Trace- and improved data processing inequalities for von Neumann algebras

Stefan Hollands[1]*
[2] *ITP, Universität Leipzig and MPI-MiS Leipzig*

February 23, 2021

Abstract

We prove a version of the data-processing inequality for the relative entropy for general von Neumann algebras with an explicit lower bound involving the measured relative entropy. The inequality, which generalizes previous work by Sutter et al. on finite dimensional density matrices, yields a bound how well a quantum state can be recovered after it has been passed through a channel. The natural applications of our results are in quantum field theory where the von Neumann algebras are known to be of type III. Along the way we generalize various multi-trace inequalities to general von Neumann algebras.

1 Introduction

The relative entropy $S(\rho|\sigma) = \mathrm{Tr}(\rho \ln \rho - \rho \ln \sigma)$ is an important operationally defined measure for the distinguishability of two statistical operators ρ, σ. A fundamental property of S is that

$$S(\rho|\sigma) - S(T(\rho)|T(\sigma)) \geqslant 0 \tag{1}$$

for a quantum channel T, i.e. completely positive linear trace preserving map[1]. The above difference represents the loss of distinguishability between σ, ρ if these are passed through the channel T.

An important general question that can be abstracted from concrete settings such as quantum communication or quantum error correction is to what extent the action of a quantum channel can be reversed, i.e. to what extent it may be possible to recover ρ from $T(\rho)$. It was understood already a long time ago by Petz that the question of recoverability is intimately linked to the case of saturation of the data processing inequality (DPI) (1), see e.g. [28]. As was understood by [17] – and has subsequently

*stefan.hollands@uni-leipzig.de

[1]In the body of the paper, we use the slightly different notation $\tilde{T}$ for the action of a channel on a density matrix (Schrödinger picture), while T denotes the dual action (Heisenberg picture) of the channel on the observables.

been generalized in various was by [25, 35, 17, 7, 10, 23, 34, 39] – explicit lower bounds in the DPI or related information theoretic inequalities can provide information how well a channel may be reversed if the inequality is e.g. nearly saturated.

The current best result in this direction appears to be that by Sutter, Berta, and Tomamichel [35]. It provides an explicit recovery channel, such that the recovered state is close to the original state ρ in a suitable information theoretic measure provided the difference in the DPI is also small. The recovery channel $\alpha_{\sigma,T}$ is called "explicit" because it is given by a concrete expression involving only reference state σ and T (not the state ρ that is to be recovered), and always perfectly recovers σ, i.e. $\alpha_{\sigma,T}(T(\sigma)) = \sigma$. In fact, it is closely related – though not precisely equal – to the channel originally proposed by Petz [29, 30, 31, 28].

The above mentioned works (though not [29, 30, 31, 28]) establish their results only for type I von Neumann algebras – in particular [35] assumes a finite-dimensional Hilbert space. While this is well-motivated by applications in quantum computing, there are cases of interest when the algebras are not of this type. A notable example of this are quantum field theoretic applications related to the "quantum null energy condition" (see e.g. [12]) where the algebras are of type III [9, 19]. With this application in mind we proved in [15] a generalization of [25] in the case when the channel T corresponds to an inclusion of general von Neumann algebras. This result has been generalized to arbitrary 2-positive channels T in [16], where the following improved DPI has been demonstrated:

$$S(\rho|\sigma) - S(T(\rho)|T(\sigma)) \geqslant \frac{1-s}{s} \int_{\mathbb{R}} dt \, \beta_0(t) \, D_s(\alpha^t_{\sigma,T}(T(\rho))|\rho). \tag{2}$$

Here, $s \in [1/2, 1)$ and D_s are the so-called "sandwiched Renyi entropies" [27, 40], which for $s = 1/2$ become the negative log squared fidelity. $\beta_0(t)dt$ is a certain explicit probability density and $\alpha^t_{\eta,T}$ is an explicit 1-parameter family of recovery channels that is a disintegration of $\alpha_{\eta,T}$ in the sense $\int dt \, \beta_0(t) \, \alpha^t_{\eta,T} = \alpha_{\eta,T}$. Using convexity of D_s and Jensen's inequality, the bound implies

$$S(\rho|\sigma) - S(T(\rho)|T(\sigma)) \geqslant \frac{1-s}{s} D_s(\alpha_{\sigma,T}(T(\rho))|\rho). \tag{3}$$

A qualitatively similar result has been proved for general von Neumann algebras by Junge and LaRacuente [26]. In their result, the sandwiched Renyi entropies are now replaced by some other information theoretic quantity with an operational meaning. Both [16, 26] lead to the same inequality for $s = 1/2$. For type I algebras and $s = 1/2$ (2) is the result by [25], but the relation for general s is unclear to the author. We also mention recent results by Gao and Wilde [18] of a roughly similar flavor but different emphasis, which apply to von Neumann algebras with a trace though not type III.

In the present paper, we provide a generalization of [35] to arbitrary (sigma-finite) von Neumann algebras. This version of the improved DPI is qualitatively similar to (3). The definition of the recovery channel is in fact identical to that in (3), but we have yet another information theoretic quantity on the right side, namely (thm. 1)

$$S(\rho|\sigma) - S(T(\rho)|T(\sigma)) \geqslant S_{\text{meas}}(\alpha_{\sigma,T}(T(\rho))|\rho). \tag{4}$$

Here, S_{meas} is the "measured relative entropy", defined as the maximum possible value of the relative entropy restricted to a commutative subalgebra. We show below (prop. 1) that for $s = 1/2$, this inequality is sharper than (3) – though not in general the inequality (2) with the integral outside – for all ρ, σ. A conceptual advantage of (4) over both (2) and (3) (and likewise to the inequalities proven in [26]) is that it is saturated in the commutative case, as noted already by [35]. So in this respect (4) is sharp unlike its predecessors.

Our proof technique is similar in several respects to that in [35] and related antecedents such as [25] in that we also use interpolation arguments for L_p-spaces. However, there are also some key differences requiring technical modifications: For instance, the operators $\ln \rho$ or $\ln \sigma$ no longer exist for general von Neumann algebras or the use of ordinary L_p (Schatten)-spaces is prohibited since a general von Neumann algebra does not have a trace. As in our previous papers [15, 16] – referred to as papers I,II – our solution to the first problem is to work entirely with Araki's relative modular operator, the log of which can roughly be viewed as a *difference* between $\ln \rho$ and $\ln \sigma$. Likewise, as in [15, 16], our solution to the second problem is to work with the Araki-Masuda non-commutative L_p-spaces [3] which are very closely related to the sandwiched relative Renyi entropies[2]. For these norms, we require a complex interpolation theory, see lem. 1, which generalizes a result in [15]. This result is then applied to a specially constructed analytic family of vectors and combined with certain cutoff-techiques for appropriately extended domains of analyticity in a similar way as in [15]. However, in [15, 16], such cutoff techniques were needed to control the limit of the Araki-Masuda norms as $p \to 2$, whereas in the present paper, it is the limit $p \to \infty$ which is relevant. The regularization is necessary here to apply the powerful technique of bounded perturbations of normal states of a von Neumann algebra, and a (somewhat modified) version of the Lie-Trotter product formula for von Neumann algebras [6]. These ideas go beyond [15, 16] and also yield various new "trace" inequalities for von Neumann algebras which could be of independent interest.

This paper is organized as follows. In sec. 2 we review some prerequisite notions from the theory of von Neumann algebras. In sec. 3 we establish an interpolation theorem for the Araki-Masuda L_p-norms, which we apply in sec. 4 to obtain generalizations of various known mutli-trace inequalities to von Neumann algebras. In sec. 5 we establish our main result, thm. 1. The definition of the L_p-norms is relegated to the appendix.

2 Von Neumann algebras and modular theory

Let $\mathcal{A} = M_n(\mathbb{C})$. The fundamental representation of this algebra is on $\mathbb{C}^n$, but one can also work in the "standard" Hilbert space ($\mathscr{H} \simeq M_n(\mathbb{C}) \simeq \mathbb{C}^n \otimes \mathbb{C}^n$). Vectors $|\zeta\rangle$ in $\mathscr{H}$ are thus identified with matrices $\zeta \in M_n(\mathbb{C})$. $\mathscr{H} \simeq M_n(\mathbb{C})$ is both a left and right module for $\mathcal{A}$,

$$l(a)\,|\zeta\rangle = |a\zeta\rangle \qquad r(b)\,|\zeta\rangle = |\zeta b\rangle\,, \tag{5}$$

and the inner product on $\mathscr{H}$ is the Hilbert-Schmidt inner product $\langle \zeta_1|\zeta_2\rangle = \mathrm{Tr}(\zeta_1^* \zeta_2)$. A mixed state, represented by a density matrix ω, gives rise to a linear functional on $\mathcal{A}$ by

$$\omega(a) = \mathrm{Tr}(\omega a), \tag{6}$$

[2][26] use a somewhat different approach to L_p spaces to circumvent the absence of a tracial state in the general von Neumann algebra setting. Their approach appears to us less natural for the purposes of this paper.

where the functional and the state is denoted by the same symbol. These linear functionals are alternatively characterized by the property $\omega(a^*a) \geqslant 0, \omega(1) = 1$.

A (σ-finite) von Neumann algebra in standard form $\mathcal{M}$ is an ultra-weakly closed linear subspace of the bounded operators on a Hilbert space $\mathscr{H}$. $\mathcal{M}$ should contain 1, be closed under products and the $*$-operation should have a cyclic and separating vector $|\psi\rangle \in \mathscr{H}$. Cyclic and separating means that $\mathcal{M}|\psi\rangle$ is dense in $\mathscr{H}$ and $m|\psi\rangle = 0$ implies $m = 0$. In the matrix example, ψ should therefore be invertible. The set of ultra-weakly continuous positive linear functionals (thus satisfying $\omega(a^*a) \geqslant 0, \omega(1) = 1$) is called $\mathscr{S}(\mathcal{M})$. For a detailed account of von Neumann algebras see [36].

Associated with a von Neumann algebra in standard form[3] is a convex cone $\mathscr{P}^\natural_\mathcal{M}$ and an anti-linear involution J, called "modular conjugation" leaving this cone invariant. A possible choice of this non-unique "natural cone" for $\mathcal{A} = M_n(\mathbb{C})$ is the subset of positive semi-definite matrices in $\mathscr{H}$, and in this case, $J|\zeta\rangle = |\zeta^*\rangle$. A general property of J which is easily verified in this example is that $J\mathcal{M}J = \mathcal{M}'$, the latter meaning the commutant of $\mathcal{M}$ on $\mathscr{H}$. Given vectors $|\psi\rangle, |\eta\rangle, |\zeta\rangle \in \mathscr{H}$ and $m \in \mathcal{M}$, one defines following Araki [1] (see also app. C of [3] for many more details)

$$S_{\eta,\psi}\left(m|\psi\rangle + (1 - \pi^{\mathcal{M}'}(\psi))|\zeta\rangle\right) = \pi^{\mathcal{M}}(\psi)m^*|\eta\rangle. \tag{7}$$

Here $\pi^{\mathcal{M}}(\psi) \in \mathcal{M}$ is the orthogonal projection onto the closure of the subspace $\mathcal{M}'|\psi\rangle$ and $\pi^{\mathcal{M}'}(\psi) \in \mathcal{M}'$ that onto the closure of $\mathcal{M}|\psi\rangle$. The definition is consistent because $m\pi^{\mathcal{M}}(\psi) = 0$ if $m|\psi\rangle = 0$. One shows that $S_{\eta,\psi}$ is a closable operator and that if $|\psi\rangle \in \mathscr{P}^\natural_\mathcal{M}$, then

$$S_{\eta,\psi} = J\Delta^{1/2}_{\eta,\psi}, \quad S^*_{\eta,\psi}\bar{S}_{\eta,\psi} = \Delta_{\eta,\psi}, \tag{8}$$

One calls the self-adjoint, non-negative operator $\Delta_{\eta,\psi}$ the "relative modular operator". Its support is $\pi^{\mathcal{M}}(\eta)\pi^{\mathcal{M}'}(\psi)$ and complex powers $\Delta^z_{\eta,\psi}$ are understood as 0 on the orthogonal complement of the support. The modular conjugation and relative modular operators of $\mathcal{A} = M_n(\mathbb{C})$ with the above choice of natural cone are:

$$J|\zeta\rangle = |\zeta^*\rangle \qquad \Delta_{\eta,\psi} = l(\omega_\eta)r(\omega_\psi^{-1}), \tag{9}$$

where we invert the density matrix ω_ψ on the range of $\pi^{\mathcal{M}'}(\psi)$ which in the case at hand is the orthogonal projector onto the complement of the null space of ω_ψ.

For a general von Neumann algebra, every positive linear functional $\omega \in \mathscr{S}(\mathcal{M})$ corresponds to one and only one vector $|\xi_\omega\rangle$ in the natural cone $\mathscr{P}^\natural_\mathcal{M}$ such that $\omega(a) = \langle\xi_\omega|a\xi_\omega\rangle$. Vice versa, any vector $|\psi\rangle$ (in the natural cone or not) gives rise to a linear functional

$$\omega_\psi(a) = \langle\psi|a\psi\rangle, \quad \text{for all } a \in \mathcal{A}. \tag{10}$$

For $\mathcal{A} = M_n(\mathbb{C})$, this linear functional is identified with the density matrix $\omega_\psi = \psi\psi^*$ and the natural cone vectors correspond to the unique positive square root of the corresponding density matrix, now thought of as pure states in the standard Hilbert space. So the vector representative of a density matrix ω in the natural cone is $|\xi_\omega\rangle = |\omega^{1/2}\rangle$. An

[3] More precisely, a standard form is actually defined by the combined structure $(\mathcal{M}, \mathscr{H}, \mathscr{P}^\natural_\mathcal{M}, J)$, which can be recovered if we have a cyclic and separating vector.

important fact used implicitly in several places below is that if two linear functionals are close in norm, then the vectors in the natural cone are as well, and vice versa:

$$\|\xi_\psi - \xi_\eta\|^2 \leqslant \|\omega_\eta - \omega_\xi\| \leqslant \|\xi_\psi + \xi_\eta\| \, \|\xi_\psi - \xi_\eta\|, \tag{11}$$

where the norm of a linear functional is $\|\omega\| = \sup\{|\omega(m)| : m \in \mathcal{M}, \|m\| = 1\}$. In the case $\mathcal{A} = M_n(\mathbb{C})$, the latter norm is $\|\omega\| = \operatorname{Tr}|\omega|$, so the above relation expresses the Powers-Störmer inequality between the trace norm and the Hilbert-Schmidt norm.

Let us finish this briefest of introduction to von Neumann algebras by summarizing (again) some of our

Notations and conventions: Calligraphic letters $\mathcal{A}, \mathcal{M}, \ldots$ denote von Neumann algebras, always assumed σ-finite. Calligraphic letters $\mathscr{H}, \mathscr{K}, \ldots$ denote complex Hilbert spaces, always assumed to be separable. $\mathscr{S}(\mathcal{M})$ denotes the set of all ultra-weakly continuous, positive, normalized linear functionals on $\mathcal{M}$ ("states"), which are in one-to-one correspondence with density matrices if $\mathcal{A} = M_n(\mathbb{C})$. $\mathcal{M}_+$ is the subset of all non-negative self-adjoint operators in $\mathcal{M}$ and $\mathcal{M}_{\text{s.a.}}$ the subset of all self-adjoint elements of the von Neumann algebra $\mathcal{M}$. We use the physicist's "ket"-notation $|\psi\rangle$ for vectors in a Hilbert space. The scalar product is written as

$$(|\psi\rangle, |\psi'\rangle)_{\mathscr{H}} =: \langle\psi|\psi'\rangle \tag{12}$$

and is anti-linear in the first entry. The norm of a vector is written simply as $\||\psi\rangle\| =: \|\psi\|$. The action of a linear operator T on a ket is sometimes written as $T|\phi\rangle = |T\phi\rangle$. In this spirit, the norm of a bounded linear operator T on $\mathscr{H}$ is written as $\|T\| = \sup_{|\psi\rangle : \|\psi\|=1} \|T\psi\|$.

3 Interpolation of non-commutative L_p norms

For the algebra $\mathcal{A} = M_n(\mathbb{C})$ the standard Hilbert space $\mathscr{H} \cong M_n(\mathbb{C})$ on which $\mathcal{A}$ acts by left multiplication can be equipped with various norms. We have already mentioned that the 2-norm (or Hilbert-Schmidt norm)

$$\|\zeta\|_2 = (\operatorname{Tr} \zeta\zeta^*)^{1/2}, \tag{13}$$

actually defines the Hilbert space norm on $\mathscr{H}$ (so the subscript "2" is generally omitted). For $p > 0$, one can generalize this to

$$\|\zeta\|_p = [\operatorname{Tr}(\zeta\zeta^*)^{p/2}]^{1/p}. \tag{14}$$

Given a faithful vector $|\psi\rangle \in \mathscr{H}$ with associated linear functional $\omega_\psi(a) = \langle\psi|a\psi\rangle = \operatorname{Tr}(a\omega_\psi)$ (Hilbert Schmidt inner product), one can also define the yet more general norms:

$$\|\zeta\|_{p,\psi} = [\operatorname{Tr}(\zeta\omega_\psi^{2/p-1}\zeta^*)^{p/2}]^{1/p}. \tag{15}$$

The faithful condition is relevant for $p > 2$ as it ensures that ω_ψ is invertible. The generalized L_p-norms $\|\zeta\|_{p,\psi}$ evidently reduce to usual L_p-norms if $\omega_\psi(a) = \operatorname{Tr}(a)/n$ is the tracial state. A general von Neumann algebra $\mathcal{M}$ in standard form need not have such a tracial state, but Araki and Masuda [3] have shown that one can still define the above "non-commuting L_p-norms" for $p \geqslant 1$ using the relative modular operators based on $|\psi\rangle$, see also [23, 24, 8]. Their basic definitions are recalled for convenience in the appendix. The following interpolation result for the Araki-Masuda L_p-norms is one of the main workhorses of this article.

Lemma 1. *Let $|G(z)\rangle$ be a $\mathscr{H}$-valued holomorphic function on the strip $\mathbb{S}_{1/2} = \{0 < \mathrm{Re} z < 1/2\}$ that is uniformly bounded in the closure, and let $|\psi\rangle \in \mathscr{H}$ a state of a σ-finite von Neumann algebra $\mathcal{M}$ in standard form acting on $\mathscr{H}$. For $0 < \theta < 1/2$, $p_0, p_1 \in [1,2]$ or $p_0, p_1 \in [2,\infty]$, let*

$$\frac{1}{p_\theta} = \frac{1-2\theta}{p_0} + \frac{2\theta}{p_1}. \tag{16}$$

Then

$$\begin{aligned} &\ln \|G(\theta)\|_{p_\theta,\psi} \\ &\leqslant \int_{-\infty}^{\infty} dt \left((1-2\theta)\alpha_\theta(t) \ln \|G(it)\|_{p_0,\psi} + (2\theta)\beta_\theta(t) \ln \|G(1/2+it)\|_{p_1,\psi} \right), \end{aligned} \tag{17}$$

where

$$\alpha_\theta(t) = \frac{\sin(2\pi\theta)}{(1-2\theta)(\cosh(2\pi t) - \cos(2\pi\theta))}, \qquad \beta_\theta(t) = \frac{\sin(2\pi\theta)}{2\theta(\cosh(2\pi t) + \cos(2\pi\theta))}. \tag{18}$$

Proof. We may assume $|\psi\rangle \in \mathscr{P}^\natural_{\mathcal{M}}$ by invariance of the L_p-norms. In parts (a1), (a2) of this proof we first apply that $|\psi\rangle$ is faithful in order to apply the results by [3].

(a1) Assume that $p_0, p_1 \in [1,2]$. This part of the proof is taken from paper I and only included for convenience. Denote the dual of a Hölder index p by p', defined so that $1/p + 1/p' = 1$. [3] have shown that the non-commutative $L_p(\mathcal{M},\psi)$-norm of a vector $|\zeta\rangle$ relative to $|\psi\rangle$ can be characterized by (dropping the superscript on the norm)

$$\|\zeta\|_{p,\psi} = \sup\{|\langle\zeta|\zeta'\rangle| : \quad \|\zeta'\|_{p',\psi} \leqslant 1\}. \tag{19}$$

They have furthermore shown ([3], thm. 3) that when $p' \geqslant 2$, any vector $|\zeta'\rangle \in L_{p'}(\mathcal{M},\psi)$ has a unique generalized polar decomposition, i.e. can be written in the form $|\zeta'\rangle = u\Delta^{1/p'}_{\phi,\psi}|\psi\rangle$, where u is a unitary or partial isometry from $\mathcal{M}$. Furthermore, they show that $\|\zeta'\|_{p',\psi} = \|\phi\|^{p'}$. We may thus choose a u and a normalized $|\phi\rangle$, so that

$$\|G(\theta)\|_{p_\theta,\psi} = \langle u\Delta^{1/p'_\theta}_{\phi,\psi}\psi|G(\theta)\rangle, \tag{20}$$

perhaps up to a small error which we can let go zero in the end. Now we define p_θ as in the statement, so that

$$\frac{1}{p'_\theta} = \frac{1-2\theta}{p'_0} + \frac{2\theta}{p'_1}, \tag{21}$$

and we define an auxiliary function $f(z)$ by

$$f(z) = \langle u\Delta^{2\bar{z}/p'_1 + (1-2\bar{z})/p'_0}_{\phi,\psi}\psi|G(z)\rangle, \tag{22}$$

noting that

$$f(\theta) = \|G(\theta)\|_{p_\theta,\psi} \tag{23}$$

by construction. By Tomita-Takesaki-theory, $f(z)$ is holomorphic in $\mathbb{S}_{1/2}$. For the values at the boundary of the strip $\mathbb{S}_{1/2}$, we estimate

$$
\begin{aligned}
|f(it)| &= |\langle u\Delta_{\phi,\psi}^{-2it(1/p_1'-1/p_0')}\Delta_{\phi,\psi}^{1/p_0'}\psi|G(it)\rangle| \\
&\leqslant \|u\Delta_{\phi,\psi}^{-2it(1/p_1'-1/p_0')}\Delta_{\phi,\psi}^{1/p_0'}\psi\|_{p_0',\psi}\|G(it)\|_{p_0,\psi} \\
&\leqslant \|\Delta_{\phi,\psi}^{-2it(1/p_1'-1/p_0')}\Delta_{\phi,\psi}^{1/p_0'}\psi\|_{p_0',\psi}\|G(it)\|_{p_0,\psi} \\
&\leqslant \|\phi\|^{p_0'}\|G(it)\|_{p_0,\psi} \\
&\leqslant \|G(it)\|_{p_0,\psi}.
\end{aligned}
\tag{24}
$$

Here we used the version of Hölder's inequality proved by [3], we used $\|a^*\zeta\|_{p_0',\psi} \leqslant \|a\|\|\zeta\|_{p_0',\psi}$ for any $a \in \mathcal{A}$, see [3], lem. 4.4, and we used $\|\Delta_{\phi,\psi}^{-2it(1/p_1'-1/p_0')}\Delta_{\phi,\psi}^{1/p_0'}\psi\|_{p_0',\psi} \leqslant \|\phi\|^{p_0'}$ which we prove momentarily. A similar chain of inequalities also gives

$$
|f(1/2+it)| \leqslant \|G(1/2+it)\|_{p_1,\psi}. \tag{25}
$$

To prove the remaining claim, let $|\zeta'\rangle = \Delta_{\phi,\psi}^{z}|\psi\rangle$ and $z = 1/p' + 2it$. Then we have, using the variational characterization by [3] of the $L_{p'}(\mathcal{M},\psi)$-norm when $p' \geqslant 2$:

$$
\begin{aligned}
\|\zeta'\|_{p',\psi} &= \sup\{\|\Delta_{\chi,\psi}^{1/2-1/p'}\Delta_{\phi,\psi}^{z}\psi\| : \|\chi\| = 1\} \\
&= \sup\{\|\Delta_{\chi,\psi}^{1/2-1/p'-2it}\Delta_{\phi,\psi}^{1/p'+2it}\psi\| : \|\chi\| = 1\} \\
&= \sup\{\|\Delta_{\chi,\psi}^{1/2-1/p'}(D\chi : D\phi)_{2t}\pi^{\mathcal{M}}(\phi)\Delta_{\phi,\psi}^{1/p'}\psi\| : \|\chi\| = 1\} \\
&\leqslant \sup\{\|\Delta_{\chi,\psi}^{1/2-1/p'}a\Delta_{\phi,\psi}^{1/p'}\psi\| : \|\chi\| = 1, a \in \mathcal{M}, \|a\| = 1\} \\
&\leqslant \sup\{\|a\Delta_{\phi,\psi}^{1/p'}\psi\|_{p',\psi} : a \in \mathcal{M}, \|a\| = 1\}.
\end{aligned}
\tag{26}
$$

Using [3], lem. 4.4, we continue this estimation as

$$
\leqslant \sup_{a\in\mathcal{M},\|a\|=1} \|a\|\|\Delta_{\phi,\psi}^{1/p'}\psi\|_{p',\psi} = \|\phi\|^{p'}, \tag{27}
$$

which gives the desired result.

Next, we use the Hirschman improvement of the Hadamard three lines theorem [21].

Lemma 2. *Let $g(z)$ be holomorphic on the strip $\mathbb{S}_{1/2}$, continuous and uniformly bounded at the boundary of $\mathbb{S}_{1/2}$. Then for $\theta \in (0, 1/2)$,*

$$
\ln|g(\theta)| \leqslant \int_{-\infty}^{\infty} \left(\beta_\theta(t)\ln|g(1/2+it)|^{2\theta} + \alpha_\theta(t)\ln|g(it)|^{1-2\theta}\right) \mathrm{d}t, \tag{28}
$$

where $\alpha_\theta(t), \beta_\theta(t)$ are as in lem. 1.

Applying this to $g = f$ gives the statement of the theorem.

(a2) Now we assume that $p_0, p_1 \in [2, \infty]$. [3] have shown that for any[4] $\zeta_+' \in L_{p'}^+(\mathcal{M},\psi) :=$ $L_{p'}$-closure of $\mathscr{P}_{\mathcal{M}}^{1/(2p')}$, $1 \leqslant p' \leqslant 2$ there is $\phi \in \mathscr{H}$ such that for all $\zeta \in L_p(\mathcal{M},\psi)$ we have

$$
\langle\zeta_+'|\zeta\rangle = \langle\Delta_{\phi,\psi}^{1/2}\psi|\Delta_{\phi,\psi}^{(1/p')-(1/2)}\zeta\rangle \tag{29}
$$

[4]The cone $\mathscr{P}_{\mathcal{M}}^{1/(2p')}$ is defined as the closure of $\Delta_{\psi}^{1/(2p')}\mathcal{M}_+|\psi\rangle$ and its properties are discussed in [3].

and such that $\|\zeta'_+\|_{p',\psi} = \|\phi\|^{2/p}$. Furthermore, by the non-commutative Hölder inequality proven in [3], there exists $\zeta' \in L_{p'_\theta}(\mathcal{M}, \psi)$ such that

$$\|G(\theta)\|_{p_\theta,\psi} = \langle \zeta' | G(\theta) \rangle, \quad \|\zeta'\|_{p'_\theta,\psi} = 1. \tag{30}$$

Thus, since by [3], thm. 3 we may write $\zeta' = u\zeta'_+, u \in \mathcal{M}$ with $u^*u \leqslant 1$ and $\zeta'_+ \in L^+_{p'_\theta}(\mathcal{M}, \psi)$, we have

$$\begin{aligned} \|G(\theta)\|_{p_\theta,\psi} =& \langle \Delta^{1/2}_{\phi,\psi}\psi | \Delta^{1/p'_\theta - 1/2}_{\phi,\psi} u^* G(\theta) \rangle \\ =& \langle \Delta^{1/2}_{\phi,\psi}\psi | \Delta^{(1-2\theta)/p'_0 + (2\theta)/p'_1 - 1/2}_{\phi,\psi} u^* G(\theta) \rangle \end{aligned} \tag{31}$$

and $\|\phi\| = 1$. Similarly to the previous case we now consider the function

$$f(z) = \langle \Delta^{1/2}_{\phi,\psi}\psi | \Delta^{(1-2z)/p'_0 + (2z)/p'_1 - 1/2}_{\phi,\psi} u^* G(z) \rangle, \tag{32}$$

which is holomorphic for $z \in \mathbb{S}_{1/2}$ and uniformly bounded on the closure. For the lower boundary value we calculate

$$\begin{aligned} |f(it)| =& |\langle \Delta^{1/2}_{\phi,\psi}\psi | \Delta^{-2it(1/p'_0 - 1/p'_1)}_{\phi,\psi} \Delta^{1/p'_0 - 1/2}_{\phi,\psi} u^* G(it) \rangle| \\ \leqslant& \|\Delta^{1/2}_{\phi,\psi}\psi\| \, \|\Delta^{1/p'_0 - 1/2}_{\phi,\psi} u^* G(it)\| \\ =& \|\phi\| \, \|\Delta^{1/2 - 1/p_0}_{\phi,\psi} u^* G(it)\| \\ \leqslant& \sup\{\|\Delta^{1/2 - 1/p_0}_{\chi,\psi} u^* G(it)\| : \|\chi\| = 1\} \\ =& \|u^* G(it)\|_{p_0,\psi} \leqslant \|u^*\| \|G(it)\|_{p_0,\psi} = \|G(it)\|_{p_0,\psi} \end{aligned} \tag{33}$$

using in the last line the variational characterization of the L_p-norms and [3], lem. 4.4. A similar chain of inequalities also gives $|f(1/2 + it)| \leqslant \|G(1/2 + it)\|_{p_1,\psi}$. The rest follows from Hirschman's improvement as in the previous case.

(b) In the remaining part of the proof, we remove the faithful condition on the state $|\psi\rangle$. Suppose that ω_ψ is non-faithful. For σ-finite $\mathcal{M}$, there exists some cyclic and separating vector $|\eta\rangle$ for $\mathcal{M}$ and we put

$$\omega_{\psi_\varepsilon} = (1 - \varepsilon) \, \omega_\psi + \varepsilon \, \omega_\eta \tag{34}$$

so that $|\psi_\varepsilon\rangle \in \mathscr{P}^\natural_\mathcal{M}$ is now faithful for $\mathcal{M}$ (and $\mathcal{M}'$). The proof is then completed by the following lemma because we can apply part (a1),(a2) to the faithful state $|\psi_\varepsilon\rangle$ and obtain b) by taking the limit $\varepsilon \to 0$ and using the dominated convergence theorem under the integral.

Lemma 3. *Let $\omega_\psi, \omega_\eta \in \mathscr{S}(\mathcal{M})$, and let $\omega_{\psi_\varepsilon} = (1 - \varepsilon) \, \omega_\psi + \varepsilon \, \omega_\eta$. Then $\lim_{\varepsilon \to 0+} \|\zeta\|_{p,\psi_\varepsilon} = \|\zeta\|_{p,\psi}$ for any $p \geqslant 1$ and $|\zeta\rangle \in \mathscr{H}$.*

Proof. (1) Case $p \geqslant 2$: Clearly $\omega_{\psi_\varepsilon} \geqslant (1 - \varepsilon)\omega_\psi$, from which it follows that $\Delta_{\phi,\psi_\varepsilon} \leqslant (1 - \varepsilon)^{-1} \Delta_{\phi,\psi}$ and thus by Löwner's theorem [20], $\Delta^\alpha_{\phi,\psi_\varepsilon} \leqslant (1 - \varepsilon)^{-\alpha} \Delta^\alpha_{\phi,\psi}$ for $\alpha \in [0, 1]$, so by the variational definition of the L_p norm (appendix):

$$\|\zeta\|_{p,\psi_\varepsilon} \leqslant (1 - \varepsilon)^{(1/p)-(1/2)} \|\zeta\|_{p,\psi} \quad \text{for } p \geqslant 2. \tag{35}$$

Therefore, by choosing $\varepsilon > 0$ sufficiently small, we can achieve that

$$\|\zeta\|_{p,\psi_\varepsilon} - \|\zeta\|_{p,\psi} < \delta \tag{36}$$

for any given $\delta > 0$. To get a similar inequality in the reverse direction, we use the following lemma.

Lemma 4. *Let $\omega_\psi, \omega_\eta, \omega_{\psi_n}, \omega_{\eta_n} \in \mathscr{S}(\mathcal{M})$ be such that $\lim_n \|\omega_\psi - \omega_{\psi_n}\| = 0, \lim_n \|\omega_\eta - \omega_{\eta_n}\| = 0$ and such that $\omega_{\eta_n} \leqslant C\omega_\eta, \omega_\psi \leqslant C\omega_{\psi_n}$ for some $C < \infty$ and all n. Then*

$$\lim_n \|(\Delta_{\eta,\psi}^{\alpha/2} - \Delta_{\eta_n,\psi_n}^{\alpha/2})\zeta\| = 0 \tag{37}$$

for any $\alpha \in [0,1), |\zeta\rangle \in \mathscr{D}(\Delta_{\eta,\psi}^{\alpha/2})$.

Proof. We use the shorthands $\Delta = \Delta_{\eta,\psi}, \Delta_n = \Delta_{\eta_n,\psi_n}$. Without loss of generality $\alpha > 0$. To deal with the powers, we employ the standard formula

$$X^\alpha = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty d\lambda\, \lambda^\alpha \left[\lambda^{-1} - (\lambda + X)^{-1}\right] \tag{38}$$

for $\alpha \in (0,1), X \geqslant 0$. We use this with $X = \Delta^{1/2}$ and $= \Delta_n^{1/2}$ giving us that

$$\begin{aligned} &\|(\Delta^{\alpha/2} - \Delta_n^{\alpha/2})\zeta\| \\ &\leqslant \int_0^\infty d\lambda\, \lambda^{\alpha-1} \left\|\left[(1+\lambda\Delta^{-1/2})^{-1} - (1+\lambda\Delta_n^{-1/2})^{-1}\right]\zeta\right\|. \end{aligned} \tag{39}$$

In the rest of the proof we denote by c any constant depending only on α, C. We split the integration domain into three parts: $(0,\delta), (\delta, L), (L, \infty)$.

(i) Range $(0,\delta)$: In this range, we use

$$\begin{aligned} &\int_0^\delta d\lambda\, \lambda^{\alpha-1} \left\|\left[(1+\lambda\Delta^{-1/2})^{-1} - (1+\lambda\Delta_n^{-1/2})^{-1}\right]\zeta\right\| \\ &= \int_0^\delta d\lambda\, \lambda^{\alpha} \left\|\left[(\lambda+\Delta^{1/2})^{-1} - (\lambda+\Delta_n^{1/2})^{-1}\right]\zeta\right\| \\ &\leqslant \int_0^\delta d\lambda\, \lambda^{\alpha} \left\{\left\|(\lambda+\Delta^{1/2})^{-1}\zeta\right\| + \left\|(\lambda+\Delta_n^{1/2})^{-1}\zeta\right\|\right\} \end{aligned} \tag{40}$$

$$\leqslant 2\|\zeta\| \int_0^\delta d\lambda\, \lambda^{\alpha-1} = c\|\zeta\|\delta^\alpha$$

using that $\Delta, \Delta_n \geqslant 0$.

(ii) Range (δ, L): By [2], II, lem. 4.1,

$$\left\|\left[(\lambda+\Delta^{1/2})^{-1} - (\lambda+\Delta_n^{1/2})^{-1}\right]\zeta\right\| \to 0 \quad \text{as } n \to \infty, \text{ when } \lambda > 0. \tag{41}$$

and the convergence is uniform for λ in the compact set $[\delta, L]$.

(iii) Range (L,∞). The domination assumption gives $\Delta_n \leqslant C^2\Delta$. The function $\mathbb{R}_+ \ni x \mapsto (\lambda + x^{-1/2})^{-2}$ is operator monotone, thus by by Löwner's theorem [20]:

$$\|(1+\lambda\Delta_n^{-1/2})^{-1}\zeta\| = \langle\zeta|(1+\lambda\Delta_n^{-1/2})^{-2}\zeta\rangle^{1/2} \leqslant \langle\zeta|(1+\lambda C^{-1}\Delta^{-1/2})^{-2}\zeta\rangle^{1/2}. \tag{42}$$

and since $C \geqslant 1$ trivially

$$\|(1+\lambda\Delta^{-1/2})^{-1}\zeta\| = \langle\zeta|(1+\lambda\Delta^{-1/2})^{-2}\zeta\rangle^{1/2} \leqslant \langle\zeta|(1+\lambda C^{-1}\Delta^{-1/2})^{-2}\zeta\rangle^{1/2}. \tag{43}$$

Using these inequalities under the integral (39) gives:

$$\begin{aligned}
&\int_L^\infty d\lambda\, \lambda^{\alpha-1} \left\| \left[(1+\lambda\Delta^{-1/2})^{-1} - (1+\lambda\Delta_n^{-1/2})^{-1} \right] \zeta \right\| \\
&\leqslant \int_L^\infty d\lambda\, \lambda^{\alpha-1} \left\{ \|(1+\lambda\Delta^{-1/2})^{-1}\zeta\| + \|(1+\lambda\Delta_n^{-1/2})^{-1}\zeta\| \right\} \\
&\leqslant 2\int_L^\infty d\lambda\, \lambda^{\alpha-1} \langle\zeta|(1+\lambda C^{-1}\Delta^{-1/2})^{-2}\zeta\rangle^{1/2} \\
&\leqslant cL^{-\alpha/2} \left\{ \int_L^\infty d\lambda\, \lambda^{-1+\alpha} \, \langle\zeta|(1+\lambda C^{-1}\Delta^{-1/2})^{-2}\zeta\rangle \right\}^{1/2} \\
&= cL^{-\alpha/2} \left\{ \langle\zeta|f(C\Delta^{1/2})\zeta\rangle \right\}^{1/2} \leqslant cL^{-\alpha/2}\|\Delta^{\alpha/4}\zeta\|,
\end{aligned} \tag{44}$$

uniformly in n. Here we have applied Jensen's inequality to the probability measure $L^\alpha\lambda^{-1-\alpha}d\lambda$ on (L,∞) in the third step. We have also defined/estimated the non-negative function

$$f(x) = \int_L^\infty d\lambda\, \lambda^{-1+\alpha}(1+x^{-1}\lambda)^{-2} \leqslant cx^\alpha. \tag{45}$$

Applying standard subharmonic analysis to the subharmonic function $z \mapsto \ln\|\Delta^{\alpha z/2}\|$ in the strip $0 \leqslant \mathrm{Re} z \leqslant 1$, we have $\|\Delta^{\alpha/4}\zeta\|^2 \leqslant \|\zeta\|\|\Delta^{\alpha/2}\zeta\|$, giving

$$\int_L^\infty d\lambda\, \lambda^{\alpha-1} \left\| \left[(1+\lambda\Delta^{-1/2})^{-1} - (1+\lambda\Delta_n^{-1/2})^{-1} \right] \zeta \right\| \leqslant c(L^{-\alpha}\|\zeta\|\|\Delta^{\alpha/2}\zeta\|)^{1/2}. \tag{46}$$

Now we choose δ, L so small/large that the contributions from (i), (iii), i.e. (40), (46) are $< \varepsilon/3$ each (independently of n) and then n so large that the contribution (ii) from (δ, L) is $< \varepsilon/3$. Then the integral (39) is $< \varepsilon$ by (i), (ii), (iii), and the proof is complete. ☐

We can now complete the proof of lem. 3. We can pick a unit $|\phi\rangle$ such that $\|\zeta\|_{\psi,p} \leqslant \|\Delta_{\phi,\psi}^{(1/2)-(1/p)}\zeta\| + \delta/2$ by the variational definition of the L_p norm for $p \geqslant 2$. Lem. 4 and the triangle inequality shows that there is an $\varepsilon > 0$ such that

$$\begin{aligned}
\|\zeta\|_{p,\psi} &\leqslant \|\Delta_{\phi,\psi}^{(1/2)-(1/p)}\zeta\| + \delta/2 \\
&\leqslant \|\Delta_{\phi,\psi_\varepsilon}^{(1/2)-(1/p)}\zeta\| + \|(\Delta_{\phi,\psi}^{(1/2)-(1/p)} - \Delta_{\phi,\psi_\varepsilon}^{(1/2)-(1/p)})\zeta\| + \delta/2 \\
&\leqslant \sup\{\|\Delta_{\chi,\psi_\varepsilon}^{(1/2)-(1/p)}\zeta\| : |\chi\rangle \in \mathscr{H}, \|\chi\| = 1\} + \delta \\
&= \|\zeta\|_{p,\psi_\varepsilon} + \delta,
\end{aligned} \tag{47}$$

and this together with (36) gives $|\, \|\zeta\|_{p,\psi} - \|\zeta\|_{p,\psi_\varepsilon} \,| < 2\delta$. Since δ is arbitrarily small, the proof of lem. 3 is complete when $p \geqslant 2$.

(2) Case $1 \leqslant p \leqslant 2$: This proof has already appeared in paper I and is only included for convenience. Since by (34) $\omega_{\psi_\varepsilon}/(1-\varepsilon) > \omega_\psi$, it now follows similarly as in part (1) of this proof that

$$\|\zeta\|_{p,\psi} \leqslant (1-\varepsilon)^{(1/p)-(1/2)} \|\zeta\|_{p,\psi_\varepsilon} \quad \text{for } 1 \leqslant p \leqslant 2. \tag{48}$$

The L_p-norms $\|\zeta\|^p_{p,\psi}$ may be considered for fixed $|\zeta\rangle$ as functionals of the state ω_ψ, and as such they are convex. Indeed, let $D_s(\omega'_\zeta|\omega'_\psi)$ be the sandwiched relative Renyi entropy relative between two functionals $\omega'_\zeta, \omega'_\psi$ on $\mathcal{M}'$ induced by vectors $|\zeta\rangle, |\psi\rangle$, related to the L_p-norms by $D_s(\omega'_\zeta|\omega'_\psi) = (s-1)^{-1} \ln \|\zeta\|^{2s}_{2s,\psi}$. The data processing inequality for this quantity (see e.g. [8], thm. 14) in combination with standard arguments as in e.g. [27], proof of prop. 1 implies joint convexity in $\omega'_\zeta, \omega'_\psi$. This gives in combination with (34) that (for $p = 2s$)

$$\|\zeta\|_{p,\psi_\varepsilon} \leqslant (1-\varepsilon) \|\zeta\|_{p,\psi} + \varepsilon \|\zeta\|_{p,\eta} . \tag{49}$$

Combining (48) with (49) implies the statement of lem. 3 in the case $1 \leqslant p \leqslant 2$. □

This completes the proof of lem. 1. □

4 Multi-trace inequalities for von Neumann algebras

As applications of lem. 2 we now prove various inequalities that reduce to "multi-trace inequalities" in the case of finite type I factors. For simplicity, it will be assumed that ω_ψ is a faithful state on the von Neumann algebra $\mathcal{M}$, meaning $\omega_\psi(m^*m) = 0$ implies $m = 0$ for all $m \in \mathcal{M}$.

Corollary 1. *Let $a_1, \dots, a_n \in \mathcal{M}_+$, $r \in (0,1], p \geqslant 2$. Then*

$$\frac{1}{r} \ln \|a_1^r \cdots a_n^r \psi\|_{p/r,\psi} \leqslant \int_{\mathbb{R}} dt\, \beta_{r/2}(t) \ln \|a_1^{1+it} \cdots a_n^{1+it} \psi\|_{p,\psi}. \tag{50}$$

Proof. We choose $p_1 = p, p_0 = \infty, \theta = r/2$ and

$$G(z) = a_1^{2z} \cdots a_n^{2z} |\psi\rangle \tag{51}$$

in lem. 2. Then $\|G(z)\|$ is uniformly bounded on $\mathbb{S}_{1/2}$ and $p_\theta = p/r$. At the lower boundary of the strip:

$$\|G(it)\|_{p_0,\psi} = \|a_1^{2it} \cdots a_n^{2it} \psi\|_{\infty,\psi} = \|a_1^{2it} \cdots a_n^{2it}\| = 1 \tag{52}$$

because a_k^{2it} are unitary operators (using the isomeric identification of $L_\infty(\mathcal{M},\psi) \ni a|\psi\rangle \mapsto a \in \mathcal{M}$ proven in [3].) Thus the term from the lower boundary does not contribute and we obtain the statement. □

Another corollary of a similar nature is:

Corollary 2. *(Araki-Lieb-Thirring inequality) For $r \geqslant 2, |\psi\rangle, |\zeta\rangle \in \mathscr{H}$ there holds*

$$\|\zeta\|^2_{r,\psi} \leqslant \|\Delta^{r/4}_{\zeta,\psi} \psi\|^{4/r}. \tag{53}$$

Proof. A proof for this has already been given in [8], thm. 12, so the only point is to show an alternative proof. We may assume that $\|\Delta_{\zeta,\psi}^{r/4}\psi\| < \infty$, otherwise the statement is trivial. Also, we may assume without loss of generality that $|\zeta\rangle$ is in the natural cone. In lem. 2, we take $G(z) = \Delta_{\zeta,\psi}^{rz/2}\psi$, $p_1 = 2, p_0 = \infty, \theta = 1/r$, so $p_\theta = r$. Then $G(z)$ is holomorphic and uniformly bounded in $\mathbb{S}_{1/2}$, see e.g. lem. 3 of [6].

On the left side of lem. 2 we obtain $\ln \|\Delta_{\zeta,\psi}^{1/2}\psi\|_{r,\psi}^r = \ln \|\zeta\|_{r,\psi}^r$. We compute at the lower boundary of the strip:

$$\|G(it)\|_{p_0,\psi} = \|\Delta_{\zeta,\psi}^{irt/2}\psi\|_{\infty,\psi} = \|\Delta_{\zeta,\psi}^{irt/2}\Delta_{\psi,\psi}^{-irt/2}\psi\|_{\infty,\psi} = \|u(rt/2)\psi\|_{\infty,\psi} = \|u(rt/2)\| = 1. \quad (54)$$

Here $u(t) = \Delta_{\zeta,\psi}^{it}\Delta_{\psi,\psi}^{-it}$ is the Connes cocycle which is a unitary from $\mathcal{M}$ and we used again the isomeric identification of $L_\infty(\mathcal{M},\psi) \ni a|\psi\rangle \mapsto a \in \mathcal{M}$ proven in [3]. Thus the term from the lower boundary does not contribute. At the upper boundary of the strip:

$$\|G(1/2+it)\|_{p_1,\psi} = \|\Delta_{\zeta,\psi}^{irt/2+r/4}\psi\|_{2,\psi} = \|\Delta_{\zeta,\psi}^{r/4}\psi\|, \quad (55)$$

which no longer depends upon t, using that the L_2 norm is equal to the Hilbert space norm [3] and that $\Delta_{\zeta,\psi}^{it}$ is a unitary operator. Since $\int dt\beta_\theta(t) = 1$ we obtain the statement. ☐

Let h be a self-adjoint element of $\mathcal{M}$ and $|\psi\rangle \in \mathscr{H}$ a normalized state vector. Following Araki [4], the non-normalized perturbed state $|\psi^h\rangle$ is defined by the absolutely convergent series

$$|\psi^h\rangle = \sum_{n=0}^{\infty} \int_0^{1/2} ds_1 \dots \int_0^{s_{n-1}} ds_n\, \Delta_\psi^{s_n} h \Delta_\psi^{s_{n-1}-s_n} h \dots \Delta_\psi^{s_1-s_2} h|\psi\rangle, \quad (56)$$

which can also be written as $e^{(\ln\Delta_\psi + h)/2}|\psi\rangle$ [6]. This technique of perturbations has been generalized to semi-bounded – instead of bounded – operators by [14], see also [28], sec. 12. The perturbations, h that would normally be in $\mathcal{M}_{\text{s.a.}}$ are in this framework generalized to so-called "extended-valued upper bounded self-adjoint operators affiliated with $\mathcal{M}$", the space of which is called $\mathcal{M}_{\text{ext}}$. More precisely, $h \in \mathcal{M}_{\text{ext}}$ if

(i) it is a linear, upper semi-continuous map $\mathscr{S}(\mathcal{M}) \ni \sigma \mapsto \sigma(h) \in \mathbb{R} \cup \{\infty\}$, and

(ii) the set $\{\sigma(h) : \sigma \in \mathscr{S}(\mathcal{M})\}$ is bounded from above.

For any "operator" $h \in \mathcal{M}_{\text{ext}}$, one shows that it is consistent to define:

Definition 1. *(see [14], thm. 3.1) If $h \in \mathcal{M}_{\text{ext}}$, the perturbed state σ^h of a normal state $\sigma \in \mathscr{S}(\mathcal{M})$, is given by the unique extremizer of the convex variational problem*

$$c(\sigma,h) = \sup\{\rho(h) - S(\rho|\sigma) : \rho \in \mathscr{S}(\mathcal{M})\} \quad (57)$$

provided the sup is not $-\infty$.

The latter is certainly the case if $h \in \mathcal{M}_{\text{s.a.}}$ is an ordinary self-adjoint element of the von Neumann algebra $\mathcal{M}$, and in this case the above "thermodynamic" definition of the perturbed state is up to normalizations equivalent to Araki's "perturbative" definition (56):

$$c(\sigma,h) = \ln\|\eta^h\|^2, \quad \sigma^h(m) = \langle\eta^h|m|\eta^h\rangle/\|\eta^h\|^2, \quad (58)$$

wherein $|\eta\rangle$ is a vector representer of the state σ, see [14], ex. 3.3. Furthermore, $h \in \mathcal{M}_{\rm ext}$ has the spectral decomposition [14], prop. 2.13 (B)

$$h = \int_{-\infty}^{c} \lambda E_h(d\lambda) - \infty \cdot q. \tag{59}$$

Here, $q \in \mathcal{M}$ is a projector onto the subspace where h is $-\infty$, and the measure $E_h(d\lambda)$ takes values in the projections in $(1-q)\mathcal{M}(1-q)$, so it commutes with q.

Corollary 3. *(Generalized Golden-Thomson inequality) For $h_i \in \mathcal{M}_{\rm ext}, |\psi\rangle \in \mathscr{H}, \|\psi\| = 1$ there holds*

$$\ln \|\psi^{h_1+\dots+h_k}\|^2 \leqslant \int_{\mathbb{R}} dt\, \beta_0(t) \ln \left\{ \| \prod_{j=1}^{k} e^{(1/2+it)h_j} \psi\| \ \| \prod_{j=k}^{1} e^{(1/2-it)h_j} \psi\| \right\}. \tag{60}$$

Proof. Case I). First we assume each $h_j \in \mathcal{M}_{\rm s.a.}$, i.e. it is bounded. We let

$$G(z) = \Delta_\psi^{z/2} e^{zh_1} \dots e^{zh_k} |\psi\rangle. \tag{61}$$

By standard results of Tomita-Takesaki theory, this family of vectors is analytic on $\mathbb{S}_{1/2}$ and uniformly bounded in the norm of $\mathscr{H}$ on the closure, for instance by the maximum of 1 and $\prod_{i=1}^k \|e^{h_i}\|$ using a standard Phragmen-Lindelöf type argument. In lem. 2, we use this with $p_1 = 2, p_0 = \infty, \theta = 1/n, n \in 2\mathbb{N}$, so $p_\theta = n$. At the lower boundary of $\mathbb{S}_{1/2}$, we get $\|G(it)\|_{2,\psi} = 1$ – the L_2-norm is the Hilbert space norm – so this does not contribute. Keeping therefore only the term from the upper boundary, we have

$$\ln \|\Delta_\psi^{1/(2n)} e^{h_1/n} \cdots e^{h_k/n} \psi\|_{\psi,n}^n \leqslant \int_{\mathbb{R}} dt\, \beta_{1/n}(t) \ln \|\Delta_\psi^{1/4} e^{(1/2+it)h_1} \cdots e^{(1/2+it)h_k} \psi\|^2. \tag{62}$$

Now we consider the left side, putting $a_n = e^{h_1/n} \cdots e^{h_k/n}$. By [3], thm. 3 (4), there exists[5] $|\phi_n\rangle \in L_n(\mathscr{H}, \psi) \cap \mathscr{P}_{\mathcal{M}}^{1/(2n)}$ such that

$$\Delta_{\phi_n,\psi}^{1/n} |\psi\rangle = \Delta_\psi^{1/(2n)} a_n |\psi\rangle, \quad \|\phi_n\|^2 = \|\Delta_\psi^{1/(2n)} a_n \psi\|_{\psi,n}^n. \tag{63}$$

It follows that

$$|\phi_n\rangle = J \Delta_{\phi_n,\psi}^{1/2} |\psi\rangle = J(\Delta_\psi^{1/(2n)} a_n \Delta_\psi^{1/(2n)})^{n/2} |\psi\rangle \tag{64}$$

by a straightforward repeated application of [3], lem. 7.7 (2); for the details see e.g. [22], lem. 4.1. Combining (62), (63), (64), we arrive at

$$\ln \|(\Delta_\psi^{1/(2n)} e^{h_1/n} \cdots e^{h_k/n} \Delta_\psi^{1/(2n)})^{n/2} \psi\|^2 \leqslant \int_{\mathbb{R}} dt\, \beta_{1/n}(t) \ln \|\Delta_\psi^{1/4} e^{(1/2+it)h_1} \cdots e^{(1/2+it)h_k} \psi\|^2. \tag{65}$$

We now take the limit $n \to \infty$ on the left side. Araki's version of the Lie-Trotter formula (suitably generalized to k operators $h_1, \dots, h_k$, using that $e^{h_1/n} \cdots e^{h_k/n} = 1 + n^{-1}(h_1 + \dots + h_k) + O(n^{-2})$ where $\|O(n^{-2})\| \leqslant Cn^{-2}$ for all $n > 0$) see [6], rem.s 1 and 2, establishes that

$$s - \lim_n (\Delta_\psi^{1/(2n)} e^{h_1/n} \cdots e^{h_k/n} \Delta_\psi^{1/(2n)})^{n/4} |\psi\rangle = |\psi^{h_1+\dots+h_k}\rangle = e^{(\ln \Delta_\psi + h_1 + \dots + h_k)/2} |\psi\rangle, \tag{66}$$

[5] The cone $\mathscr{P}_{\mathcal{M}}^{1/(2n)}$ is defined as the closure of $\Delta_\psi^{1/(2n)} \mathcal{M}_+ |\psi\rangle$ in $\mathscr{H}$.

so we get

$$\ln \|\psi^{h_1+\cdots+h_k}\|^2 \leqslant \int_{\mathbb{R}} dt\, \beta_0(t)\, \ln \|\Delta_\psi^{1/4} e^{(1/2+it)h_1} \cdots e^{(1/2+it)h_k} \psi\|^2. \tag{67}$$

On the integrand we finally use the following well-known application of the Hadamard three lines theorem ($0 \leqslant \alpha < 1/2, m \in \mathcal{M}$),

$$\|\Delta_\psi^\alpha m\psi\| \leqslant \|\Delta_\psi^{1/2} m\psi\|^{2\alpha} \|m\psi\|^{1-2\alpha} = \|m^*\psi\|^{2\alpha}\|m\psi\|^{1-2\alpha} \tag{68}$$

using that $z \mapsto \ln \|\Delta_\psi^z m\psi\|$ is subharmonic on $\mathbb{S}_{1/2}$. Using this with $\alpha = 1/4, m = e^{(1/2+it)h_1} \cdots e^{(1/2+it)h_k}$ gives the statement of the corollary.

Case II). The proof can be generalized to the case when $h_j \in \mathcal{M}_{\rm ext}$ by reducing to the case I) via an approximation argument: Elements $k \in \mathcal{M}_{\rm ext}$ can be approximated by bounded self-adjoint elements $k_n \in \mathcal{M}_{\rm s.a.}$ by introducing a cutoff in the spectral decomposition (69), as in

$$k_n = \int_{-n}^{c} \lambda E_k(d\lambda) - n \cdot q \quad ; \tag{69}$$

in fact one shows that $|\psi^{k_n}\rangle \to |\psi^k\rangle$ strongly, see [14], prop. 3.15. We perform this cutoff for every h_j obtaining a $h_{j,n}$. Since the desired inequality holds for $h_{j,n}$ by case I), the proof is completed by the fact that $e^{(1/2+it)h_{j,n}} \to e^{(1/2+it)h_j}$ as $n \to \infty$ strongly and uniformly in t (as can be seen by decomposing $\mathscr{H} = q_j\mathscr{H} + (1-q_j)\mathscr{H}$). $\square$

Examples: 1) In the previous corollary we take $k = 1, h_1 = h$. Then the norm in the integrand no longer depends upon t and we can use that $\int dt \beta_0(t) = 1$ to get:

$$\|\psi^h\| \leqslant \|e^{h/2}\psi\|, \tag{70}$$

as shown previously by [6].

2) Finite-dimensional type I algebras. Let $\mathcal{A} = M_n(\mathbb{C})$. We will work in the standard Hilbert space ($\mathscr{H} \simeq M_n(\mathbb{C}) \simeq \mathbb{C}^{n*} \otimes \mathbb{C}^n$) and identify state functionals such as ω_ψ with density matrices via $\omega_\psi(a) = \mathrm{Tr}(a\omega_\psi)$. Vectors $|\zeta\rangle$ in $\mathscr{H}$ are thus identified with matrices $\zeta \in M_n(\mathbb{C})$. We have already mentioned that the $L_p(\mathcal{A}, \psi)$-norms can be computed using the well known correspondence between these norms and the sandwiched relative entropy discussed in [8]: $\|\zeta\|_{p,\psi}^p = \mathrm{Tr}(\zeta\rho_\psi^{2/p-1}\zeta^*)^{p/2}$ where $|\zeta\rangle \in \mathscr{H}$ is identified with a matrix $\zeta \in M_n(\mathbb{C})$ as described. Let a_i be non-negative matrices. The multi-matrix inequality in cor. 1 then reads, when ω_ψ is the normalized tracial state $\omega_\psi(a) = \mathrm{Tr}(a)/n$,

$$\ln \mathrm{Tr}\, |a_1^r \cdots a_k^r|^{p/r} \leqslant \int_{\mathbb{R}} dt\, \beta_{r/2}(t)\, \ln \mathrm{Tr}\, |a_1 a_2^{1+it} \cdots a_{k-1}^{1+it} a_k|^p, \tag{71}$$

which generalzes the Araki-Lieb-Thirring inequality (corresponding to $k = 2$). This has been derived previously in [41, ?], so our result can be seen as a generalization of these results to arbitrary von Neumann algebras. Cor. 2 is another generalization of this inequality which gives nothing new in the present case. Cor. 3 gives the following inequality. Under the above identification of vectors $|\psi\rangle \in \mathscr{H}$ and matrices, the perturbed vector is

$$|\psi^h\rangle = |e^{\ln\psi + h/2}\rangle \tag{72}$$

(assuming $|\psi\rangle$ to be in the natural cone, i.e. self-adjoint and non-negative), and then choosing $|\psi = 1_n/\sqrt{n}\rangle$ as the vector representing the tracial state on $\mathcal{A}$, we have

$$\ln \operatorname{Tr} e^{h_1+\cdots+h_k} \leqslant \int_{\mathbb{R}} dt\, \beta_0(t)\, \ln \operatorname{Tr} |e^{(1/2)h_1} e^{(1/2+it)h_2} \cdots e^{(1/2+it)h_{k-1}} e^{(1/2)h_k}|^2, \tag{73}$$

for any hermitian matrices h_i. This reduces to the Golden-Thomson inequality for $k = 2$,

$$\operatorname{Tr} e^{h_1+h_2} \leqslant \operatorname{Tr}(e^{h_1} e^{h_2}), \tag{74}$$

using that the trace in the integrand no longer depends on t and $\int dt\beta_0(t) = 1$. For arbitrary number of matrices this is due to [35], who also explain the relation with Lieb's triple matrix inequality (for $k = 3$).

5 Improved DPI and recovery channels

5.1 Relative entropy and measured relative entropy

For the von Neumann algebra $\mathcal{A} = M_n(\mathbb{C})$, the relative entropy between two states (density matrices) ω_ψ, ω_η is defined by:

$$S(\omega_\psi|\omega_\eta) = \operatorname{Tr}(\omega_\psi \ln \omega_\psi - \omega_\psi \ln \omega_\eta). \tag{75}$$

This may be expressed in terms of the logarithm of the relative modular operator in (9), and this observation is the basis for Araki's approach [1, 2] to relative entropy for general von Neumann algebras. The main technical difference in the general case is that the individual terms in the above expression such as the von Neumann entropy $-\operatorname{Tr}(\omega_\psi \ln \omega_\psi)$ are usually infinite. Thus form a mathematical viewpoint, the relative- and not the absolute entropy is the primary concept.

Let $(\mathcal{M}, J, \mathscr{P}^\natural_{\mathcal{M}}, \mathscr{H})$ be a von Neumann algebra in standard form acting on a Hilbert space $\mathscr{H}$, with natural cone $\mathscr{P}^\natural_{\mathcal{M}}$ and modular conjugation J. According to [1, 2], if $\pi^{\mathcal{M}}(\eta) \geqslant \pi^{\mathcal{M}}(\psi)$, the relative entropy may be defined in terms of them by[6]

$$S(\psi|\eta) = -\lim_{\alpha\to 0^+} \frac{\langle \xi_\psi | \Delta^\alpha_{\eta,\psi} \xi_\psi\rangle - 1}{\alpha}, \tag{76}$$

otherwise, it is by definition infinite. Here, $|\xi_\psi\rangle$ denotes the unique representer of a vector $|\psi\rangle$ in the natural cone. The relative entropy only depends on the functionals ω_ψ, ω_η on $\mathcal{M}$, but not the choice of vectors $|\psi\rangle, |\eta\rangle$ that define these functionals. We will therefore use interchangeably the notations $S(\psi|\eta) = S(\omega_\psi|\omega_\eta)$. Araki's definition of $S(\omega_\psi|\omega_\eta)$ still satisfies the data processing inequality (1) [37] along with many other properties, see e.g. [28].

For $t \in \mathbb{R}$, the Connes-cocycle $(D\psi : D\eta)_t$ is the isometric operator from $\mathcal{M}$ satisfying

$$(D\psi : D\eta)_t \pi^{\mathcal{M}'}(\psi) = \Delta^{it}_{\psi,\psi} \Delta^{-it}_{\eta,\psi}. \tag{77}$$

[6]The limit exists under this condition but may be equal to $+\infty$.

It only depends on the state functionals ω_ψ, ω_η. In terms of the Connes-cocycle, the relative entropy (76) may also be defined as

$$S(\omega_\psi|\omega_\eta) \equiv S(\psi|\eta) = -i\frac{\mathrm{d}}{\mathrm{d}t}\omega_\psi((D\eta : D\psi)_t)|_{t=0}. \tag{78}$$

The last expression has the advantage that it does not require one to know the vector representative of $|\psi\rangle$ in the natural cone; in particular it shows that S only depends on the state functionals.[7]

Later we will use the following variational expression for the relative entropy [33], prop. 1,

$$S(\psi|\eta) = \sup_{h\in\mathcal{M}_{\text{s.a.}}}\{\omega_\psi(h) - \ln\|\eta^h\|^2\}, \tag{79}$$

with $\mathcal{M}_{\text{s.a.}}$ the set of self-adjoint elements of $\mathcal{M}$. A related variational quantity is the **"measured relative entropy"**, S_{meas}, defined as

$$S_{\text{meas}}(\psi|\eta) = \sup_{h\in\mathcal{M}_{\text{s.a.}}}\{\omega_\psi(h) - \ln\|e^{h/2}\eta\|^2\}. \tag{80}$$

From the Golden-Thomson inequality (70) we find

$$S_{\text{meas}}(\psi|\eta) \leqslant S(\psi|\eta). \tag{81}$$

S_{meas} can also be written in terms of the classical relative entropy $S(\mu|\nu)$ (Kullback-Leibler divergence) of two probability measures

$$S(\mu|\nu) = \int d\mu \ \ln\frac{d\mu}{d\nu} \tag{82}$$

as follows. Let $a \in \mathcal{M}_{\text{s.a.}}$ be a self-adjoint element of $\mathcal{M}$. Then it has a spectral decomposition

$$a = \int \lambda E_a(d\lambda) \tag{83}$$

with an $\mathcal{M}$-valued projection measure $E_a(d\lambda)$. Given $|\psi\rangle, |\eta\rangle \in \mathscr{H}$, we get Borel measures $d\mu_{\psi,a} = \langle\psi|E_a(d\lambda)\psi\rangle$, and likewise for $|\eta\rangle$. Physically, these correspond to the probability distributions for measument outcomes of a in the states $|\psi\rangle$ resp. $|\eta\rangle$. The relative entropy between these measures is defined (but can be $+\infty$) if $\text{supp}\mu_{\eta,a} \subset \text{supp}\mu_{\psi,a}$, wherein $d\mu_{\psi,a}/d\mu_{\eta,a}$ means the Radon-Nikodym derivative between the measures. We may perform the maximization in over $f(h)$ with[8] $f \in L^\infty(\mathbb{R};\mathbb{R})$ and $h \in \mathcal{M}_{\text{s.a.}}$ because $f(h) \in \mathcal{M}_{\text{s.a.}}$. Maximizing first for fixed h over f and using (= eq. (75) in the commutative case)

$$\sup\left\{\int f d\mu - \ln\int e^f d\nu : f \in L^\infty(\mathbb{R};\mathbb{R})\right\} = S(\mu|\nu), \tag{84}$$

[7] The derivative exists whenever $S(\psi|\eta) < \infty$ [28], thm. 5.7.

[8] More precisely, the space L^∞ is defined relative to the measure $\mu_{h,\psi}$ relative to some faithful normal state $\psi \in \mathscr{S}(\mathcal{M})$. Depending on the nature of this measure, "L^∞" means either $\ell^\infty(\{1,\ldots,n\}), \ell^\infty(\mathbb{N})$ or $L^\infty(\mathbb{R})$ or a combination thereof, wherein the counting measure is understood in the first two cases, whereas the Lebesgue measure is understood in the last case.

we can write the measured relative entropy in the following way:

$$S_{\rm meas}(\omega_\psi|\omega_\eta) = \sup\{S(\mu_{h,\psi}|\mu_{h,\eta}) : h \in \mathcal{M}_{\rm s.a.}\} \\ = \sup\{S(\omega_{\psi|\mathcal{C}}|\omega_{\eta|\mathcal{C}}) : \mathcal{C} \subset \mathcal{M} \text{ a commutative von Neumann subalgebra}\}. \tag{85}$$

This motivates the name "measured relative entropy". The second equality holds by [28], prop. 7.13, for a related discussion see also [7], lem. 1 which corresponds to counting measures on the finite set $\{1, \dots, n\}$.

For later we would like to know the relationship between $S_{\rm meas}$ and the fidelity, F. According to [38], the fidelity between two states $\omega_\eta, \omega_\psi \in \mathscr{S}(\mathcal{M})$ on a von Neumann algebra $\mathcal{M}$ in standard form may be defined as

$$F(\omega_\psi|\omega_\eta) = \sup\{|\langle\eta|u'\psi\rangle| : u' \in \mathcal{M}', \|u'\| = 1\}. \tag{86}$$

It is related to the L_1-norm relative to $\mathcal{M}'$ by $F(\omega_\psi|\omega_\eta) = \|\eta\|_{1,\psi,\mathcal{M}'}$, see e.g. paper I, lem. 3 (1). We claim:

Proposition 1. *If $\omega_\eta \in \mathscr{S}(\mathcal{M})$ is a faithful state on the von Neumann algebra $\mathcal{M}$, then $S_{\rm meas}(\omega_\psi|\omega_\eta) \geqslant -\ln F(\omega_\psi|\omega_\eta)^2$.*

Proof. We may assume at that $|\eta\rangle$ is cyclic for $\mathcal{M}$, for if not we can obtain an equivalent standard form of $\mathcal{M}$ after a GNS-construction based on ω_η and work with that standard form. Without loss of generality, $|\eta\rangle \in \mathscr{P}^\sharp_{\mathcal{M}}$. Consider in $L_1(\mathcal{M}', \eta)$ the polar decomposition $|\psi\rangle = u'^*|\psi_+\rangle$ into a $u' \in \mathcal{M}'$ such that $u'^*u' = \pi^{\mathcal{M}'}(\psi) \leqslant 1$ and $|\psi_+\rangle \in \mathscr{P}^{1/2}_{\mathcal{M}'}$, see [3], thm. 3. By definition, the cone $\mathscr{P}^{1/2}_{\mathcal{M}'}$ is the closure of $\Delta'^{1/2}_\psi \mathcal{M}'_+|\eta\rangle$, which equals the closure of $\mathcal{M}_+|\eta\rangle$, since $J\Delta'^{1/2}_\psi a'|\eta\rangle = a'|\eta\rangle$ for $a' \in \mathcal{M}'_+$, $J|\eta\rangle = |\eta\rangle$ and $J\mathcal{M}'J = \mathcal{M}$. Thus, there exists a sequence $\{a_n\} \subset \mathcal{M}_+$ such that $\lim_n a_n|\eta\rangle = u'|\psi\rangle$ strongly, so

$$\lim_n \langle\eta|a_n\eta\rangle = \langle\eta|u'\psi\rangle \in \mathbb{R}_+. \tag{87}$$

Then, with $E_{a_n}(d\lambda)$ the spectral decomposition of a_n and $d\mu_{a_n,\psi} = \langle\psi|E_{a_n}(d\lambda)\psi\rangle$, $d\mu_{a_n,\eta} = \langle\eta|E_{a_n}(d\lambda)\eta\rangle$, the definition of the measured relative entropy and Jensen's inequality applied to the convex function $-\ln$ yields

$$S_{\rm meas}(\omega_\psi|\omega_\eta) \geqslant S(\mu_{a_n,\psi}|\mu_{a_n,\eta}) \geqslant -2\ln \int \left(\frac{d\mu_{a_n,\psi}}{d\mu_{a_n,\eta}}\right)^{1/2} d\mu_{a_n,\eta} = -2\ln F(\mu_{a_n,\psi}|\mu_{a_n,\eta}), \tag{88}$$

where the Radon-Nikodym derivative is defined since $|\eta\rangle$ is faithful. The strong limit $\lim_n a_n|\eta\rangle = u'|\psi\rangle$ and $d\mu_{a_n,\psi} = \langle u'\psi|E_{a_n}(d\lambda)u'\psi\rangle$ (because $u' \in \mathcal{M}', u'^*u' = \pi^{\mathcal{M}'}(\psi)$ and E_{a_n} takes values in $\mathcal{M}$) imply that $\|\mu_{a_n,\psi} - \mu_{a_n,a_n\eta}\|_1 \leqslant \|\omega_\psi - \omega_{a_n\eta}\| \leqslant \|\psi + a_n\eta\| \, \|\psi - a_n\eta\| \to 0$ as $n \to \infty$. By paper I, lem. 11 and (11) applied to the commutative case, this gives that also

$$|F(\mu_{a_n,\psi}|\mu_{a_n,\eta}) - F(\mu_{a_n,a_n\eta}|\mu_{a_n,\eta})| \leqslant \|\mu_{a_n,\psi} - \mu_{a_n,a_n\eta}\|_1^{1/2} \to 0. \tag{89}$$

By definition,

$$\left(\frac{d\mu_{a_n,a_n\eta}(\lambda)}{d\mu_{a_n,\eta}(\lambda)}\right)^{1/2} = \lambda \quad \text{for } \lambda \in \mathbb{R}_+, \tag{90}$$

hence by (88)

$$\begin{aligned} S_{\text{meas}}(\omega_\psi|\omega_\eta) \geqslant -2\ln\lim_n \int \lambda d\mu_{a_n,\eta} = -2\ln\lim_n \int \lambda\langle\eta|E_{a_n}(d\lambda)\eta\rangle \\ = -2\ln\lim_n\langle\eta|a_n\eta\rangle = -2\ln\langle\eta|u'\psi\rangle = -2\ln|\langle\eta|u'\psi\rangle|. \end{aligned} \tag{91}$$

The right side is by definition $\geqslant -\ln F(\omega_\psi|\omega_\eta)^2$ as $\|u'\| = 1, u' \in \mathcal{M}'$, which concludes the proof. □

5.2 Petz recovery map

We now recall the definition of the Petz map in the case of general von Neumann algebras, discussed in more detail in [28], sec. 8. Let $T : \mathcal{B} \to \mathcal{A}$ be a *-preserving linear map between two von Neumann algebras $\mathcal{A}, \mathcal{B}$ in standard form acting on Hilbert spaces $\mathscr{H}, \mathscr{K}$. If

$$\left(\langle\zeta_1| \quad \langle\zeta_1|\right) T\left(\begin{bmatrix} a & b \\ c & d\end{bmatrix}\begin{bmatrix} a^* & c^* \\ b^* & d^*\end{bmatrix}\right)\begin{pmatrix}|\zeta_1\rangle \\ |\zeta_2\rangle\end{pmatrix} \geqslant 0, \quad \forall|\zeta_i\rangle \in \mathscr{H}, \quad T(1_{\mathcal{B}}) = 1_{\mathcal{A}}, \tag{92}$$

and for all $a, b, c, d \in \mathcal{B}$, then T is called 2-positive and unital. In the matrix inequality, we mean T applied to each matrix element. By duality between $\mathcal{A}$ and $\mathscr{S}(\mathcal{A})$, $T : \mathcal{B} \to \mathcal{A}$ gives a corresponding map $\tilde{T} : \mathscr{S}(\mathcal{A}) \to \mathscr{S}(\mathcal{B})$ by $\omega \mapsto \tilde{T}(\omega) := \omega \circ T$. For finite dimensional von Neumann algebras $\mathcal{A}, \mathcal{B}$ where state functionals are identified with density matrices through $\omega(a) = \mathrm{Tr}(\omega a)$, we can think of $\tilde{T}$ as the linear operator on density matrices defined by

$$\mathrm{Tr}\,\omega T(b) = \mathrm{Tr}\,\tilde{T}(\omega)b \quad \forall b \in \mathcal{B}. \tag{93}$$

This operator $\tilde{T}$ is completely positive and trace-preserving. The quantum data processing inequality (DPI) [37] states that

$$S(\omega_\psi|\omega_\eta) \geqslant S(\omega_\psi \circ T|\omega_\eta \circ T), \tag{94}$$

where the right side could also be written as $S(\tilde{T}(\omega_\psi)|\tilde{T}(\omega_\eta))$.

We recall the definition of the Petz-map. Let $|\eta_{\mathcal{A}}\rangle$ be a cyclic and separating vector in the natural cone of a von Neumann algebra $\mathcal{A}$ in standard form. Then the KMS scalar product on $\mathcal{A}$ is defined as

$$\langle a_1, a_2\rangle_\eta = \langle\eta_{\mathcal{A}}|a_1^*\Delta_\eta^{1/2}a_2\eta_{\mathcal{A}}\rangle. \tag{95}$$

Let ω_η be the normal state functional on $\mathcal{A}$ associated with $|\eta_{\mathcal{A}}\rangle$. Then its pull-back $\omega_\eta \circ T$ to $\mathcal{B}$, which is also faithful[9] has a vector representative $|\eta_{\mathcal{B}}\rangle \in \mathscr{K}$ in the natural cone. So:

$$\omega_\eta(a) = \langle\eta_{\mathcal{A}}|a\eta_{\mathcal{A}}\rangle, \quad \omega_\eta \circ T(b) = \langle\eta_{\mathcal{B}}|b\eta_{\mathcal{B}}\rangle. \tag{96}$$

$|\eta_{\mathcal{A}}\rangle$ resp. $|\eta_{\mathcal{B}}\rangle$ give KMS scalar products for $\mathcal{A}$ resp. $\mathcal{B}$, which we can use to define the adjoint $T^+ : \mathcal{A} \to \mathcal{B}$ (depending on the choices of these vectors) of the normal, unital

[9] This follows from Kadison's inequality $T(b^*b) \geqslant T(b)^*T(b)$.

and 2-positive $T : \mathcal{B} \to \mathcal{A}$, which is again normal, unital, and 2-positive, see [28] prop. 8.3. For finite dimensional matrix algebras T^+ corresponds dually to the linear operator $\tilde{T}^+$ acting on density matrices ρ for $\mathcal{B}$ given by

$$\tilde{T}^+(\rho) = \sigma_{\mathcal{A}}^{1/2} T\left(\sigma_{\mathcal{B}}^{-1/2} \rho \sigma_{\mathcal{B}}^{-1/2}\right) \sigma_{\mathcal{A}}^{1/2}, \tag{97}$$

wherein $\sigma_{\mathcal{A}}$ is the density matrix of $|\eta_{\mathcal{A}}\rangle$ and $\sigma_{\mathcal{B}} = \tilde{T}(\sigma_{\mathcal{A}})$ for $|\eta_{\mathcal{B}}\rangle$. The rotated Petz map, which we call $\alpha^t_{\eta,T} : \mathcal{A} \to \mathcal{B}$, is defined by conjugating this with the respective modular flows, i.e.

$$\alpha^t_{\eta,T} = \varsigma^t_{\eta,\mathcal{B}} \circ T^+ \circ \varsigma^{-t}_{\eta,\mathcal{A}} \tag{98}$$

where $\varsigma^t_{\eta,\mathcal{A}} = \mathrm{Ad}\Delta^{it}_{\eta,\mathcal{A}}$ is the modular flow for $\mathcal{A}, |\eta_{\mathcal{A}}\rangle$ etc. For finite dimensional matrix algebras, $\alpha^t_{\eta,T}$ gives by duality a linear operator $\tilde{\alpha}^t_{\eta,T}$ acting on density matrices ρ for $\mathcal{B}$, which is

$$\tilde{\alpha}^t_{\eta,T}(\rho) = \sigma_{\mathcal{A}}^{1/2-it} T\left(\sigma_{\mathcal{B}}^{-1/2+it} \rho \sigma_{\mathcal{B}}^{-1/2-it}\right) \sigma_{\mathcal{A}}^{1/2+it}. \tag{99}$$

An equivalent definition of the rotated Petz map is:

Definition 2. *Let $T : \mathcal{B} \to \mathcal{A}$ be a unital, normal, and 2-positive, linear map and $|\eta_{\mathcal{A}}\rangle \in \mathscr{H}$ a faithful state. Then the rotated Petz map $\alpha^t_{\eta,T} : \mathcal{A} \to \mathcal{B}$ is defined implicitly by the identity:*

$$\langle b\eta_{\mathcal{B}} | J_{\mathcal{B}} \Delta^{it}_{\eta_{\mathcal{B}}} \alpha^t_{\eta,T}(a) \eta_{\mathcal{B}} \rangle = \langle T(b)\eta_{\mathcal{A}} | J_{\mathcal{A}} \Delta^{it}_{\eta_{\mathcal{A}}} a \eta_{\mathcal{A}} \rangle, \tag{100}$$

for all $a \in \mathcal{A}, b \in \mathcal{B}$.

Closely related to the Petz map is the linear map $V_\psi : \mathscr{K} \to \mathscr{H}$ defined[10] ω_ψ by [32, 30]

$$V_\psi b|\xi^{\mathcal{B}}_\psi\rangle := T(b)|\xi^{\mathcal{A}}_\psi\rangle \quad (b \in \mathcal{B}). \tag{102}$$

It follows from Kadison's property $T(a^*a) \geqslant T(a^*)T(a)$ (which is a consequence of (92)) that V_ψ is a contraction $\|V_\psi\| \leqslant 1$, see e.g. [32], proof of thm. 4.

As in paper II, we introduce a vector valued function

$$z \mapsto |\Gamma_\psi(z)\rangle := \Delta^z_{\eta_{\mathcal{A}},\psi_{\mathcal{A}}} V_\psi \Delta^{-z}_{\eta_{\mathcal{B}},\psi_{\mathcal{B}}} |\xi^{\mathcal{B}}_\psi\rangle \quad (z \in \overline{\mathbb{S}}_{1/2}), \tag{103}$$

the existence and properties of which are established in lem.s 3, 4 in paper II. In particular, $|\Gamma_\psi(z)\rangle$ is holomorphic inside the strip $\mathbb{S}_{1/2}$ and bounded in the closure $\overline{\mathbb{S}}_{1/2}$ in norm by 1. Furthermore, the representation (24) of paper I shows in conjunction with Stone's theorem that this function is strongly continuous on the boundaries of the strip $\mathbb{S}_{1/2}$, i.e. for $\mathrm{Re}(z) = 0$ or $\mathrm{Re}(z) = 1/2$, which is used implicitly below e.g. when we consider integrals involving this quantity along these boundaries. The relation to the Petz map is as follows, paper II, lem. 2:

$$\langle \Gamma_\psi(1/2 + it) | a \, \Gamma_\psi(1/2 + it) \rangle \leqslant \omega_\psi \circ T \circ \alpha^t_{\eta,T}(a) \quad t \in \mathbb{R}, a \in \mathcal{A}_+. \tag{104}$$

[10] As it stands, the definition is actually consistent only when $|\xi^{\mathcal{B}}_\psi\rangle$ is cyclic and separating. In the general case, one can define [32] instead

$$V_\psi(b|\xi^{\mathcal{B}}_\psi\rangle + |\zeta\rangle) := T(b)|\xi^{\mathcal{B}}_\psi\rangle \quad (b \in \mathcal{B}, \pi^{\mathcal{B}'}(\psi)|\zeta\rangle = 0). \tag{101}$$

5.3 Improved DPI

Our main theorem is:

Theorem 1. *Let $T : \mathcal{B} \to \mathcal{A}$ be a two-positive, unital (in the sense (92)) linear map between two von Neumann algebras, and let ω_ψ, ω_η be normal states on $\mathcal{A}$, with ω_η faithful. Then*

$$S(\omega_\psi|\omega_\eta) - S(\omega_\psi \circ T|\omega_\eta \circ T) \geqslant S_{\text{meas}}(\omega_\psi|\omega_\psi \circ T \circ \alpha_{T,\eta}). \tag{105}$$

with the recovery channel

$$\alpha_{T,\eta} \equiv \int_{\mathbb{R}} dt\, \beta_0(t)\, \alpha^t_{T,\eta}. \tag{106}$$

Remarks: 1) The theorem should generalize to non-faithful ω_η by applying appropriate support projections in a similar way as in paper I, lem. 1.

2) For finite-dimensional type I von Neumann algebras i.e. matrices, our result is due to [35]. The recovery channel is given explicitly by (99) in this case as an operator on density matrices, where $\sigma_\mathcal{A}, \sigma_\mathcal{B}$ are the density matrices corresponding to $\omega_\eta, \omega_\eta \circ T$.

3) By prop. 1, our bound implies that given in our previous paper II for the fidelity; in fact it is stronger in many cases.

I) Proof under a majorization condition: First we consider the special case where there exists $\infty > c \geqslant 1$ such that

$$c^{-1}\omega_\eta \leqslant \omega_\psi \leqslant c\omega_\eta. \tag{107}$$

Note that this implies $c^{-1}\omega_\eta \circ T \leqslant \omega_\psi \circ T \leqslant c\omega_\eta \circ T$ as T is positive. By [28], thm. 12.11 (due to Araki), there exists a $h = h^* \in \mathcal{A}$ such that $|\psi\rangle = |\eta^h\rangle / \|\eta^h\|$ such that $\|h\| \leqslant \ln c$, and vice versa. As is well known, this furthermore implies that the Connes cocycle $[D\eta^\mathcal{B} : D\psi^\mathcal{B}]_{iz}$ is holomorphic in the two-sided strip $\{z \in \mathbb{C} : |\text{Re}(z)| < 1/2\}$ and bounded in norm (by $c^{\text{Re}(z)}$) on the closure of this strip, see e.g. paper II, lem. 5. As a consequence, we have an absolutely convergent (in the operator norm) power series expansion

$$[D\eta^\mathcal{B} : D\psi^\mathcal{B}]_{iz} = 1 + \sum_{l=1}^{\infty} z^l k_l, \tag{108}$$

with bounded operators $k_l \in \mathcal{B}$ such that $\|k_l\| \leqslant C^l$. We set

$$k := \frac{d}{idt} T([D\eta^\mathcal{B} : D\psi^\mathcal{B}]_t)|_{t=0} \in \mathcal{A}_{\text{s.a.}}. \tag{109}$$

Using [28], cor. 12.8, and the definition of the relative entropy in terms of the Connes cocycle,

$$\begin{aligned} S_\mathcal{A}(\psi|\eta^k) =& S_\mathcal{A}(\psi|\eta) - \omega_\psi(k) \\ =& S_\mathcal{A}(\psi|\eta) - \langle \psi^\mathcal{A}| \frac{d}{idt} T([D\eta^\mathcal{B} : D\psi^\mathcal{B}]_t)\psi^\mathcal{A}\rangle|_{t=0} \\ =& S_\mathcal{A}(\psi|\eta) - S_\mathcal{B}(\psi|\eta), \end{aligned} \tag{110}$$

which is one side of the inequality that we would like to prove. The variational expression (116) then gives:

$$S_\mathcal{A}(\psi|\eta) - S_\mathcal{B}(\psi|\eta) = \sup_{h \in \mathcal{A}_{\text{s.a.}}} \{\omega_\psi(h) - \ln \|\eta^{h+k}\|^2\}, \tag{111}$$

where we used $|(\eta^k)^h\rangle = |\eta^{k+h}\rangle$ see [28], thm. 12.10. To get the desired DPI we will establish an upper bound on $\ln\|\eta^{h+k}\|^2$.

In lem. 1, we take $|G(z)\rangle = e^{zh}|\Gamma_\psi(z)\rangle$, $p_0 = \infty, p_1 = 2$ where $\theta = 1/n$ with $n \in 4\mathbb{N}$ and $h = h^* \in \mathcal{A}$. At the lower boundary we have with $u_\mathcal{B}(t) := [D\eta^\mathcal{B} : D\psi^\mathcal{B}]_t \in \mathcal{B}, u_\mathcal{A}(t) := [D\eta^\mathcal{A} : D\psi^\mathcal{A}]_t \in \mathcal{A}$ the unitary Connes cocycles,

$$
\begin{aligned}
\|G(it)\|_{p_0,\psi} =& \|e^{ith}\Delta^{it}_{\eta_\mathcal{A},\psi_\mathcal{A}} V_\psi \Delta^{-it}_{\eta_\mathcal{B},\psi_\mathcal{B}} \xi^\mathcal{B}_\psi\|_{\infty,\psi} \\
=& \|e^{ith}\Delta^{it}_{\eta_\mathcal{A},\psi_\mathcal{A}} T(u_\mathcal{B}(t))\psi\|_{\infty,\psi} \\
=& \|e^{ith}\varsigma^t_\eta[T(u_\mathcal{B}(t))]u_\mathcal{A}(t)^*\psi\|_{\infty,\psi} \\
=& \|e^{ith}\varsigma^t_\eta[T(u_\mathcal{B}(t))]u_\mathcal{A}(t)^*\| \\
=& \|\varsigma^t_\eta[T(u_\mathcal{B}(t))]\| \leqslant 1,
\end{aligned}
\tag{112}
$$

where we used $\|\varsigma^t_\eta[T(b)]\| = \|T(b)\| \leqslant \|b\|$ (from the positivity of T and $\varsigma^t_\eta = \mathrm{Ad}\Delta^{it}_{\eta_\mathcal{A}}$) as well as the isomeric identification of $L^\infty(\mathcal{A},\psi) \ni a|\psi\rangle \mapsto a \in \mathcal{A}$ proven in [3]. Since $p_\theta = n$ and $\ln\|G(it)\|_{p_0,\psi} \leqslant 0$ as just shown, we get from lem. 1

$$
\begin{aligned}
\ln\|e^{h/n}\Gamma_\psi(1/n)\|^n_{\psi,n} \leqslant& \int_\mathbb{R} dt\, \beta_{1/n}(t)\, \ln\|G(1/2+it)\|_{p_1,\psi} \\
=& \int_\mathbb{R} dt\, \beta_{1/n}(t)\, \ln\|e^{h/2}\Gamma_\psi(1/2+it)\|^2 \\
\leqslant& \ln\int_\mathbb{R} dt\, \beta_{1/n}(t)\, \|e^{h/2}\Gamma_\psi(1/2+it)\|^2 \\
\leqslant& \ln\int_\mathbb{R} dt\, \beta_{1/n}(t)\, \omega_\psi \circ T \circ \alpha^t_{\eta,T}(e^h),
\end{aligned}
\tag{113}
$$

using (104) in the third line and Jensen's inequality in the second (noting that the integrand is continuous and uniformly bounded). Taking the lim-sup $n \to \infty$, we get using the definition of the recovery channel $\alpha_{T,\eta}$:

$$
\limsup_n \ln\|e^{h/n}\Gamma_\psi(1/n)\|^n_{\psi,n} \leqslant \omega_\psi \circ T \circ \alpha_{\eta,T}(e^h). \tag{114}
$$

The next lemmas give an expression for the lim-sup:

Lemma 5. *We have* $\|e^{h/n}\Gamma_\psi(1/n)\|^n_{\psi,n} = \|(e^{h/n}\Delta^{1/n}_{\eta,\psi}a_n\Delta^{1/n}_{\eta,\psi}e^{h/n})^{n/4}\psi\|^2$, *where*

$$
a_n = T([D\eta^\mathcal{B} : D\psi^\mathcal{B}]_{i/n})^* T([D\eta^\mathcal{B} : D\psi^\mathcal{B}]_{i/n}) \in \mathcal{A}_+. \tag{115}
$$

Lemma 6. *We have* $\lim_n \|(e^{h/n}\Delta^{1/n}_{\eta,\psi}a_n\Delta^{1/n}_{\eta,\psi}e^{h/n})^{n/4}\psi\|^2 = \|\eta^{h+k}\|^2$.

Combining the two lemmas with eq.s (111), (114) gives

$$
S_\mathcal{A}(\psi|\eta) - S_\mathcal{B}(\psi|\eta) \geqslant \sup_{h\in\mathcal{A}_{\mathrm{s.a.}}} \{\omega_\psi(h) - \ln\omega_\psi \circ T \circ \alpha_{\eta,T}(e^h)\} = S_{\mathrm{meas}}(\omega_\psi|\omega_\psi \circ T \circ \alpha_{T,\eta}), \tag{116}
$$

using the variational definition (5.1) of S_{meas} in the last step.

Proof of lem. 6: Since (108) is an absolutely convergent power series in the operator norm, it follows that $a_n = 1 + 2n^{-1}k + O(n^{-2})$ where $O(n^\alpha)$ denotes a family of operators such that $\|O(n^\alpha)\| \leqslant cn^\alpha$ for all $n > 0$. Since h is bounded, we also have $e^{h/n} = 1 + n^{-1}h + O(n^{-2})$. Replacing $n \to 2n$ to simplify some expressions we trivially get

$$e^{h/(2n)}\Delta_{\eta,\psi}^{1/(2n)} a_{2n} \Delta_{\eta,\psi}^{1/(2n)} e^{h/(2n)} = \Delta_{\eta,\psi}^{1/n} + n^{-1}X_n + n^{-2}Y_n \tag{117}$$

where X_n, Y_n is a finite sum of terms of the form $x_0\Delta^{s_1}_{\eta,\psi}x_1 \cdots x_l \Delta^{s_l}_{\eta,\psi} x_l$ wherein $\sum s_j = 1/n$, $s_j \geqslant 0$ and $\|x_j\| \leqslant c$ uniformly in n. X_n is given explicitly by

$$X_n = \tfrac{1}{2}h\Delta_{\eta,\psi}^{1/n} + \tfrac{1}{2}\Delta_{\eta,\psi}^{1/n}h + \Delta_{\eta,\psi}^{1/(2n)}k\Delta_{\eta,\psi}^{1/(2n)}. \tag{118}$$

By [2], II, proof of thm. 3.1, the functions

$$F(z) := x_1\Delta^{z_1}_{\eta,\psi}x_2 \cdots x_j \Delta^{z_j}_{\eta,\psi}x_{j+1}|\psi\rangle, \quad z \in \bar{\mathbb{S}}^j_{1/2} \tag{119}$$

defined for given $x_j \in \mathcal{A}$ are (strongly) analytic in the domain $\mathbb{S}^j_{1/2} := \{(z_1, \dots, z_j) \in \mathbb{C}^j : 0 < \mathrm{Re}(z_i), \sum \mathrm{Re}(z_i) < 1/2\}$ and strongly continuous on the closure. Subharmonic analysis as in [2], II, proof of thm. 3.1, or [3] furthermore gives the bound

$$\|F(z)\| \leqslant \prod_i \|x_i\|, \quad \forall z \in \bar{\mathbb{S}}^j_{1/2}. \tag{120}$$

This bound, and the elementary formula

$$(A + tB)^N = \sum_{j=0}^{N} t^j \sum_{\substack{m_0 + \dots + m_j = N-j, \\ m_j \in \mathbb{N}_0}} A^{m_0}B \cdots A^{m_{j-1}}BA^{m_j}, \tag{121}$$

shows that the difference

$$\begin{aligned} |\zeta_n\rangle = {} & (e^{h/(2n)}\Delta_{\eta,\psi}^{1/(2n)} a_{2n} \Delta_{\eta,\psi}^{1/(2n)} e^{h/(2n)})^{n/2}|\psi\rangle \\ & - \sum_{j=0}^{n/2} n^{-j} \sum_{\substack{m_0 + \dots + m_j = n/2-j, \\ m_j \in \mathbb{N}_0}} \Delta_{\eta,\psi}^{m_0/n}X_n \cdots \Delta_{\eta,\psi}^{m_{j-1}/n}X_n\Delta_{\eta,\psi}^{m_j/n}|\psi\rangle \end{aligned} \tag{122}$$

is bounded in norm by

$$\|\zeta_n\| \leqslant \big(1 + n^{-1}(\|h\| + \|k\|) + n^{-2}c\big)^{n/2} - \big(1 + n^{-1}(\|h\| + \|k\|)\big)^{n/2} \tag{123}$$

for some $c < \infty$, hence it tends to zero in norm as $n \to \infty$. Setting now

$$|\phi_{n,j}\rangle = n^{-j} \sum_{\substack{m_0 + \dots + m_j = n/2-j, \\ m_j \in \mathbb{N}_0}} \Delta_{\eta,\psi}^{m_0/n}X_n \cdots \Delta_{\eta,\psi}^{m_{j-1}/n}X_n\Delta_{\eta,\psi}^{m_j/n}|\psi\rangle, \tag{124}$$

the strong continuity of the functions F and the usual definition of the Riemann integral implies

$$\begin{aligned} & |\phi_j\rangle := \lim_n |\phi_{n,j}\rangle \\ = & \int_0^{1/2} ds_0 \dots \int_0^{s_{j-1}} ds_j \, \Delta_{\eta,\psi}^{s_0 - s_1}(h+k)\Delta_{\eta,\psi}^{s_1-s_2}(h+k) \dots \Delta_{\eta,\psi}^{s_{j-1}-s_j}(h+k)\Delta_{\eta,\psi}^{s_j}|\psi\rangle, \end{aligned} \tag{125}$$

and the usual perturbation theory by bounded operators as in [4], prop. 16 gives $\sum_{j=0}^{\infty} |\phi_j\rangle = e^{(\ln \Delta_{\eta,\psi}+h+k)/2}|\psi\rangle$. Hence,

$$\lim_n (e^{h/(2n)} \Delta_{\eta,\psi}^{1/(2n)} a_{2n} \Delta_{\eta,\psi}^{1/(2n)} e^{h/(2n)})^{n/2} |\psi\rangle = e^{(\ln \Delta_{\eta,\psi}+h+k)/2}|\psi\rangle \tag{126}$$

strongly, as argued more carefully in [6], proof of lem. 5. We have $e^{(\ln \Delta_{\eta,\psi}+h+k)/2}|\psi\rangle = e^{(\ln \Delta_{\eta,\psi}+p'h+p'k)/2}|\psi\rangle$ (here $p' = \pi^{\mathcal{A}'}(\psi) \in \mathcal{A}'$). Also, using [28], thm. 12.6., we have $\ln \Delta_{\eta,\psi} + p'h + p'k = \ln \Delta_{\eta^{h+k},\psi}$, and this gives $|\eta^{h+k}\rangle = J|\eta^{h+k}\rangle = e^{(\ln \Delta_{\eta,\psi}+h+k)/2}|\psi\rangle$ by relative modular theory. This completes the proof. □

Proof of lem. 5: From the definitions,

$$e^{h/n}\Gamma_\psi(1/n) = e^{h/n} \Delta_{\eta_{\mathcal{A}},\psi_{\mathcal{A}}}^{1/n} V_\psi \Delta_{\eta_{\mathcal{B}},\psi_{\mathcal{B}}}^{-1/n} \Delta_{\psi_{\mathcal{B}}}^{1/n} |\psi^{\mathcal{B}}\rangle = e^{h/n} \Delta_{\eta_{\mathcal{A}},\psi_{\mathcal{A}}}^{1/n} T([D\eta^{\mathcal{B}} : D\psi^{\mathcal{B}}]_{i/n})|\psi^{\mathcal{A}}\rangle, \tag{127}$$

using the definition of the Connes-cocylce and the fact that $[D\eta^{\mathcal{B}} : D\psi^{\mathcal{B}}]_{i/n} \in \mathcal{B}$ under our assumption (107), see paper II, proof of lem. 4. In the following, let $a = e^{h/n}, b = T([D\eta^{\mathcal{B}} : D\psi^{\mathcal{B}}]_{i/n}) \in \mathcal{A}$ and $|\psi^{\mathcal{A}}\rangle = |\psi\rangle, |\eta^{\mathcal{A}}\rangle = |\eta\rangle$ etc.

By the results of [3] (which hold in the present context since ω_ψ is faithful being dominated by the faithful state ω_η), the vector $b\Delta_{\eta,\psi}^{1/n} a|\psi\rangle \in L_n(\mathcal{A},\psi)$ has a polar decomposition $b\Delta_{\eta,\psi}^{1/n} a|\psi\rangle = u\Delta_{\phi_n,\psi}^{1/n}|\psi\rangle$, where $\|b\Delta_{\eta,\psi}^{1/n} a\psi\|_{n,\psi}^n = \|\phi_n\|^2$ and where $u \in \mathcal{A}$ is a partial isometry. To get an expression for $|\phi_n\rangle$, we use the formalism of "script" $\mathscr{L}_p$-spaces of [3], notation 7.6: As a vector space $\mathscr{L}_p^*(\mathcal{A},\psi), p \geqslant 1$ consists of all formal linear combinations of formal expressions of the form

$$A = x_1 \Delta_{\zeta_1,\psi}^{z_1} x_2 \dots x_n \Delta_{\zeta_n,\psi}^{z_n} x_{n+1} \tag{128}$$

wherein $\mathrm{Re}(z_i) \geqslant 0, \sum_i \mathrm{Re}(z_i) \leqslant 1 - 1/p$, $x_i \in \mathcal{A}, \zeta_i \in \mathscr{H}$, the formal adjoint of which is defined to be

$$A^* = x_{n+1}^* \Delta_{\zeta_n,\psi}^{\bar{z}_n} x_n^* \dots x_2^* \Delta_{\zeta_1,\psi}^{\bar{z}_1} x_1^*. \tag{129}$$

The notation $\mathscr{L}_{p,0}^*(\mathcal{A},\psi)$ is reserved for formal elements A such that $\sum_i \mathrm{Re}(z_i) = 1-1/p$ in addition to all other conditions. It is then clear that $\mathscr{L}_{p,0}^*(\mathcal{A},\psi)\mathscr{L}_{q,0}^*(\mathcal{M},\psi) = \mathscr{L}_{r,0}^*(\mathcal{M},\psi)$ as formal products where $1/r' = 1/p' + 1/q'$ with $1/p' = 1 - 1/p$ as usual. By [3], lem. 7.3, if $1 \leqslant p \leqslant 2$, any element $A \in \mathscr{L}_p^*(\mathcal{A},\psi)$ can be viewed as an element of $L_{p'}(\mathcal{A},\psi)$ in the sense that $|\psi\rangle \in \mathscr{D}(A)$ and $A|\psi\rangle \in L_{p'}(\mathcal{A},\psi)$.[11] Furthermore, by [3], lem. 7.7 (2), if $A_1, A_2 \in \mathscr{L}_p^*(\mathcal{A},\psi)$ correspond to the same element under this identification, then so do A_1^*, A_2^* or A_1B, A_2B or BA_1, BA_2 if $B \in \mathscr{L}_{q,0}^*(\mathcal{A},\psi)$ (as long as $1/p' + 1/q' \leqslant 1/2$, for example).

We now start with the trivial statement that $u\Delta_{\phi_n,\psi}^{1/n} = b\Delta_{\eta,\psi}^{1/n} a$ in the sense that these elements of $\mathscr{L}_{n',0}^*(\mathcal{A},\psi)$ are identified with the same element of $L_n(\mathcal{A},\psi)$. Then repeated application of [3], lem. 7.7 (2) and the definition of adjoint gives

$$u\Delta_{\phi_n,\psi}^{2/n} u^* = b\Delta_{\eta,\psi}^{1/n} aa^* \Delta_{\eta,\psi}^{1/n} b^* \quad \text{in } \mathscr{L}_{n/(n-2),0}^*(\mathcal{A},\psi). \tag{130}$$

Forming successively $n/4$ products of this equality and applying in each step [3], lem. 7.7 (2), we find that

$$u\Delta_{\phi_n,\psi}^{1/2} u^* = (b\Delta_{\eta,\psi}^{1/n} aa^* \Delta_{\eta,\psi}^{1/n} b^*)^{n/4} \quad \text{in } \mathscr{L}_{2,0}^*(\mathcal{A},\psi), \tag{131}$$

[11] In fact, $\|A\psi\|_{p',\psi} \leqslant \|x_{n+1}\| \prod_{i=1}^n (\|x_i\| \|\zeta_i\|^{\mathrm{Re}(z_i)})$.

meaning that both sides are equal as elements of $\mathscr{H} = L_2(\mathcal{A}, \psi)$ after we apply them to $|\psi\rangle$. Thus,

$$\|(b\Delta_{\eta,\psi}^{1/n} aa^* \Delta_{\eta,\psi}^{1/n} b^*)^{n/4}\psi\|^2 = \|u\Delta_{\phi_n,\psi}^{1/2} u^*\psi\|^2 = \|uJu\phi_n\|^2 = \|\phi_n\|^2 \tag{132}$$

using modular theory. Therefore

$$\|(b\Delta_{\eta,\psi}^{1/n} aa^* \Delta_{\eta,\psi}^{1/n} b^*)^{n/4}\psi\|^2 = \|b\Delta_{\eta,\psi}^{1/n} a\psi\|_{n,\psi}^n, \tag{133}$$

and the proof of the lemma is complete. □

II) Proof in general case: We will now remove the majorization condition (107). This condition has been used in an essential way in most of the arguments so far because without it, the operator k in (109) is unbounded and thus not an element of $\mathcal{A}$. For unbounded operators the Araki-Trotter product formula and the L_p-techniques are not available and it seems non-trivial extending them to an unbounded framework. We will therefore proceed in a different way and define a regularization of ω_ψ such that the majorization condition (107) holds and such that, at the same time, the desired entropy inequality can be obtained in a limit wherein the regulator is removed. However, it is clear that this regularization must be carefully chosen because the relative entropy is not continuous but only lower semi-continuous. By itself the latter is insufficient for our purposes since the desired inequality (105) has both signs of the relative entropy.

Our regularization combines a trick invented in paper I with the convexity of the relative entropy. As in paper I, we consider a function $f(t), t \in \mathbb{R}$ with the following properties.

(A) The Fourier transform of f

$$\tilde{f}(p) = \int_{-\infty}^{\infty} e^{-itp} f(t) dt \tag{134}$$

exists as a real and non-negative Schwarz-space function. This implies that the original function f is Schwarz and has finite $L_1(\mathbb{R})$ norm, $\|f\|_1 < \infty$.

(B) $f(t)$ has an analytic continuation to the upper complex half plane such that the $L_1(\mathbb{R})$ norm of the shifted function has $\|f(\cdot + i\theta)\|_1 < \infty$ for $0 < \theta < \infty$.

Such functions certainly exist (e.g. Gaussians). We also let $f_P(t) = Pf(tP)$ for our regulator $P > 0$, and we define a regulated version of $|\psi\rangle$ by

$$|\psi_P\rangle = \frac{\tilde{f}_P(\ln \Delta_{\eta,\psi})\,|\psi\rangle}{\|\tilde{f}_P(\ln \Delta_{\eta,\psi})\psi\|}. \tag{135}$$

As shown in paper I, some key properties of the regulated vectors are:

(P1) $\omega_{\psi_P} \leqslant c_P \omega_\eta$ for some $c_P > 0$ which may diverge as $P \to \infty$,

(P2) $s - \lim_{P\to\infty} |\psi_P\rangle = |\psi\rangle$ (strong convergence),

(P3) $-2\ln\left(\|f\|_1/\|\tilde{f}\|_\infty\right) + \limsup_{P\to\infty} S(\psi_P|\eta) \leqslant S(\psi|\eta)$, where the first item gives at least "half" of the domination condition (107), the second states in which sense $|\psi_P\rangle$ approximates $|\psi\rangle$ and the third gives us an upper semi-continuity property of the relative entropy opposite to the usual lower semi-continuity property which holds for generic approximations. We define for small $\varepsilon > 0$:

$$\sigma(a) = \langle \eta | a\eta \rangle \quad \rho_{P,\varepsilon}(a) = (1-\varepsilon)\langle \psi_P | a\psi_P \rangle + \varepsilon \langle \eta | a\eta \rangle. \tag{136}$$

Thus, by P1), the relative majorization condition (107) holds e.g. with $c = \max(c_P, \varepsilon^{-1})$ between $\rho_{P,\varepsilon}$ and σ. By P2), $\lim_{P\to\infty} \lim_{\varepsilon\to 0} \|\rho - \rho_{P,\varepsilon}\| = 0$. In P3), we choose a function f such that $\|f\|_1/\|\tilde{f}\|_\infty = 1$ (which must be Gaussian). The well-known convexity of the relative entropy gives together with the definition of $\rho_{\varepsilon,P}$ that ($\rho_P = \langle \psi_P | \ . \ \psi_P \rangle$)

$$S(\rho_{P,\varepsilon}|\sigma) \leqslant (1-\varepsilon)\ S(\rho_P, \sigma) + \varepsilon\ S(\sigma|\sigma) = (1-\varepsilon)\, S(\rho_P, \sigma). \tag{137}$$

Combining this with P3), we get

$$\limsup_{P\to\infty} \limsup_{\varepsilon\to 0} S(\rho_{P,\varepsilon}|\sigma) \leqslant S(\rho|\sigma). \tag{138}$$

The norm convergence $\lim_P \lim_\varepsilon \rho_{P,\varepsilon} \circ T = \rho \circ T$ by P2) also gives in combination with the usual lower semi-continuity of the relative entropy, [2], II thm. 3.7 (2), that

$$\liminf_{P\to\infty} \liminf_{\varepsilon\to 0} S(\rho_{P,\varepsilon} \circ T | \sigma \circ T) \geqslant S(\rho \circ T | \sigma \circ T). \tag{139}$$

Now we combine eq.s (138), (139) with part I of the proof applied to the states $\rho_{P,\varepsilon}$ and σ, which obey the relative majorization condition. We get:

$$S(\rho|\sigma) - S(\rho \circ T | \sigma \circ T) \geqslant \limsup_{P\to\infty} \limsup_{\varepsilon\to 0} S_{\text{meas}}(\rho_{P,\varepsilon} | \rho_{P,\varepsilon} \circ T \circ \alpha_{T,\sigma}). \tag{140}$$

The proof of part II is then finished by proving lower semi-continuity for the measured relative entropy:

Lemma 7. *If $\mu_n, \nu_n, \mu, \nu \in \mathscr{S}(\mathcal{A})$ are such that $\lim_n \mu_n = \mu$ and $\lim_n \nu_n = \nu$ in the norm sense, then $S_{\text{meas}}(\mu|\nu) \leqslant \liminf_n S_{\text{meas}}(\mu_n|\nu_n)$.*

Proof. This is a straightforward consequence of the variational definition (5.1) of S_{meas}, choosing a near optimal h. □

□

Acknowledgements: SH thanks Tom Faulkner for conversations and the Max-Planck Society for supporting the collaboration between MPI-MiS and Leipzig U., grant Proj. Bez. M.FE.A.MATN0003.

A Weighted L_p spaces [3] and variational formulae

The weighted L_p-spaces were defined by [3] relative to a fixed vector $|\psi\rangle \in \mathscr{H}$ in the a natural cone of a standard representation of a von Neumann algebra $\mathcal{M}$. For $p \geqslant 2$, the space $L_p(\mathcal{M}, \psi)$ is defined as

$$L_p(\mathcal{M}, \psi) = \{|\zeta\rangle \in \bigcap_{|\phi\rangle \in \mathscr{H}} \mathscr{D}(\Delta_{\phi,\psi}^{(1/2)-(1/p)}), \|\zeta\|_{p,\psi} < \infty\}. \tag{141}$$

Here, the norm is

$$\|\zeta\|_{p,\psi} = \sup_{\|\phi\|=1} \|\Delta_{\phi,\psi}^{(1/2)-(1/p)}\zeta\|. \tag{142}$$

For $1 \leqslant p < 2$, $L_p(\mathcal{M}, \psi)$ is defined as the completion of $\mathscr{H}$ with respect to the following norm:

$$\|\zeta\|_{p,\psi} = \inf\{\|\Delta_{\phi,\psi}^{(1/2)-(1/p)}\zeta\| : \|\phi\| = 1, \pi^{\mathcal{M}}(\phi) \geqslant \pi^{\mathcal{M}}(\psi)\}. \tag{143}$$

In [3], it is assumed for most results that $|\psi\rangle$ is cyclic and separating. When using such results in the main text, we will be in that situation. A somewhat different approach replacing the relative modular operator by the Connes spatial derivative and containing also many new results is laid out some detail in [8].

References

[1] H. Araki, "Relative Hamiltonian for faithful normal states of a von Neumann algebra," Publ. RIMS Kyoto Univ. **9**, 165-209 (1973)

[2] H. Araki, "Relative entropy of states of von Neumann algebras.I,II." Publ. RIMS Kyoto Univ. **11**, 809-833 (1976) and **13**, 173-192 (1977)

[3] H. Araki and T. Masuda, "Positive cones and L_p-spaces for von Neumann algebras," Publ. RIMS Kyoto Univ. 18, 339-411 (1982).

[4] H. Araki, "Expansionals in Banach algebras," Ann. Sci. Ecole Norm. Sup. 6 , 67 (1973)

[5] H. Araki, "On an inequality of Lieb and Thirring," Lett. Math. Phys. 19, 167-170 (1990).

[6] H. Araki, "Golden-Thompson and Peierls-Bogliubov inequalities for a general von Neumann algebra," Commun. Math. Phys. 34, 167-178 (1973)

[7] M. Berta, M. Lemm, and M. M. Wilde. "Monotonicity of quantum relative entropy and recoverability." Quant. Inf. and Comp., **15** 1333-1354 (2015)

[8] M. Berta, V. B. Scholz and M. Tomamichel, "Renyi Divergences as Weighted Non-commutative Vector-Valued L_p -Spaces," Annales Henri Poincare **19**, no. 6, 1843 (2018)

[9] D. Buchholz, K. Fredenhagen and C. D'Antoni, "The Universal Structure of Local Algebras," Commun. Math. Phys. **111**, 123 (1987)

[10] E. A. Carlen, A. Vershynina, "Recovery map stability for the Data Processing Inequality," 1710.02409

[11] C. Cecchini and D. Petz, "State extensions," Pac. J. Math **138** (1989)

[12] F. Ceyhan and T. Faulkner, "Recovering the QNEC from the ANEC," arXiv:1812.04683 [hep-th].

[13] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle and M. Walter, "Entanglement Wedge Reconstruction via Universal Recovery Channels," Phys. Rev. X **9**, no. 3, 031011 (2019)

[14] M. J. Donald, "Relative Hamiltonians which are not bounded from above," J. Func. Anal., 91, 143-173 (1990)

[15] T. Faulkner, S. Hollands, B. Swingle and Y. Wang, "Approximate recovery and relative entropy I. general von Neumann subalgebras," [arXiv:2006.08002 [quant-ph]].

[16] T. Faulkner and S. Hollands, "Approximate recoverability and relative entropy II: 2-positive channels of general von Neumann algebras," [arXiv:2010.05513 [quant-ph]].

[17] O. Fawzi and R. Renner, "Quantum Conditional Mutual Information and Approximate Markov Chains," Commun. Math. Phys. 340, 575–611 (2015)

[18] L. Gao and M. M. Wilde, "Recovery for optimized f-divergences," [arXiv:2006.08002 [quant-ph]]

[19] R. Haag, *Local quantum physics: Fields, particles, algebras,* Springer: Berlin (1992)

[20] F. Hansen: "The fast track to Löwner's theorem," Lin. Alg. Appl. **438**, 4557–4571 (2013)

[21] I.I. Hirschman, "A convexity theorem for certain groups of transformations." J, d'Anal. Math., **2** (2), 209-218 (1952)

[22] Ch. Jäkel, and F. Robl, "The Hölder inequality for KMS states," [arXiv:1103.3608]

[23] A. Jencova, "Preservation of a Quantum Renyi Relative Entropy Implies Existence of a Recovery Map," J. Phys. A 50.8, 085303 (2017)

[24] A. Jencova, "Renyi Relative Entropies and Noncommutative L_p-Spaces," Ann. H. Poincare, **19**.8, 2513-2542 (2018)

[25] M. Junge, R. Renner, D. Sutter, M. M. Wilde and A. Winter, "Universal Recovery Maps and Approximate Sufficiency of Quantum Relative Entropy," Annales Henri Poincare **19**, no. 10, 2955 (2018)

[26] M. Junge and N. LaRacuente, "Universal recovery and p-fidelity in von Neumann algebras," [arXiv:2009.11866[quant-ph]]

[27] M. Muller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel. "On quantum relative Renyi entropies: a new generalization,". J. Math. Phys, 54(12): 122203 (2013).

[28] M. Ohya, D. Petz, *Quantum entropy and its use,* Theoretical and Mathematical Physics, Springer (1993)

[29] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von Neumann algebra," Commun. Math. Phys. 105(1):123-131 (1986)

[30] D. Petz, "Sufficiency of channels over von Neumann algebras," Quart. J. of Mathematics, 39(1), 97-108 (1988)

[31] D. Petz, "Monotonicity of quantum relative entropy revisited", Rev. Math. Phys., 15(01) 79-91 (2003).

[32] D. Petz, "Quasi-entropies for States of a von Neumann Algebra," RIMS, Kyoto Univ. 21, 787-800 (1985)

[33] D. Petz, "A variational expression for the relative entropy," Comm.Math.Phys. 114, 345-349 (1988).

[34] D. Sutter, M. Tomamichel, and A. W. Harrow, "Strengthened monotonicity of relative entropy via pinched Petz recovery map," IEEE Transactions on Information Theory, 62(5), 2907-2913 (2016)

[35] D. Sutter, M. Berta, M. Tomamichel, "Multivariate trace inequalities," Commun. Math. Phys. 352, 37-58 (2017)

[36] M. Takesaki *Theory of operator algebras, I-III*, Springer (2003)

[37] A. Uhlmann, "Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory," Commun. Math. Phys., 54(1): 21-32 (1977).

[38] A. Uhlmann, "The 'transition probability' in the state space of a *-algebra," Reports on Mathematical Physics. 9 (2): 273-279 (1976)

[39] M. M. Wilde, "Recoverability in quantum information theory," Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 471(2182):20150338 (2015)

[40] M. M. Wilde, A. Winter and D. Yang, "Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Renyi Relative Entropy," Commun. Math. Phys. **331** (2014) no.2, 593-622

[41] Mark M. Wilde, "Monotonicity of p-norms of multiple operators via unitary swivels," [arXiv:1610.01262]