

RÉNYI DIVERGENCES AS WEIGHTED NON-COMMUTATIVE VECTOR VALUED L_p -SPACES

MARIO BERTA, VOLKHER B. SCHOLZ, AND MARCO TOMAMICHEL

ABSTRACT. We show that Araki and Masuda’s weighted non-commutative vector valued L_p -spaces [Araki & Masuda, Publ. Res. Inst. Math. Sci., 18:339 (1982)] correspond to an algebraic generalization of the sandwiched Rényi divergences with parameter $\alpha = \frac{p}{2}$. Using complex interpolation theory, we prove various fundamental properties of these divergences in the setup of von Neumann algebras, including a data processing inequality and monotonicity in α . We thereby also give new proofs for the corresponding finite-dimensional properties. We discuss the limiting cases $\alpha \rightarrow \{\frac{1}{2}, 1, \infty\}$ leading to minus the logarithm of Uhlmann’s fidelity, Umegaki’s relative entropy, and the max-relative entropy, respectively. As a contribution that might be of independent interest, we derive a Riesz-Thorin theorem for Araki-Masuda L_p -spaces and an Araki-Lieb-Thirring inequality for states on von Neumann algebras.

1. INTRODUCTION

Information-theoretic concepts are often developed under the assumption that the underlying physical systems are described by finite-dimensional Hilbert spaces. It is, however, of fundamental interest to understand which finite-dimensional concepts generalize to physical systems modeled by von Neumann algebras, e.g. certain quantum field theories. Moreover, the translation of finite-dimensional arguments to a more abstract theory often requires new proof ideas or at least a streamlining of the original arguments—thereby leading us to a better understanding of the finite-dimensional special case as well.

In this work we explore a generalization to the setup of von Neumann algebras of a family of divergences, the sandwiched Rényi divergences [MLDS⁺13, WWY14], which have found operational meaning and applications in quantum information theory (see, e.g., [WWY14, MO15, MO14, TWW14, Tom16]). A connection between sandwiched Rényi divergences and weighted non-commutative L_p -spaces has been exploited already in the defining works [Tom12, MLDS⁺13, WWY14] and by Beigi [Bei13] for finite-dimensional systems. In generalizing sandwiched Rényi divergences to the algebraic setting it is thus immediate to look for a connection between those and non-commutative L_p spaces defined on von Neumann algebras. This is of course complicated by the fact that when dealing with von Neumann algebras the existence of a trace is not guaranteed and states are not represented by density operators. Nevertheless, quite some progress has been achieved in the recent decades on defining and studying non-commutative L_p spaces defined with respect to von Neumann algebras (see, e.g., [PX03] for an overview). The most well-known such theory are Haagerup’s non-commutative L_p -spaces [Haa79], as well as their weighted versions (see, e.g., [HJX09]). These are based on suitably extending the algebra such that a trace becomes available again. Another generalization of the commutative theory to von Neumann algebras was

given by Araki and Masuda [AM82], who define p -norms on state vectors (corresponding to purifications in the finite-dimensional case) in the underlying Hilbert space \mathcal{H} , not directly on the state space.

We find that the latter offers a natural generalization of the sandwiched Rényi divergences with parameter $\alpha = \frac{p}{2}$, and we hereafter call this generalization Araki-Masuda divergences. Extending their work, we show various properties of this divergence, including a data-processing inequality and monotonicity in α . We will also discuss their relation with other divergences such as Umegaki's [Ume62] relative entropy (which was generalized to von Neumann algebras by Araki [Ara76]) and Petz' generalization of Rényi divergences [Pet85, Pet86] (see also [OP93] for an overview). Finally, we conjecture that the Araki-Masuda divergences characterize the strong converse exponent of binary hypothesis testing on von Neumann algebras, complementing recent work by Jakšić *et al.* [JOPP12]. We hope that our work will lead to further generalizations of information theoretic results to von Neumann algebras such as in [BFS16].

Outline. This paper is organized as follows. In Section 2 we introduce our notation and discuss a few algebraic concepts that we need in the following. In Section 3 we slightly generalize the definition of Araki and Masuda's weighted non-commutative vector valued L_p -spaces and discuss some properties. In particular, we prove a Riesz-Thorin theorem for Araki-Masuda L_p -spaces. In Section 4 we define the Araki-Masuda divergences, which correspond to an algebraic generalization of the sandwiched Rényi divergences. We then discuss various properties and show in particular an Araki-Lieb-Thirring inequality for states on von Neumann algebras. We conclude with Section 5 where we discuss our conjecture for an operational interpretation.

2. ALGEBRAIC SETUP

2.1. Notation. We tried to optimize our notation so that the manuscript is most accessible to the quantum information community, thereby sometimes disregarding established conventions in operator theory. We also comment, in framed boxes, on how to translate certain concepts into the language of finite-dimensional quantum theory (sometimes we comment on the commutative case as well).

In these notes, \mathcal{M} denotes a W^* -algebra, $(\cdot)^\dagger$ is its involution, and $\text{id} \in \mathcal{M}$ its multiplicative identity. An element $x \in \mathcal{M}$ is called positive if $x = a^\dagger a$ for some $a \in \mathcal{M}$, and the cone of positive elements of \mathcal{M} is denoted by \mathcal{M}_+ . We denote by $\mathcal{P}(\mathcal{M})$ the space of normal positive linear functionals on \mathcal{M} , i.e. elements of the predual of \mathcal{M} that map \mathcal{M}_+ onto the positive real axis. The subset $\mathcal{S}(\mathcal{M}) := \{\rho \in \mathcal{P}(\mathcal{M}) : \rho(\text{id}) = 1\}$ contains normal states.

For two functionals $\rho, \sigma \in \mathcal{P}(\mathcal{M})$ we write $\rho \ll \sigma$ and say that ρ is supported on (the support of) σ if and only if $\sigma(x) = 0 \implies \rho(x) = 0$ for all $x \in \mathcal{M}$. Strictly stronger, we say that ρ is dominated by σ and write $\rho \lll \sigma$ if and only if there exists a positive constant C such that $\rho(x) \leq C\sigma(x)$ for all $x \in \mathcal{M}_+$.

Moreover, let $\pi : \mathcal{M} \rightarrow \mathcal{B}(\mathcal{H})$ be a $*$ -representation of \mathcal{M} as bounded operators acting on a Hilbert space \mathcal{H} such that $\pi(\mathcal{M}) \subset \mathcal{B}(\mathcal{H})$ forms a von Neumann algebra. The Hilbert space \mathcal{H} has an inner product, $\langle \cdot | \cdot \rangle$, which is taken to be anti-linear in the first variable. We say that a vector $\omega \in \mathcal{H}$ implements a functional $\omega \in \mathcal{P}(\mathcal{M})$ if and only if $\omega(a) = \langle \omega | \pi(a) \omega \rangle$ for all $a \in \mathcal{M}$. We will consistently use the same Greek letter to denote functionals and vectors that are related in this way, i.e. every vector implicitly defines a corresponding functional with the same symbol. For any $\sigma \in \mathcal{P}(\mathcal{M})$ we define its support projector $P_\sigma \in \mathcal{M}$ as the

minimal projector satisfying $\sigma(P_\sigma) = 1$. We define the set $\mathcal{H}_\sigma^* := \{\rho \in \mathcal{H} : \rho \lll \sigma\}$ and its closure, the subspace \mathcal{H}_σ , which contains all vectors ρ implementing functionals with $\rho \ll \sigma$. The Gelfand-Naimark-Segal (GNS) construction provides a $*$ -representation $\pi_\sigma : \mathcal{M} \rightarrow \mathcal{B}(\mathcal{G}_\sigma)$ as well as an implementation $\sigma \in \mathcal{G}_\sigma$ of σ that is faithful in \mathcal{G}_σ . Moreover, we may embed \mathcal{G}_σ as \mathcal{H}_σ into \mathcal{H} if the latter space allows for an implementation of σ . In this case we have $\pi_\sigma(x) = \pi(P_\sigma x P_\sigma)$.

We will often draw on intuition from the finite-dimensional case. Let \mathcal{M}_n denote the algebra of $n \times n$ matrices, which we conveniently represent as acting on the first tensor factor of a Hilbert space $\mathbb{C}^n \otimes \mathbb{C}^n$, i.e. $\pi(x) = x \otimes \text{id}_n$. The trace Tr is implemented by a vector $\tau = \sum_{i=1}^n e_i \otimes e_i$ for some orthonormal basis $\{e_i\}_{i=1}^n$ of \mathbb{C}^n . We denote the transpose with regards to this basis by $(\cdot)^T$ so that $(x \otimes \text{id}_n)\tau = (\text{id}_n \otimes x^T)\tau$. For any $\omega \in \mathcal{P}(\mathcal{M}_n)$ we define first the positive element $D_\omega \in \mathcal{M}_n$ via the relation $\omega(x) = \text{Tr} D_\omega x$ and then a vector $\omega = (D_\omega^{1/2} \otimes \text{id})\tau$ implementing ω . The vector ω is commonly called the (standard) purification of the density matrix D_ω when ω is a state. The commutant in this case comprises all matrices acting on the second tensor space.

2.2. Spatial derivative and relative modular operator. Let $\sigma \in \mathcal{P}(\mathcal{M})$ and let π be any $*$ -representation of \mathcal{M} into $\mathcal{B}(\mathcal{H})$. The GNS construction further provides a $*$ -representation π_σ into $\mathcal{B}(\mathcal{H}_\sigma)$ and a vector $\sigma \in \mathcal{H}_\sigma$ implementing σ . Following [OP93, Ch. 4], for every $\rho \in \mathcal{H}_\sigma^*$, we define the map $R^\sigma(\rho) : \mathcal{H}_\sigma \rightarrow \mathcal{H}$ by

$$(1) \quad R^\sigma(\rho) : \pi_\sigma(a)\sigma \mapsto \pi(a)\rho, \quad a \in \mathcal{M}.$$

Note that this does not overly specify the map since $\pi_\sigma(a)\sigma = 0$ implies $\pi(a)\rho = 0$ for any $a \in \mathcal{M}$ when $\rho \lll \sigma$. This map is bounded, i.e. there exists a constant C such that for all $a \in \mathcal{M}$,

$$(2) \quad \|R^\sigma(\rho)\pi_\sigma(a)\sigma\|^2 = \|\pi(a)\rho\|^2 = \rho(a^\dagger a) \leq C \sigma(a^\dagger a) = C \|\pi_\sigma(a)\sigma\|^2,$$

where the norm is the one induced by the scalar product on \mathcal{H} . Furthermore, it is easy to verify that $\pi(b)R^\sigma(\rho) = R^\sigma(\rho)\pi_\sigma(b)$ for any $b \in \mathcal{M}$, and from this we can conclude that the operator $R^\sigma(\rho)R^\sigma(\rho)^\dagger$ lies in the commutant of \mathcal{M} in $\mathcal{B}(\mathcal{H})$.¹ If there exists an element $\sigma \in \mathcal{H}$ which implements σ , it follows that the map $R^\sigma(\rho)$ is a partial isometry.

For every $\omega \in \mathcal{H}$, the functional $\rho \mapsto \langle \omega | R^\sigma(\rho)R^\sigma(\rho)^\dagger \omega \rangle$ constitutes a quadratic form to which we associate a positive self-adjoint operator $\Delta(\omega/\sigma)^{\frac{1}{2}}$ on \mathcal{H} . That is, the spatial derivative is defined on its domain $\mathcal{D}(\Delta(\omega/\sigma)^{\frac{1}{2}}) = \mathcal{H}_\sigma^*$ as

$$(3) \quad \langle \rho | \Delta(\omega/\sigma) \rho \rangle = \langle \Delta(\omega/\sigma)^{\frac{1}{2}} \rho | \Delta(\omega/\sigma)^{\frac{1}{2}} \rho \rangle = \langle \omega | R^\sigma(\rho)R^\sigma(\rho)^\dagger \omega \rangle, \quad \rho \in \mathcal{H}_\sigma^*,$$

and thus it is also defined on \mathcal{H}_σ (since \mathcal{H}_σ^* is dense therein). If σ is faithful and $\mathcal{H} = \mathcal{H}_\sigma$ this is called the relative modular operator and its domain is dense in \mathcal{H} . Functions of the spatial derivative are defined via the functional calculus on its domain. In particular, the operator $\Delta(\omega/\sigma)$ is defined as the square of $\Delta(\omega/\sigma)^{\frac{1}{2}}$. We note however, that the domain of powers of $\Delta(\omega/\sigma)$ may be different than the one of $\Delta(\omega/\sigma)$. An example are solely imaginary powers, as the operator $\Delta(\omega/\sigma)^{it}$ is an isometry for $t \in \mathbb{R}$ and hence its domain is the whole space \mathcal{H} .

If $\rho \ll \sigma$ then the operator $R^\sigma(\rho)$ is embedded into $\mathbb{C}^n \otimes \mathbb{C}^n$ as $\text{id} \otimes (D_\rho^{1/2})^T (D_\sigma^{-1/2})^T$, where the inverse is taken on the support of D_σ in case σ is not faithful. The spatial derivative is represented as $\Delta(\omega/\sigma) = D_\omega \otimes (D_\sigma^{-1})^T$. If σ is faithful this is the relative modular operator.

¹The commutant \mathcal{M}' of a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is itself a von Neumann algebra and is given by all elements in $\mathcal{B}(\mathcal{H})$ which commute with \mathcal{M} , $\mathcal{M}' = \{X \in \mathcal{B}(\mathcal{H}) : [X, Y] = 0 \forall Y \in \mathcal{M}\}$.

Let us also consider the commutative case. Let (X, Σ) be a measurable space and let ρ be a probability measure and σ be a positive measure on the σ -algebra Σ . Assume that both are absolutely continuous with respect to another positive measure τ (we can take, e.g., $\tau = \rho + \sigma$). Then ρ, σ are naturally regarded as positive functionals on the commutative von Neumann algebra $\mathcal{M} = L^\infty(X, \Sigma, \tau)$, the almost-everywhere bounded functions. Performing the GNS construction with respect to τ gives rise to the Hilbert space $L^2(X, \Sigma, \tau)$ on which $L^\infty(X, \Sigma, \tau)$ acts by point-wise multiplication. The Radon-Nikodym derivatives $\frac{d\rho}{d\tau}$ and $\frac{d\sigma}{d\tau}$ of ρ and σ with respect to τ are positive elements in $L^1(X, \Sigma, \tau)$, and hence their square roots are elements of $L^2(X, \Sigma, \tau)$ which we denote with ρ and σ . Considering the action of the map $R^\sigma(\rho)$ in (1) we see that

$$(4) \quad R^\sigma(\rho) = \rho \sigma^{-1} = \left(\frac{d\rho}{d\tau} \right)^{\frac{1}{2}} \left(\frac{d\sigma}{d\tau} \right)^{-\frac{1}{2}},$$

the inverse being defined on the support of σ and put to 0 otherwise. Correspondingly, we find for the spatial derivative $\Delta(\omega/\sigma)$ for $\omega \in L^2(X, \Sigma, \tau)$ the form

$$(5) \quad \Delta(\omega/\sigma) = \omega^2 \sigma^{-2},$$

where again powers of elements in $L^2(X, \Sigma, \tau)$ are defined by the point-wise multiplication of functions.

Our arguments are mostly based on complex interpolation theory applied to products of spatial derivatives. In particular, the monotonicity in α , our extension of Araki-Lieb-Thirring inequality to von Neumann algebras as well as our version of the Riesz-Thorin theorem are applications of the following Lemma. Its proof is deferred to Appendix A. We define the complex strip $S_1 = \{z \in \mathbb{C} : 0 \leq \Re(z) \leq 1\}$.

Lemma 1. *Let \mathcal{H}, \mathcal{K} be two Hilbert spaces, $\mathcal{M} \subset \mathcal{B}(\mathcal{H}), \mathcal{N} \subset \mathcal{B}(\mathcal{K})$ two von Neumann algebras and $V : \mathcal{H} \rightarrow \mathcal{K}$ a bounded mapping. Consider two affine holomorphic functions $g(z) = g_1 z + g_0, h(z) = h_1 z + h_0$, with $g_0, g_1, h_0, h_1 \in \mathbb{R}$ and vectors $\varphi, \rho \in \mathcal{H}, \omega \in \mathcal{K}$ and $\sigma \in \mathcal{P}(\mathcal{M}), \tau \in \mathcal{P}(\mathcal{N})$ such that for $x \in \{0, 1\}$ the statement $\varphi \in \mathcal{D}(\Delta(\rho/\sigma)^{g(x)})$ holds. If the vector-valued function*

$$(6) \quad f : S_1 \rightarrow \mathcal{K}, \quad z \mapsto \Delta(\omega/\tau)^{h(z)} V \Delta(\rho/\sigma)^{g(z)} \varphi$$

satisfies $C_x := \sup_{t \in \mathbb{R}} \|f(x + it)\| < \infty$ for $x \in \{0, 1\}$ (i.e. it is uniformly bounded on the boundaries of S_1), then $f(z)$ is holomorphic in the interior of S_1 and satisfies

$$(7) \quad \|f(\theta)\| \leq C_0^{1-\theta} C_1^\theta, \quad \text{for } 0 \leq \theta \leq 1.$$

3. NON-COMMUTATIVE ARAKI-MASUDA L_p -NORMS

3.1. Definition. We first recall the definition of weighted non-commutative vector valued L_p -spaces due to Araki and Masuda [AM82] and extend it to the case of a non-faithful reference state σ by means of the spatial derivative.

Definition 2. *Let $\sigma \in \mathcal{P}(\mathcal{M})$ and let π be a $*$ -representation of \mathcal{M} in $\mathcal{B}(\mathcal{H})$. We define the following norms on elements of \mathcal{H} . For $2 \leq p \leq \infty$ we define σ -weighted p -norm of $\rho \in \mathcal{H}$ by*

$$(8) \quad \|\rho\|_{p,\sigma} := \sup_{\omega \in \mathcal{H}, \|\omega\|=1} \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p}} \rho \right\|$$

if $\rho \ll \sigma$ (which may be infinite) and $+\infty$ otherwise. For $1 \leq p < 2$ we define it as

$$(9) \quad \|\rho\|_{p,\sigma} := \inf_{\omega \in \mathcal{H}, \|\omega\|=1} \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p}} \rho \right\|$$

where the infimum is understood to run over only those vectors with $\omega \gg \rho$.

For a fixed, faithful $\sigma \in \mathcal{P}(\mathcal{M})$ these quantities constitute norms on \mathcal{G}_σ as shown in [AM82, Thm. 1]. Generally, for $1 \leq p < 2$ the quantity is only a semi-norm. We have $\|\rho\|_{2,\sigma} \leq \|\rho\|$ with equality when $\rho \ll \sigma$. For $p > 2$ the norms are finite when $\rho \ll \ll \sigma$ but they can be infinite in general, even when $\rho \ll \sigma$.

Moreover, we note that since the support projection of the spatial derivative $\Delta(\omega/\sigma)$ is contained in the support of $\pi(P_\sigma)$, we have that $\|\rho\|_{p,\sigma} = \|\pi(P_\sigma)\rho\|_{p,\sigma}$ for $1 \leq p \leq \infty$. Hence in estimating these norms we can safely assume that we are working on the Hilbert space \mathcal{H}_σ , since elements in its complement are projected out.

Let us take a closer look at these expressions in our standard representation. We find

$$(10) \quad \left\| \Delta(\omega/\sigma)^{\frac{1-p}{2p}} \rho \right\|^2 = \left\| D_\omega^{\frac{1}{2}-\frac{1}{p}} D_\rho^{\frac{1}{2}} \otimes \left(D_\sigma^{\frac{1}{p}-\frac{1}{2}} \right)^T \tau \right\|^2 = \text{Tr} D_\omega^{1-\frac{2}{p}} D_\rho^{\frac{1}{2}} D_\sigma^{\frac{2}{p}-1} D_\rho^{\frac{1}{2}},$$

which has been studied in [MLDS⁺13, Def. 5]. Taking the supremum or infimum over ω , respectively, leads to well-known σ -weighted norms,

$$(11) \quad \|\rho\|_{p,\sigma} = \left\| D_\sigma^{\frac{1}{p}-\frac{1}{2}} D_\rho^{\frac{1}{2}} \right\|_p = \left(\text{Tr} \left(D_\rho^{\frac{1}{2}} D_\sigma^{\frac{2}{p}-1} D_\rho^{\frac{1}{2}} \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} = \left(\text{Tr} \left(D_\sigma^{\frac{1}{p}-\frac{1}{2}} D_\rho D_\sigma^{\frac{1}{p}-\frac{1}{2}} \right)^{\frac{p}{2}} \right)^{\frac{1}{p}},$$

where $\|\cdot\|_p$ denotes the Schatten norm of order p . While we can determine the optimizer in the finite-dimensional case, this is not so easily done algebraically. However the variational formula around the expression (10) generalizes to the algebraic setting as seen in (8) and (9), if interpreted as a norm on vectors, i.e. as a norm on the square root of the density matrix $D_\rho^{\frac{1}{2}}$ instead of D_ρ itself.

In the commutative case, the optimization can be performed without any problems following (4) and (5) and we arrive at the following expressions

$$(12) \quad \|\rho\|_{p,\sigma} = \left(\int_X \tau(dx) \left(\frac{d\rho}{d\tau}(x) \right)^{\frac{p}{2}} \left(\frac{d\sigma}{d\tau}(x) \right)^{1-\frac{p}{2}} \right)^{\frac{2}{p}},$$

where we assumed that $\rho \ll \sigma$ if $p \geq 2$.

The following lemma shows that this definition in fact is independent of the choice of $*$ -representation. In particular, we can interpret Definition 2 as a norm on positive normal functionals on the W^* -algebra instead of a norm on the vectors implementing these functionals.

Lemma 3. *Let $\rho, \sigma \in \mathcal{P}(\mathcal{M})$, and let $\pi : \mathcal{M} \rightarrow \mathcal{B}(\mathcal{H})$ and $\tilde{\pi} : \mathcal{M} \rightarrow \mathcal{B}(\tilde{\mathcal{H}})$ be two $*$ -representations with vectors $\rho \in \mathcal{H}$ and $\tilde{\rho} \in \tilde{\mathcal{H}}$ both implementing ρ . Then, we have $\|\rho\|_{p,\sigma} = \|\tilde{\rho}\|_{p,\sigma}$ for all $1 \leq p \leq \infty$.*

Proof. The fact that both $*$ -representations allow for a vector implementing ρ means that the mapping $V : \mathcal{H} \rightarrow \tilde{\mathcal{H}}$ defined by

$$(13) \quad V : V\pi(a)\rho \mapsto \tilde{\pi}(a)\tilde{\rho}, \quad a \in \mathcal{M},$$

is an isometry satisfying $V\pi(x) = \tilde{\pi}(x)V$. From (1) follows that $R^\sigma(\tilde{\rho}) = R^\sigma(V\rho) = VR^\sigma(\rho)$ and hence we find $V^\dagger \Delta(\tilde{\omega}/\sigma)V = \Delta(V^\dagger \tilde{\omega}/\sigma)$ for any $\tilde{\omega} \in \tilde{\mathcal{H}}$. For $p \geq 2$ we have

$$(14) \quad \left\| \Delta(\tilde{\omega}/\sigma)^{\frac{1}{2}-\frac{1}{p}} \tilde{\rho} \right\|^2 = \langle \rho | V^\dagger \Delta(\tilde{\omega}/\sigma)^{1-\frac{2}{p}} V \rho \rangle$$

$$(15) \quad \leq \left\langle \rho \left| \left(V^\dagger \Delta(\tilde{\omega}/\sigma)V \right)^{1-\frac{2}{p}} \rho \right. \right\rangle = \left\| \Delta(V^\dagger \tilde{\omega}/\sigma)^{\frac{1}{2}-\frac{1}{p}} \rho \right\|^2,$$

where the sole inequality follows from Jensen's inequality [HP03] and the operator concavity of $t \mapsto t^r$ for $r \in [0, 1]$. Taking the suprema over $\tilde{\omega}$ yields the inequality $\|\tilde{\rho}\|_{p,\sigma} \leq \|\rho\|_{p,\sigma}$ and equality follows because ρ and $\tilde{\rho}$ are interchangeable in the

above argument. For $0 \leq p < 2$ a similar argument using the operator convexity of $t \mapsto t^r$ for $r \in [-1, 0)$ yields the desired result. \square

This allows us to introduce the notation $\|\rho\|_{p,\sigma}$ for any $\rho \in \mathcal{P}(\mathcal{M})$, which refers to the norm of an arbitrary implementation of ρ . If $\rho \ll \sigma$ the GNS space \mathcal{H}_σ allows for implementations of both ρ and σ and is thus a natural choice.

In the finite-dimensional case Lemma 3 simply reaffirms our freedom to choose a purification. Consider a pure state ρ and the trivial $*$ -representation $\tilde{\pi} : x \mapsto x$ of \mathcal{M}_n into itself where ρ is implemented by the vector satisfying $\tilde{\rho}\tilde{\rho}^\dagger = D_\rho$. In this representation we have $R^\sigma(\tilde{\rho})R^\sigma(\tilde{\rho})^\dagger = \text{id}_n \rho(D_\sigma^{-1})$ and $\Delta(\tilde{\omega}/\sigma) = D_\sigma^{-1}\omega(\text{id}_n)$. The optimization turns trivial in this case and the norm evaluates to

$$(16) \quad \|\rho\|_{p,\sigma} = \|\tilde{\rho}\|_{p,\sigma} = \left\| D_\sigma^{\frac{1}{p}-\frac{1}{2}} \tilde{\rho} \right\| = \rho \left(D_\sigma^{\frac{2}{p}-1} \right)^{\frac{1}{2}},$$

which is in agreement with the expression for $\|\rho\|_{p,\sigma}$ in (11) specialized for pure states.

3.2. Norm duality. Araki and Masuda [AM82, Thm. 1] show that, for any $\sigma \in \mathcal{P}(\mathcal{M})$ and two Hölder conjugates $p, q \geq 1$ with $p^{-1} + q^{-1} = 1$, the corresponding L_p - and L_q -norms are dual on \mathcal{G}_σ , namely

$$(17) \quad \|\rho\|_{p,\sigma} = \sup\{|\langle \rho | \omega \rangle| : \omega \in \mathcal{G}_\sigma, \|\omega\|_{q,\sigma} \leq 1\}.$$

This constitutes a Hölder inequality for the inner product:

$$(18) \quad |\langle \rho | \omega \rangle| \leq \|\rho\|_{p,\sigma} \|\omega\|_{q,\sigma}.$$

These norm duality statements continues to hold on \mathcal{H}_σ , with very minor changes to the original proof.

3.3. Norm interpolation and convexity. Building on Araki and Masuda's techniques, we derive the following inequality relating the L_p -norm of $\rho \in \mathcal{P}(\mathcal{M})$ for different values of p (recall that the norm only depends on the state, not on the exact vector implementing the state).

Proposition 4. *Let $\rho, \sigma \in \mathcal{P}(\mathcal{M})$ and either $p_0, p_1 \geq 2$ or $1 \leq p_0, p_1 \leq 2$ be given. Then, we have*

$$(19) \quad \|\rho\|_{p_\theta,\sigma} \leq \|\rho\|_{p_0,\sigma}^{1-\theta} \|\rho\|_{p_1,\sigma}^\theta, \quad \text{for } \frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Proof. We use the natural implementation ρ of ρ in \mathcal{H}_ρ . If the right hand-side is infinite the claim holds trivially and hence we assume that $\|\rho\|_{p_0,\sigma} < \infty$ as well as $\|\rho\|_{p_1,\sigma} < \infty$. Furthermore, we assume without loss of generality that $p_1 \geq p_0$.

Let us first consider the case $p_0, p_1 \geq 2$. The strategy is to apply Lemma 1 with the trivial choice $h(z) = 0$ (which reduces the first spatial derivative to a projector which we can safely ignore) and V the identity map on \mathcal{H} . Hence we set

$$(20) \quad f(z) = \Delta(\omega/\sigma)^{g(z)} \rho$$

for $\omega \in \mathcal{H}_\rho$ with $g(z) = \frac{1}{2} - \left(\frac{z}{p_1} + \frac{1-z}{p_0} \right)$. We confirm that ρ is in the domain of the operator $\Delta(\omega/\sigma)^{g(x)}$ for $x \in \{0, 1\}$ by assumption, and estimate

$$(21) \quad \|f(x+it)\| = \left\| \Delta(\omega/\sigma)^{it} \left(\frac{1}{p_1} - \frac{1}{p_0} \right) \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p_x}} \rho \right\| \leq \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p_x}} \rho \right\|,$$

since the imaginary power of $\Delta(\omega/\sigma)$ is a partial isometry. We find

$$(22) \quad \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p_\theta}} \rho \right\| = \|f(\theta)\| \leq \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p_0}} \rho \right\|^{1-\theta} \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{p_1}} \rho \right\|^\theta.$$

Taking the suprema over $\omega \in \mathcal{H}_\rho$ with $\|\omega\| = 1$ we arrive at the assertion.

For $1 \leq p_0, p_1 \leq 2$ we simply use the norm duality in (17) (note that we can assume $\rho \in \mathcal{H}_\sigma$) to establish

$$(23) \quad \|\rho\|_{p_\theta, \sigma} = \sup \{ \langle \rho | \omega \rangle : \omega \in \mathcal{H}_\rho, \|\omega\|_{q_\theta, \sigma} \leq 1 \}$$

$$(24) \quad \leq \sup \{ \langle \rho | \omega \rangle^{1-\theta} \langle \rho | \omega \rangle^\theta : \omega \in \mathcal{H}_\rho, \|\omega\|_{q_0, \sigma}^{1-\theta} \|\omega\|_{q_1, \sigma}^\theta \leq 1 \}$$

$$(25) \quad \leq \sup \{ \|\rho\|_{p_0, \sigma}^{1-\theta} \|\rho\|_{q_0, \sigma}^{1-\theta} \|\rho\|_{p_1, \sigma}^\theta \|\omega\|_{q_1, \sigma}^\theta : \omega \in \mathcal{H}_\rho, \|\omega\|_{q_0, \sigma}^{1-\theta} \|\omega\|_{q_1, \sigma}^\theta \leq 1 \}$$

$$(26) \quad = \|\rho\|_{p_0, \sigma}^{1-\theta} \|\rho\|_{p_1, \sigma}^\theta.$$

Alternatively the statement can also be shown by a variation of the above argument using interpolation theory. \square

The following simple corollary is noteworthy and turns the previous result into a convexity statement.

Corollary 5. *The map $p \mapsto \log \|\rho\|_{p, \sigma}^p$ is convex on $[1, 2]$ and $[2, \infty)$ for any $\rho, \sigma \in \mathcal{P}(\mathcal{M})$.*

Proof. Taking the logarithm of (19) and multiplying with p_θ , we find

$$(27) \quad \log \|\rho\|_{p_\theta, \sigma}^{p_\theta} \leq \frac{p_\theta}{p_0} (1 - \theta) \log \|\rho\|_{p_0, \sigma}^{p_0} + \frac{p_\theta}{p_1} \theta \log \|\rho\|_{p_1, \sigma}^{p_1}.$$

It remains to verify that $\frac{p_\theta}{p_0} (1 - \theta) + \frac{p_\theta}{p_1} \theta = 1$ and $\frac{p_\theta}{p_0} (1 - \theta) p_0 + \frac{p_\theta}{p_1} \theta p_1 = p_\theta$. \square

3.4. Interpolation theory of linear operators. Here we give a version of the Riesz-Thorin theorem for Araki-Masuda L_p -spaces.

Theorem 6. *Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ be two von Neumann algebras, and $\sigma \in \mathcal{P}(\mathcal{M})$, $\tau \in \mathcal{P}(\mathcal{N})$ two positive functionals. For $T : \mathcal{H} \rightarrow \mathcal{K}$ bounded we have*

$$(28) \quad \|T\|_{p_\theta, \sigma \rightarrow q_\theta, \tau} \leq \|T\|_{p_0, \sigma \rightarrow q_0, \tau}^{1-\theta} \|T\|_{p_1, \sigma \rightarrow q_1, \tau}^\theta, \quad \text{where } \|T\|_{p, \sigma \rightarrow q, \tau} := \sup_{\rho \in \mathcal{H}} \frac{\|T\rho\|_{q, \tau}}{\|\rho\|_{p, \sigma}}$$

for $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ with $p_0, p_1 \geq 2$ as well as $\frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$ with $q_0, q_1 \geq 2$.

For finite-dimensional systems, a similar statement has been proven by Beigi [Bei13, Thm. 4] for σ -weighted norms. We note, however, that while Beigi's result is formulated for the density operators, the Araki-Masuda L_p -spaces are Banach spaces for the square roots of the density operators, or equivalently, they define L_p -norms on the purifying vectors.

Proof. If the right hand-side is infinite the claim holds trivially and hence we assume in the following that it is finite. Our strategy is to prove the upper bound

$$(29) \quad \|\Delta(\chi/\tau)^{\frac{1}{2} - \frac{1}{q_\theta}} T\rho\| \leq \|T\|_{p_0, \sigma \rightarrow q_0, \tau}^{1-\theta} \|T\|_{p_1, \sigma \rightarrow q_1, \tau}^\theta,$$

for dense set of vectors ρ with definite L_p -norm $\|\rho\|_{p_\theta, \sigma} \leq 1$ and an arbitrary $\chi \in \mathcal{K}$ with $\|\chi\| \leq 1$. The assertion would then follow by taking the supremum over such ρ and χ . Hence we first invoke Lemma 17 and choose ρ to be of the form

$$(30) \quad \rho = u \Delta(\omega/\sigma)^{\frac{1}{p_\theta}} \sigma,$$

for $\omega \in \mathcal{H}$ with $\|\omega\| = 1$ and $u \in \mathcal{M}' \subset \mathcal{B}(\mathcal{H})$ with $\|u\| \leq 1$. We aim to apply Lemma 1 and hence set

$$(31) \quad f(z) = \Delta(\chi/\tau)^{\frac{1}{2} - \left(\frac{z}{q_1} + \frac{1-z}{q_0}\right)} T u \Delta(\omega/\sigma)^{\frac{z}{p_1} + \frac{1-z}{p_0}} \sigma,$$

which corresponds to the choice $g(z) = \frac{z}{p_1} + \frac{1-z}{p_0}$, and $h(z) = \frac{1}{2} - \left(\frac{z}{q_1} + \frac{1-z}{q_0}\right)$. Applying Lemma 17 gives

$$(32) \quad \left\| u\Delta(\omega/\sigma)^{\frac{1}{p_x} + it\left(\frac{1}{p_1} - \frac{1}{p_0}\right)} \right\|_{p_x, \sigma} \leq \|\omega\|_{p_x}^{\frac{2}{p_x}} = 1$$

which implies $\sigma \in \mathcal{D}(\Delta(\omega/\sigma)^{g(x)})$ for $x \in \{0, 1\}$. Moreover, we estimate

$$(33) \quad \|f(x + it)\| \leq \left\| \Delta(\chi/\tau)^{\frac{1}{2} - \frac{1}{q_x}} Tu\Delta(\omega/\sigma)^{\frac{1}{p_x} + it\left(\frac{1}{p_1} - \frac{1}{p_0}\right)} \sigma \right\|$$

$$(34) \quad \leq \left\| Tu\Delta(\omega/\sigma)^{\frac{1}{p_x} + it\left(\frac{1}{p_1} - \frac{1}{p_0}\right)} \right\|_{q_x, \tau}$$

$$(35) \quad \leq \|T\|_{p_x, \sigma \rightarrow q_x, \tau} \left\| u\Delta(\omega/\sigma)^{\frac{1}{p_x} + it\left(\frac{1}{p_1} - \frac{1}{p_0}\right)} \right\|_{p_x, \sigma}$$

$$(36) \quad \leq \|T\|_{p_x, \sigma \rightarrow q_x, \tau},$$

where we used in the first step that the imaginary power of $\Delta(\chi/\tau)$ is a partial isometry, in the second and third step the definitions of the norms $\|\cdot\|_{q_x, \tau}$ and $\|T\|_{p_x, \sigma \rightarrow q_x, \tau}$, respectively, and the estimate from (32) in the last step. The requirements of Lemma 1 are satisfied and we get the estimate

$$(37) \quad \left\| \Delta(\chi/\tau)^{\frac{1}{2} - \frac{1}{q_\theta}} Tu\Delta(\omega/\sigma)^{\frac{1}{p_\theta}} \sigma \right\| = \|f(\theta)\| \leq \|T\|_{p_0, \sigma \rightarrow q_0, \tau}^{1-\theta} \|T\|_{p_1, \sigma \rightarrow q_1, \tau}^\theta,$$

from which the assertion follows by substituting (30). \square

4. NON-COMMUTATIVE RÉNYI DIVERGENCE

4.1. Definition. We will use the non-commutative L_p -norms to define new relative entropic functionals on states of \mathcal{M} , which turn out to be an algebraic generalization of the sandwiched Rényi divergences [MLDS⁺13, WWY14].

Definition 7. Let $\sigma \in \mathcal{P}(\mathcal{M})$, $\rho \in \mathcal{S}(\mathcal{M})$ and $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$. Then we define the Araki-Masuda divergence of order α as

$$(38) \quad D_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log Q_\alpha(\rho\|\sigma), \quad Q_\alpha(\rho\|\sigma) := \|\rho\|_{2\alpha, \sigma}^{2\alpha}.$$

The quantity $D_\infty(\rho\|\sigma)$ is defined as the corresponding limit.

Recall that Lemma 3 establishes that we are free to chose any $*$ -representation on ρ implementing ρ .

With (11) it is easily seen that in the finite-dimensional case the Araki-Masuda divergences correspond to the sandwiched Rényi divergences [MLDS⁺13, WWY14]:

$$(39) \quad D_\alpha(\rho\|\sigma) = \frac{1}{\alpha - 1} \log \text{Tr} \left(D_\sigma^{\frac{1-\alpha}{2\alpha}} D_\rho D_\sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha.$$

As a first property of this divergence we show that it is continuous and monotone as a function of α .

Lemma 8. Let $\sigma \in \mathcal{P}(\mathcal{M})$, $\rho \in \mathcal{S}(\mathcal{M})$. The function $\alpha \mapsto D_\alpha(\rho\|\sigma)$ is continuous and monotonically increasing on $[\frac{1}{2}, 1) \cup (1, \infty)$.

Proof. For ease of notation we define $\phi(t) := \log Q_{1+t}(\rho\|\sigma)$. Corollary 5 implies that $\phi(t)$ is convex on $[-\frac{1}{2}, 0]$ and $[0, \infty)$, and thus continuous in these intervals' interiors. This implies continuity of $\alpha \mapsto D_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\phi(\alpha-1)$ on $[\frac{1}{2}, 1) \cup (1, \infty)$.

Moreover, we find that $\phi(0) = 0$ since $\|\rho\|_{2,\sigma} = 1$ for any state $\rho \in \mathcal{S}(\mathcal{M})$. A standard argument reveals that $\frac{1}{t}\phi(t)$ is monotonically increasing for any convex function $\phi(t)$ with $\phi(0) = 0$. More precisely, for $-\frac{1}{2} \leq \alpha < \beta < 0$, we have

$$(40) \quad \phi(\beta) \leq \frac{\beta}{\alpha}\phi(\alpha) + \left(1 - \frac{\beta}{\alpha}\right)\phi(0) = \frac{\beta}{\alpha}\phi(\alpha),$$

which yields $\frac{1}{\beta}\phi(\beta) \geq \frac{1}{\alpha}\phi(\alpha)$, and analogous reasoning applies for $0 < \alpha < \beta$. \square

4.2. Some limiting cases. The extrema $\alpha = \frac{1}{2}$ and $\alpha = \infty$ correspond to well-known quantities. The limit $\alpha \rightarrow 1$ is more delicate and will be covered in the next sub-section.

Lemma 9. *Let $\rho \in \mathcal{S}(\mathcal{M})$, $\sigma \in \mathcal{P}(\mathcal{M})$. We have $D_{\frac{1}{2}}(\rho|\sigma) = -\log F(\rho, \sigma)$ where*

$$(41) \quad F(\rho, \sigma) := \sup \{ |\langle \rho | U \sigma \rangle| : U : \mathcal{H}_\sigma \rightarrow \mathcal{H}_\rho \text{ with } \|U\| \leq 1, \\ U\pi_\sigma(x) = \pi_\rho(x)U \text{ for all } x \in \mathcal{M} \}$$

denotes the fidelity as defined by Alberti [Alb83], and

$$(42) \quad D_\infty(\rho|\sigma) = \log \inf \{ C > 0 : \rho(x) \leq C\sigma(x) \text{ for all } x \in \mathcal{M}_+ \}.$$

These relations have essentially already been established in [AM82], although in our setup we prefer to drop the assumption that σ is faithful.

Proof. To show the first identity, (42), first note that for any C such that $\rho(x) \leq C\sigma(x)$ for all $x \in \mathcal{M}_+$, we have

$$(43) \quad \|\rho\|_{\infty,\sigma}^2 = \sup_{\substack{\omega \in \mathcal{H}_\rho \\ \|\omega\|=1}} \|\Delta(\omega/\sigma)^{\frac{1}{2}}\rho\|^2 = \sup_{\substack{\omega \in \mathcal{H}_\rho \\ \|\omega\|=1}} \langle \omega | R^\sigma(\rho)R^\sigma(\rho)^\dagger \omega \rangle = \|R^\sigma(\rho)\|^2,$$

where the latter norm is induced by the vector norms on \mathcal{H} and \mathcal{H}_σ . Hence the inequality $\|\rho\|_{\infty,\sigma}^2 \leq C$ follows from (2) and holds for all such C . On the other hand, for every $x \in \mathcal{M}_+$, we have

$$(44) \quad \rho(x) = \langle \rho | \pi(x)\rho \rangle = \langle R^\sigma(\rho)\sigma | \pi(x)R^\sigma(\rho)\sigma \rangle$$

$$(45) \quad = \langle \pi_\sigma(x)^{\frac{1}{2}}\sigma | R^\sigma(\rho)^\dagger R^\sigma(\rho)\pi_\sigma(x)^{\frac{1}{2}}\sigma \rangle \leq \|R^\sigma(\rho)\|^2 \sigma(x),$$

which implies the other direction.

For the second identity, we note that we have by L_p -space duality, (17) that

$$(46) \quad \|\rho\|_{1,\sigma} = \sup \{ |\langle \rho | \psi \rangle| : \|\psi\|_{\infty,\sigma} \leq 1 \}.$$

Using now the $p = \infty$ case just considered, we see that $\psi \in \mathcal{H}_\rho$ implies $\|R^\sigma(\psi)\| \leq 1$. By definition, we have $R^\sigma(\psi)\pi_\sigma(x) = \pi_\rho(x)R^\sigma(\psi)$ for $x \in \mathcal{M}$ and hence the relation $\|\rho\|_{1,\sigma} \leq F(\rho, \sigma)$ follows since

$$(47) \quad \langle \rho | \psi \rangle = \langle \rho | R^\sigma(\psi)\sigma \rangle.$$

In order to prove ‘ \geq ’ note that for any $\omega \in \mathcal{H}_\rho$, we have

$$(48) \quad |\langle \rho | U\sigma \rangle| = |\langle \Delta(\omega/\sigma)^{-\frac{1}{2}}\rho | \Delta(\omega/\sigma)^{\frac{1}{2}}U\sigma \rangle| \leq \|\Delta(\omega/\sigma)^{-\frac{1}{2}}\rho\| \|\Delta(\omega/\sigma)^{\frac{1}{2}}U\sigma\|$$

by the Cauchy-Schwartz inequality. The second norm simplifies to

$$(49) \quad \|\Delta(\omega/\sigma)^{\frac{1}{2}}U\sigma\| = \langle \omega | R^\sigma(U\sigma)R^\sigma(U\sigma)^\dagger \omega \rangle = \omega(UU^\dagger) \leq 1,$$

which yields $|\langle \rho | U\sigma \rangle| \leq \|\Delta(\omega/\sigma)^{-\frac{1}{2}}\rho\|$. Since this holds for all such U and all $\omega \in \mathcal{H}_\rho$ we are done. \square

Evidently $D_\infty(\rho|\sigma)$ is finite if and only if $\rho \ll \sigma$. By monotonicity it follows that all $D_\alpha(\rho|\sigma)$ are finite if $\rho \ll \sigma$.

4.3. Comparison with Umegaki's and Petz' divergences. The first entropic functional that was generalized to the operator algebraic setting (again in a work of Araki [Ara76]) was the relative entropy as defined by Umegaki [Ume62]. It is also useful to contrast the definition via Araki-Masuda L_p -spaces with a generalization of Rényi divergence proposed by Petz [Pet85] (see also [OP93, Chap. 7]).

The following definitions are in terms of an implementing vector $\rho \in \mathcal{H}$ of a state $\rho \in \mathcal{S}(\mathcal{M})$ on a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$; however, the quantities are independent of the specific choice of $*$ -representation [OP93, p. 80]. In the following we thus choose the GNS Hilbert space \mathcal{G}_ρ .

Definition 10. Let $\sigma \in \mathcal{P}(\mathcal{M})$ and $\rho \in \mathcal{S}(\mathcal{M})$. Umegaki's relative entropy is defined as [Ara76],

$$(50) \quad D(\rho\|\sigma) := \langle \rho | \log \Delta(\rho/\sigma) \rho \rangle, \quad \text{with } \rho \in \mathcal{G}_\rho \text{ implementing } \rho,$$

if $\rho \ll \sigma$ (which may be infinite) and $+\infty$ otherwise. Moreover, for $\alpha \in (0,1)$ we define [Pet85],

$$(51) \quad \bar{Q}_\alpha(\rho\|\sigma) := \langle \rho | \Delta(\rho/\sigma)^{\alpha-1} \rho \rangle, \quad \text{with } \rho \in \mathcal{G}_\rho \text{ implementing } \rho.$$

For $\alpha \in (1,2)$ we use the definition in (51) when $\rho \ll \sigma$ (which may be infinite) and set $\bar{Q}_\alpha(\rho\|\sigma) = +\infty$ otherwise. Finally, Petz' Rényi divergences are given by

$$(52) \quad \bar{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1} \log \bar{Q}_\alpha(\rho\|\sigma).$$

In the finite-dimensional case Umegaki's relative entropy [Ume62] is given as

$$(53) \quad D(\rho\|\sigma) = \text{Tr } D_\rho(\log D_\rho - \log D_\sigma),$$

and the Petz divergences [Pet86] can be written as

$$(54) \quad \bar{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1} \log \text{Tr} \left(D_\rho^\alpha D_\sigma^{1-\alpha} \right).$$

We collect some useful properties of Petz' version of Rényi divergences into the following proposition. Their proof is standard and deferred to Appendix B.

Proposition 11. Let $\sigma \in \mathcal{P}(\mathcal{M})$ and $\rho \in \mathcal{S}(\mathcal{M})$. Then, we have:

- (1) The function $\alpha \mapsto \bar{D}_\alpha(\rho\|\sigma)$ is monotonically increasing on $[0,2]$.
- (2) If $\rho \ll \sigma$ then $\bar{D}_\alpha(\rho\|\sigma)$ is finite for all $\alpha \in [0,2]$.
- (3) We have $\lim_{\alpha \nearrow 1} \bar{D}_\alpha(\rho\|\sigma) = D(\rho\|\sigma)$. Moreover, if $\bar{D}_\alpha(\rho\|\sigma)$ is finite for any $\alpha \in (1,2)$, we also have

$$(55) \quad \lim_{\alpha \searrow 1} \bar{D}_\alpha(\rho\|\sigma) = D(\rho\|\sigma).$$

4.3.1. An algebraic Araki-Lieb-Thirring inequality. The existence of two different divergences generalizing the commutative case obviously begs the question on their relation to each other. In finite dimensions, an ordering is implied by the Araki-Lieb-Thirring (ALT) [LT76, Arago] inequality.

In the finite-dimensional case the ALT inequality [LT76, Arago] for density matrices implies that

$$(56) \quad \|\rho\|_{p,\sigma}^p = \text{Tr} \left(D_\rho^{\frac{1}{2}} D_\sigma^{\frac{2}{p}-1} D_\rho^{\frac{1}{2}} \right)^{\frac{p}{2}} \leq \text{Tr} D_\rho^{\frac{p}{2}} D_\sigma^{1-\frac{p}{2}} = \|\Delta(\rho/\sigma)^{\frac{p}{4}} \sigma\|^2, \quad \text{for } p \geq 2.$$

The inequality holds in the opposite direction for $1 \leq p \leq 2$.

This motivates the following extension of the ALT inequality to the setting of W^* -algebras and non-commutative vector valued L_p -spaces. We note that Kosaki already established a version of the ALT inequality for von Neumann algebras [Kos92],

albeit only for those possessing a semifinite trace. In contrast, our ALT inequality holds for any von Neumann algebra and is formulated in terms of positive functionals.

Theorem 12. *Let \mathcal{M} be a W^* -algebra. For $\rho \in \mathfrak{S}(\mathcal{M})$ and $\sigma \in \mathfrak{P}(\mathcal{M})$ we have that*

$$(57) \quad D_\alpha(\rho\|\sigma) \leq \overline{D}_\alpha(\rho\|\sigma).$$

In fact, if $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is a von Neumann algebra such that \mathcal{H} supports implementations $\rho, \sigma \in \mathcal{H}$ of ρ and σ , respectively, then we have for $p \geq 2$ that

$$(58) \quad \|\rho\|_{p,\sigma}^p \leq \|\Delta(\rho/\sigma)^{\frac{p}{4}}\sigma\|^2.$$

The inequality holds in the opposite direction for $1 \leq p \leq 2$.

We prove the second statement for $p \geq 2$ and postpone the case $1 \leq p \leq 2$ to Appendix B, due to its more complicated nature. The first statement follows because from the independence of the L_p -norm concerning the vector representative of the state (Lemma 3) as well as from the fact that $\rho = \Delta(\rho/\sigma)^{\frac{1}{2}}\sigma$ if there exists an element $\sigma \in \mathcal{H}$ implementing σ .

Proof. If the right hand-side is infinite the claim holds trivially and hence we assume that $\|\Delta(\rho/\sigma)^{\frac{p}{4}}\sigma\|^2 < \infty$, from which it follows that $\sigma \in \mathcal{D}(\Delta(\rho/\sigma)^{\frac{p}{4}})$. We will prove

$$(59) \quad \|\Delta(\omega/\sigma)^{\frac{1}{2}-\frac{1}{p}}\rho\| \leq \|\Delta(\rho/\sigma)^{\frac{p}{4}}\sigma\|^{\frac{2}{p}},$$

for $\omega \in \mathcal{H}$ with $\|\omega\| = 1$, from which the assertion follows by taking the supremum over all such vectors. Our strategy is again based on Lemma 1 applied to the function

$$(60) \quad f(z) = \Delta(\omega/\sigma)^{h(z)} \Delta(\rho/\sigma)^{g(z)} \sigma$$

with $g(z) = \frac{zp}{4}$, and $h(z) = \frac{1-z}{2}$. It follows by assumption that $\sigma \in \mathcal{D}(\Delta(\rho/\sigma)^{g(x)})$ for $x \in \{0, 1\}$. Moreover, we can estimate

$$(61) \quad \|f(x+it)\| = \left\| \Delta(\omega/\sigma)^{h(x+it)} \Delta(\rho/\sigma)^{g(x+it)} \sigma \right\| \leq \left\| \Delta(\omega/\sigma)^{\frac{1-x}{2}} \Delta(\rho/\sigma)^{\frac{(x+it)p}{4}} \sigma \right\|,$$

again using that the imaginary power of a spatial derivative is a partial isometry. For $x = 0$ we find

$$(62) \quad \left\| \Delta(\omega/\sigma)^{\frac{1}{2}} \Delta(\rho/\sigma)^{\frac{itp}{4}} \sigma \right\| \leq \left\| \Delta(\omega/\sigma)^{\frac{1}{2}-\frac{itp}{4}} \Delta(\rho/\sigma)^{\frac{itp}{4}} \sigma \right\| = \left\| \Delta(\rho/\omega)^{\frac{itp}{4}} \omega \right\| \leq 1,$$

where we used Lemma 19 in the second step as well as $\|\omega\| = 1$ in the third step. For $x = 1$ we get

$$(63) \quad \|f(it)\| \leq \left\| P_\sigma \Delta(\rho/\sigma)^{\frac{(1+it)p}{4}} \sigma \right\| \leq \left\| \Delta(\rho/\sigma)^{\frac{p}{4}} \sigma \right\|.$$

The requirements of Lemma 1 are satisfied and we get

$$(64) \quad \left\| \Delta(\omega/\sigma)^{\frac{1-\theta}{2}} \Delta(\rho/\sigma)^{\frac{\theta p}{4}} \sigma \right\| = \|f(\theta)\| \leq \left\| \Delta(\rho/\sigma)^{\frac{p}{4}} \sigma \right\|^\theta,$$

for $0 \leq \theta \leq 1$. Choosing $\theta = \frac{2}{p}$, using that $\Delta(\rho/\sigma)^{\frac{1}{2}}\sigma = \rho$, and taking the suprema over all $\omega \in \mathcal{H}_\rho$ with $\|\omega\| = 1$ implies the claim (59). \square

4.3.2. *Umegaki's relative entropy as the limit $\alpha \rightarrow 1$.* We find that the Araki-Masuda divergence always converges to Umegaki's relative entropy when α approaches 1 from below and that if $\rho \lll \sigma$ the function $\alpha \rightarrow D_\alpha(\rho||\sigma)$ can be continuously extended to the whole range $\alpha \in [\frac{1}{2}, \infty]$.

Theorem 13. *Let \mathcal{M} be a W^* -algebra, $\sigma \in \mathcal{P}(\mathcal{M})$ and $\rho \in \mathcal{S}(\mathcal{M})$. Then, we have*

$$(65) \quad \lim_{\alpha \nearrow 1} D_\alpha(\rho||\sigma) = D(\rho||\sigma).$$

If furthermore $\rho \lll \sigma$ then we also have

$$(66) \quad \lim_{\alpha \searrow 1} D_\alpha(\rho||\sigma) = D(\rho||\sigma).$$

Given the above theorem, it makes sense to define the Araki-Masuda divergence of order one as the limit $D_1(\rho||\sigma) := \lim_{\alpha \nearrow 1} D_\alpha(\rho||\sigma)$.

Proof. We start with the limit from below in (65) and first show the direction ' \geq '. We may assume that $D_\alpha(\rho||\sigma) < \infty$ for $\alpha \in [\frac{1}{2}, 1)$ as otherwise the statement is trivial. We estimate

$$(67) \quad D_1(\rho||\sigma) = \lim_{\alpha \nearrow 1} \frac{1}{\alpha - 1} \log \inf_{\omega \in \mathcal{H}, \|\omega\|=1} \left\| \Delta(\omega/\sigma)^{\frac{1}{2} - \frac{1}{2\alpha}} \rho \right\|^{2\alpha}$$

$$(68) \quad \geq \lim_{\alpha \nearrow 1} \frac{\alpha}{\alpha - 1} \log \langle \rho | \Delta(\rho/\sigma)^{1 - \frac{1}{\alpha}} \rho \rangle$$

$$(69) \quad = \lim_{\alpha \nearrow 1} \frac{\alpha}{\alpha - 1} \log \int_0^\infty t^{1 - \frac{1}{\alpha}} \langle \rho | P(dt) \rho \rangle,$$

with the measure $P(dt)$ from the spectral decomposition of $\Delta(\rho/\sigma)$. For $\beta := 1 - \frac{1}{\alpha}$ and $\mu(dt) := \langle \rho | P(dt) \rho \rangle$ we calculate with de L'Hospital's rule the left limit

$$(70) \quad \lim_{\beta \nearrow 0} \frac{\log \left(\int_0^\infty t^\beta \mu(dt) \right)}{\beta} = \int_0^\infty \log(t) \mu(dt),$$

from which it follows that

$$(71) \quad D_1(\rho||\sigma) \geq \langle \rho | \log \Delta(\rho/\sigma) \rho \rangle = D(\rho||\sigma).$$

To show the direction ' \leq ' of (65) we invoke the ALT inequality in Theorem 12 which states that $D_\alpha(\rho||\sigma) \leq \overline{D}_\alpha(\rho||\sigma)$. The claim then follows from Proposition 11, which establishes that $\lim_{\alpha \nearrow 1} \overline{D}_\alpha(\rho||\sigma) = D(\rho||\sigma)$.

Let us now proceed to the limit from above in (66). The direction ' \geq ' follows analogously to (67)–(71), where by assumption we have that $D_\alpha(\rho||\sigma)$ is finite for all $\alpha > 1$ and calculate the limit $\beta \searrow 0$.

To get the direction ' \leq ' we again invoke the ALT inequality in Theorem 12 and are left to show $\lim_{\alpha \searrow 1} \overline{D}_\alpha(\rho||\sigma) = D(\rho||\sigma)$ which follows from Proposition 11, assertions (3), the requirements of which are satisfied by our assumption $\rho \lll \sigma$, c.f. assertion (2) of Proposition 11. \square

4.4. **The data-processing inequality.** We consider two W^* -algebras \mathcal{M}, \mathcal{N} and completely positive and unital maps $\mathcal{E} : \mathcal{N} \rightarrow \mathcal{M}$. Moreover, we assume that these maps are normal, that is, they have a pre-dual \mathcal{E}_* mapping the set of normal functionals on \mathcal{M} into the set of normal functionals on \mathcal{N} . We call \mathcal{E} a quantum channel from \mathcal{N} to \mathcal{M} . Assuming $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ for some Hilbert space \mathcal{H} , there exists by Stinespring's theorem [Sti55] a Hilbert space \mathcal{K} , a $*$ -representation π of \mathcal{N} on \mathcal{K} , and an isometry $T : \mathcal{H} \rightarrow \mathcal{K}$ such that $\mathcal{E}(a) = T^\dagger \pi(a) T$ for all $a \in \mathcal{N}$. We call the triple (\mathcal{K}, π, T) a Stinespring dilation of \mathcal{E} .

Theorem 14. *Let \mathcal{M}, \mathcal{N} be two W^* -algebras, $\rho \in \mathcal{S}(\mathcal{M})$, $\sigma \in \mathcal{P}(\mathcal{M})$, and $\mathcal{E} : \mathcal{N} \rightarrow \mathcal{M}$ be a quantum channel from \mathcal{N} to \mathcal{M} . Then, we have*

$$(72) \quad D_\alpha(\rho \parallel \sigma) \geq D_\alpha(\mathcal{E}_*(\rho) \parallel \mathcal{E}_*(\sigma)) \quad \text{for all } \alpha \in \left[\frac{1}{2}, 1\right) \cup (1, \infty].$$

In fact, assuming $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, for any Stinespring dilation (\mathcal{K}, π, T) of \mathcal{E} we have for $p \geq 2$ that

$$(73) \quad \|T\rho\|_{p, \mathcal{E}_*(\sigma)} \leq \|\rho\|_{p, \sigma} \quad \text{for all } \rho \in \mathcal{H}.$$

The inequality holds in the opposite direction for $1 \leq p \leq 2$.

The first assertion of Theorem 14 follows immediately from the second by noting that $T\rho$ is an implementing vector of $\mathcal{E}_*(\rho)$ for all $\rho \in \mathcal{S}(\mathcal{M})$ with ρ implementing ρ . In the limit $\alpha \rightarrow 1$ we get Umegaki's relative entropy and the data processing inequality holds as well [LR73, Uhl77]. The case $\alpha = \frac{1}{2}$ corresponds to the data-processing inequality of the fidelity and is due to Alberti [Alb83].

In the finite-dimensional case Beigi [Bei13] used a similar argument to prove the data-processing inequality from Riesz-Thorin for $\alpha > 1$. Data-processing in the complete range $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$ was shown by Frank and Lieb [FL13]. Recently it was noticed that Beigi's proof continues to hold for (not necessarily completely) positive unital maps [MHR15] (recall that we state our results for maps defined in the Heisenberg picture). While our proof also works for $\alpha \in [\frac{1}{2}, 1)$, we rely on the Stinespring dilation in our setup and are thus restricted to completely positive unital maps.

Proof of Theorem 14. First note that the claim in (73) can be equivalently formulated as

$$(74) \quad \|T\|_{p, \sigma \rightarrow p, \tau} \leq 1 \quad \text{for } p \geq 2.$$

In the limiting case $p \rightarrow \infty$ this statement follows trivially from the expression (42) for $\|\cdot\|_{\infty, \sigma}$ and the positivity property of the quantum channel. The case $p = 2$ follows (with equality) from $\|\cdot\|_{2, \sigma} = \|\cdot\|$ and the fact that T is an isometry. The general case for $p \geq 2$ follows by applying Theorem 6 with the parameter choices $p_0 = q_0 = \infty$, $p_1 = q_1 = 2$, $\theta = \frac{2}{p}$, $\sigma = \sigma$, and $\tau = \mathcal{E}_*(\sigma)$, yielding

$$(75) \quad \|T\|_{p, \sigma \rightarrow p, \tau} \leq \|T\|_{\infty, \sigma \rightarrow \infty, \mathcal{E}_*(\sigma)}^{1 - \frac{2}{p}} \|T\|_{2, \sigma \rightarrow 2, \mathcal{E}_*(\sigma)}^{\frac{2}{p}} \leq 1.$$

For the case $1 \leq p \leq 2$, we simply use the norm duality in (17) to establish

$$(76) \quad \|\rho\|_{p, \sigma} = \sup \{ |\langle \psi | \rho \rangle| : \|\psi\|_{q, \sigma} \leq 1 \} = \sup \{ |\langle T\psi | T\rho \rangle| : \|\psi\|_{q, \sigma} \leq 1 \}$$

$$(77) \quad \leq \|T\psi\|_{q, \mathcal{N}_*(\sigma)} \|T\rho\|_{p, \mathcal{N}_*(\sigma)}$$

$$(78) \quad \leq \|T\rho\|_{p, \mathcal{N}_*(\sigma)},$$

for $\frac{1}{q} + \frac{1}{p} = 1$, and where we used $T^\dagger T = \mathbb{I}$ (since T is an isometry) as well as the statement of the lemma for $p \geq 2$. \square

5. APPLICATION TO HYPOTHESIS TESTING

We have used Araki-Masuda weighted non-commutative vector valued L_p -spaces to define an algebraic generalization of the sandwiched Rényi divergences. We have shown various properties of these divergences, including a data-processing inequality and monotonicity in the parameter $\alpha = \frac{p}{2}$. We have also shown that the Araki-Masuda divergences are lower bounds on an earlier non-commutative generalization of Rényi divergence by Petz in the range $\alpha \in [0, 2]$. The latter quantities attain operational meaning in binary hypothesis testing on von Neumann algebras, as shown by Jakšić *et al.* [JOPP12]. Our work elicits the question whether

the Araki-Masuda divergences characterize the strong converse exponent in binary hypothesis testing on von Neumann algebras. Mosonyi and Ogawa [MO15] showed that this is the case in the finite-dimensional setting.

Let us recapitulate the notation and setup in [JOPP12] and the result in [MO15] for the case of binary hypothesis testing between identical product states. (We refer the reader to these papers for a more comprehensive discussion.) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and let $\rho, \tau \in \mathcal{S}(\mathcal{M})$ be represented by vectors $\rho, \tau \in \mathcal{H}$. A (hypothesis) test T is a positive contraction in \mathcal{M} , i.e. $\{T, \text{id} - T\} \subset \mathcal{M}_+$. For any test T , we define the errors of the first and second kind as $\rho(\text{id} - T)$ and $\tau(T)$, respectively. We use the von Neumann tensor product [Tak79, Ch. 4], denoted $\mathcal{M} \overline{\otimes} \mathcal{M}$, to define a sequence of von Neumann algebras $\mathcal{M}_n = \mathcal{M}^{\overline{\otimes} n} \subset \mathcal{B}(\mathcal{H}^{\otimes n})$, and two sequences of states in $\rho_n, \tau_n \in \mathcal{S}(\mathcal{M}_n)$ determined by their implementing vectors $\rho_n = \rho^{\otimes n}$ and $\tau_n = \tau^{\otimes n}$ in $\mathcal{H}^{\otimes n}$.

Now let us consider a sequence of tests $\{T_n\}$ such that the error of the second kind satisfies $\tau_n(T_n) \leq \exp(-nr)$ for some $r > D(\rho\|\tau)$. The strong converse to Stein's lemma (see, e.g. [JOPP12]) tells us that in this case the error of the first kind, $\rho_n(\text{id} - T_n)$, converges to 1 as $n \rightarrow \infty$. In fact, in the finite-dimensional case [ON00] it is known that this convergence is exponential in n . Let us thus define the optimal strong converse exponent as

$$(79) \quad B_e^*(r) := \inf_{\{T_n\}} \left\{ \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \rho_n(T_n) \mid \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \tau_n(T_n) \geq r \right\}$$

Mosonyi and Ogawa [MO15], again in the finite-dimensional case, show that

$$(80) \quad B_e^*(r) = \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - D_\alpha(\rho\|\tau)),$$

yielding an operational interpretation of the sandwiched Rényi divergence for $\alpha \geq 1$. We conjecture that this relation also holds in the general algebraic case with the sandwiched Rényi divergence replaced by the Araki-Masuda divergence.

To support our conjecture we derive a bound in one direction. Following the footsteps of [MO15], it is easy to show using the multiplicativity of L_p -norms under tensor products² and the data-processing inequality (Theorem 14) that

$$(81) \quad D_\alpha(\rho\|\sigma) = \frac{1}{n} D_\alpha(\rho_n\|\tau_n)$$

$$(82) \quad \geq \frac{1}{n} D_\alpha \left(\left(\begin{array}{cc} \rho_n(T_n) & 0 \\ 0 & \rho_n(\text{id} - T_n) \end{array} \right) \parallel \left(\begin{array}{cc} \tau_n(T_n) & 0 \\ 0 & \tau_n(\text{id} - T_n) \end{array} \right) \right)$$

$$(83) \quad \geq \frac{1}{n(\alpha - 1)} \log \rho_n(T_n)^\alpha \tau_n(T_n)^{1-\alpha},$$

for any test T_n and any $\alpha > 1$. This implies after some manipulations that

$$(84) \quad B_e^*(r) \geq \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - D_\alpha(\rho\|\tau)).$$

Since this derivation holds for all Rényi divergences that satisfy additivity (81) and data-processing (82), our conjecture would also imply that the Araki-Masuda divergences are minimal amongst all generalizations of Rényi divergences satisfying these two properties.

²This property essentially follows from the characterization of vectors with definite L_p -norm in Lemma 17.

ACKNOWLEDGMENTS

MB and MT thank the Department of Physics at Ghent University, and MB and VBS thank the School of Physics at University of Sydney for their hospitality while part of this work was done. MB acknowledges funding by the SNSF through a fellowship, funding by the Institute for Quantum Information and Matter (IQIM), an NSF Physics Frontiers Center (NFS Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028), and funding support from the ARO grant for Research on Quantum Algorithms at the IQIM (W911NF-12-1-0521). VBS is supported by the EU through the ERC Qute. MT is funded by an ARC Discovery Early Career Researcher Award (DECRA) fellowship and acknowledges support from the ARC Centre of Excellence for Engineered Quantum Systems (EQUS).

APPENDIX A. COMPLEX POWERS OF UNBOUNDED OPERATORS

The purpose of this appendix is to collect some basic statements concerning matrix elements of complex powers of spatial derivative operators. In particular, we prove some basic analyticity properties needed for our proofs based on interpolation theory. We start with recalling a lemma about complex powers of unbounded operators, which can, e.g., be found in [Tak03, Lem. VI.2.3] and is restated here for the reader's convenience. In order to shorten the notation, we now let for the rest of the appendix the strip S_γ be defined for real $\gamma > 0$ as

$$(85) \quad S_\gamma = \{z \in \mathbb{C} : 0 \leq \Re(z) \leq \gamma\}.$$

Lemma 15. *Let A be a densely defined self-adjoint positive operator on a Hilbert space \mathcal{H} , and let $\psi \in \mathcal{D}(A^\gamma)$, the domain of A^γ . Then the \mathcal{H} -valued function $z \in S_\gamma \mapsto A^z \psi$ is continuous and bounded on the closure of S_γ and holomorphic in the interior of S_γ .*

Although this lemma is stated for a densely defined operator A , it also holds for operators with non-trivial support projection P_A , if the domain $\mathcal{D}(A)$ of A is dense in $P_A \mathcal{H}$, the support of A . This property will be exploited in the proof of the following lemma, which is a slightly generalized form of Lemma 1. Its proof idea is similar to [AM82, Lem. A].

Lemma 16. *Let A and B be two self-adjoint positive operators on a Hilbert space \mathcal{K} and \mathcal{H} such that their domains are dense in their support, a bounded operator $V : \mathcal{H} \rightarrow \mathcal{K}$ be given. Consider two affine holomorphic functions $g(z) = g_1 z + g_0$, $h(z) = h_1 z + h_0$, with $g_0, g_1, h_0, h_1 \in \mathbb{R}$ and a vector $\varphi \in \mathcal{H}$, such that for $x \in \{0, 1\}$ the statement $\varphi \in \mathcal{D}(B^{g(x)})$ holds. If the vector-valued function*

$$(86) \quad f : S_1 \rightarrow \mathcal{K}, \quad z \mapsto A^{h(z)} V B^{g(z)} \varphi$$

is uniformly bounded on the boundaries of the strip S_1 , i.e. we have for $x \in \{0, 1\}$ that

$$(87) \quad C_x := \|f(x + it)\| < \infty$$

uniform in $t \in \mathbb{R}$, then $f(z)$ is holomorphic in the interior of S_1 and satisfies

$$(88) \quad \|f(\theta)\| \leq C_0^{1-\theta} C_1^\theta, \quad \text{for } 0 \leq \theta \leq 1.$$

Proof. We choose an element $\psi \in \mathcal{H}$ such that $\psi \in \mathcal{D}(A^{h(x)})$ for $x \in \{0, 1\}$. Moreover, we have $\varphi \in \mathcal{D}(B^{g(x)})$, $x \in \{0, 1\}$, from which we can infer by Lemma 15 that the \mathcal{K} -valued functions $z \in S_1 \mapsto A^{h(z)} \psi$ and $z \in S_1 \mapsto V B^{g(z)} \varphi$ are holomorphic in the interior of S_1 and bounded on its closure. It follows by Hartog's theorem that the function

$$(89) \quad \tilde{f} : z \in S_1 \mapsto \langle A^{h(z)} \psi | V B^{g(z)} \varphi \rangle$$

is also holomorphic in the interior and bounded continuous on the closure of S_1 . We may then apply Hadamard's three-line theorem and arrive at

$$(90) \quad \left| \sup_{t \in \mathbb{R}} \tilde{f}(\theta + it) \right| \leq \left(\left| \sup_{t \in \mathbb{R}} \tilde{f}(0 + it) \right| \right)^{1-\theta} \left(\left| \sup_{t \in \mathbb{R}} \tilde{f}(1 + it) \right| \right)^\theta$$

for $\theta \in [0, 1]$. The assumption $\|f(x + it)\| < \infty$ implies that the vectors $B^{g(x+it)}\boldsymbol{\varphi}$ are actually elements of $\mathcal{D}(A^{h(x+it)})$, again for $x \in \{0, 1\}$, and hence

$$(91) \quad \tilde{f}(x + it) = \langle A^{h(x+it)}\boldsymbol{\psi} | VB^{g(x+it)}\boldsymbol{\varphi} \rangle = \langle \boldsymbol{\psi} | A^{h(x+it)}VB^{g(x+it)}\boldsymbol{\varphi} \rangle \leq \|\boldsymbol{\psi}\| C_x.$$

This yields the upper bound

$$(92) \quad |\tilde{f}(\theta + it)| \leq \|\boldsymbol{\psi}\| C_0^{1-\theta} C_1^\theta,$$

for all $t \in \mathbb{R}$ and $0 \leq \theta \leq 1$. Due to the uniform bound in terms of the $\|\boldsymbol{\psi}\|$ it then follows that $VB^{g(z)}\boldsymbol{\psi}$ is in the domain of $A^{h(z)}$ for all values $z \in S_1$ and hence we have

$$(93) \quad \langle \boldsymbol{\psi} | f(z) \rangle = \tilde{f}(z).$$

Hence the function $\langle \boldsymbol{\psi} | f(z) \rangle$ is weakly holomorphic and satisfies

$$(94) \quad |\langle \boldsymbol{\psi} | f(\theta + it) \rangle| \leq \|\boldsymbol{\psi}\| C_0^{1-\theta} C_1^\theta,$$

for all $t \in \mathbb{R}$ and $0 \leq \theta \leq 1$ and a dense set of vector $\boldsymbol{\psi}$. The assertion follows. \square

APPENDIX B. PROPERTIES OF PETZ' RÉNYI DIVERGENCE

Here we present some properties of Petz' version of Rényi divergences, which proof is quite standard and thus only sketched.

Proof of Proposition 11. We start with (1), the statement on monotonicity. Applying Lemma 1 as in the proof of Lemma 4, we find in an analogous way that the function $t \mapsto \log \|\Delta(\rho/\sigma)^{\frac{t}{2}}\rho\|$ with $\rho \in \mathfrak{G}_\rho$ is convex on $[-1, 1]$ for any $\sigma \in \mathcal{P}(\mathcal{M})$. We then proceed as in the proof of Lemma 8.

For (2) we note that by the definition of the map $R^\sigma(\rho)$ we have

$$(95) \quad \begin{aligned} \overline{D}_2(\rho|\sigma) &= \log \langle \rho | \Delta(\rho/\sigma)\rho \rangle = \log \left\| R^\sigma(\rho)^\dagger(\rho) \right\|^2 \\ &\leq \log \left\| R^\sigma(\rho)^\dagger \right\|^2 = D_\infty(\rho|\sigma), \end{aligned}$$

and the assertion follows.

Now consider the statements in (3). The limit from below is stated in [OP93, Sec. 7] (or alternatively see [JOPS12, Prop. 5.4(3)]). For the limit from above, let $P(dt)$ denote the spectral measure of the positive self-adjoint operator $\Delta(\rho/\sigma)$, and set $\mu(dt) = \langle \rho | P(dt)\rho \rangle$. Since the Petz Rényi divergences are assumed to be finite in the open interval $(1, 2)$, we may invoke L'Hospital's rule to calculate the right limit

$$(96) \quad \lim_{\alpha \searrow 1} \overline{D}_\alpha(\rho|\sigma) = \lim_{\alpha \searrow 1} \frac{\log \left(\int_0^\infty t^{\alpha-1} \mu(dt) \right)}{\alpha - 1} = \int_0^\infty \log(t) \mu(dt) = D(\rho|\sigma).$$

\square

Next we provide the proof of the ALT inequality for the case $1 \leq p \leq 2$, which is unfortunately less elegant than for $p \geq 2$.

Proof of Theorem 12 for the case $1 \leq p \leq 2$. Let $\omega \in \mathcal{H}$ with $\|\omega\| = 1$ such that $\omega \lll \sigma$ for $\omega \in \mathcal{S}(\mathcal{M})$ implemented by ω . It follows that $\sigma \in \mathcal{D}(\Delta(\omega/\sigma))$ and we also (trivially) have $\sigma \in \mathcal{D}(\Delta(\rho/\sigma)^{\frac{1}{2}})$. By Hartog's theorem and Lemma 15 the function

$$(97) \quad f(z) := \langle \Delta(\omega/\sigma)^{1-\frac{z}{p}} \sigma | \Delta(\rho/\sigma)^{\frac{z}{2}} \sigma \rangle$$

is holomorphic in the interior and bounded on the closure of S_1 . We may then apply Hadamard's three-line theorem and arrive at

$$(98) \quad \left| f\left(\frac{p}{2}\right) \right| \leq \sup_{t \in \mathbb{R}} |f(1+it)|^{\frac{p}{2}} \sup_{t \in \mathbb{R}} |f(it)|^{1-\frac{p}{2}}.$$

We estimate

$$(99) \quad |f(it)| = \left| \langle \Delta(\omega/\sigma)^{1+\frac{it}{p}} \sigma | \Delta(\rho/\sigma)^{\frac{it}{2}} \sigma \rangle \right| = \left| \langle \Delta(\omega/\sigma)^{\frac{1}{2}+\frac{it}{p}} \omega | \Delta(\rho/\sigma)^{\frac{it}{2}} \sigma \rangle \right|$$

$$(100) \quad = \left| \langle \Delta(\omega/\sigma)^{it\left(\frac{1}{p}+\frac{1}{2}\right)} \omega | \Delta(\omega/\sigma)^{\frac{1}{2}+\frac{it}{2}} \Delta(\rho/\sigma)^{\frac{it}{2}} \sigma \rangle \right|$$

$$(101) \quad \leq \|\Delta(\omega/\sigma)^{\frac{1}{2}+\frac{it}{2}} \Delta(\rho/\sigma)^{\frac{it}{2}} \sigma\|,$$

where we used multiple times that the imaginary power of the spatial derivative is an isometry. We may then invoke Lemma 19 for $z = \frac{it}{2}$ and get that

$$(102) \quad |f(it)| \leq \|\Delta(\rho/\omega)^{\frac{it}{2}} \omega\| \leq 1.$$

We are left to estimate the value of the function on the line $1+it$. This yields

$$(103) \quad |f(1+it)| = \left| \langle \Delta(\omega/\sigma)^{1-\frac{1}{p}-\frac{it}{p}} \sigma | \Delta(\rho/\sigma)^{\frac{it}{2}} \rho \rangle \right|$$

$$(104) \quad \leq \|\Delta(\omega/\sigma)^{\frac{1}{q}-\frac{it}{p}} \sigma\|_{q,\sigma} \|\Delta(\rho/\sigma)^{\frac{it}{2}} \rho\|_{p,\sigma}$$

by norm duality on the space \mathcal{H}_σ for q determined by $\frac{1}{p} + \frac{1}{q} = 1$. Invoking Lemma 17 then yields a value of one for the first factor. This is almost what we aimed for, except of the appearance of $\Delta(\rho/\sigma)^{\frac{it}{2}}$ in the second factor. However, let us consider the state on \mathcal{M} induced by this vector:

$$(105) \quad a \in \mathcal{M} \mapsto \langle \Delta(\rho/\sigma)^{\frac{it}{2}} \rho | \pi_\sigma(a) \Delta(\rho/\sigma)^{\frac{it}{2}} \rho \rangle = \langle \rho | \Delta(\rho/\sigma)^{-\frac{it}{2}} \pi_\sigma(a) \Delta(\rho/\sigma)^{\frac{it}{2}} \rho \rangle.$$

Standard facts from the modular theory of von Neumann algebras (see, e.g., [Tak03, Thm VIII.1.2 and Thm. IX.3.8]) now imply that the one-parameter group, called the modular group of the state ρ ,

$$(106) \quad a \in \mathcal{M} \mapsto d_t^\rho(a) := \Delta(\rho/\sigma)^{-\frac{it}{2}} \pi_\rho(a) \Delta(\rho/\sigma)^{\frac{it}{2}}$$

is independent of the state σ and leaves the state ρ in variant. That is, we have $\rho(d_t^\rho(a)) = \rho(a)$. It follows that the vector $\Delta(\rho/\sigma)^{\frac{it}{2}} \rho$ is in fact implementing the state ρ and by Lemma 3 we arrive at the bound

$$(107) \quad \left| \langle \omega | \Delta(\rho/\sigma)^{\frac{p}{4}} \sigma \rangle \right| = \left| \langle \Delta(\omega/\sigma)^{\frac{1}{2}} \sigma | \Delta(\rho/\sigma)^{\frac{p}{4}} \sigma \rangle \right| \leq \|\rho\|_{\frac{p}{2},\sigma}^{\frac{p}{2}}.$$

The assertion follows by taking the supremum over vectors $\omega \in \mathcal{H}$ with $\|\omega\| = 1$ such that $\omega \lll \sigma$, since vectors with this property are dense in \mathcal{H}_σ . \square

APPENDIX C. FACTS ABOUT L_p -SPACES AND SPATIAL DERIVATIVES

Here we collect a few lemmas about Araki-Masuda L_p -spaces, which were first proven in [AM82] and immediately generalize to our setting. We do not repeat the arguments here but point the reader to the relevant sections in Araki and Masuda's paper. We notify the reader that compared to [AM82], the roles of \mathcal{M} and its commutant \mathcal{M}' has to be switched, i.e. our spatial derivative equals the relative modular operator defined with respect to states on the commutant \mathcal{M}' of \mathcal{M} . We start by stating a lemma constructing a dense set of vectors with bounded Araki-Masuda L_p -norm.

Lemma 17. *Let $\sigma \in \mathcal{H}$ be a vector implementing $\sigma \in \mathcal{P}(\mathcal{M})$. Furthermore, let $\omega \in \mathcal{H}$ and let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$. Then, we have*

$$(108) \quad \left\| u \Delta(\omega/\sigma)^{\frac{1}{p}+it} \sigma \right\|_{p,\sigma} \leq \|\omega\|_{\frac{2}{p}},$$

for $1 \leq p \leq \infty$ and all $u \in \mathcal{M}'$ with $\|u\| \leq 1$. Moreover, for $2 \leq p \leq \infty$ there exist for $\rho \in \mathcal{H}$ with $\|\rho\|_{p,\sigma} < \infty$ a unique vector $\omega \in \mathcal{H}$ as well as a partial isometry $u \in \mathcal{M}'$ such that

$$(109) \quad \rho = u \Delta(\omega/\sigma)^{\frac{1}{p}} \sigma \quad \text{as well as} \quad \|\rho\|_{p,\sigma} = \|\omega\|_{\frac{2}{p}}.$$

Proof. This follows for $1 \leq p < \infty$ as in [AM82, Lem 4.1]. The case $p = \infty$ is trivial. \square

The following Lemma provides some simple estimates on the value of the L_p -norms.

Lemma 18. *Let $\rho, \sigma \in \mathcal{P}(\mathcal{M})$. For $2 \leq p \leq \infty$ we have $\|\rho\|_{p,\sigma} \geq \|\rho\| \|\sigma\|_{\frac{2}{p}-1}$ and for $1 \leq p \leq 2$ the inequality reverses.*

Proof. This follows as in [AM82, Lem. 6.1]. \square

The final lemma is often of use in estimating the L_p -norm of a vector.

Lemma 19. *Let $\rho, \sigma \in \mathcal{P}(\mathcal{M})$ and let \mathcal{H} be a Hilbert space supporting vectors ρ and σ implementing ρ and σ , respectively. Then, we have for all $\omega \in \mathcal{H}$ that*

$$(110) \quad \left\| \Delta(\omega/\sigma)^{\frac{1}{2}-z} \Delta(\rho/\sigma)^z \sigma \right\| = \|\Delta(\rho/\omega)^z \omega\|,$$

with $0 \leq \Re(z) \leq \frac{1}{2}$, and where ω is the positive functional on \mathcal{M} implemented by ω .

Proof. This follows as in [AM82, Lem. C.2]. \square

REFERENCES

- [Alb83] P. M. Alberti. A Note on the Transition Probability over C*-Algebras. *Letters in Mathematical Physics* **7**: 25–32 (1983).
- [AM82] H. Araki and T. Masuda. Positive Cones and L_p -Spaces for von Neumann Algebras. *Publications of the Research Institute for Mathematical Sciences, Kyoto University* **18**: 339–411, (1982).
- [Ara76] B. H. Araki. Relative Entropy of States of von Neumann Algebras. *Publications of the Research Institute for Mathematical Sciences, Kyoto University* **11**: 809–833, (1976).
- [Ara90] H. Araki. On an Inequality of Lieb and Thirring. *Letters in Mathematical Physics* **19**(2): 167–170 (1990).
- [Bei13] S. Beigi. Sandwiched Rényi Divergence Satisfies Data Processing Inequality. *Journal of Mathematical Physics* **54**(12): 122202 (2013).
- [BFS16] M. Berta, F. Furrer, and V. B. Scholz. The Smooth Entropy Formalism for von Neumann Algebras. *Journal of Mathematical Physics* **57**(1): 015213 (2016).
- [FL13] R. L. Frank and E. H. Lieb. Monotonicity of a Relative Rényi Entropy. *Journal of Mathematical Physics* **54**(12): 122201 (2013).

- [Haa79] U. Haagerup. L_p -Spaces Associated with an Arbitrary von Neumann Algebra. In *Proc. Colloquium Marseille 1977*, Volume 274 of *Algebres d'Opérateurs et leurs Applications en Physique Mathématique*, pages 175—184, (1979).
- [HJX09] U. Haagerup, M. Junge, and Q. Xu. A Reduction Method for Noncommutative L_p -spaces and Applications. *Transactions of the American Mathematical Society* **362**(04): 2125–2165 (2009).
- [HP03] F. Hansen and G. K. Pedersen. Jensen's Operator Inequality. *Bulletin London Mathematical Society* **35**(4): 553–564, (2003).
- [JOPP12] V. Jaksic, Y. Ogata, Y. Pautrat, and C. A. Pillet. Entropic Fluctuations in Quantum Statistical Mechanics — An Introduction. In *Quantum Theory from Small to Large Scales: Lecture Notes of the Les Houches Summer School*, Volume 95. Oxford University Press (2012).
- [JOPS12] V. Jakšić, Y. Ogata, C.-A. Pillet, and R. Seiringer. Quantum Hypothesis Testing and Non-Equilibrium Statistical Mechanics. *Reviews in Mathematical Physics* **24**(06): 1230002 (2012).
- [Kos92] H. Kosaki. An Inequality of Araki-Lieb-Thirring (Von Neumann Algebra Case). *Proceedings of the American Mathematical Society* **114**(2): 477 (1992).
- [LR73] E. H. Lieb and M. B. Ruskai. Proof of the Strong Subadditivity of Quantum-Mechanical Entropy. *Journal of Mathematical Physics* **14**(12): 1938 (1973).
- [LT76] E. H. Lieb and W. E. Thirring. *Inequalities for the Moments of the Eigenvalues of the Schrödinger Hamiltonian and Their Relation to Sobolev Inequalities*. Princeton University Press (1976).
- [MHR15] A. Müller-Hermes and D. Reeb. Monotonicity of the Quantum Relative Entropy Under Positive Maps. Preprint [arXiv: 1512.06117](https://arxiv.org/abs/1512.06117) (2015).
- [MLDS⁺13] M. Müller-Lennert, F. Dupuis, O. Szechr, S. Fehr, and M. Tomamichel. On Quantum Rényi Entropies: A New Generalization and Some Properties. *Journal of Mathematical Physics* **54**(12): 122203 (2013).
- [MO14] M. Mosonyi and T. Ogawa. Strong Converse Exponent for Classical-Quantum Channel Coding. Preprint [arXiv: 1409.3562](https://arxiv.org/abs/1409.3562) (2014).
- [MO15] M. Mosonyi and T. Ogawa. Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Rényi Relative Entropies. *Communications in Mathematical Physics* **334**(3): 1617–1648 (2015).
- [ON00] T. Ogawa and H. Nagaoka. Strong Converse and Stein's Lemma in Quantum Hypothesis Testing. *IEEE Transactions on Information Theory* **46**(7): 2428–2433 (2000).
- [OP93] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Springer (1993).
- [Pet85] D. Petz. Quasientropies for States of a von Neumann Algebra. *Publications of the Research Institute for Mathematical Sciences* **21**(4): 787–800 (1985).
- [Pet86] D. Petz. Quasi-entropies for Finite Quantum Systems. *Reports on Mathematical Physics* **23**(1): 57–65 (1986).
- [PX03] G. Pisier and Q. Xu. Non-Commutative L_p -Spaces. In *Handbook of the Geometry of Banach Spaces*, Volume 2, chapter 34, pages 1459–1517. Elsevier (2003).
- [Sti55] W. F. Stinespring. Positive Functions On C^* -Algebras. *Proceedings of the American Mathematical Society* **6**: 211–216 (1955).
- [Tak79] M. Takesaki. *Theory of Operator Algebras I*. Springer New York (1979).
- [Tak03] M. Takesaki. *Theory of Operator Algebras II*. Volume 125 of *Encyclopaedia of Mathematical Sciences*, Springer Berlin Heidelberg (2003).
- [Tom12] M. Tomamichel. Smooth entropies: A Tutorial With Focus on Applications in Cryptography, (2012).
Online: <http://www.marcotom.info/talks/QCrypt2012Tutorial.pdf>.
- [Tom16] M. Tomamichel. *Quantum Information Processing with Finite Resources — Mathematical Foundations*. Volume 5 of SpringerBriefs in Mathematical Physics, Springer International Publishing (2016).
- [TWW14] M. Tomamichel, M. M. Wilde, and A. Winter. Strong Converse Rates for Quantum Communication. Preprint [arXiv: 1406.2946](https://arxiv.org/abs/1406.2946) (2014).
- [Uhl77] A. Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics* **54**(1): 21–32 (1977).
- [Ume62] H. Umegaki. Conditional Expectation in an Operator Algebra. *Kodai Math. Sem. Rep.* **14**: 59–85, (1962).
- [WWY14] M. M. Wilde, A. Winter, and D. Yang. Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy. *Communications in Mathematical Physics* **331**(2): 593–622 (2014).

INSTITUTE FOR QUANTUM INFORMATION AND MATTER, CALIFORNIA INSTITUTE OF TECHNOLOGY,
USA

E-mail address: berta@caltech.edu

DEPARTMENT OF PHYSICS, GHENT UNIVERSITY, GHENT, BELGIUM

E-mail address: volkher.scholz@gmail.com

SCHOOL OF PHYSICS, THE UNIVERSITY OF SYDNEY, AUSTRALIA

E-mail address: marcotom.ch@gmail.com