1}{n} \frac{1}{n} E_n(i),$$

and Lemma III.6 yields

$$\text{Tr} \left( \frac{\rho_n \rho_n}{\sigma_n \sigma_n} \right)^\alpha \leq \left( \frac{\text{Tr} \rho_n \text{Tr} \sigma_n}{\text{Tr} \rho_n \text{Tr} \sigma_n} \right)^\alpha \leq \frac{1}{n} \frac{1}{n} \frac{1}{n} \ldots \frac{1}{n} \frac{1}{n} \frac{1}{n} \frac{1}{n} \frac{1}{n} \frac{1}{n} \frac{1}{n} \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{\dim \mathcal{H}} (\text{Tr} \rho_n E_n(i)) \frac{1}{n} \frac{1}{n} \frac{1}{n} \ldots \frac{1}{n} \frac{1}{n} \frac{1}{n} E_n(i).$$

Taking the logarithm, we obtain

$$F_\alpha(\rho||\sigma) \leq \frac{1}{n} F_\alpha(\hat{\rho}_n||\sigma_n) + \frac{\log \text{Tr} \rho_n \text{Tr} \sigma_n}{n} \leq \frac{1}{n} \max_{M_n} F_{\alpha}^M(\rho_n||\sigma_n) + \frac{\alpha}{n} \log \text{Tr} \rho_n \text{Tr} \sigma_n,$$

Combining this with (27), and taking the limit $n \to +\infty$, the assertion follows. 

**Theorem III.7** implies the asymptotic attainability for the Rényi relative entropies:

**Corollary III.8** For any $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ and $\alpha \geq 1$, we have

$$D_\alpha^{(\text{new})}(\rho||\sigma) = \lim_{n \to \infty} \frac{1}{n} D_\alpha^{(\text{new})}(\hat{\rho}_n||\sigma_n) = \frac{1}{n} \max_{M_n \in \mathcal{M}(\mathcal{H}^\otimes n)} D_\alpha^{(\text{new})}(\{\text{Tr} \rho_n M_n(x)\}_{x \in \mathcal{X}}||\{\text{Tr} \sigma_n M_n(x)\}_{x \in \mathcal{X}}),$$

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

**Proof:** The case where $\text{supp} \rho \subseteq \text{supp} \sigma$ is immediate from Theorem III.7. On the other hand, if $\text{supp} \rho \not\subseteq \text{supp} \sigma$ then also $\text{supp} \hat{\rho}_n \not\subseteq \text{supp} \sigma_n$, and hence, by the definition (30), $D_\alpha^{(\text{new})}(\rho||\sigma) = D_\alpha^{(\text{new})}(\hat{\rho}_n||\sigma_n) = \max_{M_n \in \mathcal{M}(\mathcal{H}^\otimes n)} D_\alpha^{(\text{new})}(\{\text{Tr} \rho_n M_n(x)\}_{x \in \mathcal{X}}||\{\text{Tr} \sigma_n M_n(x)\}_{x \in \mathcal{X}}) = +\infty$ for every $n \in \mathbb{N}$, making the assertion trivial. 

□
Remark III.9 The same statement for the relative entropy has been shown in [19].

Theorem III.7 has a number of important corollaries:

Corollary III.10 (convexity) For any fixed $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ such that $\text{supp} \rho \subseteq \text{supp} \sigma$, $F_\alpha(\rho \| \sigma)$ is a convex function of $\alpha$ for $\alpha \geq 1$.

Proof: It is easy to see (by computing its second derivative) that $F_\alpha(\rho \| \sigma) = \frac{F_\alpha(\rho) - F_\alpha(\sigma)}{\alpha - 1}$ is a pointwise limit of convex functions, and hence it is convex. □

Corollary III.11 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 $\text{supp} \rho \subseteq \text{supp} \sigma$, since otherwise $D_\alpha^{(\text{new})}(\rho \| \sigma) = +\infty$ for every $\alpha > 1$, and the assertion holds trivially. Note that $\text{supp} \rho \subseteq \text{supp} \sigma$ implies that $F_1(\rho \| \sigma) = \log \text{Tr} \rho$, and hence $D_\alpha^{(\text{new})}(\rho \| \sigma) = \frac{F_\alpha(\rho) - F_1(\rho \| \sigma)}{\alpha - 1}$. The assertion then follows from Corollary III.10. □

Corollary III.12 (monotonicity) Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ be such that $\text{supp} \rho \subseteq \text{supp} \sigma$, and let $\mathcal{F} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ 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, $\mathcal{F}_n := \mathcal{F} \otimes \mathcal{I}$ is positive for every $n \in \mathbb{N}$. Let $\mathcal{F}_n^* : \mathcal{L}(\mathcal{K}^\otimes n) \rightarrow \mathcal{L}(\mathcal{H}^\otimes n)$ be the dual (adjoint) of $\mathcal{F}_n$, defined by

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

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

$$\max_{M \in \mathcal{M}(\mathcal{K}^\otimes n)} F_\alpha^M(\mathcal{F}_n(\rho_n) \| \mathcal{F}_n(\sigma_n)) = \max_{M \in \mathcal{M}(\mathcal{K}^\otimes n)} F_\alpha^M(\rho_n \| \sigma_n) \leq \max_{M \in \mathcal{M}(\mathcal{H}^\otimes n)} F_\alpha^M(\rho_n \| \sigma_n)$$

for any $n$. Now (33) and Theorem III.7 yield the assertion. □

Corollary III.12 immediately implies the following:

Corollary III.13 The new Rényi relative entropies are monotone under CPTP maps for $\alpha > 1$. That is, if $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ and $\mathcal{F} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is a CPTP map then

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

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

Remark III.14 Note that the above monotonicity property holds 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.

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

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

This is an analogy of the quasi-entropy [23] (or quantum $f$-divergence [23]) corresponding to the function $x \mapsto x^\alpha$. However, $Q_\alpha^{(\text{new})}$ cannot be written in the form of an $f$-divergence [23, Corollary 2.10]. Corollary III.12 is equivalent to the monotonicity of $Q$:

Corollary III.15 (monotonicity of $Q$) Let $\rho, \sigma \in \mathcal{L}(\mathcal{H})_+$ be such that $\text{supp} \rho \subseteq \text{supp} \sigma$, and let $\mathcal{F} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ be a CPTP map. Then

$$Q_\alpha^{(\text{new})}(\mathcal{F}(\rho) \| \mathcal{F}(\sigma)) \leq Q_\alpha^{(\text{new})}(\rho \| \sigma), \quad \alpha > 1.$$
Corollary III.16 (joint convexity) Let \( \rho_i, \sigma_i \in \mathcal{L}(\mathcal{H})_+ \) be such that \( \text{supp} \rho_i \subseteq \text{supp} \sigma_i, i = 1, \ldots, r \), and let \( p_1, \ldots, p_r \) be a probability distribution. Then

\[
Q^{(\text{new})}_\alpha \left( \sum_{i=1}^r p_i \rho_i \| \sum_{i=1}^r p_i \sigma_i \right) \leq \sum_{i=1}^r p_i Q^{(\text{new})}_\alpha (\rho_i \| \sigma_i).
\]

Proof: Let \( \delta_1, \ldots, \delta_r \) be orthogonal rank 1 projections on \( K := \mathbb{C}^r \), 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 \( F := \text{Tr}_K \) to be the partial trace over \( K \) in Corollary III.15, the assertion follows.

Remark III.17 In Corollary III.16 we obtained the joint convexity from the monotonicity of \( Q^{(\text{new})}_\alpha \). In [13] (and also in [30, 46] for \( \alpha \in (1, 2) \)) the authors followed the opposite approach: they first established joint convexity of \( Q^{(\text{new})}_\alpha \), 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 operations.

IV. STRONG CONVERSE EXPONENT IN QUANTUM HYPOTHESIS TESTING

A. Simple Quantum Hypothesis Testing

We study the simple hypothesis testing problem for the null hypothesis \( H_0: \rho_n \) versus the alternative hypothesis \( H_1: \sigma_n \), where \( \rho_n = \rho^{\otimes n} \) and \( \sigma_n = \sigma^{\otimes n} \) are the \( n \)-fold tensor products of arbitrarily given density operators \( \rho \) and \( \sigma \) 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

\[
\alpha_n(T_n) := \text{Tr} \rho_n (I_n - T_n),
\]
\[
\beta_n(T_n) := \text{Tr} \sigma_n T_n.
\]

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,

\[
\beta_n^*(\epsilon) := \min \left\{ \beta_n(T_n) \mid T_n : \text{test}, \alpha_n(T_n) \leq \epsilon \right\}.
\]

The quantum Stein’s lemma [19, 30] states that for all \( \epsilon \in (0, 1) \),

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

where \( D(\rho\|\sigma) \) is the quantum relative entropy given in [4]. This implies the existence of a sequence of tests \( \{ T_n \} \subseteq \mathbb{N} \) such that

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

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 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

\[ B_c(r) := \sup \left\{ -\limsup_{n \to \infty} \frac{1}{n} \log \alpha_n(T_n) \mid \limsup_{n \to \infty} \frac{1}{n} \log \beta_n(T_n) \leq -r \right\} \]

\[ = \sup \left\{ R \mid \exists \{T_n\}_{n=1}^\infty 0 \leq T_n \leq I_n, \text{ s.t.} \right. \]

\[ \limsup_{n \to \infty} \frac{1}{n} \log \beta_n(T_n) \leq -r, \ \limsup_{n \to \infty} \frac{1}{n} \log \alpha_n(T_n) \leq -R \}, \] (39)

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 [13, 32] that

\[ B_c(r) = \sup_{0 \leq s < 1} -sr - \log \text{Tr} \rho^{1-s} \sigma^s = \sup_{0 < \alpha < 1} \frac{\alpha - 1}{\alpha} \left[ r - D_\alpha^{\text{(old)}}(\rho \parallel \sigma) \right] = H_r(\rho \parallel \sigma), \] (40)

where \( D_\alpha^{\text{(old)}} \) is the traditional definition of the quantum Rényi relative entropy, given in [3], and \( H_r(\rho \parallel \sigma) \) is the Hoeffding divergence defined in [4]. (Note that the roles of the type I and the type II errors are reversed here as compared to some previous work the Hoeffding bound, and hence our \( H_r(\rho \parallel \sigma) \) corresponds to \( H_r(\sigma \parallel \rho) \) in those works. ) It can be shown that \( B_c(r) > 0 \) when \( 0 < r < D(\rho \parallel \sigma) \), and \( \alpha_n(T_n) \) goes to zero exponentially with the rate \( B_c(r) \) for an optimal sequence of tests \( \{T_n\}_{n=1}^\infty \).

On the other hand, if \( \text{supp} \rho \subseteq \text{supp} \sigma \) and \( \beta_n(T_n) \leq e^{-nr} \) with \( r > D(\rho \parallel \sigma) \) then \( \alpha_n(T_n) \) inevitably goes to 1 exponentially fast [32]; this is called the strong converse property. In this case, we are interested in determining the exponent with which the success probabilities \( 1 - \alpha_n(T_n) = \text{Tr} \rho_n T_n \) go to zero. The optimal such exponent is the strong converse exponent \( B_c^*(r) \); formally,

\[ B_c^*(r) := \inf \left\{ -\liminf_{n \to +\infty} \frac{1}{n} \log \text{Tr} \rho_n T_n \mid \limsup_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n \leq -r \right\}, \] (41)

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 \( B_c^*(r) \) can be alternatively written as

\[ B_c^*(r) = \sup \left\{ R \mid \forall \{T_n\}_{n=1}^\infty, 0 \leq T_n \leq I_n, \ 
\text{limsup}_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n \leq -r \Rightarrow \liminf_{n \to +\infty} \frac{1}{n} \log \text{Tr} \rho_n T_n \leq -R \right\} \]

\[ = \inf \left\{ R \mid \exists \{T_n\}_{n=1}^\infty, 0 \leq T_n \leq I_n, \ 
\text{limsup}_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n T_n \leq -r, \ \text{liminf}_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n T_n \geq -R \right\} \] (42)

The main result of Section IV is Theorem IV.9, where we show that, in complete analogy with [40],

\[ B_c^*(r) = \sup_{0 < \alpha < 1} \frac{\alpha - 1}{\alpha} \left[ r - D_\alpha^{\text{(new)}}(\rho \parallel \sigma) \right] = H_r^*(\rho \parallel \sigma), \] (43)

where \( H_r^*(\rho \parallel \sigma) \) is the converse Hoeffding divergence [41]. The inequality \( B_c^*(r) \geq H_r^*(\rho \parallel \sigma) \) follows easily from the monotonicity of the Rényi divergences, as we show in Lemma IV.9. 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 [13, 32]. We start with this problem in Section IV.B

Remark IV.1 Note that if \( \text{supp} \rho \subseteq \text{supp} \sigma \) 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 \( \beta_n(T_n) = 0 \leq e^{-nr}, r > 0, \) and \( \alpha_n(T_n) = (\text{Tr} \rho \sigma^0)^n, n \in \mathbb{N}, \) which converges to zero exponentially fast with an exponent \( -\log \text{Tr} \rho \sigma^0 > 0. \) Hence, for the rest we will assume that \( \text{supp} \rho \subseteq \text{supp} \sigma. \)
B. Exponents for the Neyman-Pearson tests

Let \( \rho \) and \( \sigma \) be quantum states such that

\[
\text{supp } \rho \subseteq \text{supp } \sigma, \tag{44}
\]

and let \( \rho_n, \sigma_n, \) etc. be defined as in [14]. To exclude a trivial case, we assume that \( \rho \neq \sigma \). Let us define the quantum Neyman-Pearson tests by

\[
S_n(a) := \{ \rho_n - e^{na} \sigma_n > 0 \}, \tag{45}
\]

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 \( \text{Tr } \rho_n S_n, a \) and the type II error probabilities \( \text{Tr } \sigma_n S_n, a \). Note that

\[
S_n(a) = 0 \iff a \geq D_{\text{max}}(\rho \| \sigma) := \inf \{ \gamma : \rho \leq e^{\gamma} \sigma \}. \tag{46}
\]

Here \( D_{\text{max}}(\rho \| \sigma) \) is the max-relative entropy [11, 42], and it was shown in [30, Theorem 4] that

\[
D_{\text{max}}(\rho \| \sigma) = \lim_{\alpha \to +\infty} D_{\alpha}^{(\text{new})}(\rho \| \sigma). \tag{47}
\]

Thus,

\[
\text{Tr } \rho_n S_n, a = \text{Tr } \sigma_n S_n, a = 0, \quad a \geq D_{\text{max}}(\rho \| \sigma),
\]

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

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

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

For every \( s \in \mathbb{R} \), let

\[
\psi(s) := F_{s+1}(\rho \| \sigma) = \log \text{Tr } (\sigma^{\frac{s}{s+\alpha}} e^{na} \rho \sigma^{\frac{1}{s+\alpha}})^{s+1}, \tag{48}
\]

and

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

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

**Lemma IV.2** We have

\[
\psi(0) = 0, \tag{49}
\]

\[
\psi'(0) = D(\rho \| \sigma), \tag{50}
\]

\[
\lim_{s \to +\infty} \psi'(s) = D_{\text{max}}(\rho \| \sigma), \tag{51}
\]

and

\[
\phi(a) \begin{cases} 
0, & a \leq D(\rho \| \sigma) \\
> 0, & D(\rho \| \sigma) < a \leq D_{\text{max}}(\rho \| \sigma), \\
+\infty, & D_{\text{max}}(\rho \| \sigma) < a. \end{cases} \tag{52}
\]

**Proof:** The identity in [49] is immediate from the definition of \( \psi \). \( \psi(0) = 0 \) yields \( \psi'(0) = \lim_{s \to 0} \frac{1}{s} \psi(s) = \lim_{s \to 1} D_{\alpha}(\rho \| \sigma) = D(\rho \| \sigma) \), where the last identity is due to [31, Theorem 4]. Using again [30, Theorem 4] and the L’Hospital rule, \( \lim_{s \to +\infty} \psi'(s) = \lim_{s \to +\infty} \frac{1}{s} \psi(s) = \lim_{s \to +\infty} D_{\alpha}(\rho \| \sigma) = D_{\text{max}}(\rho \| \sigma) \). By Corollary [11.10] \( s \mapsto \psi(s) \) is convex, and hence [52] follows immediately from [49]–[51]. \( \square \)

**Lemma IV.3** For any \( a \in \mathbb{R} \) and \( n \in \mathbb{N} \), we have

\[
\frac{1}{n} \log \text{Tr } \rho_n S_n(a) \leq -\phi(a), \tag{53}
\]

\[
\frac{1}{n} \log \text{Tr } \sigma_n S_n(a) \leq -\{a + \phi(a)\}. \tag{54}
\]
Proof: For any $a \in \mathbb{R}$ and $s \geq 0$, we have

\[
\text{Tr} \, \rho_n S_n(a) = \{\text{Tr} \, \rho_n S_n(a)\}^{s+1} \{\text{Tr} \, \rho_n S_n(a)\}^{-s} \\
\leq e^{-nas} \{\text{Tr} \, \rho_n S_n(a)\}^{s+1} \{\text{Tr} \, \sigma_n S_n(a)\}^{-s} \\
\leq e^{-nas} \left[\left\{\text{Tr} \, \rho_n S_n(a)\right\}^{s+1} \{\text{Tr} \, \sigma_n S_n(a)\}^{-s} \right] \\
+ \{\text{Tr} \, \rho_n(I_n - S_n(a))\}^{s+1} \{\text{Tr} \, \sigma_n(I_n - S_n(a))\}^{-s} \\
\leq e^{-nas} \text{Tr} \left(\frac{\sigma}{n^{s+1}} \rho_n \sigma_n^{s+1}\right)^{s+1} \\
= e^{-nas} e^{na(s)} ,
\]

(55)

where in the first inequality we used (10), the second inequality is trivial, and the last inequality follows from Lemma III.3. Taking the logarithm and the infimum in $s$ yields the inequality in (55).

Using (10) and (55), we get

\[
\text{Tr} \, \sigma_n S_n(a) \leq e^{-na} \text{Tr} \, \rho_n S_n(a) \leq e^{-na(s+1)} e^{na(s)} ,
\]

(56)

which yields (54). \qed

Note that the bounds in (55) and (56) are trivial for $a \geq D_{\max}(\rho \parallel \sigma)$, due to (10). For $a \leq D(\rho \parallel \sigma)$ we have $\phi(a) = 0$ (cf. (52)), and hence the upper bound in (55) is trivial in this range. More detailed information about the values of $\text{Tr} \, \sigma_n S_n(a)$ in this range is given in the setting of the Hoeffding bound; Corollary 4.5 in [22] states that

\[
\lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \, \sigma_n S_n(a) = - \sup_{0 \leq t \leq 1} \{at - \log \text{Tr} \rho \sigma^{-1-t}\} \leq -a = -\{\phi(a) + a\} , \quad a < D(\rho \parallel \sigma).
\]

Theorems IV.4 and IV.5 below show that the inequalities in (55) and (56) hold asymptotically as an equality in the non-trivial range $D(\rho \parallel \sigma) < a < D_{\max}(\rho \parallel \sigma)$.

**Theorem IV.4** For any $a \in (D(\rho \parallel \sigma), D_{\max}(\rho \parallel \sigma))$, we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \, \rho_n S_n(a) = \lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \left(\rho_n - e^{na} \sigma_n\right)_+ = -\phi(a).
\]

(57)

**Proof:** For a fixed $m \in \mathbb{N}$, let $\hat{\rho}_m := E_{\sigma_m}(\rho_m)$, and define

\[
\hat{S}_{m,k}(a) := \left\{\hat{\rho}_m^{\otimes k} - e^{km} \sigma_m^{\otimes k} > 0\right\}.
\]

(58)

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

\[
\text{Tr} \, \rho_n S_n(a) = \text{Tr} \left(\rho_n - e^{na} \sigma_n\right)_+ S_n(a) + e^{na} \text{Tr} \, \sigma_n S_n(a) \\
\geq \text{Tr} \left(\rho_n - e^{na} \sigma_n\right)_+ \\
\geq \text{Tr} \left(\hat{\rho}_m^{\otimes k} - e^{km} \sigma_m^{\otimes k}\right)_+ \\
\geq \text{Tr} \left(\hat{\rho}_m^{\otimes k} - e^{km} \sigma_m^{\otimes k}\right) \hat{S}_{m,k}(b) \\
\geq \text{Tr} \hat{\rho}_m^{\otimes k} \hat{S}_{m,k}(b) - e^{ka} \text{Tr} \hat{\rho}_m^{\otimes k} \hat{S}_{m,k}(b) \\
= \left\{1 - e^{ka} e^{-km(b-a)}\right\} \text{Tr} \hat{\rho}_m^{\otimes k} \hat{S}_{m,k}(b),
\]

(59)

where (59) follows from Lemma III.3 (with the choice $F := E_{\sigma_m} \otimes \text{Tr}_{km+1,r}$), (60) follows from (59), and we used (10) in (61). Hence, by choosing $b > a$, we get

\[
\phi(a) \geq \lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n S_n(a) \geq \liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n S_n(a) \\
\geq \liminf_{k \to \infty} \frac{1}{k} \log \text{Tr} \rho_n S_n(a) \geq \frac{1}{m} \liminf_{k \to \infty} \frac{1}{k} \text{Tr} \hat{\rho}_m^{\otimes k} \hat{S}_{m,k}(b),
\]

(63)
where the first inequality is due to \[53\].

Note that \(\hat{\rho}_m\) and \(\sigma_m\) are commuting operators, and hence they can be represented as functions on some finite set. Let

\[
\Psi_m(s) := \log \text{Tr} \hat{\rho}_m e^{s \log \hat{\rho}_m} = \log \text{Tr} \hat{\rho}_m^{1+s} \sigma_m^{-s},
\]

\[
\psi_m(s) := \frac{1}{m} \Psi_m(s) = \frac{1}{m} \log \text{Tr} \hat{\rho}_m^{1+s} \sigma_m^{-s}.
\]

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

\[
m b > m D(\rho || \sigma) = D(\rho_m || \sigma_m) \geq D(\hat{\rho}_m || \sigma_m) = E_{\hat{\rho}_m} \log \frac{\hat{\rho}_m}{\sigma_m} = \Psi'_m(0),
\]

where the first inequality is due to the monotonicity of the quantum relative entropy. Hence, by Cramér’s theorem [12, Theorem 2.1.24], we have

\[
\liminf_{k \to \infty} \frac{1}{k} \log \text{Tr} \hat{\rho}_m^k \hat{S}_{m,k}(b) = \liminf_{k \to \infty} \frac{1}{k} \log \text{Tr} \hat{\rho}_m^k \left\{ \frac{1}{k} \log \frac{\hat{\rho}_m}{\sigma_m^k} > m b \right\}
\]

\[
\geq - \inf_{\kappa > m b} \sup_{s \in \mathbb{R}} \kappa s - \Psi_m(s)
\]

\[
= - \sup_{s \geq 0} \{m b s - \Psi_m(s)\}
\]

\[
= - \sup_{s \geq 0} \{m b s - \psi_m(s)\}
\]

\[
= - m \sup_{s \geq 0} \{b s - \psi_m(s)\}.
\]

Let \(\delta_m := \frac{\log v_m}{m}\). From \[30\], we obtain

\[
\psi(s) \leq \psi_m(s) + (1 + s) \delta_m,
\]

and hence,

\[
\sup_{s \geq 0} \{b s - \psi_m(s)\} \leq \sup_{s \geq 0} \{b s - \psi(s) + (1 + s) \delta_m\}
\]

\[
= \sup_{s \geq 0} \{(b + \delta_m) s - \psi(s)\} + \delta_m
\]

\[
\leq \phi(b + \delta_m) + \delta_m.
\]

Combining \[67\] and \[69\] yields

\[
\frac{1}{m} \liminf_{k \to \infty} \frac{1}{k} \log \text{Tr} \hat{\rho}_m^k \hat{S}_{m,k}(b) \geq - \{\phi(b + \delta_m) + \delta_m\}.
\]

Substituting it back to \[53\], taking the limit \(m \to +\infty\) and using that \(\lim_{m \to +\infty} \delta_m = 0\) and that \(\phi\) is continuous on \((D(\rho || \sigma), D_{\max}(\rho \parallel \sigma))\), we obtain the assertion. \(\square\)

**Theorem IV.5** For any \(a \in (D(\rho || \sigma), D_{\max}(\rho \parallel \sigma))\), we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n S_n(a) = -\{\phi(a) + a\}.
\]

**Proof:** By \[1\], we have

\[
\text{Tr}(\rho_n - e^{nb} \sigma_n)_+ \geq \text{Tr}(\rho_n - e^{nb} \sigma_n) S_n(a)
\]

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

\[
\text{Tr}(\rho_n - e^{nb} \sigma_n)_+ + e^{nb} \text{Tr} \sigma_n S_n(a) \geq \text{Tr} \rho_n S_n(a).
\]

Assume now that \(D(\rho \parallel \sigma) < a < b < D_{\max}(\rho \parallel \sigma)\). Applying Theorem \[IV.4\] to \[73\], we get

\[-\phi(a) = \liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \rho_n S_n(a) \leq \max \left\{-\phi(b), b + \liminf_{n \to \infty} \frac{1}{n} \log \text{Tr} \sigma_n S_n(a)\right\}.
\]
Note that $D (\rho \parallel \sigma) < a < b < D_{\text{max}} (\rho \parallel \sigma)$ implies $\phi (b) > \phi (a)$, and hence we have

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

(74)

Taking $b \searrow a$, we obtain

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

(75)

Now combining (74) and (75) yields the assertion.

C. The strong converse exponent

Consider the hypothesis testing problem from Section IV A. Our aim here is to prove the identity (43), i.e., that the strong converse exponent $B^* (r)$, defined in (41), is equal to the converse Hoeffding bound $H^* (\rho \parallel \sigma)$ defined in (7). We will assume that $\rho \neq \sigma$ to avoid a trivial case, and that $\text{supp} \rho \subseteq \text{supp} \sigma$ so that we actually have a strong converse (cf. Remark IV.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 [31], except that we use the new Rényi divergences instead of the old ones.

**Lemma IV.6** For any $r \geq 0$, we have

$$B^*_c (r) \geq H^*_c (\rho \parallel \sigma).$$

(76)

**Proof:** Let $T_n \in \mathcal{L}(H_n)$ be a test and let $p_n := (\text{Tr} \rho_n T_n, \text{Tr} \rho_n (I - T_n))$ and $q_n := (\text{Tr} \sigma_n T_n, \text{Tr} \sigma_n (I - T_n))$ be the post-measurement states. By the monotonicity of the Rényi relative entropies under measurements (Lemma III.3), we have, for any $\alpha > 1$,

$$D^{(\text{new})}_\alpha (\rho_n \parallel \sigma_n) \geq D^{(\text{new})}_\alpha (p_n \parallel q_n) \geq \frac{1}{\alpha - 1} \log [(\text{Tr} \rho_n T_n)^\alpha (\text{Tr} \sigma_n T_n)^{1-\alpha}] = \frac{\alpha}{\alpha - 1} \log (1 - \alpha_n (T_n)) - \log \beta_n (T_n),$$

or equivalently,

$$\frac{1}{n} \log (1 - \alpha_n (T_n)) \leq \frac{\alpha - 1}{\alpha} \left[ D^{(\text{new})}_\alpha (\rho \parallel \sigma) + \frac{1}{n} \log \beta_n (T_n) \right].$$

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

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

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

□

**Remark IV.7** Using that the old Rényi relative entropies are also monotonic under measurements [14], exactly the same argument as above yields that

$$B^*_c (r) \geq \sup_{1 < \alpha} \frac{\alpha - 1}{\alpha} \left[ r - D^{(\text{old})}_\alpha (\rho \parallel \sigma) \right].$$

(77)

This was already pointed in [36] with a restricted optimization over $\alpha \in (1, 2]$, and later extended by Hayashi to the above form [14].

Our goal in the rest of the section is to show that (77) holds as an equality. To start with, we give some alternative expressions for $H^*_c (\rho \parallel \sigma)$. Let

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

(78)
Proof. First, we consider the case \( s \geq 0 \) and equality holds in the above inequality for some \( a \). Lemma IV.8 For any \( \phi \) continuous on \( ( -\infty, a ) \), \( \psi \) convex in the general case then follows the same way as in Corollary III.10. Convexity and (80) yield

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

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

\[
\tilde{\psi}(0) = \psi(0) = 0, \quad \tilde{\psi}'(0) = \psi'(0) = D(\rho \| \sigma),
\]

and

\[
\lim_{u \to 1} \tilde{\psi}'(u) = \lim_{s \to +\infty} (s\tilde{\psi}'(s) - \psi(s)) + \lim_{s \to +\infty} \psi'(s) = \lim_{s \to +\infty} \phi(\psi'(s)) + D_{\max}(\rho \| \sigma) = \phi(a_{\max}) + a_{\max}.
\]

It is also easy to see, by computing the second derivative, that \( \tilde{\psi} \) is convex for commuting \( \rho \) and \( \sigma \); convexity in the general case then follows the same way as in Corollary III.10. Convexity and (80) yield

\[
H^*_s(\rho \| \sigma) = 0, \quad r \leq D(\rho \| \sigma).
\]

Lemma IV.8 For any \( r \geq 0 \), we have

\[
H^*_s(\rho \| \sigma) = \begin{cases} r - a_r = \phi(a_r), & r < \phi(a_{\max}) + a_{\max}, \\ r - D_{\max}(\rho \| \sigma), & r \geq \phi(a_{\max}) + a_{\max}, \end{cases}
\]

where \( a_{\max} \) and \( r_{\max} \) are defined in (79), 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

\[
\phi(a_r) \geq \frac{s r - \psi(s)}{1 + s}, \quad s \geq 0,
\]

with equality for \( s_r \), and hence

\[
\phi(a_r) = \max_{s \geq 0} \frac{s r - \psi(s)}{1 + s}.
\]

Taking into account (79), this proves the assertion.

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

\[
\lim_{s \to +\infty} \frac{s r - \psi(s)}{s + 1} = r - \lim_{s \to +\infty} \frac{\psi(s)}{s + 1} = r - D_{\max}(\rho \| \sigma),
\]

due to (84) Theorem 4]. Hence it is enough to show that

\[
\frac{s r - \psi(s)}{s + 1} \leq r - D_{\max}(\rho \| \sigma)
\]

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

\[
r - a_{\max} \geq \phi(a_{\max}) \geq a_{\max} s - \psi(s)
\]
for every \( s \geq 0 \), from which we obtain
\[
\frac{r + \psi(s)}{s + 1} \geq a_{\text{max}}.
\] (86)

Thus we have
\[
r - a_{\text{max}} \geq r - \frac{r + \psi(s)}{s + 1} = \frac{rs - \psi(s)}{s + 1},
\] (87)

and hence \( H_r^*(\rho\|\sigma) = r - \max(D_r(\rho\|\sigma), \sigma) \), as required.

Now we are ready to prove the identity \( 43 \) for the strong converse exponent.

**Theorem IV.9** For any \( r \geq 0 \), we have
\[
B_r^*(r) = H_r^*(\rho\|\sigma).
\] (88)

**Proof:** Since we have already shown \( B_r^*(r) \geq H_r^*(\rho\|\sigma) \) in Lemma IV.6, we only have to show the converse inequality \( B_r^*(r) \leq H_r^*(\rho\|\sigma) \). Due to the definition \( 12 \) of \( B_r^*(r) \) 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
\[
\limsup_{n \to \infty} \frac{1}{n} \log \tilde{T}_n \leq -r
\] and
\[
\liminf_{n \to \infty} \frac{1}{n} \log Tr \rho_n T_n \geq -R.
\] (89)

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

(i) In the case \( D(\rho\|\sigma) < r < \max \), there exists a unique \( a_r \in (D(\rho\|\sigma), \max(D_r(\rho\|\sigma))) \) satisfying \( r - a_r = \phi(a_r) \), and Theorems IV.4 and IV.5 yield
\[
\lim_{n \to \infty} \frac{1}{n} \log \tilde{T}_n = -(\phi(a_r) + a_r) = -r,
\]
where the last identity is due to Lemma IV.8.

(ii) In the case \( 0 \leq r \leq D(\rho\|\sigma) \), we have \( H_r^*(\rho\|\sigma) = 0 \), according to \( 31 \). For any \( R > 0 \), we can find an \( a \in (D(\rho\|\sigma), \max(D_r(\rho\|\sigma))) \) such that \( 0 < \phi(a) < R \). Note that \( \phi(a) + a > D(\rho\|\sigma) \), and Theorems IV.4 and IV.5 yield
\[
\lim_{n \to \infty} \frac{1}{n} \log \tilde{T}_n = -(\phi(a) + a) < -r,
\]
where the last identity is due to Lemma IV.8.

(iii) In the case \( r \geq \max \), we use a modification of the Neyman-Pearson tests, following the method of the proof of Theorem 4 in \( 33 \). For every \( a, r \in \mathbb{R} \), let
\[
T_n(r, a) := e^{-n(r - a - \phi(a))} S_n(a).
\]
Note that for \( r \geq \max \) we have \( H_r^*(\rho\|\sigma) = r - \max(D_r(\rho\|\sigma)) \) due to Lemma IV.8. Assume now that \( a \in (D(\rho\|\sigma), \max(D_r(\rho\|\sigma))) \). Then \( r > \phi(a) + a \), and hence \( 0 \leq T_n(r, a) \leq 1 \), i.e., \( T_n(r, a) \) is a test, and
\[
\lim_{n \to \infty} \frac{1}{n} \log \tilde{T}_n = -(r + a + \phi(a) - (a + \phi(a)) = -r,
\]
by Theorems IV.4 and IV.5. Now for a given \( R > H_r^*(\rho\|\sigma) = r - \max(D_r(\rho\|\sigma)) \), we can find an \( a \in (D(\rho\|\sigma), \max(D_r(\rho\|\sigma))) \) such that \( r - \max(D_r(\rho\|\sigma)) < r - a < R \), and the assertion follows.
Remark IV.10 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 \[ \lim_{n \to \infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n T_n \leq -r \] holds with $\mathcal{H}_r^*(\rho\|\sigma)$ in place of $R$, and the proof of Theorem IV.9 yields that for this sequence, we actually have
\[
\limsup_{n \to \infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n T_n \leq -r \quad \text{and} \quad \liminf_{n \to \infty} \frac{1}{n} \log \operatorname{Tr} \rho_n T_n = \mathcal{H}_r^*(\rho\|\sigma).
\]
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 IV.9 and in case (ii) this follows from the Hoeffding bound theorem \[18, 32\].

Remark IV.11 The direct region ($0 \leq r < D(\rho\|\sigma)$) 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 $B_c^*(r) \leq \mathcal{H}_r^*(\rho\|\sigma)$ in the case $0 \leq r < D(\rho\|\sigma)$ (case (ii) of the above proof) based on Stein’s lemma rather than our argument. Indeed, applying \[19\] with $a = r$, we have $\operatorname{Tr} \sigma_n S_n(a) \leq e^{-nu}$, and at the same time, the direct part of the quantum Stein’s lemma \[12\] yields $\lim_{n \to \infty} \operatorname{Tr} \rho_n S_n(a) = 1$. Thus,
\[
\limsup_{n \to \infty} \frac{1}{n} \log \operatorname{Tr} \sigma_n S_n(a) \leq -r \quad \text{and} \quad \liminf_{n \to \infty} \frac{1}{n} \log \operatorname{Tr} \rho_n S_n(a) = 0 = \mathcal{H}_r^*(\rho\|\sigma).
\]

Remark IV.12 By Theorem IV.3 and \[10\], we have
\[
B_c^*(r) = \mathcal{H}_r^*(\rho\|\sigma) = \sup_{0 \leq u < 1} \{ru - \tilde{\psi}(u)\},
\]
where $\tilde{\psi}(u)$ is a continuous convex function on $[0, 1)$. Hence, $B_c^*(r)$ is the Legendre-Fenchel transform (polar function) of $\tilde{\psi}$, and the bipolar theorem says that
\[
\sup_{r \geq 0} \{ru - B_c^*(r)\} = \tilde{\psi}(u) = \frac{\alpha - 1}{\alpha} D^{\text{(new)}}_{\alpha}(\rho\|\sigma), \quad \alpha > 1,
\]
where in the last formula we set $\alpha := 1/(1-u)$ and used the definition \[47\] of $\psi$. That is, the new Rényi relative entropies can be expressed essentially as the Legendre-Fenchel transform of the operational quantities $B_c^*(r)$, $r \geq 0$. A more direct operational interpretation is provided in the next section.

Remark IV.13 A possible proof for the following representation of the strong converse exponent:
\[
B_c^*(r) = \max_{s \geq 0} \frac{rs - \lim_{m \to \infty} \psi_m(s)}{s+1}, \quad (91)
\]
where $\psi_m$ is defined in \[85\], has been outlined in Hayashi’s book \[12\], although it seems to have not been fully worked out. Apart from identifying the limit $\lim_{m \to \infty} \psi_m(s)$ as $sD_{\alpha}^{\text{(new)}}(\rho\|\sigma)$, our approach here differs from Hayashi’s proposal 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.

To close the section, we give one more representation of $H_r^*(\rho\|\sigma)$. This is closely related to the information spectrum approach \[33\], 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 IV.14 For any $r \geq 0$, we have
\[
H_r^*(\rho\|\sigma) = \inf_{a \in \mathbb{R}} \max \{\phi(a), r - a\} \quad (92)
\]
and
\[
\inf \{ \max \{\phi(a), r - a\} | D(\rho\|\sigma) < a < D_{\max}(\rho\|\sigma) \}.
\]
Proof: Let $a_{\max}$ and $r_{\max}$ as in \eqref{8}. 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 \IV.8. 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

\[
R(r) := \inf_{a \in \mathbb{R}} \max\{\phi(a), r - a\} = \max\{\phi(a_r), r - a_r\} = \phi(a_r) = r - a_r,
\]

and \eqref{12} follows by taking into account \eqref{2}. Note that when $D(\rho \| \sigma) < r < r_{\max}$ then $D(\rho \| \sigma) < a_r < D_{\max}(\rho \| \sigma)$, and \eqref{23} is immediate from \eqref{12}. 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

\[
\inf\{\max\{\phi(a), r - a\} \mid D(\rho \| \sigma) < a < D_{\max}(\rho \| \sigma)\} = \inf\{\phi(a) \mid D(\rho \| \sigma) < a < D_{\max}(\rho \| \sigma)\} = 0 = R(r),
\]

proving \eqref{12}.

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,

\[
R(r) = \inf_{a \in \mathbb{R}} \max\{\phi(a), r - a\} = \inf\{\max\{\phi(a), r - a\} \mid D(\rho \| \sigma) < a < D_{\max}(\rho \| \sigma)\} = \inf\{r - a \mid a \leq a_{\max}\} = r - a_{\max} = r - D_{\max}(\rho \| \sigma).
\]

Taking into account \eqref{2}, we get \eqref{12} and \eqref{23}. \hfill \Box 

**D. Representation as cutoff rates**

In the setting of Section \IV.A let

\[
\alpha_{n,r} := \min\{\text{Tr } \rho_n(I - T) : T \text{ test } , \text{Tr } \sigma_n T \leq e^{-nr}\}.
\]

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

\[
\limsup_{n \to \infty} \frac{1}{n} \log(1 - \alpha_{n,r}) \leq -\kappa(r - r_0), \quad r \in \mathbb{R}.
\]

As before, we assume that $\text{supp } \rho \subseteq \text{supp } \sigma$ and $\rho \neq \sigma$. We have the following:

**Theorem IV.15** For every $\kappa \in (0,1)$, $C_\kappa(\rho\|\sigma) = D_{\max}^{(\text{new})}(\rho\|\sigma)$.

**Proof:** By Theorem \IV.9 we have

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

By definition, we have

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

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

\[
\frac{1}{\kappa} \tilde{\psi}(\kappa) = \frac{1}{\kappa}(1 - \kappa)\psi\left( \frac{\kappa}{1 - \kappa} \right) = D_{\max}^{(\text{new})}(\rho\|\sigma)
\]

is the smallest $r_0$ for which \eqref{95} holds. \hfill \Box
The above Theorem immediately yields the following operational interpretation of the new Rényi relative entropies:

**Corollary IV.16** For every $\alpha > 1$,

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

**V. DISCUSSION**

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 $\alpha$-relative entropies $D_\alpha^{(\text{new})}$ 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 correct quantum generalization of Rényi’s $\alpha$-relative entropies for $\alpha > 1$ are given by $D_\alpha^{(\text{new})}$. On the other hand, previous results [3, 18, 28, 32] show that for $\alpha < 1$, the correct quantum generalization is the traditional notion $D_\alpha^{(\text{old})}$.

Our proof for the optimality of the converse Hoeffding divergence for the strong converse rate follows immediately from the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, under measurements; this proof technique goes back to Nagaoka’s proof for the strong converse [31]. 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 the monotonicity of the relative entropy under pinching. Thus in our proof for the strong converse exponent, and the representation of the Rényi relative entropies as cutoff rates, we only used the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, under measurements. On the other hand, the cutoff rate representation implies as straightforward consequence the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, arbitrary CPTP maps. Hence, our proof for the cutoff rate representation provides a semi-operational proof of the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, arbitrary CPTP maps. This is different from the case of $D_\alpha^{(\text{old})}$, $\alpha \in (0, 1)$, for the monotonicity of which a fully operational proof exists [32], i.e., an operational interpretation can be shown without assuming any form of monotonicity, which operational interpretation then yields the monotonicity as a trivial consequence [32].

Our proof for the optimality of the converse Hoeffding divergence for the strong converse rate follows immediately from the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, under measurements; this proof technique goes back to Nagaoka’s proof for the strong converse [31]. 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 the monotonicity of the relative entropy under pinching. Thus in our proof for the strong converse exponent, and the representation of the Rényi relative entropies as cutoff rates, we only used the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, under measurements. On the other hand, the cutoff rate representation implies as straightforward consequence the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, arbitrary CPTP maps. Hence, our proof for the cutoff rate representation provides a semi-operational proof of the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, arbitrary CPTP maps. This is different from the case of $D_\alpha^{(\text{old})}$, $\alpha \in (0, 1)$, for the monotonicity of which a fully operational proof exists [32], i.e., an operational interpretation can be shown without assuming any form of monotonicity, which operational interpretation then yields the monotonicity as a trivial consequence [32].

Our proof for the optimality of the converse Hoeffding divergence for the strong converse rate follows immediately from the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, under measurements; this proof technique goes back to Nagaoka’s proof for the strong converse [31]. 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 the monotonicity of the relative entropy under pinching. Thus in our proof for the strong converse exponent, and the representation of the Rényi relative entropies as cutoff rates, we only used the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, under measurements. On the other hand, the cutoff rate representation implies as straightforward consequence the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, arbitrary CPTP maps. Hence, our proof for the cutoff rate representation provides a semi-operational proof of the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, arbitrary CPTP maps. This is different from the case of $D_\alpha^{(\text{old})}$, $\alpha \in (0, 1)$, for the monotonicity of which a fully operational proof exists [32], i.e., an operational interpretation can be shown without assuming any form of monotonicity, which operational interpretation then yields the monotonicity as a trivial consequence [32].

Monotonicity of $D_\alpha^{(\text{new})}$ have been proved by various methods, applicable to different parameter ranges, in [3, 13, 30, 46]. The strongest monotonicity result so far is the one given in [13], where monotonicity under general CPTP maps was shown for every parameter $\alpha \in [1/2, +\infty)$, while it is known that monotonicity doesn’t hold for $\alpha < 1/2$ [30]. In Section [13] we gave a new, simple proof for the monotonicity of $D_\alpha^{(\text{new})}$, $\alpha > 1$, by first showing that the $D_\alpha^{(\text{new})}$ are asymptotically attainable by measurements for $\alpha > 1$, from which the monotonicity for general CPTP maps follows as a straightforward consequence. The same attainability property for $\alpha = 1$ was the key technical tool to prove the direct part of the quantum Stein’s lemma [15], and various generalizations [7–9]. We remark that this attainability property is not known for $D_\alpha^{(\text{new})}$ for $1/2 \leq \alpha < 1$.

For a general quantum divergence $D$, one can consider the following properties:

1. Monotonicity under CPTP maps (MON);
2. Monotonicity under measurements (MMON);
3. Attainability by measurements (A);
4. Asymptotic attainability by measurements (AA);
5. Sufficiency (S).
In the above, attainability by measurements means that for any states \( \rho, \sigma \), there exists a POVM \( \{ M_x \}_{x \in \mathcal{X}} \) such that \( D(\rho \| \sigma) = D(\{ \text{Tr} \rho M_x \}_{x \in \mathcal{X}} \| \{ \text{Tr} \sigma M_x \}_{x \in \mathcal{X}}) \), and asymptotic attainability by measurements means that for any any states \( \rho, \sigma \), there exists a sequence of POVMs \( \{ M^{(n)}_x \}_{x \in \mathcal{X}_n} \) such that \( D(\rho \| \sigma) = \lim_{n \to +\infty} \frac{1}{n} D(\{ \text{Tr} \rho^{\otimes n} M^{(n)}_x \}_{x \in \mathcal{X}} \| \{ \text{Tr} \sigma^{\otimes n} M^{(n)}_x \}_{x \in \mathcal{X}}) \). Sufficiency means that equality in the monotonicity inequality can only hold in a trivial way, i.e., for any two states \( \rho, \sigma \), and any CPTP map \( \Phi \), if \( D(\Phi(\rho) \| \Phi(\sigma)) = D(\rho \| \sigma) \) then there exists another CPTP map \( \Psi \) such that \( \Psi(\Phi(\rho)) = \rho \) and \( \Psi(\Phi(\sigma)) = \sigma \). If we assume that \( D \) is additive in the sense that for any states \( \rho, \sigma \) and any \( n \in \mathbb{N} \), \( D(\rho^{\otimes n} \| \sigma^{\otimes n}) = n D(\rho \| \sigma) \), then the following implications are straightforward:

\[
(A) \implies (AA) \implies (MON) \implies (MMON).
\]

Moreover, (A) yields an even stronger monotonicity property, namely monotonicity under trace-preserving positive maps. Indeed, if \( \Phi \) is a trace-preserving positive map and \( \{ M_x \}_{x \in \mathcal{X}} \) is a POVM that achieves \( D(\Phi(\rho) \| \Phi(\sigma)) \) then we have \( D(\Phi(\rho) \| \Phi(\sigma)) = D(\{ \text{Tr} M_x \Phi(\rho) \}_{x} \| \{ \text{Tr} M_x \Phi(\sigma) \}_{x}) = D(\{ \text{Tr} \Phi^*(M_x) \rho \}_{x} \| \{ \text{Tr} \Phi^*(M_x) \sigma \}_{x}) \leq D(\rho \| \sigma) \), where in the last step we used (A) again.

The new Rényi divergences \( D^{(\text{new})}_\alpha \) satisfy (MON) for \( \alpha \in [1/2, +\infty] \), and (AA) for \( \alpha > 1 \), as we have shown in Theorem [11.7] it is an open question whether they satisfy (AA) also for \( \alpha \in [1/2, 1) \), or whether (AA) can be strengthened to (A) in general. The parameter values \( \alpha = 1/2 \) and \( \alpha = +\infty \) are special in this respect; indeed, \( \alpha = 1/2 \) is a function of the fidelity, which is known to be achievable by measurements (see, e.g., [34, Chapter 9]), while \( D^{(\text{new})}_\infty \) is the max-relative entropy [11, 30], which also satisfies (A). By the above, \( D^{(\text{new})}_\alpha \) is non-increasing under trace-preserving positive map, or equivalently, the fidelity is non-decreasing under trace-preserving positive maps, i.e., we don’t need to assume complete positivity for these monotonicity properties. The traditional quantum Rényi divergences \( D^{(\text{old})}_\alpha \) satisfy (MON) for \( \alpha \in [0, 2] \) [26, 38] and (MMON) for \( \alpha > 2 \) [17]. Since the post-measurement states are the same for the old and the new Rényi relative entropies, (AA) for \( D^{(\text{new})}_\alpha \) excludes (AA) for \( D^{(\text{old})}_\alpha \) for \( \alpha > 1 \). It is not known whether (AA) holds for \( D^{(\text{old})}_\alpha \) for \( \alpha \in (0, 1) \); however, it is known that the pinching asymptotics does not achieve \( D^{(\text{old})}_\alpha \) in general [21], i.e., \( D^{(\text{old})}_\alpha(\rho \| \sigma) = \lim_{n \to -\infty} \frac{1}{n} D^{(\text{old})}_\alpha(\hat{\rho}_n \| \sigma) \) need not hold in general for \( \alpha \in (0, 1) \) (see [1] for the notation).

The old Rényi divergences \( D^{(\text{old})}_\alpha \) for \( \alpha \in [0, 2] \) are known to have the sufficiency property [23-25, 39, 40]; the same is an open question for the new Rényi divergences, except for \( \alpha = 1/2 \) and \( \alpha = +\infty \), for which values (S) fails. This follows from the following general observations, the first of which is due to Petz [41, Lemma 4.1].

**Lemma V.1** Let \( \rho, \sigma \) be states and \( \{ M_x \}_{x \in \mathcal{X}} \) be a measurement such that

\[
D^{(\text{old})}_{1/2}(\{ \text{Tr} \rho M_x \}_{x \in \mathcal{X}} \| \{ \text{Tr} \sigma M_x \}_{x \in \mathcal{X}}) = D^{(\text{old})}_{1/2}(\rho \| \sigma).
\]

Then \( \rho \) and \( \sigma \) commute.

**Corollary V.2** No quantum divergence can satisfy (MON)+(A)+(S).

**Proof:** Assume that \( D \) satisfies (MON), (A) and (S), and let \( \rho, \sigma \) be non-commuting states. By (A), there exists a POVM \( \{ M_x \}_{x \in \mathcal{X}} \) such that \( D(\rho \| \sigma) = D(\{ \text{Tr} \rho M_x \}_{x \in \mathcal{X}} \| \{ \text{Tr} \sigma M_x \}_{x \in \mathcal{X}}) \). By (S), there exists a CPTP map \( \Psi \) such that \( \Psi(\{ \text{Tr} \rho M_x \}_{x \in \mathcal{X}}) = \rho \) and \( \Psi(\{ \text{Tr} \sigma M_x \}_{x \in \mathcal{X}}) = \sigma \). By (MON), we have [40], and by Lemma V.1, \( \rho \) and \( \sigma \) commute, which is a contradiction. \( \square \)

Due to the above, \( D^{(\text{new})}_{1/2} \) and \( D^{(\text{new})}_\infty \) do not satisfy (S). Sufficiency of \( D^{(\text{new})}_\alpha \) for other parameter values is an open question.

We close this section by pointing out an operational proof of the Lieb-Thirring inequality [27], that follows easily from our main result, Theorem [V.9]. Indeed, combining [27] with [40], we get that

\[
D^{(\text{old})}_\alpha(\rho \| \sigma) \geq D^{(\text{new})}_\alpha(\rho \| \sigma), \quad \alpha > 1,
\]

or equivalently,

\[
\text{Tr} \rho^\alpha \sigma^{1-\alpha} \geq \left( \frac{\rho^\alpha}{\text{Tr} \rho^\alpha} \frac{1}{\alpha} \sigma^{1-\alpha} \right)^\alpha, \quad \alpha > 1.
\]
Introducing $A := \rho^\frac{1}{2}$ and $B := \sigma^\frac{1}{2}$, the above can be rewritten as

$$\text{Tr } A^\alpha B^\alpha A^\alpha \geq \text{Tr } (ABA)^\alpha, \quad \alpha > 1. \tag{97}$$

Since we were interested in hypothesis testing, we only derived Theorem 4.3 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 V.3 (Lieb-Thirring inequality)** For any positive semidefinite operators $A$ and $B$, the following holds.

**ACKNOWLEDGMENT**

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 Professor 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 and Ke Li for comments on previous versions of the manuscript.

[1] H. Araki: *On an inequality of Lieb and Thirring*; Letters in Mathematical Physics; Volume 19, Issue 2, pp. 167–170, (1990)
[2] K.M.R. Audenaert, J. Calsamiglia, Ll. Masanes, R. Munoz-Tapia, A. Acin, E. Bagan, F. Verstraete.: *Discriminating states: the quantum Chernoff bound*; Phys. Rev. Lett. 98 160501, (2007)
[3] K.M.R. Audenaert, M. Nussbaum, A. Szkola, F. Verstraete: *Asymptotic error rates in quantum hypothesis testing*; Comm. Math. Phys. 279, 251–283 (2008).
[4] K.M.R. Audenaert: *On the Araki-Lieb-Thirring inequality*; Int. J. of Information and Systems Sciences 4, pp. 78–83, (2008)
[5] Salman Beigi: *Quantum Rényi divergence satisfies data processing inequality*; J. Math. Phys., 54, 122202, (2013).
[6] R. Bhatia: *Matrix Analysis*; Springer, (1997)
[7] Igor Bjelakovic, Tyll Krüger, Rainer Siegmund-Schultze, Arleta Szkola: *The Shannon-McMillan theorem for ergodic quantum lattice systems*; Inventiones Mathematicae Vol. 155, Issue 1, pp. 203–222, (2004)
[8] I. Bjelakovic, R. Sigmund-Schultze: *An ergodic theorem for the quantum relative entropy*; Comm. Math. Phys. 247, 697–712, (2004)
[9] Igor Bjelakovic, Jean-Dominique Deuschel, Tyll Krüger, Ruedi Seiler, Rainer Siegmund-Schultze, Arleta Szkola: *A quantum version of Sanov’s theorem*; Communications in Mathematical Physics Vol. 260, Issue 3, pp. 659–671, (2005)
[10] I. Csiszár: *Generalized cutoff rates and Rényi’s information measures*; IEEE Trans. Inf. Theory 41, 26–34, (1995)
[11] N. Datta: *Min- and Max-Relative Entropies and a New Entanglement Monotone*; IEEE Transactions on Information Theory, vol. 55, no. 6, pp. 2816–2826, (2009).
[12] A. Dembo, O. Zeitouni: *Large Deviations Techniques and Applications; Second ed.*, Springer, Application of Mathematics, Vol. 38, (1998)
[13] Rupert L. Frank and Elliott H. Lieb: *Monotonicity of a relative Rényi entropy*; arXiv:1306.5358, (2013).
[14] T.S. Han and K. Kobayashi: *The strong converse theorem for hypothesis testing*; IEEE Trans. Inform. Theory, vol. 35, pp. 178–180, (1989)
[15] T.S. Han: *Information-Spectrum Methods in Information Theory*, Springer, (2003)
[16] M. Hayashi: *Optimal sequence of POVM’s in the sense of Stein’s lemma in quantum hypothesis testing*; J. Phys. A: Math. Gen. 35, pp. 10759–10773, (2002).
[17] M. Hayashi: *Quantum Information Theory: An Introduction*; Springer, (2006).
[18] M. Hayashi: *Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding*; Phys. Rev. A 76, 062301, (2007).
[19] F. Hiai, D. Petz: *The proper formula for relative entropy and its asymptotics in quantum probability*; Comm. Math. Phys. 143, 99–114 (1991).
[20] F. Hiai, unpublished, (2006)
[21] F. Hiai, M. Mosonyi, T. Ogawa: Large deviations and Chernoff bound for certain correlated states on a spin chain; J. Math. Phys. 48, (2007)
[22] F. Hiai, M. Mosonyi, T. Ogawa: Error exponents in hypothesis testing for correlated states on a spin chain; J. Math. Phys. 49, 032112, (2008)
[23] F. Hiai, M. Mosonyi, D. Petz, C. Bény: Quantum f-divergences and error correction; Rev. Math. Phys., volume 23, issue 7, pp. 691 – 747, (2011)
[24] A. Jenčová, D. Petz: Sufficiency in quantum statistical inference; Commun. Math. Phys. 263, 259–276, (2006).
[25] A. Jenčová, D. Petz. Sufficiency in quantum statistical inference. A survey with examples; Infin. Dimens. Anal. Quantum Probab. Relat. Top. 9, 331–351, (2006)
[26] E.H. Lieb: Convex trace functions and the Wigner-Yanase-Dyson conjecture; Adv. Math. 11, 267–288 (1973).
[27] E.H. Lieb, W. Thirring: Studies in mathematical physics; pp. 269–297. Princeton University Press, Princeton, (1976)
[28] M. Mosonyi, F. Hiai: On the quantum Renyi relative entropies and related capacity formulas; IEEE Trans. Inf. Theory, 57, pp. 2474–2487, (2011)
[29] M. Mosonyi: Rényi divergences and the classical capacity of finite compound channels; arxiv:1310.7525, (2013)
[30] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, Marco Tomamichel: On quantum Renyi entropies: a new definition and some properties; J. Math. Phys. 54, 122203, (2013)
[31] H. Nagaoka: Strong converse theorems in quantum information theory; In Proc. of ERATO Workshop on Quantum Information Science, Tokyo, September, p. 33, 2001; (M. Hayashi ed., Asymptotic Theory of Quantum Statistical Inference, World Scientific, pp. 64–65, 2005.)
[32] H. Nagaoka: The converse part of the theorem for quantum Hoeffding bound; quant-ph/0611289
[33] H. Nagaoka and M. Hayashi: An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses; IEEE Trans. Inform. Theory, 53, 534–549, (2007).
[34] M.A. Nielsen, I.L. Chuang: Quantum Information and Quantum Computation; Cambridge University Press, Cambridge, UK, (2000)
[35] M. Nussbaum, A. Szkoła: A lower bound of Chernoff type for symmetric quantum hypothesis testing; Ann. Statist. 37, 1040–1057, (2009)
[36] T. Ogawa, H. Nagaoka: Strong converse and Stein’s lemma in quantum hypothesis testing; IEEE Trans. Inform. Theory 47, 2428–2433 (2000).
[37] T. Ogawa, M. Hayashi: On error exponents in quantum hypothesis testing; IEEE Trans. Inf. Theory, vol. 50, issue 6, pp. 1368–1372, (2004)
[38] D. Petz: Quasi-entropies for finite quantum systems; Rep. Math. Phys. 23, 57–65, (1986)
[39] D. Petz: Sufficient subalgebras and the relative entropy of states of a von Neumann algebra; Commun. Math. Phys. 105, 123–131, (1986)
[40] D. Petz: Sufficiency of channels over von Neumann algebras; Quart. J. Math. Oxford Ser. (2) 39, no. 153, 97–108, (1988)
[41] D. Petz: Monotonicity of quantum relative entropy revisited; Reviews in Mathematical Physics, vol. 15, no. 1, 79–91, (2003)
[42] R. Renner: Security of Quantum Key Distribution, PhD dissertation, Swiss Federal Institute of Technology Zurich, Diss. ETH No. 16242, (2005).
[43] A. Rényi: On measures of entropy and information; Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I, pp. 547–561, Univ. California Press, Berkeley, California, (1961).
[44] A. Uhlmann: Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory; Commun. Math. Phys. 54, 21–32, (1977)
[45] H. Umegaki: Conditional expectation in an operator algebra; Kodai Math. Sem. Rep. 14, 59–85, (1962)
[46] Mark M. Wilde, Andreas Winter, Dong Yang: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels; arXiv:1306.1586, (2013)