The Strong Converse Exponent of Discriminating Infinite-Dimensional Quantum States

Milán Mosonyi

1 MTA-BME Lendület Quantum Information Theory Research Group, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
2 Department of Analysis, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary. E-mail: milan.mosonyi@gmail.com

Received: 30 May 2022 / Accepted: 24 November 2022
Published online: 10 March 2023 – © The Author(s) 2023

Abstract: The sandwiched Rényi divergences of two finite-dimensional density operators quantify their asymptotic distinguishability in the strong converse domain. This establishes the sandwiched Rényi divergences as the operationally relevant ones among the infinitely many quantum extensions of the classical Rényi divergences for Rényi parameter $\alpha > 1$. The known proof of this goes by showing that the sandwiched Rényi divergence coincides with the regularized measured Rényi divergence, which in turn is proved by asymptotic pinching, a fundamentally finite-dimensional technique. Thus, while the notion of the sandwiched Rényi divergences was extended recently to density operators on an infinite-dimensional Hilbert space (in fact, even for states of an arbitrary von Neumann algebra), these quantities were so far lacking an operational interpretation similar to the finite-dimensional case, and it has also been open whether they coincide with the regularized measured Rényi divergences. In this paper we fill this gap by answering both questions in the positive for density operators on an infinite-dimensional Hilbert space, using a simple finite-dimensional approximation technique. We also initiate the study of the sandwiched Rényi divergences, and the related problem of the strong converse exponent, for pairs of positive semi-definite operators that are not necessarily trace-class (this corresponds to considering weights in a general von Neumann algebra setting). This is motivated by the need to define conditional Rényi entropies in the infinite-dimensional setting, while it might also be interesting from the purely mathematical point of view of extending the concept of Rényi (and other) divergences to settings beyond the standard one of positive trace-class operators (positive normal functionals in the von Neumann algebra setting). In this spirit, we also discuss the definition and some properties of the more general family of Rényi $(\alpha, z)$-divergences of positive semi-definite operators on an infinite-dimensional separable Hilbert space.
1. Introduction

In a simple binary i.i.d. quantum state discrimination problem, an experimenter is presented with several identically prepared quantum systems, all in the same state that is either described by a density operator $\rho$ on the system’s Hilbert space $\mathcal{H}$ (null-hypothesis $H_0$), or by another density operator $\sigma$ (alternative hypothesis $H_1$). The experimenter’s task is to guess which hypothesis is correct, based on the result of a 2-outcome measurement, represented by a pair of operators $(T_n(0) := T_n, T_n(1) = I - T_n)$, where $T_n \in \mathcal{B}(\mathcal{H}_n)$ is a test on $\mathcal{H}_n := \mathcal{H}^\otimes n$, i.e., $0 \leq T_n \leq I$, and $n$ is the number of identically prepared systems. If the outcome of the measurement is $k$, described by the measurement operator $T_n(k)$, the experimenter decides that hypothesis $k$ is true. The type I success probability, i.e., the probability that the experimenter correctly identifies the state to be $\rho_n$, and the type II error probability, i.e., the probability that the experimenter erroneously identifies the state to be $\sigma_n$, are given by

$$
\gamma_n(T_n|\rho_n) := \text{Tr} \rho_n T_n, \quad \beta_n(T_n|\sigma_n) := \text{Tr} \sigma_n T_n,
$$

respectively, where $\rho_n = \rho^\otimes n$, $\sigma_n = \sigma^\otimes n$.

In the asymptotic analysis of the problem, it is customary to look for the optimal asymptotics of the type I success probabilities under the constraint that the type II error probabilities decrease at least as fast as $\beta_n \sim e^{-nr}$ with some fixed $r$. It is known that if $r$ is smaller than the relative entropy of $\rho$ and $\sigma$ then the type I success probabilities converge to 1 exponentially fast, and the optimal exponent (the so-called direct exponent) is equal to the Hoeffding divergence $H_r$ of $\rho$ and $\sigma$ [3,15,26,37]. The Hoeffding divergences are defined from the Petz-type Rényi divergences $D^\alpha$ with $\alpha \in (0, 1)$, and the above result establishes the operational significance of these divergences [32,37].

On the other hand, it was shown in [33] (see also [16,36,38]) that if the Hilbert space is finite-dimensional, (equivalently, the density operators are of finite rank), and $r$ is larger than the relative entropy, then the type I success probabilities converge to 0 exponentially fast, and the optimal exponent (the so-called strong converse exponent) is equal to the Hoeffding anti-divergence $H^*_r$ of $\rho$ and $\sigma$. ($H^*_r$, as well as the various divergences mentioned below, will be precisely defined in the main text.) The Hoeffding anti-divergences are defined from the sandwiched Rényi divergences $D^*_\alpha$ with $\alpha > 1$ [35,46], and this result establishes the operational significance of these divergences.

A key step in the proof of the strong converse exponent in [33] is showing that the regularized measured Rényi divergence $\overline{D}^\text{meas}_\alpha \rho$ coincides with the sandwiched Rényi divergence $D^*_\alpha \rho$ for any $\alpha > 1$, which was proved using the pinching inequality [14], a fundamentally finite-dimensional technique. Thus, while the notion of the sandwiched Rényi divergences was extended recently to density operators on an infinite-dimensional Hilbert space (in fact, even for states of an arbitrary von Neumann algebra) in [6] and [27], these quantities were so far lacking an operational interpretation similar to the finite-dimensional case described above, and it has also been open whether they coincide with the regularized measured Rényi divergences. In this paper we fill this gap by answering both questions in the positive for density operators on an infinite-dimensional Hilbert space.

We also initiate the study of the sandwiched Rényi divergences, and the related problem of the strong converse exponents, for pairs of positive semi-definite operators that are not necessarily trace-class (this corresponds to considering weights in a general von Neumann algebra setting). This is motivated by the need to define conditional Rényi entropies in the infinite-dimensional setting, while it might also be interesting from
the purely mathematical point of view of extending the concept of Rényi (and other) divergences to settings beyond the standard one of positive trace-class operators (or positive normal functionals, in the von Neumann algebra setting). In this spirit, we also discuss the definition and some properties of the more general family of Rényi \((\alpha, z)\)-divergences \([4, 25]\) in this setting. To the best of our knowledge, this is new even for trace-class operators when the underlying Hilbert space is infinite-dimensional.

The structure of the paper is as follows. In Sect. 2 we collect some necessary preliminaries. In Sect. 3 we define the Rényi \((\alpha, z)\)-divergences for an arbitrary pair of positive semi-definite operators on a possibly infinite-dimensional Hilbert space, and establish some of their properties. The most important part of this section for the later applications is the recoverability of the sandwiched Rényi divergence from finite-dimensional restrictions, given in Proposition 3.40. Based on this, in Sect. 3.4 we show that the sandwiched Rényi divergence is equal to the regularized measured Rényi divergence for pairs of states, extending the finite-dimensional result of \([33]\) to infinite dimension. In Sect. 4.1 we consider a generalization of the state discrimination problem where the hypotheses are given by (not necessarily trace-class) positive semi-definite operators, and establish lower and upper bounds on the strong converse exponents in this setting. In particular, we show that the strong converse exponent is equal to the Hoeffding anti-divergence for quantum states, thereby giving an operational interpretation of the sandwiched Rényi divergences analogous to the finite-dimensional case. Moreover, we prove the above equality also in the case where the reference operator \(\sigma\) is only assumed to be compact, and to dominate the first operator \(\varrho\) as \(\varrho \leq \lambda \sigma\) for some \(\lambda > 0\). In Sect. 4.2, we give a direct operational interpretation to the sandwiched Rényi divergences as generalized cutoff rates, extending the analogous interpretations given previously for classical \([8]\) and finite-dimensional quantum states \([33]\). In Sect. 4.3 we use the strong converse result from Sect. 4.1 to show the monotonicity of the sandwiched Rényi divergences under the action of the dual of a normal unital completely positive map. While this follows from \([6, 27]\) for density operators, our proof is completely different, and also applies to other settings, e.g., for a compact \(\sigma\) that dominates \(\varrho\).

2. Preliminaries

Throughout the paper, \(\mathcal{H}\) and \(\mathcal{K}\) will denote separable Hilbert spaces (of finite or infinite dimension), and \(B(\mathcal{H}, \mathcal{K})\) will denote the set of everywhere defined bounded linear operators from \(\mathcal{H}\) to \(\mathcal{K}\), with \(B(\mathcal{H}, \mathcal{H}) =: B(\mathcal{H})\). We will use the notations \(B(\mathcal{H})_{\text{sa}}\) for the set of self-adjoint, and \(B(\mathcal{H})_{\geq 0}\), for the set of non-zero positive semi-definite (PSD), operators in \(B(\mathcal{H})\) respectively, and

\[B(\mathcal{H})_{[0, 1]} := \{T \in B(\mathcal{H}) : 0 \leq T \leq I\}\]

for the set of tests in \(B(\mathcal{H})\). A test \(T\) is projective if \(T^2 = T\). We will denote the set of all projections on \(\mathcal{H}\) by \(\mathbb{P}(\mathcal{H})\), and the set of finite rank projections by \(\mathbb{P}_f(\mathcal{H})\). The set of finite-rank operators on \(\mathcal{H}\) will be denoted by \(B_f(\mathcal{H})\). The set of density operators, or states, on \(\mathcal{H}\) will be denoted by \(\mathcal{S}(\mathcal{H})\). For two PSD operators \(\varrho, \sigma \in B(\mathcal{H})_{\geq 0}\), we will use the notations

\[B(\mathcal{H}, \mathcal{K})_{\varrho, \sigma}^+ := \{K \in B(\mathcal{H}, \mathcal{K}) : K \varrho K^* \neq 0, \ K \sigma K^* \neq 0\},\]

and \(B_f(\mathcal{H})_{\varrho, \sigma}^+ := B_f(\mathcal{H}) \cap B(\mathcal{H}, \mathcal{K})_{\varrho, \sigma}^+\), \(\mathbb{P}(\mathcal{H})_{\varrho, \sigma}^+ := \mathbb{P}(\mathcal{H}) \cap B(\mathcal{H}, \mathcal{K})_{\varrho, \sigma}^+\), \(\mathbb{P}_f(\mathcal{H})_{\varrho, \sigma}^+ := \mathbb{P}_f(\mathcal{H}) \cap B(\mathcal{H}, \mathcal{K})_{\varrho, \sigma}^+\), where \(B(\mathcal{H})_{\varrho, \sigma}^+ := B(\mathcal{H}, \mathcal{H})_{\varrho, \sigma}^+\).
For a (possibly unbounded) self-adjoint operator $A$ on a Hilbert space $H$, let $P^A(\cdot)$ denote its spectral PVM, and for any complex-valued measurable function $f$ defined at least on $\text{spec}(A)$, let $f(A) = \int_{\mathbb{R}} f dP^A$ be the operator defined via the usual functional calculus. We will use the relations
\[
(f(A))^* = \overline{f}(A),
\]
\[
\overline{f}(A)g(A) = (fg)(A), \quad \text{dom}(f(A)g(A)) = \text{dom}(g(A)) \cap \text{dom}((fg)(A)),
\]
where $\overline{f}$ stands for the pointwise complex conjugate of $f$, and for a closable operator $X$, $\overline{X}$ denotes its closure.

We say that a (not necessarily everywhere defined or bounded) linear operator $A$ on a Hilbert space is positive semi-definite (PSD), if it is self-adjoint, and $\text{spec}(A) \subseteq [0, +\infty)$. If $A$ is PSD then we may define its real powers as
\[
A^p := \text{id}_{(0, +\infty)}^p(A) = \int_{(0, +\infty)} \text{id}_{(0, +\infty)}^p dP^A, \quad p \in \mathbb{R}.
\]
In particular, $A^0$ is the projection onto $(\ker A)^\perp = \overline{\text{ran}} A =: \text{supp} A$,
\[
(A^p)^{-1} = (A^{-1})^p = A^{-p}, \quad p \in \mathbb{R},
\]
and
\[
A \text{ is bounded} \implies A^{-p}A^p = I, \quad A^p A^{-p} = I_{\text{ran} A^p}, \quad p > 0.
\]
Note that here we use the notation
\[
\text{id}_B : x \mapsto x
\]
for the identity map on any subset $B \subseteq \mathbb{R}$. Analogously, we will denote the characteristic function, or indicator function, of a subset $X_0$ of a set $X$ by
\[
1_{X_0} : x \mapsto \begin{cases} 1, & x \in X_0, \\ 0, & x \notin X_0. \end{cases}
\]
For any $X \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ with polar decomposition $X = V|X|$, we have $|X^*| = V|X|V^*$, whence $|X^*|^p = V|X|^pV^*$ for any $p \in \mathbb{R}$. In particular,
\[
\text{Tr}(X^*X)^p = \text{Tr}(XX^*)^p, \quad p > 0,
\]
which we will use in many proofs below without further notice. We will use the notation $\|X\|_p := (\text{Tr} |X|^p)^{1/p}$ for $X \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and $p > 0$. When $p \geq 1$, $\|\cdot\|_p$ is a norm on the Schatten $p$-class
\[
\mathcal{L}^p(\mathcal{H}) := \{X \in \mathcal{B}(\mathcal{H}) : \text{Tr} |X|^p < +\infty\}.
\]
We will denote the usual operator norm on $\mathcal{B}(\mathcal{H})$ by $\|\cdot\|_\infty$.

The following is well known; see, e.g., [18, Proposition 2.7] and [31, Theorem 2.3].
Lemma 2.1 (Hölder inequality). Let $p_0, p_1, p > 0$ be such that $\frac{1}{p_0} + \frac{1}{p_1} = \frac{1}{p}$. For any $A, B \in \mathcal{B}(\mathcal{H})$,

$$\|AB\|_p \leq \|A\|_{p_0} \|B\|_{p_1}. \quad (2.4)$$

Moreover, if $\|A\|_{p_0} \|B\|_{p_1} < +\infty$ then equality holds in (2.4) if and only if $A = \lambda B$ or $B = \lambda A$ for some $\lambda \geq 0$.

We will use the notations $\text{wo} \lim$ and $\text{so} \lim$ for limits in the weak and the strong operator topologies, respectively. The following two statements are from [13].

Lemma 2.2. Let $A \in \mathcal{L}^p(\mathcal{H})$ for some $p \geq 1$, and $B_n \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, $C_n \in \mathcal{B}(\mathcal{K}, \mathcal{H})$, $n \in \mathbb{N}$, be two sequences bounded in operator norm and converging strongly to some $B_\infty \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and $C_\infty \in \mathcal{B}(\mathcal{K}, \mathcal{H})$, respectively. Then

$$\lim_{n \to +\infty} \|B_n AC_n - B_\infty AC_\infty\|_p = 0, \quad \lim_{n \to +\infty} \|B_n AC_n\|_p = \|B_\infty AC_\infty\|_p. \quad (2.5)$$

Proof. The first limit in (2.5) is immediate from [13, Theorem 1], and the second limit follows from it trivially. \qed

The following is Theorem 2 in [13]:

Lemma 2.3. Let $p \in [1, +\infty)$ and $A, A_n \in \mathcal{L}^p(\mathcal{H})$, $n \in \mathbb{N}$, be such that (so) $\lim_n A_n = A$, (so) $\lim_n A_n^* = A^*$, and $\lim_n \|A_n\|_p = \|A\|_p$. Then $\lim_n \|A_n - A\|_p = 0$.

The following is a special case of [18, Proposition 2.11]:

Lemma 2.4. Assume that a sequence $A_n \in \mathcal{B}(\mathcal{H})$, $n \in \mathbb{N}$, converges to some $A \in \mathcal{B}(\mathcal{H})$ in the weak operator topology. For any $p \in [1, +\infty]$,

$$\|A\|_p \leq \lim \inf_{n \to +\infty} \|A_n\|_p.$$ 

We will need the following straightforward generalization of the minimax theorem from [32, Corollary A.2]. Its proof is essentially the same, which we include for readers’ convenience.

Lemma 2.5. Let $X$ be a compact topological space, $Y$ be an upward directed partially ordered set, and let $f : X \times Y \to \mathbb{R} \cup \{-\infty, +\infty\}$ be a function. Assume that

(i) $f(., y)$ is upper semicontinuous for every $y \in Y$ and
(ii) $f(x, .)$ is monotonic decreasing for every $x \in X$.

Then

$$\sup_{x \in X} \inf_{y \in Y} f(x, y) = \inf_{y \in Y} \sup_{x \in X} f(x, y), \quad (2.6)$$

and the suprema in (2.6) can be replaced by maxima.

Proof. The inequality $\sup_{x \in X} \inf_{y \in Y} f(x, y) \leq \inf_{y \in Y} \sup_{x \in X} f(x, y)$ is trivial, and for the converse inequality it is sufficient to prove that for any finite subset $Y' \subseteq Y$,

$$\sup_{x \in X} \inf_{y \in Y'} f(x, y) \geq \inf_{y \in Y} \sup_{x \in X} f(x, y),$$

according to [32, Lemma A.1] (applied to $-f$ in place of $f$). Due to $Y$ being upward directed, for any finite subset $Y' \subseteq Y$, there exists a $y^* \in Y$ such that $y \leq y^*$ for every $y \in Y'$. Since $f(x, .)$ is assumed to be monotone decreasing, we get

$$\sup_{x \in X} \inf_{y \in Y'} f(x, y) \geq \sup_{x \in X} f(x, y^*) \geq \inf_{y \in Y} \sup_{x \in X} f(x, y),$$

as required. The assertion about the maxima is straightforward from the assumed semicontinuity and the compactness of $X$. \qed
3. The Rényi $(\alpha, z)$-Divergences in Infinite Dimension

The sandwiched Rényi $\alpha$-divergences for pairs of finite-dimensional density operators were introduced in [35,46]. The Rényi $(\alpha, z)$-divergences [4,25] give a 2-parameter extension of this family, which includes both the sandwiched Rényi divergences (corresponding to $z = \alpha$) and the Petz-type, or standard Rényi divergences [40] (corresponding to $z = 1$) as special cases.

The concept of the sandwiched Rényi divergences was extended recently to pairs of positive normal linear functionals on a general von Neumann algebra in [6,27,28], while the Petz-type Rényi divergences have been studied in this more general setting for a long time [21,29,39]. These extensions require advanced knowledge of von Neumann algebras, and the details of the proofs might be difficult to verify for those who are not experts in the subject. Below we give a more pedestrian exposition of the definition and basic properties of the Rényi divergences in the simpler case where the states are represented by density operators on a possibly infinite-dimensional Hilbert space, while in the same time we also generalize the above works in this setting to the case where density operators may be replaced by arbitrary positive semi-definite operators. Since these are mostly not assumed to be trace-class, they cannot be normalized to states in the properly infinite-dimensional case. Moreover, we also consider the more general notion of Rényi $(\alpha, z)$-divergences in this setting.

The recoverability of the sandwiched Rényi divergences from finite-size restrictions, given in Proposition 3.40, seems to be new even for density operators, although in that case it follows easily from the known properties of monotonicity and lower semi-continuity of the sandwiched Rényi divergences.

3.1. Definition and basic properties. The sandwiched Rényi divergence of $\rho$ and $\sigma$ is finite according to the definition in [27] if and only if $\rho$ is in Kosaki’s interpolation space $L^\alpha(H, \sigma)$. The following lemma gives various alternative characterizations of this condition, and also an extension that we will use in the definition of the Rényi $(\alpha, z)$-divergences in this setting. The lemma is essentially a special case of Douglas’ range inclusion theorem [10] for PSD operators with $A:=\varrho^{\alpha \over z}$ and $B:=\sigma^{\alpha-1 \over 2z}$ (points (iv)–(iv)) as well as an extension with further equivalent characterizations (points (i)–(iii)), and it is inspired by a similar statement for the $\alpha = z = +\infty$ case given in [30].

Let us introduce the notation

$$A:= (1, +\infty) \times (0, +\infty) \cup \{(+\infty, +\infty)\}.$$ 

For $(\alpha, z):=(+\infty, +\infty)$, we will use the convention $\varrho^{\infty \over z}: = 1$, and define similar expressions by a formal calculus, e.g., $\varrho^{\alpha \over 2z} = 1 \over 2$, $\sigma^{1-\alpha \over 2z} = 1 \over 2$, $\varrho^{\alpha-1 \over 2z} = 1 \over 2$, etc.

Lemma 3.1. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$, and let $(\alpha, z) \in A$. The following are equivalent:

(i) There exists an $R \in \mathcal{B}(\mathcal{H})$ such that

$$\varrho^{\alpha \over z} = \sigma^{\alpha-1 \over 2z} R \sigma^{\alpha-1 \over 2z}.$$  \hspace{1cm} (3.1)

(ii) $\text{ran} \varrho^{\alpha \over z} \subseteq \text{ran} \sigma^{\alpha-1 \over 2z}$, and $\sigma^{1-\alpha \over 2z} \varrho^{\alpha \over z} \sigma^{1-\alpha \over 2z}$ is densely defined and bounded.
(iii) $\varrho^0 \leq \sigma^0$, and for any/some sequences $0 < c_n < d_n$ with $c_n \to 0$, $d_n \to +\infty$, the sequence of bounded operators

$$\sigma_n^{\frac{1-a}{2} - \varrho^{\frac{a}{2}} \sigma_n^{\frac{1-a}{2}}} \quad n \in \mathbb{N},$$

(3.2)

converges in the weak/strong operator topology, where $\sigma_n := P_n \sigma P_n$, $P_n := 1_{(c_n, d_n)}(\sigma)$. (iv) $\text{ran} \, \sigma^{\frac{a}{2}} - \varrho^{\frac{a}{2}} \subseteq \text{ran} \, \sigma^{\frac{a-1}{2}}$. (v) $\sigma^{\frac{a}{2}} \in \mathcal{B}(\mathcal{H})$. (vi) There exists a $\lambda \geq 0$ such that $\varrho^{\frac{a}{2}} \leq \lambda \sigma^{\frac{a-1}{2}}$.

Moreover, if the above hold then $\varrho^{\frac{a}{2}} \sigma^{\frac{a-1}{2}} = (\sigma^{\frac{a}{2}} \varrho^{\frac{a}{2}})^*$, and among all operators $R$ as in (3.1) there exists a unique PSD operator $R^0 \leq \sigma^0$, denoted by $\varrho_{\sigma, \alpha, z}$, which can be expressed as

$$\varrho_{\sigma, \alpha, z} = \sigma^{\frac{a}{2}} \varrho^{\frac{a}{2}} \sigma^{\frac{a-1}{2}} = (\sigma^{\frac{a}{2}} \varrho^{\frac{a}{2}} \sigma^{\frac{a-1}{2}})^* = (\sigma^{\frac{a}{2}} \varrho^{\frac{a}{2}})(\sigma^{\frac{a-1}{2}})^*$$

(3.3)

where $(\sigma_n)_{n \in \mathbb{N}}$ is any sequence as in (iii). This unique $\varrho_{\sigma, \alpha, z}$ is in the von Neumann algebra generated by $\varrho$ and $\sigma$, and its operator norm is equal to the smallest $\lambda$ for which (iv) holds.

**Proof.** Note that if $R \in \mathcal{B}(\mathcal{H})$ satisfies (3.1) then so does $\sigma^0 R \sigma^0$ as well. Moreover, any of the conditions above imply $\varrho^0 \leq \sigma^0$. Hence, we may assume without loss of generality that $\text{supp} \, \sigma = \mathcal{H}$, so that $\text{ran} \, (\sigma^{\frac{a-1}{2}}) = (\ker (\sigma^{\frac{a-1}{2}}))^{\perp} = (\ker \sigma)^{\perp} = \mathcal{H}$.

Assume that (i) holds. Then $\text{ran} \, \sigma^{\frac{a}{2}} \subseteq \text{ran} \, \sigma^{\frac{a-1}{2}}$ holds trivially, and

$$\sigma^{\frac{a}{2}} \varrho^{\frac{a}{2}} \sigma^{\frac{a-1}{2}} = \sigma^{\frac{a}{2}} \varrho^{\frac{a}{2}} \sigma^{\frac{a-1}{2}} = R \sigma^{\frac{a}{2}} \sigma^{\frac{a-1}{2}} = R \bigg|_{\text{ran} \, \sigma^{\frac{a}{2}}} = R \bigg|_{\text{ran} \, \sigma^{\frac{a-1}{2}}}$$

whence its closure is equal to $R$. This proves (ii) and the existence of the unique $\varrho_{\sigma, \alpha, z}$ with the postulated properties, as well as the first equality in (3.3). Moreover, for any $0 < c_n < d_n$ with $c_n \to 0$, $d_n \to +\infty$, we have, with $\sigma_n$ as in (iii),

$$\sigma_n^{\frac{1-a}{2} - \varrho^{\frac{a}{2}} \sigma_n^{\frac{1-a}{2}}} = \sigma_n^{\frac{1-a}{2}} \varrho^{\frac{a}{2}} \sigma_n^{\frac{1-a}{2}} = \sigma_n^{\frac{1-a}{2}} \varrho_{\sigma, \alpha, z} \sigma_n^{\frac{1-a}{2}}$$

(3.5)

where we used (2.2). This proves (iii) and the first equality in (3.4). Since $\sigma_n^{\frac{1-a}{2} - \varrho^{\frac{a}{2}} \sigma_n^{\frac{1-a}{2}}}$ is in the von Neumann algebra generated by $\varrho$ and $\sigma$, so is $\varrho_{\sigma, \alpha, z}$, according to (3.5). Obviously, (i) also implies

$$\varrho^{\frac{a}{2}} \leq \sigma^{\frac{a-1}{2}} \big( \| \varrho_{\sigma, \alpha, z} \|_{\infty} I \big) \sigma^{\frac{a-1}{2}} = \| \varrho_{\sigma, \alpha, z} \|_{\infty} \sigma^{\frac{a-1}{2}},$$

whence (vi) follows with $\lambda := \| \varrho_{\sigma, \alpha, z} \|_{\infty}$. As a consequence, $\lambda_{\text{min}} \leq \| \varrho_{\sigma, \alpha, z} \|_{\infty}$, where $\lambda_{\text{min}}$ denotes the smallest $\lambda$ for which (vi) holds. Conversely, let $\lambda$ be as in (vi). Multiplying both sides by $\sigma_n^{\frac{1-a}{2}}$ yields $\sigma_n^{\frac{1-a}{2}} \varrho^{\frac{a}{2}} \sigma_n^{\frac{1-a}{2}} \leq \lambda 1_{(c_n, d_n)}(\sigma)$, which in combination with (3.5) gives $\| \varrho_{\sigma, \alpha, z} \|_{\infty} \leq \lambda$. Thus, $\lambda_{\text{min}} = \| \varrho_{\sigma, \alpha, z} \|_{\infty}$, as stated.
Assume next that (ii) holds. Then
\[
\sigma^{a-1} \sigma^{\frac{1-a}{2}} = \sigma^{\frac{1-a}{2}} \sigma^{a-1} \rho z \sigma^{\frac{1-a}{2}} \sigma^{a-1} = \sigma, \tag{3.6}
\]
where the last equality follows from the assumption \( \text{ran } \sigma^{\frac{a}{2}} \subseteq \text{ran } \sigma^{\frac{a-1}{2}} \). Since \( \sigma^{\frac{a}{2}} \) is everywhere defined, it is actually equal to the first operator in (3.6), and thus (i) holds. Moreover, if (3.6) holds then for any \( \phi \in \mathcal{H} \),
\[
0 \leq \langle \phi, \sigma^{\frac{a}{2}} \phi \rangle = \langle \phi, \sigma^{\frac{a-1}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}} \phi \rangle = \langle \sigma^{\frac{a-1}{2}} \phi, \sigma^{\frac{1-a}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}} \phi \rangle.
\]
Since \( \sigma^{\frac{a-1}{2}} \) is dense and \( \sigma^{\frac{1-a}{2}} \sigma^{\frac{1-a}{2}} \) is bounded, it follows that \( \sigma^{\frac{a-1}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}} \) is PSD.

Assume now (iii), i.e., that for some sequences \( 0 < c_n < d_n \) with \( c_n \to 0, d_n \to +\infty \), the sequence of operators \( (\sigma_n^{\frac{a-1}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{1-a}{2}} \sigma_n^{\frac{a}{2}})_{n \in \mathbb{N}} \) converges in the weak operator topology to some operator \( R_{\sigma, \alpha, z} \). Then
\[
\sigma^{\frac{a-1}{2}} R_{\sigma, \alpha, z} \sigma^{\frac{a}{2}} = \text{(wo)} \lim_{n \to +\infty} \sigma^{\frac{a-1}{2}} \sigma_n^{\frac{1-a}{2}} \sigma^{\frac{1-a}{2}} \sigma_n^{\frac{a}{2}} = \sigma^{\frac{a}{2}},
\]
and hence (i) holds, as well as the second equality in (3.4).

The equivalence of (iv), (v), and (iv) follows from Douglas’ range inclusion theorem [10]. Note that (iv) \( \iff \) (v) is simple, as \( \sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}} \) being everywhere defined is equivalent to \( \text{ran } \sigma^{\frac{a}{2}} \subseteq \text{dom } \sigma^{\frac{1-a}{2}} = \text{ran } \sigma^{\frac{a-1}{2}} \), and boundedness of \( \sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}} \) is automatic from the boundedness of \( \sigma^{\frac{a}{2}} \) and the closedness of \( \sigma^{\frac{1-a}{2}} \), due to the closed graph theorem. Moreover, we have
\[
\sigma^{\frac{a-1}{2}} (\sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}}) = I \quad \text{ran } \sigma^{\frac{a-1}{2}} \sigma^{\frac{a}{2}} = \sigma^{\frac{a}{2}},
\]
whence
\[
\sigma^{\frac{a}{2}} \sigma^{\frac{1-a}{2}} = (\sigma^{\frac{a}{2}})^* \quad \text{ran } \sigma^{\frac{a-1}{2}} \sigma^{\frac{a}{2}} = (\sigma^{\frac{a-1}{2}})^*,
\]
which is densely defined and bounded. Thus, \( \sigma^{\frac{a}{2}} \sigma^{\frac{1-a}{2}} = (\sigma^{\frac{a-1}{2}})^* \). Finally,
\[
\sigma^{\frac{a-1}{2}} \left[ (\sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}}) \sigma^{\frac{a}{2}} \sigma^{\frac{1-a}{2}} \right] \sigma^{\frac{a}{2}} \geq \sigma^{\frac{a-1}{2}} \sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}} \sigma^{\frac{a}{2}} \sigma^{\frac{1-a}{2}} = \sigma^{\frac{a}{2}}, \tag{3.7}
\]
where the last equality follows from the assumption \( \text{ran } \sigma^{\frac{a}{2}} \subseteq \text{ran } \sigma^{\frac{a-1}{2}} \). Thus, (i) follows with \( \rho_{\sigma, \alpha, z} = (\sigma^{\frac{1-a}{2}} \sigma^{\frac{a}{2}}) \sigma^{\frac{a}{2}} \sigma^{\frac{1-a}{2}} \), and we also have the second and the third equalities in (3.3). Note that the last expression in (3.3) gives another proof for the positive semi-definiteness of \( \rho_{\sigma, \alpha, z} \). □
Definition 3.2. For $\sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$ and $(\alpha, z) \in \mathbb{A}$, let
\[
B^{\alpha,z}(\mathcal{H}, \sigma) := \left\{ \rho \in \mathcal{B}(\mathcal{H})_{\geq 0} : \exists R \in \mathcal{B}(\mathcal{H}) \text{ s.t. } \rho^z = \sigma^{z^{-1}} R \sigma^{z^{-1}} \right\}.
\]
When $\alpha = z$, we will use the shorthand notation $B^{\alpha,\alpha}(\mathcal{H}, \sigma) := B^\alpha(\mathcal{H}, \sigma)$.

Remark 3.3. Note that $\rho \in B^{\alpha,z}(\mathcal{H}, \sigma)$ if and only if it satisfies (i) in Lemma 3.1, which is equivalently characterized by all the other points in Lemma 3.1. In particular, there exists a unique PSD $\varrho_{\alpha,\alpha,z}$ satisfying $\varrho_{\alpha,\alpha,z}^0 \leq \sigma^0$ and $\varrho^z_{\alpha,\alpha,z} = \sigma^{z^{-1}} \varrho_{\alpha,\alpha,z} \sigma^{z^{-1}}$, and thus the map $\rho \mapsto \varrho_{\alpha,\alpha,z}$ is well-defined from $B^{\alpha,z}(\mathcal{H}, \sigma)$ onto $\{\tau \in \mathcal{B}(\mathcal{H})_{\geq 0} : \tau^0 \leq \sigma^0\}$, and it is also injective, hence it is a bijection. When $\alpha = z$, we will use the notation $\varrho_{\alpha,\alpha,z} =: \varrho_{\alpha,\alpha}$.

Lemma 3.4. For any $\sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$, $(0, +\infty) \ni z \mapsto B^{\alpha,z}(\mathcal{H}, \sigma)$ is increasing, i.e.,
\[
0 < z < z' \implies B^{\alpha,z}(\mathcal{H}, \sigma) \subseteq B^{\alpha,z'}(\mathcal{H}, \sigma).
\]
Proof. Let $\rho \in B^{\alpha,z}(\mathcal{H}, \sigma)$. Then, by (iv) of Lemma 3.1, $\varrho^z_{\alpha,\alpha,z} \leq \lambda \sigma^{z^{-1}}$ for some $\lambda \in (0, +\infty)$. Since $z \leq z'$, $id_{(0, +\infty)}$ is operator monotone, whence
\[
\varrho^z_{\alpha,\alpha,z} = (\varrho^z_{\alpha,\alpha,z})^\frac{z}{z'} \leq \lambda^\frac{z}{z'} (\sigma^{z^{-1}})^\frac{z}{z'} = \sigma^{\frac{z^{-1}}{z'}}.
\]
Again by (iv) of Lemma 3.1, $\rho \in B^{\alpha,z'}(\mathcal{H}, \sigma)$. \qed

Remark 3.5. By (3.3)–(3.4), for $P_n$ and $\sigma_n$ as in (3.2),
\[
P_n \varrho_{\sigma,\alpha,z} P_n = \sigma_n^{\frac{1-a}{2a}} \varrho_{\sigma,\alpha} \sigma_n^{\frac{1-a}{2a}},
\]
and if $\alpha = z$, then we further have
\[
P_n \varrho_{\sigma,\alpha} P_n = \sigma_n^{\frac{1-a}{2a}} \varrho_{\sigma,\alpha} \sigma_n^{\frac{1-a}{2a}} = (P_n \sigma P_n)^{\frac{1-a}{2a}} (P_n \varrho P_n) (P_n \sigma P_n)^{-\frac{1-a}{2a}}.
\]
Thus, with $\varrho_n := P_n \varrho P_n$,
\[
\varrho_{\sigma,\alpha} = (\text{wo}) \lim_{n \to +\infty} \sigma_n^{\frac{1-a}{2a}} \varrho_n \sigma_n^{\frac{1-a}{2a}}.
\]

Remark 3.6. Note that if $\rho \in B^{\alpha}(\mathcal{H}, \sigma)$, i.e., $\rho = \sigma^{\frac{a-1}{2a}} \varrho_{\sigma,\alpha} \sigma^{\frac{a-1}{2a}}$ with $\varrho_{\sigma,\alpha} \in \mathcal{B}(\mathcal{H})_{\geq 0}$, $\varrho_{\sigma,\alpha}^0 \leq \sigma^0$, then for any $\alpha' < \alpha$,
\[
\rho = \sigma^{\frac{a'-1}{2a'}} \varrho_{\sigma,\alpha'} \sigma^{\frac{a'-1}{2a'}} = \sigma^{\frac{a-1}{2a}} \varrho_{\sigma,\alpha} \sigma^{\frac{1-a}{2a}} - \frac{1}{2a} \sigma^{\frac{a'-1}{2a}},
\]
whence $\rho \in B^{\alpha'}(\mathcal{H}, \sigma)$, and
\[
\varrho_{\sigma,\alpha'} = \sigma^{\frac{1-a}{2a}} \varrho_{\sigma,\alpha} \sigma^{\frac{1-a}{2a}} - \frac{1}{2a}. \quad (3.8)
\]
In particular, if $\rho \in B^{\infty}(\mathcal{H}, \sigma)$, i.e., $\rho = \sigma^{1/2} \varrho_{\sigma,\infty} \sigma^{1/2}$ with some $\varrho_{\sigma,\infty} \in \mathcal{B}(\mathcal{H})_{\geq 0}$, $\varrho_{\sigma,\infty}^0 \leq \sigma^0$, then $\rho \in B^{\alpha}(\mathcal{H}, \sigma)$ for every $\alpha > 1$, and
\[
\varrho_{\sigma,\alpha} = \sigma^{\frac{1-a}{2a}} \varrho_{\sigma,\infty} \sigma^{\frac{1-a}{2a}}.$
As an immediate consequence,
\[ \cap_{\alpha>1} B^\alpha(\mathcal{H}, \sigma) \supseteq B^\infty(\mathcal{H}, \sigma) = \{ \varrho \in \mathcal{B}(\mathcal{H}) : D_{\text{max}}(\varrho\|\sigma) < +\infty\}, \]
where
\[ D_{\text{max}}(\varrho, \sigma) := \inf\{ \kappa \in \mathbb{R} : \varrho \leq e^\kappa \sigma \} \tag{3.9} \]
is the max-relative entropy of \( \varrho \) and \( \sigma \) [9, 41].

**Definition 3.7.** For \( \sigma \in \mathcal{B}(\mathcal{H}) \geq 0 \) and \( (\alpha, z) \in (1, +\infty) \times (0, +\infty) \), let
\[ \mathcal{L}^{\alpha, z}(\mathcal{H}, \sigma) := \{ \varrho \in B^{\alpha, z}(\mathcal{H}, \sigma) : \text{Tr} \varrho^{z} \varrho^{\alpha \sigma, z} < +\infty\}. \]
Again, when \( \alpha = z \), we will use the notation \( \mathcal{L}^{\alpha, \alpha}(\mathcal{H}, \sigma) := \mathcal{L}^{\alpha}(\mathcal{H}, \sigma) \).

**Remark 3.8.** Note that for \( \alpha > 1, \sigma \in \mathcal{B}(\mathcal{H}) \), and if \( z \geq 1 \) then \( \mathcal{L}^{\alpha}(\mathcal{H}) \) is an ideal in \( \mathcal{B}(\mathcal{H}) \). Thus, by (i) of Lemma 3.1, if \( \varrho \in \mathcal{L}^{\alpha, z}(\mathcal{H}, \sigma) \) then \( \varrho^{z} \in \mathcal{L}^{z}(\mathcal{H}, \sigma) \), or equivalently, \( \varrho \in \mathcal{L}^{\alpha}(\mathcal{H}) \).

Assume now that \( \sigma \) is trace-class and \( \varrho \in \mathcal{L}^{\alpha, z}(\mathcal{H}, \sigma) \) for some \( (\alpha, z) \in (1, +\infty) \times (0, +\infty) \). Then, by Lemma (i) of 3.1 and the operator Hölder inequality, \( \text{Tr} (\varrho^{z})^{\frac{\alpha}{z}} < +\infty \), where \( \frac{1}{r} = \frac{\alpha-1}{2z} + \frac{z}{z} + \frac{\alpha-1}{2z} = \frac{\alpha}{z} \), or equivalently, \( \varrho \in \mathcal{L}^{1}(\mathcal{H}) \). Thus, we get
\[ \text{trace-class} \implies \mathcal{L}^{\alpha, z}(\mathcal{H}, \sigma) \subseteq B^{\alpha, z}(\mathcal{H}, \sigma) \cap \mathcal{L}^{1}(\mathcal{H}), \quad \alpha > 1, \quad z > 0. \]

It is easy to see that the above inclusion is strict. Indeed, let \( \sigma \in \mathcal{B}(l^{2}(\mathbb{N})) \) be diagonal in the canonical basis of \( l^{2}(\mathbb{N}) \), i.e., \( \sigma = \sum_{k \in \mathbb{N}} s(k) |1_{[k]}\rangle \langle 1_{[k]}| \) for some \( s : \mathbb{N} \to (0, +\infty) \) such that \( \sum_{k \in \mathbb{N}} s(k) < +\infty \) (i.e., \( \sigma \) is trace-class) and \( \sum_{k \in \mathbb{N}} s(k) \frac{a-1}{2z} < +\infty \).

Define \( \varrho := \sum_{k \in \mathbb{N}} s(k) \frac{a-1}{2z} |1_{[k]}\rangle \langle 1_{[k]}| \). Then \( \varrho \) is trace-class, and for any sequence \( (P_{n} = I_{(c_{n}, d_{n})}(\sigma))_{n \in \mathbb{N}} \) as in Lemma 3.1,
\[ \frac{1-a}{2z} \varrho^{z} P_{n}^{\frac{1-a}{2z}} = \sum_{k : c_{n} \leq s(k) < d_{n}} |1_{[k]}\rangle \langle 1_{[k]}|, \]
which goes to \( I \) in the strong operator topology. Hence, \( \varrho \in B^{\alpha, z}(\mathcal{H}, \sigma) \cap \mathcal{L}^{1}(\mathcal{H}) \), but \( \varrho_{\sigma, \alpha, z} = I \), and therefore \( \varrho \notin \mathcal{L}^{\alpha, z}(\mathcal{H}) \).

The following is an extension of the Rényi \((\alpha, z)\)-divergences [4] to the case of infinite-dimensional PSD operators. It is also a special case of Jenčová’s definition of the sandwiched Rényi divergence [27] when \( \varrho \) and \( \sigma \) are trace-class, and \( z = \alpha \), and it is a natural extension of it otherwise.

**Definition 3.9.** For \( \varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0 \) and \( (\alpha, z) \in (1, +\infty) \times (0, +\infty) \), let
\[ Q_{\alpha, z}(\varrho\|\sigma) := \begin{cases} \text{Tr} \varrho^{z} \varrho^{\alpha \sigma, z}, & \varrho \in B^{\alpha, z}(\mathcal{H}, \sigma), \\ +\infty, & \text{otherwise}, \end{cases} \]
with \( \varrho_{\sigma, \alpha, z} \) as in Lemma 3.1. The Rényi \((\alpha, z)\)-divergence of \( \varrho \) and \( \sigma \) is defined as
\[ D_{\alpha, z}(\varrho\|\sigma) := \frac{1}{\alpha - 1} \log Q_{\alpha, z}(\varrho\|\sigma). \]
We use the notations \( Q_{*}^{\alpha} := Q_{\alpha, \alpha} \) and \( D_{*}^{\alpha} := D_{\alpha, \alpha} \), and call the latter the sandwiched Rényi \( \alpha \)-divergence.
We also define the following variants of the Rényi \((\alpha,z)\)-divergences for trace-class operators:

**Definition 3.10.** For PSD trace-class operators \(\varrho, \sigma \in \mathcal{L}^1(\mathcal{H})_{\geq 0}\) and \((\alpha, z) \in (1, +\infty) \times (0, +\infty)\), let

\[
\tilde{D}_{\alpha,z}(\varrho\|\sigma) := D_{\alpha,z}(\varrho\|\sigma) - \frac{1}{\alpha - 1} \log \text{Tr} \varrho.
\]

We also use the notation \(\tilde{D}^{*}_{\alpha} := \tilde{D}_{\alpha,\alpha}\).

**Remark 3.11.** For a convex function \(f\) on \([0, +\infty)\), the *quantum \(f\)-divergence* of a pair of positive normal functionals on a von Neumann algebra is defined using the relative modular operator; see [21,39]. In particular, it is well-defined for a pair of positive trace-class operators \(\varrho, \sigma\) on a Hilbert space and \(f_{\alpha} := \text{id}_{[0, +\infty)}\) for any \(\alpha > 1\); let it be denoted by \(Q_{f_{\alpha}}(\varrho\|\sigma)\). According to [21, Theorem 3.6],

\[
Q_{f_{\alpha}}(\varrho\|\sigma) = Q_{\alpha,1}(\varrho\|\sigma), \quad \alpha > 1.
\]

In particular, for PSD trace-class operators \(\varrho\) and \(\sigma\), \(D_{\alpha,1}(\varrho\|\sigma)\) in Definition 3.9 coincides with the Petz-type or standard quantum Rényi \(\alpha\)-divergence of \(\varrho\) and \(\sigma\), just as in the finite-dimensional case; see, e.g. [4].

**Remark 3.12.** Note that for any \(\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}\) and any \((\alpha, z) \in (1, +\infty) \times (0, +\infty)\),

\[
Q_{\alpha,z}(\varrho\|\sigma) > 0, \quad D_{\alpha,z}(\varrho\|\sigma) > -\infty,
\]

and

\[
D_{\alpha,z}(\varrho\|\sigma) < +\infty \iff Q_{\alpha,z}(\varrho\|\sigma) < +\infty \iff \varrho \in \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma).
\]

**Remark 3.13.** It is clear from their definitions that \(Q_{\alpha,z}, D_{\alpha,z}\) and \(\tilde{D}_{\alpha,z}\) satisfy the scaling properties

\[
Q_{\alpha,z}(\lambda \varrho\| \eta \sigma) = \lambda^{\alpha} \eta^{1-\alpha} Q_{\alpha,z}(\varrho\|\sigma),
\]

\[
D_{\alpha,z}(\lambda \varrho\| \eta \sigma) = D_{\alpha,z}(\varrho\|\sigma) + \frac{\alpha}{\alpha - 1} \log \lambda - \log \eta,
\]

\[
\tilde{D}_{\alpha,z}(\lambda \varrho\| \eta \sigma) = \tilde{D}_{\alpha,z}(\varrho\|\sigma) + \log \lambda - \log \eta,
\]

valid for any \(\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}\) and \(\lambda, \eta \in (0, +\infty)\).

**Remark 3.14.** According to Lemma 3.1, if \(\varrho \in \mathcal{B}^{\alpha,z}(\mathcal{H}, \sigma)\) then

\[
Q_{\alpha,z}(\varrho\|\sigma) = \text{Tr} \sigma^{\frac{\alpha}{2z}} \varrho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{2z}}
\]

which is a straightforward generalization of the formula for PSD operators on a finite-dimensional Hilbert space. Moreover, Lemma 3.1 also yields the formula

\[
Q_{\alpha,z}(\varrho\|\sigma) = \text{Tr} \left( \varrho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{2z}} \varrho^{\frac{\alpha}{2z}} \right)^{\frac{1}{z}},
\]

which generalizes the finite-dimensional expression \(\text{Tr} \left( \varrho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{2z}} \varrho^{\frac{\alpha}{2z}} \right)^{\frac{1}{z}}\). Note that by Lemma 3.1, (3.14) can also be written as

\[
Q_{\alpha,z}(\varrho\|\sigma) = \left\| \sigma^{\frac{1-\alpha}{2z}} \varrho^{\frac{\alpha}{2z}} \right\|_{2z}^{2z},
\]

where we use the notation \(\|\cdot\|_{z} = (\text{Tr} \left| \cdot \right|^{z})^{1/z}\) also for \(z \in (0, 1)\).
A further connection to the finite-dimensional formula is given by the following:

**Lemma 3.15.** Let \( q, \sigma \in B(\mathcal{H}) \geq 0 \) be such that \( q^0 \leq \sigma^0 \), and let \( (\alpha, z) \in (1, +\infty) \times [1, +\infty) \). Then \( q \in \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \), or equivalently, \( Q_{\alpha,z}(q\|\sigma) < +\infty \), if and only if for any/some sequences \( 0 < c_n < d_n \) with \( c_n \to 0 \), \( d_n \to +\infty \), \( (\sigma_n^{1-\alpha} q \sigma_n^{1-\alpha})_{n\in\mathbb{N}} \) is a convergent sequence in \( \mathcal{L}^z(\mathcal{H}) \), where \( \sigma_n \coloneqq \text{id}_{(c_n, d_n)}(\sigma) \).

Moreover, if \( q \in \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \) then

\[
\lim_{n \to +\infty} \left\| Q_{\alpha,\sigma,z} - \frac{1}{z} \sigma_n^{\frac{1}{2z}} q \sigma_n^{\frac{1}{2z}} \right\|_z = 0, \tag{3.15}
\]

and if \( q \in \mathcal{B}^{\alpha,z}(\mathcal{H}, \sigma) \) then

\[
Q_{\alpha,z}(q\|\sigma) = \lim_{n \to +\infty} \text{Tr} \left( \sigma_n^{\frac{1}{2z}} q \sigma_n^{\frac{1}{2z}} \right)^z \tag{3.16}
\]

for any sequences as above.

**Proof.** The “if” part follows since convergence in \( z \)-norm implies (so) convergence, whence \( Q_{\alpha,\sigma,z} \) exists as in Lemma 3.1, and the (so) limit coincides with the \( z \)-norm limit, whence \( Q_{\alpha,\sigma,z} \in \mathcal{L}^z(\mathcal{H}, \sigma) \).

Assume now that \( q \in \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \). Then \( \sigma_n^{1-\alpha} q \sigma_n^{1-\alpha} = P_n Q_{\alpha,\sigma,z} P_n \), with \( P_n \coloneqq 1_{(c_n, d_n)}(\sigma) \), and the “only if” part, as well as (3.15), follows from Lemma 2.2.

Note that (3.15) trivially implies (3.16) when \( q \in \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \). Assume thus that \( q \in \mathcal{B}^{\alpha,z}(\mathcal{H}, \sigma) \setminus \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \), so that \( Q_{\alpha,z}(q\|\sigma) = +\infty \). Since \( \sigma_n^{1-\alpha} q \sigma_n^{1-\alpha} = P_n Q_{\alpha,\sigma,z} P_n \) converges to \( Q_{\alpha,\sigma,z} \) in the weak operator topology, Lemma 2.4 yields that

\[
+\infty = Q_{\alpha,z}(q\|\sigma) = \left\| Q_{\alpha,\sigma,z} \right\|_z \leq \lim_{n \to +\infty} \left\| \sigma_n^{\frac{1}{2z}} q \sigma_n^{\frac{1}{2z}} \right\|_z = \lim_{n \to +\infty} \text{Tr} \left( \sigma_n^{\frac{1}{2z}} q \sigma_n^{\frac{1}{2z}} \right)^z,
\]

from which (3.16) follows.

**Proposition 3.16.** For any \( q, \sigma \in B(\mathcal{H}) \geq 0 \), and any \( \alpha \in (1, +\infty) \),

\[
Q_{\alpha,z}(q\|\sigma), \; D_{\alpha,z}(q\|\sigma), \; \tilde{D}_{\alpha,z}(q\|\sigma) \text{ are decreasing in } z.
\]

In particular, for any \( \sigma \in B(\mathcal{H}) \geq 0 \), \( \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \) is increasing in \( z \), i.e.,

\[
0 < z \leq z' \implies \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \subseteq \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma).
\]

**Proof.** It is sufficient to prove that for any \( 0 < z < z' \), \( Q_{\alpha,z}(q\|\sigma) \geq Q_{\alpha,z'}(q\|\sigma) \) holds. This is obvious when \( Q_{\alpha,z}(q\|\sigma) = +\infty \), and hence for the rest we assume the contrary, i.e., that \( q \in \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma) \). By Lemma 3.4, this implies that \( q \in \mathcal{B}^{\alpha,z}(\mathcal{H}, \sigma) \). Thus, by Lemma 3.15,

\[
Q_{\alpha,z}(q\|\sigma) = \lim_{n \to +\infty} \text{Tr} \left( \sigma_n^{\frac{1-\alpha}{2z}} q \sigma_n^{\frac{1-\alpha}{2z}} \right)^{z'}.
\]

According to Araki’s inequality [2, Theorem 2], \( \text{Tr} q(B^{1/2} A B^{1/2})^q \leq \text{Tr} q(B^{q/2} A^{q} B^{q/2}) \) for any \( A, B \in B(\mathcal{H}) \geq 0, \; q \in [1, +\infty) \), and monotone increasing continuous function \( q \).
on \([0, +\infty)\) such that \(\varphi(0) = 0\) and \(t \mapsto \varphi(e^t)\) is convex on \(\mathbb{R}\). Applying this to \(A:=\varrho_{\alpha}^{1/\alpha}\), \(B:=\sigma_{n}^{1/\alpha}\), \(q:=\frac{1}{z}\), and \(\varphi:=\text{id}\) yields

\[
\text{Tr} \left( \frac{1-\alpha}{z} q^1 \sigma_{n}^{1-\alpha} \right) z = \text{Tr} \left[ \left( \frac{1-\alpha}{z} q^1 \sigma_{n}^{1-\alpha} \right)^{z} \right] \leq \text{Tr} \left( \frac{1-\alpha}{z} q^1 \sigma_{n}^{1-\alpha} \right)^{z} \n
\]

for every \(n \in \mathbb{N}\). Thus, by (3.17),

\[
Q_{\alpha^*,\alpha}(\varrho||\sigma) \leq \lim_{n \to +\infty} \text{Tr} \left( \frac{1-\alpha}{z} q^1 \sigma_{n}^{1-\alpha} \right)^{z} = Q_{\alpha^*,\alpha}(\varrho||\sigma),
\]

where the equality is again due to Lemma 3.15.

\(\square\)

**Remark 3.17.** As a special case of Proposition 3.16, we get that for any \(\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0\),

\[
D_{\varrho}^*(\varrho||\sigma) = D_{\alpha,\alpha}(\varrho||\sigma) \leq D_{\alpha,1}(\varrho||\sigma),
\]

i.e., the sandwiched Rényi \(\alpha\)-divergence cannot be larger than the Petz-type Rényi \(\alpha\)-divergence. This has been proved for positive normal functionals on a von Neumann algebra (positive trace-class operators in our case) in [6, Theorem 12] and [27, Corollary 3.6] using different methods than in the proof of Proposition 3.16 above.

**Remark 3.18.** Assume that \(Q_{\alpha^*,\alpha}(\varrho||\sigma) < +\infty\), i.e., \(\varrho \in \mathcal{L}^{\alpha}(\varrho||\sigma)\) for some \(\alpha > 1\), and \(1 < \alpha^* < \alpha\). Then, by (3.8),

\[
Q_{\alpha^*,\alpha}(\varrho||\sigma) = \text{Tr} \varrho_{\sigma,\alpha^*}^{\alpha^*} \varrho_{\sigma,\alpha^*}^{\alpha^*} \leq \left\| \varrho_{\sigma,\alpha^*}^{\alpha^*} \right\| \left\| \varrho_{\sigma,\alpha^*}^{\alpha^*} \right\| \leq \left(\text{Tr} \sigma \right)^{\alpha^*} Q_{\alpha^*,\alpha}(\varrho||\sigma)^{\alpha^*},
\]

where the inequality follows by the operator Hölder inequality. In particular, if \(\sigma\) is trace-class then \(Q_{\alpha^*,\alpha}(\varrho||\sigma) < +\infty\). If \(\text{Tr} \sigma = 1\) then a simple rearrangement yields

\[
\frac{\alpha' - 1}{\alpha'} D_{\alpha^*,\alpha}(\varrho||\sigma) \leq \frac{\alpha - 1}{\alpha} D_{\alpha^*,\alpha}(\varrho||\sigma).
\]

Note that this is weaker than \(D_{\alpha^*,\alpha}(\varrho||\sigma) \leq D_{\alpha^*,\alpha}(\varrho||\sigma)\), which was proved in [28, Proposition 3.7].

**Remark 3.19.** Since we do not assume the second operator to be trace-class, the expression \(D_{\alpha,\alpha}(\varrho||I)\) makes sense, and we recover the following identity for the Rényi \(\alpha\)-entropy of a state \(\varrho \in \mathcal{S}(\mathcal{H})\), which is well-known in the finite-dimensional case:

\[
S_{\alpha}(\varrho) := \frac{1}{1 - \alpha} \log \text{Tr} \varrho^{\alpha} = -D_{\alpha^*,\alpha}(\varrho||I), \quad \alpha > 1.
\]

(3.18)

(In fact, this makes sense for arbitrary PSD operator \(\varrho\)).

More importantly, allowing non trace-class operators enables the definition of conditional (\(\alpha, z\))-entropies. Following [44], we define two different notions of conditional (\(\alpha, z\))-entropy between systems \(A\) and \(B\) in a state \(\varrho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)\) as

\[
S_{\alpha^*,\alpha}(A|B)^{\downarrow} := -D_{\alpha^*,\alpha}(\varrho_{AB}||I_A \otimes \varrho_B),
\]

(3.19)
where $\varrho_B = \text{Tr}_A \varrho_{AB}$ denotes the marginal of $\varrho_{AB}$ on system $B$. Again, (3.19)–(3.20) make sense even when $\varrho_{AB}$ is only assumed to be PSD. Note that while the Rényi entropies (3.18) can be defined directly for $\varrho$ without reference to any Rényi divergences, this is not the case for the conditional Rényi entropies (3.19)–(3.20), and the ability to take non-trace-class operators at least in the second argument of the divergence is crucial for the definition.

According to Proposition 3.16, for any fixed $\varrho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, and any $\alpha > 1$,

$$S_{\alpha,z}(A|B)^\dagger := - \inf_{\omega_B \in \mathcal{S}(\mathcal{H}_B)} D_{\alpha,z}(\varrho_{AB} \| I_A \otimes \omega_B),$$  (3.20)

In particular, either version of the sandwiched conditional Rényi entropy is at least as large as the corresponding version of the Petz-type conditional Rényi entropy.

**Lemma 3.20.** For any $\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0$ and $(\alpha, z) \in (1, +\infty) \times (0, +\infty)$,

$$Q_{\alpha,z}(\varrho \| \sigma) > 0, \quad D_{\alpha,z}(\varrho \| \sigma) > -\infty.$$  (3.21)

**Proof.** The assertion is trivial when $\varrho \notin \mathcal{L}^{\alpha,z}(\mathcal{H}, \sigma)$, and hence we assume the contrary. Then

$$Q_{\alpha,z}(\varrho \| \sigma) = \text{Tr} \varrho_{z}\sigma_{z}\varrho_{z} = 0 \iff \varrho_{z}\alpha,z = 0 \implies \varrho = \frac{1}{\sqrt{z}} \varrho_{z}\alpha,z = 0,$$

counter to the assumption that $\varrho \in \mathcal{B}(\mathcal{H}) \geq 0$. Hence, the inequalities in (3.21) hold. \qed

**Remark 3.21.** Stronger bounds than the ones in (3.21) are given below in Corollary 3.27 for trace-class operators.

**Lemma 3.22.** Let $\varrho_k, \sigma_k \in \mathcal{B}(\mathcal{H}_k) \geq 0$, $k = 1, 2$. For any $(\alpha, z) \in (1, +\infty) \times (0, +\infty)$,

$$\varrho_1 \otimes \varrho_2 \in \mathcal{B}^{\alpha,z}(\mathcal{H}_1 \otimes \mathcal{H}_2, \sigma_1 \otimes \sigma_2) \iff \varrho_k \in \mathcal{B}^{\alpha,z}(\mathcal{H}_k, \sigma_k), \quad k = 1, 2,$$  (3.22)

$$\varrho_1 \otimes \varrho_2 \in \mathcal{L}^{\alpha,z}(\mathcal{H}_1 \otimes \mathcal{H}_2, \sigma_1 \otimes \sigma_2) \iff \varrho_k \in \mathcal{L}^{\alpha,z}(\mathcal{H}_k, \sigma_k), \quad k = 1, 2,$$  (3.23)

and $(\varrho_1 \otimes \varrho_2)_{\alpha_1 \otimes \sigma_1, \alpha_2, \alpha, z} = (\varrho_1)_{\alpha_1, \alpha, z} \otimes (\varrho_2)_{\alpha_2, \alpha, z}$. As a consequence,

$$Q_{\alpha,z}(\varrho_1 \otimes \varrho_2 \| \sigma_1 \otimes \sigma_2) = Q_{\alpha,z}(\varrho_1 \| \sigma_1) Q_{\alpha,z}(\varrho_2 \| \sigma_2),$$  (3.24)

$$D_{\alpha,z}(\varrho_1 \otimes \varrho_2 \| \sigma_1 \otimes \sigma_2) = D_{\alpha,z}(\varrho_1 \| \sigma_1) + D_{\alpha,z}(\varrho_2 \| \sigma_2).$$  (3.25)

**Proof.** The right to left implications in (3.22, 3.23) are obvious from choosing $R := (\varrho_1)_{\alpha_1, \alpha, z} \otimes (\varrho_2)_{\alpha_2, \alpha, z}$ in (i) of Lemma 3.1. Assume that $\varrho_1 \otimes \varrho_2 \in \mathcal{B}^{\alpha,z}(\mathcal{H}_1 \otimes \mathcal{H}_2, \sigma_1 \otimes \sigma_2)$. By (iv) of Lemma 3.1, there exists a $\lambda \geq 0$ such that

$$\varrho_1^\alpha \otimes \varrho_2^\alpha \leq \lambda (\sigma_1 \otimes \sigma_2)^\alpha \otimes \sigma_2^\alpha = \lambda \sigma_1^\alpha \otimes \sigma_2^\alpha.$$  

Choose any $\psi_2 \notin \ker(\varrho_2)$. For any $\psi_1 \in \mathcal{H}_1$, we get

$$\psi_1, \varrho_1^\alpha \psi_1 \leq \lambda \psi_1, \sigma_1^\alpha \psi_1 \leq \lambda \psi_2, \varrho_2^\alpha \psi_2.$$  

\[=_{x_1 > 0} \]
Thus, \( \sigma_1^{\frac{a}{2}} \leq \lambda(\kappa_2/\kappa_1)\sigma_1^{\frac{a-1}{2}} \), and again by (iv) of Lemma 3.1, \( \varrho_1 \in B^{a,z}(\mathcal{H}_1, \sigma_1) \). An exactly analogous argument gives \( \varrho_2 \in B^{a,z}(\mathcal{H}_2, \sigma_2) \). This proves the left to right implication in (3.22), and we also get

\[
(\sigma_1 \otimes \sigma_2)^{\frac{1-a}{2}} (\varrho_1 \otimes \varrho_2)_{\frac{a}{2}} = (\sigma_1^{\frac{1-a}{2}} \varrho_1^{\frac{a}{2}}) \otimes (\sigma_2^{\frac{1-a}{2}} \varrho_2^{\frac{a}{2}}),
\]

from which \( (\varrho_1 \otimes \varrho_2)\sigma_1 \otimes \sigma_2.a \) \( = (\varrho_1)\sigma_1 \otimes (\varrho_2)\sigma_2.a \), according to (3.3), and thus (3.24) and (3.25) follow due to the multiplicativity of the trace. The left to right implication in (3.23) follows immediately from the above. \( \square \)

3.2. Variational formulas. The following variational representations of \( Q_{a,z} \) and \( D_{a,z} \) are very useful to establish their fundamental properties. We will use these variational formulas to prove monotonicity of \( Q_{a,z} \) under restrictions of the operators to subspaces (Lemma 3.32, Corollary 3.34) and to give a lower bound on the strong converse exponent (Lemma 4.2).

For \( z = a \) (the case of the sandwiched Rényi divergence), the variational formula in (3.26) was given first in [12] for finite-dimensional PSD operators, and was extended to the case of pairs of positive normal functionals on a general von Neumann algebra in [28] (see also [21, Lemma 3.19] for the case \( \alpha < 1 \)), while the variational formula in (3.27) can be obtained as an intermediate step in the proof of the first variational formula, and it was given in [5] in the finite-dimensional case.

For finite-dimensional invertible PSD operators and arbitrary \((\alpha, z) \in (1, +\infty) \times (0, +\infty)\), both variational formulas (3.26)–(3.27) follow as special cases of [47, Theorem 3.3].

The version below is an extension of the above when \((\alpha, z) \in (1, +\infty) \times (0, +\infty)\) are arbitrary, and the operators \( \varrho, \sigma \) can be PSD operators on an infinite-dimensional Hilbert space satisfying the conditions in Lemma 3.23. Our proof follows essentially that of [47, Theorem 3.3].

For any \( \sigma \in B(\mathcal{H})_{\geq 0} \), let

\[
B(\mathcal{H})_{\sigma,a,z} := \left\{ H \in B(\mathcal{H})_{\geq 0} : \text{Tr} \left( H^{1/2} \sigma^{\frac{a-1}{2}} H^{1/2} \right)^{\frac{1}{z-1}} < +\infty \right\},
\]

\[
B(\mathcal{H})^+_{\sigma,a,z} := \left\{ H \in B(\mathcal{H})_{\geq 0} : 0 < \text{Tr} \left( H^{1/2} \sigma^{\frac{a-1}{2}} H^{1/2} \right)^{\frac{1}{z-1}} < +\infty \right\}.
\]

Lemma 3.23. Let \( \varrho, \sigma \in B(\mathcal{H})_{\geq 0} \) and \((\alpha, z) \in (1, +\infty) \times (0, +\infty)\), and assume that one of the following holds: a) \( \varrho \notin B^{a,z}(\mathcal{H}, \sigma) \); b) \( \varrho \in B^{a,z}(\mathcal{H}, \sigma) \setminus L^{a,z}(\mathcal{H}, \sigma) \) and \( \sigma \) is compact; c) \( \varrho \in L^{a,z}(\mathcal{H}, \sigma) \). Then

\[
Q_{a,z}(\varrho \| \sigma) = \sup_{H \in B(\mathcal{H})_{\sigma,a,z}} \left\{ \alpha \text{Tr} \left( H^{1/2} \varrho^{\frac{a}{2}} H^{1/2} \right)^{\frac{1}{z}} \right\} + (1 - \alpha) \text{Tr} \left( H^{1/2} \sigma^{\frac{a-1}{2}} H^{1/2} \right)^{\frac{1}{z-1}}, \tag{3.26}
\]

\[
\log Q_{a,z}(\varrho \| \sigma) = \sup_{H \in B(\mathcal{H})^+_{\sigma,a,z}} \left\{ \alpha \log \text{Tr} \left( H^{1/2} \varrho^{\frac{a}{2}} H^{1/2} \right)^{\frac{1}{z}} \right\} + (1 - \alpha) \log \text{Tr} \left( H^{1/2} \sigma^{\frac{a-1}{2}} H^{1/2} \right)^{\frac{1}{z-1}}. \tag{3.27}
\]

The equality in (3.26) still holds if the supremum is taken over \( B(\mathcal{H})^+_{\sigma,a,z} \). Moreover, in cases a) and b), and in case c) if \( \sigma \) is compact, the \( H \) operators in (3.26) and (3.27) may additionally be required to be of finite rank.
Proof. For any \( H \in B(\mathcal{H}) \), let

\[
F(H) := \text{Tr} \left( H^{1/2} \sigma^{\alpha - 1} \sigma^{\alpha - 1} H^{1/2} \right), \quad G(H) := \text{Tr} \left( H^{1/2} \sigma^{\alpha - 1} H^{1/2} \right).
\]

(3.28)

Assume first that \( \rho \notin B^{\alpha,z}(\mathcal{H}, \sigma) \), and hence \( Q_{\alpha,z}(\rho \| \sigma) = \log Q_{\alpha,z}(\rho \| \sigma) = +\infty \).

By (iv) of Lemma 3.1, for every \( \lambda > 0 \) there exists a vector \( x_\lambda \in \mathcal{H} \) such that

\[
\left\langle x_\lambda, \sigma^{\alpha - 1} x_\lambda \right\rangle > \lambda \langle x_\lambda, \sigma^{\alpha - 1} x_\lambda \rangle.
\]

(3.29)

Clearly, for any \( x \in \mathcal{H}, H_x := |x\rangle \langle x| \in B(\mathcal{H})_{\sigma,\alpha,z} \cap B_f(\mathcal{H}), \) and

\[
F(H_x) = \left( x, \sigma^{\alpha - 1} x \right)^{\widehat{\sigma}}, \quad G(H_x) = \langle x, \sigma^{\alpha - 1} x \rangle^{\widehat{\sigma}}.
\]

If \( \langle x_\lambda, \sigma^{\alpha - 1} x_\lambda \rangle = 0 \) for some \( \lambda > 0 \) then let \( x_{\lambda,t} := tx_\lambda + t^{-1} y, \ t > 0, \) where \( y \in (\ker \sigma)^{\perp} \setminus \{0\} \) is some fixed vector. Then \( H_{x_{\lambda,t}} \in B(\mathcal{H})_{\sigma,\alpha,z} \cap B_f(\mathcal{H}), \) (3.29) implies that \( \left\langle x_\lambda, \sigma^{\alpha - 1} x_\lambda \right\rangle > 0, \) and it is straightforward to verify that

\[
\lim_{t \to +\infty} F(H_{x_{\lambda,t}}) = +\infty, \quad \lim_{t \to +\infty} G(H_{x_{\lambda,t}}) = 0.
\]

Thus,

\[
\lim_{t \to +\infty} \left( \alpha F(H_{x_{\lambda,t}}) + (1 - \alpha) G(H_{x_{\lambda,t}}) \right) = +\infty = Q_{\alpha,z}(\rho \| \sigma) = \log Q_{\alpha,z}(\rho \| \sigma)
\]

\[
= \lim_{t \to +\infty} \left( \alpha \log F(H_{x_{\lambda,t}}) + (1 - \alpha) \log G(H_{x_{\lambda,t}}) \right),
\]

and therefore (3.26)–(3.27) hold.

If \( \langle x_\lambda, \sigma^{\alpha - 1} x_\lambda \rangle > 0 \) for every \( \lambda > 0 \) then let \( \tilde{x}_\lambda := x_\lambda, \sigma^{\alpha - 1} x_\lambda \)^{-1/2}. Then

\[
\langle \tilde{x}_\lambda, \sigma^{\alpha - 1} \tilde{x}_\lambda \rangle = 1 = \langle \tilde{x}_\lambda, \sigma^{\alpha - 1} \tilde{x}_\lambda \rangle^{\widehat{\sigma}} = G(H_{\tilde{x}_\lambda}),
\]

(3.30)

and

\[
F(H_{\tilde{x}_\lambda}) = \left( \tilde{x}_\lambda, \sigma^{\alpha - 1} \tilde{x}_\lambda \right)^{\widehat{\sigma}} > \left( \lambda \langle \tilde{x}_\lambda, \sigma^{\alpha - 1} \tilde{x}_\lambda \rangle \right)^{\widehat{\sigma}} = \lambda \tilde{\sigma},
\]

according to (3.29) and (3.30). Thus,

\[
\lim_{\lambda \to +\infty} \left( \alpha F(H_{\tilde{x}_\lambda}) + (1 - \alpha) G(H_{\tilde{x}_\lambda}) \right) = +\infty = Q_{\alpha,z}(\rho \| \sigma) = \log Q_{\alpha,z}(\rho \| \sigma)
\]

\[
= \lim_{\lambda \to +\infty} \left( \alpha \log F(H_{\tilde{x}_\lambda}) + (1 - \alpha) \log G(H_{\tilde{x}_\lambda}) \right),
\]

and therefore (3.26)–(3.27) hold, even with the optimizations restricted to finite-rank operators.

This completes the proof of case a), and hence for the rest we assume that b) or c) holds.

Consider any sequences \( 0 < c_n < d_n \) with \( c_n \to 0, d_n \to +\infty \), and let \( P_n := 1_{(c_n, d_n)}(\sigma), \sigma_n := \text{id}_{(c_n, d_n)}(\sigma) = P_n \sigma P_n \), and

\[
H_n := \sigma_n^{1 - \alpha} \left( \sigma_n^{1 - \alpha} \sigma_n^{1 - \alpha} \right)^{\alpha - 1} \sigma_n^{1 - \alpha} \sigma_n^{1 - \alpha} = \sigma_n^{1 - \alpha} (P_n \Theta_{\sigma,\alpha,z} P_n)^{\alpha - 1} \sigma_n^{1 - \alpha}.
\]
Then
\[
F(H_n) = \text{Tr} \left( Q^{1/2} H_n^{1/2} \right)^z = \text{Tr} \left( \sum_{\alpha} \frac{1}{2} Q^{1/2} H_n^{1/2} \right)^z = \text{Tr} \left( \sum_{\alpha} \left( \sigma_n^{1/2} Q^{1/2} \sigma_n^{1/2} \right)^{z/2} \right) = \text{Tr} \left( \sigma_n^{1/2} Q^{1/2} \sigma_n^{1/2} \right)^z = \text{Tr} \left( P_n Q_{\sigma,\alpha,z} P_n \right)^z,
\]
and similarly,
\[
G(H_n) = \text{Tr} \left( H_n^{1/2} Q^{1/2} H_n^{1/2} \right)^z = \text{Tr} \left( \sigma_n^{1/2} H_n^{1/2} \right)^{z/2} = \text{Tr} \left( \sigma_n^{1/2} \right)^z = \text{Tr} \left( P_n \right)^z = \text{Tr} \left( P_n Q_{\sigma,\alpha,z} P_n \right)^z.
\]

We have
\[
\text{Tr}(P_n Q_{\sigma,\alpha,z} P_n)^z = \text{Tr}(Q_{\sigma,\alpha,z} P_n)^z \leq \text{Tr} Q_{\sigma,\alpha,z} = Q_{\alpha,z}(\sigma \| \sigma);
\]
in particular,
\[
\limsup_{n \to +\infty} \text{Tr}(P_n Q_{\sigma,\alpha,z} P_n)^z \leq Q_{\alpha,z}(\sigma \| \sigma). \tag{3.33}
\]

Moreover, if \( Q_{\alpha,z}(\sigma \| \sigma) < +\infty \), i.e., in case \( c \), or if \( \sigma \) is compact (in which case \( H_n \) and \( P_n Q_{\sigma,\alpha,z} P_n \) are of finite rank) then \( F(H_n) = G(H_n) < +\infty \), whence \( H_n \in \mathcal{B}(\mathcal{H})_{\sigma,\alpha,z} \).

Since \( \varrho \in \mathcal{B}^{\sigma,\alpha,z}(\mathcal{H}, \sigma) \) implies \( \varrho^0 \leq \sigma^0 \), it is also true that \( 0 \neq \varrho \sigma_n^{1/2} Q^{1/2} \sigma_n^{1/2} \), and hence \( H_n \in \mathcal{B}(\mathcal{H})_{\sigma,\alpha,z}^+ \), for all large enough \( n \).

When \( z \geq 1 \), Lemma 2.4 yields
\[
Q_{\alpha,z}(\sigma \| \sigma) = \text{Tr} Q_{\sigma,\alpha,z} = \left\| Q_{\sigma,\alpha,z} \right\|_z = \liminf_{n \to +\infty} \left\| P_n Q_{\sigma,\alpha,z} P_n \right\|_z = \liminf_{n \to +\infty} \text{Tr}(P_n Q_{\sigma,\alpha,z} P_n)^z. \tag{3.34}
\]

When \( z \in (0, 1) \), \( \text{id}^z_{[0, +\infty)} \) is operator concave, and hence \( (P_n Q_{\sigma,\alpha,z} P_n) \geq P_n Q_{\alpha,z} P_n \), whence
\[
\liminf_{n \to +\infty} \text{Tr}(P_n Q_{\sigma,\alpha,z} P_n)^z \geq \liminf_{n \to +\infty} \text{Tr} P_n Q_{\sigma,\alpha,z} P_n = \text{Tr} Q_{\sigma,\alpha,z} = Q_{\alpha,z}(\sigma \| \sigma). \tag{3.35}
\]
Combining (3.31)–(3.35) gives
\[
Q_{\alpha,z}(\sigma \| \sigma) = \lim_{n \to +\infty} \text{Tr} \left( \sigma_n^{1/2} Q^{1/2} \sigma_n^{1/2} \right)^z
\]
The variational formulas in (3.26)–(3.27) hold for the sandwiched quantities (3.26) to obtain (3.27).

It is interesting that one can formally take the logarithm of each term in Remark 3.24.

This completes the proof when $Q_{\alpha,z}(\varrho||\sigma) = +\infty$, i.e., in case b).

Assume for the rest that case c) holds. By the above considerations, we have LHS $\leq$ RHS in (3.26)–(3.27), and hence we only have to show the converse inequalities. By Lemma 3.1 and Definition 3.7, $\varrho^{\frac{a}{z}} = \sigma^{\frac{a-1}{z}} \varrho_{\sigma,\alpha,z}^{\sigma^{\frac{a}{z}}} \sigma^{\frac{a-1}{z}} \varrho_{\sigma,\alpha,z}$, where $\operatorname{Tr} \varrho_{\sigma,\alpha,z}^{\sigma^{\frac{a}{z}}} < +\infty$. For any $H \in \mathcal{B}(\mathcal{H})_{\sigma,\alpha,z}$, we have

$$\operatorname{Tr} \left( \frac{1}{2} \varrho_{\alpha,z}^{\frac{1}{2}} H^{1/2} \right) = \operatorname{Tr} \left( \varrho_{\alpha,z}^{\frac{1}{2}} H \varrho_{\alpha,z}^{\frac{1}{2}} \right) = \operatorname{Tr} \left( H^{1/2} \varrho_{\alpha,z}^{\frac{1}{2}} \right) = \operatorname{Tr} \left( H^{1/2} \varrho_{\alpha,z}^{\frac{1}{2}} \right)$$

$$= \left\| H^{1/2} \sigma^{\frac{a-1}{z}} \frac{1-a}{z} \varrho^{\frac{1}{z-\lambda}} \right\| \leq \left\| H^{1/2} \sigma^{\frac{a-1}{z}} \frac{1-a}{z-\lambda} \varrho^{\frac{1}{z-\lambda}} \frac{1}{z-\lambda} \right\|_2$$

$$= \left[ \operatorname{Tr} \left( \sigma^{\frac{a-1}{z}} H \sigma^{\frac{a-1}{z}} \right) \frac{1}{z-\lambda} \right] \alpha \left[ \operatorname{Tr} \left( \left( \sigma^{\frac{1-a}{z}} \varrho^{\frac{1}{z-\lambda}} \right) H \left( \sigma^{\frac{1-a}{z}} \varrho^{\frac{1}{z-\lambda}} \right) \right) \right]^{1/2}$$

$$\leq \frac{\alpha - 1}{\alpha} \operatorname{Tr} \left( \sigma^{\frac{a-1}{z}} H \sigma^{\frac{a-1}{z}} \right) \frac{1}{z-\lambda} + \frac{1}{\alpha} Q_{\alpha,z}(\varrho||\sigma),$$

where we used that ran $\varrho^{\frac{a}{z}} \subseteq \operatorname{dom} \sigma^{\frac{a-1}{z}}$ and $\sigma^{\frac{a}{z}} \varrho^{\frac{a}{z}} \in \mathcal{B}(\mathcal{H})$, according to Lemma 3.1, and the expression in (3.14) for $Q_{\alpha,z}(\varrho||\sigma)$. The first inequality above is due to the operator Hölder inequality, and the second inequality is trivial from the convexity of the exponential function. A simple rearrangement yields that LHS $\geq$ RHS in (3.26, 3.27), completing the proof.

Remark 3.24. It is interesting that one can formally take the logarithm of each term in (3.26) to obtain (3.27).

Remark 3.25. The variational formulas in (3.26)–(3.27) hold for the sandwiched quantities ($z = \alpha$) when $\varrho \in \mathcal{B}^{\alpha}(\mathcal{H}, \sigma) \setminus \mathcal{L}^\alpha(\mathcal{H}, \sigma)$ even if $\sigma$ is not compact [20]. However, we won’t need this fact in the rest of the paper.

Remark 3.26. Note that the case $z = 1$ corresponds to the Petz-type Rényi divergences. By the above, $\varrho \in \mathcal{B}^{\alpha,1}(\mathcal{H}, \sigma)$ if and only if $\varrho^{\alpha} \leq \lambda \sigma^{\alpha-1}$ with some $\lambda \geq 0$, in which case

$$Q_{\alpha}(\varrho||\sigma) := Q_{\alpha,1}(\varrho||\sigma) = \operatorname{Tr} \varrho^{\frac{1-a}{z}} \varrho^{\frac{1-a}{z}}.$$

(See [21, Theorem 3.6] for a generalization of the above in the setting of von Neumann algebras, and also for an analogous formula in the case \( \alpha \in (0, 1) \).) Moreover, we have the variational formulas

\[
Q_\alpha (\varrho \parallel \sigma) = \sup_{H \in B(\mathcal{H})_{\sigma, 1}^+} \left\{ \alpha \operatorname{Tr} (H^{1/2} \varrho^\alpha H^{1/2})^{1/\alpha} + (1 - \alpha) \operatorname{Tr} (H^{1/2} \sigma^{\alpha - 1} H^{1/2})^{1/\alpha - 1} \right\},
\]

\[
\log Q_\alpha (\varrho \parallel \sigma) = \sup_{H \in B(\mathcal{H})_{\sigma, 1}^+} \left\{ \alpha \log \operatorname{Tr} (H^{1/2} \varrho^\alpha H^{1/2})^{1/\alpha} + (1 - \alpha) \log \operatorname{Tr} (H^{1/2} \sigma^{\alpha - 1} H^{1/2})^{1/\alpha - 1} \right\},
\]

where \( B(\mathcal{H})_{\sigma, 1}^+ = \{ H \in B(\mathcal{H}) \geq 0 : 0 < \operatorname{Tr} (H^{1/2} \sigma^{\alpha - 1} H^{1/2})^{1/\alpha - 1} < +\infty \} \).

These variational expressions for the Petz-type Rényi divergences do not seem to have appeared in the literature before, even for finite-dimensional operators, although in that case they follow easily from the results of [47].

The variational formulas in Lemma 3.23 can be used to prove the following important properties of the Rényi \((\alpha, z)\)-divergences.

**Corollary 3.27.** Let \( \varrho, \sigma \in B(\mathcal{H}) \geq 0 \) be such that \( \sigma \) is trace-class. For every \((\alpha, z) \in (1, +\infty) \times (0, +\infty) \),

\[
Q_{\alpha, z} (\varrho \parallel \sigma) \geq (\operatorname{Tr} \varrho)^{\alpha} (\operatorname{Tr} \sigma)^{1 - \alpha} \geq \alpha \operatorname{Tr} \varrho + (1 - \alpha) \operatorname{Tr} \sigma. \quad (3.40)
\]

If, moreover, \( \varrho \) is trace-class then we have

\[
Q_{\alpha, z} (\varrho \parallel \sigma) = (\operatorname{Tr} \varrho)^{\alpha} (\operatorname{Tr} \sigma)^{1 - \alpha} \iff \sigma = \eta \varrho \text{ for some } \eta \in (0, +\infty), \quad (3.41)
\]

and

\[
Q_{\alpha, z} (\varrho \parallel \sigma) = \alpha \operatorname{Tr} \varrho + (1 - \alpha) \operatorname{Tr} \sigma \iff \sigma = \varrho. \quad (3.42)
\]

**Proof.** The second inequality in (3.40) follows simply from the convexity of \( \text{id}_{[0, +\infty)}^\varrho \). The first inequality is obvious when \( Q_{\alpha, z} (\varrho \parallel \sigma) = +\infty \), and hence we may assume the contrary, i.e., that \( \varrho \in \mathcal{L}^{\alpha, z}(\mathcal{H}, \sigma) \). The assumption that \( \sigma \) is trace-class yields that \( I \in B(\mathcal{H})_{\sigma, z}^+ \), and the variational formula in (3.27) with \( H := I \) yields the first inequality in (3.40). In fact, we don’t need the “full power” of the variational formula in (3.27) to obtain the first inequality in (3.40), as it follows simply from the Hölder inequality as in (3.36)–(3.39), with \( H = I \).

Assume for the rest that \( \varrho \) is trace-class. The right to left implications are straightforward to verify in both (3.41) and (3.42). Assume now that the equality on the LHS of (3.41) holds. It implies that \( Q_{\alpha, z} (\varrho \parallel \sigma) \) is finite, i.e., \( \varrho \in \mathcal{L}^{\alpha, z}(\mathcal{H}, \sigma) \), and the inequality in (3.37) holds as an equality for \( H = I \). Thus, by the characterization of the equality case in Hölder’s inequality (Lemma 2.1), \( \sigma = \lambda \left( \left( \varrho^{\frac{1}{2\alpha}} \sigma^{\frac{1}{2\alpha}} \right)^{\frac{2\alpha}{2\alpha}} \right) = \lambda \varrho_{\sigma, z}^{\frac{1}{2\alpha}} \) for some \( \lambda > 0 \). From this we get \( \sigma_n^{\frac{1}{\alpha}} = \lambda \sigma_n^{\frac{1}{2\alpha}} \varrho_{\sigma, z}^{\frac{1}{2\alpha}} \) and

\[
\sigma_n^{\frac{1}{\alpha}} = P_n \sigma^{\frac{1}{\alpha}} P_n = \lambda \sigma_n^{\frac{1}{2\alpha}} P_n \varrho_{\sigma, z}^{\frac{1}{2\alpha}} P_n = \lambda \sigma_n^{\frac{1}{2\alpha}} \varrho_{\sigma, z}^{\frac{1}{2\alpha}} \sigma_n^{\frac{1}{2\alpha}}
\]

\[
\sigma_n^{\frac{1}{\alpha}} = P_n \sigma^{\frac{1}{\alpha}} P_n = \lambda \sigma_n^{\frac{1}{2\alpha}} P_n \varrho_{\sigma, z}^{\frac{1}{2\alpha}} P_n = \lambda \sigma_n^{\frac{1}{2\alpha}} \varrho_{\sigma, z}^{\frac{1}{2\alpha}} \sigma_n^{\frac{1}{2\alpha}}
\]
for every $n \in \mathbb{N}$, where $P_n$ is as in Lemma 3.1. Rearranging yields $\sigma_n^{\frac{a}{z}} = \lambda \frac{1}{n} P_n \rho_n^{\frac{a}{z}} P_n$. Thus,

$$
\sigma^{\frac{a}{z}} = (\text{so}) \lim_n \sigma_n^{\frac{a}{z}} = \lambda \frac{1}{n} (\text{so}) \lim_n P_n \rho_n^{\frac{a}{z}} P_n = \lambda \frac{1}{z} Q^{\frac{a}{z}}.
$$

(In the last equality we use that $Q \in \mathcal{L}^{a,z}(\mathcal{H}, \sigma)$ implies $\varrho^0 \leq \sigma^0$). Hence, $\sigma = \lambda \frac{1}{z} Q$, i.e., the RHS of (3.41) holds true.

Finally, assume that the equality on the LHS of (3.42) is true. By (3.40), this implies that the equality on the LHS of (3.41) is true, and hence, by the above, $\sigma = \eta \varrho$ for some $\eta \in (0, +\infty)$. Moreover, the second equality in (3.40) holds as an equality, whence $\text{Tr} \varrho = \text{Tr} \sigma$, so we get $\varrho = \sigma$ as given on the RHS of (3.42). □

**Corollary 3.28.** For any $(\alpha, z) \in (1, +\infty) \times (0, +\infty)$, the Rényi $(\alpha, z)$-divergence $D_{\alpha,z}$ is strictly positive in the sense that for any two density operators $\varrho, \sigma \in S(\mathcal{H})$,

$$
D_{\alpha,z}(\varrho \| \sigma) \geq 0, \quad \text{with equality if and only if } \varrho = \sigma.
$$

**Proof.** Immediate from Corollary 3.27. □

**Remark 3.29.** Non-negativity of the Rényi $(\alpha, z)$-divergences has been proved in [34] in the finite-dimensional case, by different methods. Strict positivity of the sandwiched Rényi $\alpha$-divergences with $\alpha > 1$ has been proved in the general von Neumann algebra case in [27].

Finally, we prove the lower semi-continuity of $Q_{\alpha,z}$ and $D_{\alpha,z}$ on pairs of trace-class operators from the variational formula; we will use this later in the proof of Lemma 3.39.

**Corollary 3.30.** For any $\alpha > 1$ and $z \geq \alpha$, $Q_{\alpha,z}$ and $D_{\alpha,z}$ are lower semi-continuous on $\mathcal{L}^1(\mathcal{H}) \times \mathcal{L}^1(\mathcal{H})$.

**Proof.** Let $\varrho_n, \sigma_n \in \mathcal{L}^1(\mathcal{H})$, $n \in \mathbb{N}$, be convergent sequences in trace-norm, with $\varrho := \lim_{n \to +\infty} \varrho_n$, $\sigma := \lim_{n \to +\infty} \sigma_n$. Then

$$
\|\varrho^{\frac{a}{z}} - \sigma^{\frac{a}{z}}\|_{\frac{1}{a}} = (\text{Tr} \varrho_n)^{\frac{a}{z}} = \|\varrho_n\|_{\frac{1}{a}} \xrightarrow{n \to +\infty} \|\varrho\|_{\frac{1}{a}} = (\text{Tr} \varrho)^{\frac{a}{z}} = \|\varrho^{\frac{a}{z}}\|_{\frac{1}{a}}.
$$

Since $\|\varrho_n - \varrho\|_{\infty} \leq \|\varrho_n - \varrho\|_{1} \to 0$, the continuity of the functional calculus implies $\|\varrho_n^{\frac{a}{z}} - \varrho^{\frac{a}{z}}\|_{\infty} \to 0$; in particular, $\varrho_n^{\frac{a}{z}} \to \varrho^{\frac{a}{z}}$ in the strong operator topology. Hence, by Lemma 2.3, $\lim_n \|\varrho_n^{\frac{a}{z}} - \varrho^{\frac{a}{z}}\|_{\frac{1}{a}} = 0$. Thus, for any $H \in \mathcal{B}(\mathcal{H})_{\geq 0}$,

\[
\left\| H^{1/2} \varrho_n^{\frac{a}{z}} H^{1/2} - H^{1/2} \varrho^{\frac{a}{z}} H^{1/2} \right\|_{\frac{1}{a}} \leq \left\| H^{1/2} \varrho_n^{\frac{a}{z}} H^{1/2} - H^{1/2} \varrho^{\frac{a}{z}} H^{1/2} \right\|_{\frac{1}{a}} \xrightarrow{n \to +\infty} 0.
\]

This shows that for any $H \in \mathcal{B}(\mathcal{H})_{\geq 0}$, $\varrho \mapsto \text{Tr} \left( H^{1/2} \varrho^{\frac{a}{z}} H^{1/2} \right)^{\frac{1}{a}} = \|H^{1/2} \varrho^{\frac{a}{z}} H^{1/2}\|_{\frac{1}{a}}^{\frac{1}{a}}$ is continuous on $\mathcal{L}^1(\mathcal{H})$, and continuity of $\sigma \mapsto \text{Tr} \left( H^{1/2} \sigma^{a-1} H^{1/2} \right)^{\frac{1}{a-1}}$ on $\mathcal{L}^1(\mathcal{H})$ can be proved in the same way. Thus, by Lemma 3.23, $\mathcal{L}^1(\mathcal{H}) \times \mathcal{L}^1(\mathcal{H}) \ni (\varrho, \sigma) \mapsto Q_{\alpha,z}(\varrho, \sigma)$ is the supremum of continuous functions, and hence it is upper semi-continuous. The assertion about the lower semi-continuity of $D_{\alpha,z}$ follows trivially from this. □
Remark 3.31. Lower semi-continuity of the sandwiched Rényi $\alpha$-divergences for $\alpha > 1$ (i.e., $z = \alpha > 1$) was given in [27, Proposition 3.10] in the general von Neumann algebra setting, with a different proof.

3.3. Finite-dimensional approximations. Our next goal is to investigate the relation between the sandwiched Rényi divergences of finite-dimensional restrictions of the operators and those of the unrestricted operators. We start with the following:

Lemma 3.32. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$, and $K \in \mathcal{B}(\mathcal{H}, \mathcal{K})^+_{\varrho, \sigma}$ be a contraction. For any $(\alpha, z) \in (1, +\infty) \times (0, +\infty)$ with $\max\{\alpha - 1, \alpha/2\} \leq z \leq \alpha$,

$$Q_{\alpha, z}(K \varrho K^* \| K \sigma K^*) \leq Q_{\alpha, z}(\varrho \| \sigma).$$  \hfill (3.43)

Proof. By assumption, $\text{id}_{[0, +\infty)}$ is operator convex and $\text{id}_{[0, +\infty)}^{a-1}$ is operator concave, whence

$$(K \varrho K^*)^{\frac{a}{z}} \leq K \varrho^{\frac{a}{z}} K^*, \quad (K \sigma K^*)^{\frac{a-1}{z}} \geq K \sigma^{\frac{a-1}{z}} K^*,$$  \hfill (3.44)

according to the operator Jensen inequality [7, Theorem 11].

If $\varrho \notin \mathcal{L}^{a, \alpha}_{\varrho, \sigma}(\mathcal{H}, \sigma)$ then $Q_{\alpha, z}(\varrho \| \sigma) = +\infty$ by definition, and (3.43) holds trivially. Hence, for the rest we assume that $\varrho \in \mathcal{L}^{a, \alpha}_{\varrho, \sigma}(\mathcal{H}, \sigma)$. Lemma 3.1 yields the existence of some $\lambda \geq 0$ such that $\varrho^{\frac{a}{z}} \leq \lambda \sigma^{\frac{a-1}{z}}$. Thus,

$$(K \varrho K^*)^{\frac{a}{z}} \leq K \varrho^{\frac{a}{z}} K^* \leq \lambda K \sigma^{\frac{a-1}{z}} K^* \leq \lambda (K \sigma K^*)^{\frac{a-1}{z}},$$

where the first and the last inequalities are due to (3.44). Hence, again by Lemma 3.1, $K \varrho K^* \in \mathcal{L}^{a, \alpha}_{\varrho, \sigma}(\mathcal{H}, K \sigma K^*)$; in particular, the variational formulas in Lemma 3.23 hold for $K \varrho K^*$ and $K \sigma K^*$ in place of $\varrho$ and $\sigma$, respectively.

For any $H \in \mathcal{B}(\mathcal{K})_{\lambda \sigma, \alpha, z}$,

$$+\infty > \text{Tr} \left( H^{1/2} (K \sigma K^*)^{\frac{a-1}{z}} H^{1/2} \right)^{\frac{a}{a-1}} \geq \text{Tr} \left( H^{1/2} K \sigma^{\frac{a-1}{z}} K^* H^{1/2} \right)^{\frac{a}{a-1}} = \text{Tr} \left( \sigma^{\frac{a-1}{z}} K^* H K \sigma^{\frac{a-1}{z}} \right)^{\frac{a}{a-1}} = \text{Tr} \left( (K^* H K)^{1/2} \sigma^{\frac{a-1}{z}} (K^* H K)^{1/2} \right)^{\frac{a}{a-1}},$$  \hfill (3.45)

where the second inequality is due to (3.44). In particular, $K^* H K \in \mathcal{B}(\mathcal{H})_{\sigma, \alpha, z}$. Similarly,

$$\text{Tr} \left( H^{1/2} (K \varrho K^*)^{\frac{a}{z}} H^{1/2} \right)^{\frac{a}{a-1}} \leq \text{Tr} \left( H^{1/2} K \varrho^{\frac{a}{z}} K^* H^{1/2} \right)^{\frac{a}{a-1}} = \text{Tr} \left( \varrho^{\frac{a}{z}} K^* H K \varrho^{\frac{a}{z}} \right)^{\frac{a}{a-1}} = \text{Tr} \left( (K^* H K)^{1/2} \varrho^{\frac{a}{z}} (K^* H K)^{1/2} \right)^{\frac{a}{a-1}}.$$  \hfill (3.46)

Plugging (3.45)–(3.46) into the variational formula yields

$$Q_{\alpha, z}(K \varrho K^* \| K \sigma K^*) = \sup_{H \in \mathcal{B}(\mathcal{K})_{\lambda \sigma \alpha, z}} \left\{ \alpha \text{Tr} \left( H^{1/2} (K \varrho K^*)^{\frac{a}{z}} H^{1/2} \right)^{\frac{a}{a-1}} \right\}.$$
\[ + (1 - \alpha) \operatorname{Tr} \left( H^{1/2} (K \sigma K^\ast) \frac{\alpha - 1}{\alpha} H^{1/2} \right) \frac{z}{\alpha - 1} \]

\[ \leq \sup_{H \in B(K_{K^\ast,\alpha,z})} \left\{ \alpha \operatorname{Tr} \left( (K^\ast H K)^{1/2} Q_\alpha^{\sigma} (K^\ast H K)^{1/2} \right) \frac{z}{\alpha} \right\} + (1 - \alpha) \operatorname{Tr} \left( (K^\ast H K)^{1/2} \frac{\alpha - 1}{\alpha} (K^\ast H K)^{1/2} \right) \frac{z}{\alpha - 1} \]

\[ \leq \sup_{H \in B(H_{\alpha,z})} \left\{ \alpha \operatorname{Tr} \left( H^{1/2} Q_\alpha^{\sigma} H^{1/2} \right) + (1 - \alpha) \operatorname{Tr} \left( H^{1/2} \frac{\alpha - 1}{\alpha} H^{1/2} \right) \right\} \]

\[ = Q_{\alpha,z}(\varrho \| \sigma). \]

\[ \square \]

\textbf{Remark 3.33.} When \( z = \alpha \) and \( K = P \) is a projection in Lemma 3.32, one could appeal to the monotonicity of \( Q_\alpha^* \) under positive trace-preserving maps, and its additivity on direct sums [27, Proposition 3.11], to obtain the inequality (3.43) as

\[ Q_\alpha^*(\varrho \| \sigma) \geq Q_\alpha^*(P \varrho P + P \perp \varrho P \perp \| P \sigma P + P \perp \sigma P \perp) \]

\[ = Q_\alpha^*(P \varrho P \| P \sigma P) + Q_\alpha^*(P \perp \varrho P \perp \| P \perp \sigma P \perp) \]

\[ \geq Q_\alpha^*(P \varrho P \| P \sigma P). \]

Note, however, that these properties were only proved in [27] for positive normal functionals, i.e., positive trace-class operators in our setting, and hence this argument gives (3.43) in a restricted setting compared to that of Lemma 3.32, even when we only consider \( z = \alpha \) and reductions by projections.

Recall that the set of projections on \( \mathcal{H} \) is an upward directed partially ordered set w.r.t. the PSD order. Lemma 3.32 yields the following:

\textbf{Corollary 3.34.} Let \( \varrho, \sigma \in B(\mathcal{H})_{\geq 0} \) and \( (\alpha, z) \) as in Lemma 3.32. For any contraction \( K \in B(\mathcal{H})_{\varrho,\sigma}^+, \) and any projection \( P \in \mathbb{P}(\mathcal{H}) \) such that \( |K| \leq P, \)

\[ Q_{\alpha,z}(K \varrho K^\ast \| K \sigma K^\ast) \leq Q_{\alpha,z}(P \varrho P \| P \sigma P). \quad (3.47) \]

\[ \text{In particular,} \]

\[ \mathbb{P}(\mathcal{H})_{\varrho,\sigma}^+ \ni P \mapsto Q_{\alpha,z}(P \varrho P \| P \sigma P) \text{ is increasing.} \quad (3.48) \]

\textbf{Proof.} Since \( K (P \varrho P) K^\ast = K \varrho K^\ast \) and \( K (P \sigma P) K^\ast = K \sigma K^\ast, \) (3.47) follows immediately by replacing \( \varrho \) with \( P \varrho P \) and \( \sigma \) with \( P \sigma P \) in Lemma 3.32. The monotonicity in (3.48) follows immediately from this. \( \square \)

\textbf{Definition 3.35.} For any \( \varrho, \sigma \in B(\mathcal{H})_{\geq 0} \) and \( (\alpha, z) \in (1, +\infty) \times (0, +\infty), \) let

\[ Q_{\alpha,z}(\varrho \| \sigma)_{fa} := \sup_{P \in \mathbb{P}(\mathcal{H})_{\varrho,\sigma}^+} Q_{\alpha,z}(P \varrho P \| P \sigma P), \]

\[ D_{\alpha,z}(\varrho \| \sigma)_{fa} := \sup_{P \in \mathbb{P}(\mathcal{H})_{\varrho,\sigma}^+} D_{\alpha,z}(P \varrho P \| P \sigma P) = \frac{1}{\alpha - 1} \log Q_{\alpha,z}(\varrho \| \sigma)_{fa}, \]

be the finite-dimensional approximations of \( Q_{\alpha,z}(\varrho \| \sigma) \) and \( D_{\alpha,z}(\varrho \| \sigma), \) respectively. If, moreover, \( \varrho \) is trace-class then we also define

\[ \tilde{D}_{\alpha,z}(\varrho \| \sigma)_{fa} := D_{\alpha,z}(\varrho \| \sigma)_{fa} - \frac{1}{\alpha - 1} \log \operatorname{Tr} \varrho. \]
Remark 3.36. It is clear from (3.10) that for any $\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0$ and $(\alpha, z) \in (1, +\infty) \times (0, +\infty)$,

$$Q_{\alpha, z}(\varrho \parallel \sigma)_{fa} > 0, \quad D_{\alpha, z}(\varrho \parallel \sigma)_{fa} > -\infty. \quad (3.49)$$

Lemma 3.37. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0$ and $(\alpha, z)$ as in Lemma 3.32. Then

$$Q_{\alpha, z}(\varrho \parallel \sigma)_{fa} \leq Q_{\alpha, z}(\varrho \parallel \sigma), \quad (3.50)$$

and

$$Q_{\alpha, z}(\varrho \parallel \sigma)_{fa} = \sup \left\{ Q_{\alpha, z}(T \varrho T \parallel T \sigma T) : T \in \mathcal{B}(\mathcal{H})_{0,1} \cap \mathcal{B}_f(\mathcal{H})_{\varrho, \sigma}^+ \right\}.$$

Proof. Immediate from Corollary 3.34. \hfill \Box

Our next goal is to see when equality in (3.50) holds.

Lemma 3.38. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\varrho, \sigma}^+$ and $1 < \alpha = z$ be such that $\varrho \in \mathcal{B}^{\alpha, z}(\mathcal{H}, \sigma)$, let $0 < c_n < d_n, n \in \mathbb{N}$, be sequences such that $c_n \to 0, d_n \to +\infty$, and $P_n := 1(c_n, d_n)(\sigma)$. Then

$$Q_{\alpha, z}(\varrho \parallel \sigma) \leq \liminf_{n \to +\infty} Q_{\alpha, z}(P_n \varrho P_n \parallel P_n \sigma P_n). \quad (3.51)$$

Proof. Note that by assumption, $\varrho^0 \leq \sigma^0$, and for every large enough $n$, $P_n \in \mathcal{B}(\mathcal{H})_{\varrho, \sigma}^+$. By Lemma 3.1,

$$\varrho_{\alpha, \sigma, z} = (\text{wo}) \lim_{n \to +\infty} \left( P_n \sigma P_n \right)^{\frac{1-\alpha}{2\alpha}} \left( \frac{P_n \varrho P_n}{P_n \sigma P_n} \right)^{\frac{1-\alpha}{2\alpha}} \leq (P_n \varrho P_n)^{\frac{1-\alpha}{2\alpha}}, \quad (3.52)$$

where the inequality follows from the operator Jensen inequality [7, Theorem 11] due to the fact that $\alpha/z \in (0, 1]$ by assumption. Hence,

$$Q_{\alpha, z}(\varrho \parallel \sigma) = \left\| \varrho_{\alpha, \sigma, z} \right\|_{z}^{\frac{z}{z}} \leq \liminf_{n \to +\infty} \left\| \left( P_n \sigma P_n \right)^{\frac{1-\alpha}{2\alpha}} \left( P_n \varrho P_n \right)^{\frac{1-\alpha}{2\alpha}} \right\|_{z}^{\frac{z}{z}} \leq \liminf_{n \to +\infty} \left\| \left( P_n \sigma P_n \right)^{\frac{1-\alpha}{2\alpha}} \left( P_n \varrho P_n \right)^{\frac{z}{z}} \right\|_{z}^{\frac{z}{z}} = \liminf_{n \to +\infty} Q_{\alpha, z}(P_n \varrho P_n \parallel P_n \sigma P_n),$$

where the first inequality is due to Lemma 2.4, and the second inequality follows from (3.52). \hfill \Box

The range of $(\alpha, z)$ pairs to which both Lemma 3.37 and Lemma 3.38 apply is $1 < \alpha \leq z$, i.e., the case of the sandwiched Rényi divergences, and hence for the rest we restrict to this case. Fortunately, this is sufficient for the intended applications in the rest of the paper.
Lemma 3.39. Let \( \varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \) be trace-class, and \( K_n \in \mathcal{B}(\mathcal{H}, \mathcal{K})^+_{\varrho, \sigma} \), \( n \in \mathbb{N} \), be contractions such that
\[
\exists \ (\text{so}) \lim_{n \to +\infty} K_n =: K_\infty \in \mathcal{B}(\mathcal{H}, \mathcal{K})^+_{\varrho, \sigma}, \quad \exists \ (\text{so}) \lim_{n \to +\infty} K_n^*.
\]
(That is, \( (K_n)_{n \in \mathbb{N}} \) converges in the strong* operator topology.) Then
\[
Q_\alpha^*(K_\infty \varrho K_\infty^* \| K_\infty \sigma K_\infty^*) = \lim_{n \to +\infty} Q_\alpha^*(K_n \varrho K_n^* \| K_n \sigma K_n^*), \quad (3.53)
\]
\[
\leq \lim_{n \to +\infty} \sup Q_\alpha^*(K_n \varrho K_n^* \| K_n \sigma K_n^*) \leq Q_\alpha^*(\varrho \| \sigma). \quad (3.54)
\]
In particular, if \( P_n \in \mathbb{P}(\mathcal{H})_{\varrho, \sigma}^\perp \), \( n \in \mathbb{N} \), is a sequence of projections strongly converging to some \( P_\infty \) with \( P_\infty \varrho P_\infty = \varrho \) and \( P_\infty \sigma P_\infty = \sigma \) then
\[
\lim_{n \to +\infty} Q_\alpha^*(P_n \varrho P_n \| P_n \sigma P_n) = Q_\alpha^*(\varrho \| \sigma).
\]
Proof. The second inequality in (3.54) is obvious from Lemma 3.32, and the first inequality is trivial. By the assumptions and Lemma 2.2,
\[
\lim_{n \to +\infty} \| K_n \varrho K_n^* - K_\infty \varrho K_\infty^* \|_1 = 0 = \lim_{n \to +\infty} \| K_n \sigma K_n^* - K_\infty \sigma K_\infty^* \|_1. \quad (3.55)
\]
Since \( Q_\alpha^* \) is lower semi-continuous on \( \mathcal{L}^1(\mathcal{H}) \times \mathcal{L}^1(\mathcal{H}) \) (see Corollary 3.30, or [27, Proposition 3.10]), we get the inequality in (3.53). The last assertion follows obviously.

Lemmas 3.37, 3.38, and 3.39 imply immediately the following:

Proposition 3.40. Let \( \varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \), and assume that \( \varrho \) and \( \sigma \) are trace-class, or that \( \sigma \) is compact and \( \varrho \in \mathcal{B}^\sigma(\mathcal{H}, \sigma) \). Then
\[
Q_\alpha^*(\varrho \| \sigma) = Q_\alpha^*(\varrho \| \sigma)_{\text{fa}} = \lim_{P_{\mathcal{H}}(\mathcal{H})_{\varrho, \sigma}^+ \ni P \not\ni I} Q_\alpha^*(P \varrho P \| P \sigma P), \quad (3.56)
\]
\[
= \lim_{n \to +\infty} Q_\alpha^*(P_n \varrho P_n \| P_n \sigma P_n), \quad (3.57)
\]
\[
D_\alpha^*(\varrho \| \sigma) = D_\alpha^*(\varrho \| \sigma)_{\text{fa}} = \lim_{P_{\mathcal{H}}(\mathcal{H})_{\varrho, \sigma}^+ \ni P \not\ni I} D_\alpha^*(P \varrho P \| P \sigma P), \quad (3.58)
\]
\[
= \lim_{n \to +\infty} D_\alpha^*(P_n \varrho P_n \| P_n \sigma P_n), \quad (3.59)
\]
for every \( \alpha > 1 \), where the convergence in (3.56) and (3.58) is a net convergence in the strong operator topology, and (3.57) and (3.59) hold for any sequence \( (P_n)_{n \in \mathbb{N}} \) as in Lemma 3.38. If, moreover, \( \varrho \) is trace-class then
\[
\tilde{D}_\alpha^*(\varrho \| \sigma) = \tilde{D}_\alpha^*(\varrho \| \sigma)_{\text{fa}} = \lim_{P_{\mathcal{H}}(\mathcal{H})_{\varrho, \sigma}^+ \ni P \not\ni I} \tilde{D}_\alpha^*(P \varrho P \| P \sigma P), \quad (3.60)
\]
\[
= \lim_{n \to +\infty} \tilde{D}_\alpha^*(P_n \varrho P_n \| P_n \sigma P_n), \quad \alpha > 1.
\]
Remark 3.41. Finite-dimensional approximability for the standard $f$-divergences was given in [19, Theorem 4.5] in the general von Neumann algebra setting. In particular, it shows that for any two PSD trace-class operators on a Hilbert space, the standard (or Petz-type) Rényi divergences satisfy

$$D_{α,1}(φ∥σ) = D_{α,1}(φ∥σ)_{fa} := \sup_{P \in \mathcal{P}(\mathcal{H})_{0,α}^+} D_{α,1}(PφP∥PσP), \quad α \in [0, 2].$$

It is an open question whether finite-dimensional approximability holds for any $(α, z)$ pairs other than $α \in [0, 2]$ and $z = 1$, and $z = α \in (1, +∞)$.

There are cases apart from the ones treated in Proposition 3.40 where the inequality in (3.50) holds with equality. In particular, we have the following trivial case, which we will use in the proof of Proposition 4.3.

Lemma 3.42. Let $φ, σ ∈ \mathcal{B}(\mathcal{H})_{≥0}$.

If $φ^0 \nleq σ^0$ then $Q_{α, z}(φ∥σ)_{fa} = +∞ = Q_{α, z}(φ∥σ), \quad α ∈ (1, +∞), \quad z ∈ (0, +∞).

(3.61)

Proof. Assume that $φ^0 \nleq σ^0$, so that there exists a unit vector $ψ ∈ \mathcal{H}$ such that $σ^0ψ = 0$, $φ^0ψ ≠ 0$. Let $φ$ be any unit vector such that $σ^0 φ = φ$, and for every $t ∈ [0, 1]$, define $ψ_t := \sqrt{1-t}ψ + \sqrt{t}φ$, $P_t := |ψ_t⟩⟨ψ_t|$. Then

$$P_t φ P_t = |ψ_t⟩⟨ψ_t| φψ_t →_{t→0} |ψ⟩⟨ψ| φψ, \quad P_t σ P_t = |ψ_t⟩⟨ψ_t| σψ_t →_{t→0} 0,$$

while $P_t σ P_t ≠ 0$ for every $t ∈ (0, 1]$. Thus,

$$Q_{α, z}(φ∥σ)_{fa} ≥ \lim_{t ↓ 0} Q_{α, z}(P_t φ P_t∥P_t σ P_t) = \lim_{t ↓ 0} ⟨ψ_t, φψ_t⟩^α ⟨ψ_t, σψ_t⟩^{1-α} = +∞.$$

Since $φ^0 \nleq σ^0$ implies that $φ \notin \mathcal{B}^{α,z}(\mathcal{H}, σ)$ (see Lemma 3.1), we also get $Q_{α, z}(φ∥σ) = +∞$. \quad \square

The finite-dimensional approximability of the sandwiched Rényi divergences in Proposition 3.40 is the key property used in proving the main results of the paper, the equality of the sandwiched and the regularized measured Rényi divergences, and the determination of the strong converse exponent of state discrimination, in Sects. 3.4 and 4.1.

The following monotonicity result has been proved for finite-rank states in [35], and for states of a general von Neumann algebra in [6, 27]. We give a different proof of it in our setting as an illustration of the use of the finite-dimensional approximability in extending finite-dimensional results to infinite dimension. We will give yet another proof in Sect. 3.4, using a different representation of the sandwiched Rényi divergences.

Corollary 3.43. Let $φ, σ ∈ \mathcal{L}^1(\mathcal{H})_{≥0}$ be PSD trace-class operators. Then

$$(1, +∞) \ni α ↦ \tilde{D}_α^*(φ∥σ) \text{ is increasing}, \quad (3.62)$$

and

$$\lim_{α \to +∞} \tilde{D}_α^*(φ∥σ) = \lim_{α \to +∞} D_α^*(φ∥σ) = D_{max}(φ∥σ). \quad (3.63)$$
Proof. These are well-known when $\varrho$ and $\sigma$ are finite-rank [35]. Thus, by (3.60), the monotonicity in (3.62) holds. This also shows that the first limit in (3.63) exists, and it is trivial by definition that it is equal to the second limit. To show the last equality in (3.63), it is sufficient to consider the case when $\varrho$ and $\sigma$ are density operators, due to the scaling properties in Remark 3.13. Then $D^*_\alpha(\varrho\|\sigma) = \tilde{D}^*_\alpha(\varrho\|\sigma)$ for every $\alpha > 1$, and

$$\lim_{\alpha \to +\infty} D^*_\alpha(\varrho\|\sigma) = \lim_{\alpha \to +\infty} \tilde{D}^*_\alpha(\varrho\|\sigma) = \sup_{\alpha > 1} \tilde{D}^*_\alpha(\varrho\|\sigma) = \sup_{\alpha > 1} D^*_\alpha(\varrho\|\sigma)$$

$$= \sup_{\alpha > 1, P \in \mathbb{P}_f(H)_{\varrho,\sigma}} D^*_\alpha(P\varrho P\|P\sigma P)$$

$$= \sup_{\alpha > 1, P \in \mathbb{P}_f(H)_{\varrho,\sigma}} \sup \left\{ \tilde{D}^*_\alpha(P\varrho P\|P\sigma P) + \frac{1}{\alpha - 1} \log \text{Tr} P\varrho P \right\}$$

$$= \sup_{\alpha > 1, P \in \mathbb{P}_f(H)_{\varrho,\sigma}} D_{\text{max}}(P\varrho P\|P\sigma P) = D_{\text{max}}(\varrho\|\sigma).$$

Here, the first three equalities are trivial, and the fourth one follows by Proposition 3.40. The fifth equality is again trivial, and the sixth one is by definition. In the seventh equality we use that both $\alpha \mapsto \tilde{D}^*_\alpha(P\varrho P\|P\sigma P)$ and $\alpha \mapsto \frac{1}{\alpha - 1} \log \text{Tr} P\varrho P$ are increasing, and hence the supremum of their sum over $\alpha > 1$ is the sum of their limits at $\alpha \to +\infty$, which is equal to $D_{\text{max}}(P\varrho P\|P\sigma P)$, according to the known behaviour in the finite-dimensional case. The last equality is straightforward to verify. \[\square\]

3.4. Regularized measured Rényi divergence. A finite-outcome positive operator-valued measure (POVM) on a Hilbert space $H$ is a map $M : \{r\} \to B(H)$, where $\{r\} := \{1, \ldots, r\}$, all $M_i$ is PSD, and $\sum_{i=1}^r M_i = I$. (We assume without loss of generality that the set of possible outcomes is a subset of $\mathbb{N}$.) We denote the set of such POVMs by $\mathbb{P}(H, \{r\})$. For two PSD trace-class operators $\varrho, \sigma \in \mathcal{L}^1(H)_{\geq 0}$, their measured Rényi divergence is defined as

$$D^\text{meas}_\alpha(\varrho\|\sigma) := \sup_{r \in \mathbb{N}} \sup_{M \in \mathbb{P}(H,\{r\})} D_\alpha\left(\left(\text{Tr} \varrho M_i\right)_{i \in [r]}\left|\right|\left(\text{Tr} \sigma M_i\right)_{i \in [r]}\right),$$

where in the second expression we have the classical Rényi divergence [42] of the given non-negative functions on $[r]$. This is defined for $p, q \in [0, +\infty)^{[r] \setminus \{0\}}$ as

$$D_\alpha(p\|q) := \begin{cases} \frac{1}{\alpha - 1} \log \sum_{i \in [r]} p(i)^\alpha q(i)^{1-\alpha}, & \text{supp } p \subseteq \text{supp } q, \\ +\infty, & \text{otherwise}. \end{cases}$$

One might consider more general POVMs for the definition, but that does not change the value of the measured Rényi divergence; see, e.g., [21, Proposition 5.2]. The regularized measured Rényi divergence of $\varrho$ and $\sigma$ is then defined as

$$\overline{D}^\text{meas}_\alpha(\varrho\|\sigma) := \sup_{n \in \mathbb{N}} \frac{1}{n} D^\text{meas}_\alpha(\varrho^\otimes n\|\sigma^\otimes n) = \lim_{n \to \infty} \frac{1}{n} D^\text{meas}_\alpha(\varrho^\otimes n\|\sigma^\otimes n).$$

The following has been shown in [33]:
Lemma 3.44. For finite-rank PSD operators $\varrho$, $\sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$,

$$D_{\alpha}^{\text{meas}}(\varrho\|\sigma) = D_{\alpha}^{\times}(\varrho\|\sigma), \quad \alpha > 1.$$ 

In the proof of the next theorem, we will use the monotonicity of the sandwiched Rényi $\alpha$-divergences under finite-outcome measurements for $\alpha > 1$. The more general statement of monotonicity under quantum operations has been proved in [6, Theorem 14] and [27, Theorem 3.14] in the general von Neumann algebra setting. We give a different proof for trace-class operators on a Hilbert space in Corollary 4.15 below.

Theorem 3.45. Let $\varrho$, $\sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$ be trace-class, and $\alpha > 1$. Then

$$\overline{D}_{\alpha}^{\text{meas}}(\varrho\|\sigma) = D_{\alpha}^{\times}(\varrho\|\sigma).$$

Proof. The inequality $\overline{D}_{\alpha}^{\text{meas}}(\varrho\|\sigma) \leq D_{\alpha}^{\times}(\varrho\|\sigma)$ is trivial from the monotonicity of $D_{\alpha}^{\times}$ under quantum operations and its additivity under tensor products (Lemma 3.22), and hence we only need to prove the converse inequality. By Proposition 3.40, for any $c < D_{\alpha}^{\times}(\varrho\|\sigma)$ there exists a finite-rank projection $P \in \mathbb{P}_{f}(\mathcal{H})_{\geq 0}$ such that

$$c < D_{\alpha}^{\times}(P\varrho P\|P\sigma P) \leq D_{\alpha}^{\times}(\varrho\|\sigma).$$

By Lemma 3.44, there exist $n \in \mathbb{N}$, a number $r \in \mathbb{N}$, and $M_i \in \mathcal{B}(\mathcal{H}^{\otimes n})_{\geq 0}$, $M_i^0 \leq P^{\otimes n}$, $i \in [r]$, with $\sum_{i\in[r]} M_i = P^{\otimes n}$, such that

$$c < \frac{1}{n} D_{\alpha}((\operatorname{Tr}(P\varrho P)^{\otimes n} M_i)_{i\in[r]} \| (\operatorname{Tr}(P\sigma P)^{\otimes n} M_i)_{i\in[r]}).$$

(3.64) Let us define $\tilde{M}_i := M_i$, $i \in [r]$, and $\tilde{M}_{r+1} := I_{\mathcal{H}^{\otimes n}} - P^{\otimes n}$. Then $(\tilde{M}_i)_{i\in[r+1]}$ is a POVM on $\mathcal{H}^{\otimes n}$, and we have

$$c < \frac{1}{n} D_{\alpha}((\operatorname{Tr}(P\varrho P)^{\otimes n} \tilde{M}_i)_{i\in[r]} \| (\operatorname{Tr}(P\sigma P)^{\otimes n} \tilde{M}_i)_{i\in[r]})$$

$$= \frac{1}{n} D_{\alpha}((\operatorname{Tr} \varrho^{\otimes n} \tilde{M}_i)_{i\in[r]} \| (\operatorname{Tr} \sigma^{\otimes n} \tilde{M}_i)_{i\in[r]})$$

$$\leq \frac{1}{n} D_{\alpha}((\operatorname{Tr} \varrho^{\otimes n} \tilde{M}_i)_{i\in[r+1]} \| (\operatorname{Tr} \sigma^{\otimes n} \tilde{M}_i)_{i\in[r+1]})$$

$$\leq \frac{1}{n} D_{\alpha}^{\text{meas}}(\varrho^{\otimes n}\|\sigma^{\otimes n})$$

$$\leq \overline{D}_{\alpha}^{\text{meas}}(\varrho\|\sigma),$$

where the first inequality is by (3.64), the equality and the second inequality are trivial, and the third and the fourth inequalities are by definition. Thus, $c < \overline{D}_{\alpha}^{\text{meas}}(\varrho\|\sigma)$, and since the above holds for every $c < D_{\alpha}^{\times}(\varrho\|\sigma)$, the assertion follows. \hfill $\Box$

Their representation given in Theorem 3.45 distinguishes the sandwiched Rényi divergences among all quantum generalizations of the classical Rényi divergences; in particular, it gives special importance to the $\alpha = 1$ case in the family of Rényi $(\alpha, z)$-divergences, at least for $\alpha > 1$. It also allows to deduce some important properties of the sandwiched Rényi divergences from those of the classical Rényi divergences; we present such an example in Corollary 3.46. Note that the properties in Corollary 3.46 were also proved in [6,27] in the general von Neumann algebra setting, by different methods. Yet another proof was given in our setting in Corollary 3.43.
Corollary 3.46. Let \( \varrho, \sigma \in L^1(\mathcal{H}) \geq 0 \) be PSD trace-class operators. Then

\[(1, +\infty) \ni \alpha \mapsto \tilde{D}_\alpha^*(\varrho \| \sigma) \text{ is increasing}, \tag{3.65}\]

and

\[
\sup_{\alpha > 1} \tilde{D}_\alpha^*(\varrho \| \sigma) = \lim_{\alpha \to +\infty} \tilde{D}_\alpha^*(\varrho \| \sigma) = \lim_{\alpha \to +\infty} D_\alpha^*(\varrho \| \sigma) \tag{3.66}
\]

\[
= D_{\max}(\varrho \| \sigma) \tag{3.67}
\]

\[
= \log \inf \{ \lambda > 0 : \varrho \leq \lambda \sigma \} \tag{3.68}
\]

\[
= \log \sup \left\{ \frac{\operatorname{Tr} \varrho T}{\operatorname{Tr} \sigma T} : T \in B(\mathcal{H})[0,1], \operatorname{Tr} \sigma T > 0 \right\}. \tag{3.69}
\]

Proof. The increasing property in (3.65) is well-known and easy to verify for commuting finite-rank states (i.e., in the finite-dimensional classical setting). The general case follows immediately from this and Theorem 3.45. The first equality in (3.66) is immediate from the increasing property in (3.65), and the second equality is trivial by definition.

Note that the equality in (3.68) is by definition (see (3.9)), and it is clear that \( D_{\max}(\varrho \| \sigma) \) is an upper bound on (3.69). To prove the converse inequality, note first that (3.69) is equal to \(+\infty\) if \( \varrho \not\leq \sigma \), and hence for the rest we assume the contrary. Let \( 0 < \lambda < \exp(D_{\max}(\varrho \| \sigma)) \). By definition, there exists a unit vector \( \psi \in \mathcal{H} \) such that \( \langle \psi, \varrho \psi \rangle > \lambda \langle \psi, \sigma \psi \rangle \). In particular, \( \langle \psi, \varrho \psi \rangle > 0 \), and hence also \( \langle \psi, \sigma \psi \rangle > 0 \), due to the assumption that \( \varrho \not\leq \sigma \). Choosing \( T := |\psi \rangle \langle \psi| \) shows that (3.69) is lower bounded by \( \log \lambda \) for any such \( \lambda \), and hence it is also lower bounded by \( D_{\max}(\varrho \| \sigma) \). Thus, we get the equality in (3.69).

It is also straightforward to verify that the expressions in (3.66) are upper bounded by \( D_{\max}(\varrho \| \sigma) \). To prove the converse inequality, note that for any test \( T \) as in (3.69),

\[
D_\alpha^*(\varrho \| \sigma) = \tilde{D}_\alpha^\text{meas}(\varrho \| \sigma) \geq D_\alpha\left( (\operatorname{Tr} \varrho T, \operatorname{Tr} \varrho (I - T)), (\operatorname{Tr} \sigma T, \operatorname{Tr} \sigma (I - T)) \right)
\]

\[
\geq \frac{1}{\alpha - 1} \log \left( (\operatorname{Tr} \varrho T)^\alpha (\operatorname{Tr} \sigma T)^{1-\alpha} \right) = \frac{\alpha}{\alpha - 1} \log \operatorname{Tr} \varrho T - \log \operatorname{Tr} \sigma T,
\]

whence

\[
\lim_{\alpha \to +\infty} D_\alpha^*(\varrho \| \sigma) \geq \log \frac{\operatorname{Tr} \varrho T}{\operatorname{Tr} \sigma T}.
\]

Taking the supremum over \( T \) yields that \( \lim_{\alpha \to +\infty} D_\alpha^*(\varrho \| \sigma) \) is lower bounded by (3.69), which in turn is equal to \( D_{\max}(\varrho \| \sigma) \) by the above.

\[\square\]

3.5. The Hoeffding anti-divergences. For any \( \varrho, \sigma \in B(\mathcal{H}) \geq 0 \), let

\[
\psi^*(\varrho \| \sigma | \alpha) := \log Q_\alpha^*(\varrho \| \sigma) = (\alpha - 1) D_\alpha^*(\varrho \| \sigma), \quad \alpha > 1,
\]

\[
\tilde{\psi}^*(\varrho \| \sigma | u) := (1 - u) \psi^*(\varrho \| \sigma | (1 - u)^{-1}), \quad u \in (0, 1).
\]

We will need these quantities to define the Hoeffding anti-divergences, which will give the strong converse exponent of state discrimination in Sect. 4.
Lemma 3.47. Let $\varrho, \sigma \in B(\mathcal{H})_{\geq 0}$ with $\varrho^0 \leq \sigma^0$.

(i) For any finite-rank projection $P \in \mathbb{P}_f (\mathcal{H})_{\varrho,\sigma}$, \( \psi^* (P \varrho P \| P \sigma P) \) and \( \tilde{\psi}^* (P \varrho P \| P \sigma P) \) are finite-valued convex functions on $(1, +\infty)$ and $(0, 1)$, respectively, and hence they are continuous. Moreover,

\[
\psi^* (P \varrho P \| P \sigma P |1):= \lim_{u \searrow 0} \tilde{\psi}^* (P \varrho P \| P \sigma P |u) = \log \text{Tr} \ P \varrho P,
\]

and the so extended functions $\psi^* (P \varrho P \| P \sigma P \cdot)$ and $\tilde{\psi}^* (P \varrho P \| P \sigma P \cdot)$ are convex and continuous on $[1, +\infty)$ and on $[0, 1]$, respectively.

(ii) For every $\alpha \in [1, +\infty]$, $P \mapsto \psi^* (P \varrho P \| P \sigma P |u)$ is monotone increasing on $\mathbb{P}_f (\mathcal{H})_{\varrho,\sigma}$.

(iii) For every $u \in [0, 1]$, $P \mapsto \tilde{\psi}^* (P \varrho P \| P \sigma P |u)$ is monotone increasing on $\mathbb{P}_f (\mathcal{H})_{\varrho,\sigma}$.

Proof. By [33, Corollary 3.11], $\psi^* (P \varrho P \| P \sigma P \cdot)$ is a finite-valued convex function on $(1, +\infty)$. Hence, it can be written as $\psi^* (P \varrho P \| P \sigma P |\alpha) = \sup_{t \in [c; \alpha + d]}$ for some $[c_1, d_1] \subseteq \mathbb{R}$, $\alpha \in (1, +\infty)$, and with some $t \in [c_1, d_1]$. This implies that $\psi^* (P \varrho P \| P \sigma P |u) = (1-u) \sup_{t \in [c_1, d_1]}$ for $u \leq 1$, and therefore $\tilde{\psi}^* (P \varrho P \| P \sigma P \cdot)$ is also convex and finite-valued on $(0, 1)$, and thus it is continuous as well. The limits in (3.70)–(3.71) follow by a straightforward computation, using in the second limit that $\lim_{\alpha \to +\infty} D_n^*(\omega \| \tau) = D_{\max}(\omega \| \tau)$ for finite-rank states $\omega, \tau$ (see [35] or Corollary 3.46). Convexity and continuity of the extensions are obvious from the definitions. Monotonicity in (ii) and (iii) are immediate from Corollary 3.34. \hfill \Box

Corollary 3.48. For any $\varrho, \sigma \in B(\mathcal{H})_{\geq 0}$, the functions

\[
\psi^* (\varrho \| \sigma |\alpha)_{\text{fa}} := \begin{cases} 
\sup_{P \in \mathbb{P}_f (\mathcal{H})_{\varrho,\sigma}} \psi^* (P \varrho P \| P \sigma P |\alpha), & \alpha \in [1, +\infty], \\
+\infty, & \alpha \in (-\infty, 1),
\end{cases}
\]

\[
\tilde{\psi}^* (\varrho \| \sigma |u)_{\text{fa}} := \begin{cases} 
\sup_{P \in \mathbb{P}_f (\mathcal{H})_{\varrho,\sigma}} \tilde{\psi}^* (P \varrho P \| P \sigma P |u), & u \in [0, 1], \\
+\infty, & u \in \mathbb{R} \setminus [0, 1],
\end{cases}
\]

\[
\tilde{\psi}^* (\varrho \| \sigma |u)_{\text{ra}} := \begin{cases} 
\sup_{n \in \mathbb{N}} \frac{1}{n} \tilde{\psi}^* (\varrho^{\otimes n} \| \sigma^{\otimes n} |u)_{\text{fa}}, & u \in [0, 1], \\
+\infty, & u \in \mathbb{R} \setminus [0, 1],
\end{cases}
\]

are convex and lower semi-continuous on $\mathbb{R}$ (and on $\mathbb{R} \cup (+\infty)$ in the case of $\psi^* (\varrho \| \sigma \cdot |\alpha)_{\text{fa}}$).

Proof. If $\varrho^0 \not\leq \sigma^0$ then all three functions are easily seen to be constant $+\infty$ on $\mathbb{R}$, and hence for the rest we assume that $\varrho^0 \leq \sigma^0$. Since the supremum of convex functions is again convex, and the supremum of lower semi-continuous functions is again lower semi-continuous, both properties hold for the above functions on $[1, +\infty)$, $[0, 1]$, and $[0, 1]$, respectively, according to Lemma 3.47, and it is trivial to verify that the same is true on the whole of $\mathbb{R}$. \hfill \Box

Remark 3.49. It is clear that

$\tilde{\psi}^* (\varrho \| \sigma |u)_{\text{fa}} = (1-u) \psi^* (\varrho \| \sigma |1/(1-u))_{\text{fa}} = (1-u) \log Q^*_1/(1-u) (\varrho \| \sigma)_{\text{fa}}, \quad u \in (0, 1)$,
and (3.70)–(3.71) yield
\[ \tilde{\psi}^* (\varrho \| \sigma | 0)_{fa} = \log \text{Tr} \varrho = \tilde{\psi}^* (\varrho \| \sigma | 0)_{\tilde{\varrho}}, \] (3.72)
\[ \tilde{\psi}^* (\varrho \| \sigma | 1)_{fa} = D_{\max} (\varrho \| \sigma) = \tilde{\psi}^* (\varrho \| \sigma | 1)_{\tilde{\varrho}}. \] (3.73)

This motivates to define
\[ \tilde{\psi}^* (\varrho \| \sigma | 0) := \log \text{Tr} \varrho, \] (3.74)
\[ \tilde{\psi}^* (\varrho \| \sigma | 1) := D_{\max} (\varrho \| \sigma). \] (3.75)

**Remark 3.50.** By Corollary 3.48, if \( \varrho \) and \( \sigma \) are such that \( Q^*_\alpha (\varrho \| \sigma) = Q^*_\alpha (\varrho \| \sigma)_{fa}, \alpha > 1 \), then \( \psi^* (\varrho \| \sigma | \cdot) \) and \( \tilde{\psi}^* (\varrho \| \sigma | \cdot) \) are convex and lower semi-continuous on \((1, +\infty)\) and on \((0, 1)\), respectively. In particular, this holds when both \( \varrho \) and \( \sigma \) are trace-class, according to Proposition 3.40.

Recall the definition of the finite-dimensional approximation of the sandwiched Rényi divergences as a special case of Definition 3.35: For \( \varrho, \sigma \in B(\mathcal{H}) \geq 0 \),
\[ D^*_\alpha (\varrho \| \sigma)_{fa} := \sup_{\varrho, \sigma \in \mathcal{P}(\mathcal{H})_{0, \sigma}} D^*_\alpha (P \varrho P \| P \sigma P) = \frac{1}{\alpha - 1} \psi^* (\varrho \| \sigma | \alpha)_{fa}. \]

Analogously, we define
\[ D^*_\alpha (\varrho \| \sigma)_{\tilde{\varrho}} := \sup_{\varrho, \sigma \in \mathcal{P}(\mathcal{H})_{0, \sigma}} D^*_\alpha (\tilde{\varrho} \| \sigma) = \frac{1}{\alpha - 1} \tilde{\psi}^* (\varrho \| \sigma | \alpha)_{\tilde{\varrho}}. \]

**Definition 3.51.** For \( \varrho, \sigma \in B(\mathcal{H}) \geq 0 \) and \( r \in \mathbb{R} \), let
\[ H^*_r (\varrho \| \sigma) := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left[ r - D^*_\alpha (\varrho \| \sigma) \right] = \sup_{u \in (0, 1)} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u) \right\} \] (3.76)
\[ H^*_r (\varrho \| \sigma)_{fa} := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left[ r - D^*_\alpha (\varrho \| \sigma)_{fa} \right] = \sup_{u \in (0, 1)} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u)_{fa} \right\} \] (3.77)
\[ H^*_r (\varrho \| \sigma)_{\tilde{\varrho}} := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left[ r - D^*_\alpha (\tilde{\varrho} \| \sigma) \right] = \sup_{u \in (0, 1)} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u)_{\tilde{\varrho}} \right\} \] (3.78)
\[ \hat{H}^*_r (\varrho \| \sigma) := \sup_{u \in [0, 1]} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u) \right\} \] (3.79)
\[ \hat{H}^*_r (\varrho \| \sigma)_{fa} := \max_{u \in [0, 1]} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u)_{fa} \right\} = \max_{u \in \mathbb{R}} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u)_{fa} \right\} \] (3.80)
\[ \hat{H}^*_r (\varrho \| \sigma)_{\tilde{\varrho}} := \max_{u \in [0, 1]} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u)_{\tilde{\varrho}} \right\} = \max_{u \in \mathbb{R}} \left\{ ur - \tilde{\psi}^* (\varrho \| \sigma | u)_{\tilde{\varrho}} \right\}. \] (3.81)

Here, \( H^*_r (\varrho \| \sigma) \) and \( \hat{H}^*_r (\varrho \| \sigma) \) are two different versions of the Hoeffding anti-divergence of \( \varrho \) and \( \sigma \) with parameter \( r \in \mathbb{R} \), and the rest of the quantities are different finite-dimensional approximations.

**Remark 3.52.** \( H^*_r (\varrho \| \sigma) \) and \( \hat{H}^*_r (\varrho \| \sigma) \) are called anti-divergences because for trace-class operators they are monotone non-decreasing under quantum operations; this is immediate from the monotone non-increasing property of \( D^*_\alpha \) under such maps for \( \alpha > 1 \); see [6, Theorem 14], [27, Theorem 3.14], or Theorem 4.14.
The Hoeffding anti-divergences are defined as Legendre-Fenchel transforms (polar functions). For some of them this transformation can be reversed as follows; this will be used in Theorem 4.14.

**Lemma 3.53.** For any \( \varrho, \sigma \in B(\mathcal{H}) \geq 0 \),

\[
\tilde{\psi}^\ast(\varrho \parallel \sigma | u) = \sup_{r \in \mathbb{R}} \left\{ ur - H_r^\ast(\varrho \parallel \sigma) \right\}, \quad u \in \mathbb{R} \setminus \{0, 1\},
\]

\[
\tilde{\psi}^\ast(\varrho \parallel \sigma | u) = \sup_{r \in \mathbb{R}} \left\{ ur - \hat{H}_r^\ast(\varrho \parallel \sigma) \right\}, \quad u \in \mathbb{R}.
\]

**Proof.** By Corollary 3.48, \( \tilde{\psi}^\ast(\varrho \parallel \sigma | \cdot) \) and \( \tilde{\psi}^\ast(\varrho \parallel \sigma | \cdot) \) are convex and lower semi-continuous on \( \mathbb{R} \), and hence (3.83)–(3.84) follow from (3.80)–(3.81) according to the bipolar theorem (see, e.g., [11, Proposition 4.1]). Likewise, \( r \mapsto H_r^\ast(\varrho \parallel \sigma) \) is the polar function of \( f(u) := \tilde{\psi}^\ast(\varrho \parallel \sigma | u) + (\pm \infty) \mathbf{1}_{\{0, 1\}}(u), u \in \mathbb{R} \), and hence, by [11, Proposition 4.1], its polar function is the largest convex and lower semi-continuous minorant of \( f \), which is exactly \( \hat{\psi}^\ast(\varrho \parallel \sigma | \cdot) \). This proves (3.82). \( \Box \)

The different variants of the Hoeffding anti-divergence defined above appear naturally in different bounds on the strong converse exponents; see Sect. 4. Our next goal is to explore their relations; in particular, to find sufficient conditions for some or all of them to coincide. Note that this is not always the case, as shown in Examples 3.59–3.60.

**Lemma 3.54.** Let \( \varrho, \sigma \in B(\mathcal{H}) \geq 0 \). For any \( u \in [0, 1] \) and any \( r \in \mathbb{R} \),

\[
\tilde{\psi}^\ast(\varrho \parallel \sigma | u) \leq \tilde{\psi}^\ast(\varrho \parallel \sigma | u) \leq \tilde{\psi}^\ast(\varrho \parallel \sigma | u),
\]

\[
H_r^\ast(\varrho \parallel \sigma) \leq H_r^\ast(\varrho \parallel \sigma) \leq H_r^\ast(\varrho \parallel \sigma),
\]

\[
\hat{H}_r^\ast(\varrho \parallel \sigma) \leq \hat{H}_r^\ast(\varrho \parallel \sigma) \leq \hat{H}_r^\ast(\varrho \parallel \sigma),
\]

and

\[
Q_{\alpha}(\varrho \parallel \sigma) = Q_{\alpha}(\varrho \parallel \sigma), \quad \alpha > 1 \iff \tilde{\psi}^\ast(\varrho \parallel \sigma | u) = \tilde{\psi}^\ast(\varrho \parallel \sigma | u), \quad u \in [0, 1]
\]

\[
\implies \begin{cases} H_r^\ast(\varrho \parallel \sigma) = H_r^\ast(\varrho \parallel \sigma), & r \in \mathbb{R}, \\ \hat{H}_r^\ast(\varrho \parallel \sigma) = \hat{H}_r^\ast(\varrho \parallel \sigma), & r \in \mathbb{R}. \end{cases}
\]

In particular, if \( \varrho \) and \( \sigma \) are trace-class, or \( \sigma \) is compact and \( \varrho \in B^\infty(\mathcal{H}, \sigma) \), then all equalities in (3.85)–(3.86) hold.

**Proof.** The inequalities are immediate from (3.50) and the definitions of the given quantities. The equivalence in (3.85) is trivial by definition, as is the implication in (3.86). The last assertion follows from Proposition 3.40. \( \Box \)
Proposition 3.56. Let \( \varrho, \sigma \in B(\cal H)_{\geq 0} \) be such that \( D_{\alpha_0}^*(\varrho \| \sigma) < +\infty \) (equivalently, \( \varrho \in L_{\alpha_0}(\cal H, \sigma) \)) for some \( \alpha_0 \in (1, +\infty) \). Then

\[
H_r^*(\varrho \| \sigma)_{fa} = \tilde{H}_r^*(\varrho \| \sigma)_{fa}, \quad r \in \mathbb{R}.
\] (3.87)

Proof. It is enough to prove that

\[
\sup_{u \in [0, 1]} \left\{ ur - \tilde{\psi}^*(\varrho \| \sigma | u)_{fa} \right\} = \left\{ ur - \tilde{\psi}^*(\varrho \| \sigma | u)_{fa} \right\} \sup_{u \in [0, 1]} \left\{ ur - \tilde{\psi}^*(\varrho \| \sigma | u)_{fa} \right\}.
\] (3.88)

We prove the first equality, as the second one follows the same way. If \( \tilde{\psi}^*(\varrho \| \sigma | 0)_{fa} = +\infty \) then there is nothing to prove, and hence we assume the contrary. Also by assumption,

\[
+\infty > (\alpha_0 - 1) D_{\alpha_0}^*(\varrho \| \sigma) = \psi^*(\varrho \| \sigma | \alpha_0) \geq \psi^*(\varrho \| \sigma | \alpha_0)_{fa} = \frac{\tilde{\psi}^*(\varrho \| \sigma | u_0)_{fa}}{1 - u_0},
\]

where \( u_0 := (\alpha_0 - 1)/\alpha_0 \). By Corollary 3.48, \( \tilde{\psi}^*(\varrho \| \sigma | \cdot)_{fa} \) is convex on \([0, 1]\), and finiteness at 0 and \( u_0 \) implies \( \tilde{\psi}^*(\varrho \| \sigma | u)_{fa} < +\infty, u \in [0, u_0] \). By Lemma 3.20, we also have \( \tilde{\psi}^*(\varrho \| \sigma | u)_{fa} = -\infty, u \in [0, u_0] \). Hence, \( u \mapsto ur - \tilde{\psi}^*(\varrho \| \sigma | u)_{fa} \) is a finite-valued concave and upper semi-continuous function on \([0, u_0]\), whence it is also continuous on \([0, u_0] \). This proves the asserted equality. \( \square \)

Proposition 3.56. Let \( \varrho, \sigma \in B(\cal H)_{\geq 0} \) be such that \( D_{\alpha_0}^*(\varrho \| \sigma) < +\infty \) for some \( \alpha_0 \in (1, +\infty) \), and \( Q_{\alpha}^*(\varrho \| \sigma)_{fa} = Q_{\alpha}^*(\varrho \| \sigma), \alpha > 1 \). Then, for every \( r \in \mathbb{R} \),

\[
H_r^*(\varrho \| \sigma) = H_r^*(\varrho \| \sigma)_{fa} = \tilde{H}_r^*(\varrho \| \sigma)_{fa},
\]

\[
\tilde{H}_r^*(\varrho \| \sigma) = \tilde{H}_r^*(\varrho \| \sigma)_{fa}.
\] (3.89)

Proof. Immediate from Lemmas 3.54 and 3.55. \( \square \)

Remark 3.57. Some further properties of, and relations among, the different Hoeffding anti-divergences are given in Appendix A. While these are not used in the rest of the paper, they might give some extra insight into the different bounds given in Proposition 4.5.

We close this section with some statements on the possible values of the Hoeffding anti-divergences. For these, we will need the notion of the Umegaki relative entropy [45]. For two finite-rank PSD operators \( \varrho, \sigma \in B(\cal H)_{\geq 0} \), it is defined as

\[
D(\varrho \| \sigma) := \left\{ \begin{array}{ll}
\text{Tr} \varrho (\hat{\log} \varrho - \hat{\log} \sigma), & \varrho^0 \leq \sigma^0, \\
+\infty, & \text{otherwise},
\end{array} \right.
\]

where \( \hat{\log} x := \log x, x > 0, \) and \( \hat{\log} 0 := 0 \). For positive normal functionals on a von Neumann algebra, it may be defined using the relative modular operator [1]. In the simple case of PSD trace-class operators \( \varrho, \sigma \) on a separable Hilbert space \( \cal H \), their relative entropy may be expressed equivalently as [19, Theorem 4.5]

\[
D(\varrho \| \sigma) = \lim_{\varrho \in B(\cal H)_{\geq 0} \ni P \searrow I} D(P \varrho P \| P \sigma P) = \lim_{n \to +\infty} D(P_n \varrho P_n \| P_n \sigma P_n).
\]
where the second equality holds for any increasing sequence $P_n \in \mathbb{P}_f(\mathcal{H})^+_{\varrho,\sigma}$, $n \in \mathbb{N}$, converging strongly to $I$. For non-zero PSD trace-class operators $\varrho, \sigma$ and $\lambda, \eta \in (0, +\infty)$, the scaling laws

$$D^*_\alpha(\varrho \| \sigma) = D^*_\alpha(\varrho \| \sigma) + \frac{\alpha}{\alpha - 1} \log \lambda - \log \eta,$$

$$H^*_r(\varrho \| \sigma) = H^*_r(\varrho \| \sigma) - \log \lambda,$$

$$D(\varrho \| \sigma) = \lambda D(\varrho \| \sigma) + (\text{Tr} \varrho) \lambda \log \lambda,$$

$$D_{\text{max}}(\varrho \| \sigma) = D_{\text{max}}(\varrho \| \sigma) + \log \lambda - \log \eta,$$

are easy to verify from the definitions (see also Remark 3.13). It was shown in [6,27] that

$$\exists \alpha_0 > 1 : D^*_\alpha(\varrho \| \sigma) < +\infty \implies \lim_{\alpha \searrow 1} \tilde{D}^*_\alpha(\varrho \| \sigma) = \inf_{\alpha > 1} \tilde{D}^*_\alpha(\varrho \| \sigma) = \frac{1}{\text{Tr} \varrho} D(\varrho \| \sigma).$$

**Lemma 3.58.** Let $\varrho, \sigma \in \mathcal{L}^1(\mathcal{H})_{\geq 0}$ be PSD trace-class operators.

(i) For every $r \in \mathbb{R}$,

$$H^*_r(\varrho \| \sigma) \geq r - D_{\text{max}}(\varrho \| \sigma).$$

(ii) If there exists an $\alpha_0 \in (1, +\infty)$ such that $D^*_\alpha(\varrho \| \sigma) < +\infty$ then

$$H^*_r(\varrho \| \sigma) = H^*_r(\varrho \| \sigma)_{\tilde{\alpha}} = H^*_r(\varrho \| \sigma)_{\tilde{\alpha}} = \tilde{H}^*_r(\varrho \| \sigma) = \tilde{H}^*_r(\varrho \| \sigma)_{\tilde{\alpha}} = \tilde{H}^*_r(\varrho \| \sigma)_{\tilde{\alpha}},$$

$$\begin{cases}
- \log \text{Tr} \varrho, \\
\in \left(- \log \text{Tr} \varrho, r - \frac{1}{\text{Tr} \varrho} D(\varrho \| \sigma) - \log \text{Tr} \varrho, \right), \\
r > \frac{1}{\text{Tr} \varrho} D(\varrho \| \sigma) - \log \text{Tr} \varrho.
\end{cases}$$

(iii) If $D^*_\alpha(\varrho \| \sigma) = +\infty$ for every $\alpha \in (1, +\infty)$ then

$$H^*_r(\varrho \| \sigma) = H^*_r(\varrho \| \sigma)_{\tilde{\alpha}} = H^*_r(\varrho \| \sigma)_{\tilde{\alpha}} = -\infty$$

$$< - \log \text{Tr} \varrho = \tilde{H}^*_r(\varrho \| \sigma) = \tilde{H}^*_r(\varrho \| \sigma)_{\tilde{\alpha}} = \tilde{H}^*_r(\varrho \| \sigma)_{\tilde{\alpha}}.$$

**Proof.** (i) By the scaling laws (3.90–3.92),

$$H^*_r(\varrho \| \sigma) = H^*_r(\text{log Tr} \varrho \left( \frac{\varrho}{\text{Tr} \varrho} \left( \frac{\sigma}{\text{Tr} \sigma} \right) \right) - \log \text{Tr} \varrho$$

$$= \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left[ r + \log \text{Tr} \sigma - D^*_\alpha \left( \frac{\varrho}{\text{Tr} \varrho} \left( \frac{\sigma}{\text{Tr} \sigma} \right) \right) \right] - \log \text{Tr} \varrho.$$ (3.100)

According to Corollary 3.46, $\lim_{\alpha \to +\infty} D^*_\alpha \left( \frac{\varrho}{\text{Tr} \varrho} \left( \frac{\sigma}{\text{Tr} \sigma} \right) \right) = D_{\text{max}} \left( \frac{\varrho}{\text{Tr} \varrho} \left( \frac{\sigma}{\text{Tr} \sigma} \right) \right) = D_{\text{max}}(\varrho \| \sigma) - \log \text{Tr} \varrho + \log \text{Tr} \sigma$, and hence,

$$H^*_r(\varrho \| \sigma) \geq \lim_{\alpha \to +\infty} \frac{\alpha - 1}{\alpha} \left[ r + \log \text{Tr} \sigma - D^*_\alpha \left( \frac{\varrho}{\text{Tr} \varrho} \left( \frac{\sigma}{\text{Tr} \sigma} \right) \right) \right] - \log \text{Tr} \varrho$$

$$= r - D_{\text{max}}(\varrho \| \sigma),$$

proving (3.95).
(ii) The equalities in (3.96) follow from Proposition 3.56. Using the assumption $D_{\alpha_0}^*(\varrho\|\sigma) < +\infty$, (3.94) and (3.92) give

$$\inf_{\alpha > 1} D_{\alpha}^*\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right) = \lim_{\alpha \downarrow 1} D_{\alpha}^*\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right) = D\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right) = \frac{1}{\text{Tr}\varrho} D(\varrho\|\sigma) - \log \text{Tr}\varrho + \log \text{Tr}\sigma.$$  

(3.101)

In particular, the above limit is finite, and thus

$$-\log \text{Tr}\varrho = \lim_{\alpha \downarrow 1} \frac{\alpha - 1}{\alpha} \left[r + \log \text{Tr}\sigma - D_{\alpha}^*\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right)\right] - \log \text{Tr}\varrho \leq H_r^*(\varrho\|\sigma),$$

where in the second expression we used (3.100), and the inequality is by definition. On the other hand, (3.100) shows that $H_r^*(\varrho\|\sigma) > -\log \text{Tr}\varrho$ holds if and only if

$$r + \log \text{Tr}\sigma > \inf_{\alpha > 1} D_{\alpha}^*\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right) = D\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right) = \frac{1}{\text{Tr}\varrho} D(\varrho\|\sigma) - \log \text{Tr}\varrho + \log \text{Tr}\sigma,$$

(3.102)

where the equalities are due to (3.101). Note that (3.102) is exactly the condition in the second line of (3.97), and hence we obtain the first line in (3.97). Assume now that $r$ is as in (3.102). Then

$$\frac{\alpha - 1}{\alpha} \left[r + \log \text{Tr}\sigma - D_{\alpha}^*\left(\frac{\varrho}{\text{Tr}\varrho} \bigg\| \frac{\sigma}{\text{Tr}\sigma}\right)\right] - \log \text{Tr}\varrho < r - \frac{1}{\text{Tr}\varrho} D(\varrho\|\sigma),$$

proving the second line of (3.97).

(iii) By Lemma 3.54, $\tilde{\psi}^*(\varrho\|\sigma|u) = \tilde{\psi}^*(\varrho\|\sigma|u)_{\text{fa}} = \tilde{\psi}^*(\varrho\|\sigma|u)_{\text{fa}} = +\infty$ for every $u \in (0, 1)$, whence $H_r^*(\varrho\|\sigma) = H_r^*(\varrho\|\sigma)_{\text{fa}} = H_r^*(\varrho\|\sigma)_{\text{fa}} = -\infty$. On the other hand, $\tilde{\psi}^*(\varrho\|\sigma|0) = \tilde{\psi}^*(\varrho\|\sigma|0)_{\text{fa}} = \tilde{\psi}^*(\varrho\|\sigma|0)_{\text{fa}} = \log \text{Tr}\varrho$, according to Remark 3.49, and $\tilde{\psi}^*(\varrho\|\sigma|1) = \tilde{\psi}^*(\varrho\|\sigma|1)_{\text{fa}} = \tilde{\psi}^*(\varrho\|\sigma|1)_{\text{fa}} = D_{\text{max}}(\varrho\|\sigma) = +\infty$, where the last equality follows from Corollary 3.46. Hence, $H_r^*(\varrho\|\sigma) = H_r^*(\varrho\|\sigma)_{\text{fa}} = H_r^*(\varrho\|\sigma)_{\text{fa}} = -\log \text{Tr}\varrho$. \hfill \qed

Example 3.59. Let $(e_n)_{n \in \mathbb{N}}$ be an orthonormal basis in $\mathcal{H}$, and $\varrho := c_1 \sum_{n=1}^{+\infty} n^{-\beta} |e_n\rangle \langle e_n|$, $\sigma := c_2 \sum_{n=1}^{+\infty} n^{-\gamma} |e_n\rangle \langle e_n|$, with some $\beta > 1$ and $\gamma > 0$, where $c_1$ and $c_2$ are choosen so that $\varrho$ and $\sigma$ are density operators. Obviously, $\varrho$ and $\sigma$ are commuting (classical). For $P_N := \sum_{n=1}^{N} |e_n\rangle \langle e_n|$, we have

$$Q_{\alpha}^*(P_N \varrho P_N \| P_N \sigma P_N) = c_1 c_2^{-\alpha} \sum_{n=1}^{N} n^{-\alpha \beta + (1-\alpha)\gamma} \xrightarrow{N \to +\infty} +\infty, \quad \alpha \in (1, +\infty),$$

whence

$$D_{\alpha}^*(\varrho\|\sigma) = +\infty, \quad \alpha \in (1, +\infty).$$
and

\[ H_r^*(\varrho \parallel \sigma) = H_r^*(\varrho \parallel \sigma)_{\text{fa}} = H_r^*(\varrho \parallel \sigma)_{\text{fa}} = -\infty \]

\[ < -\log \text{Tr} \varrho = \hat{H}_r^*(\varrho \parallel \sigma) = \hat{H}_r^*(\varrho \parallel \sigma)_{\text{fa}} = \hat{H}_r^*(\varrho \parallel \sigma)_{\text{fa}}, \]

according to Lemma 3.58. Note also that

\[ \tilde{\psi}^*(\varrho \parallel \sigma | u) = \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}} = \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}} = +\infty, \quad u \in (0, 1), \]

\[ \tilde{\psi}^*(\varrho \parallel \sigma | 0)_{\text{fa}} = \tilde{\psi}^*(\varrho \parallel \sigma | 0)_{\text{fa}} = \log \text{Tr} \varrho = 0 < \lim_{u \downarrow 0} \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}}, \]

\[ \tilde{\psi}^*(\varrho \parallel \sigma | 1)_{\text{fa}} = \tilde{\psi}^*(\varrho \parallel \sigma | 1)_{\text{fa}} = D_{\max}(\varrho \parallel \sigma) = +\infty = \lim_{u \uparrow 1} \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}}. \]

For the relative entropy we get

\[ D(\varrho \parallel \sigma) = c_1 \sum_{n=1}^{+\infty} \frac{1}{n^\beta} \log \frac{c_1 n^\gamma}{c_2 n^\beta} = \log \frac{c_1}{c_2} + c_1 \sum_{n=1}^{+\infty} \frac{(n^\gamma - \beta) \log n}{n^\beta} < +\infty, \]

if \( \beta > \gamma + 1 \). Hence, assuming that \( D(\varrho \parallel \sigma) < +\infty \) is not sufficient for Lemma 3.55 and Lemma 3.58.

This also gives an example where

\[ D(\varrho \parallel \sigma) < +\infty = \lim_{\alpha \searrow 1} D_{\alpha}^*(\varrho \parallel \sigma), \]

which is contrary to the case where \( D_{\alpha_0}^*(\varrho \parallel \sigma) < +\infty \) for some \( \alpha_0 > 1 \); see (3.94). This kind of behaviour was already pointed out in [19, Remark 5.4].

**Example 3.60.** Let \( \varrho = \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \) be such that \( \varrho \) is not trace-class. Then \( \varrho = \varrho^{\frac{a-1}{2}} = \varrho^{\frac{a-1}{2}} \), whence \( \varrho \in \mathcal{B}^a(\mathcal{H}, \sigma) \) with \( \varrho_{\sigma, \alpha} = \varrho^{\frac{1}{\alpha}} \notin \mathcal{L}^a(\mathcal{H}) \), and

\[ \tilde{\psi}^*(\varrho \parallel \sigma | u) = \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}} = \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}} = +\infty, \quad u \in (0, 1), \]

\[ \tilde{\psi}^*(\varrho \parallel \sigma | 0)_{\text{fa}} = \tilde{\psi}^*(\varrho \parallel \sigma | 0)_{\text{fa}} = \log \text{Tr} \varrho = +\infty = \lim_{u \downarrow 0} \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}}, \]

\[ \tilde{\psi}^*(\varrho \parallel \sigma | 1)_{\text{fa}} = \tilde{\psi}^*(\varrho \parallel \sigma | 1)_{\text{fa}} = D_{\max}(\varrho \parallel \sigma) = 0 < \lim_{u \uparrow 1} \tilde{\psi}^*(\varrho \parallel \sigma | u)_{\text{fa}}. \]

Thus, for every \( r \in \mathbb{R} \),

\[ H_r^*(\varrho \parallel \sigma) = H_r^*(\varrho \parallel \sigma)_{\text{fa}} = H_r^*(\varrho \parallel \sigma)_{\text{fa}} = -\infty \]

\[ < r = \hat{H}_r^*(\varrho \parallel \sigma) = \hat{H}_r^*(\varrho \parallel \sigma)_{\text{fa}} = \hat{H}_r^*(\varrho \parallel \sigma)_{\text{fa}}. \]

In particular, this holds also when \( \varrho \) is compact, and obviously \( \varrho \in \mathcal{B}^{\infty}(\mathcal{H}, \varrho) \). This shows that the assumption \( D_{\alpha_0}^*(\varrho \parallel \sigma) < +\infty \) for some \( \alpha_0 < +\infty \) is also important in this case of Proposition 3.56.

Note also that this is an example where

\[ \exists \lim_{\alpha \to +\infty} D_{\alpha}^*(\varrho \parallel \sigma) (= +\infty) \neq D_{\max}(\varrho \parallel \sigma). \]

This cannot happen when \( \varrho \) and \( \sigma \) are both trace-class, according to Corollary 3.43 or Corollary 3.46.
4. The Strong Converse Exponent

4.1. The strong converse exponents and the Hoeffding anti-divergences. Before restricting our attention to the i.i.d. case in the main result, we first consider a generalization of the binary state discrimination problem described in the Introduction. First, we do not assume the problem to be i.i.d. In the most general case, a simple asymptotic binary operator discrimination problem is specified by a sequence of Hilbert spaces $\mathcal{H}_n$, $n \in \mathbb{N}$, and for each $n \in \mathbb{N}$, a pair $\varrho_n, \sigma_n \in \mathcal{B}(\mathcal{H}_n)_{\geq 0}$, representing the null and the alternative hypotheses, respectively. Since the operators are not assumed to be trace-class, the expressions in (1.1) may not make sense, and need to be modified as

$$\gamma_n(T_n | \varrho_n) := \text{Tr} \left( T_n^{1/2} \varrho_n T_n^{1/2} \right) = \text{Tr} \left( \varrho_n^{1/2} T_n \varrho_n^{1/2} \right),$$

$$\beta_n(T_n | \sigma_n) := \text{Tr} \left( T_n^{1/2} \sigma_n T_n^{1/2} \right) = \text{Tr} \left( \sigma_n^{1/2} T_n \sigma_n^{1/2} \right),$$

to define the generalized type I success and type II errors, respectively. These expressions are equal to those in (1.1) when $\varrho_n$ and $\sigma_n$ are trace-class.

**Definition 4.1.** Let $\tilde{\varrho} := (\varrho_n)_{n \in \mathbb{N}}, \tilde{\sigma} := (\sigma_n)_{n \in \mathbb{N}}$ be as above. The strong converse exponents of the simple asymptotic binary operator discrimination problem $H_0 : \tilde{\varrho}$ vs. $H_1 : \tilde{\sigma}$ with type II exponent $r \in \mathbb{R}$ are defined as

$$\text{sc}_r(\tilde{\varrho} \parallel \tilde{\sigma}) := \inf \left\{ \liminf_{n \to +\infty} -\frac{1}{n} \log \gamma_n(T_n | \varrho_n) : \liminf_{n \to +\infty} -\frac{1}{n} \log \beta_n(T_n | \sigma_n) \geq r \right\},$$

$$\text{sc}_r(\tilde{\varrho} \parallel \tilde{\sigma}) := \inf \left\{ \limsup_{n \to +\infty} -\frac{1}{n} \log \gamma_n(T_n | \varrho_n) : \limsup_{n \to +\infty} -\frac{1}{n} \log \beta_n(T_n | \sigma_n) \geq r \right\},$$

$$\text{sc}_r(\tilde{\varrho} \parallel \tilde{\sigma}) := \inf \left\{ \lim_{n \to +\infty} -\frac{1}{n} \log \gamma_n(T_n | \varrho_n) : \lim_{n \to +\infty} -\frac{1}{n} \log \beta_n(T_n | \sigma_n) \geq r \right\},$$

where the infima are taken along all test sequences $T_n \in \mathcal{B}(\mathcal{H}_n)_{[0,1]}$, $n \in \mathbb{N}$, satisfying the indicated condition, and in the last expression also that the limit exists.

We will need an extension of the notion of the Hoeffding anti-divergence in the above setting. Let

$$\psi^*(\tilde{\varrho} \parallel \tilde{\sigma} | \alpha) := \limsup_{n \to +\infty} \frac{1}{n} \psi^*(\varrho_n \parallel \sigma_n | \alpha), \quad \alpha \in (1, +\infty),$$

$$\tilde{\psi}^*(\tilde{\varrho} \parallel \tilde{\sigma} | u) := \limsup_{n \to +\infty} \frac{1}{n} \tilde{\psi}^*(\varrho_n \parallel \sigma_n | u)$$

$$= \begin{cases} (1 - u) \psi^*(\tilde{\varrho} \parallel \tilde{\sigma} | (1 - u)^{-1}), & u \in (0, 1), \\ \limsup_{n \to +\infty} \frac{1}{n} \log \text{Tr} \varrho_n, & u = 0, \\ \limsup_{n \to +\infty} \frac{1}{n} D_{\max}(\varrho_n \parallel \sigma_n), & u = 1, \end{cases}$$

where we used (3.74)–(3.75), and

$$\hat{H}_r^*(\tilde{\varrho} \parallel \tilde{\sigma}) := \sup_{u \in [0,1]} \{ u r - \tilde{\psi}^*(\tilde{\varrho} \parallel \tilde{\sigma} | u) \}.$$
The inequality in the following lemma is called the optimality part of the Hoeffding bound. For trace-class operators, it can be easily obtained from the monotonicity of the sandwiched Rényi divergence under measurements; see [6,33,36]. If we do not assume \( \sigma_n \) to be trace-class, we can still obtain it using the variational formula in (3.27), as we show below.

**Proposition 4.2.** For every \( r \in \mathbb{R} \),
\[
\hat{H}_r^* (\tilde{\varrho} \| \tilde{\sigma}) \leq \text{sc}_{\tau}(\tilde{\varrho} \| \tilde{\sigma}) \leq \text{sc}_{\tau}(\tilde{\varrho} \| \tilde{\sigma}) \leq \text{sc}_{\tau}(\tilde{\varrho} \| \tilde{\sigma}).
\]  
(4.1)

**Proof.** All the inequalities are trivial by definition, except for the first one. Thus, we need to show that for any \( r \in \mathbb{R} \) and any \( u \in [0, 1] \),
\[
\text{sc}_{\tau}(\tilde{\varrho} \| \tilde{\sigma}) \geq ur - \tilde{\psi}^*(\tilde{\varrho} \| \tilde{\sigma} | u).
\]  
(4.2)

Let us fix \( r \in \mathbb{R} \) for the rest. First, note that for any test \( T_n \),
\[
\gamma_n(T_n|\varrho_n) = \text{Tr}(\varrho_n^{1/2}T_n\varrho_n^{1/2}) \leq \text{Tr} \varrho_n.
\]
Thus, for any sequence of tests \((T_n)_{n \in \mathbb{N}}\),
\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \gamma_n(T_n|\varrho_n) \geq -\limsup_{n \to +\infty} \frac{1}{n} \log \text{Tr} \varrho_n = 0 \cdot r - \tilde{\psi}^*(\tilde{\varrho} \| \tilde{\sigma} | 0),
\]
proving (4.2) for \( u = 0 \). For \( u = 1 \), (4.2) is trivial when \( \psi^*(\tilde{\varrho} \| \tilde{\sigma} | 1) = +\infty \), and hence we assume the contrary; in particular, \( D_{\max}(\varrho_n\|\sigma_n) < +\infty \) for any large enough \( n \). Let \((T_n)_{n \in \mathbb{N}}\) be a test sequence such that
\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \beta_n(T_n|\sigma_n) \geq r.
\]
Then for any \( r' < r \) and any large enough \( n \), \( \beta_n(T_n|\sigma_n) \leq \exp(-nr') \), whence
\[
\gamma_n(T_n|\varrho_n) = \text{Tr}(T_n^{1/2}\varrho_nT_n^{1/2}) \leq \exp(D_{\max}(\varrho_n\|\sigma_n)) \text{Tr}(T_n^{1/2}\sigma_nT_n^{1/2})
\]
\[
\leq \exp(D_{\max}(\varrho_n\|\sigma_n) - nr').
\]
Thus,
\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \gamma_n(T_n|\varrho_n) \geq r' - \limsup_{n \to +\infty} \frac{1}{n} D_{\max}(\varrho_n\|\sigma_n) = r' - \tilde{\psi}^*(\tilde{\varrho} \| \tilde{\sigma} | 1).
\]
This gives (4.2) for \( u = 1 \).

For the rest, let us fix an \( u \in (0, 1) \), and corresponding \( \alpha = 1/(1 - u) > 1 \). If \( \tilde{\psi}^*(\tilde{\varrho} \| \tilde{\sigma} | u) = +\infty \) then
\[
ur - \tilde{\psi}^*(\tilde{\varrho} \| \tilde{\sigma} | u) = -\infty \leq \text{sc}_{\tau}(\tilde{\varrho} \| \tilde{\sigma})
\]  
(4.3)
holds trivially. Hence, we assume that \( \tilde{\psi}^*(\varrho_n\|\sigma_n|u) < +\infty \), or equivalently, \( \varrho_n \in \mathcal{L}^\alpha(\mathcal{H}_n, \sigma_n) \) for every large enough \( n \). In particular, the variational formula (3.27) holds (with \( z = \alpha \)).
Consider now a sequence of tests \( (T_n)_{n \in \mathbb{N}} \) such that \( \liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr}(T_{n}^{1/2} \sigma_n T_{n}^{1/2}) \geq r \). Then \( \text{Tr}(T_{n}^{1/2} \sigma_n T_{n}^{1/2}) < +\infty \) for every large enough \( n \), and we have

\[
\text{Tr} \left( T_{n}^{1/2} \sigma_n^{a-1} T_{n}^{1/2} \right)^{\frac{a}{a-1}} \leq \text{Tr} \left( T_{n}^{1/2} \sigma_n T_{n}^{1/2} \right) < +\infty, \tag{4.4}
\]

where the first inequality is due to the operator Jensen inequality \([7, \text{Theorem 11}]\). Hence, \( T_n \in \mathcal{B}(\mathcal{H}_n)_{\sigma_n, \alpha, \alpha} \). If \( \text{Tr} \left( T_{n}^{1/2} \sigma_n^{a-1} T_{n}^{1/2} \right)^{\frac{a}{a-1}} > 0 \) then the variational formula (3.27) yields

\[
\psi^*(\varrho_n \| \sigma_n | \alpha) \geq \alpha \log \text{Tr} \left( T_{n}^{1/2} \varrho_n T_{n}^{1/2} \right) + (1 - \alpha) \log \text{Tr} \left( T_{n}^{1/2} \sigma_n^{-1} T_{n}^{1/2} \right)^{\frac{a}{a-1}}
\]

\[
\geq \alpha \log \text{Tr} \left( T_{n}^{1/2} \varrho_n T_{n}^{1/2} \right) + (1 - \alpha) \log \text{Tr} \left( T_{n}^{1/2} \sigma_n T_{n}^{1/2} \right),
\]

where the second inequality is due to (4.4). In particular, we also have \( \text{Tr} \left( T_{n}^{1/2} \varrho_n T_{n}^{1/2} \right) < +\infty \). By a simple rearrangement, we get

\[
-\frac{1}{n} \log \text{Tr} \left( T_{n}^{1/2} \varrho_n T_{n}^{1/2} \right) \geq \frac{\alpha - 1}{\alpha} \left( -\frac{1}{n} \right) \log \text{Tr} \left( T_{n}^{1/2} \sigma_n T_{n}^{1/2} \right) - \frac{1}{\alpha n} \psi^*(\varrho_n \| \sigma_n | \alpha). \tag{4.5}
\]

If \( \text{Tr} \left( T_{n}^{1/2} \sigma_n^{a-1} T_{n}^{1/2} \right)^{\frac{a}{a-1}} = 0 \) then \( T_{n}^{1/2} \sigma_n^{a-1} T_{n}^{1/2} = 0 \). Since \( \varrho_n \in \mathcal{L}^\alpha(\mathcal{H}_n, \sigma_n) \), this implies \( T_{n}^{1/2} \varrho_n T_{n}^{1/2} = 0 \), according to Lemma 3.1, and therefore (4.5) holds trivially, with both sides equal to \(+\infty\).

Taking the liminf in (4.5) yields

\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \left( T_{n}^{1/2} \varrho_n T_{n}^{1/2} \right) \geq \frac{\alpha - 1}{\alpha} r - \frac{1}{\alpha} \psi^*(\varrho \| \sigma | \alpha) = ur - \tilde{\psi}^*(\varrho \| \sigma | u).
\]

Since this holds for every test sequence as above, we get \( ur - \tilde{\psi}^*(\varrho \| \sigma | u) \leq sc(\varrho \| \sigma) \), as required.

For the rest, we restrict our attention to the i.i.d. case, where

\[
\mathcal{H}_n = \mathcal{H}^\otimes n, \quad \varrho_n = \varrho^\otimes n, \quad \sigma_n = \sigma^\otimes n, \quad n \in \mathbb{N},
\]

for some Hilbert space \( \mathcal{H} \) and \( \varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \). Note that by Lemma 3.22, \( \tilde{\psi}^*\left( (\varrho^\otimes n)_{n \in \mathbb{N}} \| (\sigma^\otimes n)_{n \in \mathbb{N}} | u \right) = \tilde{\psi}^*(\varrho | \sigma | u), \quad u \in (0, 1), \quad n \in \mathbb{N}, \) and the same identity is straightforward to verify for \( u = 0, 1 \), whence

\[
H^*_r\left( (\varrho^\otimes n)_{n \in \mathbb{N}} \| (\sigma^\otimes n)_{n \in \mathbb{N}} \right) = H^*_r(\varrho \| \sigma), \quad r \in \mathbb{R}.
\]

We replace the notations \( \tilde{\varrho} \) and \( \tilde{\sigma} \) with \( \varrho \) and \( \sigma \), respectively, in the strong converse exponents introduced above. Let

\[
sc_r(\varrho \| \sigma)_f, \quad sc_f(\varrho \| \sigma)_f, \quad \text{and} \quad sc_f(\varrho \| \sigma)_f
\]

be defined the same way as \( sc(\varrho \| \sigma), \quad sc(\varrho \| \sigma), \) and \( sc(\varrho \| \sigma) \), respectively, but with the restrictions that only finite-rank tests are used. Obviously,

\[
sc_r(\varrho \| \sigma) \leq sc_r(\varrho \| \sigma)_f, \quad sc_f(\varrho \| \sigma) \leq sc_f(\varrho \| \sigma)_f, \quad sc_r(\varrho \| \sigma) \leq sc_f(\varrho \| \sigma)_f.
\]

The following lower bound follows by a straightforward adaptation of Nagaoka’s method \([36]\).
Proposition 4.3. Let $\rho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$. For every $r \in \mathbb{R}$,

$$
\hat{H}_r^*(\rho \| \sigma)_{\text{fa}} \leq \text{sc}_r(\rho \| \sigma)_r.
$$

Proof. Let us fix an $r \in \mathbb{R}$. We need to prove that for every $u \in [0, 1]$,

$$
\text{sc}_r(\rho \| \sigma)_r \geq ur - \tilde{\psi}^*(\rho \| \sigma | u)_{\text{fa}}.
$$

The cases $u = 0$ and $u = 1$ can be proved exactly the same way as in the proof of Proposition 4.2 above. For the rest, let us fix an $u \in (0, 1)$, with corresponding $\alpha = 1/(1 - u) > 1$. If $\tilde{\psi}^*(\rho \| \sigma | u)_{\text{fa}} = +\infty$ then (4.6) holds trivially, and hence for the rest we assume that $\tilde{\psi}^*(\rho \| \sigma | u)_{\text{fa}} < +\infty$. In particular, we have $\rho^0 \leq \sigma^0$, according to Lemma 3.42.

Let $T_n \in \mathcal{B}(\mathcal{H}^n)_{[0,1]}$, $n \in \mathbb{N}$, be a sequence of finite-rank tests such that

$$
\liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} T_n^{1/2} \sigma^0 T_n^{1/2} \geq r.
$$

Assume first that $T_n^{1/2} \rho^0 T_n^{1/2} \neq 0$, whence, by the assumption that $\rho^0 \leq \sigma^0$, we also have $T_n^{1/2} \sigma^0 T_n^{1/2} \neq 0$. By Lemma 3.37,

$$
Q^*_\alpha(\rho^0 \| \sigma^0)_{\text{fa}} \geq Q^*_\alpha(T_n^{1/2} \rho^0 T_n^{1/2} \| T_n^{1/2} \sigma^0 T_n^{1/2})
\geq (\text{Tr} T_n^{1/2} \rho^0 T_n^{1/2})^\alpha (\text{Tr} T_n^{1/2} \sigma^0 T_n^{1/2})^{1-\alpha},
$$

where the second inequality follows from Corollary 3.27. A simple rearrangement yields

$$
-\frac{1}{n} \log \text{Tr} T_n^{1/2} \rho^0 T_n^{1/2} \geq \frac{\alpha - 1}{\alpha} \left( -\frac{1}{n} \log \text{Tr} T_n^{1/2} \sigma^0 T_n^{1/2} \right) - \frac{1}{\alpha n} \tilde{\psi}^*(\rho^0 \| \sigma^0 | u)_{\text{fa}}
\geq u \left( -\frac{1}{n} \log \text{Tr} T_n^{1/2} \sigma^0 T_n^{1/2} \right) - \frac{1}{n} \tilde{\psi}^*(\rho^0 \| \sigma^0 | u)_{\text{fa}} \leq \tilde{\psi}^*(\rho \| \sigma | u)_{\text{fa}}.
$$

These inequalities also hold (trivially, with the leftmost expression being $+\infty$) when $T_n^{1/2} \rho^0 T_n^{1/2} = 0$. Thus, we get

$$
\liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} T_n^{1/2} \rho^0 T_n^{1/2} \geq ur - \tilde{\psi}^*(\rho \| \sigma | u)_{\text{fa}}.
$$

Since this holds for every test sequences as above, (4.6) follows. □

Lemma 4.4. For finite-rank PSD operators $\rho, \sigma$ on a Hilbert space, with $0 \neq \rho^0 \leq \sigma^0$, we have

$$
\text{sc}_r(\rho \| \sigma) \leq H_r^*(\rho \| \sigma), \quad r \in \mathbb{R}.
$$

Proof. The inequality in (4.7) was proved in [33, Theorem 4.10] for finite-rank density operators, under the implicit assumption that $D(\rho \| \sigma) \neq D_{\max}(\rho \| \sigma)$, and it was proved in [22] in the case $D(\rho \| \sigma) = D_{\max}(\rho \| \sigma)$. The case of general PSD operators follows easily by replacing $\rho$ and $\sigma$ with $\rho/\text{Tr} \rho$ and $\sigma/\text{Tr} \sigma$, respectively, and using the scaling laws (3.91) and $\text{sc}_r(\lambda \rho \| \eta \sigma) = \text{sc}_{r+\log \eta}(\rho \| \sigma) - \log \lambda$. □
Proposition 4.5. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0$ be such that $\varrho^0 \leq \sigma^0$. For every $r \in \mathbb{R}$,

$$
\hat{H}_r^*(\varrho \| \sigma) \leq \overline{\text{sc}}_r(\varrho \| \sigma) \leq \overline{\text{sc}}_r(\varrho \| \sigma) \leq \text{sc}_r(\varrho \| \sigma)
$$

$$
\hat{H}_r^*(\varrho \| \sigma)_{\text{fa}} \leq \overline{\text{sc}}_r(\varrho \| \sigma)_{\text{f}} \leq \overline{\text{sc}}_r(\varrho \| \sigma)_{\text{f}} \leq \text{sc}_r(\varrho \| \sigma)_{\text{f}} \leq \hat{H}_r^*(\varrho \| \sigma)_{\text{fa}}.
$$

Proof. By Propositions 4.2 and 4.3, we only need to prove $\text{sc}_r(\varrho \| \sigma)_{\text{f}} \leq \hat{H}_r^*(\varrho \| \sigma)_{\text{fa}}$. Let $P \in \mathbb{P}_f(\mathcal{H})_{0,\sigma}^+$. According to Lemma 4.4, there exists a sequence of tests $(S_{P,n})_{n \in \mathbb{N}}$ such that $S_{P,n} \leq P \otimes^n$, and

$$
\liminf_{n \to +\infty} -\frac{1}{n} \log \operatorname{Tr}(P \sigma P) \otimes^n S_{P,n} \geq r, \quad (4.8)
$$

$$
\lim_{n \to +\infty} -\frac{1}{n} \log \operatorname{Tr}(P \varrho P) \otimes^n S_{P,n} \leq H_r^*(P \varrho P \| P \sigma P)
$$

$$
= \max_{u \in [0,1]} \left\{ ur - \tilde{\psi}^*(P \varrho P \| P \sigma P | u) \right\}, \quad (4.9)
$$

where the equality is due to Propositions 3.56 and 3.40. Note that

$$
\operatorname{Tr}(P \sigma P) \otimes^n S_{P,n} = \operatorname{Tr} \sigma^n \left( P \otimes^n S_{P,n} P \otimes^n \right),
$$

$$
\operatorname{Tr}(P \varrho P) \otimes^n S_{P,n} = \operatorname{Tr} \varrho^n \left( P \otimes^n S_{P,n} P \otimes^n \right),
$$

and therefore (4.8)–(4.9) yield

$$
\text{sc}_r(\varrho \| \sigma)_{\text{f}} \leq \max_{u \in [0,1]} \left\{ ur - \tilde{\psi}^*(P \varrho P \| P \sigma P | u) \right\}.
$$

Thus,

$$
\text{sc}_r(\varrho \| \sigma)_{\text{f}} \leq \inf_{P \in \mathbb{P}_f(\mathcal{H})_{0,\sigma}^+} \max_{u \in [0,1]} \left\{ ur - \tilde{\psi}^*(P \varrho P \| P \sigma P | u) \right\}. \quad (4.10)
$$

By Lemma 3.47, $u \mapsto ur - \tilde{\psi}^*(P \varrho P \| P \sigma P | u)$ is continuous on the compact set $[0,1]$ for every $P \in \mathbb{P}_f(\mathcal{H})_{0,\sigma}^+$. On the other hand, $\mathbb{P}_f(\mathcal{H})_{0,\sigma}^+$ is an upward directed partially ordered set with respect to the PSD order, and for any $u \in [0,1]$, $P \mapsto ur - \tilde{\psi}^*(P \varrho P \| P \sigma P | u)$ is monotone decreasing on $\mathbb{P}_f(\mathcal{H})$, again by Lemma 3.47. Hence, by Lemma 2.5, we may exchange the inf and the max in (4.10). Thus, we get the upper bound

$$
\text{sc}_r(\varrho \| \sigma)_{\text{f}} \leq \max_{u \in [0,1]} \left\{ ur - \sup_{P \in \mathbb{P}_f(\mathcal{H})} \tilde{\psi}^*(P \varrho P \| P \sigma P | u) \right\}
$$

$$
= \max_{u \in [0,1]} \left\{ ur - \tilde{\psi}^*(\varrho \| \sigma | u)_{\text{fa}} \right\} = \hat{H}_r^*(\varrho \| \sigma)_{\text{fa}},
$$

as required. \(\square\)
Theorem 4.6. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$ be such that $D_\alpha^*(\varrho\|\sigma) < +\infty$ for some $\alpha \in (1, +\infty)$, and $Q_\alpha^*(\varrho\|\sigma)_{\text{fn}} = Q_\alpha^*(\varrho\|\sigma)$, $\alpha > 1$. Then

$$s_{\alpha}(\varrho\|\sigma) = \overline{s}_{\alpha}(\varrho\|\sigma) = s_{\alpha}(\varrho\|\sigma)_f = \frac{s_{\alpha}(\varrho\|\sigma)}{f} = s_{\alpha}(\varrho\|\sigma)_f = H_\alpha^*(\varrho\|\sigma)_{\text{fn}} / H_\alpha^*(\varrho\|\sigma)_{fa}$$

$$= H_{\alpha}^*(\varrho\|\sigma) = \hat{H}_\alpha^*(\varrho\|\sigma)_{\text{fn}} = \hat{H}_\alpha^*(\varrho\|\sigma)_{fa}, \quad r \in \mathbb{R}.$$ 

On the other hand, if $\varrho, \sigma$ are trace-class and $D_\alpha^*(\varrho\|\sigma) = +\infty$ for all $\alpha \in (1, +\infty)$, then

$$s_{\alpha}(\varrho\|\sigma) = \overline{s}_{\alpha}(\varrho\|\sigma) = s_{\alpha}(\varrho\|\sigma) = \frac{s_{\alpha}(\varrho\|\sigma)}{f} = s_{\alpha}(\varrho\|\sigma)_f = H_\alpha^*(\varrho\|\sigma)_{\text{fn}} / H_\alpha^*(\varrho\|\sigma)_{fa}$$

$$> -\infty = H_{\alpha}^*(\varrho\|\sigma) = \hat{H}_\alpha^*(\varrho\|\sigma)_{\text{fn}} = \hat{H}_\alpha^*(\varrho\|\sigma)_{fa}, \quad r \in \mathbb{R}.$$ 

Proof. Immediate from Propositions 4.5, 3.56, and Lemma 3.58. \qed

As a special case of Theorem 4.6, we get the exact characterization of the strong converse exponent of discriminating quantum states on a separable Hilbert space, as follows:

Corollary 4.7. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$ be density operators. For every $r \in \mathbb{R}$,

$$s_{\alpha}(\varrho\|\sigma) = \overline{s}_{\alpha}(\varrho\|\sigma) = s_{\alpha}(\varrho\|\sigma) = \frac{s_{\alpha}(\varrho\|\sigma)}{f} = s_{\alpha}(\varrho\|\sigma)_f = H_\alpha^*(\varrho\|\sigma)_{\text{fn}} / H_\alpha^*(\varrho\|\sigma)_{fa} \quad (4.11)$$

$$= \hat{H}_\alpha^*(\varrho\|\sigma) \geq 0, \quad (4.12)$$

and

$$\hat{H}_\alpha^*(\varrho\|\sigma) > 0 \iff \exists \alpha > 1 : D_\alpha^*(\varrho\|\sigma) < +\infty \text{ and } r > D(\varrho\|\sigma). \quad (4.13)$$

Proof. The equalities in (4.11) are immediate from Theorem 4.6, and the characterization of positivity in (4.13) follows from Lemma 3.58. \qed

Remark 4.8. Let $\varrho$ and $\sigma$ be density operators. According to the direct part of the quantum Stein’s lemma [24, 26], for every $r < D(\varrho\|\sigma)$ there exists a test sequence $T_n \in \mathcal{B}(\mathcal{H}^{\otimes n})_{[0,1]}$, $n \in \mathbb{N}$, such that

$$\lim_{n \to +\infty} \frac{1}{n} \log \text{Tr} \varrho^{\otimes n}(I - T_n) = 0, \quad \liminf_{n \to +\infty} \frac{1}{n} \log \text{Tr} \sigma^{\otimes n} T_n \geq r. \quad (4.14)$$

It was shown in [36, 38] that in the finite-dimensional case, for any test sequence $T_n \in \mathcal{B}(\mathcal{H}^{\otimes n})_{[0,1]}$, $n \in \mathbb{N}$,

$$r := \liminf_{n \to +\infty} \frac{1}{n} \log \text{Tr} \sigma^{\otimes n} T_n > D(\varrho\|\sigma) \implies \lim_{n \to +\infty} \text{Tr} \varrho^{\otimes n}(I - T_n) = 1. \quad (4.15)$$

That is, if the type II error decreases with an exponent larger than the relative entropy, then the type I error goes to 1; this is called the strong converse to Stein’s lemma. The optimal (lowest) speed of convergence to 1 is exponential, with the exponent being equal to the Hoeffding anti-divergence $H^*_\alpha(\varrho\|\sigma)$, according to [33]. Corollary 4.7 generalizes this to the infinite-dimensional case, with one important difference. While in the finite-dimensional case finiteness of the relative entropy implies strict positivity of $H^*_\alpha(\varrho\|\sigma)$ for every $r > D(\varrho\|\sigma)$, and hence the strong converse property, in the infinite-dimensional case it might happen that $D(\varrho\|\sigma) < +\infty$, yet $H^*_r(\varrho\|\sigma) = 0$ for every $r \in \mathbb{R}$, and
hence the type I error sequence does not converge to 1 with an exponential speed along a test sequence \((T_n)_{n \in \mathbb{N}}\), even if \(\lim_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \sigma^\otimes_n T_n > D(\varrho \| \sigma)\). According to Corollary 4.7, this happens if and only if \(D^*_\alpha(\varrho \| \sigma) = +\infty\) for every \(\alpha > 1\). It is an open question what kind of behaviour can occur in this case; if the type II exponent is above the relative entropy, do the type I error probabilities still go to 1 (strong converse property) but with a sub-exponential speed, or may it happen that the strong converse property does not hold, i.e., (4.15) is not satisfied? Note that the monotonicity of the relative entropy under measurements implies that

\[
\lim_{n \to +\infty} \text{Tr} \varrho^\otimes_n (I - T_n) = 0 \implies \limsup_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \sigma^\otimes_n T_n \leq D(\varrho \| \sigma);
\]

see, e.g., the proof of (2.4) in [24], or [23, Proposition 5.2]. In particular, (4.14) cannot hold with \(r > D(\varrho \| \sigma)\) for any test sequence.

4.2. Generalized cutoff rates. Corollary 4.7 gives an operational interpretation to the Hoeffding anti-divergences, but not directly to the sandwiched Rényi divergences. To get such an operational interpretation, one can consider the following quantity, introduced originally in [8] for the finite-dimensional classical case:

**Definition 4.9.** Let \(\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \succeq 0\) and \(\kappa \in (0, 1)\). The generalized \(\kappa\)-cutoff rate \(C_\kappa(\varrho \| \sigma)\) is defined to be the infimum of all \(r_0 \in \mathbb{R}\) such that \(\text{sc}_r(\varrho \| \sigma) \geq \kappa (r - r_0)\) holds for every \(r \in \mathbb{R}\). Analogously, \(C_\kappa(\varrho \| \sigma)_{\text{fa}}\) is defined to be the infimum of all \(r_0 \in \mathbb{R}\) such that \(\text{sc}_r(\varrho \| \sigma)_{\text{fa}} \geq \kappa (r - r_0)\) holds for every \(r \in \mathbb{R}\).

**Proposition 4.10.** Let \(\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \succeq 0\).

(i) For any \(\kappa \in (0, 1)\),

\[
C_\kappa(\varrho \| \sigma)_{\text{fa}} \leq C_\kappa(\varrho \| \sigma) \leq D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma). \tag{4.16}
\]

(ii) If \(\kappa\) is such that there exist \(0 < \kappa_1 < \kappa < \kappa_2 < 1\) for which \(D^*_{\frac{1}{1-\kappa_j}}(\varrho \| \sigma)_{\text{fa}} < +\infty\), \(j = 1, 2\), then

\[
D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma)_{\text{fa}} \leq C_\kappa(\varrho \| \sigma)_{\text{fa}} \leq C_\kappa(\varrho \| \sigma) \leq D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma). \tag{4.17}
\]

If, moreover, \(D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma)_{\text{fa}} = D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma)\), then all the inequalities in (4.17) hold as equalities.

**Proof.** (i) The first inequality in (4.16) is trivial by definition. If \(D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma) = +\infty\) then the second inequality in (4.16) holds trivially, and hence we assume the contrary. By Proposition 4.2,

\[
\text{sc}_r(\varrho \| \sigma) \geq \tilde{H}^*_r(\varrho \| \sigma) = \sup_{u \in [0, 1]} \{ur - \tilde{\psi}^*(\varrho \| \sigma |u)\} \geq \kappa r - \tilde{\psi}^*(\varrho \| \sigma |\kappa)
\]

\[
= \kappa \left( r - \frac{1}{\kappa} \tilde{\psi}^*(\varrho \| \sigma |\kappa) \right) = D^*_{\frac{1}{1-\kappa}}(\varrho \| \sigma).
\]
from which the second inequality in (4.16) follows by definition.

(ii) By the assumptions, \( \tilde{\psi}^*(\varrho\|\sigma|)_{fa} < +\infty, \) \( j = 1, 2, \) and \( \) hence \( \tilde{\psi}^*(\varrho\|\sigma|)_{fa} \) is a finite-valued convex function on \( [\kappa_1, \kappa_2] \), according to Remark 3.36 and Corollary 3.48. Thus, in particular, its left and right derivatives \( \partial^-\tilde{\psi}^*(\varrho\|\sigma|\kappa) \) and \( \partial^+\tilde{\psi}^*(\varrho\|\sigma|\kappa) \) at \( \kappa \) exist and are finite: \(-\infty < \partial^-\tilde{\psi}^*(\varrho\|\sigma|\kappa)_{fa} \leq \partial^+\tilde{\psi}^*(\varrho\|\sigma|\kappa)_{fa} < +\infty. \) Moreover, \( \varrho_0 \leq \sigma_0 \), according to Lemma 3.42. For any \( r \in [\partial^-\tilde{\psi}^*(\varrho\|\sigma|\kappa)_{fa}, \partial^+\tilde{\psi}^*(\varrho\|\sigma|\kappa)_{fa}], \)

\[
\text{sc}_r(\varrho\|\sigma)_V \leq \hat{H}^*_r(\varrho\|\sigma)_{fa} = \max_{u \in [0, 1]} \{ ur - \tilde{\psi}^*(\varrho\|\sigma|u)_{fa} \} = r - \tilde{\psi}^*(\varrho\|\sigma|\kappa)_{fa} = \kappa \left( r - \frac{1}{\kappa} \tilde{\psi}^*(\varrho\|\sigma|\kappa)_{fa} \right) = D^*_{\frac{1}{1-\kappa}}(\varrho\|\sigma)_{fa}
\]

where the first inequality is due to Proposition 4.5. This yields the first inequality in (4.17), and the rest have already been proved in the previous point. \( \square \)

Proposition 4.10 and Corollary 3.40 yield immediately the following:

**Theorem 4.11.** Let \( \varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \), be such that \( \varrho \) and \( \sigma \) are trace-class, or \( \sigma \) is compact and \( \varrho \in \mathcal{B}^\infty(\mathcal{H}, \sigma) \). Let \( \kappa \in (0, 1) \), and assume that \( D^*_\alpha(\varrho\|\sigma) < +\infty \) for \( \alpha \) in a neighborhood of \( \alpha_0 := 1/(1 - \kappa) \). Then

\[
C_\kappa(\varrho\|\sigma)_{fa} = C_\kappa(\varrho\|\sigma) = D^*_\frac{1}{1-\kappa}(\varrho\|\sigma), \quad \text{or equivalently,} \quad D^*_\alpha(\varrho\|\sigma) = C_{\alpha_0^{-1}}(\varrho\|\sigma) = C_{\alpha_0^{-1}}(\varrho\|\sigma)_{fa}.
\]

### 4.3. Monotonicity of the Rényi divergences

The operational representation of the Hoeffding anti-divergences in Sect. 4.1 can be used to obtain the monotonicity of the sandwiched Rényi divergences under quantum operations.

In the Heisenberg picture, a quantum operation from a system with Hilbert space \( \mathcal{H} \) to a system with Hilbert space \( \mathcal{K} \) is given by a unital normal completely positive map \( \Phi : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H}) \), which can be written as

\[
\Phi : \mathcal{B}(\mathcal{K}) \ni A \mapsto V^*(A \otimes I_E)V = \sum_{i \in \mathcal{I}} V_i^*AV_i, \quad \text{(4.18)}
\]

where \( V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}_E \) is an isometry, \( V_i := (I_{\mathcal{K}} \otimes |e_i\rangle\langle e_i|)V \) for some ONB \( (e_i)_{i \in \mathcal{I}} \) in the auxiliary Hilbert space \( \mathcal{H}_E \), and the sum in (4.18) converges in the strong operator topology [17,43]. As in everywhere in the paper, we assume that \( \mathcal{H}, \mathcal{K} \) are separable, in which case the auxiliary Hilbert space \( \mathcal{H}_E \) can be chosen to be separable, and the index set \( \mathcal{I} \) in (4.18) countable.

In the Schrödinger picture, a density operator \( \varrho \in \mathcal{S}(\mathcal{H}) \) is transformed by the predual map

\[
\Phi^*(\varrho) := \sum_{i \in \mathcal{I}} V_i^*\varrho V_i^* = \sum_{i \in \mathcal{I}} (I_{\mathcal{K}} \otimes |e_i\rangle\langle e_i|)\varrho V^*(I_{\mathcal{K}} \otimes |e_i\rangle) = \text{Tr}_E V\varrho V^*, \quad \text{(4.19)}
\]

where the sum converges in trace-norm, and the result is a density operator on \( \mathcal{K} \). Note that the predual map of \( \Phi \) maps from \( L^1(\mathcal{H}) \) to \( L^1(\mathcal{K}) \), and is usually denoted by \( \Phi^* \).

On the other hand, if \( \mathcal{B}(\mathcal{H}) \) and \( \mathcal{B}(\mathcal{K}) \) are both equipped with the ultraweak topology
then the corresponding dual map of $\Phi$ coincides with the predual map, which justifies the notation $\Phi^*$.

If $\varrho$ is PSD but not trace-class then the sum in (4.19) need not converge (in the weak, equivalently, in the strong operator topology), but it may, in which case we say that $\Phi^*$ is defined on $\varrho$, and define $\Phi^*(\varrho) := \sum_{i \in I} V_i \varrho V_i^*$. A trivial case where $\Phi^*$ is defined on every $\varrho \in \mathcal{B}(\mathcal{H})_{\geq 0}$ is when $\Phi$ has only finitely many operators in its Kraus decomposition, or equivalently, $\mathcal{H}_E$ is finite-dimensional.

Lemma 4.12. Let $\varrho \in \mathcal{B}(\mathcal{H})_{\geq 0}$ and $\Phi : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ be a unital normal completely positive map. If $\Phi^*$ is defined on $\varrho$ then

$$\text{Tr} \ A^{1/2} \Phi^*(\varrho) A^{1/2} = \text{Tr} \ \Phi(A)^{1/2} \varrho \Phi(A)^{1/2}, \quad A \in \mathcal{B}(\mathcal{K})_{\geq 0}.$$ 

**Proof.** Let $(f_j)_{j \in J}$ be an orthonormal basis in $\mathcal{K}$. Then

$$\text{Tr} \ A^{1/2} \Phi^*(\varrho) A^{1/2} = \sum_{j \in J} \left( A^{1/2} f_j, \Phi^*(\varrho) A^{1/2} f_j \right) = \sum_{i \in I} \left( A^{1/2} f_j, \varrho V_i^* A^{1/2} f_j \right)$$

$$= \sum_{i \in I} \sum_{j \in J} \left( A^{1/2} f_j, V_i \varrho V_i^* A^{1/2} f_j \right) = \text{Tr} A^{1/2} V_i \varrho V_i^* A^{1/2} = \text{Tr} \varrho V_i^* AV_i \varrho^{1/2}$$

$$= \sum_{i \in I} \text{Tr} \varrho^{1/2} V_i^* AV_i \varrho^{1/2}$$

$$= \sum_{i \in I} \sum_{j \in J} \left( \varrho^{1/2} f_j, V_i^* AV_i \varrho^{1/2} f_j \right)$$

$$= \sum_{j \in J} \sum_{i \in I} \left( \varrho^{1/2} f_j, V_i^* AV_i \varrho^{1/2} f_j \right) = \text{Tr} \varrho^{1/2} \Phi(A) \varrho^{1/2} = \text{Tr} \Phi(A)^{1/2} \varrho \Phi(A)^{1/2}.$$

$\square$

The transformation on multiple systems is given by

$$\Phi^*_{\otimes n} : \mathcal{B}(\mathcal{K}^\otimes n) \ni A \mapsto (V^\otimes n)^*(A \otimes I_E)^n V^\otimes n$$

$$= \sum_{i \in I^n} (V_i \otimes \ldots \otimes V_{i_n})^* A (V_i \otimes \ldots \otimes V_{i_n}). \quad (4.20)$$

If $\Phi^*$ is defined on $\varrho \in \mathcal{B}(\mathcal{H})_{\geq 0}$ then $(\Phi^*_{\otimes n})^*$ is defined on $\varrho_{\otimes n}$, and $(\Phi^*_{\otimes n})^*(\varrho_{\otimes n}) = (\Phi^*(\varrho))^\otimes n$.

In the context of operator discrimination, a transformation $\Phi$ effectively reduces the available tests for discriminating $\varrho$ and $\sigma$, thereby increasing the strong converse exponent, as expressed by the following:

Lemma 4.13. Let $\varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0}$ and $\Phi : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ be a unital normal completely positive map such that $\Phi^*$ is defined on $\varrho$ and $\sigma$. Then

$$\mathcal{S}_r(\Phi^*(\varrho) \| \Phi^*(\sigma)) \geq \mathcal{S}_r(\varrho \| \sigma), \quad \overline{\mathcal{S}}_r(\Phi^*(\varrho) \| \Phi^*(\sigma)) \geq \overline{\mathcal{S}}_r(\varrho \| \sigma), \quad r \in \mathbb{R}.$$
Proof. We only prove the assertion for \( \bar{\mathcal{C}}_r \), as the proof for \( \overline{\mathcal{S}}_r \) goes the same way. We have

\[
\bar{\mathcal{C}}_r(\Phi^*(\varrho)\|\Phi^*(\sigma)) = \inf \left\{ \liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \, T_n^{1/2} (\Phi^*(\varrho)) \otimes n T_n^{1/2} : \right.
\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \, T_n^{1/2} (\Phi^*(\sigma)) \otimes n T_n^{1/2} \geq r \right\}
\]

\[
= \inf \left\{ \liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr}(\Phi^\otimes n(T_n))^{1/2} \varrho^\otimes n (\Phi^\otimes n(T_n))^{1/2} : \right.
\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr}(\Phi^\otimes n(T_n))^{1/2} \sigma^\otimes n (\Phi^\otimes n(T_n))^{1/2} \geq r \right\}
\]

\[
\geq \inf \left\{ \liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \, S_n^{1/2} \varrho^\otimes n S_n^{1/2} : \right.
\[
\liminf_{n \to +\infty} -\frac{1}{n} \log \text{Tr} \, S_n^{1/2} \sigma^\otimes n S_n^{1/2} \geq r \right\}
\]

\[
= \bar{\mathcal{C}}_r(\varrho\|\sigma),
\]

where the first two infima are taken over sequences of tests \( T_n \in \mathcal{B}(\mathcal{K}^\otimes n)_{[0,1]} \), \( n \in \mathbb{N} \), satisfying the given conditions, the third infimum is taken over sequences of tests \( S_n \in \mathcal{B}(\mathcal{H}^\otimes n)_{[0,1]} \), \( n \in \mathbb{N} \), satisfying the given condition, the first equality is by definition, the second equality follows from Lemma 4.12, the inequality is obvious from the fact that \( T_n \in \mathcal{B}(\mathcal{K}^\otimes n)_{[0,1]} \implies \Phi^\otimes n(T_n) \in \mathcal{B}(\mathcal{H}^\otimes n)_{[0,1]} \), and the last equality is again by definition.

The proof of the following monotonicity result is similar to the proof of the analogous result given in [37, Remark 2] for the monotonicity of the Petz-type Rényi divergences and finite-dimensional density operators. The main ideas in the proof are using the bounds on the strong converse exponents given in Proposition 4.5, the monotonicity of the strong converse exponents given in Lemma 4.13, and the fact that the Rényi divergences can be expressed from the Hoeffding anti-divergences by Legendre-Fenchel transformation, i.e., Lemma 3.53.

Theorem 4.14. Let \( \varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \), be such that

\[
D_\alpha^*(\varrho\|\sigma) = D_\alpha^*(\varrho\|\sigma), \quad \alpha > 1,
\]

and let \( \Phi : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H}) \) be a unital normal completely positive linear that is defined on \( \varrho \) and \( \sigma \). Then

\[
D_\alpha^*(\Phi^*(\varrho)\|\Phi^*(\sigma)) = D_\alpha^*(\varrho\|\sigma), \quad \alpha > 1,
\]

Proof. By assumption, \( \tilde{\mathcal{H}}_r^*(\varrho\|\sigma) = \mathcal{H}_r^*(\varrho\|\sigma), \, r \in \mathbb{R} \), and

\[
\mathcal{H}_r^*(\varrho\|\sigma) = \mathcal{H}_r^*(\varrho\|\sigma) \leq \bar{\mathcal{C}}_r(\varrho\|\sigma) \leq \bar{\mathcal{C}}_r(\Phi^*(\varrho)\|\Phi^*(\sigma)) \]

\[
\leq \tilde{\mathcal{H}}_r^*(\Phi^*(\varrho)\|\Phi^*(\sigma)), \quad r \in \mathbb{R},
\]

where the first and the last inequalities follow from Proposition 4.5, and the second inequality from Lemma 4.13. Hence, by Lemma 3.53,

\[
\tilde{y}^*(\varrho\|\sigma|u) = \sup_{r \in \mathbb{R}} \left\{ ur - \mathcal{H}_r^*(\varrho\|\sigma) \right\} \geq \sup_{r \in \mathbb{R}} \left\{ ur - \tilde{\mathcal{H}}_r^*(\Phi^*(\varrho)\|\Phi^*(\sigma)) \right\}
\]

Theorem 4.14. Let \( \varrho, \sigma \in \mathcal{B}(\mathcal{H})_{\geq 0} \), be such that

\[
D_\alpha^*(\varrho\|\sigma) = D_\alpha^*(\varrho\|\sigma), \quad \alpha > 1,
\]

and let \( \Phi : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H}) \) be a unital normal completely positive linear that is defined on \( \varrho \) and \( \sigma \). Then

\[
D_\alpha^*(\Phi^*(\varrho)\|\Phi^*(\sigma)) \leq D_\alpha^*(\varrho\|\sigma), \quad \alpha > 1.
\]

Proof. By assumption, \( \tilde{\mathcal{H}}_r^*(\varrho\|\sigma) = \mathcal{H}_r^*(\varrho\|\sigma), \, r \in \mathbb{R} \), and

\[
\mathcal{H}_r^*(\varrho\|\sigma) = \mathcal{H}_r^*(\varrho\|\sigma) \leq \bar{\mathcal{C}}_r(\varrho\|\sigma) \leq \bar{\mathcal{C}}_r(\Phi^*(\varrho)\|\Phi^*(\sigma)) \]

\[
\leq \tilde{\mathcal{H}}_r^*(\Phi^*(\varrho)\|\Phi^*(\sigma)), \quad r \in \mathbb{R},
\]

where the first and the last inequalities follow from Proposition 4.5, and the second inequality from Lemma 4.13. Hence, by Lemma 3.53,

\[
\tilde{y}^*(\varrho\|\sigma|u) = \sup_{r \in \mathbb{R}} \left\{ ur - \mathcal{H}_r^*(\varrho\|\sigma) \right\} \geq \sup_{r \in \mathbb{R}} \left\{ ur - \tilde{\mathcal{H}}_r^*(\Phi^*(\varrho)\|\Phi^*(\sigma)) \right\}
\]
\[ = \tilde{\psi}^*(\Phi^*(\varrho)\|\Phi^*(\sigma)|u)_{fa}, \quad u \in \mathbb{R}, \]

which is equivalent to (4.22).

**Corollary 4.15.** Let \( \varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0 \), and let \( \Phi : \mathcal{B}(\mathcal{K}) \rightarrow \mathcal{B}(\mathcal{H}) \) be a unital normal completely positive map. Assume that

(a) \( \varrho \) and \( \sigma \) are both trace-class,

or

(b) \( \Phi^* \) is defined on \( \varrho \) and \( \sigma \), and \( \Phi^*(\varrho) \) and \( \Phi^*(\sigma) \) are compact, and \( \varrho \in \mathcal{B}_\infty(\mathcal{H}, \sigma) \).

Then

\[ D^*_\alpha(\Phi^*(\varrho)\|\Phi^*(\sigma)) \leq D^*_\alpha(\varrho\|\sigma), \quad \alpha > 1. \] (4.23)

**Proof.** If \( \varrho \) and \( \sigma \) are both trace-class then \( \Phi^* \) is automatically defined on them, and \( \Phi^*(\varrho) \) and \( \Phi^*(\sigma) \) are both trace-class. Note that \( \varrho \in \mathcal{B}_\infty(\mathcal{H}, \sigma) \iff \varrho \leq \lambda\sigma \) for some \( \lambda \geq 0 \), whence \( \Phi^*(\varrho) \leq \lambda\Phi^*(\sigma) \), i.e., \( \Phi^*(\varrho) \in \mathcal{B}_\infty(\mathcal{K}, \Phi^*(\sigma)) \). By Remark 3.6, \( \varrho \in \mathcal{B}^\alpha(\mathcal{H}, \sigma) \) and \( \Phi^*(\varrho) \in \mathcal{B}^\alpha(\mathcal{K}, \Phi^*(\sigma)), \alpha > 1 \). Thus, by Lemma 3.40, the assumptions guarantee that \( D^*_\alpha(\varrho\|\sigma)_{fa} = D^*_\alpha(\varrho\|\sigma) \) and \( D^*_\alpha(\Phi^*(\varrho)\|\Phi^*(\sigma))_{fa} = D^*_\alpha(\Phi^*(\varrho)\|\Phi^*(\sigma)), \alpha > 1 \), and therefore (4.23) follows immediately from Theorem 4.14. \( \square \)

**Remark 4.16.** Monotonicity of the form (4.23) in the case where both \( \varrho \) and \( \sigma \) are trace-class is a special case of [6, Theorem 14] and [27, Theorem 3.14], where monotonicity was proved in the more general setting of normal positive linear functionals on a von Neumann algebra. Our proof above is completely different from the proofs given in [6] and [27].

5. Conclusion

We have shown that for any \( \alpha > 1 \), the sandwiched Rényi \( \alpha \)-divergence of infinite-dimensional density operators has the same operational interpretation in the context of state discrimination as in the finite-dimensional case, and also that it coincides with the regularized measured Rényi \( \alpha \)-divergence, again analogously to the finite-dimensional case. Our results can be extended to more general operator algebraic settings, as shown in [22].

It is worth noting that while in [33] the equality of the sandwiched Rényi divergence and the regularized measured Rényi divergence was an important ingredient of showing the equality of the strong converse exponent and the Hoeffding anti-divergence, the extensions to the infinite-dimensional case can be done separately, building in each problem only on the corresponding finite-dimensional result and the recoverability of the sandwiched Rényi divergences from finite-dimensional restrictions.

We also considered the extension of the sandwiched Rényi divergences (and more generally, Rényi \((\alpha, z)\)-divergences) to pairs of not necessarily trace-class positive semi-definite operators, and established some properties of this extension. Related to this, we considered a generalization of the state discrimination problem, where the hypotheses may be represented by general positive semi-definite operators. We gave bounds on the strong converse exponent in this problem, and showed that at least in some cases, the equality between the strong converse exponent and the Hoeffding anti-divergence still holds in this generalized setting.

There are a number of interesting problems left open in the paper. Probably the most important is clarifying whether \( Q^*_\alpha(\varrho\|\sigma) = Q^*_\alpha(\varrho\|\sigma)_{fa} \) holds for every pair of PSD operators \( \varrho, \sigma \) and every \( \alpha > 1 \), and if not then whether there exist other examples for
which it holds, apart from the ones given in Proposition 3.40 and Lemma 3.42. Such examples would extend the applicability of Proposition 4.5 and Theorem 4.14, among others. While less relevant for the problem of operator discrimination, the same question may be asked for more general Rényi \((\alpha, z)\)-divergences, which seems interesting from the matrix analysis point of view. Finally, from the point of view of quantum information theory, the most important question seems to be to clarify the optimal asymptotics of the type I error probability when the type II exponent is strictly above the relative entropy of the two states (assuming that the latter is finite), while all their sandwiched Rényi \(\alpha\)-divergences are \(+\infty\) for \(\alpha > 1\); see Remark 4.8.

Acknowledgements. This work was partially funded by the National Research, Development and Innovation Office of Hungary via the research grants K124152 and KH129601, and by the Ministry of Innovation and Technology and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary. The author is grateful to Péter Vrana for discussions at the early stage of the project, to Ludovico Lami for comments on the strong converse property that led to Remark 4.8, and to Fumio Hiai for numerous helpful comments. The author is also grateful to an anonymous referee for several comments that helped to improve the presentation.

Funding Open access funding provided by Budapest University of Technology and Economics.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A: Some Further Properties of the Hoeffding Anti-divergences

Lemma A.1. For any \(\varrho, \sigma \in \mathcal{B}(\mathcal{H}) \geq 0\), and for every \(u \in \mathbb{R}\),

\[
\begin{align*}
  n \mapsto \tilde{\psi}^\ast(\varrho \otimes^n \| \sigma \otimes^n |u\rangle_{\text{fa}}) & \quad \text{is superadditive,} \\
  n \mapsto \frac{1}{2^n} \tilde{\psi}^\ast(\varrho \otimes 2^n \| \sigma \otimes 2^n |u\rangle_{\text{fa}}) & \quad \text{is monotone increasing,}
\end{align*}
\]

and

\[
\tilde{\psi}^\ast(\varrho \| \sigma |u\rangle_{\text{fa}}) = \lim_{n \to +\infty} \frac{1}{n} \tilde{\psi}^\ast(\varrho \otimes^n \| \sigma \otimes^n |u\rangle_{\text{fa}}) = \lim_{n \to +\infty} \frac{1}{2^n} \tilde{\psi}^\ast(\varrho \otimes 2^n \| \sigma \otimes 2^n |u\rangle_{\text{fa}}) = \sup_{n \in \mathbb{N}} \frac{1}{2^n} \tilde{\psi}^\ast(\varrho \otimes 2^n \| \sigma \otimes 2^n |u\rangle_{\text{fa}}). \quad (A.1)
\]

Proof. The assertions are trivial when \(u \in \mathbb{R} \setminus [0, 1]\), because then all quantities above are equal to \(+\infty\). For \(u \in (0, 1)\), superadditivity is obvious from restricting to projections of the form \(P_1 \otimes P_2, P_1 \in \mathbb{P}_f(\mathcal{H}^\otimes n)_{\text{fa}}^+\), \(P_2 \in \mathbb{P}_f(\mathcal{H}^\otimes m)_{\text{fa}}^+\) in the definition
of $\tilde{\psi}^*(q^{\otimes (n+m)} \| \sigma^{\otimes (n+m)} | u)_{fa}$, according to Lemma 3.22, and the cases $u \in [0, 1]$ follow by taking limits (in fact, even additivity holds there, according to (3.72)–(3.73)). The rest of the assertions are straightforward consequences.

Lemma A.2. For any $q, \sigma \in B(H) \geq 0$, and every $n \in \mathbb{N}$,

$$\tilde{\psi}^*(q^{\otimes n} \| \sigma^{\otimes n} | u) = n \tilde{\psi}^*(q \| \sigma | u), \quad u \in (0, 1),$$

$$\tilde{\psi}^*(q^{\otimes n} \| \sigma^{\otimes n} | u)_{fa} = n \tilde{\psi}^*(q \| \sigma | u)_{fa}, \quad u \in \mathbb{R}.$$

Proof. The first equality is immediate from the multiplicativity of $Q^*_\alpha$, given in Lemma 3.22, and the second one follows from (A.1), as

$$\tilde{\psi}^*(q^{\otimes n} \| \sigma^{\otimes n} | u)_{fa} = \lim_{k \to +\infty} \frac{1}{k} \tilde{\psi}^*(q^{\otimes kn} \| \sigma^{\otimes kn} | u)_{fa} = n \lim_{k \to +\infty} \frac{1}{nk} \tilde{\psi}^*(q^{\otimes kn} \| \sigma^{\otimes kn} | u)_{fa} = \tilde{\psi}^*(q \| \sigma | u)_{fa}.$$ □

Lemma A.3. For any $q, \sigma \in B(H) \geq 0$ and $r \in \mathbb{R}$,

$$n \mapsto H^*_r(q^{\otimes n} \| \sigma^{\otimes n}) \text{ is subadditive,}$$

where $H^*_r$ stands for any of the Hoeffding anti-divergences in Definition 3.51. Moreover,

$$H^*_r(q^{\otimes n} \| \sigma^{\otimes n}) = n H^*_r(q \| \sigma), \quad \hat{H}^*_r(q^{\otimes n} \| \sigma^{\otimes n})_{fa} = n \hat{H}^*_r(q \| \sigma)_{fa}, \quad n \in \mathbb{N}. \quad (A.3)$$

Proof. (A.3) is immediate from Lemma A.2, and it trivially implies subadditivity for the quantities defined in (3.76), (3.78), (3.79), (3.81). By Lemma A.1, $n \mapsto \tilde{\psi}^*(q^{\otimes n} \| \sigma^{\otimes n} | u)_{fa}$ is superadditive for every $u \in \mathbb{R}$, which implies the subadditivity of $n \mapsto H^*_r(q^{\otimes n} \| \sigma^{\otimes n})_{fa}$ and $n \mapsto \hat{H}^*_r(q^{\otimes n} \| \sigma^{\otimes n})_{fa}$. □

Proposition A.4. For any $q, \sigma \in B(H) \geq 0$ and $r \in \mathbb{R}$,

$$\hat{H}^*_r(q \| \sigma)_{fa} = \inf_{n \in \mathbb{N}} \frac{1}{n} \hat{H}^*_r(q^{\otimes n} \| \sigma^{\otimes n})_{fa} = \lim_{n \to +\infty} \frac{1}{n} \hat{H}^*_r(q^{\otimes n} \| \sigma^{\otimes n})_{fa} = \inf_{n \in \mathbb{N}} \frac{1}{2n} \hat{H}^*_r(q^{\otimes 2n} \| \sigma^{\otimes 2n})_{fa}. \quad (A.4)$$

Proof. The second equality in (A.4) follows from the subadditivity of $n \mapsto \hat{H}^*_r(q^{\otimes n} \| \sigma^{\otimes n})_{fa}$ given in Lemma A.3. To see the other two equalities in (A.4)–(A.5), note that by definition,

$$\hat{H}^*_r(q \| \sigma)_{fa} = \max_{u \in [0, 1]} \inf_{n \in \mathbb{N}} \frac{1}{n} \left\{ unr - \tilde{\psi}^*(q^{\otimes n} \| \sigma^{\otimes n} | u)_{fa} \right\}$$

$$\leq \inf_{n \in \mathbb{N}} \max_{u \in [0, 1]} \frac{1}{n} \left\{ unr - \tilde{\psi}^*(q^{\otimes n} \| \sigma^{\otimes n} | u)_{fa} \right\} = \inf_{n \in \mathbb{N}} \frac{1}{n} \hat{H}^*_n(q^{\otimes n} \| \sigma^{\otimes n} | u)_{fa}$$

$$\leq \inf_{n \in \mathbb{N}} \max_{u \in [0, 1]} \frac{1}{2n} \left\{ u2^n r - \tilde{\psi}^*(q^{\otimes 2n} \| \sigma^{\otimes 2n} | u)_{fa} \right\} = \inf_{n \in \mathbb{N}} \frac{1}{2n} \hat{H}^*_2(q^{\otimes 2n} \| \sigma^{\otimes 2n} | u)_{fa}.$$
On the other hand, by Corollary 3.48 and Lemma A.1, \( \frac{1}{2n} \left\{ u^{2n} r - \tilde{\psi}^* (\rho \otimes \sigma | u)_{fa} \right\} \) is upper semi-continuous in \( u \) on the compact set \([0, 1]\), and monotone decreasing in \( n \), whence

\[
\inf_{n \in \mathbb{N}} \frac{1}{2n} \tilde{H}^n_{2n,r} (\rho \otimes \sigma | u)_{fa} = \inf_{n \in \mathbb{N}} \max_{u \in [0, 1]} \frac{1}{2n} \left\{ u^{2n} r - \tilde{\psi}^* (\rho \otimes \sigma | u)_{fa} \right\} = \max_{u \in [0, 1]} \left\{ u r - \tilde{\psi}^* (\rho | u)_{\mathbb{F}} \right\} = \hat{H}^n_r (\rho | \sigma)_{\mathbb{F}},
\]

where the second equality is due to Lemma 2.5, and the third equality follows from Lemma A.1.

References

1. Araki, H.: Relative entropy of states of von Neumann algebras. Publ. RIMS Kyoto Univ. 11, 809–833 (1976)
2. Araki, H.: On an inequality of Lieb and Thirring. Lett. Math. Phys. 19, 167–170 (1990)
3. Audenaert, K.M.R., Nussbaum, M., Szkoła, A., Verstraete, F.: Asymptotic error rates in quantum hypothesis testing. Commun. Math. Phys. 279, 251–283 (2008). arXiv:0708.4282
4. Audenaert, K.M.R., Datta, N.: \( \alpha \)-\( z \)-relative Renyi entropies. J. Math. Phys. 56, 022202 (2015). arXiv:1310.7178
5. Berta, M., Fawzi, O., Tomamichel, M.: On variational expressions for quantum relative entropies. Lett. Math. Phys. 107(12), 2239–2265 (2017). arXiv:1512.02615
6. Berta, M., Scholz, V.B., Tomamichel, M.: Renyi divergences as weighted non-commutative vector-valued \( L_p \)-spaces. Ann. Henri Poincaré 19, 1843–1867 (2018). arXiv:1608.05317
7. Brown, L.G., Kosaki, H.: Jensen’s inequality in semi-finite von Neumann algebras. J. Oper. Theory 23(1), 3–19 (1990)
8. Csiszár, I.: Generalized cutoff rates and Rényi’s information measures. IEEE Trans. Inf. Theory 41(1), 26–34 (1995)
9. Datta, N.: Min- and max-relative entropies and a new entanglement monotone. IEEE Trans. Inf. Theory 55(6), 2816–2826 (2009)
10. Douglas, R.G.: On majorization, factorization, and range inclusion of operators on Hilbert space. Proc. Am. Math. Soc. 47(2), 413–415 (1966)
11. Ekeland I, Témam R: Convex Analysis and Variational Problems. SIAM (1999)
12. Frank, R.L., Lieb, E.H.: Monotonicity of a relative Rényi entropy. J. Math. Phys. 54(12), 122201 (2013). arXiv:1306.5358
13. Grömm, H.: Two theorems about \( C_p \). Rep. Math. Phys. 4(3), 211–215 (1973)
14. 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)
15. Hayashi, M.: Error exponent in asymptotic quantum hypothesis testing and its application to classical-quantum channel coding. Phys. Rev. A 76(6), 062301 (2007). arXiv:quant-ph/0611013
16. Hayashi, M.: Quantum information theory: mathematical foundation, 2nd ed. In Graduate Texts in Physics. Springer (2017)
17. Hellwig, K.-E., Kraus, K.: Operations and measurements: II. Commun. Math. Phys. 16, 142–147 (1970)
18. Hiai, F.: Log-majorizations and norm inequalities for exponential operators. Banach Center Publ. 38, 119–181 (1997)
19. Hiai, F.: Quantum \( f \)-divergences in von Neumann algebras: I-standard \( f \)-divergences. J. Math. Phys. 59, 102202 (2018)
20. Hiai, F.: Private communication (2021)
21. Hiai, F.: Quantum \( f \)-Divergences in von Neumann Algebras. Springer, Berlin (2021)
22. Hiai, F., Mosonyi, M.: Quantum Rényi divergences and the strong converse exponent of state discrimination in operator algebras. Annales Henri Poincaré, 2022. arXiv:2110.07320
23. Hiai, F., Mosonyi, M., Ogawa, T.: Error exponents in hypothesis testing for correlated states on a spin chain. J. Math. Phys. 49, 032112 (2008)
24. Hiai, F., Petz, D.: The proper formula for relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. 143(1), 99–114 (1991)
25. Jaksic, V., Ogata, Y., Pautrat, Y., Pillet, C.-A.: Entropic fluctuations in quantum statistical mechanics. an introduction. In: Quantum Theory from Small to Large Scales, August 2010, volume 95 of Lecture Notes of the Les Houches Summer School. Oxford University Press (2012)
26. Jaksic, V., Ogata, Y., Pillet, C.-A., Seiringer, R.: Quantum hypothesis testing and non-equilibrium statistical mechanics. Rev. Math. Phys. 24(6), 1230002 (2012). arXiv:1109.3804
27. Jenčová, A.: Rényi relative entropies and noncommutative $L_p$-spaces. Ann. Henri Poincaré 19, 2513–2542 (2018). arXiv:1609.08462
28. Jenčová, A.: Rényi relative entropies and noncommutative $L_p$-spaces II. Ann. Henri Poincaré 22, 3235–3254 (2021). arXiv:1707.00047
29. Kosaki, H.: Interpolation theory and the Wigner-Yanase-Dyson-Lieb concavity. Commun. Math. Phys. 87, 315–329 (1982)
30. Li, Y., Gao, S., Hao, H.: The sandwiched Rényi divergence and quantum positive evidence order in infinite-dimensional Hilbert space. Rep. Math. Phys. 88(2), 175–193 (2021)
31. McCarthy, C.A.: $C_p$. Israel J. of Math. 5, 249–271 (1967)
32. Mosonyi, M., Hiai, F.: On the quantum Rényi relative entropies and related capacity formulas. IEEE Trans. Inf. Theory 57(4), 2474–2487 (2011)
33. Mosonyi, M., Ogawa, T.: Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. Commun. Math. Phys. 334(3), 1617–1648 (2015). arXiv:1309.3228
34. Mosonyi, M., Ogawa, T.: Divergence radii and the strong converse exponent of classical-quantum channel coding with constant compositions. IEEE Trans. Inf. Theory 67(3), 1668–1698 (2021). arXiv:1811.10599
35. Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., Tomamichel, M.: On quantum Rényi entropies: a new generalization and some properties. J. Math. Phys. 54(12), 122203 (2013). arXiv:1306.3142
36. Nagaoka, H.: Strong converse theorems in quantum information theory. In: Proceedings of ERATO Workshop on Quantum Information Science, page 33, 2001. Also appeared in Asymptotic Theory of Quantum Statistical Inference, ed. M. Hayashi, World Scientific (2005)
37. Nagaoka, H.: The converse part of the theorem for quantum Hoeffding bound. arXiv:quant-ph/0611289 (2006)
38. Ogawa, T., Nagaoka, H.: Strong converse and Stein’s lemma in quantum hypothesis testing. IEEE Trans. Inf. Theory 46(7), 2428–2433 (2000). arXiv:quant-ph/9906090
39. Petz, D.: Quasi-entropies for states of a von Neumann algebra. Publ. RIMS Kyoto Univ. 21, 787–800 (1985)
40. Petz, D.: Quasi-entropies for finite quantum systems. Rep. Math. Phys. 23, 57–65 (1986)
41. Renner, R.: Security of Quantum Key Distribution. PhD thesis, Swiss Federal Institute of Technology Zurich (2005). Diss. ETH No. 16242
42. Rényi, A.: On measures of entropy and information. In Proc. 4th Berkeley Sympos. Math. Statist. and Prob., volume I, pages 547–561. Univ. California Press, Berkeley, California (1961)
43. Stinespring, W.F.: Positive functions on $C^*$-algebras. Proc. Am. Math. Soc. 6, 211–216 (1955)
44. Tomamichel, M., Berta, M., Hayashi, M.: Relating different quantum generalizations of the conditional Rényi entropy. J. Math. Phys. 55, 082206 (2014). arXiv:1311.3887
45. Umegaki, H.: Conditional expectation in an operator algebra, IV (entropy and information). Kodai Math. Sem. Rep. 14, 59–85 (1962)
46. Wilde, M.M., Winter, A., Yang, D.: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. Commun. Math. Phys. 331(2), 593–622 (2014). arXiv:1306.1586
47. Zhang, H.: From Wigner-Yanase-Dyson conjecture to Carlen-Frank-Lieb conjecture. Adv. Math. 365, 107053 (2020). arXiv:1811.01205

Communicated by G. Chiribella