Available online at www.sciencedirect.com
SciVerse ScienceDirect
Advances in Mathematics 234 (2013) 667–696
ADVANCES IN Mathematics
www.elsevier.com/locate/aim

Concavification of free entropy

Philippe Biane[a,*], Yoann Dabrowski[b]

[a] CNRS, IGM, Université Paris-Est, Champs-sur-Marne, France
[b] Université de Lyon, Université Lyon 1, Institut Camille Jordan UMR 5208, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France

Received 5 March 2012; accepted 6 November 2012
Available online 13 December 2012

Communicated by Dan Voiculescu

Abstract

We introduce a modification of Voiculescu's free entropy which coincides with the lim inf variant of Voiculescu's free entropy on extremal states, but is a concave upper semi-continuous function on the trace state space. We also extend the orbital free entropy of Hiai et al. (2009) [8] to non-hyperfinite multivariables and prove freeness in the case of additivity of Voiculescu's entropy (or vanishing of our extended orbital entropy).

MSC: primary 46L54; secondary 60G15; 94A17

Keywords: Free entropy; Orbital entropy; Additivity of entropy; Poulsen simplex; Concentration of measure; Csiszar's projection

1. Introduction

Voiculescu has introduced a free entropy quantity, for tracial states on a von Neumann algebra generated by n self-adjoint elements, which has been very useful for the solution of many long standing open problems in von Neumann algebra theory. It turns out that free entropy satisfies an unusual property for an entropy quantity which is a "degenerate convexity" property, i.e. the entropy of any nonextremal state is $-\infty$, which is in sharp contrast with the usual concavity and

* Corresponding author.
E-mail addresses: biane@univ-mlv.fr (P. Biane), dabrowski@math.univ-lyon1.fr (Y. Dabrowski).

0001-8708/$ - see front matter
doi:10.1016/j.aim.2012.11.003

upper semi-continuity property of classical entropy. Recently Hiai [7] defined a free analogue of pressure and considered its Legendre transform. He obtained a quantity which is concave and upper semi-continuous, and majorizes Voiculescu's free entropy. It is not clear whether this quantity coincides with Voiculescu's free entropy on extremal states. In this paper we introduce a modified definition, through random matrix approximations, which yields a quantity which is both concave upper semi-continuous, and coincides with the lim inf variant of Voiculescu's free entropy on extremal states. Our main argument is the simple observation that a probability measure on a compact convex set, whose barycenter is close to an extremal point, has most of its mass concentrated near this point (see Lemma 6.1 below). This is obvious in finite dimension, but requires further clarification in infinite dimension. In this paper we rely on the fact that the convex set we consider is a Poulsen simplex.

We use an analogous idea to generalize the definition of free orbital entropy, due to Hiai, Miyamoto and Ueda [8]. In this paper, the authors introduced, via a microstates approach, an entropy quantity $\chi_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n})$, where each $\mathbf{X_i}$ is a finite set of noncommutative random variables generating a hyperfinite algebra. They used this quantity to generalize Voiculescu's additivity result [20], namely: for noncommutative random variables $X_1, \ldots, X_n$, if

$$\chi(X_1, \ldots, X_n) = \chi(X_1) + \cdots + \chi(X_n)$$

and these quantities are finite, then the X_i are free. More generally, they showed that $\chi_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n}) = 0$ is equivalent to freeness in the hyperfinite context above even though the finiteness of entropy fails in general in this case. They recover the previous result since they also show:

$$\chi(X_1, \ldots, X_n) = \chi_{orb}(X_1, \ldots, X_n) + \chi(X_1) + \cdots + \chi(X_n),$$

in case these quantities are finite.

In Section 7, we introduce a definition of $\chi_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n})$, for arbitrary finite sets $\mathbf{X_i}$ of noncommutative random variables, obtained by replacing microstates by probability measures. We show that many of the arguments of [8] have analogues in this setting, and we obtain the full generalization of the additivity result when random variables X_i are replaced by arbitrary finite sets $\mathbf{X_i}$.

This paper is organized as follows. We start by recalling some well known facts on trace states and on Legendre transform and classical entropy (including Csiszar's projections result) in Sections 2 and 3. Then we prove the main result about concavification in Sections 4 and 6, after a few preliminaries about Poulsen simplices in Section 5. In Section 7 we extend the definition of orbital entropy, and prove freeness in the case of additivity of Voiculescu's entropy, in Corollary 7.4. Finally, after a few more preliminaries in Section 8, Section 9 is devoted to some further variants and extensions of our definitions, which might prove useful for future applications.

2. The set of trace states

Let $\mathbf{C}\langle X_1, \ldots, X_n\rangle$ be the free $*$-algebra with unit generated by $n \geq 1$ self-adjoint elements $X_1, \ldots, X_n$, which we identify with the space of noncommutative polynomials in the indeterminates $X_1, \ldots, X_n$. We consider the set $\mathcal{S}_c^n$ of trace states on $\mathbf{C}\langle X_1, \ldots, X_n\rangle$. This set consists in all positive, tracial $*$-linear maps $\tau : \mathbf{C}\langle X_1, \ldots, X_n\rangle \to \mathbf{C}$ such that $\tau(1) = 1$ and, for any $P \in \mathbf{C}\langle X_1, \ldots, X_n\rangle$ there exists some constant $R_P > 0$ such that

$$\tau((P^*P)^k) \leq R_P^{2k} \quad \text{for } k \geq 0 \tag{2.1}$$

Let us denote by $\mathcal{S}_R^n$ the set of all trace states such that $\max(R_{X_1}, \dots, R_{X_n}) \leq R$.

Especially, for $R \geq T$ we have $\mathcal{S}_R^n \supset \mathcal{S}_T^n$. Moreover, $\mathcal{S}_c^n = \cup_{R \geq 0} S_R^n$. Finally for $\tau \in \mathcal{S}_c^n$, we define $\mathcal{R}(\tau) = \inf\{R, \tau \in S_R^n\}$ so that obviously $\tau \in S_{\mathcal{R}(\tau)}^n$.

The set $\mathcal{S}_R^n$ can be identified with the set of trace states on the free product C^*-algebra $*_{i=1}^n C([-R, R])$, cf [7]. It is a compact convex set for the weak* topology. By the reduction theory for von Neumann algebras, it is a Choquet simplex, and its extreme points (for $n \geq 2$) are the factor states [16]. Note that, as a consequence, an extreme point in $\mathcal{S}_R^n$ is still an extreme point in $\mathcal{S}_T^n$ for $T \geq R$. Moreover, the second author proved in [5, Corollary 5] that, for $n > 1$, $\mathcal{S}_R^n$ is a Poulsen Simplex, i.e. the unique metrizable Choquet simplex with a dense set of extreme points (cf. [12]). If $\mathcal{A}$ is a von Neumann algebra equipped with a tracial state φ, and $(X_1, \dots, X_n) \in \mathcal{A}$ an n-tuple such that $\sup_i \|X_i\| \leq R$, one defines a state $\tau_{X_1,\dots,X_n} \in \mathcal{S}_R^n$ by the formula

$$\tau_{X_1,\dots,X_n}(P) = \varphi(P(X_1, \dots, X_n)).$$

In particular, if $\mathcal{A} = M_N(\mathbf{C})$ and $\varphi = \frac{1}{N}\mathrm{Tr}$ the normalized trace, we denote by H_N^R the set of hermitian matrices of size N, whose operator norm is less than R, then an n-tuple $(M_1, \dots, M_n) \in (H_N^R)^n$ defines a state $\tau_{M_1,\dots,M_n} \in \mathcal{S}_R^n$, by

$$\tau_{M_1,\dots,M_n}(P) = \frac{1}{N}\mathrm{Tr}(P(M_1, \dots, M_n)).$$

Similarly, a probability measure μ on $(H_N^R)^n$ (always assumed Borel) defines a random state in $\mathcal{S}_R^n$, whose barycenter τ_μ, defined by

$$\tau_\mu(P) = \int_{(H_N^R)^n} \frac{1}{N}\mathrm{Tr}(P(M_1, \dots, M_n))d\mu(M_1, \dots, M_n),$$

is again an element of $\mathcal{S}_R^n$.

For $\tau \in \mathcal{S}_R^n$, let $V_{\epsilon,K}(\tau)$ be the set of states $\sigma \in \mathcal{S}_R^n$ such that, for all monomials m of degree less than K, we have:

$$|\tau\,(m(X_1, \dots, X_n)) - \sigma\,(m(X_1, \dots, X_n))| < \epsilon.$$

The sets $(V_{\epsilon,K}(\tau); \epsilon, K > 0)$ form a basis of neighborhoods of τ in the weak* topology.

3. Classical entropy, its Legendre transform and Csiszar's projection

Recall that the entropy of a probability measure μ on $\mathbf{R^p}$ is the quantity

$$\mathrm{Ent}(\mu) = \begin{cases} -\displaystyle\int_{\mathbf{R^p}} f(x)\log f(x)dx & \text{if } \mu(dx) = f(x)dx, \log(f) \in L^1(\mu) \\ -\infty & \text{otherwise.} \end{cases}$$

The entropy is a concave upper semi-continuous function of μ.

Moreover, there is also a well known notion of relative entropy of two probability measures (also called Kullback–Leibler divergence, cf. [11]).

$$\mathrm{Ent}(\mu|\nu) = \begin{cases} -\displaystyle\int_{\mathbf{R^p}} f(x)\log f(x)d\nu(x) & \text{if } \mu(dx) = f(x)d\nu(x), \\ -\infty & \text{if } \mu \text{ is not absolutely continuous with respect to } \nu. \end{cases}$$

Note that, by Jensen inequality, $\mathrm{Ent}(\mu|\nu) \le 0$. The relative entropy satisfies the following key property: For any measurable map T, if $T_*\mu$ is the pushforward measure of μ, we have (cf. [11, Chapter 2 Theorem 4.1]):

$$\mathrm{Ent}(T_*\mu|T_*\nu) \ge \mathrm{Ent}(\mu|\nu). \tag{3.1}$$

If $E \subset \mathbf{R}^{\mathbf{p}}$ is a subset with positive Lebesgue measure, and μ is the normalized Lebesgue measure on E, then

$$\mathrm{Ent}(\mu) = \log(\mathrm{Leb}(E)).$$

Actually this is the maximum value of Ent on the set of all probability measures supported by E. Analogously, if μ is the restriction of ν to E, renormalized into a probability measure, then

$$\mathrm{Ent}(\mu|\nu) = \log(\nu(E))$$

and again this is the maximum value of $\mathrm{Ent}(\cdot|\nu)$ on the set of all probability measures supported by E. From this we deduce the following estimates.

Lemma 3.1. *Let μ be supported by E and $F \subset E$ a measurable subset, then*

$$\begin{aligned}\mathrm{Ent}(\mu) \le\ & \mu(F)\log Leb(F) + \mu(E\setminus F)\log Leb(E\setminus F)\\ & - \mu(F)\log\mu(F) - (1-\mu(F))\log(1-\mu(F))\end{aligned} \tag{3.2}$$

and

$$\begin{aligned}\mathrm{Ent}(\mu|\nu) \le\ & \mu(F)\log\nu(F) + \mu(E\setminus F)\log\nu(E\setminus F)\\ & - \mu(F)\log\mu(F) - (1-\mu(F))\log(1-\mu(F)).\end{aligned} \tag{3.3}$$

Proof.

$$\begin{aligned}\mathrm{Ent}(\mu) &= -\int_F f(x)\log f(x)dx - \int_{E\setminus F} f(x)\log f(x)dx\\ &= -\mu(F)\int_F \frac{f(x)}{\mu(F)}\log\frac{f(x)}{\mu(F)}dx - \mu(E\setminus F)\int_{E\setminus F}\frac{f(x)}{\mu(E\setminus F)}\log\frac{f(x)}{\mu(E\setminus F)}dx\\ &\quad - \mu(F)\log\mu(F) - \mu(E\setminus F)\log\mu(E\setminus F)\\ &\le \mu(F)\log\mathrm{Leb(F)} + \mu(E\setminus F)\log\mathrm{Leb}(E\setminus F)\\ &\quad - \mu(F)\log\mu(F) - (1-\mu(F))\log(1-\mu(F)).\end{aligned}$$

The proof of the other inequality is similar (cf. [11, Chapter 2 Corollary 3.2]). □

We shall need another characterization of entropy, through its Legendre transform. Indeed we have, for any probability measure μ supported by a set E, of finite Lebesgue measure,

$$\mathrm{Ent}(\mu) = \inf_{\phi\in C_b(E)}\left(\log\left(\int_E \exp\phi(x)dx\right) - \int_E \phi(x)\mu(dx)\right)$$

where $C_b(E)$ is the space of bounded, real valued continuous functions on E. Likewise (see e.g. [6, Section 6.2]) for any probability measures μ, ν supported on E,

$$\mathrm{Ent}(\mu|\nu) = \inf_{\phi\in C_b(E)}\left(\log\left(\int_E \exp\phi(x)d\nu(x)\right) - \int_E \phi(x)\mu(dx)\right). \tag{3.4}$$

It follows that if $f_1, \ldots, f_p$ are real valued bounded measurable functions on E, then we have

$$\inf_{\lambda \in \mathbb{R}^p} \left(\log \int_E e^{\sum_i \lambda_i f_i(x)} dx - \sum_i a_i \lambda_i \right)$$
$$= \sup \left\{ \mathrm{Ent}(\mu) \mid \mu \text{ supported on } E;\ \int f_i(x)\mu(dx) = a_i, i = 1, \ldots, p \right\} \tag{3.5}$$

where the sup is defined as $-\infty$ if there is no such probability measure.

We will apply these considerations to the case where the set E is a product of balls H_N^R, i.e. balls of radius R for the operator norm in the space of $N \times N$ hermitian matrices, with Lebesgue measure, and the functions $f_1, \ldots, f_p$ are traces of selfadjoint polynomials in noncommuting indeterminates, of the form

$$f(M_1, \ldots, M_n) = N\mathrm{Tr}(P(M_1, \ldots, M_n)).$$

Let us define

$$I_N(P) = \int_{(H_N^R)^n} e^{-N\mathrm{Tr}(P(M_1,\ldots,M_n))} dM_1 \ldots dM_n,$$

for P a self-adjoint element of $\mathbf{C}\langle \mathbf{X_1}, \ldots, \mathbf{X_n} \rangle$.

Definition 3.2. For $\tau \in \mathcal{S}_R^n$, we define $\rho_{N,K}(\tau)$ as the maximum of the entropy of (Borel) probability measures μ on $(H_N^R)^n$ whose barycenter coincides with τ on monomials of degree less than K, and $\rho_{N,K}(\tau) = -\infty$ if there is no such measure. Equivalently, if $P[(H_N^R)^n]$ is the above set of Borel probability measures, we have:

$$\rho_{N,K}(\tau) = \sup_{\substack{\mu \in P[(H_N^R)^n] \\ \tau_\mu \in \cap_{\epsilon>0} V_{\epsilon,K}(\tau)}} \mathrm{Ent}(\mu).$$

We have, by (3.5):

$$\rho_{N,K}(\sigma) = \inf_{\substack{P \in \mathbf{C}\langle X_1,\ldots,X_n \rangle \\ P=P^*, \deg(P) \leq K}} \left(\log I_N(P) + N^2 \sigma(P) \right), \tag{3.6}$$

which is therefore a concave upper semi-continuous function of σ.

Even though we will not need it before Section 9, it may be enlightening to use the language of Csiszar's I-projections (cf. [4], see also [13, Chapter 10] for an exposition). Let us recall the basics. Let $\mathcal{E}$ be a closed convex set of probability distributions then, by the strict concavity of relative entropy, there exists a unique probability measure realizing $\sup_{\mu \in \mathcal{E}} \mathrm{Ent}(\mu|\nu)$. This probability distribution, denoted C, is called Csiszar's I-projection of the probability distribution ν on the convex set $\mathcal{E}$. Csiszar [4] first proved its existence when $\mathcal{E}$ is variation closed and contains a μ with $\mathrm{Ent}(\mu|\nu) > -\infty$. Moreover C is characterized by:

$$\mathrm{Ent}(\mu|\nu) \leq \mathrm{Ent}(\mu|C) + \mathrm{Ent}(C|\nu),$$

for every $\mu \in \mathcal{E}$. We can infer from this that $\rho_{N,K}(\tau)$, if finite, is the entropy of Csiszar's I-projection $C_{N,0,K}(\tau)$ of normalized Lebesgue measure (on $(H_N^R)^n$) on the set of measures whose mean agrees with τ on monomials of order less than K. It is a well known result about exponential families (see e.g. [4, Theorem 3.1] or [13, Theorem 10.2]) that $C_{N,0,K}(\tau)$ has a density with respect to normalized Lebesgue measure on $(H_N^R)^n$ of the form $\frac{1}{Z}e^{-\mathrm{Tr}(V(X))}$ for a non commutative polynomial V of degree less than K. Especially, $\rho_{N,K}(\tau)$ is the entropy of a well-studied unitary invariant random matrix model.

4. Voiculescu's free entropy and its modification

Let $\tau \in \mathcal{S}_R^n$, let $\epsilon > 0$ be a real number and K, N be positive integers. We denote by $\Gamma_R(\tau, \epsilon, K, N)$ the set of n-tuples of hermitian matrices $M_1, \ldots, M_n \in H_N^R$ such that for all monomials $m(X_1, \ldots, X_n) = X_{i_1} \ldots X_{i_k}$ of degree less than K we have:

$$\left|\tau\left(m(X_1, \ldots, X_n)\right) - \frac{1}{N}\mathrm{Tr}\left(m(M_1, \ldots, M_n)\right)\right| < \epsilon.$$

Equivalently $\Gamma_R(\tau, \epsilon, K, N)$ is the set of n-tuples of hermitian matrices $M_1, \ldots, M_n \in H_N^R$ whose associated state $\tau_{M_1,\ldots,M_n}$ is in $V_{\epsilon,K}(\tau)$.

Definition 4.1 (*[19]*)**.** Define for $\tau \in \mathcal{S}_R^n$:

$$\chi_R(\tau) = \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty} \left(\frac{1}{N^2}\log\left(\mathrm{Leb}(\Gamma_R(\tau, \epsilon, K, N))\right) + \frac{n}{2}\log N\right).$$

The free entropy of a tracial state $\tau \in \mathcal{S}_c^n$ is:

$$\chi(\tau) = \sup_{R \geq \mathcal{R}(\tau)} \chi_R(\tau).$$

It is known that, if τ is not an extreme point of $\mathcal{S}_R^n$, then $\chi(\tau) = -\infty$, cf [18]. Furthermore, if τ is considered as a state in $\mathcal{S}_{R'}^n$ for some $R' > R > \mathcal{R}(\tau)$ then $\chi_{R'}(\tau) = \chi_R(\tau)$. Since it is not known whether the lim sup in the definition is a limit, it has been useful to define:

$$\underline{\chi}_R(\tau) = \lim_{K\to\infty,\epsilon\to 0} \liminf_{N\to\infty} \left(\frac{1}{N^2}\log\left(\mathrm{Leb}(\Gamma_R(\tau, \epsilon, K, N))\right) + \frac{n}{2}\log N\right)$$

and, for a nontrivial ultrafilter ω on $\mathbf{N}$:

$$\chi_R^\omega(\tau) = \lim_{K\to\infty,\epsilon\to 0} \lim_{N\to\omega} \left(\frac{1}{N^2}\log\left(\mathrm{Leb}(\Gamma_R(\tau, \epsilon, K, N))\right) + \frac{n}{2}\log N\right).$$

In [21], a state τ for which these limits coincide is called *regular*.

We are now going to concavify the previous definition in the following way.

Definition 4.2. We define the concavified free entropy of a tracial state $\tau \in \mathcal{S}_R^n$ by:

$$\underline{\tilde{\chi}}_R(\tau) = \lim_{K\to\infty,\epsilon\to 0} \liminf_{N\to\infty} \left(\frac{1}{N^2}\left[\sup_{\sigma \in V_{\epsilon,K}(\tau)} \rho_{N,K}(\sigma)\right] + \frac{n}{2}\log N\right),$$

and likewise $\tilde{\chi}_R(\tau)$ with a lim sup and $\tilde{\chi}_R^\omega(\tau)$ with a limit to ω.

Finally, we put for $\tau \in \mathcal{S}_c^n$:

$$\underline{\tilde{\chi}}(\tau) = \sup_{R \geq \mathcal{R}(\tau)} \underline{\tilde{\chi}}_R(\tau)$$

and likewise for $\tilde{\chi}(\tau)$, $\tilde{\chi}^\omega(\tau)$.

We thus have, as for Voiculescu's free entropy, three variants, but we do not know whether they all coincide. Note that, since $\{\mu \in P[(H_N^R)^n] : \tau_\mu \in V_{\epsilon,K}(\tau)\} = \cup_{\sigma \in V_{\epsilon,K}(\tau)} \{\mu \in P[(H_N^R)^n] : \tau_\mu \in \cap_{\eta>0} V_{\eta,K}(\sigma)\}$, we have the alternative formula:

$$\underline{\tilde{\chi}}_R(\tau) = \lim_{K\to\infty,\epsilon\to 0} \liminf_{N\to\infty} \left(\frac{1}{N^2} \left[\sup_{\substack{\mu \in P[(H_N^R)^n] \\ \tau_\mu \in V_{\epsilon,K}(\tau)}} \mathrm{Ent}(\mu) \right] + \frac{n}{2} \log N \right).$$

We have the fundamental properties:

Proposition 4.3. *The quantity $\underline{\tilde{\chi}}_R(\tau)$ is a concave upper semi-continuous function of τ. So is $\tilde{\chi}_R^\omega(\tau)$. Furthermore, we have:*

$$\underline{\tilde{\chi}}_R(\tau) \geq \underline{\chi}_R(\tau), \qquad \tilde{\chi}_R(\tau) \geq \chi_R(\tau), \qquad \tilde{\chi}_R^\omega(\tau) \geq \chi_R^\omega(\tau),$$

and $\tilde{\chi}_R, \tilde{\chi}_R^\omega$ are subadditive: if τ_1, τ_2 are the marginal states giving the noncommutative distributions of $X_1, \dots, X_m$ and $X_{m+1}, \dots, X_n$ respectively, then

$$\tilde{\chi}_R(\tau) \leq \tilde{\chi}_R(\tau_1) + \tilde{\chi}_R(\tau_2), \qquad \tilde{\chi}_R^\omega(\tau) \leq \tilde{\chi}_R^\omega(\tau_1) + \tilde{\chi}_R^\omega(\tau_2).$$

Proof. According to (3.6), we have:

$$\rho_{N,K}(\sigma) = \inf_{\substack{P \in \mathbb{C}_{sa}\langle X_1,\dots,X_n\rangle \\ \deg(P)\leq K}} \left(\log I_N(P) + N^2 \sigma(P) \right)$$

which is therefore a concave upper semi-continuous function of σ. Let τ_1 and τ_2 be states, and let $\sigma_1 \in V_{\epsilon,K}(\tau_1), \sigma_2 \in V_{\epsilon,K}(\tau_2)$, then

$$\lambda\sigma_1 + (1-\lambda)\sigma_2 \in V_{\epsilon,K}(\lambda\tau_1 + (1-\lambda)\tau_2)$$

therefore by concavity,

$$\sup_{\sigma \in V_\epsilon(\lambda\tau_1+(1-\lambda)\tau_2)} \rho_{N,K}(\sigma) \geq \lambda\rho_{N,K}(\sigma_1) + (1-\lambda)\rho_{N,K}(\sigma_2).$$

Since this is true for all σ_1, σ_2 we get:

$$\sup_{\sigma \in V_\epsilon(\lambda\tau_1+(1-\lambda)\tau_2)} \rho_{N,K}(\sigma) \geq \lambda \sup_{\sigma_1 \in V_\epsilon(\tau_1)} \rho_{N,K}(\sigma_1) + (1-\lambda) \sup_{\sigma_2 \in V_\epsilon(\tau_2)} \rho_{N,K}(\sigma_2).$$

The reader may have noted this is also a consequence of the expression of the left hand side as the entropy of Csiszar's I-projection on the set of measures having mean in $V_{\epsilon,K}(\lambda\tau_1+(1-\lambda)\tau_2)$. Thus $\sup_{\sigma\in V_{\epsilon,K}(\tau)} \rho_{N,K}(\sigma)$ is a concave function of τ, and taking a liminf we see that:

$$\liminf_{N\to\infty} \left(\frac{1}{N^2} \left[\sup_{\sigma \in V_{\epsilon,K}(\tau)} \rho_{N,K}(\sigma) \right] + \frac{n}{2} \log N \right)$$

is again concave in τ.

It is easy to check that taking the limit as ϵ goes to zero gives an upper semi-continuous function. Since it is nonincreasing in K, the limit as $K \to \infty$ is again concave and upper semi-continuous.

Subadditivity follows from the subadditivity of classical entropy. Note that we cannot deduce it for the lim inf variant, since in general the inequality $\liminf(a_n+b_n) \leq \liminf(a_n)+\liminf(b_n)$

fails. Of course if all variants of the free entropy actually coincide, subadditivity would follow in this case. □

Remark 4.4. We notice that the state of maximal $\tilde{\chi}$ entropy in $\mathcal{S}_R^n$ is the distribution of a free family of arc-sine distributed self-adjoint operators, where the arcsine distribution is on $[-R, R]$. It corresponds to taking the limit of barycenters of normalized Lebesgue measure on $(H_N^R)^n$. In particular, this quantity is finite. (The reader may also be referred to [9, Section 5.6] for this finiteness.)

Remark 4.5. As the referee reminded us, Voiculescu suggested in [19, section 7.1] several alternative definitions of free entropy. We discuss here the relation with our definition. The first variant $\chi^{(1)}(\tau)$ has been studied in [1] and the second variant $\chi^{(2)}(\tau)$ happens to be by definition exactly our $\tilde{\chi}(\tau)$. The first part of this paper may thus be seen as a study of this suggestion of Voiculescu. Recall the definition:

$$\chi^{(1)}(\tau) = \sup_{R\geq\mathcal{R}(\tau)} \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty} \left(\frac{1}{N^2} \left[\sup_{\substack{\mu\in P[(H_N^R)^n] \\ E_\mu\left(\left|\frac{1}{N}\mathrm{Tr}(P)-\tau(P)\right|\right)<\epsilon \\ \forall P \text{ monomial, } deg(P)\leq K}} \mathrm{Ent}(\mu) \right] + \frac{n}{2}\log N \right).$$

In [1], Belinschi proved $\chi^{(1)}(\tau) = \chi(\tau)$. for any $\tau \in \mathcal{S}_c^n$. We want to point out that the nonlinearity of the condition in $\frac{1}{N}\mathrm{Tr}(P)$ under law μ is the key why this equality is valid here (as in Hiai's second variant of entropy [7, Section 6]). In the variant $\tilde{\chi}(\tau)$ we only have a condition on τ_μ, and this allows us to get a concavification,; this is also what makes it harder to prove equality with $\chi(\tau)$ in the factorial case.

We may also compare our definition with the quantity obtained by [7] using the Legendre transform of free pressure. Define, for $P = P^* \in \mathbf{C}\langle X_1, \ldots, X_n\rangle$:

$$\pi_R(P) = \limsup_{N\to\infty} \frac{1}{N^2}\log I_N(P) + \frac{n}{2}\log N.$$

Hiai defines the entropy by:

$$\eta_R(\tau) = \inf_{P=P^*\in\mathbf{C}\langle X_1,\ldots,X_n\rangle} \tau(P) + \pi_R(P).$$

By (3.5), for any P monomial of degree less than K, $\sigma \in V_{\epsilon,K}(\tau)$, we have:

$$\frac{1}{N^2}\rho_{N,K}(\sigma) \leq \frac{1}{N^2}\log I_N(P) + \tau(P) + \epsilon.$$

Thus, taking a supremum, a limsup (or liminf), and then the limit in ϵ, K, we get:

$$\tilde{\chi}_R(\tau) \leq \tau(P) + \pi_R(P),$$

so that taking an infimum over P we also get:

$$\tilde{\chi}_R(\tau) \leq \eta_R(\tau).$$

We do not know when there is actually an equality, but in the one variable case ($n = 1$), it is known $\eta_R(\tau) = \chi(\tau)$ and thus $\eta_R(\tau) = \chi(\tau) = \tilde{\chi}(\tau) = \underline{\chi}(\tau)$ for R large enough.

In this article, we mainly study $\underline{\tilde{\chi}}_R(\tau)$ instead of $\underline{\tilde{\chi}}(\tau)$. This is motivated by the following result, really similar to [19, Proposition 2.4].

Proposition 4.6. *Consider $\tau \in \mathcal{S}_c^n$. For any $T > R > \mathcal{R}(\tau)$ we have:*

$$\underline{\tilde{\chi}}_T(\tau) = \underline{\tilde{\chi}}_R(\tau) = \underline{\tilde{\chi}}(\tau), \qquad \tilde{\chi}_T(\tau) = \tilde{\chi}_R(\tau) = \tilde{\chi}(\tau), \qquad \tilde{\chi}_T^{\omega}(\tau) = \tilde{\chi}_R^{\omega}(\tau) = \tilde{\chi}^{\omega}(\tau).$$

Proof. We only prove the lim inf variant, and of course it will suffice to prove for $T > R > \mathcal{R}(\tau)$, $\underline{\tilde{\chi}}_T(\tau) \leq \underline{\tilde{\chi}}_R(\tau)$ (the other inequality is obvious). Let $S = \frac{\mathcal{R}(\tau)+R}{2}$. Define the continuous piecewise linear function $h : [-T, T] \to \mathbb{R}$ by $h(t) = \alpha$ for $t \in [-T, -R] \cup [R, T]$, $h(t) = 1$ for $t \in [-S, S]$, $h(t) = \alpha + (1-\alpha)\frac{t+R}{R-S}$ the linear interpolation for $t \in [-R, -S]$ and $h(t) = \alpha + (1-\alpha)\frac{-t+R}{R-S}$, with $\alpha = \frac{R-S}{2T-(R+S)} < 1$ since $T > R$.

In this way, if we define a continuous increasing function $g : [-T, T] \to [-R, R]$ by $g(t) = -R + \int_{-T}^{t} h(s)ds$ we have $g(T) = R$, $g(t) = t$ for $t \in [-S, S]$ and $g'(t) \in [\alpha, 1]$. Let also $G : (H_N^T)^n \to (H_N^R)^n$ defined by $G(A_1, \ldots, A_n) = (g(A_1), \ldots, g(A_n))$. Especially for a state $\tau \in \mathcal{S}_T^n$, we get a state $G_*\tau \in \mathcal{S}_R^n$, so that $\tau_{G_*\mu} = G_*\tau_\mu$, defined by:

$$(G_*\tau)(P(X_1, \ldots, X_n)) = \tau(P(g(X_1), \ldots, g(X_n))).$$

Fix $\epsilon > 0$, $K \in \mathbb{N}^*$, $\tau \in \mathcal{S}_T^n$, we will choose $\delta_1, \delta_2 > 0$ small enough later. First, as in the proof of [19, Proposition 2.4], we get $0 < \epsilon_1 < \epsilon/2$, $K_1 > K$ such that for any $\sigma \in V_{\epsilon_1,K_1}(\tau) \cap \mathcal{S}_T^n$ (with $E(X_j, B)$ the spectral projection of the self-adjoint element X_j (computed in its GNS representation) on the set $B \subset \mathbb{R}$):

$$\sigma(E(X_j, [-T, -S] \cup [S, T])) \leq \delta_1\delta_2,$$
$$\sigma(|g(X_j) - X_j|) \leq \delta_2.$$

This implies $G_*\sigma \in V_{\epsilon,K}(\tau \cap \mathcal{S}_R^n)$, for δ_2 small enough (e.g. $\delta_2 < \epsilon/2KT^{K-1}$).

Consider $\mu \in P[(H_N^R)^n]$ such that $\tau_\mu \in V_{\epsilon_1,K_1}(\tau)$, we can estimate by Chebyshev's inequality:

$$P_\mu\left(\frac{1}{N}\mathrm{Tr}(E(X_j, [-T, -S] \cup [S, T])) \geq \delta_2\right) \leq \frac{E_\mu\left(\frac{1}{N}\mathrm{Tr}(E(X_j, [-T, -S] \cup [S, T]))\right)}{\delta_2} \leq \delta_1.$$

We can also compute $\frac{dG_*\mu}{dLeb} = \left(\frac{d\mu}{dLeb} \circ G^{-1}\right) \times |\det(Jac(G^{-1}))|$. If we write ∂g the two variable function $\partial g(A, B) = (g(A) - g(B))/(A - B)$, $A \neq B$ extended by $\partial g(B, B) = g'(B)$ on the diagonal, the jacobian of g is given by ∂g applied by functional calculus so that:

$$\mathrm{Ent}(G_*\mu) = \mathrm{Ent}(\mu) + \sum_j E_\mu\left(\frac{1}{2}(\mathrm{Tr} \otimes \mathrm{Tr})(\log|\partial g(X_j \otimes 1, 1 \otimes X_j)|^2)\right).$$

In the proof of [19, Proposition 2.4], Voiculescu showed that, for a matrix $X_j \in (H_N^T)^n$ such that $\frac{1}{N}\mathrm{Tr}(E(X_j, [-T, -S] \cup [S, T])) \leq \delta_2$, the positive determinant of the jacobian of g is bounded below so that:

$$\left|\frac{1}{2}(\mathrm{Tr} \otimes \mathrm{Tr})(\log|\partial g(X_j \otimes 1, 1 \otimes X_j)|^2)\right| \leq (N + N^2 - (N(1-\delta_2))^2)|\log\alpha|.$$

Moreover for any matrix $X_j \in (H_N^R)^n$, we have: $\left|\frac{1}{2}(\mathrm{Tr}\otimes\mathrm{Tr})(\log|\partial g(X_j\otimes 1, 1\otimes X_j)|^2)\right| \leq N^2|\log\alpha|$.

As a consequence, we get:

$$\mathrm{Ent}(G_*\mu) \geq \mathrm{Ent}(\mu) - n(N + N^2(2\delta_2 - \delta_2^2))|\log\alpha| - n\delta_1 N^2|\log\alpha|.$$

Taking suprema and liminf, we get:

$$\begin{aligned}&\liminf_{N\to\infty}\left(\frac{1}{N^2}\left[\sup_{\substack{\nu\in P[(H_N^R)^n]\\ \tau_\nu\in V_{\epsilon,K}(\tau)}}\rho_{N,K}(\sigma)\right] + \frac{n}{2}\log N\right)\\ &\quad\geq \liminf_{N\to\infty}\left(\frac{1}{N^2}\left[\sup_{\substack{\mu\in P[(H_N^T)^n]\\ \tau_\mu\in V_{\epsilon_1,K_1}(\tau)}}\rho_{N,K}(\sigma)\right] + \frac{n}{2}\log N\right)\\ &\qquad + n(2\delta_2 - \delta_2^2)\log\alpha + n\delta_1\log\alpha.\end{aligned}$$

Since δ_1, δ_2 can be made arbitrarily small choosing ϵ_1, K_1, we get the desired inequality. □

5. A preliminary separation result

In order to prove that Voiculescu's entropy coincides with its modification on extremal states, we will need a separation result. We gather here references to the literature. Recall that for K a convex subset of the dual E^* of a complex topological vector space, an $x \in E$ is said to expose f in K if $f \in K$ and $\Re g(x) < \Re f(x)$ for all $g \in K$ other than f. Those f which are so exposed by elements of E are weak-* exposed points of K. We now state a result of Sidney [15] (attributed by Asplund to Bishop in the Banach space case)

Proposition 5.1. *Let E be a separable Fréchet space and K a non-empty convex weak* compact subset of its topological dual E^*. Then K is the weak* closed convex hull of the set of its weak* exposed points (this set is thus non empty).*

Since it is proved in [5] that $\mathcal{S}_R^n$ is a Poulsen simplex, we will use the following result [12] of homogeneity.

Proposition 5.2. *Let S_1 and S_2 be metrizable simplices with $\overline{Ext\ S_i} = S_i$ (i.e. Poulsen simplices), for $i = 1, 2$. Let F_i be a proper closed face of S_i, $i = 1, 2$, and let φ be an affine homeomorphism which maps F_2 onto F_1. Then φ can be extended to an affine homeomorphism which maps S_2 onto S_1.*

Applying those two results, the second to move any extremal point to a weak-* exposed point, which exists via the first result, one easily gets:

Proposition 5.3. *Let E be a separable Fréchet space and K a non-empty convex weak* compact subset of its topological dual E^*, which is a Poulsen simplex. Then any extreme point of K is a weak* exposed point.*

Corollary 5.4. *Let τ be an extremal state in $\mathcal{S}_R^n$, $n > 1$, and $\epsilon > 0$. For any $\eta > 0$, there exists a self adjoint polynomial $Q_\eta \in \mathbb{C}\langle X_1, \ldots, X_n\rangle$ such that for every $\sigma \in \mathcal{S}_R^n$ we have:*

$$\tau(Q_\eta) > \sigma(Q_\eta) - \eta,$$

and for all $\sigma \notin V_{\epsilon,K}(\tau)$ *one has:*

$$\sigma(Q_\eta) < \tau(Q_\eta) - 1.$$

Proof. Take $\eta < 1/2$. Since τ is weak-* exposed, first take Q in $*_{i=1}^n C([-R, R])$ exposing it in $\mathcal{S}_R^n$, one can assume Q self adjoint. After multiplication by a scalar one can assume, since $V^c = V_{\epsilon,K}(\tau)^c \cap \mathcal{S}_R^n$ is a compact set, that $\sup_{\sigma\in V^c} \sigma(Q) \le \tau(Q) - 2$. Let Q_η be a self-adjoint polynomial such that $\|Q - Q_\eta\|_R \le \eta/2$. For any state σ we have $|\sigma(Q_\eta) - \sigma(Q)| \le \eta/2$, thus if $\sigma \ne \tau$:

$$\sigma(Q_\eta) \le \sigma(Q) + \eta/2 < \tau(Q) + \eta/2 \le \tau(Q_\eta) + \eta$$

and:

$$\sup_{\sigma\in V^c} \sigma(Q_\eta) \le \tau(Q_\eta) - 2 + \eta < \tau(Q_\eta) - 1. \quad \square$$

6. Extremal states

We first prove a concentration lemma.

Lemma 6.1. *If* τ *is an extremal state in* $\mathcal{S}_R^n$, $n > 1$, *then for any* $\eta, \epsilon, K > 0$ *there exists* $\delta, L > 0$ *such that, for any probability measure* μ *on* $(H_N^R)^n$, *whose barycenter is in* $V_{\delta,L}(\tau)$, *we have:*

$$\mu(\Gamma_R(\tau, \epsilon, K, N)) \ge 1 - \eta.$$

Proof. Let $\eta \in]0, 1/4[$. Then, by Corollary 5.4, we can find some self adjoint polynomial $Q_\eta \in \mathbf{C}\langle X_1, \dots, X_n\rangle$ such that for every $\sigma \in \mathcal{S}_R^n$ we have:

$$\tau(Q_\eta) > \sigma(Q_\eta) - \eta/2$$

and for all $\sigma \notin V_{\epsilon,K}(\tau)$ one has:

$$\sigma(Q_\eta) < \tau(Q_\eta) - 1.$$

Let us now choose $L = \deg(Q_\eta)$, and δ small enough so that for all $\sigma \in V_{\delta,L}(\tau)$ we have:

$$|\tau(Q_\eta) - \sigma(Q_\eta)| < \eta/2.$$

If μ is a probability measure on $(H_N^R)^n$ whose barycenter τ_μ is in $V_{\delta,L}(\tau)$ then we have

$$\begin{aligned}
\tau(Q_\eta) - \eta/2 &\le \tau_\mu(Q_\eta) \\
&= \int_{\Gamma_R(\tau,\epsilon,K,N)} \frac{1}{N}\mathrm{Tr}(Q_\eta)d\mu + \int_{(H_N^R)^n\setminus\Gamma_R(\tau,\epsilon,K,N)} \frac{1}{N}\mathrm{Tr}(Q_\eta)d\mu \\
&\le \mu(\Gamma_R(\tau, \epsilon, K, N))(\tau(Q_\eta) + \eta/2) \\
&\quad + \mu((H_N^R)^n \setminus \Gamma_R(\tau, \epsilon, K, N))(\tau(Q_\eta) - 1) \\
&\le \tau(Q_\eta) + \eta/2 - (1 - \mu(\Gamma_R(\tau, \epsilon, K, N))).
\end{aligned}$$

Therefore

$$\mu(\Gamma_R(\tau, \epsilon, K, N)) \ge 1 - \eta. \quad \square$$

Proposition 6.2. *For any factor state τ in $\mathcal{S}_R^n$, $n > 1$:*

$$\underline{\tilde{\chi}}_R(\tau) = \underline{\chi}_R(\tau).$$

Likewise $\tilde{\chi}_R^{\omega}(\tau) = \chi_R^{\omega}(\tau)$, $\tilde{\chi}_R(\tau) = \chi_R(\tau)$.

Proof. Consider an extremal state $\tau \in \mathcal{S}_R^n$, and $\eta, \epsilon, K > 0$. We can choose δ, L as in Lemma 6.1, so that we can estimate the entropy of μ using (3.2) and the variations of $x \mapsto x \log x + (1-x)\log(1-x)$:

$$\begin{aligned}
\mathrm{Ent}(\mu) &\le \log \mathrm{Leb}\left(\Gamma_R(\tau, \epsilon, K, N)\right) + \mu(\Gamma_R(\tau, \epsilon, K, N)^c) \log \frac{\mathrm{Leb}\left((H_N^R)^n\right)}{\mathrm{Leb}\left(\Gamma_R(\tau, \epsilon, K, N)\right)} \\
&\quad - \eta \log(\eta) - (1-\eta)\log(1-\eta) \\
&\le (1-\eta)\log \mathrm{Leb}\left(\Gamma_R(\tau, \epsilon, K, N)\right) + \eta \log \mathrm{Leb}\left((H_N^R)^n\right) \\
&\quad - \eta \log(\eta) - (1-\eta)\log(1-\eta)
\end{aligned}$$

This inequality holds for all probability measures with barycenter in $V_{\delta,L}(\tau)$, therefore the right hand side is a majorant of $\sup_{\sigma \in V_{\delta,L}(\tau)} \rho_{N,L}(\sigma)$. Now multiply both sides of this inequality by $1/N^2$, add $\frac{n}{2}\log N$ and take lim inf (or lim sup or a limit to ω) then infimum over, successively, δ, L, ϵ, K, to get the result. $\square$

7. Orbital free entropy and freeness in case of additivity of entropy

7.1. Motivation

In this section, we extend the definition of orbital free entropy of [8] to not necessarily hyperfinite multivariables. Let us explain the main ideas before entering into technical details. Orbital entropy aims at measuring the lack of freeness in the same way that relative entropy of a measure μ with respect to the tensor product of its marginals (also called mutual information) does measure the lack of independence in the classical case. In the non-microstate context, Voiculescu first introduced in [22] a notion of mutual information $i^*(W^*(X_1, \ldots, X_m), W^*(X_{m+1}, \ldots, X_{m+n}))$ measuring this lack of freeness using conjugation by a free unitary brownian motion and proved, using this tool, that additivity of non-microstate free entropy implies freeness. In the microstate context, [8] defined $\chi_{orb}(X_1; \ldots; X_n)$, which measures the lack of freeness of $W^*(X_1), \ldots, W^*(X_n)$ (and a variant where the $W^*(X_i)$ are replaced by hyperfinite algebras) relying on the fact that (at least at the level of measure spaces) the space of hermitian matrices can be factored into eigenbasis and eigenvalues, allowing to build microstates in a product of unitary groups. This idea however breaks down when one tries to replace X_i by sets of variables generating non hyperfinite algebras, since in this case there does not exist a good description of the microstates. Our idea here is to overcome this lack of a microstates model by using entropies of measures instead of volumes of microstates. At this point, we have several possible candidates for a generalization. We will use one of them in this section, in order to reach our goal, the result that additivity of free entropy implies freeness. We will explore further possibilities, in order to lay the ground for future investigations, in the last section. Finally, note that we prove this result about additivity only for extremal states. It seems likely that for nonextremal states freeness should be replaced by a kind of freeness with amalgamation with respect to some commutative central algebra, but we do not investigate this in the present paper.

7.2. The orbital free entropy of Hiai, Miyamoto and Ueda

We consider finite sets of non-commutative random variables $\mathbf{X_i} = \{X_{i1}, \ldots, X_{iP_i}\}$ for $i = 1, \ldots, n$, and $\bar{n} = \sum_i P_i$, $\tilde{P} = \max_i P_i$ with joint non-commutative (tracial) distribution $\tau_{\mathbf{X_1};\ldots;\mathbf{X_n}} \in \mathcal{S}_c^{\bar{n}}$. When each set $\mathbf{X_i}$ generates a hyperfinite algebra, Hiai, Miyamoto and Ueda [8] defined orbital free entropy $\chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n})$. Let us recall their definition. Let $(\Xi_i(N))_{i=1\ldots n}; N \to \infty$ be a sequence of matrix sets of size N ($\Xi_i = \{\xi_{i1}, \ldots, \xi_{iP_i}\}$) which approximates $(\mathbf{X_i})_{i=1\ldots n}$ in mixed moments as $N \to \infty$. For $U \in U(N)$ we denote $U\Xi_i(N)U = \{U\xi_{i1}U^*, \ldots, U\xi_{iP_i}U^*\}$. Let

$$\Gamma_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \Xi_1(N), \ldots \Xi_n(N), N, K, \epsilon)$$

be the set of $(U_1, \ldots, U_n) \in U(N)^n$ such that the conjugated sets $(U_i\Xi_i(N)U_i^*)_{i=1\ldots n}$ approximate the mixed moments of $(\mathbf{X_i})_{i=1\ldots n}$ up to an error of ϵ and for degrees less than K or, in other words, such that $\tau_{U_1\Xi_1(N)U_1^*,\ldots,U_n\Xi_n(N)U_n^*} \in V_{\epsilon,K}(\tau_{\mathbf{X_1};\ldots;\mathbf{X_n}})$. Let $\mathcal{H}_N^n$ be the Haar measure on $U(N)^n$ (which, in the sequel, we will always assume normalized to be a probability), and

$$\gamma_{N,\Xi(N),\epsilon,K} = \mathcal{H}_N^n(\Gamma_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \Xi_1(N), \ldots, \Xi_n(N), N, K, \epsilon)) \tag{7.1}$$

then the orbital free entropy is defined as:

$$\chi_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n}) = \lim_{\epsilon\to 0, K\to\infty} \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{N,\Xi(N),\epsilon,K}.$$

It is proved in [8, Lemma 4.2], that this quantity does not depend on the chosen sequence $\Xi(N)$. This relies on Jung's Lemma [10], (see also Lemma 1.2 in [8]) which we recall here for future reference.

Lemma 7.1. *Let $\tau = \tau_{X_1,\ldots,X_m}$ where the variables $X_1, \ldots, X_m$ generate a hyperfinite algebra. Denote by $\|\cdot\|_p$ the p-norm associated with τ. For every $\epsilon > 0$ there exists L, δ such that, for every $\Xi = (\xi_1, \ldots, \xi_m)$ and $\Xi' = (\xi_1', \ldots, \xi_m')$ in $(H_N)^n$, satisfying $\tau_\Xi, \tau_{\Xi'} \in V_{\delta,L}(\tau)$, there exists some unitary $U \in U(N)$ such that*

$$\|U\xi_iU^* - \xi_i'\|_p < \epsilon, \quad \textit{for } i = 1, \ldots, n.$$

Furthermore, Hiai, Miyamoto and Ueda proved that free orbital entropy depends only on the W^*-algebras $W_i = \mathbf{X_i}''$ generated by each set, i.e.

$$\chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) = \chi_{orb}(\mathbf{Y_1}; \ldots; \mathbf{Y_n}),$$

for any other choice of finite sets $\mathbf{Y_1}; \ldots; \mathbf{Y_n}$ such that $W_i = \mathbf{Y_i}''$ (note that one does not assume that $\mathbf{X_i}, \mathbf{Y_i}$ contain the same number of elements). Also they proved the formula relating orbital free entropy to Voiculescu's free entropy:

$$\chi(\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}) = \chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) + \chi(\mathbf{X_1}) + \cdots + \chi(\mathbf{X_n}), \tag{7.2}$$

and proved that, for a set with finite free entropy, additivity of free entropy, i.e.

$$\chi(\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}) = \chi(\mathbf{X_1}) + \cdots + \chi(\mathbf{X_n}), \tag{7.3}$$

which is equivalent to $\chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) = 0$ by (7.2), holds if and only if $\mathbf{X_1}, \ldots, \mathbf{X_n}$ are free.

7.3. Orbital free entropy for arbitrary multivariables

In the following, we give a definition of $\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n})$ for arbitrary finite sets $\mathbf{X_1}; \ldots; \mathbf{X_n}$, which coincides with the previous definition when the sets of multivariables are hyperfinite and $\tau_{\mathbf{X_1};\ldots;\mathbf{X_n}}$ is a factor state.

Let $\mu \in P(H_N^R)^{\bar{n}}$ be a (Borel) probability measure on $(H_N^R)^{\bar{n}} = \prod_i (H_N^R)^{P_i}$, considered as the joint distribution of sets of random matrices $\mathbf{M_1}; \ldots; \mathbf{M_n}$, with $\mathbf{M_i} = \{M_{i1}, \ldots, M_{iP_i}\}$. We denote by $U\mu$ the probability measure on $(H_N^R)^{\bar{n}}$, obtained by conjugating the sets $\mathbf{M_i}$ by independent Haar unitaries from $U(N)$, i.e. $U\mu$ is the joint distribution of the sets $U_i\mathbf{M_i}U_i^* = \{U_i M_{i1} U_i^*, \ldots, U_i M_{iP_i} U_i^*\}$, where $U_1, \ldots, U_n$ are independent unitary matrices, all distributed according to (normalized) Haar measure on $U(N)$. Equivalently, if

$$\Phi_N : U(N)^n \times (H_R)^{\bar{n}} \to (H_R)^{\bar{n}}$$

is the map given by conjugation:

$$(U_i, \mathbf{X_i})_{i=1,\ldots,n} \mapsto (U_i \mathbf{X_i} U_i^*)_{i=1,\ldots,n}$$

then $U\mu$ is given by the pushforward measure:

$$U\mu = \Phi_{N*}(\mathcal{H}_N^n \otimes \mu).$$

Definition 7.2. Let $\mathbf{X_1}; \ldots; \mathbf{X_n}$ be finite sets of noncommutative random variables as above, their *orbital entropy* is defined as:

$$\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) = \sup_{R \geq \mathcal{R}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})} \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty} \left(\frac{1}{N^2} \sup_{\substack{\mu \in P(H_N^R)^{\bar{n}} \\ \tau_\mu \in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})}} \mathrm{Ent}(\mu | U\mu) \right).$$

Similarly we define the lim inf and ultrafilter variants $\underline{\tilde{\chi}}_{orb}$ and $\tilde{\chi}_{orb}^{\omega}$.

Note that, in this definition, limits in ϵ, K are actually infima.

Recall from [21, Def 3.1] that a state is said to have *finite-dimensional approximants* if for every K, ϵ there exists N_0 such that for $N \geq N_0$, $\Gamma_R(\tau, \epsilon, K, N) \neq \emptyset$.

Theorem 7.3. *The orbital free entropy satisfies the following properties.*

(1) *(Negativity)*

$$\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) \leq 0.$$

(2) *(Vanishing for one multivariable)*

$$\tilde{\chi}_{orb}(\mathbf{X}) = 0,$$

for any single multivariable $\mathbf{X} = \{X_1, \ldots, X_m\}$ *having finite-dimensional approximants.*

(3) *(Monotonicity)*

$$\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) \leq \tilde{\chi}_{orb}(\mathbf{Y_1}; \ldots; \mathbf{Y_n}),$$

if $\mathbf{Y}_i \subset \mathbf{X}_i$ *for* $1 \leq i \leq n$.

(4) *(Subadditivity)*

$$\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_m}; \mathbf{X_{m+1}}; \ldots; \mathbf{X_n}) \le \tilde{\chi}_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_m}) + \tilde{\chi}_{orb}(\mathbf{X_{m+1}}; \ldots; \mathbf{X_n}).$$

(5) *(Connection with free entropy)*

$$\tilde{\chi}(\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}) \le \tilde{\chi}_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n}) + \tilde{\chi}(\mathbf{X_1}) + \cdots + \tilde{\chi}(\mathbf{X_n}).$$

(6) *(Agreement with previous definition)*
Assume $\mathbf{X_i}$ are hyperfinite multivariables and let $\chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n})$ denote the orbital free entropy of [8]*, then*

$$\chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) \le \tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}).$$

Moreover, if $\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ is extremal then

$$\chi_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) = \tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}).$$

(7) *(Alternative microstates formula in the extremal case)*
If $\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ is extremal then

$$\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) = \sup_R \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty} \sup_{\Xi,\tau_\Xi \in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})} \left(\frac{1}{N^2}\log \gamma_{N,\Xi,\epsilon,K}\right).$$

(8) *(Dependence on algebras)*
If $\mathbf{X_1}, \ldots, \mathbf{X_n}, \mathbf{Y_1}, \ldots, \mathbf{Y_n}$ are multi-variables such that $\mathbf{Y_i} \subset W^*(\mathbf{X_i})$ for $1 \le i \le n$, then

$$\tilde{\chi}_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n}) \le \tilde{\chi}_{orb}(\mathbf{Y_1}, \ldots, \mathbf{Y_n}).$$

In particular, $\tilde{\chi}_{\mathrm{orb}}(\mathbf{X_1}, \ldots, \mathbf{X_n})$ depends only upon $W^*(\mathbf{X_1}), \ldots, W^*(\mathbf{X_n})$.

(9) *(Orbital Talagrand's inequality and Characterization of Freeness)*
For $\tau = \tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ extremal, let $\tau_{free} = \tau_{\mathbf{X_1}} * \cdots * \tau_{\mathbf{X_n}}$ the free product of its marginals, then:

$$d_W(\tau, \tau_{free}) \le 4R\sqrt{-\tilde{P}\,\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n})},$$

where d_W is the 2-Wasserstein distance of [2]*. As a consequence, if τ is extremal and has finite-dimensional approximants, then $\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n}) = 0$ if and only if $\tau = \tau_{free}$.*

Corollary 7.4. *If $\chi^\omega(\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}) > -\infty$, then*

$$\chi^\omega(\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}) = \chi^\omega(\mathbf{X_1}) + \cdots + \chi^\omega(\mathbf{X_n})$$

if and only if $\mathbf{X_1}, \ldots, \mathbf{X_n}$ are free. The only if part also holds for the limsup variant χ.

Proof of corollary. Assume $\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}$ has finite entropy, and

$$\chi(\mathbf{X_1}, \ldots, \mathbf{X_n}) = \chi(\mathbf{X_1}) + \cdots + \chi(\mathbf{X_n}).$$

By finiteness of Voiculescu's entropy we know that $\tau_{\mathbf{X_1}\cup\cdots\cup\mathbf{X_n}}$ is an extremal state and, by Proposition 6.2, $\chi(\mathbf{X_1}, \ldots, \mathbf{X_n}) = \tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n})$.

Assume also for contradiction $\tilde{\chi}_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n}) < 0$. From (5) of Theorem 7.3, we get:

$$\chi(\mathbf{X_1}) + \cdots + \chi(\mathbf{X_n}) < \tilde{\chi}(\mathbf{X_1}) + \cdots + \tilde{\chi}(\mathbf{X_n}).$$

By the general inequality in Proposition 4.3, there exists an i with $\chi(\mathbf{X_i}) < \tilde{\chi}(\mathbf{X_i})$. By the end of Remark 4.5, the set $\mathbf{X_i}$ contains at least two variables, so that by Proposition 6.2 again, $\tau_{\mathbf{X_i}}$

cannot be extremal, which implies $\chi(\mathbf{X_i}) = -\infty$ by Voiculescu's result [18], a contradiction with $\chi(\mathbf{X_1} \cup \cdots \cup \mathbf{X_n}) > -\infty$.

We thus deduce, using (1) of Theorem 7.3, $\tilde{\chi}_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n}) = 0$.

Then $\mathbf{X_1}, \ldots, \mathbf{X_n}$ are free by point (9) of the same theorem. The ultrafilter variant is similar. The converse statement is due to Voiculescu [21]. $\square$

Proof of Theorem 7.3. In the following we say that μ is an approximating measure for τ if τ_μ belongs to $V_{\epsilon,K}(\tau)$ for some $\epsilon, K > 0$.

(1) Negativity follows from the negativity of relative entropy.

(2) For a single multivariable, if μ is an approximating measure, then $\nu = U\mu$ also approximates with the same precision, and obviously $U\nu = \nu$, therefore $\mathrm{Ent}(\nu|U\nu) = 0$.

(3) If $\mathbf{Y}_i \subset \mathbf{X}_i$ for $1 \le i \le n$ and μ is an approximating measure for the $\mathbf{X}_i$, then its image by the projection map q on the marginal distribution of the $\mathbf{Y}_i$ is an approximating measure for the $\mathbf{Y}_i$, furthermore $qU\mu = Uq\mu$, therefore by (3.1) we have

$$\mathrm{Ent}(\mu|U\mu) \le \mathrm{Ent}(q\mu|qU\mu) = \mathrm{Ent}(q\mu|Uq\mu),$$

and taking limits gives the required inequality.

(4) If μ is an approximating measure for $\mathbf{X_1}; \ldots; \mathbf{X_n}$ let μ_1 and μ_2 denote the marginal distributions of $\mathbf{X_1}; \ldots; \mathbf{X_m}$ and $\mathbf{X_{m+1}}; \ldots; \mathbf{X_n}$, then $U\mu_1$ and $U\mu_2$ are the marginal distributions under $U\mu$, therefore, by subadditivity of relative entropy:

$$\mathrm{Ent}(\mu|U\mu) \le \mathrm{Ent}(\mu_1|U\mu_1) + \mathrm{Ent}(\mu_2|U\mu_2).$$

The inequality follows by taking limits.

(5) Let μ be an approximating measure on $(H_N^R)^{\bar{n}} = (H_N^R)^{P_1} \times \cdots \times (H_N^R)^{P_n}$, with finite entropy, and consider the action of $U(N)^n$ by conjugation on $(H_N^R)^{\bar{n}}$, then $U\mu$ is the average of $(U_1, \ldots, U_n) \cdot \mu$ with respect to Haar measure on $U(N)^n$. Let f be the density of μ with respect to Lebesgue measure on $(H_N^R)^{\bar{n}}$, then f_U, the density of $U\mu$ is the average of $f((U_1, \ldots, U_n)\cdot)$ with respect to Haar measure. It follows that:

$$\begin{aligned}
\mathrm{Ent}(\mu) &= -\int_{(H_N^R)^{\bar{n}}} f \log f \, dM \\
&= -\int_{(H_N^R)^{\bar{n}}} \frac{f}{f_U} \log \frac{f}{f_U} f_U \, dM - \int_{(H_N^R)^{\bar{n}}} f \log f_U dM \\
&= \mathrm{Ent}(\mu|U\mu) - \int_{(H_N^R)^{\bar{n}}} f \log f_U dM \\
&= \mathrm{Ent}(\mu|U\mu) - \int_{(H_N^R)^{\bar{n}}} f_U \log f_U dM \quad \text{by } U(N)^n \text{ invariance of } dM \\
&= \mathrm{Ent}(\mu|U\mu) + \mathrm{Ent}(U\mu).
\end{aligned}$$

Now we can use the subadditivity of $\mathrm{Ent}(U\mu)$ with respect to the projections on the spaces $(H_N^R)^{P_i}$, which gives

$$\mathrm{Ent}(U\mu) \le \mathrm{Ent}(p_1 U\mu) + \cdots + \mathrm{Ent}(p_n U\mu)$$

Letting $N \to \infty$, $K \to \infty$, $\epsilon \to 0$ gives the required inequality.

(6) Let $\Xi(N)$ be an approximating sequence, as in the definition of (hyperfinite) orbital free entropy. Let $\nu_{\Xi(N)}$ be the probability measure obtained by restricting $\mathcal{H}_N^n$ to $\Gamma_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n} :$

$\Xi_1(N), \ldots \Xi_n(N), N, K, \epsilon)$ and normalizing (if the orbital entropy is finite and N is sufficiently large, this measure is well defined), then (recall (7.1))

$$\log \gamma_{N,\Xi(N),\epsilon,K} = \mathrm{Ent}(\nu_{\Xi(N)}|\mathcal{H}_N^n).$$

Let $\Psi_{\Xi(N)} : U(N)^n \to (H_R)^{\bar{n}}$ the map given by conjugation:

$$(U_i)_{i=1,\ldots,n} \mapsto (U_i \Xi_i(N) U_i^*)_{i=1,\ldots,n}$$

we have $U\Psi_{\Xi(N)*}(\nu_{\Xi(N)}) = \Psi_{\Xi(N)*}(\mathcal{H}_N^n)$ therefore, by (3.1)

$$\frac{1}{N^2}\mathrm{Ent}(\nu_{\Xi(N)}|\mathcal{H}_N^n) \le \frac{1}{N^2}\mathrm{Ent}(\Psi_{\Xi(N)*}(\nu_{\Xi(N)})|U\Psi_{\Xi(N)*}(\nu_{\Xi(N)})).$$

Since $\Psi_{\Xi(N)*}(\nu_{\Xi(N)})$ is an approximating measure for $\mathbf{X_1}, \ldots, \mathbf{X_n}$, the right hand side, after taking limits in N, ϵ, K, is bounded by $\tilde{\chi}_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n})$. The inequality $\chi_{orb} \le \tilde{\chi}_{orb}$ follows.

Let us now assume that $\tau := \tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ is extremal. Fix $\eta, \epsilon > 0$ and an integer $K > 0$. Using Jung's Lemma, and following the proof of Lemma 4.2 in [8], we can take $\delta \le \epsilon/2$, $L \ge K$ such that, for all families of sets $(\Theta_i)_{i=1,\ldots,n}$ of $N \times N$ hermitian matrices such that for all i we have $\tau_{(\Theta_i)} \in V_{\delta,L}(p_i\tau)$ (a fortiori if $\tau_{(\Theta_i)_{i=1,\ldots,n}} \in V_{\delta,L}(\tau)$) we have, for N large enough:

$$\gamma_{N,\Theta,\epsilon/2,K} \le \gamma_{N,\Xi(N),\epsilon,K}. \tag{7.4}$$

Note also the elementary equality for any $U_1, \ldots, U_n$ unitaries coming from invariance of the Haar measure:

$$\gamma_{N,\Phi_N(U_1,\ldots,U_n,\Theta),\epsilon/2,K} = \gamma_{N,\Theta,\epsilon/2,K}. \tag{7.5}$$

Then using Lemma 6.1, if we take $\delta' > 0$ sufficiently small and L' sufficiently large, for any measure μ on $(H_N^R)^{\bar{n}}$ such that $\tau_\mu \in V_{\delta',L'}(\tau)$, we get:

$$\mu(\Gamma_R(\tau, \delta, L, N)) \ge 1 - \eta.$$

Therefore, by (3.3),

$$\mathrm{Ent}(\mu|U\mu) \le (1-\eta)\log[U\mu(\Gamma_R(\tau, \delta, L, N))] - f(\eta), \tag{7.6}$$

with $f(\eta) = \eta\log\eta + (1-\eta)\log(1-\eta)$.

Let $U\Gamma_R(\tau, \delta, L, N) = \{\Theta \mid \exists(U_1, \ldots, U_n)\ \Phi_N(U_1, \ldots, U_n, \Theta) \in \Gamma_R(\tau, \delta, L, N)\}$. The measure $U\mu$ is the image of $\mathcal{H}_N^n \otimes \mu$ by the conjugation map Φ_N, and the set $\Phi_N^{-1}(\Gamma_R(\tau, \delta, L, N))$ is the union over matrix sets:

$$\cup_{\Theta\in U\Gamma_R(\tau,\delta,L,N)}\Gamma_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \Theta_1, \ldots \Theta_n, N, L, \delta) \times \{\Theta\}.$$

It follows that:

$$\begin{aligned} U\mu(\Gamma_R(\tau, \delta, L, N)) &= \int_{U\Gamma_R(\tau,\delta,L,N)} \gamma_{N,\Theta,\delta,L}\, d\mu(\Theta) \\ &\le \int_{U\Gamma_R(\tau,\delta,L,N)} \gamma_{N,\Theta,\epsilon/2,K}\, d\mu(\Theta) \quad \text{by } \delta \le \epsilon/2,\ L \ge K \\ &\le \gamma_{N,\Xi(N),\epsilon,K} \quad \text{by (7.4) and (7.5).} \end{aligned} \tag{7.7}$$

Then combining (7.6), (7.7) and taking limits yields the inequality:

$$\tilde{\chi}_{orb} \le (1-\eta)\chi_{orb}.$$

Since η is arbitrary, we are done.

(7) The proof is a variant of the one in (6). First take some family Ξ of hermitian matrices with $\tau_\Xi \in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})$. Replacing $\Xi(N)$ by Ξ in the arguments of the first part of (6) we deduce:

$$\log \gamma_{N,\Xi,\epsilon,K} = \mathrm{Ent}(\nu_\Xi|\mathcal{H}_N^n) \leq \mathrm{Ent}(\Psi_{\Xi*}(\nu_\Xi)|U\,\Psi_{\Xi*}(\nu_\Xi)).$$

Since $\tau_{\Psi_{\Xi*}(\nu_\Xi)} \in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})$ we obtain the following inequality:

$$\log \gamma_{N,\Xi,\epsilon,K} \leq \sup_{\substack{\mu\in P(H_N^R)^{\bar{n}} \\ \tau_\mu\in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})}} \mathrm{Ent}(\mu|U\mu).$$

This implies the lower bound in the statement.

Assume now that $\tau = \tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ extremal. Fix η, ϵ, K choose δ, L as in Lemma 6.1. For any $\mu \in P(H_N^R)^{\bar{n}}$ with $\tau_\mu \in V_{\delta,L}(\tau)$ we have:

$$\mathrm{Ent}(\mu|U\mu) \leq (1-\eta)\log\left[U\mu(\Gamma_R(\tau,\epsilon,K,N))\right] - f(\eta).$$

With the same computation as in the proof of (7.7) we get the inequality:

$$U\mu(\Gamma_R(\tau,\epsilon,K,N)) \leq \sup_{\Xi,\tau_\Xi\in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})} \gamma_{N,\Xi,\epsilon,K}.$$

The second inequality of the statement follows.

(8) Let $\mathbf{X_i} = \{X_{i1},\ldots,X_{iP_i}\}$ and $\mathbf{Y_i} = \{Y_{i1},\ldots,Y_{iQ_i}\}$, with $\bar{m} = \sum_i Q_i$. By Kaplansky density theorem, for each i, one can find a set of non-commutative polynomials $P_{ij}(\mathbf{X_i})$, $j = 1,\ldots,Q_i$ as close as we want in distribution to the set $\mathbf{Y_i}$. For such a family, we write:

$$\mathbf{P_i}(\mathbf{X_i}) = (P_{i1}(\mathbf{X_i}),\ldots,P_{iQ_i}(\mathbf{X_i})),$$
$$\mathbf{P}(\mathbf{X_1},\ldots,\mathbf{X_n}) = (\mathbf{P_1}(\mathbf{X_1}),\ldots,\mathbf{P_n}(\mathbf{X_n})).$$

Let $\epsilon, K > 0$. One can find polynomials $P_{ij}(\mathbf{X_i})$, $j = 1,\ldots,Q_i$, a real $\delta > 0$ sufficiently small and an integer L sufficiently large such that for all μ probability measure on $(H_N^R)^{\bar{n}}$ in $V_{\delta,L}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})$ we have $\mathbf{P}_\star\mu \in V_{\epsilon,K}(\tau_{\mathbf{Y_1},\ldots,\mathbf{Y_n}})$.

Since $\Phi_n((U_1,\ldots,U_n),(\mathbf{P_1}(\mathbf{X_1}),\ldots,\mathbf{P_n}(\mathbf{X_n}))) = (\mathbf{P_1}(U_1\mathbf{X_1}U_1^*),\ldots,\mathbf{P_n}(U_n\mathbf{X_n}U_n^*))$, it is clear that $U\mathbf{P}_\star\mu = \mathbf{P}_\star U\mu$. By (3.1) we have $\mathrm{Ent}(\mathbf{P}_\star\mu|U\mathbf{P}_\star\mu) \geq \mathrm{Ent}(\mu|U\mu)$, therefore

$$\frac{1}{N^2}\sup_{\nu,\tau_\nu\in V_{\epsilon,K}(\tau_{\mathbf{Y_1},\ldots,\mathbf{Y_n}})} \mathrm{Ent}(\nu|U\nu) \geq \frac{1}{N^2}\sup_{\mu,\tau_\mu\in V_{\delta,L}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})} \mathrm{Ent}(\mu|U\mu).$$

Now take a lim sup then infimum over, successively, δ, L, ϵ, K, to get the result.

(9) First, choose a subsequence N_m and μ_m probability measures on $(H_{N_m}^R)^{\bar{n}}$ such that (τ_{μ_m}) converges weakly to τ and:

$$\tilde{\chi}_{orb}(\mathbf{X_1};\ldots;\mathbf{X_n}) = \lim_{m\to\infty}\left(\frac{1}{N_m^2}\mathrm{Ent}(\mu_m|U\mu_m)\right).$$

Without loss of generality, we assume that this orbital entropy is finite. We follow arguments close to the proof of Lemma 3.4 in [8]. Remark that in the definition of $U\mu$ we can replace the unitary group $U(N)$ by $SU(N)$, since $U(N)$ acts by conjugation. Let $f_m(\mathbf{M})$ be the density of μ_m with respect to $U\mu_m$ (which exists if m is sufficiently large). Then for almost all values of $\mathbf{M}$, the function

$$g_m(U_1,\ldots,U_n,\mathbf{M}) = f_m(U_1\mathbf{M}_1U_1^*,\ldots,U_n\mathbf{M}_nU_n^*)$$

is a probability density in the variables $U_1, \ldots, U_n$, with respect to the Haar measure $\mathcal{SH}^n_{N_m}$ on $SU(N_m)^n$. For $\mathbf{M}$ let $\pi_{\mathbf{M}}$, be a probability measure on $SU(N_m)^n \times SU(N_m)^n$, which is an optimal coupling between $g_m(U_1, \ldots, U_n, \mathbf{M})\mathcal{SH}^n_{N_m}$ and $\mathcal{SH}^n_{N_m}$ for the geodesic distance on $SU(N_m)^n$. This means that the marginals of the measure $\pi_{\mathbf{M}}$ on the two components of $SU(N_m)^n \times SU(N_m)^n$ are the measures $g_m(U_1, \ldots, U_n, \mathbf{M})\mathcal{SH}^n_{N_m}$ and $\mathcal{SH}^n_{N_m}$, and the squared Wasserstein distance between the measures $g_m(U_1, \ldots, U_n, \mathbf{M})\mathcal{SH}^n_{N_m}$ and $\mathcal{SH}^n_{N_m}$ is

$$\int_{SU(N_m)^n \times SU(N_m)^n} \left[d_{\text{geod}}((U_1, \ldots, U_n), (V_1, \ldots, V_n))\right]^2 d\pi_{\mathbf{M}}(U, V).$$

Such a measure can be constructed measurably with respect to $\mathbf{M}$ (see e.g. Corollary 5.22 in [17]). We thus deduce an estimate for the non-commutative 2-Wasserstein distance:

$$\begin{aligned}
&d_W(\tau_{\mu_m}, \tau_{U\mu_m})^2 \\
&\quad \le \int dU\mu_m(\mathbf{M}) \int d\pi_{\mathbf{M}}(\mathbf{U}, \mathbf{V}) \sum_{i=1}^{n} \sum_{j=1}^{P_i} \frac{1}{N_m} \|U_i \mathbf{M}_{ij} U_i^* - V_i \mathbf{M}_{ij} V_i^*\|^2_{HS} \\
&\quad \le \int dU\mu_m(\mathbf{M}) \int d\pi_{\mathbf{M}}(\mathbf{U}, \mathbf{V}) 4R^2 \tilde{P} \frac{1}{N_m} \sum_{i=1}^{n} \|\mathbf{U}_i - \mathbf{V}_i\|^2_{HS} \\
&\quad \le \int dU\mu_m(\mathbf{M}) \int d\pi_{\mathbf{M}}(\mathbf{U}, \mathbf{V}) 4R^2 \tilde{P} \frac{1}{N_m} \left[d_{\text{geod}}((U_1, \ldots, U_n), (V_1, \ldots, V_n))\right]^2,
\end{aligned}$$

where we used the fact that the Hilbert–Schmidt distance can be majorized by the geodesic distance. Now using the Talagrand inequality of [14] on $SU(N_m)^n$, as in Proposition 3.5 of [8] we get:

$$\begin{aligned}
d_W(\tau_{\mu_m}, \tau_{U\mu_m})^2 \le (4R)^2 \tilde{P} \int dU\mu_m(\mathbf{M}) \frac{-1}{N_m^2} &\mathrm{Ent}(g(U_1, \ldots, U_n, \mathbf{M}) \\
&\times \mathcal{SH}^n_{N_m}(U) | \mathcal{SH}^n_{N_m}(U)).
\end{aligned}$$

Now we can use the fact that $U\mu_m$ is invariant by the action of $U(N)^n$ and interchange the order of integration to get

$$\begin{aligned}
&\int \mathrm{Ent}(g(U_1, \ldots, U_n, \mathbf{M})\mathcal{SH}^n_{N_m}(U) | \mathcal{SH}^n_{N_m}(U)) dU\mu_m(\mathbf{M}) \\
&\quad = \int\int f_m(U_1 \mathbf{M}_1 U_1^*, \ldots, U_n \mathbf{M}_n U_n^*) \log f_m(U_1 \mathbf{M}_1 U_1^*, \ldots, U_n \mathbf{M}_n U_n^*) \\
&\quad\quad \times d\mathcal{H}^n_{N_m}(\mathbf{U}) dU\mu_m(\mathbf{M}) \\
&\quad = \mathrm{Ent}(\mu_m | U\mu_m),
\end{aligned}$$

thus

$$d_W(\tau_{\mu_m}, \tau_{U\mu_m})^2 \le \frac{-(4R)^2 \tilde{P}}{N_m^2} \mathrm{Ent}(\mu_m | U\mu_m).$$

By our choice of μ_m, the noncommutative distribution of the random matrix sets $\mathbf{M}$ under μ_m converges weakly to τ as $N \to \infty$. Let us check that similarly, under $U\mu_m$ this noncommutative distribution converges weakly to τ_{free}. This is a consequence of Remark 3.2 in [3]. Indeed, there it is proved that $\mathbf{M}$ is asymptotically free from $\{U_1\}, \ldots, \{U_n\}$ (independent Haar unitaries) provided the distribution of $\mathbf{M}$ concentrates around its mean. But this concentration is provided by Lemma 6.1. We leave the easy but tedious details to the reader.

As a consequence of Talagrand's inequality, the only if part of the characterization of freeness is obvious. Now assume $\tau = \tau_{free}$ and take μ_m as above so that now $\tau_{U\mu_m}$ tends weakly to $\tau = \tau_{free}$. Thus for m large enough so that $\tau_{U\mu_m}\epsilon, K$ approximates τ we have

$$0 = \left(\frac{1}{N_m^2}\mathrm{Ent}(U\mu_m|U\mu_m)\right) \le \frac{1}{N_m^2} \sup_{\substack{\mu\in P(H_{N_m}^R)^{\bar{n}} \\ \tau_\mu\in V_{\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})}} \mathrm{Ent}(\mu|U\mu).$$

As a consequence taking a limit in m and then in ϵ, K since they are arbitrary in the argument above, we get $\tilde{\chi}_{orb}(\mathbf{X_1},\ldots,\mathbf{X_n}) = 0$. □

8. Preliminaries about entropy, marginals and unitary invariant versions of a measure

Before giving several other generalizations of orbital entropy, we start with some preliminary results.

Let $\mu \in P((H_R^N)^{\bar{n}})$, considered as the probability distribution of a family of random matrices $(\mathbf{A_1},\ldots,\mathbf{A_n})$, where each $\mathbf{A_i}$ consists in a bunch of variables like $\mathbf{X_i}$. Again $U\mu$ is then the law of $(U_1\mathbf{A_1}U_1^*,\ldots,U_n\mathbf{A_n}U_n^*)$ where U_i are independent variables distributed with respect to the Haar measure $\mathcal{H}_N$ of the unitary group $U(N)$. More generally, we consider partial conjugations in the following way: if $\pi : [1,n] \to [1,\ell]$ is a surjective map (equivalently, we can consider the partition $\Pi = \{\pi^{-1}(i)\}_{i=1,\ldots,\ell}$ of $[1,n]$ which it defines) we denote $U^\pi\mu (= U^\Pi\mu)$ the law of $(U_{\pi(1)}\mathbf{A_1}U_{\pi(1)}^*,\ldots,U_{\pi(n)}\mathbf{A_n}U_{\pi(n)}^*)$ where the U_i, $i = 1\ldots,\ell$ are independent Haar random unitary matrices. We will write U^G for the global unitary invariant version, corresponding to $\Pi = \{\{1,\ldots,n\}\}$. It is clear that for any absolutely continuous measure μ the measure $U^\pi\mu$ is absolutely continuous.

Lemma 8.1. (i) *Let $\mu,\nu \in P((H_R^N)^{\bar{n}})$ with $U^\pi\nu = \nu$, then*

$$\mathrm{Ent}(\mu|\nu) = \mathrm{Ent}(\mu|U^\pi\mu) + \mathrm{Ent}(U^\pi\mu|\nu).$$

(ii) *Let $\mu \in P((H_R^N)^{\bar{n}})$ and $\nu = \bigotimes_i \nu_i$, $\nu_i \in P((H_R^N)^{\tilde{P}_i})$. We denote $q_1\mu$ and $q_2\mu$ the marginals for the bunch of variables corresponding, respectively, to $(\mathbf{A_1},\ldots,\mathbf{A_m})$ and $(\mathbf{A_{m+1}},\ldots,\mathbf{A_n})$, then*

$$\mathrm{Ent}(\mu|\nu) = \mathrm{Ent}(\mu|q_1\mu\otimes q_2\mu) + \mathrm{Ent}(q_1\mu\otimes q_2\mu|\nu).$$

(iii) *With the notations of* (ii) *and $V = U^\Pi$ for $\Pi = \{\{1,\ldots,m\},\{m+1,\ldots,n\}\}$ we have:*

$$\mathrm{Ent}(V\mu|U\mu) \le \mathrm{Ent}(q_1U^G\mu|Uq_1\mu) + \mathrm{Ent}(q_2U^G\mu|Uq_2\mu).$$

Proof. (i) This is a generalization to relative entropy of an equality in the proof of Theorem 7.3(5) above. Without loss of generality we assume $\mu \ll \nu$ since if we do not have both $\mu \ll U^\pi\mu$ and $U^\pi\mu \ll \nu$, the right hand side is $-\infty$ and the equality is true if we do not have $\mu \ll \nu$, so that in any case we can assume $\mu \ll \nu$. Consider $\rho = \frac{d\mu}{d\nu}$. Since ν is unitarily invariant, we have $U^\pi\mu \ll U^\pi\nu = \nu$. Moreover,

$$\begin{aligned}\rho_U(\mathbf{A_1},\ldots,\mathbf{A_n}) &:= \frac{dU^\pi\mu}{d\nu}(\mathbf{A_1},\ldots,\mathbf{A_n}) \\ &= \int d\mathcal{H}_N^\ell(U_1,\ldots,U_\ell)\rho(U_{\pi(1)}\mathbf{A_1}U_{\pi(1)}^*,\ldots,U_{\pi(n)}\mathbf{A_n}U_{\pi(n)}^*).\end{aligned}$$

Using (3.1) we have $\mathrm{Ent}(\mu|\nu) \leq \mathrm{Ent}(U^{\pi}\mu|\nu)$ and we can thus compute:

$$\begin{aligned}\mathrm{Ent}(\mu|\nu) &= -\int \rho \ln(\rho) d\nu = -\int \rho \ln(\rho_U) d\nu - \int \rho \ln\left(\frac{\rho}{\rho_U}\right) d\nu \\ &= -\int \rho_U \ln(\rho_U) d\nu + \mathrm{Ent}(\mu|U\mu) \\ &= \mathrm{Ent}(U\mu|\nu) + \mathrm{Ent}(\mu|U\mu),\end{aligned}$$

where, in the third line, we used unitary invariance to replace ρ by ρ_U. The reverse implication starting from finiteness of the left hand side is also clear.

(ii) The proof is similar to (i). In order to solve finiteness issues, one can again use (3.1) to get $\mathrm{Ent}(\mu|\nu) \leq \mathrm{Ent}(q_i\mu|q_i\nu)$.

(iii) The inequality comes from subadditivity of entropy. Indeed consider, without loss of generality, ρ_V the density of $V\mu$ with respect to $U\mu$. Using unitary invariance of $U\mu$ we get:

$$\begin{aligned}\mathrm{Ent}(V\mu|U\mu) &= -\int \rho_V \ln(\rho_V) dU\mu \\ &= -\int dU\mu(\mathbf{A}) \int R(U_1, \ldots, U_n, \mathbf{A}) \\ &\quad \times \ln(R(U_1, \ldots, U_n, \mathbf{A})) d\mathcal{H}_N^n(U_1, \ldots, U_n),\end{aligned}$$

where, for a.e. $\mathbf{A} = (\mathbf{A_1}, \ldots, \mathbf{A_n})$, the quantity

$$R(U_1, \ldots, U_n, \mathbf{A}) = \rho_V(U_1\mathbf{A_1}U_1^*, \ldots, U_n\mathbf{A_n}U_n^*)$$

is a probability density on $U(N)^n$. Let R_1, R_2 be the densities of marginals, namely, with obvious notations,

$$\begin{aligned}R_1(\mathbf{U_2}, \mathbf{A}) &= \int d\mathcal{H}_N^m(\mathbf{U_1}) R(\mathbf{U_1}, \mathbf{U_2}, \mathbf{A}), \\ R_2(\mathbf{U_1}, \mathbf{A}) &= \int d\mathcal{H}_N^{n-m}(\mathbf{U_2}) R(\mathbf{U_1}, \mathbf{U_2}, \mathbf{A}),\end{aligned}$$

we have

$$R_2(\mathbf{U_1}, \mathbf{A}) = R_2(\mathbf{I}, U_1\mathbf{A_1}U_1^*, \ldots, U_m\mathbf{A_m}U_m^*, \mathbf{A_{m+1}}, \ldots, \mathbf{A_n}).$$

Moreover

$$R_2(\mathbf{I}, \mathbf{A}) dU\mu(\mathbf{A}) = dV\mu(\mathbf{A})$$

is a probability measure with marginal $q_1V\mu = q_1U^G\mu$. Using the subadditivity of ordinary entropy relative to a product measure, we get:

$$\begin{aligned}\mathrm{Ent}(V\mu|U\mu) &\leq -\int dU\mu(\mathbf{A}) R_2(\mathbf{I}, \mathbf{A}) \ln(R_2(\mathbf{I}, \mathbf{A})) \\ &\quad - \int dU\mu(\mathbf{A}) R_1(\mathbf{I}, \mathbf{A}) \ln(R_1(\mathbf{I}, \mathbf{A})) \\ &\leq \mathrm{Ent}(q_1U^G\mu|Uq_1\mu) + \mathrm{Ent}(q_2U^G\mu|Uq_2\mu). \quad \square\end{aligned}$$

9. Variants and extensions

9.1. Overview

The main drawback of our definition of orbital free entropy is that we are unable to prove equality in part (5) of Theorem 7.3. In order to overcome this problem, as mentioned at the beginning of Section 7, several other generalizations of orbital entropy may be considered. We will describe below two variants which we call maximal mutual entropy and I-mutual entropy. The last one satisfies the required additivity property however we lose the fact that it depends only on the subalgebras generated by the subset of variables. Let us describe briefly the content of this section. First, we can consider, as for Voiculescu's entropy, a variant of free entropy in the presence of another set of variables, which plays a dummy role in the definition. This will be considered in Section 9.2. Instead of using the relative entropy of μ, an approximating measure, with respect to its unitary invariant mean $U\mu$, we can consider the relative entropy with respect to the product of the marginal distributions of $U\mu$ with respect to the subsets. This yields a quantity which we call maximal mutual entropy, and which we consider in Section 9.3. Again this quantity depends only on the W^* algebras generated by the subsets, and is subadditive. Another alternative is to use Ciszar's I-projection first and then to take the relative entropy of this specific measure with respect to the tensor product of its marginals (which are automatically unitary invariant in this case). This gives what we call I-mutual entropy, studied in Section 9.4. This quantity satisfies a strong additivity property (property below), which generalizes the additivity of the orbital entropy of [8]. Unfortunately, we are not able to prove that it depends only on the W^* algebras generated by the subsets. All these entropies coincide with orbital entropy defined in [8] in the context they define it. It is plausible that they always coincide, although we do not have a proof of this fact at this stage.

9.2. Orbital entropy in the presence of other variables

As in Section 6, we consider finite sets of non-commutative random variables $\mathbf{X_i} = \{X_{i1}, \ldots, X_{iP_i}\}$ for $i = 1, \ldots, n$, and $\bar{n} = \sum_i P_i$, while $\mathbf{Y} = \{Y_1, \ldots, Y_t\}$ is likewise a multivariable containing t variables. Their joint non-commutative (tracial) distribution is $\tau = \tau_{\mathbf{X_1};\ldots;\mathbf{X_n};\mathbf{Y}} \in \mathcal{S}_R^{\bar{n}+t}$. We will use the notation:

$$A_{N,\epsilon,K}(\tau) = \{\mu \in P((H_R^N)^{\bar{n}+t}) \mid \tau_\mu \in V_{\epsilon,K}(\tau)\}.$$

Also we denote $p\mu$ the marginal distribution of μ on the $\mathbf{X}$ variables.

Definition 9.1. The *free orbital entropy of* $\mathbf{X_1}, \ldots, \mathbf{X_n}$ *in the presence of* $\mathbf{Y}$ *is, if* $\tau = \tau_{\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y}}$:

$$\tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) = \sup_{R \geq \mathcal{R}(\tau)} \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty} \left(\frac{1}{N^2} \sup_{\mu \in A_{N,\epsilon,K}(\tau)} \mathrm{Ent}(p\mu | Up\mu) \right).$$

The orbital free entropy in the presence of other variables satisfies properties similar to the ones of Theorem 7.3, the proofs being easy variations on the proofs for the orbital free entropy. We state here only an improved version of the additivity property.

Theorem 9.2.

$$\begin{aligned}&\tilde{\chi}_{orb}(\mathbf{X_1};\ldots;\mathbf{X_m};\mathbf{X_{m+1}};\ldots;\mathbf{X_n}:\mathbf{Y}) \le \tilde{\chi}_{orb}(\mathbf{X_1}\cup\cdots\cup\mathbf{X_m};\mathbf{X_{m+1}}\cup\cdots\cup\mathbf{X_n}:\mathbf{Y})\\ &\quad+\tilde{\chi}_{orb}(\mathbf{X_1};\ldots;\mathbf{X_m}:\mathbf{X_{m+1}}\cup\cdots\cup\mathbf{X_n}\cup\mathbf{Y})\\ &\quad+\tilde{\chi}_{orb}(\mathbf{X_{m+1}};\ldots;\mathbf{X_n}:\mathbf{X_1}\cup\cdots\cup\mathbf{X_m}\cup\mathbf{Y}).\end{aligned}$$

Proof. Write for μ in $A_{N,\epsilon,K}(\tau)Vp\mu$ as in Lemma 8.1(iii) the unitary invariant variant for blocks. Note that $pU^G\mu = U^G p\mu$ and:

$$\mathrm{Ent}(p\mu|Up\mu) = \mathrm{Ent}(p\mu|U^G p\mu) + \mathrm{Ent}(U^G p\mu|Up\mu) \le \mathrm{Ent}(U^G p\mu|Up\mu),$$

(from Lemma 8.1(i)) so that, since $U^G\mu \in A_{N,\epsilon,K}(\tau)$, we may assume $\mu = U^G\mu$ when we bound orbital entropy. Applying Lemma 8.1(i) and (iii) we get the concluding estimate

$$\begin{aligned}\mathrm{Ent}(p\mu|Up\mu) &= \mathrm{Ent}(p\mu|Vp\mu) + \mathrm{Ent}(Vp\mu|Up\mu)\\ &\le \mathrm{Ent}(p\mu|Vp\mu) + \mathrm{Ent}(q_1\mu|Uq_1\mu) + \mathrm{Ent}(q_2\mu|Uq_2\mu). \quad \square\end{aligned}$$

9.3. Maximal mutual entropy

We use the same notations as in the preceding section, and denote $p_1,\ldots,p_n$ the projections on the sets of variables $\mathbf{X_1};\ldots;\mathbf{X_n}$.

Definition 9.3. The *free maximal mutual entropy of* $\mathbf{X_1},\ldots,\mathbf{X_n}$ *in the presence of* $\mathbf{Y}$ is, if $\tau = \tau_{\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y}}$:

$$\begin{aligned}&\tilde{\chi}_{Mmut}(\mathbf{X_1};\ldots;\mathbf{X_n}:\mathbf{Y})\\ &\quad= \sup_{R\ge\mathcal{R}(\tau)} \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty}\left(\frac{1}{N^2}\sup_{\mu\in A_{N,\epsilon,K}(\tau)} \mathrm{Ent}(p\mu|p_1Up\mu\otimes\cdots\otimes p_nUp\mu)\right).\end{aligned}$$

If $\mathbf{Y}$ is empty we just write $\tilde{\chi}_{Mmut}(\mathbf{X_1};\ldots;\mathbf{X_n})$.

Note that the limits in ϵ, K are actually infima. We also define a notion of relative entropy to state the best subadditivity result. We compare it in the next subsection, but note already that it coincides with the definition of Section 4 when $\mathbf{Y}=\emptyset$.

Definition 9.4. We define, for $\tau = \tau_{\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y}}$, a *random microstate free entropy in the presence* of $\mathbf{Y}$ as:

$$\begin{aligned}&\tilde{\chi}(\mathbf{X_1};\ldots;\mathbf{X_n}:\mathbf{Y})\\ &\quad= \sup_{R\ge\mathcal{R}(\tau)} \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty}\left(\frac{1}{N^2}\sup_{\mu\in A_{N,\epsilon,K}(\tau)} \mathrm{Ent}(p\mu) + \frac{n}{2}\log N\right).\end{aligned}$$

Theorem 9.5. *The free maximal mutual entropy satisfies the following properties:*

(1) *(Vanishing for one variable)*

$$\tilde{\chi}_{Mmut}(\mathbf{X_1}) = 0,$$

for any single multivariable having finite-dimensional approximants.

(2) *(Improved Subadditivity)*

$$\begin{aligned}&\tilde{\chi}_{Mmut}(\mathbf{X_1};\ldots;\mathbf{X_m};\mathbf{X_{m+1}};\ldots;\mathbf{X_n}:\mathbf{Y})\\&\quad\le \tilde{\chi}_{Mmut}(\mathbf{X_1}\cup\cdots\cup\mathbf{X_m};\mathbf{X_{m+1}}\cup\cdots\cup\mathbf{X_n}:\mathbf{Y})\\&\qquad+\tilde{\chi}_{Mmut}(\mathbf{X_1};\ldots;\mathbf{X_m}:\mathbf{X_{m+1}}\cup\cdots\cup\mathbf{X_n}\cup\mathbf{Y})\\&\qquad+\tilde{\chi}_{Mmut}(\mathbf{X_{m+1}};\ldots;\mathbf{X_n}:\mathbf{X_1}\cup\cdots\cup\mathbf{X_m}\cup\mathbf{Y}).\end{aligned}$$

(3) *(Improved subadditivity of entropy)*

$$\tilde{\chi}(\mathbf{X_1},\mathbf{X_2}:\mathbf{Y})\le\tilde{\chi}_{Mmut}(\mathbf{X_1};\mathbf{X_2}:\mathbf{Y})+\tilde{\chi}(\mathbf{X_1}:\mathbf{X_2}\cup\mathbf{Y})+\tilde{\chi}(\mathbf{X_2}:\mathbf{X_1}\cup\mathbf{Y}).$$

(4) *(Agreement with previous definition) If $\mathbf{X_i}$ are hyperfinite multivariables then*

$$\chi_{orb}(\mathbf{X_1};\ldots;\mathbf{X_n})\le\tilde{\chi}_{Mmut}(\mathbf{X_1};\ldots;\mathbf{X_n})$$

(where free orbital entropy is in the sense of [8]). If moreover $\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ is extremal then

$$\tilde{\chi}_{Mmut}(\mathbf{X_1};\ldots;\mathbf{X_n})=\chi_{orb}(\mathbf{X_1};\ldots;\mathbf{X_n}).$$

(5) *(Dependence on algebras) If $\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y_1},\ldots,\mathbf{Y_n}$ are multi-variables such that $\mathbf{Y_i}\subset W^*(\mathbf{X_i})$ for $1\le i\le n$, then*

$$\tilde{\chi}_{Mmut}(\mathbf{X_1},\ldots,\mathbf{X_n})\le\tilde{\chi}_{Mmut}(\mathbf{Y_1},\ldots,\mathbf{Y_n}).$$

In particular, $\tilde{\chi}_{Mmut}(\mathbf{X_1},\ldots,\mathbf{X_n})$ depends only upon $W^*(\mathbf{X_1}),\ldots,W^*(\mathbf{X_n})$.

Proof. The proofs of (1), (5) are similar to the corresponding properties of $\tilde{\chi}_{orb}$.

(2) Let p be the projection on the $\mathbf{X}$ variables, p_i the projection $\mathbf{X_i}$, and q_1,q_2 the projections on $\mathbf{X_1},\ldots,\mathbf{X_m}$ and $\mathbf{X_{m+1}},\ldots,\mathbf{X_n}$, respectively. Let $\mu\in A_{N,\epsilon,K}(\tau)$, we may assume, as in the proof of Theorem 9.2, $\mu=U^G\mu$, so that we have $p_iUp\mu=p_i\mu$ and $q_iV\mu=q_i\mu$. Applying Lemma 8.1(ii) we get:

$$\begin{aligned}\mathrm{Ent}\left(p\mu\,\middle|\,\bigotimes_i p_iUp\mu\right)&=\mathrm{Ent}(p\mu|q_1Vp\mu\otimes q_2Vp\mu)\\&\quad+\mathrm{Ent}\left(q_1\mu\otimes q_2\mu\,\middle|\,\bigotimes_i p_iUp\mu\right)\end{aligned}$$

And we have:

$$\begin{aligned}\mathrm{Ent}\left(q_1\mu\otimes q_2\mu\,\middle|\,\bigotimes_i p_iUp\mu\right)&=\mathrm{Ent}\left(q_1\mu\,\middle|\,\bigotimes_{i=1,\ldots,m} p_iUp\mu\right)\\&\quad+\mathrm{Ent}\left(q_2\mu\,\middle|\,\bigotimes_{i=m+1,\ldots,n} p_iUp\mu\right).\end{aligned}$$

Taking suprema and limits yields the inequality.

(3) With a similar notation as in the previous point, we take $\mu=U^G\mu$ in $A_{N,\epsilon,K}(\tau)$, then:

$$\mathrm{Ent}(p\mu|Leb)=\mathrm{Ent}(p\mu|q_1V\mu\otimes q_2V\mu)+\mathrm{Ent}(q_1\mu\otimes q_2\mu|Leb)$$

and again we may take suprema and limits to get the required conclusion.

(4) This follows from Theorem 7.3(6), as well as Theorem 9.9(4) and Proposition 9.10 to be proved below. □

9.4. I-mutual entropy

In order to extend again in this subsection [8] for (not necessarily hyperfinite) multivariables, we consider multivariables $\mathbf{X_i} = (\mathbf{X_{i1}}, \ldots, \mathbf{X_{iP_i}})$ where each $\mathbf{X_{ij}}$ is itself a family of hyperfinite multivariables, i.e. $\mathbf{X_{ij}} = \{X_{ij1}, \ldots, X_{ijQ_{ij}}\}$ and $\tilde{P}_i = \sum_{j=1}^{P_i} Q_{ij}, \bar{n} = \sum_{i=1}^{n} \tilde{P}_i$. For the definition of free entropy in presence we consider also analogously $\mathbf{Y} = (\mathbf{Y_1}, \ldots, \mathbf{Y_P})$ containing $\bar{t}$ variables. For technical reasons (in order to get values agreeing with those of [8] in the hyperfinite case) we will let the approximations depend on doubled parameters $\epsilon = (\epsilon_1, \epsilon_2)$, $K = (K_1, K_2)$.

We first consider $\sigma_{ij} = \sigma_{N,\epsilon_2,K_2}(\tau_{\mathbf{X_{ij}}})$ the normalized restriction of Lebesgue measure to the set $V_{\epsilon_2,K_2}(p_{ij}\tau)$ of ϵ_2, K_2 approximations of $p_{ij}\tau = \tau_{\mathbf{X_{ij}}}$ where p_{ij} gives the marginals on the ij-th bunch of hyperfinite variables. We denote by $C_{N,\epsilon,K}(\tau)$ Csiszar's I-projection of $S_{N,\epsilon_2,K_2}(\tau) := \bigotimes_{i,j} \sigma_{ij}$ on:

$$A_{N,\epsilon,K}(\tau) = \{\mu \in P((H_R^N)^{\bar{n}+\bar{t}}) \mid \tau_\mu \in V_{\epsilon_1,K_1}(\tau)\ \forall i, j\ p_{ij}\tau_\mu \in V_{\epsilon_2,K_2}(p_{ij}\tau)\}.$$

Thus, we allow us to approximate better the hyperfinite marginals. This will be used to define an I-mutual entropy with good additivity properties, which was a motivation for Voiculescu's non-microstates mutual information and for Hiai–Miyamoto–Ueda's microstate variant. However the other variants seem to be better behaved in every other respects. We will use not only a free ultrafilter ω on the integers but also a point θ in the boundary of the Stone–Čech compactification of $(0, 1]$. If $A_{N,\epsilon,K}(\tau)$ does not contain elements of finite entropy, any entropy involving $C_{N,\epsilon,K}$ (thus undefined) is by convention $-\infty$. Likewise, a sup over an empty set is $-\infty$.

Definition 9.6. Let $\tau = \tau_{\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y}}$, we define *I-mutual entropy* as

$$\tilde{\chi}_{Imut}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) = \sup_{R \geq \mathcal{R}(\tau)} \limsup_{\epsilon_1 \to 0} \limsup_{K_1 \to \infty} \limsup_{\epsilon_2 \to 0} \limsup_{K_2 \to \infty} \limsup_{N \to \infty} \times \left(\frac{1}{N^2} \mathrm{Ent}(pC_{N,\epsilon,K}(\tau) | p_1 U p C_{N,\epsilon,K}(\tau) \otimes \cdots \otimes p_n U p C_{N,\epsilon,K}(\tau)) \right),$$

where p_i is the projection on submultivariables $\mathbf{X_i}$ and p on $\mathbf{X_1}, \ldots, \mathbf{X_n}$. We write $\tilde{\chi}_{Imut}(\mathbf{X_1}; \ldots; \mathbf{X_n})$ when $\mathbf{Y} = \emptyset$. Likewise we define $\underline{\tilde{\chi}}_{Imut}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y})$ a liminf variant (with respect to N, ϵ, K) of I-mutual entropy and an ultrafilter variant $\tilde{\chi}_{Imut}^{\omega,\theta}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y})$ (with $\lim_{1/R \to \theta} \lim_{\epsilon_1 \to \theta} \lim_{K_1 \to \omega} \lim_{\epsilon_2 \to \theta} \lim_{K_2 \to \omega} \lim_{N \to \omega}$).

We will also need a notion of free I-entropy in the presence of other variables to get additivity properties with I-mutual entropy. Instead of maximizing the entropy of the projection of measures also approximating $\mathbf{Y}$, which would be more natural in the spirit of Voiculescu's definition and correspond to the definition taken in the previous subsection, we take Csiszar's projection including approximation of $\mathbf{Y}$, we project and take entropy.

Definition 9.7. We define *free I-entropy in the presence* of $\mathbf{Y}$ as:

$$\tilde{\chi}_I(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) = \sup_{R \geq \mathcal{R}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y}})} \limsup_{K_1 \to \infty, \epsilon_1 \to 0} \limsup_{K_2 \to \infty, \epsilon_2 \to 0} \limsup_{N \to \infty} \times \left(\frac{1}{N^2} \mathrm{Ent}(pC_{N,\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n},\mathbf{Y}})) + \frac{n}{2} \log N \right),$$

and likewise $\tilde{\chi}_I^{\omega,\theta}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y})$, $\underline{\tilde{\chi}}_I(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y})$.

We have inequalities, as in Sections 4 and 6, given in the following lemma.

Lemma 9.8. *We have:*

$$\tilde{\chi}_I(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y}) \le \tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y}),$$
$$\chi(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) \le \tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y}),$$
$$\chi(\mathbf{X_1}; \ldots; \mathbf{X_n}) \le \tilde{\chi}_I(\mathbf{X_1}, \ldots, \mathbf{X_n}),$$

and corresponding ultrafilter, liminf variants.

If $\tau_{\mathbf{X_1};\ldots;\mathbf{X_n},\mathbf{Y}}$ is extremal we also have:

$$\chi(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y}) = \tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y}).$$

Especially, if $\tau_{\mathbf{X_1};\ldots;\mathbf{X_n}}$ is extremal we have:

$$\chi(\mathbf{X_1}; \ldots; \mathbf{X_n}) = \tilde{\chi}_I(\mathbf{X_1}; \ldots; \mathbf{X_n}) = \tilde{\chi}(\mathbf{X_1}; \ldots; \mathbf{X_n}).$$

Proof. Let $\tau = \tau_{\mathbf{X_1};\ldots;\mathbf{X_n}:\mathbf{Y}}$ Since $C_{N,\epsilon,K} \in A_{N,(\epsilon_1,\epsilon_1),(K_1,K_1)}(\tau)$ by definition we obtain $\tilde{\chi}_I(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) \le \tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y})$.

The inequalities between χ and $\tilde{\chi}$ are similar to those in Sections 4 and 6. Let us merely outline the proofs for the reader's convenience. First, recall Voiculescu's definition from [18]:

$$\chi(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) = \sup_{R \ge \mathcal{R}(\tau)} \lim_{K\to\infty,\epsilon\to 0} \limsup_{N\to\infty} \left(\frac{1}{N^2} \log(p\Gamma_R(\tau,\epsilon,K,N)) + \frac{n}{2}\log N \right),$$

where $pA \in (H_R^N)^{\bar{n}}$ is now the projection of the set $A \in (H_R^N)^{\bar{n}+t}$.

Fix $\epsilon, K > 0$. For $\mathbf{M} \in p\Gamma_R(\tau,\epsilon,K,N)$, we consider the fiber:

$$\Gamma_{R,\mathbf{M}} = (\{\mathbf{M}\} \times (H_R^N)^t) \cap \Gamma_R(\tau,\epsilon,K,N).$$

We define a probability measure μ with support in $\Gamma_R(\tau,\epsilon,K,N)$ (so that $\tau_\mu \in V_{\epsilon,K}(\tau)$), on a measurable set $A \in (H_R^N)^{\bar{n}+t}$ by:

$$\mu(A) = \frac{1}{\mathrm{Leb}(p\Gamma_R(\tau,\epsilon,K,N))} \times \int_{p\Gamma_R(\tau,\epsilon,K,N)} d\mathrm{Leb}_{(H_R^N)^{\bar{n}}}(\mathbf{M}) \frac{(\delta_{\mathbf{M}} \times \mathrm{Leb}_{(H_R^N)^t})(A \cap \Gamma_{R,\mathbf{M}})}{(\delta_{\mathbf{M}} \times \mathrm{Leb}_{(H_R^N)^t})(\Gamma_{R,\mathbf{M}})}.$$

By definition, we get $p\mu(B) = \frac{1}{\mathrm{Leb}(p\Gamma_R(\tau,\epsilon,K,N))}\mathrm{Leb}(B \cap p\Gamma_R(\tau,\epsilon,K,N))$, so that:

$$\log(p\Gamma_R(\tau,\epsilon,K,N)) = \mathrm{Ent}(p\mu) \le \sup_{\mu\in A_{N,\epsilon,K}(\tau)} \mathrm{Ent}(p\mu).$$

We conclude $\chi(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) \le \tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y})$.

Conversely, assume τ extremal. Fix $\eta, \epsilon, K > 0$ and choose δ, L as in Lemma 6.1 so that, if $\mu \in A_{N,\delta,L}(\tau)$, $\mu(\Gamma_R(\tau,\epsilon,K,N)) \ge 1-\eta$. Note that we have:

$$p\mu(p\Gamma_R(\tau,\epsilon,K,N)) = \mu(p\Gamma_R(\tau,\epsilon,K,N) \times (H_R^N)^t) \ge \mu(\Gamma_R(\tau,\epsilon,K,N)) \ge 1-\eta.$$

Thus, as in Proposition 6.2, we get $\tilde{\chi}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y}) \le (1-\eta)\chi(\mathbf{X_1}, \ldots, \mathbf{X_n} : \mathbf{Y})$.

Consider now the case without $\mathbf{Y}$, the only remaining inequality is $\chi(\mathbf{X_1}; \dots; \mathbf{X_n}) \leq \tilde{\chi}_I(\mathbf{X_1}, \dots, \mathbf{X_n})$. First, note that:

$$\mathrm{Ent}(C_{N,\epsilon,K}) = \mathrm{Ent}(C_{N,\epsilon,K}|S_{N,\epsilon_2,K_2}) + \mathrm{Ent}(S_{N,\epsilon_2,K_2}).$$

Indeed by its definition as I-projection of the measure S_{N,ϵ_2,K_2}, we know that $C_{N,\epsilon,K}$ has a density with respect to S_{N,ϵ_2,K_2}, and since S_{N,ϵ_2,K_2} is Lebesgue measure normalized on some set, the density with respect to Lebesgue measure does not change except for a constant and the equality above is thus easy.

We can also consider R_{N,ϵ_2,K_2} the normalized Lebesgue measure on $\Gamma_R(\tau, \epsilon_2, K_2, N)$ so that:

$$\begin{aligned}\log(\mathrm{Leb}(\Gamma_R(\tau, \epsilon_2, K_2, N))) &= \mathrm{Ent}(R_{N,\epsilon_2,K_2}) = \mathrm{Ent}(R_{N,\epsilon_2,K_2}|S_{N,\epsilon_2,K_2}) \\ &\quad + \mathrm{Ent}(S_{N,\epsilon_2,K_2}),\end{aligned}$$

the last equality coming from inclusion of the support of R in the support of S, both being normalized Lebesgue measure on subsets. Finally, by definition of I-projection, we get the inequality:

$$\mathrm{Ent}(C_{N,\epsilon,K}|S_{N,\epsilon_2,K_2}) \geq \mathrm{Ent}(R_{N,\epsilon_2,K_2}|S_{N,\epsilon_2,K_2}).$$

As a consequence, we also get:

$$\frac{1}{N^2}\log(\mathrm{Leb}(\Gamma_R(\tau, \epsilon_2, K_2, N))) + \frac{n}{2}\log N \leq \frac{1}{N^2}\mathrm{Ent}(C_{N,\epsilon,K}) + \frac{n}{2}\log N,$$

and we can take successively limits in $N, K_2, \epsilon_2, K_1, \epsilon_1, R$ to conclude.

Note that it is not obvious that in general we could have $\chi(\mathbf{X_1}; \dots; \mathbf{X_n} : \mathbf{Y}) \leq \tilde{\chi}_I(\mathbf{X_1}, \dots, \mathbf{X_n} : \mathbf{Y})$. $\square$

Theorem 9.9. (1) *(Vanishing for one variable)*

$$\tilde{\chi}^a_{Imut}(\mathbf{X_1}) = 0,$$

for $\mathbf{X_1}$ having finite-dimensional approximants.

(2) *(Improved Subadditivity)*

$$\begin{aligned}&\tilde{\chi}_{Imut}(\mathbf{X_1}; \dots; \mathbf{X_m}; \mathbf{X_{m+1}}; \dots; \mathbf{X_n} : \mathbf{Y}) \\ &\quad\leq \tilde{\chi}_{Imut}(\mathbf{X_1} \cup \dots \cup \mathbf{X_m}; \mathbf{X_{m+1}} \cup \dots \cup \mathbf{X_n} : \mathbf{Y}) \\ &\qquad + \tilde{\chi}_{Imut}(\mathbf{X_1}; \dots; \mathbf{X_m} : \mathbf{X_{m+1}} \cup \dots \cup \mathbf{X_n} \cup \mathbf{Y}) \\ &\qquad + \tilde{\chi}_{Imut}(\mathbf{X_{m+1}}; \dots; \mathbf{X_n} : \mathbf{X_1} \cup \dots \cup \mathbf{X_m} \cup \mathbf{Y}), \\ &\tilde{\chi}^{\omega,\theta}_{Imut}(\mathbf{X_1}; \dots; \mathbf{X_m}; \mathbf{X_{m+1}}; \dots; \mathbf{X_n} : \mathbf{Y}) \\ &\quad= \tilde{\chi}^{\omega,\theta}_{Imut}(\mathbf{X_1} \cup \dots \cup \mathbf{X_m}; \mathbf{X_{m+1}} \cup \dots \cup \mathbf{X_n} : \mathbf{Y}) \\ &\qquad + \tilde{\chi}^{\omega,\theta}_{Imut}(\mathbf{X_1}; \dots; \mathbf{X_m} : \mathbf{X_{m+1}} \cup \dots \cup \mathbf{X_n} \cup \mathbf{Y}) \\ &\qquad + \tilde{\chi}^{\omega,\theta}_{Imut}(\mathbf{X_{m+1}}; \dots; \mathbf{X_n} : \mathbf{X_1} \cup \dots \cup \mathbf{X_m} \cup \mathbf{Y}).\end{aligned}$$

(3) *(Improved subadditivity of entropy)*

$$\begin{aligned}&\tilde{\chi}_I(\mathbf{X_1}, \mathbf{X_2} : \mathbf{Y}) \leq \tilde{\chi}_{Imut}(\mathbf{X_1}; \mathbf{X_2} : \mathbf{Y}) + \tilde{\chi}_I(\mathbf{X_1} : \mathbf{X_2} \cup \mathbf{Y}) + \tilde{\chi}_I(\mathbf{X_2} : \mathbf{X_1} \cup \mathbf{Y}), \\ &\tilde{\chi}^{\omega,\theta}_I(\mathbf{X_1}, \mathbf{X_2} : \mathbf{Y}) = \tilde{\chi}^{\omega,\theta}_{Imut}(\mathbf{X_1}; \mathbf{X_2} : \mathbf{Y}) + \tilde{\chi}^{\omega,\theta}_I(\mathbf{X_1} : \mathbf{X_2} \cup \mathbf{Y}) + \tilde{\chi}^{\omega,\theta}_I(\mathbf{X_2} : \mathbf{X_1} \cup \mathbf{Y}).\end{aligned}$$

(4) *(Agreement with previous definition)*
If $\mathbf{X_i}$ *are hyperfinite multivariables (more accurately* $\mathbf{P_i} = 1$*) then*

$$\chi_{orb}(\mathbf{X_1}; \ldots ; \mathbf{X_n}) \leq \tilde{\chi}_{Imut}(\mathbf{X_1}; \ldots ; \mathbf{X_n}).$$

(χ_{orb} *in the sense of* [8]*). If moreover* $\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}}$ *is extremal then*

$$\tilde{\chi}_{Imut}(\mathbf{X_1}; \ldots ; \mathbf{X_n}) = \chi_{orb}(\mathbf{X_1}; \ldots ; \mathbf{X_n}).$$

Proof. (1) Similar to $\tilde{\chi}_{orb}$.

(2), (3) These follow from equalities in the corresponding proofs for $\tilde{\chi}_{Mmut}$.

(4) After using Theorem 7.3(6) in the case of extremality and relating inequalities of our variants (Proposition 9.10), it remains to prove: $\chi_{orb}(\mathbf{X_1}; \ldots ; \mathbf{X_n}) \leq \tilde{\chi}_{Imut}(\mathbf{X_1}; \ldots ; \mathbf{X_n})$.

We take notations of [8] especially $\Xi_i(N)$ (as in Lemma 4.2 and Definition 4.1 there) is a sequence approximating the hyperfinite variables $\mathbf{X_i}$ in mixed moments. We now show that, for every ϵ_1, K_1, there exists δ, L such that, for every $\epsilon = (\epsilon_1, \epsilon_2), \epsilon_2 \leq \delta,\ K = (K_1, K_2), K_2 \geq L$:

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{N,\Xi(N),\epsilon_1/2,K} \leq \limsup_{N\to\infty} \frac{1}{N^2} \mathrm{Ent}(C_{N,\epsilon,K} | D_{N,\epsilon,K}),$$

where $C_{N,\epsilon,K} = C_{N,\epsilon,K}(\tau_{\mathbf{X_1},\ldots,\mathbf{X_n}})$, $D_{N,\epsilon,K} = p_1 UpC_{N,\epsilon,K} \otimes \cdots \otimes p_n UpC_{N,\epsilon,K} = p_1 C_{N,\epsilon,K} \otimes \cdots \otimes p_n C_{N,\epsilon,K}$. First, we use Jung's Lemma and follow the proof of Lemma 4.2 in [8]. We can thus take δ, L such that, for all families of sets $(\Theta_i)_{i=1,\ldots,n}$ of $N \times N$ hermitian matrices, for N large enough, with $\tau_{(\Theta_i)} \in V_{\delta,L}(p_i\tau)$ for all i, we have:

$$\gamma_{N,\Theta,\epsilon_1,K_1} \geq \gamma_{N,\Xi(N),\epsilon_1/2,K_1}. \tag{9.1}$$

Moreover, using again Lemma 8.1(ii),

$$\begin{aligned}\mathrm{Ent}(C_{N,\epsilon,K} | D_{N,\epsilon,K}) &= \mathrm{Ent}(C_{N,\epsilon,K} | S_{N,\epsilon_2,K_2}(\tau)) - \mathrm{Ent}(D_{N,\epsilon,K} | S_{N,\epsilon_2,K_2}(\tau)) \\ &\geq \mathrm{Ent}(C_{N,\epsilon,K} | S_{N,\epsilon_2,K_2}(\tau)).\end{aligned}$$

In order to use the definition of Csizar's projection, we have to take a specific measure in $A_{N,\epsilon,K}$. Note that we have considered Csizar's projection with respect to $S_{N,\epsilon_2,K_2}(\tau)$, in order to have a measure with support included in a set where hyperfinite variables for marginals will be of the form Ξ', for which we can apply the relation (9.1) above. Let

$$dT_{N,\epsilon,K}(\Xi') = \frac{1_{\Xi' \in \Gamma_R(\mathbf{X_1},\ldots,\mathbf{X_n},N,K,\epsilon_1)}}{\gamma_{N,\Xi',\epsilon_1,K_1}} d(S_{N,\epsilon_2,K_2}(\tau))(\Xi').$$

This is a probability measure: since S_{N,ϵ_2,K_2} is an $U(N)^n$ invariant probability we can compute the total mass by integrating the density over unitaries and by definition

$$\begin{aligned}\mathcal{H}_N^n(1_{U\Xi'U^* \in \Gamma_R(\mathbf{X},N,K_1,\epsilon_1)}) &= \mathcal{H}_N^n(\Gamma_{orb}(\mathbf{X_1}, \ldots, \mathbf{X_n} : \Xi_1', \ldots \Xi_n', N, K_1, \epsilon_1)) \\ &= \gamma_{N,\Xi',\epsilon_1,K_1}.\end{aligned}$$

From this and since its support is in $\Gamma_R(\mathbf{X_1}, \ldots, \mathbf{X_n}, N, K_1, \epsilon_1)$ we deduce that $T_{N,\epsilon,K} \in A_{N,\epsilon,K}$.

It follows, by definition of C as Csiszar's projection of S, that

$$\begin{aligned}\mathrm{Ent}(C_{N,\epsilon,K} | S_{N,\epsilon_2,K_2}) &\geq \mathrm{Ent}(T_{N,\epsilon,K} | S_{N,\epsilon_2,K_2}) = T_{N,\epsilon,K}(\log(\gamma_{N,.,\epsilon_1,K_1})) \\ &\geq T_{N,\epsilon,K}(\log(\gamma_{N,\Xi(N),\epsilon_1/2,K_1})) = \log(\gamma_{N,\Xi(N),\epsilon_1/2,K_1}).\end{aligned}$$

The second inequality comes from (9.1) since $\epsilon_2 \leq \delta$, $K_2 \geq L$. This concludes. $\square$

9.5. Comparison of the various entropies

Beyond the case of equality in the context of [8], we have the following general inequality.

Proposition 9.10 (*Relating Inequalities*).

$$\tilde{\chi}_{Imut}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) \leq \tilde{\chi}_{Mmut}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) \leq \tilde{\chi}_{orb}(\mathbf{X_1}; \ldots; \mathbf{X_n} : \mathbf{Y}) \leq 0.$$

Proof. Negativity comes from negativity of relative entropy. The first inequality follows from $C_{N,\epsilon,K}(\tau) \in A_{N,\epsilon,K}(\tau) \subset A_{N,\epsilon_1,K_1}(\tau)$ (for $K_1 \leq K_2, \epsilon_1 \geq \epsilon_2$) and our conventions in case this is empty.

Finally, applying Lemma 8.1(i) for any $\mu \in A_{N,\epsilon,K}(\tau)$ we get the inequality:

$$\mathrm{Ent}(p\mu | p_1 U p\mu \otimes \cdots \otimes p_n U p\mu) \leq \mathrm{Ent}(p\mu | U p\mu).$$

The second inequality follows. □

Added in proof: Y. Ueda now developed in [23] a way of defining orbital entropy for non-hyperfinite algebras without using random microstates, as suggested by our alternative formula in Theorem 7.3(7). The same additivity issue appears in this new approach but it enables to develop an interesting orbital entropy dimension for non-hyperfinite algebras.

Acknowledgments

The authors want to thank respectively U. Haagerup and D. Shlyakhtenko for fruitful discussions. The authors are grateful to the Erwin Schrödinger Institute where part of this work has been completed. They also thank the organizers of the workshop on "Random Matrix, Operator Algebra, and Mathematical Physics Aspects" in the semester on Bialgebras in free Probability having taken place there. Finally the authors would like to thank an anonymous referee for some very helpful comments and suggestions.

The research of P. Biane was partially supported by ANR grant GranMa 08-BLAN-0311-01.

References

[1] S. Belinschi, On an equivalent definition of free entropy, Math. Res. Lett. 10 (2003) 185–189.

[2] P. Biane, D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, Geom. Funct. Anal. 11 (6) (2001) 1125–1138.

[3] B. Collins, Moments and Cumulants of Polynomial random variables on unitary groups, the Itzykson–Zuber integral and free probability, Int. Math. Res. Not. 17 (2003) 953–982.

[4] I. Csiszar, I-divergence geometry of probability distributions and minimization problems, Ann. Probab. 3 (1) (1975) 146–158.

[5] Y. Dabrowski, A note about proving non-Γ under a finite non-microstates free Fisher information assumption, J. Funct. Aanal. 258 (11) (2010) 3662–3674.

[6] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, in: Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010, Corrected reprint of the second (1998) edition.

[7] F. Hiai, Free analog of pressure and its Legendre transform, Comm. Math. Phys. 255 (1) (2005) 229–252.

[8] F. Hiai, T. Miyamoto, Y. Ueda, Orbital approach to microstate free entropy, Int. J. Math. 20 (2) (2009) 227–273.

[9] F. Hiai, D. Petz, The Semicircle Law, Free Random Variables and Entropy, in: Mathematical Surveys and Monographs, vol. 77, AMS, 2000.

[10] K. Jung, Amenability, tubularity, and embeddings into R^{ω}, Math. Ann. 338 (2007) 241–248.

[11] S. Kullback, Information Theory and Statistics, in: Wiley publications in Statistics, 1959.

[12] J. Lindenstrauss, G. Olsen, Y. Sternfeld, The Poulsen simplex, Ann. Inst. Fourier 28 (1) (1978) 91–114.

[13] M. Nagasawa, Schrödinger Equations and Diffusion Theory, Birkhäuser, Basel, 1993.

[14] F. Otto, C. Villani, Generalization of an inequality by talagrand and links with the logarithmic Sobolev inequality, J. Funct. Anal. 173 (2000) 361–400.

[15] S.J. Sidney, Exposed points for duals of separable Frechet spaces, in: Banach Spaces, Harmonic Analysis, and Probability Theory Proceedings of the Special Year in Analysis, Held at the University of Connecticut 1980–1981, 166–168.

[16] M. Takesaki, M. Theory of operator algebras. I, in: Encyclopaedia of Mathematical Sciences, 124, in: Operator Algebras and Non-commutative Geometry, vol. 5, Springer-Verlag, Berlin, 2002, Reprint of the first (1979) edition.

[17] C. Villani, Optimal Transport, Old and New, Springer, 2008.

[18] D.V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory III: the absence of Cartan subalgebras, Geom. Funct. Anal. 6 (1) (1986) 172–199.

[19] D.V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory II, Invent. Math. 118 (3) (1994) 411–440.

[20] D.V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory IV: maximum entropy and freeness, in: D.V. Voiculescu (Ed.), Free Probability Theory, in: Fields Inst. Commun., vol. 12, Amer. Math. Soc, 1997, pp. 293–302.

[21] D.V. Voiculescu, A strengthened asymptotic freeness result for random matrices with applications to free entropy, Int. Math. Res. Not. (1) (1998) 41–63.

[22] D.V. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory VI: liberation and mutual free information, Adv. Math. 146 (1999) 101–166.

[23] Y. Ueda, Orbital free entropy, revisited. Preprint 2012: arxiv:1210.6421.