Variational approach to relative entropies with an application to QFT

Stefan Hollands[1]

Abstract
We define a new divergence of von Neumann algebras using a variational expression similar in nature to Kosaki's formula for Umegaki's relative entropy. Our divergence satisfies several of the usual desirable properties, upper bounds the sandwiched Renyi entropy and reduces to the fidelity in a limit. As an illustration, we use the formula in quantum field theory to compute our divergence between the vacuum in a bipartite system and an "orbifolded"—in the sense of a conditional expectation—system in terms of the Jones index. We take the opportunity to point out an entropic certainty relation associated with an inclusion of von Neumann factors related to the relative entropy. This certainty relation has an equivalent formulation in terms of error correcting codes.

1 Introduction

The relative entropy between two density matrices ω_η, ω_ψ, defined as

$$S(\omega_\psi|\omega_\eta) = \mathrm{Tr}(\omega_\psi \ln \omega_\psi) - \mathrm{Tr}(\omega_\psi \ln \omega_\eta), \qquad (1)$$

is an asymptotic measure of their distinguishability. Classically, $e^{-NS(\{p_i\}|\{q_i\})}$ approaches for large N the probability for a sample of size N of letters, distributed according to the true distribution $\{p_i\}$, when calculated according to an incorrect guess $\{q_i\}$. In the setting of general von Neumann algebras $\mathcal{A}$, density matrices should be

✉ Stefan Hollands
stefan.hollands@uni-leipzig.de

[1] Institute for Theoretical Physics, Universitaet Leipzig, Bruederstrasse 16, D-04103 Leipzig, Germany

replaced by algebraic states, i.e., positive, linear, normalized and ultra-weakly continuous functionals $\omega_\eta, \omega_\psi : \mathcal{A} \to \mathbb{C}$. The correspondence to density matrices is given by $\omega_\psi(a) = \mathrm{Tr}(a\omega_\psi), a \in \mathcal{A}$ in the case of matrix algebras $\mathcal{A} = M_n(\mathbb{C})$. It has been shown how to generalize the relative entropy to general von Neumann algebras by Araki [2,3]. Instead of $\ln\omega_\psi$ and $\ln\omega_\eta$, which have no obvious counterparts in a general von Neumann algebra, he used relative modular operators.

By far the most fundamental property of S—from which in fact essentially all others follow—is its monotonicity under a channel. A channel $T : \mathcal{B} \to \mathcal{A}$ between von Neumann algebras is a completely positive, unital normal linear map. It induces a corresponding action on states by pull-back, $\omega_\psi \mapsto \omega_\psi \circ T$. [36] has shown for general von Neumann algebras that

$$S(\omega_\psi|\omega_\eta) \geq S(\omega_\psi \circ T|\omega_\eta \circ T). \tag{2}$$

In quantum information theory, T is related to data processing, so (2) is sometimes called the data-processing inequality (DPI).

S plays an important role when characterizing the entanglement between subsystems. Over the years, several generalizations of S with different operational meaning have therefore been given, see e.g., [19]. One such generalization is the one-parameter family of "sandwiched relative Renyi divergences (entropies)" $D_s, s \in [1/2, 1) \cup (1, \infty)$ proposed by [32,38], which generalize the classical alpha-Renyi entropies and have several interesting properties. For example, like the relative entropy S, they can be defined for general von Neumann algebras [4,6], have an operational meaning [31], satisfy the DPI [5,6,15], and interpolate between S ($s \to 1$) and the fidelity F ($s \to 1/2$) [6,32,38].

The purpose of this note is to point out a related variational expression, $\Phi_s(\omega_\psi|\omega_\eta), s \in (1/2, 1)$, [Eq. (41)] inspired by a corresponding characterization of S due to Kosaki [23]. Φ_s reduces to a multiple of the fidelity in the classical case. Thus, it cannot be seen as a generalization of the classical Renyi entropies, but it is shown to provide an upper bound for the sandwiched relative Renyi entropies $D_s, s \in (1/2, 1)$. Furthermore, Φ_s is shown to have several other desirable properties. For example, it satisfies the DPI, and it is ordered with respect to the states that it depends on.

By construction, our divergence $\Phi_s(\omega_\psi|\omega_\eta)$ is defined for arbitrary von Neumann algebras, thus in particular type III. As is well-known [9], the algebras of observables in algebraic quantum field theory (QFT) [16] are of this type. In the second half of this note, we will give an application of Φ_s to QFT. We consider a QFT described by a Haag–Kastler net [16] $\mathcal{F} = \{\mathcal{F}(A)\}$ and a net of subfactors $\mathcal{A} = \{\mathcal{A}(A)\}$ in the sense of [28]. If A_n, B_n are disjoint regions separated by a corridor of size $\sim 1/n$, we can consider a conditional expectation "$E_n = E_{A_n} \otimes E_{B_n}$" projecting $\mathcal{F}(A_n) \vee \mathcal{F}(B_n)$ to $\mathcal{A}(A_n) \vee \mathcal{A}(B_n)$. The partial state of the vacuum with respect to the subsystem $\mathcal{F}(A_n) \vee \mathcal{F}(B_n)$ is called ω_Ω. We show [Theorem 1]

$$\lim_{n \to \infty} \Phi_s(\omega_\Omega|\omega_\Omega \circ E_n) = \ln[\mathcal{F} : \mathcal{A}], \tag{3}$$

which yields a formula (60) for F (fidelity) as a limiting case. Here, $[\mathcal{F} : \mathcal{A}]$ is the Jones index [21,24] of the subnet [28], whose values are restricted to $\{4\cos^2(\pi/n) :$

$n = 3, 4, \dots\} \cup [4, \infty]$. An example is a subtheory $\mathcal{A} \subset \mathcal{F}$ of charge neutral operators under a finite gauge group G, in which case $[\mathcal{F} : \mathcal{A}] = |G|$. Similar results could be obtained in analogous settings in higher dimensions.

We also point out a dual result for the inclusion $\mathcal{F}' \subset \mathcal{A}'$ and the dual conditional expectations E_n' in the case of the fidelity. This last result is a consequence of an "entropic (un)certainty relation", given in Corollaries 1, 3, which generalize a result by [30] to general von Neumann algebras. A noteworthy special case of Corollary 3 is the following. Consider a finite index inclusion $\mathcal{M} \supset \mathcal{N}$ of factors and $E : \mathcal{M} \to \mathcal{N}$ the corresponding minimal conditional expectation with dual conditional expectation $E' : \mathcal{N}' \to \mathcal{M}'$. Then, we have

$$F_{\mathcal{M}}(\omega_\psi | \omega_\psi \circ E) \cdot F_{\mathcal{N}'}(\omega_\psi' | \omega_\psi' \circ E') \geq \frac{1}{\sqrt{[\mathcal{M} : \mathcal{N}]}}. \tag{4}$$

Here, ω_ψ is the state on $\mathcal{M}$ induced by a vector $|\psi\rangle$ ("purification") in a natural cone of a standard representation of $\mathcal{M}$ and ω_ψ' the corresponding state on $\mathcal{N}'$ induced by the same vector. F is the fidelity between two states. Such relations remind one of the Heisenberg uncertainty principle. We plan to come back to this in the future.

Notations and conventions Calligraphic letters $\mathcal{A}, \mathcal{M}, \dots$ denote von Neumann algebras. Calligraphic letters $\mathscr{H}, \mathscr{K}, \dots$ denote linear spaces. We use the physicist's "ket"-notation $|\psi\rangle$ for vectors in a Hilbert space. The scalar product is written as $\langle\psi|\psi'\rangle$ and is anti-linear in the first entry. The norm of a vector is written simply as $\||\psi\rangle\| =: \|\psi\|$. Each vector $|\psi\rangle \in \mathscr{H}$ gives rise to a positive definite linear functional on the von Neumann algebra $\mathcal{M}$ acting on $\mathscr{H}$ via

$$\omega_\psi(m) = \langle\psi|m\psi\rangle, \quad m \in \mathcal{M}. \tag{5}$$

The commutant of $\mathcal{M}$ is denoted as $\mathcal{M}'$ and consists of those bounded operators commuting with all elements of $\mathcal{M}$.

2 von Neumann algebras and relative entropy

2.1 Relative modular theory and entropy

Let $(\mathcal{M}, J, \mathscr{P}_{\mathcal{M}}^\natural, \mathscr{H})$ be a von Neumann algebra in standard form acting on a Hilbert space $\mathscr{H}$, with natural cone $\mathscr{P}_{\mathcal{M}}^\natural$ and modular conjugation J (for an explanation of these terms, see [8,35] as general references). We will use relative modular operators $\Delta_{\psi,\zeta}$ associated with two vectors $|\zeta\rangle \in \mathscr{P}_{\mathcal{M}}^\natural$, $|\psi\rangle \in \mathscr{H}$. The nonnegative, self-adjoint operator $\Delta_{\psi,\zeta}$ is characterized by

$$J\Delta_{\psi,\zeta}^{1/2}(a|\zeta\rangle + |\chi\rangle) = \pi^{\mathcal{M}}(\zeta)a^*|\psi\rangle, \quad \forall\, a \in \mathcal{M}, \ |\chi\rangle \in (1 - \pi^{\mathcal{M}'}(\zeta))\mathscr{H}. \tag{6}$$

Here, $\pi^{\mathcal{M}'}(\psi)$ is the support projection of the vector $|\psi\rangle$, defined as the orthogonal projection onto the closure of the subspace $\mathcal{M}|\psi\rangle \subset \mathscr{H}$. The nonzero support of $\Delta_{\psi,\zeta}$ is $\pi^{\mathcal{M}}(\psi)\pi^{\mathcal{M}'}(\zeta)$, and complex powers such as $\Delta_{\psi,\zeta}^{z}$ are understood via the functional calculus on this support and are defined as 0 on $1 - \pi^{\mathcal{M}}(\psi)\pi^{\mathcal{M}'}(\zeta)$. $\Delta_{\psi,\zeta}^{1/2}$ depends on $|\psi\rangle$ only via the associated state functional (5).

According to [2,3], if the support projections satisfy $\pi^{\mathcal{M}}(\psi) \geq \pi^{\mathcal{M}}(\zeta)$, the relative entropy may be defined by

$$S(\zeta|\psi) = -\lim_{\alpha \to 0^+} \frac{\langle\zeta|\Delta_{\psi,\zeta}^{\alpha}\zeta\rangle - 1}{\alpha}, \tag{7}$$

; otherwise, it is by definition infinite. The relative entropy may be viewed as a function of the functionals (5) ω_ψ, ω_ζ on $\mathcal{M}$. So one can write instead also $S(\omega_\zeta|\omega_\psi)$ without ambiguity. In the case of the matrix algebra $\mathcal{M} = M_n(\mathbb{C})$, where ω_ζ and ω_ψ are identified with density matrices as $\omega_\psi(a) = \mathrm{Tr}(a\omega_\psi)$, etc., the relative entropy is the usual expression (1).

Kosaki [23] has given the following variational formula for two normalized state functionals ω_ψ, ω_ζ on $\mathcal{M}$:

$$S(\omega_\zeta|\omega_\psi) = \sup_{n\in\mathbb{N}} \sup_{x:(1/n,\infty)\to\mathcal{M}}$$
$$\left\{ \ln n - \int_{1/n}^{\infty} [\omega_\zeta(x(t)^*x(t)) + t^{-1}\omega_\psi(y(t)y(t)^*)]t^{-1}\mathrm{d}t \right\}, \tag{8}$$

where the second supremum is over all step functions $x(t)$ valued in $\mathcal{M}$ with finite range where $y(t) = 1 - x(t)$. (8) no longer makes explicit reference to modular theory, and the dependence on the state functionals (as opposed to vectors) is manifest. Some further explanations and uses of Kosaki's formula are discussed, e.g., in [33], ch. 5.

2.2 Conditional expectations, index, and relative entropy

Let $\mathcal{M}, \mathcal{N}$ be two von Neumann algebras. A linear operator $T : \mathcal{M} \to \mathcal{N}$ is called a channel if it is ultra-weakly continuous ("normal"), unital $T(1) = 1$, and completely positive. The latter means that the induced mapping $T \otimes id_n : \mathcal{M} \otimes M_n(\mathbb{C}) \to \mathcal{N} \otimes M_n(\mathbb{C})$, maps positive elements to positive elements for all $n \geq 1$.

If $\mathcal{N} \subset \mathcal{M}$ is a von Neumann sub-algebra, then a quantum channel $E : \mathcal{M} \to \mathcal{N}$ is called a conditional expectation if

$$E(n_1 m n_2) = n_1 E(m) n_2 \tag{9}$$

for $m \in \mathcal{M}, n_i \in \mathcal{N}$. The space of such conditional expectations is called $C(\mathcal{M}, \mathcal{N})$. A faithful normal operator valued weight is an unbounded and unnormalized positive linear map $N : \mathcal{M} \to \mathcal{N}$ with the same bimodule property and with dense domain $\mathcal{M}_+ (= \text{non-negative elements of } \mathcal{M})$ [17]. The space of such operator-valued weights

is denoted $P(\mathcal{M}, \mathcal{N})$, and clearly $C(\mathcal{M}, \mathcal{N})$ is a subset thereof. Both $C(\mathcal{M}, \mathcal{N})$ and $P(\mathcal{M}, \mathcal{N})$ may be empty.

Let $\mathcal{M}, \mathcal{N}$ be factors. If there exists $E \in C(\mathcal{M}, \mathcal{N})$, then the best constant $\lambda > 0$ such that

$$E(m^*m) \geq \lambda^{-1} m^* m \quad \text{for all } m \in \mathcal{M} \tag{10}$$

is called $ind(E)$, the index of E. If there is any conditional expectation at all, then there is one for which λ is minimal [18]. This $\lambda = [\mathcal{M} : \mathcal{N}]$ is the Jones–Kosaki index of the inclusion [21,24,34].

Haagerup [17] has established a canonical correspondence $N \in P(\mathcal{M}, \mathcal{N}) \leftrightarrow N^{-1} \in P(\mathcal{N}', \mathcal{M}')$ satisfying $(N^{-1})^{-1} = N$, $(N_1 \circ N_2)^{-1} = N_2^{-1} \circ N_1^{-1}$. One can connect this to the notion of a "spatial derivative" [10]. To this end, let $\mathcal{M}$ be a von Neumann algebra acting on $\mathscr{H}$, let $|\zeta\rangle, |\psi\rangle \in \mathscr{H}$, and let $B(\mathscr{H})$ be the von Neumann algebra of all bounded operators on $\mathscr{H}$. Applying (5) to $\mathcal{M}$ and the commutant $\mathcal{M}'$, we get state functionals ω_ζ', respectively, ω_ψ on $\mathcal{M}'$, respectively, $\mathcal{M}$. Now, the functional $\omega_\psi : \mathcal{M} \to \mathbb{C}$ is a special case of a conditional expectation, so the dual conditional expectation ω_ψ^{-1} is in $P(B(\mathscr{H}), \mathcal{M}')$. Thus, $\omega_\zeta' \circ \omega_\psi^{-1}$ is a weight on $B(\mathscr{H})$. Such a weight defines a densely defined positive definite (sesquilinear) quadratic form on $\mathscr{H}$ by

$$q_{\psi,\zeta}(\phi_1, \phi_2) = \omega_\zeta' \circ \omega_\psi^{-1}(|\phi_2\rangle\langle\phi_1|), \tag{11}$$

and the operator on $\mathscr{H}$ representing $q_{\psi,\zeta}$ is called the "spatial derivative," $\Delta_{\mathcal{M}}(\omega_\zeta'/\omega_\psi)$. It can be seen to only depend on the functionals ω_ζ', respectively, ω_ψ on $\mathcal{M}'$, respectively, $\mathcal{M}$. $\Delta_{\mathcal{M}}(\omega_\zeta'/\omega_\psi)$ equals the relative modular operator $\Delta_{\mathcal{M};\zeta,\psi}$ in case $|\psi\rangle \in \mathscr{P}_{\mathcal{M}}$. It follows that if $|\zeta\rangle$ is in the form domain of $\ln \Delta_{\mathcal{M}}(\omega_\zeta'/\omega_\psi)$, then the relative entropy may also be written as $S(\zeta|\psi) = \langle\zeta| \ln \Delta_{\mathcal{M}}(\omega_\zeta'/\omega_\psi)\zeta\rangle$. This representation and the structures established by [10,17] have an immediate corollary for a conditional expectation $E : \mathcal{M} \to \mathcal{N}$. First, by [10], thm. 9, the spatial derivative has the duality property

$$\Delta_{\mathcal{M}}(\omega_\zeta'/\omega_\psi) = \Delta_{\mathcal{M}'}(\omega_\psi/\omega_\zeta')^{-1}. \tag{12}$$

Furthermore, $\omega_\psi' \circ (\omega_\psi \circ E)^{-1} = (\omega_\psi' \circ E^{-1}) \circ \omega_\psi^{-1}$, so [24]

$$\Delta_{\mathcal{M}}(\omega_\psi'/\omega_\psi \circ E) = \Delta_{\mathcal{N}}(\omega_\psi' \circ E^{-1}/\omega_\psi) = \Delta_{\mathcal{N}'}(\omega_\psi/\omega_\psi' \circ E^{-1})^{-1}. \tag{13}$$

Taking a log and the expectation value with respect to the vector $|\psi\rangle$ then gives:

$$S_{\mathcal{M}}(\omega_\psi|\omega_\psi \circ E) + S_{\mathcal{N}'}(\omega_\psi'|\omega_\psi' \circ E^{-1}) = 0. \tag{14}$$

Note that E^{-1} is not normalized unless $E = id$. If $\mathcal{M}$ is a factor such that $ind(E) = \lambda < \infty$ is finite, then it can be shown from (10) that 1 is in the domain of E^{-1} and $\lambda 1 = E^{-1}(1)$. Therefore,

$$E' = \lambda^{-1} E^{-1} \tag{15}$$

is a (normalized) conditional expectation $E' \in C(\mathcal{N}', \mathcal{M}')$ [24]. In fact, if E is minimal, then also E' is. Using the standard scaling properties of the relative entropy thereby gives the following trivial corollary which generalizes [30] who have considered by an explicit method the special case of finite dimensional type I von Neumann algebras:

Corollary 1 *Let $\mathcal{N} \subset \mathcal{M}$ be a an inclusion of von Neumann factors with finite index $[\mathcal{M} : \mathcal{N}] < \infty$, acting standardly on a Hilbert space $\mathscr{H}$. Assume that $E \in C(\mathcal{M}, \mathcal{N})$ is the minimal conditional expectation, $E' \in C(\mathcal{N}', \mathcal{M}')$ the dual minimal conditional expectation. For $|\psi\rangle \in \mathscr{H}$, we have*

$$S_{\mathcal{M}}(\omega_\psi | \omega_\psi \circ E) + S_{\mathcal{N}'}(\omega'_\psi | \omega'_\psi \circ E') = \ln[\mathcal{M} : \mathcal{N}]. \tag{16}$$

(Note that ω'_ψ in the second expression means the functional (5) on $\mathcal{N}'$, etc.)

Results of a similar flavor have also been given by [39]. Interesting physical applications of the above "certainty relation" (16) involving Wilson–'t Hooft operators in four-dimensional quantum Yang-Mills theory have recently been pointed out by [12,30]. In such a situation, the algebras are expected to be of type III [9]. Then, the minimal conditional expectation E and its dual E' can be described more explicitly using Q-systems [27]. In this framework, $\mathcal{M}$ is generated by $\mathcal{N}$ together with a single operator, v, [see "Appendix", Eq. (77)] and $\mathcal{N}'$ is generated by $\mathcal{M}'$ together with a single operator, v'. The operators $w = j_{\mathcal{N}}(v') \in \mathcal{N}$, $w' = j_{\mathcal{M}}(v) \in \mathcal{M}'$ and the "canonical" endomorphsms

$$\gamma = j_{\mathcal{N}} j_{\mathcal{M}} : \mathcal{M} \to \mathcal{N}, \quad \gamma' = j_{\mathcal{M}} j_{\mathcal{N}} : \mathcal{N}' \to \mathcal{M}' \tag{17}$$

can be defined, where $j_{\mathcal{N}}(n) = J_{\mathcal{N}} n J_{\mathcal{N}}$ and $J_{\mathcal{N}}$ is the modular conjugation[1] of $\mathcal{N}$, etc. The expectations E, E' are then given by (here $d = [\mathcal{M} : \mathcal{N}]^{1/2}$)

$$E(m) = \frac{1}{d} w^* \gamma(m) w, \quad E'(n') = \frac{1}{d} w'^* \gamma'(n') w'. \tag{18}$$

Another property is that $J_{\mathcal{M}} v' = v' J_{\mathcal{N}}$, $J_{\mathcal{M}} v = v J_{\mathcal{N}}$.

The operator v' is closely related to the idea of quantum error correcting codes as described by [14]: For the sake of easier comparison, define

$$V := v'/\sqrt{d}, \quad V' := v/\sqrt{d}, \tag{19}$$

with the normalizations made such that V, V' are isometries. For any $|\psi\rangle, |\zeta\rangle \in \mathscr{H}$ we define states ω_ζ on $\mathcal{N}$ and ω'_ζ on $\mathcal{N}'$ by (5). Then, we have the implications

$$\begin{cases} \omega'_\zeta|_{\mathcal{N}'} = \omega'_\psi|_{\mathcal{N}'} \implies \omega'_{V\zeta}|_{\mathcal{M}'} = \omega'_{V\psi}|_{\mathcal{M}'} \\ \omega_\zeta|_{\mathcal{N}} = \omega_\psi|_{\mathcal{N}} \implies \omega_{V\zeta}|_{\mathcal{M}} = \omega_{V\psi}|_{\mathcal{M}}, \end{cases} \tag{20}$$

---

[1] With respect to a fixed natural cone $\mathscr{P}_{\mathcal{N}}^\sharp$.

so $\mathcal{M}$ is "standardly c-reconstructible" from $\mathcal{N}$ in the terminology [14]. In the context of holography, $\mathcal{N}$ would be a bulk observable algebra, $\mathcal{M}$ a corresponding CFT algebra and the subspace $V\mathcal{H} \subset \mathcal{H}$ the "code subspace." Dually, the operator V' is used in a similar way to "standardly c-reconstruct" $\mathcal{N}'$ from $\mathcal{M}'$, with similar relations. While the existence and properties of the operator V are equivalent to the existence of some conditional expectation $E : \mathcal{M} \to \mathcal{N}$ alone [14], thm. 7, the existence of the operator V' for the dual code does not follow from these results but requires a finite index (and minimal conditional expectation).

These facts can be used to give an "error correction version" of the certainty relation expressed by Corollary 1. We simply observe the equalities

$$E(m) = \frac{1}{d} w^* \gamma(m) w = \frac{1}{d} j_{\mathcal{N}}(v'^*) j_{\mathcal{N}} j_{\mathcal{M}}(m) j_{\mathcal{N}}(v') = \frac{1}{d} J_{\mathcal{N}} v'^* J_{\mathcal{M}} m J_{\mathcal{M}} v' J_{\mathcal{N}} = V^* m V \quad (21)$$

we get $E'(n') = V'^* n' V'$ for $n' \in \mathcal{N}'$. This gives in view of Corollary 1:

Corollary 2 (Error correcting code version) *Let $\mathcal{M} \supset \mathcal{N}$ be an inclusion of type III von Neumann factors with finite index and let $|\psi\rangle \in \mathcal{H}$. Let V be a code operator as in (20) and V' the dual code operator. Then,*

$$S_{\mathcal{M}}(\omega_\psi | \omega_{V\psi}) + S_{\mathcal{N}'}(\omega'_\psi | \omega'_{V'\psi}) = \ln[\mathcal{M} : \mathcal{N}]. \quad (22)$$

2.3 Sandwiched Renyi divergence

A family of entropy functionals for von Neumann algebras extrapolating the relative entropy are the "sandwiched Renyi divergences (entropies)" [32,38]. In the general von Neumann algebra setting, they can be defined in terms of certain L_p norms [5,6], which were defined by [4] (see also [6,20]) relative to a fixed cyclic and separating vector $|\psi\rangle \in \mathcal{H}$ in the natural cone of a standard representation of a von Neumann algebra $\mathcal{M}$.

More precisely, for $1 \le p \le 2$, $L_p(\mathcal{M}, \psi)$ is defined as the completion of $\mathcal{H}$ with respect to the following pseudo-norm[2]:

$$\|\zeta\|_{p,\psi} = \inf\{\|\Delta_{\phi,\psi}^{(1/2)-(1/p)} \zeta\| : \|\phi\| = 1, \pi^{\mathcal{M}}(\phi) = \pi^{\mathcal{M}}(\zeta)\}. \quad (23)$$

We deviate slightly from [4] by imposing $\pi^{\mathcal{M}}(\phi) = \pi^{\mathcal{M}}(\zeta)$ instead of $\pi^{\mathcal{M}}(\phi) \ge \pi^{\mathcal{M}}(\zeta)$, but this makes no difference for the resulting formula in the case of finite-dimensional von Neumann algebras, nor the results that we invoke below. It is well-known that any state $|\zeta'\rangle \in \mathcal{H}$ such that $\omega'_{\zeta'} = \omega'_\zeta$ as state functionals on $\mathcal{M}'$ is related to $|\zeta\rangle \in \mathcal{H}$ by $|\zeta'\rangle = u|\zeta\rangle$, where $u \in \mathcal{M}$, $u^* u = \pi^{\mathcal{M}}(\zeta)$. Furthermore, it is known [4], thm. 5 that $\|m\zeta\|_{p,\psi} \le \|m\|\|\zeta\|_{p,\psi}$ for any $m \in \mathcal{M}$, and from this one can easily see that $\|\zeta\|_{p,\psi}$ depends on $|\zeta\rangle \in \mathcal{H}$ only via the induced functional ω'_ζ on $\mathcal{M}'$. Vice versa, to get an invariant for state functionals on $\mathcal{M}$, one should therefore consider the L_p norms relative to $\mathcal{M}'$ as in the following definition.

---

[2] If $|\psi\rangle$ is such that $|\zeta\rangle$ is not in the domain of $\Delta_{\phi,\psi}^{(1/2)-(1/p)}$, we set $\|\Delta_{\phi,\psi}^{(1/2)-(1/p)} \zeta\| = \infty$.

Definition 1 Let $\mathcal{M}$ be a von Neumann algebra in standard form acting on $\mathcal{H}$, $|\zeta\rangle \in \mathcal{H}$. The "sandwiched Renyi divergences" [32,38] D_s, $s \in [1/2, 1)$ are defined by

$$D_s(\omega_\zeta|\omega_\psi) = (s-1)^{-1} \ln \|\zeta\|_{2s,\psi,\mathcal{M}'}^{2s} \tag{24}$$

with norm taken relative to $\mathcal{M}'$.

The L_p norms for $p > 2$ and corresponding sandwiched Renyi entropies for $s > 1$ may be defined by duality [4], thm. 1, but the precise definition is not needed for the purposes of this paper. The generalization of the definition to non-faithful state functionals ω_ψ whose representing vector $|\psi\rangle$ is not separating is given in [6], who also prove key properties of D_s [5,15,32,38] in the general von Neumann setting. States on the finite-dimensional type I factor $\mathcal{A} = M_n(\mathbb{C})$ correspond to density matrices via $\omega_\psi(a) = \text{Tr}(a\omega_\psi)$. In that case, the definition gives

$$D_s(\omega_\zeta|\omega_\psi) = (s-1)^{-1} \ln \text{Tr}(\omega_\psi^{(1-s)/(2s)} \omega_\zeta \omega_\psi^{(1-s)/(2s)})^s. \tag{25}$$

The sandwiched Renyi divergences extrapolate the relative entropy S which can be recovered as the limit $s \to 1^-$ [6], thm. 13. At the other end, for $s \to 1/2^+$, one recovers the negative log squared fidelity. In fact, the L_1 norm relative to $\mathcal{M}'$ is related to the fidelity [1,37] relative to $\mathcal{M}$ by

$$\|\zeta\|_{1,\psi,\mathcal{M}'} = \sup\{|\langle\zeta|a'\psi\rangle| : a' \in \mathcal{M}', \|a'\| = 1\} = F_\mathcal{M}(\omega_\zeta|\omega_\psi), \tag{26}$$

see [13], lem. 3 (1), which generalizes [4], lem. 5.3 when ψ is not necessarily faithful.

It has been shown that $D_s \leq S$ for $s \leq 1$ by [6], prop. 4. Therefore, Corollary 1 implies:

Corollary 3 *For a finite index inclusion $\mathcal{N} \subset \mathcal{M}$ of von Neumann factors with minimal conditional expectation $E : \mathcal{M} \to \mathcal{N}$, we have*

$$D_s^\mathcal{M}(\omega_\psi|\omega_\psi \circ E) + D_s^{\mathcal{N}'}(\omega_\psi'|\omega_\psi' \circ E') \leq \ln[\mathcal{M} : \mathcal{N}] \tag{27}$$

for any $s \in [1/2, 1)$ and any vector $|\psi\rangle \in \mathcal{H}$ with induced state functionals ω_ψ' on $\mathcal{N}'$ and ω_ψ on $\mathcal{M}$.

A noteworthy special case arises for $s = 1/2$:

$$F_\mathcal{M}(\omega_\psi|\omega_\psi \circ E) \cdot F_{\mathcal{N}'}(\omega_\psi'|\omega_\psi' \circ E') \geq \frac{1}{\sqrt{[\mathcal{M} : \mathcal{N}]}}. \tag{28}$$

There are also evident error correcting code formulations of this analogous to Corollary 2.

3 Variational formulas

Here, we point out a variational expression related to the L_p norms in the range $p \in (1, 2)$ similar to Kosaki's formula [23] for the relative entropy. To simplify the discussion, we assume that the fixed reference vector $|\psi\rangle \in \mathscr{P}^\natural_{\mathcal{M}}$ in the definition of the L_p norm is cyclic and separating for the von Neumann algebra $\mathcal{M}$ in standard form acting on $\mathscr{H}$. Let $\Delta' \equiv \Delta'_{\psi,\phi}$ be the relative modular operator defined using $\mathcal{M}'$ and the unit vector $|\phi\rangle \in \mathscr{P}^\natural_{\mathcal{M}}$. Then, the support of Δ' is $\pi^{\mathcal{M}}(\phi)$. $\Delta'^{1/2}$ is a nonnegative self-adjoint operator whose domain we denote by $\mathscr{D}(\phi)$ (it is equal the closure in the graph norm of $\mathcal{M}'|\phi\rangle \oplus (1 - \pi^{\mathcal{M}}(\phi))\mathscr{H}$).

We can apply [33], lem. 5.8, showing that, for $t > 0$, $|\zeta\rangle \in \mathscr{D}(\phi)$, we have

$$\langle \Delta'(\Delta' + t)^{-1}\zeta|\zeta\rangle = \inf\{\|\xi\|^2 + t^{-1}\|\Delta'^{1/2}\eta\|^2 : |\xi\rangle + |\eta\rangle = |\zeta\rangle, \quad |\xi\rangle, |\eta\rangle \in \mathscr{D}(\phi)\}. \quad (29)$$

Now, we note the well-known formula

$$\lambda^\alpha = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \frac{\lambda}{t + \lambda} t^{\alpha-1} \mathrm{d}t \quad (30)$$

when $\lambda > 0, \alpha \in (0, 1)$, which is commonly used to trade the power Δ'^α for the resolvent $\Delta'(\Delta' + t)^{-1}$, for which we have the variational expression (29). Then, arguing in the same way as in the proof of [33], prop. 5.10 gives

$$\|\Delta'^{\alpha/2}\zeta\|^2 = \frac{\sin(\pi\alpha)}{\pi} \inf_{\xi,\eta:\mathbb{R}_+ \to \mathscr{D}(\phi)} \int_0^\infty [\|\xi(t)\|^2 + t^{-1}\|\Delta'^{1/2}\eta(t)\|^2]t^{\alpha-1}\mathrm{d}t, \quad (31)$$

where $|\zeta\rangle \in \mathscr{D}(\phi)$ and where the infimum is taken over all step functions $\xi, \eta : [0, \infty] \to \mathscr{D}(\phi)$ with finite range such that $|\xi(t)\rangle = |\zeta\rangle$ for sufficiently small $t > 0$, such that $|\eta(t)\rangle = |\zeta\rangle$ for sufficiently large t, and such that $|\xi(t)\rangle + |\eta(t)\rangle = |\zeta\rangle$. The support of Δ' is $\pi^{\mathcal{M}}(\phi)$, thus we have $\|\Delta'^{1/2}\eta(t)\| = \|\Delta'^{1/2}\pi^{\mathcal{M}}(\phi)\eta(t)\|$. Taking into account the definition of the L_1 norm relative to $\mathcal{M}$ and the relation $\Delta^{-1}_{\phi,\psi} = \Delta'_{\psi,\phi}$, see e.g., [4], thm. C. 1, we get $\|\Delta'^{1/2}\eta(t)\| \geq \|\pi^{\mathcal{M}}(\phi)\eta(t)\|_{1,\psi,\mathcal{M}}$. The last expression is also equal to the fidelity $F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\eta(t)}|\omega'_\psi)$ relative to $\mathcal{M}'$ by (26). Taking furthermore into account the trivial fact that $\|\pi^{\mathcal{M}}(\phi)\xi(t)\| \leq \|\xi(t)\|$ we get

$$\|\Delta'^{\alpha/2}\zeta\|^2 \geq \frac{\sin(\pi\alpha)}{\pi} \inf_{\xi,\eta:\mathbb{R}_+ \to \mathscr{D}(\phi)} \int_0^\infty$$
$$[\omega'_{\pi^{\mathcal{M}}(\phi)\xi(t)}(1) + t^{-1}F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\eta(t)}|\omega'_\psi)^2]t^{\alpha-1}\mathrm{d}t, \quad (32)$$

where still $|\zeta\rangle \in \mathscr{D}(\phi)$ and the infimum is still taken over all step functions as described above. We want to enlarge the domain of $|\zeta\rangle$ for which (32) is valid.

Lemma 1 (32) *holds for* $|\zeta\rangle$ *in the domain of* $\Delta'^{\alpha/2}$ *provided that we take the infimum over functions* ξ, η *that are now valued in* $\mathscr{H}$ *and have analogous properties otherwise.*

Proof First, by applying an increasing sequence of spectral projections we define $|\zeta_n\rangle := E_{[0,n]}(\Delta')|\zeta\rangle \in \mathscr{D}(\phi)$. Then, $|\zeta_n\rangle \to |\zeta\rangle$, $\Delta'^{\alpha/2}|\zeta_n\rangle \to \Delta'^{\alpha/2}|\zeta\rangle$ strongly, and the above inequality (32) holds for $|\zeta_n\rangle$.

We next consider a sufficiently large n such that $\|\zeta - \zeta_n\|$ and $\|\Delta'^{\alpha/2}\zeta - \Delta'^{\alpha/2}\zeta_n\|$ are $< \varepsilon$. For this fixed n, we then consider step functions $\xi_n, \eta_n : [0, \infty] \to \mathscr{D}(\phi)$ with finite range such that $|\xi_n(t)\rangle = |\zeta_n\rangle$ for sufficiently small $t > 0$, such that $|\eta_n(t)\rangle = |\zeta_n\rangle$ for sufficiently large t, such that $|\xi_n(t)\rangle + |\eta_n(t)\rangle = |\zeta_n\rangle$, and such that the infimum on the right side of (32), applied to $|\zeta_n\rangle$, is nearly achieved by the functions ξ_n, η_n up to a small tolerance, ε. We also define new functions $\tilde{\xi}_n, \tilde{\eta}_n$ by

$$|\tilde{\eta}_n(t)\rangle = \begin{cases} |\eta_n(t)\rangle & \text{for } t < 1, \\ |\zeta\rangle - |\xi_n(t)\rangle & \text{for } t \geq 1, \end{cases} \quad |\tilde{\xi}_n(t)\rangle = \begin{cases} |\xi_n(t)\rangle & \text{for } t \geq 1, \\ |\zeta\rangle - |\eta_n(t)\rangle & \text{for } t < 1. \end{cases} \quad (33)$$

The new functions have analogous properties as the old ones with $|\zeta_n\rangle$ replaced by $|\zeta\rangle$. Furthermore, for $t < 1$, we have $\|\tilde{\eta}_n(t) - \eta_n(t)\| = 0$, $\|\tilde{\xi}_n(t) - \xi_n(t)\| = \|\zeta - \zeta_n\|$, whereas for $t \geq 1$, we have $\|\tilde{\xi}_n(t) - \xi_n(t)\| = 0$, $\|\tilde{\eta}_n(t) - \eta_n(t)\| = \|\zeta - \zeta_n\|$. We now claim that

$$\|\Delta'^{\alpha/2}\zeta\|^2 \geq \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty [\omega'_{\pi^{\mathcal{M}}(\phi)\tilde{\xi}_n(t)}(1) + t^{-1}F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\tilde{\eta}_n(t)}|\omega'_\psi)^2]t^{\alpha-1}\mathrm{d}t - c\varepsilon, \quad (34)$$

constant only depending on α, $|\zeta\rangle$ but not n.

By construction such an equation holds with $c = 1$ for $|\zeta_n\rangle$ and the functions ξ_n, η_n, and we want to deduce (34) from that. Also, by construction $\|\Delta'^{\alpha/2}\zeta\|^2$ differs from $\|\Delta'^{\alpha/2}\zeta_n\|^2$ by $c\varepsilon$. Then, if we split the integral in (34) into a contribution from $t \in [0, 1)$ and $t \in [1, \infty)$, we can see that the integral for the functions ξ_n, η_n is equal to the integral for the functions $\tilde{\xi}_n, \tilde{\eta}_n$ up to an error which is bounded in terms of the sum of the integrals

$$\int_0^1 |\|\pi^{\mathcal{M}}(\phi)\xi_n(t)\|^2 - \|\pi^{\mathcal{M}}(\phi)\tilde{\xi}_n(t)\|^2|t^{\alpha-1}\mathrm{d}t,$$
$$\int_1^\infty |F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\eta_n(t)}|\omega'_\psi)^2 - F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\tilde{\eta}_n(t)}|\omega'_\psi)^2|t^{\alpha-2}\mathrm{d}t, \quad (35)$$

times numerical coefficients depending only on α. The second integral is shown to be of the order $\|\zeta - \zeta_n\| < \varepsilon$ as follows. The fidelity is continuous under strong limits (see e.g. [13], lem. 11) which in the case at hand gives

$$|F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\eta_n(t)}|\omega'_\psi) - F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\tilde{\eta}_n(t)}|\omega'_\psi)| \leq \|\tilde{\eta}_n(t) - \eta_n(t)\|, \quad (36)$$

where we can use $\|\tilde{\eta}_n(t) - \eta_n(t)\| = \|\zeta - \zeta_n\| < \varepsilon$. In the second integral, we actually have the squares of the fidelities, so we use the elementary identity $|a^2 - \tilde{a}^2| \leq |a - \tilde{a}|^2 + 2|a||a - \tilde{a}|$, where a and $\tilde{a}$ are the fidelities under the second integral (35) associated with η_n and $\tilde{\eta}_n$. For the term corresponding to $2|a||a - \tilde{a}|$, we then need a bound on $\int_1^\infty F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\eta_n(t)}|\omega'_\psi)t^{\alpha-2}\mathrm{d}t$. Such a bound can be obtained

immediately from (32), the Cauchy–Schwarz inequality, and the construction of the functions ξ_n, η_n:

$$\left(\int_1^\infty F_{\mathcal{M}'}(\omega'_{\pi^{\mathcal{M}}(\phi)\eta_n(t)} | \omega'_\psi) t^{\alpha-2} \mathrm{d}t \right)^2 \leq c(\|\Delta'^{\alpha/2} \zeta_n\|^2 + \varepsilon). \tag{37}$$

Then, by construction $\|\Delta'^{\alpha/2} \zeta_n\| \leq \varepsilon + \|\Delta'^{\alpha/2} \zeta\|$ which in turn is uniformly bounded in n. Also the first integral in (35) is shown to be of the order $\|\zeta - \zeta_n\| < \varepsilon$. Here, we use first the reverse triangle inequality

$$\big| \|\pi^{\mathcal{M}}(\phi) \xi_n(t)\| - \|\pi^{\mathcal{M}}(\phi) \tilde{\xi}_n(t)\| \big| \leq \|\tilde{\xi}_n(t) - \xi_n(t)\|, \tag{38}$$

where we can use $\|\tilde{\xi}_n(t) - \xi_n(t)\| = \|\zeta - \zeta_n\| < \varepsilon$. In the first integral, we actually have the squares of the norms, so we use the elementary identity $|a^2 - \tilde{a}^2| \leq |a - \tilde{a}|^2 + 2|a||a - \tilde{a}|$, where a and $\tilde{a}$ are now the norms under the first integral (35) associated with ξ_n and $\tilde{\xi}_n$. For the term corresponding to $2|a||a - \tilde{a}|$, we then need a bound on $\int_0^1 \|\pi^{\mathcal{M}}(\phi) \xi_n(t)\| t^{\alpha-2} \mathrm{d}t$. Such a bound can be obtained immediately from (32), the Cauchy–Schwarz inequality, and the construction of the functions ξ_n, η_n:

$$\left(\int_0^1 \|\pi^{\mathcal{M}}(\phi) \xi_n(t)\| t^{\alpha-1} \mathrm{d}t \right)^2 \leq c(\|\Delta'^{\alpha/2} \zeta_n\|^2 + \varepsilon). \tag{39}$$

Again, $\|\Delta'^{\alpha/2} \zeta_n\|$ is uniformly bounded in n. Thus, we see that the integrals (35) have an upper bound of the form $c\varepsilon$. Hence, our argument implies that (34) holds. Trivially, the inequality (32) then also holds up to a tolerance of order $c\varepsilon$ if we take the infimum over the set of all step functions $\xi, \eta : [0, \infty] \to \mathcal{H}$ with finite range such that $|\xi(t)\rangle = |\zeta\rangle$ for sufficiently small $t > 0$ such that $|\eta(t)\rangle = |\zeta\rangle$ for sufficiently large t, and such that $|\xi(t)\rangle + |\eta(t)\rangle = |\zeta\rangle$. Since ε was arbitrarily small, the statement follows. $\qquad \square$

Suppose now that $\pi^{\mathcal{M}}(\phi) = \pi^{\mathcal{M}}(\zeta)$. The structure of the right side of (32) implies range of the step functions may be restricted to $\pi^{\mathcal{M}}(\phi)\mathcal{H} = \pi^{\mathcal{M}}(\zeta)\mathcal{H}$. Suppose we have step functions $\xi, \eta : [0, \infty] \to \pi^{\mathcal{M}}(\zeta)\mathcal{H}$ such that the infimum in (32) is attained up to a small tolerance. Then, we can pick step functions $x', y' : [0, \infty] \to \mathcal{M}'$ such that $x'(t) + y'(t) = 1$, $x'(t) = 1$ for sufficiently small $t > 0$, $y'(t) = 1$ for sufficiently large $t > 0$, and such that $\|x'(t)\zeta - \xi(t)\|$, $\|y'(t)\zeta - \eta(t)\|$ is small for all $t > 0$. Because the fidelity is continuous, see e.g., [13], lem. 11, we may replace the infimum in (32) by the infimum over all functions $|\xi(t)\rangle, |\eta(t)\rangle$ of the form $x'(t)|\zeta\rangle, y'(t)|\zeta\rangle$, where $x'(t) + y'(t) = 1$, $x', y' : \mathbb{R}_+ \to \mathcal{M}'$ step functions with finite range such that $x'(t) = 1$ for sufficiently small $t > 0$ and $y'(t) = 1$ for sufficiently large $t > 0$. Taking $\alpha = 2/p - 1 \in (0, 1)$, and defining p', c_p as in (42), this yields

$$\|\Delta'^{(1/p)-(1/2)}\zeta\|^2 \geq c_p \inf_{x',y':\mathbb{R}_+\to\mathcal{M}'} \int_0^\infty [\omega'_\zeta(x'(t)^*x'(t))$$
$$+ t^{-1} F_{\mathcal{M}'}(y'(t)\omega'_\zeta y'(t)^*|\omega'_\psi)^2] t^{-2/p'} dt, \tag{40}$$

for all $|\zeta\rangle$ in the domain of $\Delta'^{(1/p)-(1/2)}$ (which depends upon the unit vector $|\phi\rangle \in \mathscr{P}_{\mathcal{M}}^\natural$) such that $\pi^{\mathcal{M}}(\phi) = \pi^{\mathcal{M}}(\zeta)$.

We now take into account the relation $\Delta_{\phi,\psi}^{-1} = \Delta'_{\psi,\phi}$, see e.g., [4], thm. C. 1, on the left side of (40). We note that the right side no longer depends on the choice of $|\phi\rangle$, whereas on the the left side, we can drop the condition that $|\phi\rangle$ is in the natural cone, because $\Delta_{\phi,\psi}$ is unchanged if we replace $|\phi\rangle$ by $u'|\phi\rangle$, $u' \in \mathcal{M}'$, $u'^*u' = \pi^{\mathcal{M}'}(\phi)$ and such a replacement preserves $\pi^{\mathcal{M}}(\phi) = \pi^{\mathcal{M}}(\zeta)$.

Then, we minimize the left side of (40) for fixed $|\zeta\rangle$ over unit vectors $|\phi\rangle \in \mathscr{H}$ such that $\pi^{\mathcal{M}}(\phi) = \pi^{\mathcal{M}}(\zeta)$ and such that $|\zeta\rangle$ is in the domain of $\Delta_{\phi,\psi}^{(1/2)-(1/p)}$, which gives the L_p norm (23) of $|\zeta\rangle$ relative to $\mathcal{M}$. As a consequence, the following proposition follows after reversing the roles of $\mathcal{M}$ and $\mathcal{M}'$:

Proposition 1 *Let $\mathcal{M}$ be a von Neumann algebra in standard form with cyclic and separating vector $|\psi\rangle$ in the natural cone. For any $1 < p < 2$, and $|\zeta\rangle \in \mathscr{H}$, we have the variational formula*

$$\|\zeta\|_{p,\psi,\mathcal{M}'}^2 \geq c_p \inf_{x:\mathbb{R}_+\to\mathcal{M}} \int_0^\infty [\omega_\zeta(x(t)^*x(t)) + t^{-1}F_\mathcal{M}(y(t)\omega_\zeta y(t)^*|\omega_\psi)^2]t^{-2/p'} dt, \tag{41}$$

for the L_p-norm relative to $\mathcal{M}'$ and $|\psi\rangle$. $F_\mathcal{M}$ is the fidelity relative to $\mathcal{M}$,

$$c_p = -\frac{\sin(2\pi/p)}{\pi} > 0, \quad \frac{1}{p} + \frac{1}{p'} = 1, \tag{42}$$

$y(t) = 1 - x(t)$, $x : \mathbb{R}_+ \to \mathcal{M}$ a step function with finite range such that $x(t) = 1$ for sufficiently small $t > 0$ and $x(t) = 0$ for sufficiently large $t > 0$, and we use the notation $(x\omega x^*)(a) = \omega(x^*ax)$.

Remarks Both sides of the inequality (41) only depend on $|\zeta\rangle \in \mathscr{H}$ via the functional ω_ζ on $\mathcal{M}$.

We will now start to investigate the variational expression in the proposition in its own right. For easier reference, we make the following definition where p corresponds to $2s$.

Definition 2 Let $\mathcal{M}$ be a von Neumann algebra in standard form acting on $\mathcal{H}$, $s \in (1/2, 1)$. The "generalized fidelity" is defined by

$$\Phi_s(\omega_\zeta|\omega_\psi) = \ln\left\{c_{2s} \inf_{x:\mathbb{R}_+\to\mathcal{M}} \int_0^\infty [\omega_\zeta(x(t)^*x(t)) + t^{-1}F(y(t)\omega_\zeta y(t)^*|\omega_\psi)^2]t^{\frac{s-1}{s}}\frac{dt}{t}\right\}^{\frac{s}{s-1}} \tag{43}$$

with the infimum and notations as defined in Proposition 1.

Remarks (1) The normalizations of Φ_s are chosen in such a way that $\Phi_s \geq D_s$ by Proposition 1.

(2) The terminology "generalized fidelity" is due to the following observation. Consider $\mathcal{M} = M_n(\mathbb{C})$ and diagonal (normalized) density matrices $\omega_\zeta = diag(p_1, \ldots, p_n)$, $\omega_\psi = diag(q_1, \ldots, q_n)$. We use the abbreviation $F = F(\omega_\zeta | \omega_\psi) = \sum_i \sqrt{p_i q_i}$ for the fidelity. By considering the variational expression in the definition of Φ_s with diagonal $x(t) = diag(x_1(t), \ldots, x_n(t))$, one can easily convince oneself that the infimum can be reached by approximations of

$$x_i(t) = \sqrt{\frac{q_i}{p_i}} \frac{F}{t+1} \tag{44}$$

by step functions. Inserting this into the variational formula one gets $\Phi_s \geq -\frac{s}{1-s} \ln F^2$. Corollary 6 shows that an inequality of this type with a worse constant is true generally. On the other hand, as given in Corollary 6, we always have the reverse inequality which implies that $\Phi_s = -\frac{s}{1-s} \ln F^2$ in the present case. This becomes (minus log of) the squared fidelity when $s = 1/2$.

Using this formula for Φ_s in the commutative case, the inequality $\Phi_s \geq D_s$ (Proposition 1) is seen to be equivalent to Hölder's inequality applied to the fidelity $F = \sum_i \sqrt{p_i q_i}$. Indeed, taking the Hölder exponents to be $p = 2s$, $p' = (2s)/(2s-1)$, one has

$$\sum_i \sqrt{p_i q_i} = \sum_i (p_i^{1/2} q_i^{(1-s)/(2s)}) q_i^{(2s-1)/(2s)} \leq \left(\sum_i p_i^s q_i^{1-s} \right)^{1/(2s)} \tag{45}$$

using also $\sum_i q_i = 1$. Then, applying log to this inequality and using that $D_s = (s-1)^{-1} \ln \sum_i p_i^s q_i^{1-s}$ in the commutative case, we get $-\frac{s}{1-s} \ln F^2 = \Phi_s \geq D_s$, i.e., Proposition 1 in the commutative case.

(3) The properties shown in the following indicate that Φ_s has many of the desired properties of a divergence. To the best of our knowledge, Φ_s is a new generalization of the negative log squared fidelity.

We now investigate some properties of Φ_s. First, consider $|\zeta_1\rangle$, $|\zeta_2\rangle$ such that $\omega_{\zeta_1} \leq \omega_{\zeta_2}$ in the sense of functionals on the von Neumann algebra $\mathcal{M}$. It is well-known that such a condition implies the existence of $a' \in \mathcal{M}'$ such that $|\zeta_1\rangle = a'|\zeta_2\rangle$ and $\|a'\| \leq 1$. Then, (26) immediately gives:

$$\begin{aligned} F_{\mathcal{M}}(y\omega_{\zeta_1} y^*, \omega_\psi) &= \sup\{|\langle y\zeta_1 | b'\psi\rangle| : b' \in \mathcal{M}', \|b'\| = 1\} \\ &= \sup\{|\langle ya'\zeta_2 | b'\psi\rangle| : b' \in \mathcal{M}', \|b'\| = 1\} \\ &= \sup\{|\langle y\zeta_2 | a'^* b'\psi\rangle| : b' \in \mathcal{M}', \|b'\| = 1\} \\ &\leq \sup\{|\langle y\zeta_2 | c'\psi\rangle| : c' \in \mathcal{M}', \|c'\| = 1\} \\ &= F_{\mathcal{M}}(y\omega_{\zeta_2} y^*, \omega_\psi) \end{aligned} \tag{46}$$

for any $y \in \mathcal{M}$, since $\|a'^* b'\| \leq 1$ so the sup in the fourth line is over a larger set. But then the variational formula (41) gives without difficulty $\Phi_s(\omega_{\zeta_1} | \omega_\psi) \geq$

$\Phi_s(\omega_{\zeta_2}|\omega_\psi)$. Similarly, consider $|\psi_1\rangle$, $|\psi_2\rangle$ such that $\omega_{\psi_1} \leq \omega_{\psi_2}$. By the same argument $F(y\omega_\zeta y^*, \omega_{\psi_1}) \leq F(y\omega_\zeta y^*, \omega_{\psi_2})$, and the variational formula (41) thereby gives the following corollary:

Corollary 4 *For normal positive functionals on a von Neumann algebra $\omega_{\zeta_1} \leq \omega_{\zeta_2}$ and $\omega_{\psi_1} \leq \omega_{\psi_2}$, we have also $\Phi_s(\omega_{\zeta_1}|\omega_{\psi_1}) \geq \Phi_s(\omega_{\zeta_2}|\omega_{\psi_2})$ when $1 > s > 1/2$.*

As an application, consider an inclusion of von Neumann factors $\mathcal{N} \subset \mathcal{M}$ together with a conditional expectation $E : \mathcal{M} \to \mathcal{N}$ and unit vector $|\zeta\rangle$ such that $ind(E) = \lambda < \infty$. Then by definition, $\omega_\zeta \circ E \geq \lambda^{-1}\omega_\zeta$. The identity $\Phi_s(\omega_\zeta|\lambda^{-1}\omega_\psi) = \Phi_s(\omega_\zeta|\omega_\psi) + \ln\lambda$ [Corollary 6,3)] and the previous corollary trivially give

$$\Phi_s(\omega_\zeta|\omega_\zeta \circ E) \leq \ln\lambda \tag{47}$$

because $\Phi_s(\omega_\psi|\omega_\psi) = 0$.

We can also prove the DPI for Φ_s in the context of properly infinite von Neumann algebras using only properties of the fidelity in the range $1/2 \leq s \leq 1$.

Corollary 5 *Let $\mathcal{M}, \mathcal{N}$ be properly infinite von Neumann algebras and $T : \mathcal{M} \to \mathcal{N}$ a channel. Then, for two normal state functionals ω_ζ, ω_ψ we have $\Phi_s(\omega_\zeta \circ T|\omega_\psi \circ T) \leq \Phi_s(\omega_\zeta|\omega_\psi)$ for $s \in (1/2, 1)$.*

Proof By [25], thm. 2.10 (which assumes properly infinite von Neumann algebras), T can be written in Stinespring form $T(b) = v^*\rho(b)v$, where $v \in \mathcal{M}$, $v^*v = 1$, $vv^* = q$ (q a projection) and $\rho : \mathcal{N} \to \mathcal{M}$ a homomorphism of von Neumann algebras. Then, it is sufficient to prove the theorem separately for the case (i) $T_1(a) = v^*av$ and the case (ii) $T_2(b) = \rho(b)$.

(i) Using (26) with $\mathcal{M}'$ in place of $\mathcal{M}$, we have for $y \in \mathcal{M}$:

$$\begin{aligned}
F_\mathcal{M}(y\omega_{v\zeta}y^*|\omega_{v\psi}) &= \sup\{|\langle yv\zeta|x'v\psi\rangle| : \|x'\| = 1, x' \in \mathcal{M}'\} \\
&= \sup\{|\langle yv\zeta|vx'\psi\rangle| : \|x'\| = 1, x' \in \mathcal{M}'\} \\
&= \sup\{|\langle v^*yv\zeta|x'\psi\rangle| : \|x'\| = 1, x' \in \mathcal{M}'\} \\
&= F_\mathcal{M}((v^*yv)\omega_\zeta(v^*yv)^*|\omega_\psi).
\end{aligned} \tag{48}$$

Furthermore,

$$\omega_{v\zeta}(x^*x) = \omega_\zeta(v^*x^*xv) \geq \omega_\zeta((v^*xv)^*v^*xv). \tag{49}$$

Then, we have

$$\begin{aligned}
c_p &\inf_{x:\mathbb{R}_+ \to \mathcal{M}} \int_0^\infty [\omega_{v\zeta}(x(t)^*x(t)) + t^{-1}F_\mathcal{M}(y(t)\omega_{v\zeta}y(t)^*|\omega_{v\psi})^2]t^{-2/p'}dt \\
&= c_p \inf_{x:\mathbb{R}_+ \to \mathcal{M}} \int_0^\infty [\omega_{v\zeta}(x(t)^*x(t)) + t^{-1}F_\mathcal{M}((v^*y(t)v)\omega_{v\zeta}(v^*y(t)v)^*|\omega_\psi)^2]t^{-2/p'}dt \\
&\geq c_p \inf_{x:\mathbb{R}_+ \to \mathcal{M}} \int_0^\infty [\omega_\zeta(X(t)^*X(t)) + t^{-1}F_\mathcal{M}(Y(t)\omega_\zeta Y(t)^*|\omega_\psi)^2]t^{-2/p'}dt,
\end{aligned} \tag{50}$$

where $Y(t) = v^* y(t) v$, $X(t) = v^* x(t) v$. Note that these are particular examples of piecewise constant functions valued in $\mathcal{M}$ with finite range such that $X(t) + Y(t) = 1$ and such that $Y(t) = 0$ for sufficiently small t and $X(t) = 0$ for sufficiently large t. Thus, we can make the right side at most smaller by taking the infimum over *all* such functions. This results in $\Phi_s(\omega_{v\zeta} | \omega_{v\psi}) \leq \Phi_s(\omega_\zeta | \omega_\psi)$ using the definition of Φ_s (43) (where $2s = p$).

(ii) We have

$$c_p \inf_{x:\mathbb{R}_+ \to \rho(\mathcal{N})} \int_0^\infty [\omega_\zeta(x(t)^* x(t)) + t^{-1} F_{\rho(\mathcal{N})}(y(t)\omega_\zeta y(t)^* | \omega_\psi)^2] t^{-2/p'} dt$$

$$\geq c_p \inf_{X:\mathbb{R}_+ \to \mathcal{M}} \int_0^\infty [\omega_\zeta(X(t)^* X(t)) + t^{-1} F_{\rho(\mathcal{N})}(Y(t)\omega_\zeta Y(t)^* | \omega_\psi)^2] t^{-2/p'} dt$$

$$\geq c_p \inf_{X:\mathbb{R}_+ \to \mathcal{M}} \int_0^\infty [\omega_\zeta(X(t)^* X(t)) + t^{-1} F_{\mathcal{M}}(Y(t)\omega_\zeta Y(t)^* | \omega_\psi)^2] t^{-2/p'} dt, \quad (51)$$

where in the first step we took the infimum over the larger set of piecewise constant functions X valued in $\mathcal{M}$ with finite range such that $1 - X(t) = Y(t) = 0$ for sufficiently small t and $X(t) = 0$ for sufficiently large t. In the second step, we used the monotonicity $F_{\rho(\mathcal{N})} \geq F_{\mathcal{M}}$ since $\rho(\mathcal{N})$ is a von Neumann subalgebra of $\mathcal{M}$, by (26). This yields $\Phi_s(\omega_\zeta \circ \rho | \omega_\psi \circ \rho) \leq \Phi_s(\omega_\zeta | \omega_\psi)$. $\qquad \square$

Applying the DPI to the channel $\mathcal{A} \to \mathcal{A} \oplus \cdots \oplus \mathcal{A}, a \mapsto a \oplus \cdots \oplus a$ and the states $\rho = \oplus_i \lambda_i \omega_{\psi_i}, \sigma = \oplus_i \lambda_i \omega_{\zeta_i}$ implies that Φ_s is jointly convex by a standard argument, see e.g., [32], proof of prop. 1,

$$\sum_i \lambda_i \Phi_s(\omega_{\zeta_i} | \omega_{\psi_i}) \geq \Phi_s\left(\sum_i \lambda_i \omega_{\zeta_i} \Big| \sum_j \lambda_j \omega_{\psi_j}\right) \quad (52)$$

where the sum is finite and $\lambda_i \geq 0$, $\sum \lambda_i = 1$. Next, we obtain the following corollary:

Corollary 6 *Let $\mathcal{M}$ be a von Neumann algebra and $s \in (1/2, 1)$.*

(1) *We have for $\|\zeta\| = 1$*

$$\Phi_s(\omega_\zeta | \omega_\psi) \geq -\ln F(\omega_\zeta | \omega_\psi)^2. \quad (53)$$

(2) *We have for $\|\psi\| = 1$*

$$\Phi_s(\omega_\zeta | \omega_\psi) \leq -\frac{s}{1-s} \ln F(\omega_\zeta | \omega_\psi)^2. \quad (54)$$

(3) $\Phi_s(\omega_\zeta | \lambda \omega_\psi) = \Phi_s(\omega_\zeta | \omega_\psi) - \ln \lambda$ *for $\lambda > 0$.*
(4) *We have for $\|\psi\| = 1 = \|\zeta\|$ that $\lim_{s \to (1/2)^+} \Phi_s(\omega_\zeta | \omega_\psi) = -\ln F(\omega_\zeta | \omega_\psi)^2$.*
(5) $\Phi_s(\omega_\zeta | \omega_\psi) \geq 0$ *for $\|\psi\| = 1 = \|\zeta\|$ with equality iff $\omega_\zeta = \omega_\psi$.*

Proof For 1), we choose an approximation of

$$x(t) = \frac{F(\omega_\zeta|\omega_\psi)^2}{t + F(\omega_\zeta|\omega_\psi)^2}1 \tag{55}$$

by step functions. Then, we apply the variational definition of Φ_s (43) and the integral formula (30) upon which the result follows by a simple calculation using $\|\zeta\| = 1$.

For 2), we first use the supremum characterization of the fidelity (26), by which have $F(y\omega_\zeta y^*, \omega_\psi)^2 \geq |\langle\psi|y\zeta\rangle|^2 = \|P_\psi y\zeta\|^2$, where $P_\psi = |\psi\rangle\langle\psi|$ is a projector because $\|\psi\| = 1$. Then ($p = 2s$),

$$
\begin{aligned}
c_p \inf_{x:\mathbb{R}_+ \to \mathcal{M}} &\int_0^\infty [\omega_\zeta(x(t)^*x(t)) + t^{-1}F(y(t)\omega_\zeta y(t)^*|\omega_\psi)^2]t^{-2/p'}dt \\
&\geq c_p \inf_{x:\mathbb{R}_+ \to \mathcal{M}'} \int_0^\infty [\|x(t)\zeta\|^2 + t^{-1}\|P_\psi y(t)\zeta\|^2]t^{-2/p'}dt \\
&= c_p \int_0^\infty \langle\zeta|P_\psi(t + P_\psi)^{-1}\zeta\rangle t^{-2/p'}dt \\
&= c_p\|P_\psi\zeta\|^2 \int_0^\infty (t+1)^{-1}t^{-2/p'}dt = |\langle\zeta|\psi\rangle|^2.
\end{aligned}
\tag{56}
$$

This remains true if we change $|\zeta\rangle \to u'|\zeta\rangle$ for any unitary u' from $\mathcal{M}'$, thus giving

$$
\begin{aligned}
c_p \inf_{x:\mathbb{R}_+ \to \mathcal{M}} &\int_0^\infty [\omega_\zeta(x(t)^*x(t)) + t^{-1}F(y(t)\omega_\zeta y(t)^*|\omega_\psi)^2]t^{-2/p'}dt \\
&\geq \sup\{|\langle u'\zeta|\psi\rangle|^2 : u' \in \mathcal{M}' \text{ unitary}\} = F(\omega_\zeta|\omega_\psi)^2,
\end{aligned}
\tag{57}
$$

using a well-known characterization [1] of the fidelity in the last step. The rest then follows from the definition (43) of Φ_s.

For (3), we use the homogeneity of the fidelity $F(\lambda y(t)\omega_\psi y(t)^*|\omega_\zeta) = \sqrt{\lambda}F(y(t)\omega_\psi y(t)^*|\omega_\zeta)$ and apply the change of variables $t' = t/\lambda$ in the integral (43).

Item (4) is a combination of (1) and (2).

Item (5) follows from the properties $F(\omega_\zeta|\omega_\psi) \leq 1$, $F(\omega_\zeta|\omega_\psi) = 1$ iff $\omega_\zeta = \omega_\psi$, and 1), 2). $\qquad\square$

4 Application to quantum field theory

Here, we consider an application of Φ_s to quantum field theory inspired by [26]. For simplicity and concreteness, we consider chiral conformal quantum field theories (CFTs) on a single lightray (real line) or equivalently the circle in the conformally compactified picture. But the arguments are of a rather general nature and would apply with some fairly obvious modifications to general quantum field theories in higher dimensions under appropriate hypotheses.

We assume axioms that are standard in algebraic quantum field theory [16]. According to this axiom scheme, fulfilled by many examples, a chiral CFT is an assignment $\mathcal{A} : I \mapsto \mathcal{A}(I)$, wherein $I \subset S^1$ is an open interval and $\mathcal{A}(I)$ a von Neumann algebra acting on a fixed Hilbert space $\mathscr{H}$ with the following properties:

1. (Isotony) If $I_1 \subset I_2$ then $\mathcal{A}(I_1) \subset \mathcal{A}(I_2)$.
2. (Commutativity) If $I_1 \cap I_2$ is empty, then $[\mathcal{A}(I_1), \mathcal{A}(I_2)] = \{0\}$.
3. (Möbius covariance) There is a strongly continuous unitary representation U on $\mathscr{H}$ of the Möbius group $G = \widetilde{SL_2(\mathbb{R})}/\mathbb{Z}_2$ which is consistent with the standard action of this group on the circle by fractional linear transformations, in the sense $U(g)\mathcal{A}(I)U(g)^* = \mathcal{A}(gI)$ for all $g \in G$.
4. (Positive energy) The one-parameter subgroup of rotations has a positive generator L_0 under the representation U.
5. (Vacuum) There is a unique vector $|\Omega\rangle \in \mathcal{H}$, called the vacuum, which is invariant under all $U(g)$, $g \in G$.
6. (Additivity) Let I and I_n be intervals such that $I = \cup_n I_n$. Then, $\mathcal{A}(I) = \vee_n \mathcal{A}(I_n)$ (strong closure).

The special situation we would like to study here are two chiral CFTs $\mathcal{A}$, $\mathcal{F}$ in the above sense such that $\mathcal{A}(I) \subset \mathcal{F}(I)$ is an inclusion of von Neumann factors acting on the same Hilbert space $\mathscr{H}$ for any open interval I, and transforming under the same representation, U. By general arguments, these factors have to be of type III [9]. A typical example is when $\mathcal{A}$ is the Virasoro net (operator algebras generated by the stress energy tensor) and $\mathcal{F}$ is an extension of finite index as classified in [22]. For further details on such a setting, see e.g., [27,28]. We will also assume that the Jones-Kosaki index $\lambda \equiv [\mathcal{F}(I) : \mathcal{A}(I)]$ is finite (hence independent of I by [28]). By [27], lemma 13, this implies that for each I there is a conditional expectation $E_I : \mathcal{F}(I) \to \mathcal{A}(I)$, satisfying the Pimsner–Popa inequality (10). We assume that E_I leaves the vacuum vector invariant, $\omega_\Omega \circ E_I = \omega_\Omega$ for all intervals I. Furthermore, these conditional expectations are assumed to be consistent in the sense $E_I|_{\mathcal{F}(J)} = E_J$ for $J \subset I$ [28]. Consider two sets of intervals (identifying S^1 with the real line via a stereographic projection):

$$A_n = (a, -1/n), \quad B_n = (1/n, b), \tag{58}$$

wherein n is a natural number and $a < 0, b > 0$. We consider the von Neumann algebra inclusion $\mathcal{A}(A_n) \vee \mathcal{A}(B_n) \subset \mathcal{F}(A_n) \vee \mathcal{F}(B_n)$, and we let E_n be the conditional expectation $\mathcal{F}(A_n) \vee \mathcal{F}(B_n) \to \mathcal{A}(A_n) \vee \mathcal{A}(B_n)$ such that

$$E_n(a_n b_n) = E_{A_n}(a_n) E_{B_n}(b_n) \quad \forall a_n \in \mathcal{F}(A_n), b_n \in \mathcal{F}(B_n). \tag{59}$$

Thus, E_n only projects out degrees of freedom of the individual parts of the system in (58) separately.[3] In the limit as $n \to \infty$ (denoted as $\lim_n$ in the following), these systems touch each other. We can show the following theorem.

Theorem 1 *We have $\lim_n \Phi_s(\omega_\Omega|\omega_\Omega \circ E_n) = \ln[\mathcal{F} : \mathcal{A}]$ for $s \in [1/2, 1)$.*

---

[3] Somewhat formally $E_n = E_{A_n} \otimes E_{B_n}$, which holds rigorously if the split property holds in the CFT.

Corollary 7 *We have*

$$\lim_n F(\omega_\Omega | \omega_\Omega \circ E_n) = [\mathcal{F} : \mathcal{A}]^{-1/2}. \tag{60}$$

Proof Obvious in view of Corollary 6, 4) and Theorem 1. □

Proof of Theorem 1 The proof strategy is similar to that of a result by Longo and Xu [26] who have considered the relative entropy S instead of the divergence Φ_s. As their proof, we make use of the variational definition of the divergence Φ_s.

First assume that $1/2 < s < 1$. We use the notation $d^2 = \lambda \equiv [\mathcal{F}(I) : \mathcal{A}(I)] < \infty$ which is independent of I [28]. Let $|\psi_n\rangle$ be a vector such that $\omega_{\psi_n} = \omega_\Omega \circ E_n$, as a functional on $\mathcal{F}(A_n) \vee \mathcal{F}(B_n)$. □

Lemma 2 *There exists a sequence $\{f_n\} \subset \mathcal{F}(A_n) \vee \mathcal{F}(B_n)$ such that $f_n \to 1$ strongly and*

$$\lim_n \omega_\Omega(f_n) = 1, \quad \lim_n \omega_\Omega(f_n^* f_n) = 1, \quad \lim_n \omega_{\psi_n}(f_n^* f_n) = \lambda^{-1}. \tag{61}$$

Proof The proof is given in [26], prop. 4.5. However we rephrase it somewhat in preparation to the discussions in the next section. A finite index inclusion $\mathcal{N} \subset \mathcal{M}$ of properly infinite von Neumann factors is characterized uniquely by its associated Q-system [7,29] (x, w, θ), wherein $x, w \in \mathcal{N}$ obey certain relations relative to the canonical endomorphism θ of $\mathcal{N}$, see appendix A.

Applying this structure to the inclusions $\mathcal{A}(A_n) \subset \mathcal{F}(A_n)$ we get $v_{A_n} \in \mathcal{F}(A_n)$ and similarly for B_n. These are fixed uniquely demanding that the corresponding conditional expectations, see appendix A, be given by the $|\Omega\rangle$ preserving conditional expectation E_{A_n} etc. By translation-dilation covariance, this implies for example that $v_{A_n} \to v_A$ strongly as $n \to \infty$. Another standard result in this setting, shown, e.g., in [26], lemma 2.9, is that v_{A_n} can be "transported" to v_{B_n} in the sense that there is a unitary $u_{B_n A_n} \in \mathcal{A}(a, b) \cap \text{Hom}(\theta_{B_n}, \theta_{A_n})$, such that $v_{B_n} = u_{B_n A_n} v_{A_n}$. By additivity, we may find a sequence of unitaries $a_{n,k} \in \mathcal{A}(A_n)$, $b_{n,k} \in \mathcal{B}(B_n)$ such that $\sum_{k=1}^{N(n)} b_{n,k}^* a_{n,k} - u_{B_n A_n} \to 0$ as $n \to \infty$, in the strong sense. Then, let

$$V_{A_n,k} = \frac{1}{\sqrt{d}} a_{n,k} v_{A_n} \in \mathcal{F}(A_n), \quad V_{B_n,k}^* = \frac{1}{\sqrt{d}} v_{B_n}^* b_{n,k}^* \in \mathcal{F}(B_n). \tag{62}$$

Finally, let

$$f_n = \sum_{k=1}^{N(n)} V_{B_n,k}^* V_{A_n,k}. \tag{63}$$

Then, it follows that $f_n \to d^{-1} v_B^* v_B = 1$ strongly by construction and the relations of Q-systems, see appendix A. This already implies the first two of the claimed limits

in (61). On the other hand,

$$
\begin{aligned}
\omega_\Omega \circ E_n(f_n^* f_n) &= \sum_{k,l} \omega_\Omega \circ E_n(V_{A_n,k}^* V_{B_n,k} V_{B_n,l}^* V_{A_n,l}) \\
&= \sum_{k,l} \omega_\Omega \circ E_n(V_{A_n,k}^* V_{A_n,l} V_{B_n,k} V_{B_n,l}^*) \\
&= \sum_{k,l} \omega_\Omega(E_{A_n}(V_{A_n,k}^* V_{A_n,l}) E_{B_n}(V_{B_n,k} V_{B_n,l}^*)) \\
&= \sum_{k,l} \omega_\Omega(V_{A_n,k}^* V_{A_n,l} E_{B_n}(V_{B_n,k} V_{B_n,l}^*)) \\
&= d^{-3} \sum_{k,l} \omega_\Omega(v_{A_n}^* a_{n,k}^* a_{n,l} v_{A_n} b_{n,k} b_{n,l}^*) \\
&= d^{-3} \sum_{k,l} \omega_\Omega(v_{A_n}^* a_{n,k}^* b_{n,k} a_{n,l} b_{n,l}^* v_{A_n}) \to d^{-2}
\end{aligned}
\tag{64}
$$

using commutativity in the first line, the definition of E_n in the second line, $E_I|_{\mathcal{F}(J)} = E_J$ for $J \subset I$ and $\omega_\Omega \circ E_I = \omega_\Omega$ in the third line, the bimodule property of E_{B_n} as well as $E_{B_n}(v_{B_n} v_{B_n}^*) = d^{-1}$ by properties of the Q-system in the fourth line, commutativity again in the fifth line, and $\sum a_{n,k}^* b_{n,k} a_{n,l} b_{n,l}^* \to 1$ strongly as $n \to \infty$ and $v_{A_n}^* v_{A_n} = d \cdot 1$ in the last line (using properties of the Q-system). $\qquad\square$

Next, we define

$$
x_n(t) = \begin{cases} 1 - \frac{t}{t+\lambda^{-1}} f_n & \text{if } 1/k \le t \le k \\ 1 & \text{if } t > k \\ 0 & \text{if } t < 1/k. \end{cases}
\tag{65}
$$

Using the properties (61) of f_n, we have for $t \in (1/k, k)$:

$$
\begin{aligned}
\lim_n \omega_\Omega(x_n(t)^* x_n(t)) &= \frac{\lambda^{-2}}{(t+\lambda^{-1})^2} \\
\limsup_n F(y_n(t)\omega_\Omega y_n(t)^* | \omega_{\psi_n}) &\le \limsup_n \|y_n(t)^* \psi_n\| = \frac{\lambda^{-1} t^2}{(t+\lambda^{-1})^2},
\end{aligned}
\tag{66}
$$

using in the second line the Cauchy–Schwarz inequality in order to estimate the fidelity characterized through (26). Therefore, for fixed k, we have

$$
\begin{aligned}
&\limsup_n \int_{1/k}^k \left[\omega_\Omega(x_n(t)^* x_n(t)) + t^{-1} F(y_n(t)\omega_\Omega y_n(t)^* | \omega_{\psi_n})^2 \right] t^{-(2s-1)/s} dt \\
&\le \int_{1/k}^k \left[\frac{\lambda^{-2}}{(t+\lambda^{-1})^2} + \frac{\lambda^{-1} t}{(t+\lambda^{-1})^2} \right] t^{-(2s-1)/s} dt \\
&= c_{2s}^{-1} \lambda^{(s-1)/s} - \frac{s}{1-s} k^{(s-1)/s} - \frac{s}{2s-1} k^{-(2s-1)/s}
\end{aligned}
$$

$$+s \sum_{m=1}^{\infty} (-1)^m \left\{ \left(\frac{\lambda}{k}\right)^m \frac{1}{ms + (1-s)} + \left(\frac{1}{\lambda k}\right)^{m+1} \frac{1}{ms + (2s-1)} \right\}, \quad (67)$$

using the integral (30) and the definition of c_p from prop. 1 in the last step. The last sum is of order $O(k^{-1})$ uniformly in $s \in [1/2, 1]$. On the other hand, using the definition of $x_n(t)$ in the range $t < 1/k$, we have

$$\limsup_n \int_0^{1/k} \left[\omega_\Omega(x_n(t)^* x_n(t)) + t^{-1} F(y_n(t)\omega_\Omega y_n(t)^* |\omega_{\psi_n})^2 \right] t^{-(2s-1)/s} dt \quad (68)$$
$$= \int_0^{1/k} t^{-(2s-1)/s} dt = \frac{s}{1-s} k^{(s-1)/s},$$

while using the definition of $x_n(t)$ in the range $t > k$, we have

$$\limsup_n \int_k^{\infty} \left[\omega_\Omega(x_n(t)^* x_n(t)) + t^{-1} F(y_n(t)\omega_\Omega y_n(t)^* |\omega_{\psi_n})^2 \right] t^{-(2s-1)/s} dt \quad (69)$$
$$= \int_k^{\infty} t^{-(2s-1)/s-1} dt = \frac{s}{2s-1} k^{-(2s-1)/s}.$$

Consequently, when $s = p/2$, the variational expression (41) is estimated by[4]

$$\limsup_n c_{2s} \int_0^{\infty} \left[\omega_\Omega(x_n(t)^* x_n(t)) + t^{-1} F(y_n(t)\omega_\Omega y_n(t)^* |\omega_{\psi_n})^2 \right] t^{-2/(2s)'} dt \quad (70)$$
$$\leq \lambda^{(s-1)/s} + O(k^{-1})$$

for any $k > 0$ where $O(k^{-1})$ is a term bounded in norm by Ck^{-1} uniformly in $s \in [1/2, 1]$. Letting $k \to \infty$ this term disappears, and then using the definition of Φ_s and of $|\psi_n\rangle$ gives

$$\liminf_n \Phi_s(\omega_\Omega|\omega_\Omega \circ E_n) \geq \ln \lambda. \quad (71)$$

On the other hand, we have already seen before in (47) that $\Phi_s(\omega_\Omega|\omega_\Omega \circ E_n) \leq \ln \lambda$. The proof of the theorem is therefore complete for the case $1/2 < s < 1$.

Now we turn to the limiting case $s \to (1/2)^+$. We go back to the proof and investigate the limit as $s \to (1/2)^+$. By inspection it can be seen that in order to obtain an expression in (70) not exceeding $\lambda^{(s-1)/s} + O(k^{-1}) + \varepsilon$ for some $\varepsilon > 0$, we need $n \geq n_0(k, \varepsilon)$, where n_0 does not depend on $s \in [1/2, 1]$. Furthermore, we have argued in the proof that $O(k^{-1})$ is uniform in $s \in [1/2, 1]$. Thus, the limit $s \to 1/2^+$ may be taken and we learn that $F(\omega_\Omega|\omega_{\psi_n})^2 \leq \lambda^{-1} + O(k^{-1}) + \varepsilon$ when $n \geq n_0(k, \varepsilon)$. Thus, $\limsup_n F(\omega_\Omega|\omega_{\psi_n})^2 \leq \lambda^{-1}$ and the rest is as before.

---

[4] Note that the variational expression holds by continuity also for strongly continuous families such as $x_n(t)$.

Remarks Cor. 1 for $s = 1/2$ gives the a dual formulation of this result when applied to $\mathcal{M}_n = \mathcal{F}(A_n) \vee \mathcal{F}(B_n)$, $\mathcal{N}_n = \mathcal{A}(A_n) \vee \mathcal{A}(B_n)$ and $E_n' : (\mathcal{A}(A_n) \vee \mathcal{A}(B_n))' \to (\mathcal{F}(A_n) \vee \mathcal{F}(B_n))'$, which is the dual conditional expectation. Indeed, if we combine cor.s 1, 7 and 6, 4), we immediately find that

$$\lim_n F(\omega_\Omega' | \omega_\Omega' \circ E_n') = 1. \tag{72}$$

5 Conclusions

We end this paper by commenting on the physical significance of the result in sec. 4. For this, it is instructive to have in mind the example of a Haag–Kastler QFT [16], $\mathcal{F}$, containing charged fields. These map the vacuum $|\Omega\rangle$ to states with net (flavor) charge. The subset of charge neutral operators is $\mathcal{A}$. On the full Hilbert space $\mathcal{H}$ (including charged states), the gauge group G acts by global unitaries which transform the charged fields and leave the vacuum invariant. The conditional expectation $E_I : \mathcal{F}(I) \to \mathcal{A}(I)$ is the Haar-average over G and projects onto the charge neutral operators ("observables") in a given region I, which is left invariant because gauge transformations commute with translations by the Coleman–Mandula theorem (see e.g., [16]). Assuming that G is a finite group with $|G|$ elements, the index is $|G| = [\mathcal{F} : \mathcal{A}]$.

Given two spacelike related regions A_n and B_n separated by a finite corridor of size $\sim 1/n$, the conditional expectation E_n defined by (59) is basically the tensor product $E_{A_n} \otimes E_{B_n}$. $\Phi_s(\omega_\Omega | \omega_\Omega \circ E_n)$ in a sense accounts for the correlations between A_n and B_n that are visible using charge operators only in both subsystem. This interpretation becomes more and more precise when the regions move together. The above intuitive argument has been substantiated (in a somewhat heuristic way) in the very lucid paper by [11], in the case of the relative entropy S—such that we should use Kosaki's variational formula for S (8) instead of the variational definition of Φ_s (43). They first argue using known properties of S in connection with conditional expectations that the mutual information between A_n and B_n in the vacuum state satisfies

$$I_{\mathcal{F}}(A_n | B_n) - I_{\mathcal{A}}(A_n | B_n) = S(\omega_\Omega | \omega_\Omega \circ E_n). \tag{73}$$

When $n \to \infty$, it is plausible that the mutual information on the left side is dominated by correlations between charge carrying operators localized very near the edges where A_n and B_n approach each other. Furthermore, although each term in $I_{\mathcal{F}}(A_n | B_n) - I_{\mathcal{A}}(A_n | B_n)$ is expected to diverge, the difference ought to be a finite number related to the order of G. In fact, by investigating more closely the right side of the equation, they argue that $S(\omega_\Omega | \omega_\Omega \circ E_n)$ converges to $\ln |G|$ when $n \to \infty$.

Actually, the core of the argument by [11] has a similar flavor to that given in the proof of Theorem 1, in the following sense. Going to our proof, a key step is the construction of the "vertex operators" which have in a sense maximal correlation across the separating corridor between A_n and B_n as stated in lem. 2. To simplify, let us take half lines A_n, B_n separated by a corridor of width $2/n$ symmetrically around the origin. Proceeding somewhat informally to simplify the discussion, we consider instead the isometric vertex operators $V_n = u_{C_n B_n} v_{B_n} / \sqrt{d}$ where $C_n = (1/n, 2/n)$

and $u_{C_n B_n}$ is a unitary charge transporter from B_n to C_n. Then, V_n is localized in $(1/n, 2/n)$, and it creates an incoherent superposition of all irreducible charges in this interval by the Q-system construction, see app. A. Letting $J = J_{\mathcal{F}}$ be the modular conjugation associated with the half-line $(0, \infty)$, we can say that $\bar{V}_n = J V_n J$ creates the opposite charges in the opposite interval $(-2/n, -1/n)$ because J is basically the PCT operator exchanging A_n with B_n, and particle with anti-particle (Bisognano–Wichmann theorem [16]).

Thus, the correlation function which we want to maximize similar to Lemma 2 is

$$1 \geq \langle \Omega | \bar{V}_n V_n \Omega \rangle = \langle \Omega | V_n \Delta^{1/2} V_n^* \Omega \rangle, \tag{74}$$

where the inequality is simply the Cauchy–Schwarz inequality. The modular flow Δ^{it} corresponds to dilations by e^t (Bisognano–Wichmann theorem), and $V_n | \Omega \rangle$ should be approximately dilation invariant moving ever closer to the edge of B_n when $n \to \infty$. Thus, the limit of $\langle \Omega | \bar{V}_n V_n \Omega \rangle$ should indeed be 1. Arguing just as in Lemma 2, one can likewise see at least formally that $\langle \psi_n | \bar{V}_n V_n \psi_n \rangle$ should tend to λ^{-1}.

Thus, in this sense, the quantity $S(\omega_\Omega | \omega_\Omega \circ E_n)$ is dominated in the limit $n \to \infty$ by particle anti-particle pair correlations very close to the edges across the corridor in accordance with the intuitive picture proposed by [11].

A Q-systems, subfactors and OPE [7,28,29]

A Q-system is a way to encode an inclusion of properly infinite von Neumann factors $\mathcal{N} \subset \mathcal{M}$ possessing a minimal conditional expectation $E : \mathcal{M} \to \mathcal{N}$ such that the index, denoted here by d^2, is finite. An important point is that the data in the Q-system only refer to the smaller factor, $\mathcal{N}$.

Central to the construction is the notion of an endomorphism of $\mathcal{N}$, which is an ultra-weakly continuous $*$-homomorphism of $\mathcal{N}$ such that $\theta(1) = 1$. Given two endomorphisms ρ, θ, one says that a linear operator is an intertwiner, written $T \in \text{Hom}(\rho, \theta)$, if $T \rho(n) = \theta(n) T$ for all $n \in \mathcal{N}$. Two endomorphisms are called equivalent if there is a unitary intertwiner and irreducible if there is no non-trivial self-intertwiner. One writes $\theta \cong \oplus_i \rho_i$ if there is a finite set of irreducible and

mutually inequivalent endomorphisms ρ_i and isometries $w_i \in \mathrm{Hom}(\rho_i, \theta)$ such that $\theta(n) = \sum_i w_i^* \rho_i(n) w_i$ for all $n \in \mathcal{N}$ and such that $w_i w_j^* = \delta_{ij} 1$.

Definition 3 A Q-system is a triple (θ, x, w) where: $\theta \cong \oplus_i \rho_i$ is an endomorphism of $\mathcal{N}$, $w \in \mathrm{Hom}(\theta, id) \cap \mathcal{N}$ and $x \in \mathrm{Hom}(\theta^2, \theta) \cap \mathcal{N}$ such that

$$w^* x = \theta(w^*) x = 1, \quad x^2 = \theta(x) x, \quad \theta(x^*) x = x x^* = x^* \theta(x), \tag{75}$$

as well as

$$w^* w = d \cdot 1, \quad x^* x = d \cdot 1. \tag{76}$$

Given a Q-system, one defines an extension $\mathcal{M}$ as follows. As a set, $\mathcal{M}$ consists of all symbols of the form nv, where $n \in \mathcal{N}$ with the product, $*$-operation, and unit defined by, respectively

$$n_1 v n_2 v = n_1 \theta(n_2) x v, \quad (nv)^* = w^* x^* \theta(n^*) v, \quad 1 = w^* v. \tag{77}$$

Associativity and consistency with the $*$-operation follow from the defining relations of a Q-system, which also give $vn = \theta(n)v$. The conditional expectation is related to the data by $E(nv) = d^{-1} n w$ and is used to induce the operator norm on $\mathcal{M}$. The bimodule property of E follows again from the defining relations of the Q-system, which also give for instance $E(vv^*) = d^{-1} 1$.

Conversely, given an inclusion of properly infinite factors $\mathcal{N} \subset \mathcal{M}$, the data of the Q-system and $v \in \mathcal{M}$ can be found by a canonical procedure and $d^2 = [\mathcal{M} : \mathcal{N}]$.

Let ρ_i and $w_i \in \mathrm{Hom}(\rho_i, \theta)$ be the endomorphisms and intertwiners corresponding to the decomposition $\theta \cong \oplus_i \rho_i$ into irreducibles. Next, define

$$\psi_i = w_i^* v. \tag{78}$$

The relations in Definition 3 imply that the following relations hold. Define:

$$c_{i,j}^k = w_i^* \theta(w_j^*) x w_k, \quad w_0 = w, \tag{79}$$

and let $\rho_0 = id$ be the trivial endomorphism of $\mathcal{N}$. Then,

- (Operator product expansion): $\psi_i \psi_j = \sum_k c_{i,j}^k \psi_k$.
- ($*$-operation) $\psi_k^* = c_{\bar{k},k}^0 {}^* \psi_{\bar{k}}$ and $c_{j,k}^0 = \delta_{j,\bar{k}} R_k$, where $R_k \in \mathrm{Hom}(\rho_0, \bar{\rho}_k \rho_k)$ is the intertwiner characterizing the "conjugate sector."
- (Unit) $\psi_0 = 1$.

In the QFT context, one not only has one inclusion, but a net of inclusions $\mathcal{A}(I) \subset \mathcal{F}(I)$ [28]. Furthermore, in chiral CFT, $\mathcal{A}(I)$ is often taken to be the algebra generated by the smeared stress tensor inside I ("Virasoro-net"). From this, one should be able to construct an operator product expansion in the usual sense in the physics literature, although to establish the connection in full precision/generality remains an open problem.

The basic idea is to consider the "fields" $\psi_{i,I}$ for each interval I. To obtain a pointlike vertex operator, we should shrink $I \to \{x\}$, while at the same time subtract the vacuum expectation value $\langle \Omega | \psi_{I,i} \Omega \rangle 1$ and rescaling[5] by $|I|^{-h_i}$ to obtain a finite limit, $V_i(x)$. These "primary" fields obey an OPE with "coefficients" $c^k_{i,j}(x, y)$ that are still operators in the Virasoro net. We should think of them as operator valued functions $c^k_{i,j}(x, y) = c^k_{i,j}(x, y, \{L_n\})$. When formally expanded out as a power series in the Virasoro generators $\{L_n\}_{n \in \mathbb{Z}}$, this ought to give the operator product expansion with certain numerical coefficients containing on the right side the primary vertex operators $V_k(y)$ as well as their descendants $\phi_{k,\{n\}}(y) = [L_{n_1}, [\cdots L_{n_m}, V_k(y)]]$, where $n_1 < n_2 < \cdots < 0$. This is the form of the operator product expansion usually given in the physics literature. Representation theoretic considerations then formally determine the scaling of the numerical OPE coefficients. Such partly heuristic claims are at the basis of our discussion in Sect. 5.

References

1. Alberti, P.M.: A note on the transition probability over C^* algebras. J. Math. Phys. **7**, 25–32 (1983)
2. Araki, H.: Relative Hamiltonian for faithful normal states of a von Neumann algebra. Publ. RIMS Kyoto Univ. **9**, 165–209 (1973)
3. Araki, H.: Relative entropy of states of von Neumann algebras. I, II. Publ. RIMS Kyoto Univ. **11**, 809–833 (1976) and **13**, 173–192 (1977)
4. Araki, H., Masuda, T.: Positive cones and L_p-spaces for von Neumann algebras. Publ. RIMS Kyoto Univ. **18**, 339–411 (1982)
5. Beigi, S.: Sandwiched relative Renyi entropy satisfies data processing inequality. J. Math. Phys. **54**, 122202 (2013)
6. Berta, M., Scholz, V.B., Tomamichel, M.: Renyi divergences as weighted non-commutative vector-valued L_p-Spaces. Ann. Henri Poincare **19**(6), 1843 (2018)
7. Bischoff, M., Longo, R., Rehren, K.-H.: Tensor Categories and Endomorphisms of von Neumann Algebras (with Applications to Quantum Field Theory) Springer Briefs in Mathematical Physics (2015)
8. Bratteli, O., Robinson, D.W.: Operator Algebras and Quantum Statistical Mechanics I. Springer (1987). Bratteli, O., Robinson, D.W.: Operator Algebras and Quantum Statistical Mechanics II. Springer (1997)
9. Buchholz, D., Fredenhagen, K., D'Antoni, C.: The universal structure of local algebras. Commun. Math. Phys. **111**, 123 (1987)
10. Connes, A.: Spatial theory of von Neumann algebras. J. Funct. Anal. **35**, 153–164 (1980)
11. Casini, H., Huerta, M., Magan, J.M., Pontello, D.: Entanglement entropy and superselection sectors. Part I. Global symmetries. JHEP **2002**, 14 (2020)
12. Casini, H., Huerta, M., Magan, J.M., Pontello, D.: Entropic order parameters for the phases of QFT. J. High Energ. Phys. **2021**, 277 (2021). https://doi.org/10.1007/JHEP04(2021)277
13. Faulkner, T., Hollands, S., Swingle, B., Wang, Y.: Approximate recovery and relative entropy I. General von Neumann subalgebras. arXiv:2006.08002 [quant-ph]
14. Faulkner, T.: The holographic map as a conditional expectation. arXiv:2008.04810 [hep-th]
15. Frank, R., Lieb, E.: Monotonicity of a relative Renyi entropy. J. Math. Phys. **54**, 122201 (2013)
16. Haag, R.: Local Quantum Physics: Fields, Particles, Algebras. Springer, Berlin (1992)
17. Haagerup, U.: Operator valued weights in von Neumann algebras. I + II. J. Funct. Anal. **32**, 175–206 (1979) and **33**, 339–361 (1979)
18. Hiai, F.: Minimizing indices of conditional expectations onto a subfactor. Publ. Res. Inst. Math. Sci. Kyoto Univ. **24**, 673–678 (1988)
19. Hiai, F.: Quantum f-divergences in von Neumann Algebras. Springer, Berlin (2021)
20. Jencova, A.: Renyi relative entropies and noncommutative L_p-spaces. Ann. Henri Poincare **19**(8), 2513–2542 (2018)
21. Jones, V.: Index for subfactors. Invent. Math. **72**, 1–25 (1983)

---

[5] h_i is expected to equal the highest weight provided by the irreducible Virasoro representation ρ_i.

22. Kawahigashi, Y., Longo, R.: Classification of two-dimensional local conformal nets with c less than 1 and 2 cohomology vanishing for tensor categories. Commun. Math. Phys. **244**, 63 (2004)
23. Kosaki, H.: Relative entropy for states: a variational expression. J. Oper. Theory **16**, 335–348 (1986)
24. Kosaki, H.: Extension of Jones theory on index to arbitrary factors. J. Funct. Anal. **66**, 123–140 (1986)
25. Longo, R.: (2018). On Landauer's principle and bound for infinite systems. Commun. Math. Phys. **363**(2), 53
26. Longo, R., Xu, F.: Relative entropy in CFT. Adv. Math. **337**, 139 (2018)
27. Longo, R.: Conformal subnets and intermediate subfactors. Commun. Math. Phys. **237**(1–2), 7–30 (2003)
28. Longo, R., Rehren, K.-H.: Nets of subfactors. Rev. Math. Phys. **7**, 567–597 (1995)
29. Longo, R.: A duality theory for Hopf algebras and for subfactors. Commun. Math. Phys. **159**, 133–150 (1994)
30. Magan, J.M., Pontello, D.: QQuantum complementarity through entropic certainty principles. Phys. Rev. A **103**, 012211 (2021). https://doi.org/10.1103/PhysRevA.103.012211
31. Mosonyi, M., Ogawa, T.: Quantum hypothesis testing and the operation meaning of quantum relative Renyi entropies. Commun. Math. Phys. **334**(3), 1617–1648 (2015)
32. Muller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., Tomamichel, M.: On quantum relative Renyi entropies: a new generalization. J. Math. Phys. **54**(12), 122203 (2013)
33. Ohya, M., Petz, D.: Quantum Entropy and Its Use, Theoretical and Mathematical Physics. Springer, Berlin (1993)
34. Pimsner, M., Popa, S.: Entropy and index for subfactors. Ann. Sci. Ec. Norm. Sup. **19**, 57–106 (1986)
35. Takesaki, M.: Theory of Operator Algebras, I–III. Springer, Berlin (2003)
36. Uhlmann, A.: Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory. Commun. Math. Phys. **54**(1), 21–32 (1977)
37. Uhlmann, A.: The 'transition probability' in the state space of a *-algebra. Rep. Math. Phys. **9**(2), 273–279 (1976)
38. Wilde, M.M., Winter, A., Yang, D.: Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Renyi relative entropy. Commun. Math. Phys. **331**(2), 593–622 (2014)
39. Xu, F.: On relative entropy and global index. Trans. Am. Math. Soc. **373**(5), 3515 (2020)