A functorial characterization of von Neumann entropy

Arthur PARZYGNAT

Institut des Hautes Études Scientifiques
35, route de Chartres
91440 – Bures-sur-Yvette (France)

Septembre 2020

IHES/P/20/11

A functorial characterization of von Neumann entropy

Arthur J. Parzygnat

September 15, 2020

Abstract

We classify the von Neumann entropy as a certain concave functor from finite-dimensional non-commutative probability spaces and state-preserving $*$-homomorphisms to real numbers. This is made precise by first showing that the category of non-commutative probability spaces has the structure of a Grothendieck fibration with a fiberwise convex structure. The entropy difference associated to a $*$-homomorphism between probability spaces is shown to be a functor from this fibration to another one involving the real numbers. Furthermore, the von Neumann entropy difference is classified by a set of axioms similar to those of Baez, Fritz, and Leinster characterizing the Shannon entropy difference. The existence of disintegrations for classical probability spaces plays a crucial role in our classification.

Contents

2020 Mathematics Subject Classification. 18D30, 81P17 (Primary); 18C40, 46L53, 81R15, 94A17 (Secondary).

Key words and phrases. Convex category, disintegration, Grothendieck fibration, optimal hypothesis, quantum entropy

1 Introduction and outline

In 2011, Baez, Fritz, and Leinster (BFL) characterized the Shannon entropy (difference) of finite probability distributions as the only non-vanishing continuous affine functor **FinProb** $\to \mathbb{BR}_{\geqslant 0}$ from finite probability spaces to non-negative numbers up to an overall non-negative constant [4]. Here, **FinProb** is the category of finite sets equipped with probability measures as objects and probability-preserving functions as morphisms. The codomain category, $\mathbb{BR}_{\geqslant 0}$, is the category consisting of a single object and whose morphisms from that object to itself are all non-negative real numbers equipped with addition as the composition. Each of these categories has a convex structure allowing one to take convex combinations of objects and morphisms, all of which will be reviewed in this manuscript.

A natural follow-up question is whether the von Neumann (or finite-dimensional Segal) entropy can be characterized in a similar manner by replacing **FinProb** with **NCFinProb**, the category of finite quantum probability spaces, consisting of unital finite-dimensional C^*-algebras equipped with states as objects and state-preserving unital $*$-homomorphisms as morphisms. Although this question was explored by Baez and Fritz [2], a suitably similar set of axioms was never obtained. The present manuscript accomplishes this task.

There are two main difficulties with extending BFL's result to the quantum setting. The first issue is that the difference of von Neumann entropies need not have a fixed sign. Namely there are state-preserving unital $*$-homomorphisms that *decrease* the entropy as well as *increase* the entropy. The root of this stems from the fact that when one attempts to make an observation of a quantum system in a pure state, then the outcomes are generally probabilistic (though there are also examples not of this kind, which have recently found applications in quantum information theory [21, 22]). Using our axioms, we show that the existence of disintegrations [42] (called optimal hypotheses in [3]) implies the non-negativity of the entropy difference. Since disintegrations always exist for finite-dimensional classical systems, this proves one of the main assumptions of BFL in their functorial characterization of the Shannon entropy [4]. The second difficulty is that the objects of **NCFinProb** are not convex generated by any single object in that category. Note that this occurs for **FinProb**, where an arbitrary probability spaces (X, p), with X a finite set and p a probability measure on X, can be decomposed into a convex sum as

$$(X, p) \cong \bigoplus_{x \in X} p_x \mathbf{1}, \tag{1.1}$$

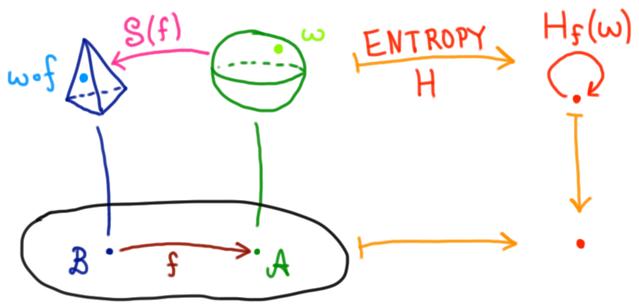
where $\mathbf{1}$ is the (essentially) unique probability space consisting of a single element and p_x is the probability of $x \in X$. In **NCFinProb**, a non-commutative probability space such as $(\mathcal{M}_m, \omega)$ cannot be expressed as a convex combination of lower-dimensional probability spaces (since $\mathcal{M}_m$ is a *factor*). Here, $m \in \mathbb{N}$, $\mathcal{M}_m$ is the C^*-algebra of $m \times m$ matrices, and ω is a state on $\mathcal{M}_m$ (which can be uniquely represented by a density matrix).

In this manuscript, we simultaneously address both these issues and provide a functorial characterization of the von Neumann entropy. This is done by introducing Grothendieck fibrations of convex categories and fibred affine functors. The category **NCFinProb** forms a fibration over **fdC*-Alg**, the category of finite-dimensional unital C^*-algebras and unital $*$-homomorphisms, by sending each quantum probability space $(\mathcal{A}, \omega)$ to the underlying C^*-

algebra $\mathcal{A}$. The von Neumann entropy (difference) provides a functor

$$\begin{tikzcd} \mathbf{NCFinProb} \arrow[r, "\text{Entropy}"] \arrow[d] & \mathbb{BR} \arrow[d] \\ \mathbf{fdC^*\text{-}Alg} \arrow[r] & \underline{\mathbf{1}} \end{tikzcd}, \tag{1.2}$$

where $\underline{\mathbf{1}}$ is the category consisting of a single object and just the identity morphism, $\mathbb{BR}$ is the one-object category whose morphisms consist of all real numbers with composition rule given by addition, and the left vertical arrow is the fibration just mentioned. The fibres of the left and right fibrations are convex categories. On the left, one has over each C^*-algebra $\mathcal{A}$, the convex set of states $\mathcal{S}(\mathcal{A})$ on $\mathcal{A}$, which is viewed as a discrete convex category. A morphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ of C^*-algebras gets lifted to the morphism $\mathcal{S}(f) : \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$, which acts as the pullback on states sending $\omega \in \mathcal{S}(\mathcal{A})$ to $\omega \circ f$. On the right, $\mathbb{BR}$ is also a convex category, with convex combinations of real numbers as the convex operation. The entropy difference is in fact a concave (lax affine) fibred functor, denoted by H, and which can be visually represented as



in a more bundle-theoretic fashion.

This entropy difference functor sends a state $\omega \in \mathcal{S}(\mathcal{A})$ together with a morphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ (which corresponds to a morphism in the total space of the fibration) to a real number $H_f(\omega)$. Given another state $\xi \in \mathcal{S}(\mathcal{A})$ and a number $\lambda \in [0,1]$, one obtains the inequality

$$H_f\big(\lambda\omega + (1-\lambda)\xi\big) \geqslant \lambda H_f(\omega) + (1-\lambda)H_f(\xi), \tag{1.3}$$

which is of fundamental importance in quantum information theory. The non-negativity of the quantity

$$\chi_{\lambda;f}(\omega,\xi) := H_f\big(\lambda\omega + (1-\lambda)\xi\big) - \lambda H_f(\omega) - (1-\lambda)H_f(\xi) \tag{1.4}$$

is closely related to the *monotonicity of entropy under partial trace*, which is known to be equivalent (in a certain sense) to *strong subadditivity* [54]. A special case of this inequality, when $f := !_{\mathcal{A}} : \mathbb{C} \to \mathcal{A}$ is the unique unital $*$-homomorphism into $\mathcal{A}$, leads to the fact that mixing always increases entropy. It is actually only this, much weaker, assumption that will play a role in our current characterization.

For more general algebras, if ω and ξ have orthogonal supports, and $f : \mathcal{B} \to \mathcal{A}$ preserves this orthogonality, then equality in (1.3) is obtained. This condition, which we call *orthogonal affinity*, is what replaces the affine assumption of entropy difference made by BFL. However, orthogonal affinity and (1.3) are not enough to guarantee that $H_{\mathcal{A}}(\omega) := H_{!_{\mathcal{A}}}(\omega)$ vanishes on pure states ω. If one imposes this additional assumption, one can show that it is no longer necessary to assume $\chi_{\lambda;f}(\omega,\xi) \geqslant 0$ for all inputs. Instead, one can demand the simpler assumption that $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states ω. In other words, one can replace BFL's non-negativity assumption for classical entropy *difference* with the assumption that $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states ω on C^*-algebras $\mathcal{A}$, with equality for pure states. The relationships between these assumptions will be made precise in the body of the present manuscript. Our main theorem can then be phrased as follows.

Theorem 1.5 [A functorial characterization of quantum entropy (Theorem 4.31 in body)]

Let $H : \mathbf{NCFinProb} \to \mathbb{B}\mathbb{R}$ be a continuous and orthogonally affine fibred functor

$$
\begin{tikzcd}
\mathbf{NCFinProb} \arrow[r, "H"] \arrow[d] & \mathbb{B}\mathbb{R} \arrow[d] \\
\mathbf{fdC^*\text{-}Alg} \arrow[r] & \underline{\mathbf{1}}
\end{tikzcd}
$$

for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$. Then there exists a constant $c \geqslant 0$ such that

$$H_f(\omega) = c\Big(S(\omega) - S(\omega \circ f)\Big)$$

for all morphisms $\mathcal{B} \xrightarrow{f} \mathcal{A}$ of C^*-algebras and states $\omega \in \mathcal{S}(\mathcal{A})$.

In this theorem, $S(\omega)$ is the von Neumann (Segal) entropy of ω, which is given by

$$S(\omega) = -\mathrm{tr}(\rho \log \rho) \tag{1.6}$$

in the special case when ω is a state on $\mathcal{M}_m$, which is always represented by a unique density matrix ρ via $\omega = \mathrm{tr}(\rho\ \cdot\ )$, with tr the (un-normalized) trace. More generally, when $\mathcal{A} := \bigoplus_{x\in X} \mathcal{M}_{m_x}(\mathbb{C})$, a state ω on $\mathcal{A}$ can be described by a collection of states $\omega_x : \mathcal{M}_{m_x}(\mathbb{C}) \rightsquigarrow \mathbb{C}$ and a probability measure $p : \{\bullet\} \rightsquigarrow X$ such that $\omega(A_x) = p_x\omega_x(A_x)$ for $A_x \in \mathcal{M}_{m_x}(\mathbb{C})$ [42, Lemma 5.50]. In this case, the von Neumann entropy of ω is given by

$$S(\omega) = -\sum_{x\in X} p_x \log(p_x) - \sum_{x\in X} p_x \mathrm{tr}(\rho_x \log \rho_x). \tag{1.7}$$

Since all finite-dimensional unital C^*-algebras are of this form (up to isomorphism), this specifies the functor H everywhere (since entropy is invariant under isomorphism).

The goal of this paper is to set up all the relevant terminology in the statement of this theorem, prove the theorem, and then describe how our geometric setup using fibrations of convex categories has the potential to offer a generalized notion of entropy, though we describe the latter in forthcoming work.

The present manuscript is broken up into several parts. We begin by reviewing states, mutual orthogonality, and the standard entropy in Section 2. Section 3 provides (what we hope is) a leisurely introduction to Grothendieck fibrations, fiberwise convex structures, fibered functors, and continuity of fibered functors. Section 4 contains our main result and several others of potential interest. In particular, we prove that our axioms imply the non-negativity of $H_f(\omega)$ for commutative C^*-algebras by using the fact that disintegrations exist for morphisms of commutative probability spaces. More generally, we prove that if a disintegration of (f, ω) exists for arbitrary non-commutative probability spaces, then $H_f(\omega) \geqslant 0$. We also include a brief historical account of axiomatizations of the von Neumann entropy and how our characterization compares with them. We review facts regarding relative entropy in Appendix A and explain how the difference of von Neumann entropies is a concave functor. Although concavity was not used in our characterization of the quantum entropy, we record these facts in the event that they may be useful for further investigations of convex categories and entropy-like functors.

Some comments are to be made on the layout and format of the manuscript. Certain environments are labelled with colors for organizational purposes so that the reader may quickly navigate results catered to their interests. Definitions are green, statements (lemmas, propositions, theorems) are red, examples and physics notes are violet, remarks are yellow, and questions/open problems are blue. The shorthand "iff" stands for "if and only if" and is used exclusively in definitions. Blackboard font will be used for fields, so that $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ are the natural numbers (not including 0), the integers, the real numbers, and the complex numbers, respectively.

2 States on finite-dimensional C^*-algebras

In this section, we set up notation and compile several standard facts that will be used throughout. In particular, since a physicist usually works with density matrices instead of states on C^*-algebras, we review the standard tools in this language. All C^*-algebras will be unital and finite-dimensional and all $*$-homomorphisms will be unital unless stated otherwise. Since all our C^*-algebras will be finite-dimensional, they will always be $*$-isomorphic to direct sums of matrix algebras, so that the reader unfamiliar with C^*-algebras should feel somewhat relieved since most of our analysis will involve only linear algebra. An especially suitable reference including more than enough background is Farenick's linear algebra text [11] (see Theorem 5.20 and Proposition 5.26 in [11] for the statement regarding all finite-dimensional C^*-algebras).

Definition 2.1 [Preliminary definitions]

Let $n \in \mathbb{N}$. A ***density matrix*** on $\mathbb{C}^n$ is a positive matrix (self-adjoint and non-negative eigenvalues) ρ with trace 1. Given a C^*-algebra $\mathcal{A}$, an element $a \in \mathcal{A}$ is ***positive*** iff there exists an $x \in \mathcal{A}$ such that $a = x^*x$. The set of positive elements in $\mathcal{A}$ is denoted by $\mathcal{A}^+$. An element $a \in \mathcal{A}$ is ***self-adjoint*** iff $a^* = a$. An element $p \in \mathcal{A}$ is a ***projection*** iff $p^*p = p$ (equivalently, a self-adjoint nilpotent element). The ***orthogonal complement*** of a projection $p \in \mathcal{A}$ is the element $p^\perp := 1_\mathcal{A} - p$ (and is also a projection). Positivity defines a partial order on self-adjoint elements and one writes $a \geqslant a'$ or $a' \leqslant a$ iff $a - a' \in \mathcal{A}^+$. Given another C^*-algebra $\mathcal{B}$, a ***positive map***[a] $\varphi : \mathcal{B} \rightsquigarrow \mathcal{A}$ is a linear map such that $\varphi(\mathcal{B}^+) \subseteq \mathcal{A}^+$. A ***weight*** on a C^*-algebra $\mathcal{A}$ is a positive map $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$. A weight is called a ***state*** iff it is unital. The set of states on a C^*-algebra $\mathcal{A}$ are denoted by $\mathcal{S}(\mathcal{A})$. A ***non-commutative probability space***[b] is a pair $(\mathcal{A}, \omega)$ consisting of a C^*-algebra together with a state $\omega \in \mathcal{S}(\mathcal{A})$. A ***state-preserving*** map (a $*$-homomorphism or a positive map) from one non-commutative probability space $(\mathcal{B}, \xi)$ to another $(\mathcal{A}, \omega)$ is a map $f : \mathcal{B} \rightsquigarrow \mathcal{A}$ such that $\xi = \omega \circ f$. Given two weights $\omega, \chi : \mathcal{A} \rightsquigarrow \mathbb{C}$, one writes $\chi \leqslant \omega$ iff $\chi(a) \leqslant \omega(a)$ for all $a \in \mathcal{A}^+$. One says χ is ***absolutely continuous*** with respect to ω, written $\chi \preceq \omega$, iff $\omega(a) = 0$ with $a \in \mathcal{A}^+$ implies $\chi(a) = 0$. A state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ is ***pure*** iff it cannot be expressed as a non-trivial convex combination of some pair of distinct states. For the C^*-algebra of $m \times m$ matrices $\mathcal{M}_m$, which is referred to as a ***matrix algebra***, the involution is the conjugate transpose and is denoted by $\dagger$ instead of $*$. If $m = 1$, then $\mathcal{M}_1 \cong \mathbb{C}$ and $\bar{z}$ is used, instead of $z^\dagger$, to denote the complex conjugate of $z \in \mathbb{C}$.

[a] We use the convention that $*$-homomorphisms are always drawn with straight arrows $\rightarrow$ and linear maps between algebras are drawn with squiggly arrows $\rightsquigarrow$. We hope this helps the reader visually distinguish between deterministic maps and stochastic maps. Note that there are some quantum information communities where positive unital maps are called *deterministic* and where the *non-unital* maps are called *non-deterministic*. For them, a deterministic map is one whose Hilbert–Schmidt dual (to be discussed in Example 2.12) sends density matrices to density matrices. We find this terminology potentially misleading because there is a precise correspondence between positive unital maps and stochastic maps (Markov kernels) that restricts to a correspondence between $*$-homomorphisms and functions [37].

[b] In this case "non-commutative" should be read as "not necessarily commutative."

Example 2.2 [Density matrices, states, and expectation values]

Positivity of an $m \times m$ matrix coincides with the C^*-algebraic definition of positivity on $\mathcal{M}_m$ (to see this, use the spectral theorem from linear algebra). Every state ω on $\mathcal{M}_m$ can be expressed as $\omega = \mathrm{tr}(\rho\ \cdot\ )$ for some unique density matrix $\rho \in \mathcal{M}_m$. Here, and everywhere else in this manuscript, tr denotes the un-normalized trace. In this way, the density matrix is viewed as the functional providing the expectation values of observables. More generally, when $\mathcal{A} := \bigoplus_{x\in X} \mathcal{M}_{m_x}$ with X a finite set and $m_x \in \mathbb{N}$, a state ω on $\mathcal{A}$ can be described by a collection of states $\omega_x : \mathcal{M}_{m_x} \rightsquigarrow \mathbb{C}$ and a probability measure $p : \{\bullet\} \rightsquigarrow X$ such that $\omega(A_x) = p_x\omega_x(A_x)$ for $A_x \in \mathcal{M}_{m_x}$ [42, Lemma 5.50]. Here, and elsewhere in the manuscript, p_x is used to denote the probability of x with respect to p. Since each state ω_x corresponds to a density matrix $\rho_x \in \mathcal{M}_{m_x}$, ω can equivalently be expressed by $\omega(A_x) = p_x\mathrm{tr}(\rho_x A_x)$ for $A_x \in \mathcal{M}_{m_x}(\mathbb{C})$. We will also use all of the following notations

$$\omega \equiv \sum_{x\in X} p_x\omega_x \equiv \sum_{x\in X} p_x\mathrm{tr}(\rho_x\ \cdot\ )$$

to indicate the same state. In this way, states generally encode the data of families of expectation values. Since every C^*-algebra $\mathcal{A}$ is isomorphic to a finite direct sum of matrix algebras, this is a full description of states on C^*-algebras. The usefulness of using C^*-algebras as opposed to just matrix algebras is to allow for a combination of classical and quantum setups, such as measurement. Furthermore, direct sums of matrix algebras are used in describing superselection sectors [46, 60], while ensembles, preparations, instruments, etc. are all naturally described by positive maps between certain C^*-algebras that are not just matrix algebras [43, Section 4].

Lemma 2.3 [The support of a weight]

Associated to every weight ω on a C^*-algebra $\mathcal{A}$ is a projection $P_\omega \in \mathcal{A}$ satisfying

$$\omega(P_\omega A) = \omega(AP_\omega) = \omega(P_\omega AP_\omega) = \omega(A) \qquad \forall\ A \in \mathcal{A}$$

and such that $P \leqslant Q$ for every other projection Q satisfying this condition (with Q replacing P_ω).

Definition 2.4 [Supports and mutually orthogonal weights]

The projection P_ω in Lemma 2.3 is called the ***support*** of ω. Two weights ω, ξ on a finite-dimensional C^*-algebra $\mathcal{A}$ are ***mutually orthogonal***, written $\omega \perp \xi$, iff any of the following equivalent conditions hold.[a]

1. If for any weight χ on $\mathcal{A}$ such that $\chi \leqslant \omega$ and $\chi \leqslant \xi$, then $\chi = 0$.
2. $P_\omega P_\xi = 0$ (which implies $P_\omega P_\xi = P_\xi P_\omega$).

A $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ ***preserves the mutual orthogonality*** $\omega \perp \xi$ iff $(\omega \circ f) \perp (\xi \circ f)$.

[a] The thermodynamic meaning of mutual orthogonality of states is summarized in [46, Section 2].

Remark 2.5 [Techniques on positivity]

Any one of the three equalities in Lemma 2.3 implies the other two. Although this is well-documented, the proof illustrates some simple, but useful, techniques, which we may often use without explicitly saying so. For example, suppose $\omega(P_\omega A P_\omega) = \omega(P_\omega A)$ for all $A \in \mathcal{A}$. Then

$$\omega(P_\omega A) = \omega\big((A^* P_\omega)^*\big) = \overline{\omega(A^* P_\omega)} = \overline{\omega(P_\omega A^* P_\omega)} = \omega(P_\omega A P_\omega)$$

since $\omega(A^*) = \overline{\omega(A)}$ for all $A \in \mathcal{A}$ (this is a consequence of positivity of ω). For the last equality, first note that if $A \geqslant 0$, then

$$\begin{aligned} 0 &\leqslant \omega\big(P_\omega^\perp A P_\omega^\perp\big) && \text{since } \omega \text{ is a positive functional and } P_\omega^\perp A P_\omega^\perp \geqslant 0 \\ &\leqslant \|A\| \omega\big(P_\omega^\perp\big) && \text{since } A \leqslant \|A\| 1_{\mathcal{A}} \text{ and } \omega \text{ is linear } (\|\cdot\| \text{ is the norm on } \mathcal{A}) \\ &= 0 && \text{by the previous equality.} \end{aligned}$$

This proves $\omega\big(P_\omega^\perp A P_\omega^\perp\big) = 0$ for $A \geqslant 0$. This equality also holds for arbitrary A since every A can always be expressed as a linear combination of at most four positive elements.[a] Therefore,

$$\omega(A) = \omega\big(AP_\omega + P_\omega A P_\omega^\perp + P_\omega^\perp A P_\omega^\perp\big) = \omega(AP_\omega) + \omega\big(P_\omega A P_\omega^\perp\big) + \omega\big(P_\omega^\perp A P_\omega^\perp\big) = \omega(AP_\omega).$$

[a] Break A up into a self-adjoint and skew-adjoint element, and then spectrally decompose the two results and split the negative and positive terms apart.

Remark 2.6 [The orthogonal complement of the support]

Another equivalent definition of the support of ω is in terms of its orthogonal complement $P_\omega^\perp$. Namely, $P_\omega^\perp$ is the (necessarily unique) largest projection satisfying $\omega(P_\omega^\perp) = 0$, i.e. for any other projection Q with $\omega(Q) = 0$, then $P_\omega^\perp \geqslant Q$.

Lemma 2.7 [The image of a support]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be a $*$-homomorphism and let $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ be a state. Then

$$f(P_{\omega\circ f}^\perp) \leqslant P_\omega^\perp \qquad \text{and} \qquad f(P_{\omega\circ f}) \geqslant P_\omega.$$

Proof. The first inequality follows from the fact that f sends projections to projections and $f(\mathcal{N}_{\omega\circ f}) \subseteq \mathcal{N}_\omega$, where $\mathcal{N}_\xi$ denotes the null-space associated to the state ξ and is given by

$$\mathcal{N}_\xi := \{A \in \mathcal{A} \ : \ \xi(A^* A) = 0\}. \tag{2.8}$$

This is proved in Proposition 3.106 in [42] using the Kadison–Schwarz inequality. The two inequalities are equivalent because

$$f(P_{\omega\circ f}) = f(1_{\mathcal{B}} - P^{\perp}_{\omega\circ f}) = f(1_{\mathcal{B}}) - f(P^{\perp}_{\omega\circ f}) = 1_{\mathcal{A}} - f(P^{\perp}_{\omega\circ f}) = f(P^{\perp}_{\omega\circ f})^{\perp} \geqslant P_{\omega}, \tag{2.9}$$

where the last inequality used $f(P^{\perp}_{\omega\circ f}) \leqslant P^{\perp}_{\omega}$. Note that we have also used (in the 4th equality) the fact that f applied to any projection is again a projection since f is a $*$-homomorphism. A similar calculation shows the converse. ■

Example 2.10 [External convex sums for finite probability spaces]

Let X, X' and Y, Y' be two finite sets, let p and q be probability measures on X and Y, respectively, and let $X \xrightarrow{\phi} X'$ and $Y \xrightarrow{\psi} Y'$ be two functions. Let $\lambda p \oplus (1-\lambda)q$ denote the measure on $X \amalg Y$ (the disjoint union) given by

$$(\lambda p \oplus (1-\lambda)q)_z := \begin{cases} p_z & \text{if } z \in X \\ q_z & \text{if } z \in Y \end{cases}.$$

Set $\mathcal{A} := \mathbb{C}^X$ and $\mathcal{B} := \mathbb{C}^Y$ to be the C^*-algebras of functions on X and Y, and similarly $\mathcal{A}' := \mathbb{C}^{X'}$ and $\mathcal{B}' := \mathbb{C}^{Y'}$. Let ω and ξ be the states on $\mathcal{A}$ and $\mathcal{B}$ associated to p and q (the associated expectation value functionals for the probability measures). Namely, $\omega(A) = \sum_{x\in X} p_x A(x)$ for all $A \in \mathbb{C}^X$ (and similarly for ξ and q). Let $\mathcal{A}' \xrightarrow{f} \mathcal{A}$ and $\mathcal{B}' \xrightarrow{g} \mathcal{B}$ be the $*$-homomorphisms associated to ϕ and ψ via pullback. Namely, if $A' \in \mathbb{C}^{X'}$ is a function on X', then $f(A') := A' \circ \phi$. The disjoint union function $X \amalg Y \xrightarrow{\phi\amalg\psi} X' \amalg Y'$ corresponds to the direct sum $*$-homomorphism $\mathbb{C}^{X'\amalg Y'} \cong \mathcal{A}' \oplus \mathcal{B}' \xrightarrow{f\oplus g} \mathcal{A} \oplus \mathcal{B} \cong \mathbb{C}^{X\amalg Y}$. Let $\widetilde{\omega}$ and $\widetilde{\xi}$ denote the states on $\mathcal{A} \oplus \mathcal{B}$ given by $\widetilde{\omega}(A \oplus B) := \omega(A)$ and $\widetilde{\xi}(A \oplus B) := \xi(B)$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. From these definitions, the state on $\mathcal{A} \oplus \mathcal{B}$ associated to $\lambda p \oplus (1-\lambda)q$ is $\lambda\widetilde{\omega} + (1-\lambda)\widetilde{\xi}$. Furthermore, $\widetilde{\omega} \perp \widetilde{\xi}$ holds and $f \oplus g$ preserves $\widetilde{\omega} \perp \widetilde{\xi}$. This construction of convex sums is one of the main ingredients in BFL's characterization of entropy [4].

Notation 2.11 [Internal direct sum]

Let $m \in \mathbb{N}$ and let $\{n_y\}_{y\in Y}$ be a collection of natural numbers indexed by a finite set Y satisfying $m = \sum_{y\in Y} n_y$. Let $\{B_y \in \mathcal{M}_{n_y}\}_{y\in Y}$ be a collection of matrices. Given an ordering of the elements of Y, set

$$\boxplus_{y\in Y} B_y := \begin{bmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_{|Y|} \end{bmatrix} \equiv \operatorname{diag}(B_1, \dots, B_{|Y|}) \in \mathcal{M}_m.$$

This notation will be frequently used, sometimes without explicitly stating that an order has been chosen. Since the left-hand-side is most compact, we prefer to use it more often.[a]

[a] This is not to be confused with the (external) direct sum $\bigoplus_{y\in Y} B_y \in \bigoplus_{y\in Y} \mathcal{M}_{n_y}$, which does not use an ordering on Y and, more importantly, is an element of a different (non-isomorphic) algebra.

Example 2.12 [The partial trace]

Since we are working with *unital* $*$-homomorphisms between C^*-algebras, we are working in the Heisenberg picture description of quantum mechanics, as opposed to the more commonly used Schrödinger picture in the quantum information theory community. The relationship between the two goes roughly as follows. If $\mathcal{B} = \mathcal{M}_n$, $\mathcal{A} = \mathcal{M}_m$, and $\mathcal{B} \xrightarrow{f} \mathcal{A}$ is a $*$-homomorphism, then there exists a $p \in \mathbb{N}$ such that $m = pn$ and a unitary $U \in \mathcal{M}_m$ such that $f = \mathrm{Ad}_U \circ g$, where $\mathrm{Ad}_U(A) := UAU^\dagger$ for all $A \in \mathcal{A}$ and where g is

$$\mathcal{B} \ni B \mapsto g(B) := \begin{bmatrix} B & & 0 \\ & \ddots & \\ 0 & & B \end{bmatrix}$$

(cf. [1], [58, Lecture 10]). There is an inner product on the vector space of linear maps between $\mathcal{A}$ and $\mathcal{B}$, known as the ***Hilbert–Schmidt*** or ***Frobenius*** inner product. The adjoint, g^*, of g with respect to this inner product is given by

$$\mathcal{A} \cong \mathcal{M}_p \otimes \mathcal{M}_n \ni A \otimes B \mapsto g^*(A \otimes B) = \mathrm{tr}(A)B,$$

which is often written as $\mathrm{tr}_{\mathcal{M}_p}$, since it is precisely the ***partial trace*** operation from quantum information theory (see [42, Remark 3.15] or [34, Section 2.4.3] for more details). The adjoint of f is $g^* \circ \mathrm{Ad}_{U^\dagger}$. The downside of working in this picture is that the partial trace is not a $*$-homomorphism (in general, the Hilbert–Schmidt adjoint of a $*$-homomorphism need not be a $*$-homomorphism). Since we think of $*$-homomorphisms as deterministic maps [37] and since these morphisms form a suitable category, we prefer to express our results in the Heisenberg picture for mathematical ease. However, to make the prose flow more smoothly, we will often use the terminology of partial trace from the Schrödinger picture even when referring to the associated $*$-homomorphism in the Heisenberg picture.

Lemma 2.13 [The partial trace on direct sums]

Let $\mathcal{B} := \bigoplus_{y\in Y} \mathcal{M}_{n_y} \xrightarrow{f} \bigoplus_{x\in X} \mathcal{M}_{m_x} =: \mathcal{A}$ be a $*$-homomorphism and let $\omega = \sum_{x\in X} p_x \mathrm{tr}(\rho_x \,\cdot\,)$ be a state on $\mathcal{A}$ (cf. Example 2.2). Then the following facts hold.

1. There exist non-negative numbers c_{xy}, called the ***multiplicity*** of the factor $\mathcal{M}_{n_y}$ inside $\mathcal{M}_{m_x}$ associated to f, such that $m_x = \sum_{y\in Y} c_{xy} n_y$ for all $x \in X$.

2. There exist unitaries $U_x \in \mathcal{M}_{m_x}$ such that f is of the form

$$\bigoplus_{y\in Y} \mathcal{M}_{n_y} \ni \bigoplus_{y\in Y} B_y \overset{f}{\mapsto} \bigoplus_{x\in X} U_x \left(\boxplus_{y\in Y} \operatorname{diag}(\overbrace{B_y, \cdots, B_y}^{c_{yx} \text{ times}}) \right) U_x^\dagger.$$

3. The pullback state $\omega \circ f$ can be expressed as

$$\omega \circ f = \sum_{y\in Y} q_y \operatorname{tr}(\sigma_y \,\cdot\,), \quad \text{where} \quad q_y\sigma_y = \sum_{x\in X} p_x f^*_{xy}(\rho_x) \qquad \forall\, y \in Y$$

and f^*_{xy} denotes the (Hilbert–Schmidt) adjoint of $f_{xy} : \mathcal{M}_{n_y} \to \mathcal{M}_{m_x}$, which is the component of f mapping between the factors as indicated.

Proof. The first two claims are standard facts (cf. Section 1.1.2 and 1.1.3 in Fillmore [12] or Theorem 5.6 in Farenick [11] for proofs). Given a chosen ordering on the elements of X and Y, such that $s := |X|$ and $t := |Y|$, the image of f on an element B would look like

$$f\left(\begin{pmatrix} B_1 \\ \vdots \\ B_t \end{pmatrix}\right) := \begin{pmatrix} U_1 \operatorname{diag}(\overbrace{B_1, \ldots, B_1}^{c_{11}\text{ times}}, \ldots, \overbrace{B_t, \ldots, B_t}^{c_{1t}\text{ times}}) U_1^\dagger \\ \vdots \\ U_s \operatorname{diag}(\underbrace{B_1, \ldots, B_1}_{c_{s1}\text{ times}}, \ldots, \underbrace{B_t, \ldots, B_t}_{c_{st}\text{ times}}) U_s^\dagger \end{pmatrix} \tag{2.14}$$

in terms of more standard matrix notation (the vector notation is used to indicate the components in the direct sums). The last claim is proved in [43, Lemma 6.7]. ■

Remark 2.15 [Justifying the terminology 'partial trace' for direct sums]

Using the notation from Lemma 2.13, but setting the unitaries to be identities, the dual $\mathcal{A} \overset{f^*}{\rightsquigarrow} \mathcal{B}$ can be viewed as a partial trace in the following sense. First, note that f^* has yx component given by $(f^*)_{yx} = f^*_{xy} : \mathcal{M}_{m_x} \rightsquigarrow \mathcal{M}_{n_y}$. Given a matrix $A_x \in \mathcal{M}_{m_x}$, it can be broken up into a $t \times t$ matrix

$$A_x = \begin{bmatrix} A_{x;11} & \cdots & A_{x;1t} \\ \vdots & & \vdots \\ A_{x;t1} & \cdots & A_{x;tt} \end{bmatrix},$$

where each $A_{x;kl}$ is a $c_{xk}n_k \times c_{xl}n_l$ matrix. In particular, $A_{x;yy} \in \mathcal{M}_{c_{xy}n_y} \cong \mathcal{M}_{c_{xy}} \otimes \mathcal{M}_{n_y}$. Using this notation,

$$f^*_{xy}(A_x) = \operatorname{tr}_{\mathcal{M}_{c_{xy}}}(A_{x;yy}),$$

which justifies calling the dual a partial trace for direct sums.

Lemma 2.16 [$*$-isomorphisms preserve mutual orthogonality]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be $*$-isomorphism and let ω, ξ be any two states on $\mathcal{A}$. Then

$$\omega \perp \xi \qquad \Longrightarrow \qquad (\omega \circ f) \perp (\xi \circ f).$$

Furthermore, under the same assumption on f, a state ζ on $\mathcal{A}$ is pure if and only if $\zeta \circ f$ is pure.

Proof. If P_ω and P_ξ are the supports of ω and ξ, respectively, then the claim will follow if we prove that $f^{-1}(P_\omega)$ and $f^{-1}(P_\xi)$ are the supports of $\omega \circ f$ and $\xi \circ f$, respectively, because $f^{-1}(P_\omega)f^{-1}(P_\xi) = f^{-1}(P_\omega P_\xi) = f^{-1}(0) = 0$. It suffices to focus on ω. First, note that $f^{-1}(P_\omega)$ is a projection since f^{-1} is a $*$-homomorphism. Furthermore,

$$(\omega \circ f)\big(f^{-1}(P_\omega)B\big) = \omega\big(P_\omega f(B)\big) = \omega\big(f(B)\big) = (\omega \circ f)(B) \qquad \forall\, B \in \mathcal{B}, \tag{2.17}$$

which proves that $f^{-1}(P_\omega)$ satisfies the first condition of a support for $\omega \circ f$ in Lemma 2.3 (cf. Remark 2.5). Suppose that Q is another projection satisfying $(\omega \circ f)(QB) = (\omega \circ f)(B)$ for all $B \in \mathcal{B}$. Then $f(Q)$ satisfies

$$\omega\big(f(Q)A\big) = (\omega \circ f)\big(Qf^{-1}(A)\big) = (\omega \circ f)\big(f^{-1}(A)\big) = \omega(A) \qquad \forall\, A \in \mathcal{A}. \tag{2.18}$$

Hence, since P_ω is the minimal such projection, $P_\omega \leqslant f(Q)$. Since $*$-homomorphisms preserve the $\leqslant$ order structure, $f^{-1}(P_\omega) \leqslant Q$. ■

Example 2.19 [Channels that do not preserve orthogonality]

There are certainly examples of $*$-homomorphisms $\mathcal{B} \to \mathcal{A}$ that do not always preserve mutual orthogonality. A simple example is $!_{\mathbb{C}^2} : \mathbb{C} \to \mathbb{C}^2$, where every pair of mutually orthogonal states gets pulled back to 1. A non-classical example is the $*$-homomorphism $\mathcal{M}_2 \to \mathcal{M}_2 \otimes \mathcal{M}_2$, sending B to $B \otimes \mathbb{1}_2$, and any two density matrices on $\mathbb{C}^2 \otimes \mathbb{C}^2$ corresponding to any two orthogonal Bell states [34, Section 2.3]. In this case, the pullback state is $\frac{1}{2}\mathrm{tr}$.

Lemma 2.20 [Overlapping states remain overlapping under evolution]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be $*$-homomorphism and let ω, ξ be two states on $\mathcal{A}$ that are *not* mutually orthogonal. Then $\omega \circ f$ and $\xi \circ f$ are also not mutually orthogonal.

Proof. Suppose, to the contrary, that $P_{\omega\circ f}P_{\xi\circ f} = 0$. Then

$$0 = f(0) = f(P_{\omega\circ f}P_{\xi\circ f}) = f(P_{\omega\circ f})f(P_{\xi\circ f}). \tag{2.21}$$

But, by Lemma 2.7, $f(P_{\omega\circ f}) \geqslant P_\omega$ and $f(P_{\xi\circ f}) \geqslant P_\xi$ so that their product cannot vanish by the assumption $P_\omega P_\xi \neq 0$. Thus, $0 \neq 0$, a contradiction. ■

Physics 2.22 [Evolving states with overlapping supports]

The interpretation of Lemma 2.20 is that if two states have overlapping supports, then no quantum operation will ever completely separate them. In contrast, Lemma 2.16 says that reversible dynamics (such as unitary evolution) cannot mix states.

Now that we have defined the objects and morphisms of interest, we can define entropy and its generalizations to matrix algebras and C^*-algebras.

Definition 2.23 [Shannon, von Neumann, and Segal entropy]

Let p be a probability measure on a finite set X whose value at $x \in X$ is denoted by p_x. Then the ***Shannon entropy*** of p is the non-negative number

$$S_{\mathrm{Sh}}(p) := -\sum_{x \in X} p_x \log(p_x)$$

(with the convention $0 \log 0 := 0$). Let ρ be a density matrix on $\mathbb{C}^n$. The ***von Neumann entropy*** of ρ is the non-negative number

$$S_{\mathrm{vN}}(\rho) := -\mathrm{tr}\big(\rho \log \rho\big),$$

which is defined by the functional calculus. Let ω be a state on $\mathcal{A} = \bigoplus_{x \in X} \mathcal{M}_{m_x}$ with decomposition given by $\omega = \sum_{x \in X} p_x \mathrm{tr}(\rho_x \,\cdot\,)$ as in Example 2.2. The ***Segal entropy*** of ω is the non-negative number

$$S_{\mathrm{Se}}(\omega) := S_{\mathrm{Sh}}(p) + \sum_{x \in X} p_x S_{\mathrm{vN}}(\rho_x).$$

On occasion, the letter S will exclusively be used to refer to any of these three definitions, using the input to distinguish which formula should be used. As such, ***entropy*** will refer to any of these three, while ***quantum entropy*** will refer to either of the last two.[a]

[a] The Segal entropy was actually defined much more generally for certain infinite-dimensional systems [48]. Therefore, since we're working in finite dimensions, this could also be called the von Neumann entropy (see Remark 2.25 for some more details).

To motivate the definition of the Segal entropy, we recall the following useful fact about the entropy of convex combinations.

Lemma 2.24 [Inequalities for concavity of entropy]

Let $\{\rho_x\}_{x \in X}$ be a collection of density matrices on a Hilbert space indexed by a finite set X. Then

$$\sum_{x \in X} p_x S_{\mathrm{vN}}(\rho_x) \leqslant S_{\mathrm{vN}}\left(\sum_{x \in X} p_x \rho_x\right) \leqslant S_{\mathrm{Sh}}(p) + \sum_{x \in X} p_x S_{\mathrm{vN}}(\rho_x)$$

for any nowhere-vanishing probability distribution p on X. Furthermore, the second inequality becomes an equality if and only if $\rho_x \perp \rho_{x'}$ for all distinct $x, x' \in X$.

Proof. The first inequality is the concavity of the von Neumann entropy. Proofs of these claims can be found in [34, Theorem 11.8 (4)] as well [27, Corollary pg 247] and [28, equation (2.2)]. ∎

Remark 2.25 [Motivating the Segal entropy]

Given a state $\omega = \sum_{x\in X} p_x \mathrm{tr}(\rho_x\ \cdot\ )$ as in Definition 2.23, the Segal entropy also equals

$$S_{\mathrm{Se}}(\omega) = S_{\mathrm{vN}}(\rho), \quad \text{where} \quad \rho := \boxplus_{x\in X} p_x\rho_x \in \mathcal{M}_m \quad \text{and} \quad m := \sum_{x\in X} m_x$$

(upon choosing an ordering of the elements of X). This follows from the mutual orthogonality between $p_x\rho_x$ and $p_{x'}\rho_{x'}$ for distinct $x, x' \in X$ by Lemma 2.24.

We now come to the crucial definition for the entropy change along a morphism.

Definition 2.26 [The entropy change along a morphism]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be a $*$-homomorphism of C^*-algebras and let ω be a state on $\mathcal{A}$. The ***entropy change of*** ω ***along*** f is the number

$$S_f(\omega) := S_{\mathrm{Se}}(\omega) - S_{\mathrm{Se}}(\omega \circ f).$$

The following lemma contains a crucial observation that distinguishes the entropy change along a morphism between commutative C^*-algebras and the entropy change along a morphism where at least one of the C^*-algebras is necessarily not commutative.

Lemma 2.27 [The entropy change along certain morphisms]

Recall the notation from Definition 2.26.

1. If f is a $*$-isomorphism, then $S_f(\omega) = 0$.
2. If $\mathcal{A}$ and $\mathcal{B}$ are commutative C^*-algebras, then $S_f(\omega) \geqslant 0$ for all states ω and $*$-homomorphisms $\mathcal{B} \xrightarrow{f} \mathcal{A}$.
3. If $\mathcal{A}$ is not commutative and f is not a $*$-isomorphism, then there exists a state $\omega \in \mathcal{S}(\mathcal{A})$ such that $S_f(\omega) < 0$.[a]

[a] If $\mathcal{B}$ is not commutative, then a $*$-homomorphism $\mathcal{B} \to \mathcal{A}$ does not exist if $\mathcal{A}$ is commutative. This is related to the fact that there is no *hard evidence* in quantum mechanics (see [41] for further details).

Proof.

1. Let f be a $*$-isomorphism and write decompositions $\mathcal{B} := \bigoplus_{y\in Y} \mathcal{M}_{n_y}$ and $\mathcal{A} := \bigoplus_{x\in X} \mathcal{M}_{m_x}$. Set $\xi := \omega \circ f$ and let $\omega = \sum_{x\in X} p_x \mathrm{tr}(\rho_x\ \cdot\ )$ be as in Example 2.2. Similarly, write

$\xi = \sum_{y \in Y} q_y \mathrm{tr}(\sigma_y \ \cdot \ )$. Since f is a $*$-isomorphism, there exists a bijection $X \xrightarrow{\phi} Y$ and a collection of unitaries $U_x \in \mathcal{M}_{m_x}$ such that

$$m_x = n_{\phi(x)} \quad \text{and} \quad p_x U_x \rho_x U_x^\dagger = q_{\phi(x)} \sigma_{\phi(x)} \qquad \forall\, x \in X. \tag{2.28}$$

This last claim follows from the form of $*$-homomorphisms between C^*-algebras (see Lemma 2.13 above or Lecture 10 in [58] for example). The claim $S_f(\omega) = 0$ then follows from the functional calculus and Definition 2.23.

2. The claim that $S_f(\omega) \geqslant 0$ when $\mathcal{A}$ and $\mathcal{B}$ are commutative follows from the fact that every commutative finite-dimensional C^*-algebra is isomorphic to functions on a finite set as described in Example 2.10. In this case, the Segal entropy becomes the Shannon entropy. If p is the probability measure on X corresponding to ω and q is the probability measure on Y corresponding to $\omega \circ f$, then

$$S_f(\omega) = S_{\mathrm{Se}}(\omega) - S_{\mathrm{Se}}(\omega \circ f) = S_{\mathrm{Sh}}(p) - S_{\mathrm{Sh}}(q), \tag{2.29}$$

which is shown to be non-negative in [4] (we will also provide a self-contained proof of this result more abstractly in Proposition 4.8 using disintegrations).

3. If $\mathcal{A}$ is not commutative, then it has some matrix algebra $\mathcal{M}_m$ as a factor with $m > 1$. Let ρ be a rank 1 density matrix in $\mathcal{A}$ with support in $\mathcal{M}_m$ (so that ρ is a pure state). Let A be a self-adjoint $m \times m$ matrix that does not commute with ρ (such a matrix necessarily exists because the center of $\mathcal{M}_m$ consists of multiples of the identity). Let $\sigma(A)$ denote the spectrum of A. Let $\mathcal{B} := \mathbb{C}^{\sigma(A)} \xrightarrow{f} \mathcal{A}$ send e_λ, the function on $\sigma(A)$ whose value at λ is 1 and is 0 elsewhere, to P_λ in $\mathcal{M}_m$, the projection onto the λ-eigenspace. Then $\omega \circ f$ is not a pure state, in the sense that the associated measure on $\sigma(A)$ is not a Dirac measure. Thus, the entropy change is $S_f(\omega) = S_{\mathrm{Se}}(\omega) - S_{\mathrm{Se}}(\omega \circ f) = 0 - S_{\mathrm{Se}}(\omega \circ f) \leqslant 0$. ■

Item 2 in Lemma 2.27 was used as an axiom by BFL to characterize the entropy change in the classical setting. Since it fails when one includes non-commutative C^*-algebras, we will have to replace this axiom with one that more accurately reflects the properties of entropy in quantum mechanics.

Physics 2.30 [The uncertainty principle and negative conditional entropy]

The proof of item 3 in Lemma 2.27 illustrates a crucial feature of measurement in quantum mechanics. Even if one begins with a pure state on $\mathcal{A}$, measuring an observable A that does not commute with the density matrix associated to the pure state necessarily gives rise to a non-deterministic probability distribution on the eigenvalues (allowed observed values for the observable). This is one way of looking at the uncertainty principle. As another example illustrating the validity of item 3 in Lemma 2.27 using only matrix algebras, take the state ω on $\mathcal{M}_2 \otimes \mathcal{M}_2 \cong \mathcal{M}_4$ to be the EPR state and let $\mathcal{M}_2 \xrightarrow{f} \mathcal{M}_2 \otimes \mathcal{M}_2$ be the inclusion into one of the factors. Then $S_f(\omega) = -\log(2)$ (cf. Example 2.19). More generally, set $\mathcal{A} := \mathcal{M}_m$, $\mathcal{B} := \mathcal{M}_n$, $\mathcal{A} \xrightarrow{f} \mathcal{A} \otimes \mathcal{B}$ the standard inclusion, and $\omega = \mathrm{tr}(\rho_{\mathcal{A}\mathcal{B}} \ \cdot \ )$, where $\rho_{\mathcal{A}\mathcal{B}}$ is a density matrix in $\mathcal{A} \otimes \mathcal{B}$ with marginals $\rho_{\mathcal{A}} := \mathrm{tr}_{\mathcal{B}}(\rho_{\mathcal{A}\mathcal{B}})$ and $\rho_{\mathcal{B}} := \mathrm{tr}_{\mathcal{A}}(\rho_{\mathcal{A}\mathcal{B}})$ (cf. Example 2.12). Then the entropy difference $S_f(\omega) = S_{\mathrm{vN}}(\rho_{\mathcal{A}\mathcal{B}}) - S_{\mathrm{vN}}(\rho_{\mathcal{A}})$ is exactly the quantum conditional entropy, which, if negative, necessarily implies that $\rho_{\mathcal{AB}}$ is entangled (see near equation (21) in [23]).[a] The negativity of this expression has been recently given an operational meaning in terms of state merging [22]. The example we chose in the proof of Lemma 2.27 is meant to illustrate that entanglement is not necessary for $S_f(\omega)$ to be negative.

[a]The reader will also be able to prove the claim on their own using the additivity property of entropy and Lieb and Ruskai's theorem on the concavity of the entropy difference associated to a partial trace, which we review at the beginning of the proof of Theorem A.1 later.

We now end this section with a summary of the categories that will be used throughout.

Definition 2.31 [Categories used in this work]

In all categories that follow, except the very last one, the composition rule with be function composition.

1. **FinSet** is the category whose objects are finite sets and whose morphisms are functions.
2. **FinProb** is the category whose objects are ***finite probability spaces***, which are pairs (X, p) with X a finite set and p a probability measure on X. A morphism from (X, p) to (Y, q) is a ***probability-preserving function***, i.e. a function $X \xrightarrow{\phi} Y$ such that $q = \phi_* p$ is the pushforward of p along ϕ. In terms of probabilities at points, $q_y = \sum_{x \in \phi^{-1}(\{y\})} p_x$ for all $y \in Y$, where $\phi^{-1}(\{y\}) := \{x \in X : \phi(x) = y\}$.
3. **fdC*-Alg** is the category whose objects are (finite-dimensional unital) C^*-algebras and morphisms are (unital) $*$-homomorphisms.
4. **NCFinProb** is the category whose objects are (finite-dimensional) non-commutative probability spaces and whose morphisms are state-preserving (unital) $*$-homomorphisms.
5. $\mathbb{BR}$ ($\mathbb{BR}_{\geqslant 0}$) is the category consisting of a single object and whose morphisms from that object to itself are all real numbers (non-negative real numers) equipped with addition as the composition rule.

Finally, here are some additional categorical notations and terminologies that will be used.

- If $\mathcal{C}$ is a category, $\mathcal{C}^{\mathrm{op}}$ denotes its opposite. This means if x and y are objects in $\mathcal{C}$, then a morphism from x to y in $\mathcal{C}^{\mathrm{op}}$ is a morphism from y to x in $\mathcal{C}$.
- Given two categories $\mathcal{C}$ and $\mathcal{D}$, let $\mathcal{C} \times \mathcal{D}$ denote their cartesian product. Let $\mathcal{C} \times \mathcal{D} \xrightarrow{\gamma} \mathcal{D} \times \mathcal{C}$ be the functor that swaps the two inputs. Let $\mathcal{C} \xrightarrow{\Delta} \mathcal{C} \times \mathcal{C}$ be the diagonal functor sending an object x to (x, x) and similarly for morphisms. There are two projection functors, denoted by $\mathcal{C} \times \mathcal{D} \xrightarrow{\pi_1} \mathcal{C}$ and $\mathcal{C} \times \mathcal{D} \xrightarrow{\pi_2} \mathcal{D}$.
- A category is ***discrete*** iff its morphisms consist only of identities.

3 Fibrations and local convex structures

We have found that the language of fibrations is a convenient and geometrically appealing setting to formulate the notion of entropy change as a functor. Non-commutative probability spaces form a discrete fibration over C^*-algebras and the real numbers viewed as a one-object category form an ordinary (Grothendieck) fibration over the trivial category. The fibre over each algebra is the space of states, which comes equipped with a convex structure. Since real numbers have a convex structure as well, one can make sense of convexity, concavity, or affinity of the functor that computes the entropy change along a morphism of non-commutative probability spaces. The references for fibrations that we follow include [20, 30, 32]. We first review discrete fibrations since their definition is simpler.

Definition 3.1 [Discrete fibration]

A functor $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ is a ***discrete fibration***[a] iff for each morphism $x \xrightarrow{f} y$ in $\mathcal{X}$ and for each object v in $\mathcal{E}$ such that $\pi(v) = y$, there exists a unique morphism $u \xrightarrow{\beta} v$ such that $\pi(\beta) = f$. A morphism $u \xrightarrow{\beta} v$ such that $\pi(\beta) = f$ is called a ***lift*** of f.

[a] The reason for the word "discrete" in "discrete fibration" will be explained in Example 3.8.

Remark 3.2 [Lifts for fibrations]

It is convenient to visually represent the unique lifting in Definition 3.1 as

$$\begin{tikzcd} \mathcal{E} \arrow[d, "\pi"'] & u \arrow[r, dashed, "\exists!\beta"] & v \\ \mathcal{X} & x \arrow[r, "f"'] & y \end{tikzcd} .$$

Example 3.3 [The discrete fibration of non-commutative probability spaces]

The functor $\pi : \mathbf{NCFinProb} \to \mathbf{fdC^*\text{-}Alg}$, that sends a non-commutative probability space $(\mathcal{A}, \omega)$ to the underlying C^*-algebra $\mathcal{A}$ and a state-preserving $*$-homomorphism $(\mathcal{B}, \xi) \xrightarrow{f} (\mathcal{A}, \omega)$ to the underlying $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$, is a discrete fibration, which we visualize as

$$\begin{tikzcd} \mathbf{NCFinProb} \arrow[d, "\pi"'] & (\mathcal{B}, \xi) \arrow[r, "f"] & (\mathcal{A}, \omega) \\ \mathbf{fdC^*\text{-}Alg} & \mathcal{B} \arrow[r, "f"'] & \mathcal{A} \end{tikzcd} .$$

Indeed, given $\omega \in \mathcal{S}(\mathcal{A})$ and $\mathcal{B} \xrightarrow{f} \mathcal{A}$, the unique lift is f itself together with the state on $\mathcal{B}$

given by $\xi = \omega \circ f$.

Example 3.4 [The discrete fibration of finite probability spaces]

Another closely related example is the functor $\mathbf{FinProb}^{\mathrm{op}} \to \mathbf{FinSet}^{\mathrm{op}}$, which sends a probability space (X, p) to X and a probability-preserving function to the underlying function between sets. The $^{\mathrm{op}}$ in the categories is to guarantee that the fibration property holds. In this case, it looks like

$$
\begin{tikzcd}
\mathbf{FinProb}^{\mathrm{op}} \arrow[d] \\
\mathbf{FinSet}^{\mathrm{op}}
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
(X, p) \arrow[r, "\phi"] & (Y, q) \\
X \arrow[r, "\phi"'] & Y
\end{tikzcd}
.
$$

Such a fibration is sometimes called a ***discrete opfibration*** due to the reversal of arrows.

Definition 3.5 [Cartesian morphisms]

Let $\mathcal{E}$ and $\mathcal{X}$ be two categories and let $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ be a functor. A morphism $u \xrightarrow{\beta} v$ in $\mathcal{E}$ is ***cartesian*** iff for any morphism $x \xrightarrow{f} \pi(u)$ in $\mathcal{X}$ and any morphism $w \xrightarrow{\gamma} v$ in $\mathcal{E}$ such that $\pi(\beta) \circ f = \pi(\gamma)$, there exists a unique morphism $w \xrightarrow{\alpha} u$ in $\mathcal{E}$ such that

$$\pi(\alpha) = f \qquad \text{and} \qquad \beta \circ \alpha = \gamma.$$

Remark 3.6 [Visualizing cartesian morphisms]

A visualization of the data in Definition 3.5 is often helpful:[a]

$$
\begin{tikzcd}
\mathcal{E} \arrow[d, "\pi"'] \\
\mathcal{X}
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
w \arrow[dr, dashed, "\exists!\alpha"'] \arrow[drr, bend left, "\gamma"] & & \\
& u \arrow[r, "\boldsymbol{\beta}"] & v \\
x \arrow[dr, "f"'] \arrow[drr, bend left, "\pi(\gamma)"] & & \\
& \pi(u) \arrow[r, "\pi(\beta)"] & \pi(v)
\end{tikzcd}
.
$$

Note that the conditions necessitate $\pi(w) = x$ so that one can think of the morphisms in $\mathcal{E}$ as lying above the morphisms in $\mathcal{X}$ via the functor π.

[a] The morphism β has been bolded to emphasize that it is the morphism that is cartesian.

Definition 3.7 [Fibration]

Let $\pi : \mathcal{E} \to \mathcal{X}$ be as in Definition 3.5. Let $\mathcal{E}_x$ be the subcategory of $\mathcal{E}$ consisting of the objects u in $\mathcal{E}$ such that $\pi(u) = x$ and $\pi(\beta) = \mathrm{id}_x$ for all morphisms $u \xrightarrow{\beta} v$ with $\pi(u) = x = \pi(v)$. The category $\mathcal{E}_x$ is called the ***fibre*** of π over x and the morphisms in $\mathcal{E}_x$ are called ***vertical morphisms*** of π over x. Given a morphism $x \xrightarrow{f} y$ in $\mathcal{X}$ and an object v in $\mathcal{E}_y$, a ***cartesian lifting*** of f with target v is a cartesian morphism $u \xrightarrow{\beta} v$ such that $\pi(\beta) = f$. A functor $\pi : \mathcal{E} \to \mathcal{X}$ is a ***fibration*** iff for any morphism $x \xrightarrow{f} y$ in $\mathcal{X}$ and an object v in $\mathcal{E}_y$, a cartesian lifting exists. When π is a fibration, $\mathcal{E}$ is called the ***total category*** and $\mathcal{X}$ is called the ***base***. A fibration for which a cartesian lifting has been chosen for every pair (f, v), with f a morphism in $\mathcal{X}$ and v an object in $\mathcal{E}_y$, is called a ***cloven*** fibration.

Example 3.8 [Examples of fibrations]

(a) Every discrete fibration is a cloven fibration (in a unique way). In fact, the fibre over any object in a discrete fibration is a discrete category, hence the name. In particular, **NCFinProb** $\to$ **fdC*-Alg** is a fibration.

(b) Let $\mathcal{X}$ and $\mathcal{F}$ be two categories. The standard projection $\pi : \mathcal{X} \times \mathcal{F} \to \mathcal{X}$ is a fibration. Indeed, if $x \xrightarrow{f} y$ is a morphism in $\mathcal{X}$ and b is an object in $\mathcal{F}$, then the morphism $(x, b) \xrightarrow{(f, \mathrm{id}_b)} (y, b)$ is a cartesian lift of f. Such a fibration is called a ***trivial fibration***.

(c) A special case of a trivial fibration is when $\mathcal{X} = \bullet$ (a trivial one object category) and $\mathcal{F} = \mathbb{BR}$. Another example is when $\mathcal{F} = \mathbb{BR}_{\geqslant 0}$. In both cases, all morphisms are vertical. In $\mathbb{BR}$, every morphism is cartesian, while in $\mathbb{BR}_{\geqslant 0}$ only 0 is. Note that $\mathbb{BR} \to \bullet$ is not a discrete fibration because $\mathbb{BR}$ is not a discrete category.

Lemma 3.9 [The reindexing functor]

Let $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ be a cloven fibration for which the choice of cartesian lift for any (f, v), with $x \xrightarrow{f} y$ a morphism in $\mathcal{X}$ and v an object in $\mathcal{E}_y$, is written as $f^*(v) \xrightarrow{f_v} v$. These data determine a canonical functor $\mathcal{E}_x \xleftarrow{f^*} \mathcal{E}_y$ sending v to $f^*(v)$. For each morphism $w \xrightarrow{\kappa} v$ in $\mathcal{E}_y$, let $f^*(w) \xrightarrow{f^*(\kappa)} f^*(v)$ be the unique morphism in $\mathcal{E}_x$ obtained by the universal property of f_v being a cartesian morphism

$$\begin{tikzcd} \mathcal{E} \arrow[d, "\pi"'] \\ \mathcal{X} \end{tikzcd} \qquad \begin{tikzcd} f^*(w) \arrow[dr, dashed, "\exists! f^*(\kappa)"'] \arrow[drr, bend left, "\kappa \circ f_w"] & & \\ & f^*(v) \arrow[r, "f_v"] & v \\ x \arrow[dr, "\mathrm{id}_x"'] \arrow[drr, bend left, "f"] & & \\ & x \arrow[r, "f"] & y \end{tikzcd}$$

and where $f^*(w) \xrightarrow{f_w} v$ is the chosen cartesian lift of (f, w). Then f^* defines a functor, called the ***reindexing functor*** associated to f.

Proof. This is a standard fact that follows from the uniqueness in the universal property of cartesian morphisms. The details are left as an exercise. ■

We now discuss convex structures on our main examples. For this, we introduce the notion of a (strict) convex category, affine functors, fibrewise convex structures on fibrations, and certain convex functors, some of which have been introduced in [36, Chapter 4]. However, what follows is a self-contained and simplified presentation suitable for our purposes.

Notation 3.10 [Successive convex combinations]

Given two numbers $\lambda, \mu \in [0, 1]$ set

$$\lambda \llcorner \mu := \lambda\mu \qquad \& \qquad \lambda \lrcorner \mu := \begin{cases} \frac{\lambda(1-\mu)}{1-\lambda\mu} & \text{if } \lambda\mu \neq 1 \\ \text{arbitrary} & \text{if } \lambda = \mu = 1 \end{cases},$$

where "arbitrary" means that one can assign any value to the quantity.

If the following definition of a convex category is too abstract on a first reading, the reader is encouraged to look at Remark 3.12 first.

Definition 3.11 [Convex category]

A ***(strict) convex category*** (or more generally a ***strict convex object*** in some cartesian monoidal category) is a category $\mathcal{C}$ (or object) together with a family of functors $F_\lambda : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ (or morphisms) indexed by $\lambda \in [0, 1]$ such that the following diagrams

$$\begin{tikzcd} \mathcal{C} \times \mathcal{C} \arrow[r, bend left, "F_0"] \arrow[r, bend right, "\pi_2"'] & \mathcal{C} \end{tikzcd}, \qquad \begin{tikzcd} \mathcal{C} \times \mathcal{C} \arrow[rr, "\gamma"] \arrow[dr, "F_\lambda"'] & & \mathcal{C} \times \mathcal{C} \arrow[dl, "F_{1-\lambda}"] \\ & \mathcal{C} & \end{tikzcd},$$

$$\begin{tikzcd} \mathcal{C} \times \mathcal{C} \arrow[rr, "\gamma"] \arrow[dr, "F_\lambda"'] & & \mathcal{C} \times \mathcal{C} \arrow[dl, "F_{1-\lambda}"] \\ & \mathcal{C} & \end{tikzcd}, \text{ and } \qquad \begin{tikzcd} \mathcal{C} \times \mathcal{C} \times \mathcal{C} \arrow[r, "F_\mu \times \mathrm{id}_\mathcal{C}"] \arrow[d, "\mathrm{id}_\mathcal{C} \times F_{\lambda \lrcorner \mu}"'] & \mathcal{C} \times \mathcal{C} \arrow[d, "F_\lambda"] \\ \mathcal{C} \times \mathcal{C} \arrow[r, "F_{\lambda \llcorner \mu}"'] & \mathcal{C} \end{tikzcd}$$

commute for all $\lambda, \mu \in [0, 1]$ (cf. Definition 2.31 for notation) Commutativity of the last diagram is called ***parametric associativity***. All convex categories here will be strict so that the adjective will be henceforth dropped. A convex category is typically denoted by $(\mathcal{C}, \{F_\lambda\})$, $(\mathcal{C}, F)$, or sometimes abusively as just $\mathcal{C}$.

Remark 3.12 [Convex categories allow convex combinations of objects and morphisms]

Using the notation from Definition 3.11, it will be convenient to occasionally write

$$\lambda x + (1-\lambda)y := F_{\lambda}(x,y)$$

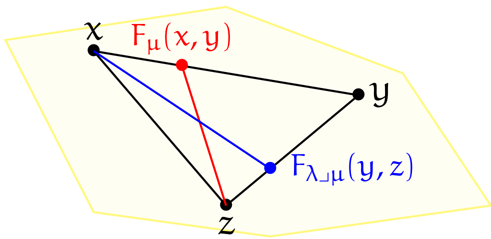
for objects x, y in $\mathcal{C}$ (and similarly for morphisms). In this case, the laws take a more familiar form:

$$\begin{aligned} 0x + 1y &= y \\ \lambda x + (1-\lambda)x &= x \\ \lambda x + (1-\lambda)y &= (1-\lambda)y + \lambda x \\ \lambda\Big(\mu x + (1-\mu)y\Big) + (1-\lambda)z &= (\lambda\llcorner\mu)x + (1-\lambda\llcorner\mu)\Big((\lambda\lrcorner\mu)y + (1-\lambda\lrcorner\mu)z\Big). \end{aligned}$$

The last formula reads

$$\lambda\Big(\mu x + (1-\mu)y\Big) + (1-\lambda)z = \lambda\mu x + (1-\lambda\mu)\left(\frac{\lambda(1-\mu)}{1-\lambda\mu}y + \frac{1-\lambda}{1-\lambda\mu}z\right)$$

if we plug in the definitions and assume $\lambda\mu \neq 1$. The formula for parametric associativity becomes significantly more transparent when one draws a picture.



The expression $F_{\lambda}\Big(F_{\mu}(x,y),z\Big)$ says to begin at x and travel $(1-\mu)$-th the way from x to y and then to travel $(1-\lambda)$-th the way from $F_{\mu}(x,y)$ to z. Meanwhile, the expression $F_{\lambda\llcorner\mu}\Big(x, F_{\lambda\lrcorner\mu}(y,z)\Big)$ says to begin at z and travel $\lambda\lrcorner\mu$-th the way from z to y and then $\lambda\llcorner\mu$-th the way from $F_{\lambda\lrcorner\mu}(y,z)$ to x. One arrives at the same destination in either case. One can also view these convex combinations in terms of degree of mixing [18], which is particularly appropriate for states on physical systems.

Example 3.13 [Examples of convex categories]

We list several examples here that will be used throughout.

(a) Every convex set is a convex category when viewed as a discrete category.

(b) In particular, if $\mathcal{A}$ is a C^*-algebra, then $\mathcal{S}(\mathcal{A})$, the set of states on $\mathcal{A}$, is a convex category.

(c) Let $\mathbb{B}\mathbb{R}$ be the one object category whose set of morphisms is $\mathbb{R}$, the set of real numbers, together with addition as the composition of morphisms. Here, the convex combination morphism $F_\lambda : \mathbb{B}\mathbb{R} \times \mathbb{B}\mathbb{R} \to \mathbb{B}\mathbb{R}$ is a functor. Since $\mathbb{B}\mathbb{R}$ has only one object, this functor acts as the identity on objects. Given $x, y \in \mathbb{R}$ (morphisms in $\mathbb{B}\mathbb{R}$), it acts as the usual convex combination $F_\lambda(x, y) := \lambda x + (1-\lambda)y$. Functoriality itself says $F_\lambda(w+y, x+z) = F_\lambda(w, x) + F_\lambda(y, z)$, which is a form of distributivity. If $\mathbb{R}_{\geqslant 0} := \{r \in \mathbb{R} : r \geqslant 0\}$, then $\mathbb{B}\mathbb{R}_{\geqslant 0}$ is also a convex category.

Note, however, that the convex categories of BFL [4] are not examples of Definition 3.11 (see Remark 3.34 for more details).

Remark 3.14 [Abstracting the notion of a convex space]

The definition of a strict convex category was obtained by internalizing the definition of a convex space, which we learned about in [14] and [13]. However, the idea of an abstract convex space goes back to work by W. Neumann, Gudder, and Świrszcz in the early 1970's [18, 19, 33, 50], Stone from 1939 [49], and von Neumann and Morgenstern from the late 1920's to early 1940's [56]. Although this list is by no means complete, it is meant to illustrate that the abstract notions of convex sets have been independently rediscovered several times.

Definition 3.15 [Affine functors]

An *affine functor* from one convex category $(\mathcal{C}, \{F_\lambda\})$ to another one $(\mathcal{D}, \{G_\lambda\})$ is a functor $S : \mathcal{C} \to \mathcal{D}$ such that the diagram

$$
\begin{tikzcd}
\mathcal{C} \times \mathcal{C} \arrow[r, "S \times S"] \arrow[d, "F_\lambda"'] & \mathcal{D} \times \mathcal{D} \arrow[d, "G_\lambda"] \\
\mathcal{C} \arrow[r, "S"'] & \mathcal{D}
\end{tikzcd}
$$

commutes for all $\lambda \in [0, 1]$.

Example 3.16 [The pullback of states is an affine functor]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be a $*$-homomorphism (or a positive map) between C^*-algebras. Then the pullback $\mathcal{S}(\mathcal{A}) \xrightarrow{\mathcal{S}(f)} \mathcal{S}(\mathcal{B})$, sending ω to $\omega \circ f$, is an affine functor (when $\mathcal{S}(\mathcal{A})$ and $\mathcal{S}(\mathcal{B})$ are viewed as discrete categories) since

$$
\big(\lambda\omega + (1-\lambda)\xi\big) \circ f = \lambda(\omega \circ f) + (1-\lambda)(\xi \circ f) \qquad \forall\, \lambda \in [0, 1],\ \omega, \xi \in \mathcal{S}(\mathcal{A}).
$$

Example 3.17 [Entropy is almost affine]

Given a $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$, the assignment $\mathcal{S}(\mathcal{A}) \xrightarrow{S_f} \mathbb{B}\mathbb{R}$ sending ω to $S_f(\omega)$ from

Definition 2.26 is *not* affine. However, Theorem A.1 shows that the inequality

$$S_f\big(\lambda\omega + (1-\lambda)\xi\big) \geqslant \lambda S_f(\omega) + (1-\lambda)S_f(\xi)$$

holds. Nevertheless, and more importantly for our characterization theorem, equality *does* hold when $\omega \perp \xi$ and $(\omega \circ f) \perp (\xi \circ f)$. The proof of this will be given in Proposition 3.27.

Definition 3.18 [Fibrewise convex structures]

A ***fibrewise convex structure*** on a fibration $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ is a cloven fibration where each fibre is a convex category and each reindexing functor (as described in Lemma 3.9) $\mathcal{E}_x \xleftarrow{f^*} \mathcal{E}_y$, associated to a morphism $x \xrightarrow{f} y$ in $\mathcal{X}$, is an affine functor. A cloven fibration equipped with a fibrewise convex structure will be called a ***fibrewise convex fibration***.

Example 3.19 [Examples of fibrewise convex structures]

(a) The discrete fibration **NCFinProb** $\to$ **fdC*-Alg** has $\mathcal{S}(\mathcal{A})$ as the fibre over each C^*-algebra $\mathcal{A}$. The set of states $\mathcal{S}(\mathcal{A})$ on a C^*-algebra $\mathcal{A}$ has a natural convex structure. Furthermore, each $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ has the pullback $\mathcal{S}(\mathcal{B}) \xleftarrow{\mathcal{S}(f)} \mathcal{S}(\mathcal{A})$ as its reindexing functor. This functor is affine, as discussed in Example 3.16.

(b) By a similar argument, $\mathbf{FinProb}^{\mathrm{op}} \to \mathbf{FinSet}^{\mathrm{op}}$ has a natural fibrewise convex structure coming from the convex combination of probability measures and the fact that the pushforward of measures is linear. The fibre over a finite set X is isomorphic to the standard simplex $\Delta^{|X|-1}$.

(c) The fibration $\mathbb{B}\mathbb{R} \to \bullet$ has a convex structure on the only fiber $\mathbb{B}\mathbb{R}$ over the single object in the base, as described in Example 3.13.

Definition 3.20 [Morphisms of fibrations]

Let $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ and $\mathcal{F} \xrightarrow{\rho} \mathcal{Y}$ be fibrations. A ***fibred functor***[a] from π to ρ is a pair of functors $\mathcal{E} \xrightarrow{\Phi} \mathcal{F}$ and $\mathcal{X} \xrightarrow{\phi} \mathcal{Y}$ such that

$$\begin{tikzcd} \mathcal{E} \arrow[r, "\Phi"] \arrow[d, "\pi"'] & \mathcal{F} \arrow[d, "\rho"] \\ \mathcal{X} \arrow[r, "\phi"'] & \mathcal{Y} \end{tikzcd}$$

commutes and such that $\Phi(\beta)$ is cartesian for every cartesian β.

[a] Our terminology differs from that of [32], who use "functor" when the base category is fixed ($\phi = \mathrm{id}$) and "1-cell" for when the base category changes.

Remark 3.21 [Fibrewise convex structures as internalized convex objects]

One can equivalently define a fibrewise convex structure as an internal convex object in the category of fibrations over a fixed based, analogous to the fibrewise monoidal structure in [32, Section 3.1]. Briefly, a convex object is an internalization of Definition 3.11 in an arbitrary cartesian monoidal category (as opposed to the specific one of categories and functors). Such an object $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ in the category of fibrations over a fixed based $\mathcal{X}$ provides the data of a family of fibred functors $F_\lambda : \mathcal{E} \times_\pi \mathcal{E} \to \mathcal{E}$ with a fixed based, i.e.

$$\begin{tikzcd} \mathcal{E} \times_\pi \mathcal{E} \arrow[rr, "F_\lambda"] \arrow[dr, "\pi"'] & & \mathcal{E} \arrow[dl, "\pi"] \\ & \mathcal{X} & \end{tikzcd}$$

commutes for all $\lambda \in [0,1]$. Here, $\mathcal{E} \times_\pi \mathcal{E}$ is the category whose objects are pairs (u,v) with $\pi(u) = \pi(v)$ and whose morphisms are pairs $(t \xrightarrow{\alpha} u, v \xrightarrow{\beta} w)$ with $\pi(\alpha) = \pi(\beta)$ (in particular, $\pi(t) = \pi(v)$ and $\pi(u) = \pi(w)$). Composition is induced from composition in $\mathcal{E} \times \mathcal{E}$.[a] Consequently, such a family of functors $\mathcal{E} \times_\pi \mathcal{E} \xrightarrow{F_\lambda} \mathcal{E}$, indexed by $\lambda \in [0,1]$, defines a convex category structure for every fibre $\mathcal{E}_x$ with x an object of $\mathcal{X}$. In addition, the F_λ's also provide an assignment on morphisms since a pair $(t \xrightarrow{\alpha} u, v \xrightarrow{\beta} w)$ over a morphism $x \xrightarrow{f} y$ gets sent to a morphism

$$\lambda t + (1-\lambda)v \xrightarrow{\lambda\alpha + (1-\lambda)\beta \equiv F_\lambda(\alpha,\beta)} \lambda u + (1-\lambda)w$$

over the morphism $x \xrightarrow{f} y$. This assignment guarantees that the associated reindexing functor $\mathcal{E}_x \xleftarrow{f^*} \mathcal{E}_y$ from Lemma 3.9 can be chosen to be affine as in Definition 3.18. This is because if one chooses cartesian lifts $f^*(u) \xrightarrow{f_u} u$ and $f^*(v) \xrightarrow{f_v} v$ of u and v, respectively, over $x \xrightarrow{f} y$, then the functor F_λ allows one to set

$$\lambda f^*(u) + (1-\lambda) f^*(v) \xrightarrow{\lambda f_u + (1-\lambda) f_v} \lambda u + (1-\lambda)v$$

to be the lift of $\lambda u + (1-\lambda)v$ over f (it is a cartesian lift because F_λ is assumed to take cartesian lifts to cartesian lifts). This says that the reindexing functor sends $\lambda u + (1-\lambda)v$ to $\lambda f^*(u) + (1-\lambda) f^*(v)$, which exactly says the reindexing functor is affine. For example, in the fibrewise convex fibration **NCFinProb** $\to$ **fdC*-Alg**, if $(\mathcal{B},\eta) \xrightarrow{g} (\mathcal{A},\omega)$ and $(\mathcal{B},\zeta) \xrightarrow{h} (\mathcal{A},\xi)$ are two morphisms over $\mathcal{B} \xrightarrow{f} \mathcal{A}$, then $g = h = f$ and their convex combination, $\lambda g + (1-\lambda)h$, is just f. In the fibrewise convex fibration $\mathbb{BR} \to \bullet$, the convex combination of objects in the fibre is trivial, while the convex combination of morphisms (elements in $\mathbb{R}$) is the usual convex combination of real numbers.

[a] For those familiar with the terminology, this is just a (standard) pullback.

Definition 3.22 [Convergence in NCFinProb]

A sequence $\mathbb{N} \ni n \mapsto \Big((\mathcal{B}_n, \xi_n) \xrightarrow{f_n} (\mathcal{A}_n, \omega_n) \Big)$ ***converges to*** $(\mathcal{A}, \xi) \xrightarrow{f} (\mathcal{B}, \omega)$ in the category **NCFinProb** iff there exists an $N \in \mathbb{N}$ such that

$$\mathcal{A}_n = \mathcal{A}, \quad \mathcal{B}_n = \mathcal{B}, \quad f_n = f \quad \forall\, n \geqslant N, \quad \lim_{n\to\infty} \omega_n = \omega, \quad \text{and} \quad \lim_{n\to\infty} \xi_n = \xi,$$

where the last two limits are with respect to the standard topologies on the state spaces $\mathcal{S}(\mathcal{A})$ and $\mathcal{S}(\mathcal{B})$, respectively.[a]

[a] If $\mathcal{A} = \bigoplus_{x\in X} \mathcal{M}_{m_x}$, with X a finite set and $m_x \in \mathbb{N}$, the topology on $\mathcal{S}(\mathcal{A})$ is the subspace topology inherited from $\mathcal{A}$ to the subset of $|X|$-tuples of positive matrices $\{\tau_x\}$ such that $\sum_{x\in X} \operatorname{tr}(\tau_x) = 1$, since $\mathcal{S}(\mathcal{A})$ can be identified with this set by the Hilbert–Schmidt inner product (cf. Example 2.2 and Example 2.12).

Remark 3.23 [Justifying the definition of convergence of sequences in NCFinProb]

The definition of convergence of a sequence of morphisms in **NCFinProb** is motivated by the one in **FinProb** from [4, page 4]. However, some justification seems to be necessary, particularly regarding why the morphisms are assumed to stabilize, i.e. are equal after some $N \in \mathbb{N}$. Recall, in the case of **FinProb**, a sequence $(X_n, p_n) \xrightarrow{f_n} (Y_n, q_n)$ ***converges to*** $(X, p) \xrightarrow{f} (Y, q)$ iff the sets X_n, Y_n and the underlying set functions f_n stabilize after a finite natural number in the sequence [4]. The sets must stabilize because their associated simplices of probability distributions are distinct and the cardinality of the set dictates which simplex one is using for the space of probability distributions. The functions must stabilize because the set of functions between two finite sets is also a finite set, which has the discrete topology. However, the probability distributions p_n on X and q_n on Y may continue to vary as long as they converge to p and q in the topology associated with the simplices $\Delta^{|X|-1}$ and $\Delta^{|Y|-1}$.

In the case of C^*-algebras, the collection $\operatorname{hom}(\mathcal{B}, \mathcal{A})$ of (unital) $*$-homomorphisms from $\mathcal{B}$ to $\mathcal{A}$ is *not* just a discrete set. In general, it has a non-trivial topology. Nevertheless, without loss of generality, one can assume the f_n eventually stabilize. To make sense of this topology and the condition we have demanded, since $\mathcal{A}$ and $\mathcal{B}$ are finite-dimensional C^*-algebras, it suffices to assume $\mathcal{A} = \bigoplus_{x\in X} \mathcal{M}_{m_x}$ and $\mathcal{B} = \bigoplus_{y\in Y} \mathcal{M}_{n_y}$ for some finite sets X and Y and $m_x, n_y \in \mathbb{N}$. In this case, an arbitrary $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ is described by its multiplicities and by a unitary (or direct sum of unitaries), as in Lemma 2.13. The multiplicities entail the constraint $m_x = \sum_{y\in Y} c_{xy} n_y$, but there could be several such multiplicities satisfying these constraints. Indeed, if[a]

$$s_x := \left| \left\{ Y \ni y \mapsto c_{xy} \in \mathbb{Z}_{\geqslant 0} \;:\; m_x = \sum_{y\in Y} c_{xy} n_y \right\} \right|$$

denotes the number of such solutions, then the number of connected components in $\operatorname{hom}(\mathcal{B}, \mathcal{A})$ is $s := \prod_{x\in X} s_x$ (for example, if $\mathcal{B} = \mathcal{M}_n$ is a matrix algebra, there is only one such component). Furthermore, since the collection of unitary matrices has a non-trivial topology, $\mathrm{hom}(\mathcal{B},\mathcal{A})$ is the disjoint union of s-many non-trivial topological spaces. Due to the connected components, one knows that a sequence of $*$-homomorphisms converging to another one must therefore necessarily have multiplicities that stabilize. Within such a component, since $\omega \circ f = \omega \circ (\mathrm{Ad}_U \circ \mathrm{Ad}_{U^\dagger}) \circ f = (\omega \circ \mathrm{Ad}_U) \circ (\mathrm{Ad}_{U^\dagger} \circ f)$ for every unitary U, one can always choose f to be of the form

$$\bigoplus_{y\in Y} \mathcal{M}_{n_y} \ni \bigoplus_{y\in Y} B_y \mapsto \bigoplus_{x\in X}\left(\boxplus_{y\in Y} \mathrm{diag}(\overbrace{B_y,\cdots,B_y}^{c_{yx}\text{ times}})\right)$$

by conjugating with some appropriate unitary U (cf. Lemma 2.13). This unitary can then be transferred to the state. Therefore, it suffices to assume the algebras and $*$-homomorphisms stabilize in a convergent sequence, but not necessarily the states.

[a] The vertical bars denote the cardinality.

Definition 3.24 [Continuous fibred functors]

A ***continuous fibred functor*** from **NCFinProb** $\to$ **fdC*-Alg** to $\mathbb{BR} \to \bullet$ is a fibred functor

$$\begin{tikzcd} \mathbf{NCFinProb} \arrow[r, "H"] \arrow[d] & \mathbb{BR} \arrow[d] \\ \mathbf{fdC^*\text{-}Alg} \arrow[r] & \underline{\mathbf{1}} \end{tikzcd}$$

such that to every sequence $\mathbb{N} \ni n \mapsto \left((\mathcal{B}_n,\xi_n) \xrightarrow{f_n} (\mathcal{A}_n,\omega_n)\right)$ converging to $(\mathcal{A},\xi) \xrightarrow{f} (\mathcal{B},\omega)$ in the category **NCFinProb**,

$$\lim_{n\to\infty} H\left((\mathcal{B}_n,\xi_n) \xrightarrow{f_n} (\mathcal{A}_n,\omega_n)\right) = H\left((\mathcal{B},\xi) \xrightarrow{f} (\mathcal{A},\omega)\right),$$

where the convergence is for a sequence of real numbers.

Notation 3.25 [The function $H_f : \mathcal{S}(A) \to \mathbb{R}$]

For fibred functors H : **NCFinProb** $\to \mathbb{BR}$, we will occasionally write

$$H_f(\omega) := H\left((\mathcal{B},\xi) \xrightarrow{f} (\mathcal{A},\omega)\right)$$

for the image of H along a morphism f in **NCFinProb** to indicate the data involved more clearly and to condense the notation. This is also justified because for a fixed $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$, the functor H defines a function $H_f : \mathcal{S}(\mathcal{A}) \to \mathbb{R}$.

The next definition is the appropriate quantum generalization of the affinity condition used by BFL in their characterization of Shannon entropy [4]. Why this is so will be explained towards the end of this section as well as Proposition 4.17 and Remark 4.24.

Definition 3.26 [Orthogonally affine fibred functor]

A fibred functor H from **NCFinProb** $\to$ **fdC*-Alg** to $\mathbb{BR} \to \bullet$ is ***orthogonally affine*** iff to each pair of C^*-algebras $\mathcal{B}$ and $\mathcal{A}$, each pair of mutually orthogonal states $\omega, \xi \in \mathcal{S}(\mathcal{A})$, and each $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ such that $(\omega \circ f) \perp (\xi \circ f)$,

$$H_f\big(\lambda\omega + (1-\lambda)\xi\big) = \lambda H_f(\omega) + (1-\lambda)H_f(\xi) \qquad \forall\, \lambda \in [0,1].$$

Proposition 3.27 [Entropy difference is continuous and orthogonally affine]

The entropy change functor from Definition 2.26 is a continuous and orthogonally affine fibred functor. In fact, if for any C^*-algebra $\mathcal{A}$ and any pair ω, ξ of mutually orthogonal states on $\mathcal{A}$, a $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ preserves the orthogonality $\omega \perp \xi$ if and only if

$$S_f\big(\lambda\omega + (1-\lambda)\xi\big) = \lambda S_f(\omega) + (1-\lambda)S_f(\xi) \qquad \forall\, \lambda \in [0,1].$$

Before proving this, we introduce a shorthand for the deviation from S_f being affine on the states ω and ξ.

Definition 3.28 [The Holevo information change along a morphism]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be a $*$-homomorphism of C^*-algebras, let ω, ξ be two states on $\mathcal{A}$, and let $\lambda \in [0,1]$. The ***Holevo information change along*** f associated to ω, ξ and λ, is the number

$$\chi_f(\lambda; \omega, \xi) := S_f\big(\lambda\omega + (1-\lambda)\xi\big) - \lambda S_f(\omega) - (1-\lambda)S_f(\xi).$$

Proposition 3.27 says, in particular, that this deviation vanishes when $\omega \perp \xi$ and $(\omega \circ f) \perp (\xi \circ f)$.

Remark 3.29 [Why have we called the deviation the Holevo information change?]

Although we initially defined $\chi_f(\lambda; \omega, \xi)$ for an arbitrary $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$, by taking $\mathcal{B} = \mathbb{C}$ and $f :=!_{\mathcal{A}}$ to be the unique $*$-homomorphism sending $\lambda \in \mathbb{C}$ to $\lambda 1_{\mathcal{A}}$, one obtains the quantity

$$\chi_{\mathcal{A}}(\lambda; \omega, \xi) := S\big(\lambda\omega + (1-\lambda)\xi\big) - \lambda S(\omega) - (1-\lambda)S(\xi).$$

As a result, for a general $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$,

$$\chi_f(\lambda; \omega, \xi) = \chi_{\mathcal{A}}(\lambda; \omega, \xi) - \chi_{\mathcal{B}}(\lambda; \omega \circ f, \xi \circ f).$$

The expression $\chi_{\mathcal{A}}(\lambda; \omega, \xi)$ is the ***Holevo information*** associated to two states ω and ξ that are mixed with the probability distribution $(\lambda, 1-\lambda)$ on the fixed algebra $\mathcal{A}$ [34, Section 12.1.1]. Because of this, we have decided to call $\chi_f(\lambda;\omega,\xi)$ the Holevo information change along f.

Proof of Proposition 3.27. Continuity of the entropy change follows from continuity of the von Neumann entropy (due to the finite-dimensionality assumption) [34, Section 11.3], [10]. To prove the statement regarding orthogonal affinity, suppose $\omega \perp \xi$. Let $\omega' := \omega \circ f$ and $\xi' := \xi \circ f$. If f preserves the mutual orthogonality, then $\omega' \perp \xi'$ and

$$
\begin{aligned}
\chi_f(\lambda;\omega,\xi) &= S\big(\lambda\omega + (1-\lambda)\xi\big) - S\big(\lambda\omega' + (1-\lambda)\xi'\big) - \lambda S_f(\omega) - (1-\lambda)S_f(\xi) \\
&\overset{\text{Lem 2.24}}{=\!=\!=} S(\lambda,1-\lambda) + \lambda S(\omega) + (1-\lambda)S(\xi) \\
&\qquad - S(\lambda,1-\lambda) - \lambda S(\omega') - (1-\lambda)S(\xi') \\
&\qquad\qquad - \lambda S_f(\omega) - (1-\lambda)S_f(\xi) \\
&= 0,
\end{aligned}
\tag{3.30}
$$

where $S(\lambda,1-\lambda)$ is the Shannon entropy of the probability measure $(\lambda,1-\lambda)$ on a two element set. Conversely, suppose $\chi_f(\lambda;\omega,\xi) = 0$. Since $\omega \perp \xi$, a similar calculation gives

$$
0 = \chi_f(\lambda;\omega,\xi) \overset{\text{Lem 2.24}}{=\!=\!=} S(\lambda,1-\lambda) + \lambda S(\omega') + (1-\lambda)S(\xi') - S\big(\lambda\omega' + (1-\lambda)\xi'\big), \tag{3.31}
$$

which gives $\omega' \perp \xi'$ by the "only if" part of Lemma 2.24. ∎

Remark 3.32 [Orthogonal affinity for an arbitrary mixture of orthogonal states]

A consequence of Proposition 3.27 is that if $\{\omega_x\}_{x\in X}$ is a collection of (pairwise) mutually orthogonal states on a C^*-algebra $\mathcal{A}$, indexed by a finite set X, and if $\mathcal{B} \xrightarrow{f} \mathcal{A}$ is a $*$-homomorphism that preserves the orthogonality of all these states, then

$$
S_f\left(\sum_{x\in X} p_x\omega_x\right) = \sum_{x\in X} p_x S_f(\omega_x)
$$

for all probability measures p on X. Notice how this contrasts with the equality condition in Lemma 2.24 for the entropy on a *fixed* algebra. This is because we have considered the entropy *difference* and the $S_{Sh}(p)$ terms cancel upon taking the difference.

In the last part of this section, we recall the convex combinations and affine functors introduced by BFL [4]. By the next section, we will have proved enough facts to provide a more precise relationship between BFL's definition and ours. The category **FinProb** naturally comes equipped with a family of functors F_λ : **FinProb** $\times$ **FinProb** $\to$ **FinProb** for each $\lambda \in [0,1]$ defined as follows.

Definition 3.33 [An external convex structure on FinProb]

For every $\lambda \in [0,1]$, define the convex sum F_λ : **FinProb** $\times$ **FinProb** $\to$ **FinProb** on objects by

$$\lambda(X,p)\oplus(1-\lambda)(Y,q) := \big(X \amalg Y, \lambda p \oplus (1-\lambda)q\big),$$
$$\text{where} \quad \big(\lambda p \oplus (1-\lambda)q\big)_z := \begin{cases} \lambda p_z & \text{if } z \in X \\ (1-\lambda)q_z & \text{if } z \in Y, \end{cases}$$

and where $X \amalg Y$ denotes the disjoint union. The convex sum of morphisms $(X,p) \xrightarrow{\phi} (X',p')$ and $(Y,q) \xrightarrow{\psi} (Y',q')$ is defined to be

$$\big(\lambda\phi \oplus (1-\lambda)\psi\big)(z) := \phi \amalg \psi := \begin{cases} \phi(z) & \text{if } z \in X \\ \psi(z) & \text{if } z \in Y. \end{cases}$$

The collection of functors $\{F_\lambda\}_{\lambda\in[0,1]}$ is called the ***external convex structure*** on **FinProb**.

The motivation for calling this an *external* convex structure comes from the distinction between internal and external monoidal fibrations [32, Section 3.1], as will be explained shortly.

Remark 3.34 [The external convex structure on FinProb is not a convex category]

FinProb with this family of functors is *not* a convex category in the sense of Definition 3.11. It is, however, a *weak* convex category (called a convex category in [36]) satisfying a form of parametric associativity. There are many variants of such weak convex categories depending on whether one demands $\lambda x \oplus (1-\lambda)x \cong x$ or $0x \oplus 1y \cong y$ or even weaker versions of these types of equations. Indeed, for objects (X,p) and (Y,q) in **FinProb**, the convex sum $0(X,p) \oplus 1(Y,q)$ has $X \amalg Y$ as its underlying set with q supported on Y as its probability measure. Therefore, it is impossible for this to be isomorphic to (Y,q). Nevertheless, one can fix this in at least two ways. One option is to use a.e. equivalence classes of measure-preserving maps to guarantee the existence of isomorphisms such as $0(X,p) \oplus 1(Y,q) \cong (Y,q)$. Another option is to not use a.e. equivalence classes of morphisms but instead *define*

$$0(X,p) \oplus 1(Y,q) := \big(\varnothing \amalg Y, 0p + 1q\big) \cong (Y,q).$$

Notice, however, that with either of these two options, $\lambda(X,p) \oplus (1-\lambda)(X,p)$ is *not* isomorphic to (X,p) for any $\lambda \in (0,1)$. Although we have a natural function $\lambda(X,p) \oplus (1-\lambda)(X,p) \to (X,p)$ given by sending x in either factor to x (so it is a two-to-one map), there is no natural map back.[a]

Nevertheless, with this weak convex structure, $\mathbf{FinProb}^{\mathrm{op}} \to \mathbf{FinSet}^{\mathrm{op}}$ is a weak *external* convex fibration[b] with the projection of $\lambda(X,p) \oplus (1-\lambda)(Y,q)$ in the total category to $X \amalg Y$ in the base category, and similarly for morphisms.

[a] There is, however, a natural stochastic section, but adding such morphisms would dramatically change the category. Such a stochastic section is an example of a disintegration [42], which is a crucial ingredient in the categorical classification of classical *relative* entropy [3, 16]. We will also discuss disintegrations in the next section, as they are used in our main characterization theorem.

[b] The $^{\text{op}}$ is needed for this to be a fibration; otherwise it is called an opfibration (cf. Example 3.4).

A completely analogous definition can be made for the fibration **NCFinProb** $\to$ **fdC*-Alg** using the (external) direct sum of C^*-algebras.

Definition 3.35 [An external convex structure on NCFinProb]

For every $\lambda \in [0,1]$, define the convex sum F_λ : **NCFinProb** $\times$ **NCFinProb** $\to$ **NCFinProb** on objects by $\lambda(\mathcal{A},\omega) \oplus (1-\lambda)(\mathcal{B},\xi) := \big(\mathcal{A}\oplus\mathcal{B}, \lambda\omega \oplus (1-\lambda)\xi\big)$, where $\big(\lambda\omega\oplus(1-\lambda)\xi\big)(A \oplus B) := \lambda\omega(A) + (1-\lambda)\xi(B)$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$. The convex sum of morphisms is the direct sum.

This convex structure on **NCFinProb** restricts to the one on **FinProb** on the subcategory of commutative C^*-algebras since $\mathbb{C}^{X \amalg Y} \cong \mathbb{C}^X \oplus \mathbb{C}^Y$.

Definition 3.36 [Externally affine functor]

A functor H : **NCFinProb** $\to \mathbb{BR}$ is ***externally affine*** iff

$$H\big(\lambda f \oplus (1-\lambda) g\big) = \lambda H(f) + (1-\lambda) H(g)$$

for all morphisms f, g in **NCFinProb** and all $\lambda \in [0,1]$.

Example 3.37 [Examples of externally affine functors]

(a) The difference of Shannon entropies studied by BFL [4] is a continuous externally affine functor **FinProb** $\to \mathbb{BR}$. In fact, it is characterized as the unique one whose image always lands in $\mathbb{BR}_{\geqslant 0}$ (see Theorem 3.38 below for the precise statement).

(b) An example of a continuous externally affine functor S : **NCFinProb** $\to \mathbb{BR}$ is the difference of Segal entropies from Definition 2.26.

(c) Another example is the following. If $\omega = \sum_{x\in X} p_x \omega_x$ and $\xi = \sum_{y \in Y} q_y \xi_y$ are states on $\mathcal{A} = \bigoplus_{x\in X} \mathcal{M}_{m_x}$ and $\mathcal{B} = \bigoplus_{y\in Y}\mathcal{M}_{n_y}$, respectively, and $f : \mathcal{B} \to \mathcal{A}$ is a state-preserving $*$-homomorphism, then $K_f(\omega) := S(p) - S(q)$, the difference of the Shannon entropies associated to the probability distributions,[a] defines a continuous externally affine functor K : **NCFinProb** $\to \mathbb{BR}$.

Notice that both K and S agree on the subcategory of commutative algebras, and they in fact agree with the Shannon entropy difference on the equivalent category **FinProb**$^{\text{op}}$, yet they are *not* proportional.[b]

[a] Note that this is well-defined, i.e. independent of the decomposition of ω and ξ into $\sum_{x\in X} p_x\omega_x$ and $\sum_{y\in Y} q_y \xi_y$, respectively.

[b] This answers a question of John Baez in the negative [2] (see specifically the original post as well as the post on June 7, 2011 at 8:12 AM). In more detail, the existence of these two distinct continuous (externally) affine functors (that are not proportional to each other) illustrates that continuous affine functors **NCFinProb** $\to \mathbb{BR}$ are not characterized by their values on **FinProb**$^{\text{op}}$ (when viewed as a subcategory of **NCFinProb**). In particular, this condition does not characterize the von Neumann entropy difference. The main reason for this, as we will later see, is because the external convex structure misses the internal convex

structure of quantum states.

For reference, we recall BFL's characterization theorem [4].

Theorem 3.38 [BFL's functorial characterization of the Shannon entropy]

If $H : \mathbf{FinProb} \to \mathbb{B}\mathbb{R}_{\geqslant 0}$ is a continuous externally affine functor, then there exists a constant $c \geqslant 0$ such that

$$H_{\phi}(p) = c\Big(S(p) - S(q)\Big)$$

for every probability-preserving function $(X, p) \xrightarrow{\phi} (Y, q)$.

We will abstractly (without reference to the entropy formulas from Definition 2.26) relate the two notions of affinity later in Proposition 4.17 after developing some general results.

4 Classifying entropy

This section contains our main result, Theorem 4.31, which is a functorial classification of the entropy difference in the non-commutative setting. Continuity and orthogonal affinity alone are not quite enough to characterize the von Neumann entropy difference, though they come quite close. By Lemma 2.27, we cannot assume that $S_f(\omega) \geqslant 0$ for all $*$-homomorphisms f and states ω on the codomain of f, since this inequality fails for non-commutative C^*-algebras. We propose a close replacement, namely $S_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on pure states, for all C^*-algebras $\mathcal{A}$. While this may sound quite different, we will explain that this assumption is a consequence of, but does not imply, BFL's assumption of $S_f(\omega) \geqslant 0$ on *commutative* C^*-algebras. Nevertheless, in Proposition 4.8, we prove the crucial observation that the non-negativity of entropy difference for *commutative* C^*-algebras is a *consequence* of the fact that state-preserving $*$-homomorphisms between commutative C^*-algebras always have *disintegrations*, also known as *optimal hypotheses* [3], which we briefly review. More generally, we show that the existence of disintegrations (with non-commutative probability spaces included) implies the non-negativity of entropy difference.

Notation 4.1 [$!_{\mathcal{A}}$ and $H_{\mathcal{A}}$]

In what follows, if $\mathcal{A}$ is a C^*-algebra, then $\mathbb{C} \xrightarrow{!_{\mathcal{A}}} \mathcal{A}$ will always refer to the unique (unital) $*$-homomorphism. If $H : \mathbf{NCFinProb} \to \mathbb{BR}$ is a functor, we implement the notation $H_{\mathcal{A}} := H_{!_{\mathcal{A}}}$ for the rest of this section.

Lemma 4.2 [H is a coboundary]

Given any $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ and a state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$, any functor $H : \mathbf{NCFinProb} \to \mathbb{BR}$ satisfies

$$H_f(\omega) = H_{\mathcal{A}}(\omega) - H_{\mathcal{B}}(\omega \circ f).$$

Proof. This follows from functoriality of H since the diagram

$$\begin{tikzcd} & \mathcal{B} \arrow[dr, "f"] & \\ \mathbb{C} \arrow[ur, "!_{\mathcal{B}}"] \arrow[rr, "!_{\mathcal{A}}"'] & & \mathcal{A} \end{tikzcd} \tag{4.3}$$

commutes ($\mathbb{C}$ is an initial object in **fdC*-Alg**). ∎

Lemma 4.4 [Non-negativity of H_f implies vanishing of $H_{\mathcal{A}}$ on pure states]

Suppose $H : \mathbf{FinProb}^{\mathrm{op}} \to \mathbb{BR}$ is a functor satisfying $H_f(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$ and $*$-homomorphisms $\mathcal{B} \xrightarrow{f} \mathcal{A}$ between commutative C^*-algebras.[a]

1. If f has a left or right inverse, then $H_f(\omega) = 0$ for all $\omega \in \mathcal{S}(\mathcal{A})$.
2. $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states.

[a] $\mathbf{FinProb}^{\mathrm{op}}$ is viewed as the full subcategory of **NCFinProb** consisting of commutative probability spaces.

Proof.

1. Suppose f has a right inverse $\mathcal{A} \xrightarrow{g} \mathcal{B}$. Then functoriality of H implies $0 = H_{\mathrm{id}_{\mathcal{A}}}(\omega) = H_g(\omega \circ f) + H_f(\omega)$ by Lemma 4.2. Since each term is non-negative by assumption, $H_f(\omega) \geqslant 0$. A similar calculation proves the same inequality if f has a left inverse.
2. First, $H_{\mathcal{A}}(\omega) \geqslant 0$ is immediate because $H_{!_{\mathcal{A}}}(\omega) \geqslant 0$ by assumption. By invariance of under $*$-isomorphisms (due to the first item), it suffices to consider a commutative C^*-algebra of the form $\mathcal{A} = \mathbb{C}^X$. Given any pure state ξ on $\mathbb{C}^X$, such a pure state is necessarily supported on some $x \in X$. Let $\mathbb{C}^X \xrightarrow{\pi_x} \mathbb{C}$ be the projection onto that component. Then π_x pulls the unique state 1 on $\mathbb{C}$ back to ξ on $\mathbb{C}^X$ and the composite $\mathbb{C} \xrightarrow{!_{\mathbb{C}^X}} \mathbb{C}^X \xrightarrow{\pi_x} \mathbb{C}$ equals $\mathrm{id}_{\mathbb{C}}$. Thus, $H_{\mathbb{C}^X}(\xi) = 0$ by the first item. ■

There is a sort of converse to Lemma 4.4, which illustrates that our axioms for entropy change imply those of BFL. We first prove a lemma about invariance under $*$-isomorphisms given *our* axioms. This is quite a different proof from the one we just gave in Lemma 4.4, and it uses the internal convex structure in a crucial way.

Lemma 4.5 [H is invariant under $*$-isomorphisms]

Suppose $H : \mathbf{NCFinProb} \to \mathbb{B}\mathbb{R}$ is an orthogonally affine fibred functor for which $H_{\mathcal{A}}(\xi) = 0$ for all pure states ξ on $\mathcal{A}$ and all C^*-algebras $\mathcal{A}$. If $\mathcal{B} \xrightarrow{f} \mathcal{A}$ is a $*$-isomorphism, then $H_f(\omega) = 0$ for all $w \in \mathcal{S}(\mathcal{A})$.

Proof. Let ω be a state on $\mathcal{A}$. Then there exists a convex decomposition $\omega = \sum_{x \in X} p_x \omega_x$ of ω in terms of mutually orthogonal pure states ω_x and a nowhere-vanishing probability measure p on some finite set X. Thus,

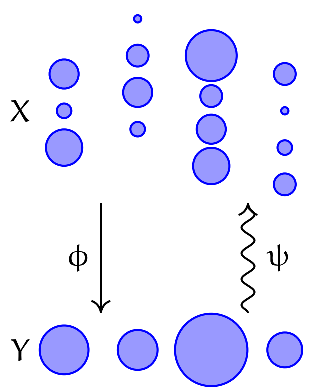
$$H_f(\omega) \overset{\text{Lem 2.16}}{\underset{\text{Defn 3.26}}{=\!=\!=}} \sum_{x \in X} p_x H_f(\omega_x) \overset{\text{Lem 4.2}}{=\!=\!=} \sum_{x \in X} p_x \Big(H_{\mathcal{A}}(\omega_x) - H_{\mathcal{B}}(\omega_x \circ f) \Big) = 0 \tag{4.6}$$

since $\omega_x \circ f$ is pure by Lemma 2.16. ■

Definition 4.7 [Disintegrations on finite probability spaces]

Let (X, p) and (Y, q) be probability spaces and let $\phi : X \to Y$ be a probability-preserving function, i.e. $q = \phi \circ p$. A ***disintegration*** of (ϕ, p, q) (or simply of ϕ if p and q are clear from context) is a stochastic map $Y \overset{\psi}{\rightsquigarrow} X$ such that

$$\begin{tikzcd} & \{\bullet\} \arrow[dl, squiggly, "p"'] \arrow[dr, squiggly, "q"] \arrow[d, phantom, "="] & \\ X & & Y \arrow[ll, squiggly, "\psi"] \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} & X \arrow[dl, "\phi"'] \arrow[d, phantom, "="] & \\ Y & & Y \arrow[ul, squiggly, "\psi"'] \arrow[ll, "\mathrm{id}_Y"] \end{tikzcd},$$



the latter diagram signifying commutativity q-a.e.[a]

[a]The graphic on the right depicts probability measures as collections of water droplets with total volume 1 (in some appropriate units). The map ϕ depicts a function preserving probability (volume), which in turn combines water droplets [17], while the disintegration ψ splits the water droplets back into their original sizes. We personally find this picture helpful in arguments such as the proof of Proposition 4.8 below. The reason ψ is written using squiggly arrow notation is because it corresponds to a positive map on the associated algebras [36].

The main fact we will use about disintegrations on finite probability spaces is that they always exist [42, Section 2].

Proposition 4.8 [Positivity of entropy difference on commutative C^*-algebras]

Suppose $H : \mathbf{NCFinProb} \to \mathbb{B}\mathbb{R}$ is an orthogonally affine fibred functor for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$.[a] Then for *commutative* C^*-algebras $\mathcal{A}$ and $\mathcal{B}$,

$$H_f(\omega) \geqslant 0$$

for all states $\omega \in \mathcal{A}$ and all $*$-homomorphisms $\mathcal{B} \xrightarrow{f} \mathcal{A}$.

[a]Technically, this assumption is only needed on the subcategory of commutative C^*-algebras, for this specific result. Nevertheless, we include it here because we will use this assumption later in our characterization theorem.

Proof. By invariance of H for $*$-isomorphisms (Lemma 4.5), it suffices to suppose $\mathcal{B} = \mathbb{C}^Y$ and $\mathcal{A} = \mathbb{C}^X$ for finite sets X and Y. In this case, let ω be represented by a probability measure p on X, let $X \xrightarrow{\phi} Y$ be the function associated to $\mathcal{B} \xrightarrow{f} \mathcal{A}$, and let $q := \phi \circ p$ be the pushforward measure corresponding to $\omega \circ f =: \xi$ (cf. Example 2.10). Every such probability measure is decomposed as

$$q = \sum_{y \in Y} q_y \delta_y, \tag{4.9}$$

where δ_y is the Dirac delta measure at y (in this case, it's just the Kronecker delta) defined by $\delta_y(y') \equiv \delta_{yy'}$, which is 1 if $y' = y$ and 0 otherwise. This expresses q as a convex sum of mutually orthogonal measures since $\delta_y \perp \delta_{y'}$ for all $y \neq y'$. Set

$$N_q := \{y \in Y \ : \ q_y = 0\} \tag{4.10}$$

and let $Y \overset{\psi}{\rightsquigarrow} X$ be a disintegration of (ϕ, p, q). Then p also decomposes as

$$p = \sum_{y \in Y} q_y \psi_y \equiv \sum_{y \in Y \setminus N_q} q_y \psi_y, \tag{4.11}$$

where the set of probability measures $\{\psi_y\}_{y \in Y \setminus N_q}$ are mutually orthogonal, i.e.

$$\psi_y \perp \psi_{y'} \qquad \forall\, y \neq y' \in Y \setminus N_q. \tag{4.12}$$

Furthermore, ϕ preserves the mutual orthogonality of these measures

$$(\phi \circ \psi_y) \perp (\phi \circ \psi_{y'}) \qquad \forall\, y \neq y' \in Y \setminus N_q, \tag{4.13}$$

since $\phi \circ \psi_y = \delta_y$ for all $y \in Y \setminus N_q$. All of this can be expressed in terms of the C^*-algebras by setting ω_y to be the state corresponding to ψ_y. Doing so gives

$$\begin{aligned}
H_f(\omega) &\overset{(4.11)}{=\!=} H_f\left(\sum_{y \in Y\setminus N_q} q_y \omega_y\right) \\
&\overset{(4.13)}{=\!=} \sum_{y \in Y\setminus N_q} q_y H_f(\omega_y) \\
&\overset{\text{Lem } 4.2}{=\!=\!=} \sum_{y \in Y\setminus N_q} q_y \Big(H_{\mathcal{A}}(\omega_y) - H_{\mathcal{B}}(\underbrace{\omega_y \circ f}_{\delta_y})\Big) \\
&= \sum_{y \in Y\setminus N_q} q_y H_{\mathcal{A}}(\omega_y) \\
&\geqslant 0,
\end{aligned} \tag{4.14}$$

where the second last equality holds by the assumption that $H_{\mathcal{B}}$ vanishes on pure states and the last inequality holds by the assumption that $H_{\mathcal{A}}$ is always non-negative. ■

Proposition 4.8 shows that our axioms *imply* the (seemingly strong) axiom of non-negativity for entropy difference used by BFL in their functorial characterization of Shannon entropy (Theorem 3.38). Combining this fact with Lemma 4.4 suggests that it is reasonable to replace the BFL axiom of non-negativity for entropy difference by non-negativity of $H_{\mathcal{A}}$ and equality to zero on pure states. In fact, a corollary of Proposition 4.8 and BFL's characterization is an alternative functorial characterization of Shannon entropy that does not explicitly use the non-negativity for entropy difference assumption. However, we still need one more important fact relating our notion of orthogonal affinity to BFL's notion of (external) affinity. After proving the following lemma, we will first show that our notion for a functor being orthogonally affine is equivalent to BFL's notion of a functor being externally affine on commutative C^*-algebras in Proposition 4.17. We will then use this towards building the final fact used in our characterization theorem.

Lemma 4.15 [Invariance under adjoining zero]

Suppose $H : \mathbf{NCFinProb} \to \mathbb{B}\mathbb{R}$ is an orthogonally affine fibred functor for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$. Let X and Y be finite sets and let $\iota : X \hookrightarrow X \amalg Y$ be the inclusion with associated $*$-homomorphism $\pi : \mathbb{C}^{X \amalg Y} \twoheadrightarrow \mathbb{C}^X$ (the projection). Then $H_\pi(\omega) = 0$ for all states $\omega \in \mathcal{S}(\mathbb{C}^X)$.

Proof. The map ι has a retract (a left inverse). For example, fixing any $x_0 \in X$, define

$$X \amalg Y \ni z \mapsto \begin{cases} z & \text{if } z \in X \\ x_0 & \text{if } z \in Y \end{cases}. \tag{4.16}$$

Thus π has a right inverse. By Proposition 4.8 and Lemma 4.4, $H_\pi(\omega) = 0$ for all $\omega \in \mathcal{S}(\mathbb{C}^X)$. ■

Proposition 4.17 [External versus orthogonal affinity]

Let $H : \mathbf{FinProb}^{\mathrm{op}} \to \mathbb{BR}$ be a fibred functor for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$ (cf. Lemma 4.4). Then H is orthogonally affine (Definition 3.26) if and only if H is externally affine (Definition 3.36).

Proof. ($\Rightarrow$) Suppose H is orthogonally affine. By Lemma 4.5, it suffices to consider state-preserving $*$-homomorphisms of the form $(\mathbb{C}^{X'}, \omega') \xrightarrow{f} (\mathbb{C}^X, \omega)$ and $(\mathbb{C}^{Y'}, \xi') \xrightarrow{g} (\mathbb{C}^Y, \xi)$. Their external convex sum defines a morphism

$$\left(\mathbb{C}^{X' \amalg Y'}, \lambda\widetilde{\omega}' + (1-\lambda)\widetilde{\xi}'\right) \xrightarrow{k:=f\oplus g} \left(\mathbb{C}^{X\amalg Y}, \lambda\widetilde{\omega} + (1-\lambda)\widetilde{\xi}\right), \tag{4.18}$$

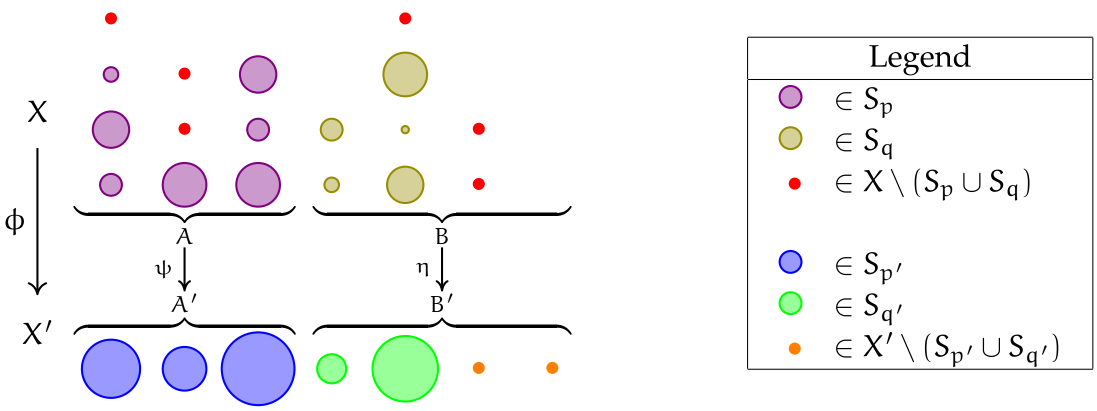
where the tildes denote the states as viewed on the direct sum (cf. Example 2.10). For example, $\widetilde{\omega}$ is the state on $\mathbb{C}^{X\amalg Y} \cong \mathbb{C}^X \oplus \mathbb{C}^Y$ obtained by the pullback of ω along the projection $\pi_X : \mathbb{C}^{X\amalg Y} \twoheadrightarrow \mathbb{C}^X$ (so that $(\mathbb{C}^X \oplus \mathbb{C}^Y, \widetilde{\omega}) \xrightarrow{\pi_X} (\mathbb{C}^X, \omega)$ is in **NCFinProb**). Furthermore,

$$\widetilde{\omega} \circ k = \widetilde{\omega}', \qquad \widetilde{\xi} \circ k = \widetilde{\xi}', \qquad \widetilde{\omega} \perp \widetilde{\xi}, \qquad \text{and} \qquad \widetilde{\omega}' \perp \widetilde{\xi}', \tag{4.19}$$

which says that $f \oplus g$ preserves the orthogonality of $\widetilde{\omega}$ and $\widetilde{\xi}$. Since H is orthogonally affine,

$$\begin{aligned} H(k) \equiv H_{f\oplus g}\left(\lambda\widetilde{\omega} + (1-\lambda)\widetilde{\xi}\right) &\overset{\text{Defn 3.26}}{=\!=} \lambda H_{f\oplus g}(\widetilde{\omega}) + (1-\lambda) H_{f\oplus g}(\widetilde{\xi}) \\ &\overset{\text{Lem 4.2}}{=\!=} \lambda\Big(H_{\mathbb{C}^{X\amalg Y}}(\widetilde{\omega}) - H_{\mathbb{C}^{X'\amalg Y'}}(\widetilde{\omega}')\Big) + (1-\lambda)\Big(H_{\mathbb{C}^{X\amalg Y}}(\widetilde{\xi}) - H_{\mathbb{C}^{X'\amalg Y'}}(\widetilde{\xi}')\Big) \\ &\overset{\text{Lem 4.15}}{=\!=} \lambda\Big(H_{\mathbb{C}^X}(\omega) - H_{\mathbb{C}^{X'}}(\omega')\Big) + (1-\lambda)\Big(H_{\mathbb{C}^Y}(\xi) - H_{\mathbb{C}^{Y'}}(\xi')\Big) \\ &\overset{\text{Lem 4.2}}{=\!=} \lambda H_f(\omega) + (1-\lambda) H_g(\xi) \equiv \lambda H(f) + (1-\lambda) H(g). \end{aligned} \tag{4.20}$$

($\Leftarrow$) Now suppose H is externally affine. As opposed to working with the C^*-algebras, it suffices to work with the associated finite sets and probability measures. As such, let p and q be probability measures on X and let p' and q' be probability measures on X'. Let $X \xrightarrow{\phi} X'$ be a function that preserves both pairs of probability measures, i.e. $\phi \circ p = p'$ and $\phi \circ q = q'$. Suppose $p \perp q$ as well as $p' \perp q'$. In what follows, we will first show that there exist morphisms $(A, p_{\restriction A}) \xrightarrow{\psi} (A', p'_{\restriction A'})$ and $(B, q_{\restriction B}) \xrightarrow{\eta} (B', q'_{\restriction B'})$ such that $\lambda\psi \oplus (1-\lambda)\eta = \phi$. Let S_r denote the support of $r \in \{p, q, p', q'\}$ (viewed as a subset of X or X' depending on the subscript). By assumption, $S_p \cap S_q = \varnothing$ and $S_{p'} \cap S_{q'} = \varnothing$. Furthermore, ϕ can be visualized as



where the indicated sets are defined by

$$A' := S_{p'}, \qquad B' := S_{q'} \cup \big(X \setminus (S_{p'} \cup S_{q'})\big), \qquad A := \phi^{-1}(A'), \qquad B := \phi^{-1}(B'), \tag{4.21}$$

and the functions $A \xrightarrow{\psi} A'$ and $B \xrightarrow{\eta} B'$ are defined by restricting ϕ to A and B, respectively. If we also define the probability measures $p_{\restriction A}, q_{\restriction B}, p'_{\restriction A'}$, and $q'_{\restriction B'}$ on A, B, A', and B', respectively, then $(A, p_{\restriction A}) \xrightarrow{\psi} (A', p'_{\restriction A'})$ and $(B, q_{\restriction B}) \xrightarrow{\eta} (B', q'_{\restriction B'})$ are morphisms in **FinProb** and most importantly,

$$\lambda \left(\begin{array}{c} (A, p_{\restriction A}) \\ \big\downarrow \psi \\ (A', p'_{\restriction A'}) \end{array} \right) \oplus (1-\lambda) \left(\begin{array}{c} (B, q_{\restriction B}) \\ \big\downarrow \eta \\ (B', q'_{\restriction B'}) \end{array} \right) = \begin{array}{c} (X, \lambda p + (1-\lambda) q) \\ \big\downarrow \phi \\ (X', \lambda p' + (1-\lambda) q') \end{array} . \tag{4.22}$$

Thus,

$$\begin{aligned} H_\phi\big(\lambda p + (1-\lambda) q\big) &\equiv H\big(\lambda\psi \oplus (1-\lambda)\eta\big) \\ &\overset{\text{Defn 3.36}}{=\!=} \lambda H(\psi) + (1-\lambda) H(\eta) \\ &= \lambda\Big(1H(\psi) + 0H(\eta)\Big) + (1-\lambda)\Big(0H(\psi) + 1H(\eta)\Big) \\ &\overset{\text{Defn 3.36}}{=\!=} \lambda H(1\psi \oplus 0\eta) + (1-\lambda) H(0\psi \oplus 1\eta) \\ &\equiv \lambda H_\phi(p) + (1-\lambda) H_\phi(q), \end{aligned} \tag{4.23}$$

which completes the proof. ■

Remark 4.24 [External affinity ignores the internal structure of quantum states]

The objects of **FinProb** are convex generated by the single object $\mathbf{1}$, which is the (essentially) unique probability space consisting of a single element. Indeed, an arbitrary finite probability space (X, p) can be decomposed into a convex sum as

$$(X, p) \cong \bigoplus_{x \in X} p_x \mathbf{1}.$$

However, in **NCFinProb**, a non-commutative probability space such as $(\mathcal{M}_m, \omega)$ cannot be expressed as an external convex combination of lower-dimensional probability spaces. Therefore, the statement "if H is externally affine (on all C^*-algebras), then H is orthogonally affine" is false.[a] The third example in Example 3.37 is a counter-example because it is not orthogonally affine. This, together with Proposition 4.17 provides some motivation for our choice of defining convex structures *internally* on the fibres over C^*-algebras.

[a] Although the converse is still true, as can be seen by a minor modification of the proof of the $(\Rightarrow)$ direction in Proposition 4.17.

Corollary 4.25 [Characterizing the Shannon entropy on commutative C^*-algebras]

Suppose $H : \mathbf{NCFinProb} \to \mathbb{B}\mathbb{R}$ is a continuous orthogonally affine fibred functor for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$. Then there exists a constant $c \geqslant 0$ such that $H_f = cS_f$ for all $*$-homomorphisms f between *commutative* C^*-algebras.

Proof. The claim follows from showing that the assumptions on H imply the assumptions of BFL for their characterization theorem (Theorem 3.38). Continuity and functoriality are already assumed. Non-negativity of $H_f(\omega)$ for all states ω and $*$-homomorphisms between commutative C^*-algebras was proved in Proposition 4.8. Finally, the notion of affine orthogonality of H is equivalent to external affinity for commutative C^*-algebras by Proposition 4.17. By BFL's characterization theorem, H is the functor giving the difference of entropies on the subcategory of commutative C^*-algebras up to an overall non-negative constant. ■

The orthogonally affine assumption for *all* C^*-algebras will provide the last fact needed to prove our characterization theorem.

Lemma 4.26 [Affine orthogonality determines entropy]

Let $H : \mathbf{NCFinProb} \to \mathbb{B}\mathbb{R}$ be a continuous and orthogonally affine fibred functor for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$. If $\omega = \sum_{x\in X} p_x\omega_x$ is any state on $\mathcal{A} := \bigoplus_{x\in X}\mathcal{M}_{m_x}$ with $\omega_x \in \mathcal{S}(\mathcal{M}_{m_x})$ and p a probability distribution on the finite set X, then there exists a constant $c \geqslant 0$ (independent of the algebras and states) such that

$$H_{\mathcal{A}}(\omega) = c\left(S(p) + \sum_{x\in X} p_x S(\omega_x)\right).$$

Proof. Let $N_p := \{x \in X : p_x = 0\}$ be the nullspace of p. For each $x \in X \setminus N_p$, decompose ω_x into a convex sum $\omega_x = \sum_{y\in Y_x} \psi_{yx}\omega_{yx}$ of pure states $\omega_{yx} \in \mathcal{S}(\mathcal{M}_{m_x})$, where Y_x is a finite set and $\{\psi_{yx}\}_{y\in Y_x}$ defines a nowhere-vanishing probability measure on Y_x so that $|Y_x|$ equals the rank (dimension of the range) of the support of ω_x. Thus, $X \setminus N_p \overset{\psi}{\rightsquigarrow} \coprod_{x\in X\setminus N_p} Y_x$ defines a stochastic map. Let $P_{yx} \in \mathcal{M}_{m_x}$ denote the one-dimensional projection associated to the pure state ω_{yx}. If P_x denotes the support of ω_x, then $P_x = \sum_{y\in Y_x} P_{yx}$ for all $x \in X \setminus N_p$. Set

$$\mathcal{B} := \left(\bigoplus_{x\in X\setminus N_p} \mathbb{C}^{Y_x}\right) \oplus \mathbb{C}^{\{\bullet\}}, \tag{4.27}$$

where $\mathbb{C}^{\{\bullet\}} \cong \mathbb{C}$ and $\bullet$ serves as a label to distinguish it from the rest of the algebra. Define a $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ by

$$\mathbb{C}^{Y_x} \ni e_y \overset{f}{\mapsto} \left(\bigoplus_{x'\in X\setminus\{x\}} 0\right) \oplus P_{yx} \quad\text{and}\quad \mathbb{C}^{\{\bullet\}} \ni e_\bullet \overset{f}{\mapsto} \left(\bigoplus_{x\in X\setminus N_p} (\mathbb{1}_{m_x} - P_x)\right) \oplus \bigoplus_{x\in N_p} \mathbb{1}_{m_x}, \tag{4.28}$$

where the first expression involving P_{yx} is simply meant that P_{yx} is viewed now as an element of $\mathcal{B}$ (with 0's on all factors other than $\mathcal{M}_{m_x}$). Then f is a (unital) $*$-homomorphism that preserves the orthogonality of *all* the ω_{yx} states with $y \in Y_x$ and $x \in X \setminus N_p$ (by viewing all the ω_{yx} as states on $\mathcal{A}$ via Lemma 4.15). Therefore,

$$\begin{aligned} H_{\mathcal{A}}(\omega) - H_{\mathcal{B}}(\omega \circ f) = H_f(\omega) &= \sum_{x \in X \setminus N} \sum_{y \in Y_x} p_x \psi_{yx} H_f(\omega_{yx}) \\ &= \sum_{x \in X \setminus N} \sum_{y \in Y_x} p_x \psi_{yx} \big(H_{\mathcal{A}}(\omega_{yx}) - H_{\mathcal{B}}(\omega_{yx} \circ f) \big) = 0 \end{aligned} \tag{4.29}$$

because $\omega_{yx} \circ f$ is a pure state. Consequently,

$$\begin{aligned} H_{\mathcal{A}}(\omega) &= H_{\mathcal{B}}(\omega \circ f) \\ &= -c \sum_{x \in X \setminus N_p} \sum_{y \in Y_x} p_x \psi_{yx} \log(p_x \psi_{yx}) \\ &= -c \sum_{x \in X \setminus N_p} \underbrace{\sum_{y \in Y_x} \psi_{yx}}_{1} p_x \log(p_x) + c \sum_{x \in X \setminus N_p} p_x \left(- \sum_{y \in Y_x} \psi_{yx} \log(\psi_{yx}) \right) \\ &= c \left(S(p) + \sum_{x \in X} p_x S(\omega_x) \right), \end{aligned} \tag{4.30}$$

where in the second equality we have used Corollary 4.25, which says H restricts to the Shannon entropy on commutative C^*-algebras up to a non-negative constant c. The last equality follows from the definition of the Shannon entropy for the $S(p)$ term and Lemma 2.24 for the $S(\omega_x)$ term. ■

Theorem 4.31 [A functorial characterization of quantum entropy]

Let $H : \mathbf{NCFinProb} \to \mathbb{BR}$ be a continuous and orthogonally affine fibred functor

$$\begin{tikzcd} \mathbf{NCFinProb} \arrow[r, "H"] \arrow[d] & \mathbb{BR} \arrow[d] \\ \mathbf{fdC^*\text{-}Alg} \arrow[r] & \mathbf{\underline{1}} \end{tikzcd}$$

for which $H_{\mathcal{A}}(\omega) \geqslant 0$ for all states $\omega \in \mathcal{S}(\mathcal{A})$, with equality on all pure states, for all C^*-algebras $\mathcal{A}$. Then there exists a constant $c \geqslant 0$ such that

$$H_f(\omega) = c \big(S(\omega) - S(\omega \circ f) \big)$$

for all morphisms $\mathcal{B} \xrightarrow{f} \mathcal{A}$ of C^*-algebras and states $\omega \in \mathcal{S}(\mathcal{A})$.

Proof. Since $H_f(\omega) = H_{\mathcal{A}}(\omega) - H_{\mathcal{B}}(\omega \circ f)$ by Lemma 4.2, Lemma 4.5 and Lemma 4.26 together show this equals the entropy difference up to the same constant c. ■

It is interesting that the notion of a disintegration was used in the proof of Proposition 4.8. Note that in the category of states on (finite-dimensional) C^*-algebras and state-preserving $*$-homomorphisms, disintegrations do not always exist [42]. Nevertheless, their existence is well understood, and when they exist, they imply $H_f(\omega) \geqslant 0$, as the following proposition shows. Since the definition of a non-commutative disintegration is not needed anywhere else in this work, the reader is referred to [42] for definitions and other facts assumed in the proof.

Proposition 4.32 [If a disintegration for (f, ω) exists, then $S_f(\omega) \geqslant 0$]

Let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be a $*$-homomorphism and $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$ a state on $\mathcal{A}$. If (f, ω) has a disintegration, then $S_f(\omega) \geqslant 0$.

Proof. By isomorphism invariance of S, it suffices to consider the case where $\mathcal{A}, \mathcal{B}, \omega$, and f are as in Lemma 2.13 (without the unitaries U_x). Write $\xi := \omega \circ f = \sum_{y \in Y} q_y \operatorname{tr}(\sigma_y \ \cdot\ )$ and $\omega = \sum_{x \in X} p_x \operatorname{tr}(\rho_x \ \cdot\ )$ as in Lemma 2.13 as well. Let $N_p \subset X$ and $N_q \subset Y$ be the null-spaces of p and q, respectively. Assume that a disintegration of (f, ω, ξ) exists. By the non-commutative disintegration theorem [42, Theorem 5.108], for each $x \in X$ and $y \in Y$ there exist non-negative matrices $\tau_{yx} \in \mathcal{M}_{c_{xy}}$ such that

$$\operatorname{tr}\left(\sum_{x \in X} \tau_{yx}\right) = 1 \qquad \forall\, y \in Y \setminus N_q \tag{4.33}$$

and

$$p_x \rho_x = \boxplus_{y \in Y} \tau_{yx} \otimes q_y \sigma_y \qquad \forall\, x \in X. \tag{4.34}$$

One more fact that will be needed is the equality

$$(C \otimes D) \log(C \otimes D) = C \log(C) \otimes D + C \otimes D \log(D) \tag{4.35}$$

for all non-negative square matrices C, D (possibly of different sizes), which can be proved using the spectral theorem for matrices. Computing $S_{\mathcal{A}}(\omega)$ first gives

$$\begin{aligned} S_{\mathcal{A}}(\omega) &\overset{\text{Rmk 2.25}}{=\!=} -\sum_{x \in X} \operatorname{tr}\big(p_x \rho_x \log(p_x \rho_x)\big) \\ &\overset{(4.34)}{=\!=} -\sum_{x \in X} \operatorname{tr}\left(\boxplus_{y \in Y \setminus N_q} (\tau_{yx} \otimes q_y \sigma_y) \log\left(\boxplus_{y' \in Y \setminus N_q} (\tau_{y'x} \otimes q_{y'} \sigma_{y'})\right)\right) \\ &= -\sum_{x \in X} \sum_{y \in Y \setminus N_q} \operatorname{tr}\big((\tau_{yx} \otimes q_y \sigma_y) \log(\tau_{yx} \otimes q_y \sigma_y)\big) \\ &\overset{(4.35)}{=\!=} -\sum_{x \in X} \sum_{y \in Y \setminus N_q} \operatorname{tr}\big(\tau_{yx} \log(\tau_{yx}) \otimes q_y \sigma_y + \tau_{yx} \otimes q_y \sigma_y \log(q_y \sigma_y)\big) \\ &\overset{(4.33)}{=\!=} \sum_{y \in Y \setminus N_q} q_y S\left(\boxplus_{x \in X} \tau_{yx}\right) + S_{\mathcal{B}}(\xi), \end{aligned} \tag{4.36}$$

where $\boxplus_{x\in X}\tau_{yx}$ is viewed as a density matrix on $\mathcal{M}_{s_x}$, where $s_x := \sum_{y\in Y\setminus N_q} c_{yx}$. Thus,

$$S_f(\omega) = S_{\mathcal{A}}(\omega) - S_{\mathcal{B}}(\xi) = \sum_{y\in Y\setminus N_q} q_y S\left(\boxplus_{x\in X}\tau_{yx}\right) \geqslant 0. \tag{4.37}$$

■

Remark 4.38 [Having a disintegration is not necessary for $S_f(\omega)\geqslant 0$]

If $S_f(\omega)\geqslant 0$, it is not necessarily the case that a disintegration of (f,ω) exists. A simple counter-example is the inclusion $f:\mathcal{B}\to\mathcal{A}\otimes\mathcal{B}$ sending $B\in\mathcal{B}$ to $1_{\mathcal{A}}\otimes B$ and where $\mathcal{B}:=\mathcal{M}_2$ and $\mathcal{A}:=\mathcal{M}_2$. Take ω to be represented by the density matrix

$$\rho = \begin{bmatrix} p_1 & 0 & 0 & 0\\ 0 & p_2 & 0 & 0\\ 0 & 0 & p_3 & 0\\ 0 & 0 & 0 & p_4\end{bmatrix},$$

where $p_1,p_2,p_3,p_4\geqslant 0$ satisfy $p_1+p_2+p_3+p_4=1$, $p_1+p_3>0$, and $p_2+p_4>0$. Then

$$S_f(\omega) = p_1\log\left(\frac{p_1+p_3}{p_1}\right) + p_2\log\left(\frac{p_2+p_4}{p_2}\right) + p_3\log\left(\frac{p_1+p_3}{p_3}\right) + p_4\log\left(\frac{p_2+p_4}{p_4}\right) \geqslant 0,$$

while a disintegration exists if and only if $p_1p_4 = p_2p_3$ [42, Theorem 4.19].

Remark 4.39 [A brief history and comparison of axiomatizations of quantum entropy]

Quantum entropy and its variants were often built upon the classical versions, whose many axiomatizations are reviewed in Csiszar's survey [7]. In 1932, von Neumann obtained a phenomelogical characterization of entropy [55, Chapter V. Section 2]. In 1968, Ingarden and Kossakowski characterized the von Neumann entropy using dimensional partial Boolean rings of projections in Hilbert space [24]. In 1974, Ochs provided a characterization using partial isometric invariance, additivity, subadditivity, and continuity (plus some additional technical axioms) [35]. In 1975, Thirring [51] characterized the von Neumann entropy using axioms closely related to those implemented by Fadeev in his characterization of the Shannon entropy [8, 9], the latter of which was simplified by Renyi [47].[a] Thirring's characterization is therefore most closely related to ours and it is worth taking the time to spell out his assumptions, which read as follows.

(i) $S(\rho)$ is a continuous function of the eigenvalues of ρ;

(ii) $S(\frac{1}{2}\mathbb{1}_2) = \log 2$;

(iii) If $\mathcal{H} = \bigoplus_{n=1}^N \mathcal{H}_n$ is a direct sum of Hilbert spaces and if $\rho = \bigoplus_{n=1}^N p_n\rho_n$ is a weighted direct sum of density matrices, where $\{p_n\}_{n\in\{1,\dots,N\}}$ is a probability distribution on Δ^N, then $S(\rho) = S(p) + \sum_{n=1}^N p_n S(\rho_n)$, where p is viewed as a diagonal matrix on $\mathbb{C}^N$ with entries given by the p_n.

There are actually several implicitly hidden assumptions within these three. For example, the dependence on eigenvalues means $S(\rho) = S(\mathrm{Ad}_U \rho)$ for all unitaries U, i.e. $S(\rho)$ is invariant under $*$-isomorphisms. The second item is merely a normalization condition, which we have ignored (it specifies the constant c). The third item is close to our orthogonal affine assumption. However, an implicit assumption is made, which can be expressed as saying that $S(\rho_n)$ is equal to $S(0 \oplus \cdots \rho_n \oplus \cdots 0)$, i.e. S is invariant under the non-unital inclusion of one matrix algebra into a direct sum. This is closely related to Och's partial isometry invariance assumption, which has been criticized as being an unnatural assumption from the physical perspective [5]. We don't necessarily agree with the criticism as we prefer not to place any demand on what axiom seems natural or not. Nevertheless, we find it quite satisfying that this assumption *as well as* invariance under $*$-isomorphisms (which corresponds to Axiom B in [5]) are consequences of our axioms.

[a] Much of Thirring's lectures on mathematical physics were translated into English, and his statement and proof can be found in [52, (2.2.4) pages 58–61]; however, it seems that the first written account of his proof in English appears in Wehrl's review [57, pages 238–239].

Question 4.40 [Other convex categories and functorial entropy]

The presentation we have given in terms of fibrations is dual to an indexed category formulation, which is prominent in some earlier work [38,39]. The latter would have involved more higher category theory, which is why we have chosen the more geometric perspective in terms of fibrations. In either case, these viewpoints suggest several points of generalization. For one, we could imagine replacing the fibration $\mathbb{BR} \to \underline{\mathbf{1}}$ from Theorem 4.31 with another fibration $\mathcal{E} \to \mathcal{X}$ of cone categories [36, Section 4.5.1] and *define* a notion of $\mathcal{E}$-valued entropy. Many of the ideas can be generalized to this broader setting, and we hope to make some of this development available in forthcoming work. However, it is not clear if there is a need for such generalizations in physics. Are there other interesting examples of convex categories appearing naturally in in the structure of physical systems?

A The Holevo information change and relative entropy

The monotonicity of relative entropy for density matrices under partial trace is a well-known and deep result in quantum information theory. Here, we phrase some of these ideas from the functorial perspective in terms of the deviation from S_f being affine on *all* states. This deviation is well-controlled by the *change* in the Holevo information associated to a $*$-homomorphism, and it turns out that this change is *always* non-negative. Although this appendix is not needed for the main results proved in this manuscript, it serves to place some of the results in a broader context and it also provides a glimpse of work that is in preparation.

Theorem A.1 [Holevo information change is non-negative]

Using the notation from Definition 3.28, the Holevo information change satisfies

$$\chi_f(\lambda;\omega,\xi) \geqslant 0 \qquad \forall\, \lambda \in [0,1].$$

Furthermore, if $\omega \perp \xi$, then a $*$-homomorphism $\mathcal{B} \xrightarrow{f} \mathcal{A}$ preserves the orthogonality $\omega \perp \xi$ if and only if

$$\chi_f(\lambda;\omega,\xi) = 0$$

for any (and hence all) $\lambda \in (0,1)$.

Proof. Note that the first claim $\chi_f(\lambda;\omega,\xi) \geqslant 0$ is clearly true if $\lambda \in \{0,1\}$ (it equals zero). Hence, fix $\lambda \in (0,1)$. The inequality $\chi_f(\lambda;\omega,\xi) \geqslant 0$ was first proven for the special case where $\mathcal{B}$ and $\mathcal{A}$ are matrix algebras by Lieb and Ruskai [26, Theorem 1]. To see the relationship between their statement and ours, their claim is that the assignment

$$\rho \mapsto S\big(\mathrm{tr}_{\mathcal{M}_p}(\rho)\big) - S(\rho), \tag{A.2}$$

which is expressed in the Schrödinger picture, is convex on density matrices, where ρ is a density matrix on $\mathcal{A} = \mathcal{M}_p \otimes \mathcal{M}_n$, $\mathcal{B} = \mathcal{M}_n$, and $\mathrm{tr}_{\mathcal{M}_p} : \mathcal{A} \rightsquigarrow \mathcal{B}$ is the Hilbert–Schmidt dual of the inclusion $f : \mathcal{B} \to \mathcal{A}$ into the second factor (cf. Example 2.12). This expression is therefore given by $S(\omega \circ f) - S(\omega)$ in our Heisenberg picture, where $\omega = \mathrm{tr}(\rho\ \cdot\ )$. Setting $\zeta := \lambda\omega + (1-\lambda)\xi$, convexity of this quantity then says

$$S(\zeta \circ f) - S(\zeta) \leqslant \lambda S(\omega \circ f) + (1-\lambda)S(\xi \circ f) - \lambda S(\omega) - (1-\lambda)S(\xi). \tag{A.3}$$

Rearranging this and using Remark 3.29 gives the desired claim for matrix algebras.

When $\mathcal{A}$ is not necessarily a matrix algebra, $\mathcal{B}$ is a matrix algebra, and f has multiplicity 1 for each subfactor of $\mathcal{A}$, this was proved by Lindblad [28, Lemma 3] (though one needs to use Lemma A.23 below to see this).

For the most general case, let $\bigoplus_{y\in Y}\mathcal{M}_{n_y} \xrightarrow{f} \bigoplus_{x\in X}\mathcal{M}_{m_x}$ be a $*$-homomorphism. Then f can be decomposed as the following composite

$$\bigoplus_{y\in Y}\mathcal{M}_{n_y} \xrightarrow{F} \bigoplus_{x\in X}\bigoplus_{y\in Y}\mathcal{M}_{c_{xy}n_y} \xhookrightarrow{L:=\bigoplus_{x\in X}L_x} \bigoplus_{x\in X}\mathcal{M}_{m_x} \xrightarrow{\varphi:=\bigoplus_{x\in X}\mathrm{Ad}_{U_x}} \bigoplus_{x\in X}\mathcal{M}_{m_x}, \tag{A.4}$$

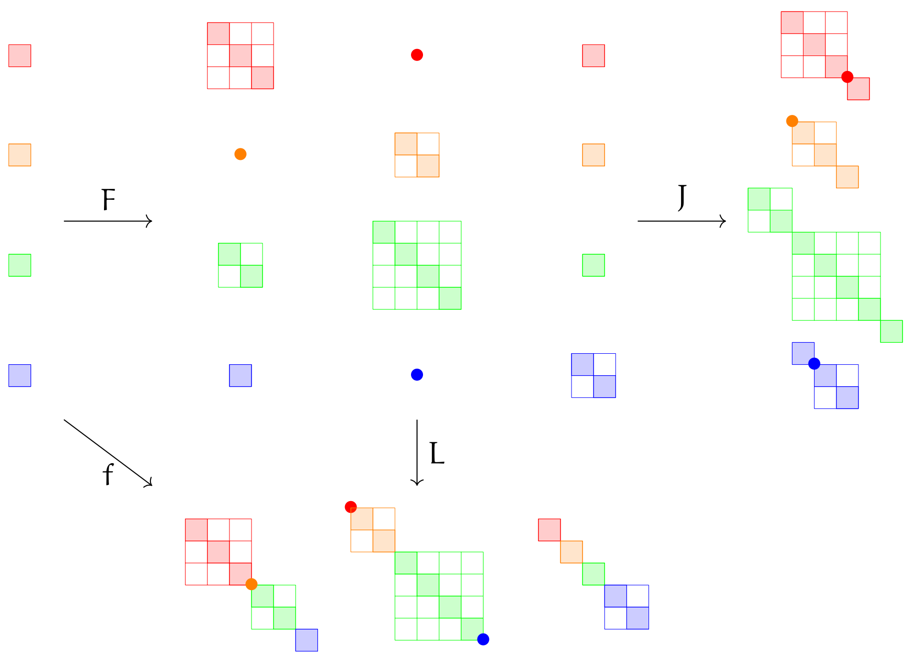
where each of these maps will be described presently. First, each c_{xy} is a non-negative integer satisfying $m_x = \sum_{y \in Y} c_{xy} n_y$ (the multiplicity of $\mathcal{M}_{n_y}$ inside $\mathcal{M}_x$ under f). The map $F : \bigoplus_{y \in Y} \mathcal{M}_{n_y} \to \bigoplus_{y \in Y} \bigoplus_{x \in X} \mathcal{M}_{c_{xy} n_y}$ is determined by the assignment

$$\mathcal{M}_{n_y} \ni B_y \mapsto \bigoplus_{x \in X} \operatorname{diag}(\overbrace{B_y, \cdots, B_y}^{c_{yx} \text{ times}}) \in \bigoplus_{x \in X} \mathcal{M}_{c_{xy} n_y}. \tag{A.5}$$

The map $L_x : \bigoplus_{y \in Y} \mathcal{M}_{c_{xy} n_y} \hookrightarrow \mathcal{M}_{m_x}$ is the inclusion sending $\bigoplus_{y \in Y} C_y$ to $\boxplus_{y \in Y} C_y$; by choosing an ordering on these indices, this latter map looks like

$$\bigoplus_{y \in |Y|} C_y \xmapsto{L_x} \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_{|Y|} \end{bmatrix} \in \mathcal{M}_{m_x}. \tag{A.6}$$

Finally, the last map in (A.4) is conjugation by a unitary so that the composite can be used to describe an arbitrary $*$-homomorphism. Since the conjugation does not change the entropy, it will be henceforth ignored. In case the notation is a bit overwhelming, these maps (together with another map J that will be defined in (A.11)) can be visualized as



with the x and y axes being the horizontal and vertical axis, respectively, where each filled box represents a matrix algebra, the bullets represent the fact that the multiplicity is zero in those cases, and a space between blocks is used to signify a direct sum. Now let ω and ξ be states on $\mathcal{A}$ as in the proof of Lemma 2.27 item 1. By Remark 3.29,

$$\chi_f(\lambda; \omega, \xi) = \chi_L(\lambda; \omega, \xi) + \chi_F(\lambda; \omega \circ L, \xi \circ L). \tag{A.7}$$

The first term satisfies $\chi_L(\lambda; \omega, \xi) \geqslant 0$ because each L_x is precisely a map of the form covered by Lindblad's result. In more detail, setting

$$r_x := \lambda p_x + (1 - \lambda) q_x, \tag{A.8}$$

one has

$$\chi_L(\lambda;\omega,\xi) = \sum_{x\in X} r_x \chi_{L_x}\left(\frac{\lambda p_x}{r_x};\rho_x,\sigma_x\right) \geqslant 0 \tag{A.9}$$

(dropping any terms for which $r_x = 0$), the first equality of which hinges on the fact that

$$S\big(\lambda\omega + (1-\lambda)\xi\big) \xlongequal{\text{Defn } 2.23} S(r) + \sum_{x\in X} r_x S\left(\frac{\lambda p_x}{r_x}\rho_x + \frac{(1-\lambda)q_x}{r_x}\sigma_x\right). \tag{A.10}$$

As for $\chi_F(\lambda;\omega\circ L,\xi\circ L)$, this term breaks up into two parts. To describe this, set $s_y := \sum_{x\in X} c_{xy}$ and define the map $J := \bigoplus_{y\in Y} J_y$, where

$$\begin{aligned} \bigoplus_{x\in X} \mathcal{M}_{c_{xy}n_y} &\xrightarrow{J_y} \mathcal{M}_{s_y n_y} \\ \bigoplus_{x\in X} C_{xy} &\mapsto \boxplus_{x\in X} C_{xy}. \end{aligned} \tag{A.11}$$

Then let ρ_{xy} and σ_{xy} be the unique set of matrices in $\mathcal{M}_{c_{xy}n_y}$ satisfying

$$\bigoplus_{y\in Y} \rho_{xy} = L_x^*(\rho_x) \quad \text{and} \quad \bigoplus_{y\in Y} \sigma_{xy} = L_x^*(\sigma_x). \tag{A.12}$$

Define

$$\pi_y := \boxplus_{x\in X} p_x\rho_{xy}, \quad \tau_y := \boxplus_{x\in X} q_x\sigma_{xy}, \quad \pi := \bigoplus_{y\in Y} \pi_y, \quad \tau := \bigoplus_{y\in Y} \tau_y, \tag{A.13}$$

and let α and β be the states on $\bigoplus_{y\in Y} \mathcal{M}_{s_y n_y}$ defined by

$$\alpha := \sum_{y\in Y} \operatorname{tr}(\pi_y\,\cdot\,) \quad \text{and} \quad \beta := \sum_{y\in Y} \operatorname{tr}(\tau_y\,\cdot\,). \tag{A.14}$$

If it helps the reader, we have merely taken the different blocks from our density matrices associated to $\omega\circ L$ and $\xi\circ L$ and have rearranged them (it is unfortunate that describing this rigorously is quite complicated and we hope the above figure helps illustrate this rearrangement more clearly). All this is to say that

$$\chi_F(\lambda;\omega\circ L,\xi\circ L) = \chi_F(\lambda;\omega\circ L,\xi\circ L) + \chi_J(\lambda;\alpha,\beta) \xlongequal{\text{Rmk } 3.29} \chi_{J\circ F}(\lambda;\alpha,\beta) \tag{A.15}$$

since $\chi_J(\lambda;\alpha,\beta) = 0$ (the entropy does not change since there has essentially only been a permutation of matrices). However, $J\circ F$ is a direct sum of $*$-homomorphisms of the kind considered by Lieb and Ruskai, so that $\chi_{J\circ F}(\lambda;\alpha,\beta) \geqslant 0$. In more detail, if $(J\circ F)_y : \mathcal{M}_{n_y} \to \mathcal{M}_{s_y n_y}$ is the y-component of the map $J\circ F$ (which is now precisely the Hilbert–Schmidt dual of a partial trace) and if we set

$$a_y := \operatorname{tr}(\pi_y), \quad b_y := \operatorname{tr}(\tau_y), \quad \text{and} \quad d_y := \lambda a_y + (1-\lambda)b_y, \tag{A.16}$$

then

$$\chi_{J\circ F}(\lambda;\alpha,\beta) = \sum_{y\in Y} d_y \chi_{(J\circ F)_y}\left(\frac{\lambda a_y}{d_y};\frac{\pi_y}{a_y},\frac{\tau_y}{b_y}\right) \geqslant 0, \tag{A.17}$$

where the last inequality is the one by Lieb and Ruskai. Thus,

$$\chi_f(\lambda;\omega,\xi) = \chi_{L\circ F}(\lambda;\omega,\xi) = \chi_L(\lambda;\omega,\xi) + \chi_{J\circ F}(\lambda;\alpha,\beta) \geqslant 0, \tag{A.18}$$

concluding the proof. ■

As mentioned in Example 3.17, Theorem A.1 says that the entropy change is a *concave* functor. Rather than defining concave fibred functors in full generality (which would involve the introduction of cone categories [36, Section 4.5.1] and order structures), we specialize to the case where the codomain is $\mathbb{BR}$, where the cone and order structure is the standard one.

Definition A.19 [Concave fibred functors]

Let $\mathcal{E} \xrightarrow{\pi} \mathcal{X}$ be a fibrewise convex fibration. A ***concave fibred functor*** from π to $\mathbb{BR} \to \bullet$ is a fibred functor

$$\begin{tikzcd} \mathcal{E} \arrow[r, "H"] \arrow[d, "\pi"'] & \mathbb{BR} \arrow[d] \\ \mathcal{X} \arrow[r] & \underline{\mathbf{1}} \end{tikzcd}$$

such that

$$H\big(\lambda\alpha + (1-\lambda)\beta\big) \geqslant \lambda H(\alpha) + (1-\lambda)H(\beta)$$

for every pair of morphisms $(t \xrightarrow{\alpha} u, v \xrightarrow{\beta} w)$ with $\pi(\alpha) = \pi(\beta)$ (cf. Remark 3.21).

This notion is part of a developing field of categorical convex analysis [36,44]. The inequality in Theorem A.1 is also closely related to the monotonicity under partial trace of Umegaki's relative entropy [53, Section 4].

Definition A.20 [Umegaki's relative entropy]

Let $\omega = \sum_{x\in X} p_x \mathrm{tr}(\rho_x\ \cdot\ )$ and $\xi = \sum_{x\in X} q_x \mathrm{tr}(\sigma_x\ \cdot\ )$ be two states on $\mathcal{A} = \bigoplus_{x\in X} \mathcal{M}_{m_x}$ such that $\omega \preceq \xi$, i.e. $P_\omega \leqslant P_\xi$. The ***entropy of*** ω ***relative to*** ξ (also just called the ***relative entropy***) is the number

$$S(\omega||\xi) \equiv S(\rho||\sigma) := \mathrm{tr}(\rho\log\rho - \rho\log\sigma),$$

where

$$\rho := \bigoplus_{x\in X} p_x\rho_x \qquad \text{and} \qquad \sigma := \bigoplus_{x\in X} q_x\sigma_x.$$

We can write the relative entropy in Definition A.20 more explicitly in terms of the constituent probabilities and density matrices from the direct sum.

Lemma A.21 [Alternative expression for the relative entropy of states]

Given $\omega \preceq \xi$ for states on $\mathcal{A}$, there always exists a decomposition of ω and ξ as in Definition A.20 such that $\rho_x \preceq \sigma_x$ for all $x \in X$. With such choices,

$$S(\omega\|\xi) = S(p\|q) + \sum_{x\in X} p_x S(\rho_x\|\sigma_x),$$

where $S(p\|q)$ is the Kullback–Leibler divergence (aka classical relative entropy)

$$S(p\|q) := \sum_{x\in X} (p_x \log p_x - p_x \log q_x).$$

Proof. Note that $p \preceq q$ holds automatically since $\omega \preceq \xi$. If $p_x = 0$, then one can choose *any* density matrix ρ_x for that value of $x \in X$. In particular, one may set $\rho_x = \sigma_x$. A more thorough discussion justifying this can be found at the beginning of Section 6 in [43]. The rest of the proof of this is a straightforward calculation:

$$\begin{aligned} S(\rho\|\sigma) &= \mathrm{tr}\left(\bigoplus_{x\in X}\Big(p_x\rho_x \log(p_x\rho_x) - p_x\rho_x\log(q_x\sigma_x)\Big)\right) \\ &= \mathrm{tr}\left(\bigoplus_{x\in X}\Big(p_x\log(p_x)\rho_x + p_x\rho_x\log(\rho_x) - p_x\log(q_x)\rho_x - p_x\rho_x\log(\sigma_x)\Big)\right) \\ &= \sum_{x\in X}\Big(p_x\log p_x - p_x\log q_x + p_x\mathrm{tr}(\rho_x\log\rho_x) - p_x\mathrm{tr}(\rho_x\log\sigma_x)\Big). \end{aligned} \tag{A.22}$$

■

Lemma A.23 [Expressing the Holevo information change in terms of relative entropy]

Let $\mathcal{A}$ be a C^*-algebra. Then

$$\chi_{\mathcal{A}}(\lambda;\omega,\xi) = \lambda S\big(\omega\big\|\lambda\omega + (1-\lambda)\xi\big) + (1-\lambda)S\big(\xi\big\|\lambda\omega + (1-\lambda)\xi\big)$$

for all states ω,ξ on $\mathcal{A}$ and for all $\lambda \in (0,1)$. When $\lambda \in \{0,1\}$, one has $\chi_{\mathcal{A}}(\lambda;\omega,\xi) = 0$ for all $\omega,\xi \in \mathcal{S}(\mathcal{A})$.

The expression in Lemma A.23 is guaranteed to be well-defined for *all* states because

$$\omega \preceq \lambda\omega + (1-\lambda)\xi \qquad \text{and} \qquad \xi \preceq \lambda\omega + (1-\lambda)\xi \tag{A.24}$$

for all $\lambda \in (0,1)$. What is perhaps less clear is if the expression in terms of the relative entropy is continuous at $\lambda \in \{0,1\}$. It turns out the expression is more than just continuous, as will be discussed in Lemma A.26.

Proof of Lemma A.23. Let $\mathcal{A}$, ρ, and σ be as in Definition A.20. The claim then follows from a straightforward calculation:

$$\begin{aligned} &\chi_{\mathcal{A}}(\lambda;\omega,\xi) = S\big(\lambda\omega + (1-\lambda)\xi\big) - \lambda S(\omega) - (1-\lambda)S(\xi) \\ &= -\mathrm{tr}\Big(\big(\lambda\rho + (1-\lambda)\sigma\big)\log\big(\lambda\rho + (1-\lambda)\sigma\big)\Big) + \lambda\mathrm{tr}(\rho\log\rho) + (1-\lambda)\mathrm{tr}(\sigma\log\sigma) \\ &= \lambda\mathrm{tr}\Big(\rho\log\rho - \rho\log\big(\lambda\rho + (1-\lambda)\sigma\big)\Big) + (1-\lambda)\mathrm{tr}\Big(\sigma\log\sigma - \sigma\log\big(\lambda\rho + (1-\lambda)\sigma\big)\Big). \end{aligned} \tag{A.25}$$

The fact that $\chi_{\mathcal{A}}(\lambda;\omega,\xi) = 0$ for $\lambda \in \{0,1\}$ is immediate from the definition of χ. ■

Lemma A.26 [Lindblad's lemma]

Let $\omega, \xi : \mathcal{A} \rightsquigarrow \mathbb{C}$ be two states on a C^*-algebra with $\omega \preceq \xi$. Then

$$S(\omega\|\xi) = \lim_{\lambda\to 0}\left(\frac{\chi_{\mathcal{A}}(\lambda;\omega,\xi)}{\lambda}\right).$$

Proof. Lindblad proved this for the case of matrix algebras in [28, Lemma 4] by expressing χ in terms of the relative entropy as in Lemma A.23 above (see also [29]). Here, we extend the proof to states on C^*-algebras. In what follows, it is convenient to set

$$r_x := \lambda p_x + (1-\lambda)q_x, \qquad s_x := \frac{\lambda p_x}{r_x}, \qquad \text{and} \qquad \tau_x := s_x\rho_x + (1-s_x)\sigma_x, \tag{A.27}$$

where we have also used our standard notation for ω and ξ as in Definition A.20 for example. Of course, s_x is defined only when $r_x \neq 0$ (it will not need to be defined for other values of x). Then

$$\begin{aligned}
\chi_{\mathcal{A}}(\lambda;\omega,\xi) &\overset{\text{Rmk 3.29}}{=\!=} S\big(\lambda\omega + (1-\lambda)\xi\big) - \lambda S(\omega) - (1-\lambda)S(\xi)\\
&\overset{\text{Defn 2.23}}{=\!=} S(r) + \sum_{x\in X} r_x S(\tau_x) - \lambda\left(S(p) + \sum_{x\in X} p_x S(\rho_x)\right) - (1-\lambda)\left(S(q) + \sum_{x\in X} q_x S(\sigma_x)\right)\\
&= \chi_X(\lambda;p,q) + \sum_{x\in X} r_x\Big(S(\tau_x) - s_x S(\rho_x) - (1-s_x)S(\sigma_x)\Big),
\end{aligned} \tag{A.28}$$

where $\chi_X(\lambda;p,q) := S(\lambda p + (1-\lambda)q) - \lambda S(p) - (1-\lambda)S(q)$. Dividing by λ and taking the $\lambda \to 0$ limit gives

$$\lim_{\lambda\to 0}\left(\frac{\chi_{\mathcal{A}}(\lambda;\omega,\xi)}{\lambda}\right) = \lim_{\lambda\to 0}\left(\frac{\chi_X(\lambda;p,q)}{\lambda}\right) + \sum_{x\in X}\lim_{\lambda\to 0}\left(r_x\left(\frac{S(\tau_x) - s_x S(\rho_x) - (1-s_x)S(\sigma_x)}{\lambda}\right)\right). \tag{A.29}$$

The first term is $S(p\|q)$, as one can easily check. Noting that

$$\lim_{\lambda\to 0} s_x(\lambda) = 0 \qquad \text{and} \qquad \lambda = \frac{q_x s_x}{p_x + s_x(q_x - p_x)}, \tag{A.30}$$

the second term splits into two terms, the first of which is

$$\lim_{\lambda\to 0} r_x = q_x, \tag{A.31}$$

and the second of which is

$$\begin{aligned}
&\lim_{\lambda\to 0}\left(\frac{S(\tau_x) - s_x S(\rho_x) - (1-s_x)S(\sigma_x)}{\lambda}\right)\\
&\quad = \lim_{s_x\to 0}\left(\frac{p_x + s_x(q_x - p_x)}{q_x}\right)\left(\frac{S(\tau_x) - s_x S(\rho_x) - (1-s_x)S(\sigma_x)}{s_x}\right)\\
&\quad = \frac{p_x}{q_x} S(\rho_x\|\sigma_x),
\end{aligned} \tag{A.32}$$

where in the last step we used the established result for matrix factors. Note that one never has to worry about dividing by zero in any of these expressions because of the $\omega \preceq \xi$ assumption. Plugging this result back into (A.29) and using Lemma A.21 gives the desired conclusion. ∎

Lemma A.26 and Lemma A.23 allows one to conveniently go back and forth between relative entropy and the Holevo information change.

Theorem A.33 [Monotonicity of relative entropy under partial trace]

Let $\omega, \xi : \mathcal{A} \rightsquigarrow \mathbb{C}$ be two states on a C^*-algebra with $\omega \preceq \xi$ and let $\mathcal{B} \xrightarrow{f} \mathcal{A}$ be a $*$-homomorphism. Then

$$S(\omega \circ f \| \xi \circ f) \leqslant S(\omega \| \xi).$$

Proof. By Theorem A.1 and then Remark 3.29,

$$0 \leqslant \chi_f(\lambda; \omega, \xi) = \chi_{\mathcal{A}}(\lambda; \omega, \xi) - \chi_{\mathcal{B}}(\lambda; \omega \circ f, \xi \circ f). \tag{A.34}$$

Dividing both sides by $\lambda \in (0,1)$, taking $\lim_{\lambda \to 0}$, and using Lemma A.26 gives the desired inequality. ∎

Question A.35 [A functorial characterization of quantum relative entropy]

The relative entropy naturally appears as a byproduct of Lindblad's Lemma and our functorial characterization of the quantum entropy. However, this is not quite a functorial characterization of the quantum relative entropy, at least not in the spirit of the recent functorial characterizations of the relative Shannon entropy (or Kullback–Leibler divergence) existing in the literature [3, 16, 25]. Therefore, a natural question to ask is if the quantum relative entropy has a functorial characterization, generalizing the characterization of Baez and Fritz [3]. Preliminary calculations suggest this may be possible, and these results will be presented elsewhere. Some of the current characterizations of the quantum relative entropy utilize Theorem A.33 as one of the main axioms [31, 59], though we suspect this can be avoided by re-examining Petz' characterization [45] from a categorical perspective. What is less clear is if there is a unifying principle encapsulating all these categorical results. For example, can these results be viewed from the abstract perspective of categories that encode information processing (CD categories, Markov categories, etc.) [6, 15, 40]?

Acknowledgements

The work presented in this manuscript began while the author was at the CUNY Graduate Center, continued when the author was at the University of Connecticut, and was completed at the Institut des Hautes Études Scientifiques. First, the author thanks Tobias Fritz, who explained several aspects of his work with John Baez and Tom Leinster [4] and who provided insight and additional references. The author has benefited from conversations with Jonathan Ben-Benjamin, Lewis Bowen, Tai-Danae Bradley, Brian Dressner, Brian Hall, Azeem ul Hassan, Mark Hillery, Chris Heunen, Manas Kulkarni, Franklin Lee, Jamie Lennox, William Mayer, Philip Parzygnat, George Poppe, Xing Su, Josiah Sugarman, Dennis Sullivan, Steven Vayl, Scott O. Wilson, Cody Youmans, and Lai-Sang Young. The author also thanks Tom for his LaTeX code on frames, which was accessed from https://texblog.org/2015/09/30/fancy-boxes-for-theorem-lemma-and-proof-with-mdframed/ on November 1st, 2018. The author thanks Yung Bae for motivation. Finally, and most importantly, the author is especially thankful to V. P. Nair, who provided several crucial suggestions during the earlier stages of this work. This work was partially supported by NSF grant PHY-1213380 and the Capelloni Dissertation Fellowship.

References

[1] Stéphane Attal, *Lecture 6: Quantum channels*, 2014. Last accessed at http://math.univ-lyon1.fr/~attal/Quantum_Channels.pdf on January 27, 2020. ↑10

[2] John C. Baez, *Category-theoretic characterizations of entropy III*, 2011. Blog post available at https://golem.ph.utexas.edu/category/2011/06/categorytheoretic_characteriza_2.html (last accessed on June 18, 2020). ↑2, 30

[3] John C. Baez and Tobias Fritz, *A Bayesian characterization of relative entropy*, Theory Appl. Categ. **29** (2014), No. 16, 422–457, available at 1402.3067. ↑2, 29, 32, 49

[4] John C. Baez, Tobias Fritz, and Tom Leinster, *A characterization of entropy in terms of information loss*, Entropy **13** (2011), no. 11, 1945–1957, available at 1106.1791. ↑2, 9, 15, 22, 25, 27, 28, 30, 31, 49

[5] Bernhard Baumgartner, *Characterizing entropy in statistical physics and in quantum information theory*, Foundations of Physics **44** (2014), no. 10, 1107–1123. ↑42

[6] Kenta Cho and Bart Jacobs, *Disintegration and Bayesian inversion via string diagrams*, Math. Struct. Comp. Sci. (2019), 1–34, available at 1709.00322. ↑49

[7] Imre Csiszár, *Axiomatic characterizations of information measures*, Entropy **10** (2008), no. 3, 261–273. ↑41

[8] D. K. Faddeev, *On the concept of entropy of a finite probabilistic scheme*, Uspekhi Matematicheskikh Nauk **11** (1956), no. 1, 227–231. English translation by Arina Zinovyeva available at https://arrowtheory.com/pub/notes/025-faddeev-entropy.html. ↑41

[9] ______, *Zum begriff der entropie einer endlichen wahrscheinlichkeitsschemas*, Arbeiten zur Informationstheorie I. Deutscher Verlag der Wissenschaften (1957), 85–90. ↑41

[10] Mark Fannes, *A continuity property of the entropy density for spin lattice systems*, Comm. Math. Phys. **31** (1973), no. 4, 291–294. ↑28

[11] Douglas R. Farenick, *Algebras of linear transformations*, Universitext, Springer-Verlag, New York, 2001. ↑6, 11

[12] Peter A. Fillmore, *A user's guide to operator algebras*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, Inc., New York, 1996. A Wiley-Interscience Publication. ↑11

[13] Joe Flood, *Semiconvex geometry*, J. Austral. Math. Soc. Ser. A **30** (1980/81), no. 4, 496–510. ↑22

[14] Tobias Fritz, *Convex spaces I: Definition and examples*, 2009. arXiv preprint: 0903.5522 [math.MG]. ↑22

[15] ______, *A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics*, Adv. Math. **370** (2020), 107239, available at 1908.07021. ↑49

[16] Nicolas Gagné and Prakash Panangaden, *A categorical characterization of relative entropy on standard Borel spaces*, The Thirty-third Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIII), 2018, pp. 135–153. ↑29, 49

[17] Mikhail Gromov, *Probability, symmetry, linearity*, Institut des Hautes Études Scientifiques, 2014. Lecture recording available at https://www.youtube.com/watch?v=aJAQVletzdY. ↑34

[18] Stanley P. Gudder, *Convex structures and operational quantum mechanics*, Comm. Math. Phys. **29** (1973), no. 3, 249–264. ↑21, 22

[19] ———, *A general theory of convexity*, Rendiconti del Seminario Matematico e Fisico di Milano **49** (1979), no. 1, 89–96. ↑22

[20] Yonatan Harpaz, *(co)Cartesian fibrations*, 2016. Notes available at: https://www.math.univ-paris13.fr/~harpaz/#notes (last accessed on July 17, 2020). ↑17

[21] Michal Horodecki, Jonathan Oppenheim, and Andreas Winter, *Partial quantum information*, Nature **436** (2005), no. 7051, 673. ↑2

[22] ———, *Quantum state merging and negative information*, Comm. Math. Phys. **269** (2007), no. 1, 107–136. ↑2, 16

[23] Ryszard Horodecki and Pawel Horodecki, *Quantum redundancies and local realism* **194** (1994), no. 3, 147–152. ↑16

[24] Roman S. Ingarden and Andrzej Kossakowski, *An axiomatic definition of information in quantum mechanics*, Bulletin L'Académie Polonaise des Sciences **XVI** (1968), no. 1, 61–65. ↑41

[25] Tom Leinster, *A short characterization of relative entropy*, J. Math. Phys. **60** (2019), no. 2, 023302. ↑49

[26] Elliott H. Lieb and Mary Beth Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, J. Math. Phys. **14** (1973), 1938–1941. ↑43

[27] Göran Lindblad, *An entropy inequality for quantum measurements*, Comm. Math. Phys. **28** (1972), no. 3, 245–249. ↑14

[28] ———, *Entropy, information and quantum measurements*, Comm. Math. Phys. **33** (1973), no. 4, 305–322. ↑14, 43, 48

[29] ———, *Completely positive maps and entropy inequalities*, Comm. Math. Phys. **40** (1975), no. 2, 147–151. ↑48

[30] Fosco Loregian and Emily Riehl, *Categorical notions of fibration*, Expo. Math. (2019), available at 1806.06129. ↑17

[31] Keiji Matsumoto, *Reverse test and characterization of quantum relative entropy*, 2010. arXiv preprint: 1010.1030 [quant-ph]. ↑49

[32] Joe Moeller and Christina Vasilakopoulou, *Monoidal Grothendieck construction*, Theory Appl. Categ. **35** (2020), no. 31, 1159–1207, available at 1809.00727. ↑17, 23, 24, 29

[33] Walter D. Neumann, *On the quasivariety of convex subsets of affine spaces*, Arch. Math. **21** (1970), no. 1, 11–16. ↑22

[34] Michael A. Nielsen and Isaac L. Chuang, *Quantum computation and quantum information*, 10th Anniversary Edition, Cambridge University Press, New York, NY, USA, 2011. ↑10, 12, 14, 27, 28

[35] Wilhelm Ochs, *A new axiomatic characterization of the von Neumann entropy*, Rep. Math. Phys. **8** (1975), no. 1, 109–120. ↑41

[36] Arthur J. Parzygnat, *Some 2-categorical aspects in physics*, Ph.D. Thesis, 2016. Ph.D. Thesis *CUNY Academic Works*. ↑20, 29, 34, 42, 46

[37] ———, *Discrete probabilistic and algebraic dynamics: a stochastic Gelfand–Naimark theorem*, 2017. arXiv preprint: 1708.00091 [math.FA]. ↑6, 10

[38] ———, *From observables and states to Hilbert space and back: A 2-categorical adjunction*, Appl. Categ. Structures **26** (2018), 1123–1157, available at 1609.08975. ↑42

[39] ———, *Stinespring's construction as an adjunction*, Compositionality **1** (2019), available at 1807.02533. ↑42

[40] ———, *Inverses, disintegrations, and Bayesian inversion in quantum Markov categories*, IHÉS, 2020. arXiv preprint: 2001.08375 [quant-ph]. ↑49

[41] ———, *Jeffrey conditioning and Bayesian inference in quantum mechanics*, 2020. In preparation. ↑14

[42] Arthur J. Parzygnat and Benjamin P. Russo, *Non-commutative disintegrations: existence and uniqueness in finite dimensions*, 2019. arXiv preprint: 1907.09689 [quant-ph]. ↑2, 4, 7, 9, 10, 29, 34, 40, 41

[43] ———, *A non-commutative Bayes' theorem*, 2020. arXiv preprint: 2005.03886 [quant-ph]. ↑7, 11, 47

[44] Paolo Perrone, *Categorical probability and stochastic dominance in metric spaces*, Ph.D. Thesis, 2018. Available as `https://paoloperrone.org/phdthesis.pdf` (last accessed on June 20, 2020). ↑46

[45] Dénes Petz, *Characterization of the relative entropy of states of matrix algebras*, Acta Math. Hung. **59** (1992), no. 3-4, 449–455. ↑49

[46] ———, *Entropy, von Neumann and the von Neumann entropy*, John von Neumann and the foundations of quantum physics, 2001, pp. 83–96. ↑7, 8

[47] Alfréd Rényi, *On measures of entropy and information*, Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I, 1961, pp. 547–561. ↑41

[48] Irving E. Segal, *A note on the concept of entropy*, J. Math. Mech. **9** (1960), no. 4, 623–629. ↑13

[49] Marshall Harvey Stone, *Postulates for the barycentric calculus*, Annali di Matematica **29** (1949), no. 1, 25–30. ↑22

[50] Tadeusz Świrszcz, *Monadic functors and convexity*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **22** (1974), 39–42. ↑22

[51] Walter E. Thirring, *Vorlesungen über mathematische physik. 8. quantenmechanik großer systeme*, Inst. für Theoret. Physik d. Univ. Wien, 1975. ↑41

[52] ———, *A course in mathematical physics: Volume 4: Quantum mechanics of large systems*, Springer-Verlag Wien, 1983. ↑42

[53] Hisaharu Umegaki, *Conditional expectation in an operator algebra. IV. entropy and information*, Kodai Math. Sem. Rep. **14** (1962), no. 2, 59–85. ↑46

[54] A. Vershynina, E. A Carlen, and E. H. Lieb, *Strong Subadditivity of Quantum Entropy*, Scholarpedia **8** (2013), no. 4, 30920. revision #169781. ↑3

[55] John von Neumann, *Mathematical foundations of quantum mechanics: New edition*, Princeton University Press, 2018. ↑41

[56] John von Neumann and Oskar Morgenstern, *Theory of games and economic behavior*, Princeton University Press, 2007. ↑22

[57] Alfred Wehrl, *General properties of entropy*, Rev. Mod. Phys. **50** (1978), 221–260. ↑42

[58] Reinhard Werner, *Mathematical methods of quantum information theory*, Leibniz Universitt Hannover, 2017. Lecture recordings available at `https://www.youtube.com/playlist?list=PLDfPUNusx1EoBAn8vXYjcF95R7mI_eR6o`. ↑10, 15

[59] Henrik Wilming, Rodrigo Gallego, and Jens Eisert, *Axiomatic characterization of the quantum relative entropy and free energy*, Entropy **19** (2017), no. 6, 241. ↑49

[60] Edward Witten, *APS medal for exceptional achievement in research: Invited article on entanglement properties of quantum field theory*, Rev. Mod. Phys. **90** (2018), 045003. ↑7

A. Parzygnat, INSTITUT DES HAUTES ÉTUDES SCIENTIFIQUES, 35 ROUTE DE CHARTRES 91440, BURES-SUR-YVETTE, FRANCE

E-mail address, A. Parzygnat: `parzygnat@ihes.fr`