Elementary proof of $\chi \leq \chi^*$

David Jekel and Jennifer Pi

University of California, San Diego

2023-06-01 Berkeley Free Probability and Operator Algebras Seminar

Introduction

Voiculescu defined several versions of free entropy for an m-tuple of self-adjoint operators:

- Free microstate entropy $\chi(X_1, \dots, X_m)$ looks at the Lebesgue measure of spaces of matrix approximations.
- Non-microstate entropy $\chi^*(X_1, \dots, X_m)$ looks at the interaction between X_1, ..., X_m and the free difference quotient and semicircular perturbations.
- Liberation mutual information and entropy is similar to χ^* but through a deformation to freely independent copies of the individual variables.

Voiculescu used χ to show that free group von Neumann algebras have no Cartan subalgebras, lack property Gamma, and more. Since then, there have been many similar applications for indecomposability results for free products.

Introduction

(Part of) the **unification problem for free entropy:** When does $\chi(X) = \chi^*(X)$?

- Biane, Capitaine, Guionnet 2003: $\chi \leq \chi^*$.
- Dabrowski 2017 and Jekel 2018-2020: $\chi(X) = \chi^*(X)$ when X satisfies a free Gibbs law for a convex potential with certain regularity / growth conditions.
- Ji, Natarajan, Vidick, Yuen, and Wright announced a negative solution of the Connes embedding problem. This implies that there exist some X with $\chi(X) = -\infty$ and $\chi^*(X) > -\infty$.

Introduction

We will give an elementary proof that $\chi \leq \chi^*$ as well as generalize it to the conditional setting. The talk will be organized as follows:

Part 1:

- Conditional entropies of Voiculescu and Shlyakhtenko $\chi(X \mid \mathcal{B})$.
- Conditional microstates entropy w.r.t. $\iota : \mathcal{B} \to \prod_{n \to \mathcal{U}} M_n(\mathbb{C})$.
- Relating $\chi(X \mid \mathcal{B}, \iota)$ with Shlyakhtenko's version.
- Relating conditional microstate entropy and classical entropy.

Part 2:

- Conditional non-microstates entropy of Voiculescu.
- Classical Fisher information and free Fisher information.
- Proof of $\chi \leq \chi^*$.

Laws and Microstate Spaces

$\Sigma_{m,R}$ will denote the space of non-commutative laws, that is, linear maps $\mu : \mathbb{C}\langle x_1, \dots, x_m \rangle \to \mathbb{C}$ satisfying

- $\mu(1) = 1$,
- $\mu(p^* p) \geq 0$,
- $\mu(pq) = \mu(qp)$,
- $|\mu(x_{i_1} \dots x_{i_k})| \leq R^k$ for i_1, ..., $i_k \in [m]$.

Given (M, τ) a tracial von Neumann algebra and $X = (X_1, \dots, X_m)$ self-adjoint from M, the law of X is $\mu_X(p) = \tau(p(X))$.

We equip $\Sigma_{m,R}$ with the weak-$*$ topology. This topology captures "convergence in moments." It will be used to define what we mean by matrix approximation.

Note this definition also makes sense for $m = \infty$.

Conditional microstate spaces

We will consider conditional microstate spaces. Here $X = (X_1, \dots, X_m)$ and $Y = (Y_1, Y_2, \dots)$ will represent self-adjoint elements of some (M, τ) and $\mathbb{C}\langle \mathbf{x}, \mathbf{y} \rangle$ is the formal polynomial $*$-algebra.

Given a neighborhood $\mathcal{O}$ of $\mathrm{law}(X, Y)$ and a tuple $Y^{(n)}$ from $M_n(\mathbb{C})$, we set

$$\Gamma_R^{(n)}(\mathcal{O} \mid Y^{(n)} \rightsquigarrow Y) = \{X^{(n)} \in M_n(\mathbb{C})_{\mathrm{sa}}^m : \mathrm{law}(X^{(n)}, Y^{(n)}) \in \mathcal{O}\}$$

This represents the space of matrix approximations for X that are compatible with a *given* choice of matrix approximations for Y.

Voiculescu's conditional free entropy

Lebesgue measure: $(M_n(\mathbb{C})^d, \langle \cdot, \cdot \rangle_{\mathrm{tr}_n})$ is linearly isometric to $\mathbb{C}^{dn^2}$, so transfer the Lebesgue measure through the isometry.

Here we assume that Y is a finite tuple and $\|X_j\|$, $\|Y_j\| \leq R$. Let $\mathcal{O}(k, \epsilon)$ be the microstate space given by the conditions that the moments up to order k are within ϵ. Then

$$\chi_{\mathrm{avg}}(X \mid Y) := \inf_{k,\epsilon} \limsup_{n \to \infty} \fint_{Y^{(n)} \in \Gamma_R^{(n)}(\mathcal{O}_Y(k,\epsilon))} \left(\frac{1}{n^2} \log \mathrm{vol}\, \Gamma^{(n)}(\mathcal{O}_{X,Y}(k,\epsilon) \mid Y^{(n)} \rightsquigarrow Y) + m \log n \right).$$

This gives us the *average* of the normalized Lebesgue measure of conditional microstate spaces.

Shlyakhtenko's conditional free entropy

As usual, you can take the sup over R at the end, but the answer is already correct if R is larger than the norms of the X_j's and Y_j's.

Instead of taking the average, Shlyakhtenko took the supremum of the measures over all the choices of $Y^{(n)}$. This approach adapts easily to Y being an infinite tuple.

$$\overline{\chi}(X \mid Y) := \inf_{k,\epsilon} \limsup_{n \to \infty} \sup_{Y^{(n)} \in \Gamma_R^{(n)}(\mathcal{O}_Y(k,\epsilon))} \left(\frac{1}{n^2} \log \operatorname{vol} \Gamma^{(n)}(\mathcal{O}_{X,Y}(k,\epsilon) \mid Y^{(n)} \rightsquigarrow Y) + m \log n \right).$$

Conditional entropy for a fixed choice of $Y^{(n)}$

Rather than taking the supremum or average, we can fix a choice of microstates for Y. Then we would have a conditional entropy that depends on the choice of microstates.

Fix a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$ and a sequence of microstates $Y^{(n)}$ such that $\lim_{n \to \mathcal{U}} \operatorname{law}(Y^{(n)}) = \operatorname{law}(Y)$. Then set

$$\chi^{\mathcal{U}}(X \mid Y^{(n)} \rightsquigarrow Y) := \inf_{\mathcal{O}} \limsup_{n \to \infty} \left(\frac{1}{n^2} \log \operatorname{vol} \Gamma^{(n)}(\mathcal{O} \mid Y^{(n)} \rightsquigarrow Y) + m \log n \right)$$

Conditional entropy for an embedding

Let $\mathcal{B} = \mathrm{W}^*(Y)$. The choice of microstates $Y^{(n)}$ gives rise to an embedding $\iota : \mathcal{B} \to \prod_{n\to\mathcal{U}} M_n(\mathbb{C})$ given by $\iota(Y_j) = [Y_j^{(n)}]$.

Lemma (J&P)

Suppose $\mathrm{W}^*(Y) = \mathrm{W}^*(Z) = \mathcal{B}$. If $Y^{(n)} \rightsquigarrow Y$ and $Z^{(n)} \rightsquigarrow Z$ give rise to the same embedding $\iota : \mathcal{B} \to \prod_{n\to\mathcal{U}} M_n(\mathbb{C})$, then $\chi^{\mathcal{U}}(X \mid Y^{(n)} \rightsquigarrow Y) = \chi^{\mathcal{U}}(X \mid Z^{(n)} \rightsquigarrow Z)$. Hence, we define $\chi^{\mathcal{U}}(X \mid \mathcal{B}, \iota)$ to be the common value for any such choice of microstates for $Y^{(n)}$.

Note Shlyakhtenko proved a similar invariance result over the choice of Y for his conditional entropy.

The proof goes by way approximating Y by polynomials (or more general functions) of Z and then getting inclusions of the relative microstate spaces.

Relationship with Shlyakhtenko's entropy

Proposition (J&P)

Let $(\mathcal{M}, \tau)$ be a tracial von Neumann algebra $X = (X_1, \ldots, X_m)$ self-adjoint elements with $\|X_j\| \leq R$.

$$\overline{\chi}(X \mid \mathcal{B}) = \sup_{\mathcal{U}, \iota} \chi^{\mathcal{U}}(X \mid \mathcal{B}, \iota).$$

This proposition represents exchanging a sup with a limit since on the right we take the supremum over $Y^{(n)}$ before taking the $\limsup$ as $n \to \infty$ and on the right, we first take an ultralimit and then take the sup over the choices of $Y^{(n)}$ and $\mathcal{U}$.

The proof (on the next slide) is conceptually a diagonalization argument.

Relationship with Shlyakhtenko's entropy

The easier direction is $\overline{\chi}(X \mid \mathcal{B}) \geq \sup_{\mathcal{U},\iota} \chi^{\mathcal{U}}(X \mid \mathcal{B}, \iota)$.

Indeed, for each $\mathcal{U}$ and ι, we have $\overline{\chi}(X \mid \mathcal{B}) \geq \chi^{\mathcal{U}}(X \mid \mathcal{B}, \iota)$ because the measure of the relative microstate space for a particular choice of $Y^{(n)}$ is less than or equal to the sup, and ultralimit is less than or equal to lim sup.

For the other direction, let $\mathcal{O}_k$ be a sequence of neighborhoods of $\operatorname{law}(\mathbf{X}, \mathbf{Y})$ in Σ_R such that $\mathcal{O}_0 = \Sigma_R$ and $\mathcal{O}_{k+1} \subseteq \mathcal{O}_k$ and $\bigcap_{k \in \mathbb{N}} \mathcal{O}_k = \{\operatorname{law}(\mathbf{X}, \mathbf{Y})\}$.

Relationship with Shlyakhtenko's entropy

Define $A_0 = \mathbb{N}$ and for $k \geq 1$,

$$A_k = \Bigg\{ n \geq k : \frac{1}{n^2} \log \sup_{\mathbf{Z}^{(n)} \in \Gamma_R^{(n)}(\pi_2(\mathcal{O}_k))} \operatorname{vol}(\Gamma_R^{(n)}(\mathcal{O}_k \mid \mathbf{Z}^{(n)} \rightsquigarrow \mathbf{Y})) + m \log n > \overline{\chi}_R(\mathbf{X} \mid \mathbf{Y}) - \frac{1}{k} \Bigg\}$$

Note $A_{k+1} \subseteq A_k$ since $\mathcal{O}_{k+1} \subseteq \mathcal{O}_k$ and $1/k > 1/(k+1)$.

A_k is nonempty by definition of $\overline{\chi}_R(\mathbf{X} \mid \mathbf{Y})$. But $\bigcap_{k \in \mathbb{N}} A_k = \varnothing$.

So $\exists \mathcal{U} \in \beta\mathbb{N} \setminus \mathbb{N}$ such that $A_k \in \mathcal{U}$ for all k.

Relationship with Shlyakhtenko's entropy

For each $n \in A_k \setminus A_{k+1}$, let $\mathbf{Y}^{(n)} \in \Gamma_R^{(n)}(\pi_2(\mathcal{O}_k))$ such that

$$\frac{1}{n^2} \log \operatorname{vol}(\Gamma_R^{(n)}(\mathcal{O}_k \mid \mathbf{Y}^{(n)} \rightsquigarrow \mathbf{Y})) + m \log n > \overline{\chi}_R(\mathbf{X} \mid \mathbf{Y}) - \frac{1}{k}.$$

Note that $\lim_{n \to \mathcal{U}} \operatorname{law}(\mathbf{Y}^{(n)}) = \operatorname{law}(\mathbf{Y})$.

Using monotonicity, in fact for $n \in A_k$, the same inequality holds.

So

$$\lim_{n \to \mathcal{U}} \frac{1}{n^2} \log \operatorname{vol}(\Gamma_R^{(n)}(\mathcal{O}_k \mid \mathbf{Z}^{(n)} \rightsquigarrow \mathbf{Y})) + m \log n \geq \overline{\chi}_R(\mathbf{X} \mid \mathbf{Y}) - \frac{1}{k}.$$

Taking the limit as $k \to \infty$ gives us $\chi^{\mathcal{U}}(\mathbf{X} \mid \mathcal{B}, \iota) \geq \overline{\chi}(\mathbf{X} \mid \mathcal{B})$.

Classical entropy

If $\mathbf{X}$ is a random variable in $\mathbb{R}^d$ with probability density ρ, then $h(\mathbf{X}) = -\int \rho \log \rho \, dx$.

Note if $\rho \, dx$ is the uniform measure on a set S, then $h(\rho) = \log \operatorname{vol}(S)$. Also, an argument from Jensen's inequality shows that this is the largest entropy out of all probability measures supported on S.

In the non-conditional setting, Voiculescu's microstate free entropy is basically the limit of normalized classical entropies of uniform measures on microstate spaces.

This leads to the principle that the microstate entropy is the supremum of limits of classical entropies of random matrix models (J., appendix to Shlyakhtenko and Tao's paper). This will be a key ingredient in our argument.

Classical entropy

Let us consider this in the conditional setting.

Recall that for random variable (X, Y) in $\mathbb{R}^d \times \Omega$ with $d\mu(x,y) = \rho(x,y)\,dx\,d\sigma(y)$ the conditional entropy is defined as

$$h(X \mid Y) = \int_\Omega \left(\int_{\mathbb{R}^d} -\rho(x,y) \log \rho(x,y)\,dx \right) d\sigma(y).$$

This is the average of the classical entropies of the conditional distribution of **X** given **Y**.

If (X, Y) has the uniform distribution on some set S, then we get

$$h(X \mid Y) \int_\Omega \log \operatorname{vol}(S_y)\,d\sigma(y).$$

Remark: Voiculescu's conditional entropy is the limit of quantities like this.

Classical entropy and free entropy

Now let's state the relationship between conditional χ and classical entropy. First, we handle $\chi^{\mathcal{U}}(\mathbf{X} \mid \mathcal{B}, \iota)$.

Proposition (J&P)

Suppose $\mathcal{B} = \mathrm{W}^*(\mathbf{Y})$, and fix microstates $\mathbf{Y}^{(n)}$ for $\mathbf{Y}$ which induce an embedding ι. Then $\chi^{\mathcal{U}}(\mathbf{X} \mid \mathcal{B}, \iota)$ is the supremum of

$$\lim_{n \to \mathcal{U}} \left(\frac{1}{n^2} h(\mathbf{X}^{(n)}) + m \log n \right)$$

over all random matrix tuples $\mathbf{X}^{(n)}$ satisfying

- $\operatorname{law}(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \to \operatorname{law}(\mathbf{X}, \mathbf{Y})$ in probability as $n \to \mathcal{U}$.
- $\lim_{n \to \mathcal{U}} \|\mathbf{X}^{(n)}\|_\infty \leq R$ in probability.
- (Tail bound) There are $C > 0$ and $K > 0$ such that for $n \in \mathbb{N}$,

$$\mathbb{P}(\|\mathbf{X}^{(n)}\|_2 \geq C + \delta) \leq e^{-Kn^2\delta^2} \text{ for all } \delta > 0.$$

Classical entropy and free entropy

The reason we used (2) and (3) instead of just assuming it is bounded in operator norm is that we want to apply this proposition to Gaussian perturbations of our models. These will not be bounded in operator norm, but will have good tail bounds like (3).

In order to show that for such $\mathbf{X}^{(n)}$,

$$\lim_{n \to \mathcal{U}} \left(\frac{1}{n^2} h(\mathbf{X}^{(n)}) + m \log n \right) \leq \chi^{\mathcal{U}}(\mathbf{X} \mid \mathcal{B}, \iota),$$

we estimate the contribution to h from the "tail" by a partitioning argument. Then we compare the contribution from the part supported on a microstate space by comparing with the uniform measure on the microstate space.

For the other direction, we can find a sequence of random matrices whose classical entropies realize $\chi^{\mathcal{U}}(\mathbf{X} \mid \mathcal{B}, \iota)$ by choosing uniform measures on microstate spaces by a diagonalization argument.

Classical entropy and free entropy

Now we get a similar characterization for $\overline{\chi}(\mathbf{X} \mid \mathcal{B})$.

Proposition (J&P)

Suppose $\mathcal{B} = \mathrm{W}^*(\mathbf{Y})$. Then $\overline{\chi}^{\mathcal{U}}(\mathbf{X} \mid \mathcal{B})$ is the supremum of

$$\lim_{n \to \mathcal{U}} \left(\frac{1}{n^2} h(\mathbf{X}^{(n)} \mid \mathbf{Y}^{(n)}) + m \log n \right)$$

over all random matrix tuples $\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}$ satisfying

- $\mathrm{law}(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \to \mathrm{law}(\mathbf{X}, \mathbf{Y})$ in probability as $n \to \mathcal{U}$.
- $\lim_{n \to \mathcal{U}} \|(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})\|_\infty \leq R$ almost surely.
- (Tail bound) There are $C > 0$ and $K > 0$ such that for $n \in \mathbb{N}$ and for all y,

$$\mathbb{P}(\|\mathbf{X}^{(n)}\|_2 \geq C + \delta \mid \mathbf{Y}^{(n)} = y) \leq e^{-Kn^2\delta^2} \text{ for all } \delta > 0.$$

We can deduce this from the previous statement. The $\leq$ argument is based on the fact that for almost every choice of $\mathbf{Y}^{(n)}$'s, the $\mathbf{X}^{(n)}$'s conditional distribution satisfies the hypotheses of the previous proposition, and then we use a Fatou's lemma argument.

The $\geq$ direction follows because taking $\mathbf{Y}^{(n)}$ to be deterministic as in the previous proposition is a valid choice satisfying the conditions in this proposition too.