Entropy and Wasserstein distance in free probability

David Jekel
University of Copenhagen

IPAM long program on Non-commutative Optimal Transport
Workshop II Dynamics of Density Operators
2025-04-30

Notation

- $\mathbb{M}_n$ is the algebra of $n \times n$ complex matrices.
- $\mathrm{tr}_n = (1/n)\,\mathrm{Tr}_n$ is the normalized trace.
- $\|X\|_{\mathrm{tr}_n} = \mathrm{tr}_n(X^2)^{1/2}$ is the normalized Hilbert-Schmidt norm.
- $\langle X, Y \rangle_{\mathrm{tr}_n} = \mathrm{tr}_n(X^* Y)$.
- $(\mathbb{M}_n)_{\mathrm{sa}}$ is the subspace of self-adjoint matrices.
- Lebesgue measure on $(\mathbb{M}_n)_{\mathrm{sa}}$ is defined through an isometric transformation of $(\mathbb{M}_n)_{\mathrm{sa}}$ to $\mathbb{R}^{n^2}$ (i.e. by fixing an orthonormal basis).

Matrix models: Single variable case

Let $f : \mathbb{R} \to \mathbb{R}$ such that f grows sufficiently fast at ∞. Let $\mu_f^{(n)}$ be the probability measure on $(\mathbb{M}_n)_{\mathrm{sa}}$ given by

$$d\mu_f^{(n)}(x) = \frac{1}{Z^{(n)}} e^{-n^2 \operatorname{tr}_n(f(x))}\, dx.$$

In terms of eigenvalues, $n^2 \operatorname{tr}_n(f(x)) = n \sum_{j=1}^n f(\lambda_j)$, so this is similar to an n-particle approximation of a Gibbs state in statistical mechanics.

Let $X^{(n)}$ be a random element of $(\mathbb{M}_n)_{\mathrm{sa}}$ with distribution $\mu_f^{(n)}$. The probability measure $\mu_f^{(n)}$ is invariant under unitary conjugation, or $UX^{(n)}U^* \sim X^{(n)}$.

The *empirical spectral distribution* of $X^{(n)}$ is $(1/n) \sum_{j=1}^n \delta_{\lambda_j}$. This is a probability distribution on $\mathbb{R}$ which is also random.

Matrix models: Single variable case

How does the empirical spectral distribution of $X^{(n)}$ behave as $n \to \infty$?

Theorem [8]

Suppose that the functional

$$\mathcal{P}(\mathbb{R}) \ni \mu \mapsto \iint_{\mathbb{R}^2} \log|s-t|\, d\mu(s)\, d\mu(t) - \int_{\mathbb{R}} f(t)\, d\mu(t)$$

has a unique maximizer (which often happens. . .). Then almost surely the empirical spectral distribution of $X^{(n)}$ weakly converges to μ as $n \to \infty$.

Example: If $f_0(x) = (1/2)x^2$, then μ is Wigner's semicircle law $d\mu(x) = (1/2\pi)\mathbf{1}_{[-2,2]}(x)\sqrt{4-x^2}\, dx$ [37].

Remarks

1. The maximizer μ is a version of a Gibbs law, where the role of entropy is played by the logarithmic energy [6].
2. In fact, the logarithmic energy of μ describes the large-n behavior of the classical differential entropy $(-\int \rho \log \rho)$ for $\mu_f^{(n)}$ [34].
3. This logarithmic energy also describes the large deviations theory for the empirical spectral distribution.

Main motivation for us: *What happens in the non-self-adjoint and the multi-matrix setting?*

Unitarily invariant functions

What are the correct functions to replace $\mathrm{tr}_n(f(x))$*?*
We focus on the polynomial case first.

Theorem (Specht, Jing [28])

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_m \in \mathbb{M}_n$. Suppose that $\mathrm{tr}_n(p(X_1, \ldots, X_m)) = \mathrm{tr}_n(p(Y_1, \ldots, Y_m))$ for all $*$-polynomials p. Then there exists a unitary U such that $Y_j = UX_jU^*$ for $j = 1, \ldots, m$.

Theorem (Procesi [31])

Let $f : \mathbb{M}_n^m \to \mathbb{C}$ be a polynomial function of the real and imaginary parts of the matrix entries. Suppose $f(UX_1U^*, \ldots, UX_mU^*) = f(X_1, \ldots, X_m)$ for unitaries U. Then f is a *trace polynomial*, that is, f is in the algebra generated by functions of the form $\mathrm{tr}_n(p(X_1, \ldots, X_m))$ for $*$-polynomials p.

Non-commutative polynomials

- $\mathbb{C}\langle t_1, \dots, t_m \rangle$ denotes the algebra of non-commutative polynomials in formal variables t_1, ..., t_m, i.e.

$$\mathbb{C}\langle t_1, \dots, t_m \rangle = \operatorname{Span}(t_{i_1} \dots t_{i_\ell} : i_1, \dots, i_\ell \in [m]).$$

- $\mathbb{C}^*\langle t_1, \dots, t_m \rangle$ denotes the algebra of non-commutative $*$-polynomials, or $\mathbb{C}\langle t_1, \dots, t_m, t_1^*, \dots, t_m^* \rangle$. We have a $*$-operation satisfying $(p^*)^* = p$ and $(pq)^* = q^* p^*$ and $(t_i)^* = t_i^*$.
- The algebra of *trace polynomials* is the space of linear combinations of expressions of the form $\operatorname{tr}(p_1) \dots \operatorname{tr}(p_k)$.

Each of these algebraic objects can be *evaluated* on some matrices $(x_1, \dots, x_m)$ by plugging in x_j for t_j, x_j^* for t_j^*, and tr_n for tr.

Multi-matrix models

Let V be a real-valued trace polynomial, and consider $\mu_V^{(n)} \in \mathcal{P}(\mathbb{M}_n^m)$ given by

$$d\mu_V^{(n)}(x) = \frac{1}{Z_V^{(n)}} e^{-n^2 V(x)}\, dx,$$

and let $\mathbf{X}^{(n)} = (X_1^{(n)}, \dots, X_m^{(n)})$ be a random variable in $\mathbb{M}_n^m$ with distribution $\mu_V^{(n)}$.

Goals:

- When does $f(\mathbf{X}^{(n)})$ converge almost surely for trace polynomials f?
- What kind of object describes the limit?
- What is the large-n behavior of the entropy $h(\mu_{V^{(n)}})$?
- What is the large-n behavior of the Wasserstein distance $d_W(\mu_{V_0^{(n)}}, \mu_{V_1^{(n)}})$?

Non-commutative laws

For a single matrix, the values of $\mathrm{tr}_n(f(x))$ are described by the empirical spectral distribution, which is just a measure, but in the multi-matrix setting we cannot use a classical measure because the polynomials are non-commutative.

Definition (Non-commutative laws)

Let $\Sigma_{m,R}$ be the set of linear functionals $\mu : \mathbb{C}^*\langle t_1, \dots, t_m \rangle \to \mathbb{C}$ satisfying

1. $\mu(1) = 1$.
2. $\mu(p^* p) \geq 0$.
3. $\mu(pq) = \mu(qp)$.
4. $|\mu(t_{i_1}^{\delta_1}, \dots, t_{i_\ell}^{\delta_\ell})| \leq R^\ell$ where $\delta_j \in \{1, *\}$.

We see $\Sigma_{m,R}$ as a non-commutative analog of $\mathcal{P}(\mathbb{D}_R^{\times m})$.

Non-commutative probability spaces

Definition

A **tracial von Neumann algebra** is a pair (M, τ) where $M \subseteq B(H)$ is a von Neumann algebras (closed under $+$, $\cdot$, $*$, and limits in weak operator topology) and $\tau : M \to \mathbb{C}$ is a linear map satisfing

- $\tau(1) = 1$
- $\tau(ab) = \tau(ba)$ for all $a, b \in A$.
- $\tau(a^* a) \geq 0$.
- $\tau(a^* a) = 0$ implies $a = 0$.

The elements of M represent "bounded random variables" that don't commute under multiplication.

Example 1: The matrix algebra $(\mathbb{M}_n, \mathrm{tr}_n)$

Example 2: $M = L^\infty(\Omega, P)$ and $\tau(f) = \int f \, dP$.

NC probability spaces and laws

If (M, τ) is a tracial von Neumann algebra, and $x = \mathbf{x} = (x_1, \dots, x_m)$ is in M and $R > \max_j \|x_j\|$, then there is a non-commutative law $\operatorname{law}(\mathbf{x})$ given by

$$\operatorname{law}(\mathbf{x})(p) = \tau(p(\mathbf{x})),$$

where $p(\mathbf{x}) \in M$ is the evaluation of p on $\mathbf{x}$.

Conversely, every non-commutative law can be realized in this way, using a version of the GNS construction.

We equip $\Sigma_{m,R}$ with the weak-$*$ topology, that is, $\mu_i \to \mu$ if and only if $\mu_i(p) \to \mu(p)$ for all p. This makes it a compact Hausdorff space (and in fact metrizable).

Free Wasserstein distance (Biane–Voiculescu)

Wasserstein distance (Biane-Voiculescu 2001 [4])

Given two non-commutative laws μ and ν, the free Wasserstein distance $d_{W,\mathrm{free}}(\mu,\nu)$ is the infimum of $\|\mathbf{x}-\mathbf{y}\|_{L^2(M,\tau)^d}$ over all tracial von Neumann algebras (M,τ) and elements $\mathbf{x}=(x_1,\dots,x_m)$ and $\mathbf{y}=(y_1,\dots,y_m)$ with $\mathrm{law}(\mathbf{x})=\mu$ and $\mathrm{law}(\mathbf{y})=\nu$.

Warnings: [14]

1. $\Sigma_{m,R}$ is not separable with respect to $d_{W,\mathrm{free}}$. (See also [30].)
2. The Wasserstein distance gives a strictly stronger topology on $\Sigma_{m,R}$ than the weak-$*$ topology.
3. Due to MIP* $=$ RE [27], the laws $\Sigma_{m,R,\mathrm{fd}}$ coming from finite-dimensional M are not weak-$*$ dense in $\Sigma_{m,R}$.
4. The Wasserstein closure of $\Sigma_{m,R,\mathrm{fd}}$ is laws which come from amenable tracial von Neumann algebras, which is much smaller than the weak-$*$ closure of $\Sigma_{m,R,\mathrm{fd}}$.

Free entropy

Voiculescu defined several versions of entropy for non-commutative laws. Here I focus on χ which is defined via matrix approximations [35]. Rather than giving a definition, here is a useful characterization.

Proposition (J. 2022 [32])

Let $\mathbf{X}^{(n)}$ be some sequence of random matrix tuples satisfying that $\text{law}(\mathbf{X}^{(n)}) \to \mu$ weak-$*$ almost surely and $\mathbb{P}(\|X_j^{(n)}\| \geq C + \delta) \leq e^{-cn^2\delta}$. Then

$$\chi(\mu) \geq \limsup_{n\to\infty} \left[\frac{1}{n^2} h(\mathbf{X}^{(n)}) + 2m \log n \right].$$

Moreover, equality is achieved for some sequence of random matrix models.

Warning: We don't know whether taking $\liminf$ instead of $\limsup$ gives a different answer.

Back to the multi-matrix models

Consider again the random matrix models

$$d\mu_V^{(n)}(x) = \frac{1}{Z_V^{(n)}} e^{-n^2 V(x)}\, dx,$$

- If there is a unique NC law maximizing $\chi(\mu) - (\mu, V)$ and if the $\limsup$ is actually a limit, $\mathrm{law}(\mathbf{X}^{(n)})$ converges almost surely to μ.
- This happens if V is strongly convex and satisfies suitable growth bounds at ∞. [18, 21, 26]
- If $\mathrm{law}(\mathbf{X}^{(n)})$ converges almost surely to μ, then [21]

$$\chi(\mu) = \limsup_{n \to \infty} \left[\frac{1}{n^2} h(\mathbf{X}^{(n)}) + 2m \log n \right].$$

Wasserstein distance and random matrices

Proposition (J.–Li–Shlyakhtenko [26])

If V_0 is quadratic and V_1 is a small perturbation of it, then $d_W(\mu_{V_0}^{(n)}, \mu_{V_1}^{(n)}) \to d_{W,\mathrm{free}}(\mu, \nu)$. The optimal transport map was constructed earlier by Guionnet and Shlyakhtenko [19]; see also [13], [10].

It is not well-understood what happens to $d_W(\mu_{V_0}^{(n)}, \mu_{V_1}^{(n)})$ as $n \to \infty$ when V_0 and V_1 are not convex.

We know at least that MIP* $=$ RE provides an obstruction [14]. The optimal coupling for the limit (if it exists) sometimes does not admit f.d. approximations. So we can consider modifying $d_{W,\mathrm{free}}$ to restrict to couplings in tracial von Neumann algebras that admit f.d. approximations.

A fool's hope

Coming at this topic naïvely, we want for something like this to be true:

Conjecture

For non-commutative laws μ and ν, there are random matrix models $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$ such that

1. $\mathrm{law}(\mathbf{X}^{(n)}) \to \mu$ and $\mathrm{law}(\mathbf{Y}^{(n)}) \to \nu$ in probability.
2. $d_W(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \to d_{W,\mathrm{free}}(\mu, \nu)$ (where we restrict to couplings with f.d. approximations).
3. $n^{-2} h(\mathbf{X}^{(n)}) + 2m \log n \to \chi(\mu)$ and $n^{-2} h(\mathbf{Y}^{(n)}) + 2m \log n \to \chi(\nu)$.

Note that (1) and (2) alone or (1) and (3) alone can be arranged.

Consequence: Suppose (X, Y) is a free optimal couplings of (μ, ν) and μ_t is the law of $(1-t)X + tY$, so that μ_t is a geodesic with respect to $d_{W,\mathrm{free}}$. If the above conjecture is true, then $t \mapsto \chi(\mu_t)$ would be concave as a consequence of McCann's result for classical entropy [29].

A fool's hope

THE CONJECTURE IS FALSE! [25] In fact, it is false even if we only demand that $\mathbf{X}^{(n)}$ has asymptotically the correct entropy as $n \to \infty$, and not $\mathbf{Y}^{(n)}$.

I will skip the details and focus on the phenomenon that this counterexample derives from: *The entropy of the random matrix models does not behave well under marginals.*

Consider a non-commutative law μ of (X_1, X_2) where X_1 is a free circular element (limit of non-self-adjoint Gaussian matrix), $\tau(X_2) = 0$, $\|X_2\| \leq 1$, $\|X_2\|_2 \geq 1/2$ and $\|[X_1, X_2]\|_2 \leq \varepsilon$ for a small ε.

If $(X_1^{(n)}, X_2^{(n)})$ is *any* random matrix model with $\text{law}(\mathbf{X}^{(n)}) \to \text{law}(\mathbf{X})$ in probability, then

$$\limsup_{n\to\infty} [n^{-2} h(X_1^{(n)}) + m \log n] < \chi(X_1).$$

A fool's hope

The reason is the following: For $A \in \mathbb{M}_n$, let

$$\varphi^{(n)}(A) = \inf_{\|B\| \leq 1} \left[\|[A,B]\|_2 + |\operatorname{tr}_n(B)| + \max(0, 1/2 - \|B\|_2) \right].$$

Then $\{A : \varphi^{(n)}(A) < \varepsilon\}$ has very small volume as $\varepsilon \to 0$, and therefore, the probability distribution of $X_1^{(n)}$ must have small entropy [36].

The volume is estimated as follows: If $\varphi^{(n)}(A) < \varepsilon$, find some B that achieves the infimum. The formula demands that B is bounded away from multiples of the identity. There is some projection P in the $*$-algebra generated by B with $[A,P]$ small and $\operatorname{tr}_n(P)$, $1 - \operatorname{tr}_n(P)$ bounded away from zero, so A is "almost block diagonal" with respect to P.

For each P, the block diagonal matrices have small volume compared to all matrices. And results of Szarek [33] allow us to estimate the cardinality of an ε-dense set of P's.

Defeating the dragon is hard, but you have to try.

From *Nowe Ateny*, first Polish-language encyclopedia (1745–1746).

Monge–Kantorovich duality?

Coming back to the main question of how to unify free entropy, free Wasserstein distance, and random matrix models, a key question is how to adapt MK duality to play well with the class of trace-polynomial functions.

Classical Monge Kantorovich duality

Let μ and ν be probability measures on $\mathbb{R}^m$ with finite second moment. Then there exist convex functions $\varphi, \psi : \mathbb{R}^m \to (-\infty, \infty]$ such that $\varphi(x) + \psi(y) \geq \langle x, y \rangle$ and when (X, Y) is an optimal coupling of (μ, ν) we have

$$\mathbb{E}\varphi(X) + \mathbb{E}\psi(Y) = \mathbb{E}\langle X, Y \rangle.$$

Moreover, ψ can be taken to be the Legendre transform of φ, that is,

$$\psi(y) = \sup_x [\langle x, y \rangle - \varphi(x)].$$

Monge–Kantorovich duality?

Now suppose φ is a real-valued trace polynomial function of m variables. For each von Neumann algebra M, we can define a function

$$\psi^M(y) = \sup_{x \in M^m} \left[\operatorname{Re}\langle x, y \rangle - \varphi^M(x) \right].$$

Is the function ψ a nice function that can be uniformly approximated by trace polynomials?

No, in fact, $\psi^M(y)$ can depend on more than just $\operatorname{law}(y)$. Sometimes, you can embed M into a larger algebra $\tilde{M}$ such that $\psi^{\tilde{M}}(y) > \psi^M(y)$. (You can find examples even just with M and $\tilde{M}$ finite-dimensional.)

A paradox

Let $f(x, y)$ be a trace polynomial that is real-valued, and let

$$\psi^M(x) = \sup_{\|y\| \leq 1} f^M(x, y).$$

If we fix n, then $\psi^{\mathbb{M}_n}$ is a function that is invariant under simultaneous unitary conjugation.

We can therefore approximate it by polynomial functions of the matrix entries that are invariant under unitary conjugation. So by Procesi's result, $\psi^{\mathbb{M}_n}$ can be approximated by trace polynomials.

However, the approximation is not uniform in n. It becomes harder and harder to approximate it by trace polynomials as $n \to \infty$.

Lack of quantifier elimination

Theorem (Farah [11])

Let (M, τ) be a II_1 factor (tracial von Neumann algebra with trivial center that is infinite-dimensional). Then (M, τ) does not have quantifier elimination. Therefore, more concretely, there exist some trace polynomials $\varphi(x_1, \dots, x_m, y)$ such that

$$\psi^M(x_1, \dots, x_m) = \sup_{\|y\| \leq 1} \varphi^M(x_1, \dots, x_m, y)$$

cannot be approximated by trace polynomials uniformly on the operator norm ball of M.

Remark: If $(M, \tau) = (L^\infty[0,1], \int(\cdot))$, then you *can* approximate such a sup formula with quantifier-free formulas [3].

Remark: Another result of Farah–J.–Pi shows that for many (M, τ), if you have a sup-inf formula, it cannot be approximated by inf-formulas [12].

Formulas

If we want a vector space of scalar functions of non-commuting variables that are closed under partial suprema and infima, we are led inexorably to the formulas from the model theory of metric structures [2].

These are like formulas from first-order logic but with the predicates having real values rather than true/false values, and the quantifiers being sup and inf over given domains.

For tracial von Neumann algebras, such a formula in prenex form would be

$$\varphi(\mathbf{x}) = \sup_{\|z_1\| \leq 1} \inf_{\|z_2\| \leq 1} \dots \sup_{\|z_{2k-1}\| \leq 1} \inf_{\|z_{2k}\| \leq 1} f(\mathbf{x}, \mathbf{z}),$$

where f is a real-valued trace polynomial and $\mathbf{x} = (x_1, \dots, x_m)$ and $\mathbf{z} = (z_1, \dots, z_{2k})$.

Definable predicates and types

Evaluation of formulas: Such a formula can be *evaluated* or *interpreted* in any tracial von Neumann algebra (M, τ) by plugging in a particular $(x_1, \dots, x_m)$ from M and evaluating the suprema and infima over the unit ball of M. This evaluation is denoted $\varphi^M(\mathbf{x})$.

Definable predicates: The space of *definable predicates* is the completion of the vector space of formulas with respect to uniform convergence on operator norm balls (that is uniform for all M).

Types: Now consider the analog of non-commutative laws but using formulas rather than only polynomials. The *type* of a tuple $\mathbf{x}$ is the map $\mathrm{tp}^M(\mathbf{x})$ from formulas to $\mathbb{R}$ given by $\varphi \mapsto \varphi^M(\mathbf{x})$. The space of types $\mathbb{S}_m$ is equipped with the weak-$*$ topology.

Challenges

When attempting to apply this framework to random matrices, we run into a basic question.

Open Question

Consider formulas φ with no free variables, e.g.

$$\varphi = \sup_{\|z_1\| \leq 1} \inf_{\|z_2\| \leq 1} \dots \sup_{\|z_{2k-1}\| \leq 1} \inf_{\|z_{2k}\| \leq 1} f(\mathbf{z}).$$

Then does $\lim_{n \to \infty} \varphi^{\mathbb{M}_n}$ exist for all φ?

Remark: The question in fact has a *negative answer* if some of the suprema and infima are over the unit ball in $\mathbb{M}_n$ and some in the unit ball of diagonal matrices. This follows after some argument from the work of Alekseev and Thom on the analogous question for permutation groups rather than matrix algebras [1].

Uncomputability

There is no hope of numerical computation for these objects with error guarantees. This is a consequence of the MIP* = RE paper [27].

Theorem (Goldbring–Hart [16])

There is no algorithm that, given a real-valued trace polynomial f with coefficients in $\mathbb{Q}[i]$ and $k \in \mathbb{N}$, outputs a rational number r such that

$$\left| \lim_{n \to \infty} \inf_{\|z_1\|, \dots, \|z_k\| \leq 1} f^{\mathbb{M}_n}(\mathbf{z}) - r \right| < \frac{1}{k}.$$

(The limit as $n \to \infty$ does exist when we only have inf's and no sup's.)

Matrix algebras *really are amazing creatures, as I have said before. You can learn all there is to know about their ways in a month, and yet after a hundred years, they can still surprise you in a pinch.* [J.R.R. Tolkien]

Wasserstein distance for types

For two types $\mu, \nu \in \mathbb{S}_m$, define $d_{W,\mathrm{tp}}(\mu,\nu)$ as the infimum of $\|\mathbf{x}-\mathbf{y}\|_2$ over $\mathbf{x}$ and $\mathbf{y}$ with types μ and ν in any tracial von Neumann algebra.

An embedding $\iota : M \to N$ of tracial von Neumann algebras is *elementary* if for every formula φ, we have $\varphi^N \circ \iota = \varphi^M$.

Fact

Given a tracial von Neumann algebra M, there exists an elementary embedding $M \to N$ such that for $\mathbf{x}, \mathbf{y} \in N^m$,

$$d_{W,\mathrm{tp}}(\mathrm{tp}^N(\mathbf{x}), \mathrm{tp}^M(\mathbf{y})) = \inf_{\alpha \in \mathrm{Aut}(N)} \|\alpha(\mathbf{x}) - \mathbf{y}\|_2.$$

Remark: For $M = L^\infty[0,1]$ and $\mathbf{X}, \mathbf{Y} \in M^m$ (random variables taking values in $\mathbb{C}^m$), we have already that $\inf_{\alpha \in \mathrm{Aut}(M)} \|\alpha(\mathbf{X}) - \mathbf{Y}\|_{L^2}$ is the Wasserstein distance.

Monge-Kantorovich duality for types

Theorem (J. 2024 [24])

Let $\mu, \nu \in \mathbb{S}_m$ be types in tracial von Neumann algebras. Then there exist convex definable predicates φ and ψ such that

$$\varphi^M(\mathbf{x}) + \psi^M(\mathbf{y}) \geq \mathrm{Re}\langle \mathbf{x}, \mathbf{y} \rangle_\tau \text{ for all } M, \mathbf{x}, \mathbf{y},$$

with equality when $(\mathbf{x}, \mathbf{y})$ is an optimal coupling of (μ, ν).

Remark: Gangbo–J.–Nam–Shlyakhtenko [14] showed a MK duality for non-commutative laws. The convex functions in this setting were E-convex functions, defined as functions $\phi^M(\mathbf{x})$ that only depend on $\mathrm{law}(\mathbf{x})$ and satisfy that if $M \subseteq N$ and $E : N \to M$ is the conditional expectation, then $\phi^M \circ E \leq \phi^N$.

These functions are not necessarily weak-$*$ continuous functions of $\mathrm{law}(\mathbf{x})$, but in the setting of types, we can arrange weak-$*$ continuity.

Free entropy for types

Free entropy χ for types can be defined similarly as for non-commutative laws [23]. Because we don't know if the limits exist as $n \to \infty$, we take limits with respect to an ultrafilter (something like a subsequential limit).

In this setting, we still don't know if there $\chi^{\mathcal{U}}$ is concave along Wasserstein geodesics, but we can obtain 1-sided bounds.

Theorem (J. 2025 [25])

Let μ and ν be types. (For the below to be nontrivial, take types that are realized in an ultraproduct $\prod_{n \to \mathcal{U}} \mathbb{M}_n$.) Let $(\mathbf{x}, \mathbf{y})$ be an optimal coupling of these types, and let μ_t be the type of $(1-t)\mathbf{x} + t\mathbf{y}$. Then

$$\chi^{\mathcal{U}}(\mu_t) \geq \max(\chi^{\mathcal{U}}(\mu) + 2m \log(1-t), \chi^{\mathcal{U}}(\nu) + 2m \log t).$$

Free entropy for types

The reason that this works is that a one-sided version of the earlier naïve conjecture holds for types.

Proposition (J. 2025 [25])

For types μ and ν from $\prod_{n \to \mathcal{U}} \mathbb{M}_n$, there are random matrix models $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$ such that

1. $\mathrm{tp}^{\mathbb{M}_n}(\mathbf{X}^{(n)}) \to \mu$ and $\mathrm{tp}^{\mathbb{M}_n}(\mathbf{Y}^{(n)}) \to \nu$ weak-$*$ in probability as $n \to \mathcal{U}$.
2. $d_W(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \to d_{W,\mathrm{tp}}(\mu, \nu)$ as $n \to \mathcal{U}$.
3. $n^{-2} h(\mathbf{X}^{(n)}) + 2m \log n \to \chi^{\mathcal{U}}(\mu)$ as $n \to \mathcal{U}$.

As seen above, the analogous statement for non-commutative laws is false.

The key point is that weak-$*$ continuity of the functions (φ, ψ) in the MK duality allows us to use them to build potentials V which produce the desired random matrix models.

Future work

Coming back to the question of the limiting behavior of entropy of $\mu_V^{(n)}$, what we need is not only to be able to discuss the infimum of some trace polynomial, but to study a stochastic optimization problem.

Proposition (special case of e.g. Boué–Dupuis [7])

Let $f : \mathbb{R}^m \to \mathbb{R}$ be a bounded function. Let B_t be a Brownian motion in $\mathbb{R}^m$ with

$$-\log \mathbb{E}[e^{-f(B_1)}] = \inf_\alpha \mathbb{E}\left[\int_0^1 \|\alpha_t\|^2 \, dt + f\left(B_1 + \int_0^1 \alpha_t \, dt\right)\right],$$

where α ranges over all control processes adapted to the given filtration associated to the Brownian motion.

Future work

Corollary

Let $\mathbf{S}_t^{(n)}$ be a normalized Brownian motion on $\mathbb{M}_n^m$ and f a real-valued trace polynomial in resolvents of x_j (for instance). Then

$$-\frac{1}{n^2}\log\frac{1}{Z^{(n)}}\int_{\mathbb{M}_n^m} e^{-n^2(\frac{1}{2}\|\mathbf{x}\|_2^2+f(\mathbf{x}))}\,d\mathbf{x} = \inf\left[\int_0^1 \|\alpha_t\|^2\,dt + f\left(\mathbf{S}_1^{(n)} + \int_0^1 \alpha_t\,dt\right)\right]$$

where $Z^{(n)}$ is the normalizing constant for the Gaussian measure and α_t ranges over control processes in $\mathbb{M}_n^m$.

This is closely related to the large-deviations rate function of Biane–Capitaine–Guionnet motivated by Malliavin calculus [5].

Future work

So understanding the large-n behavior of entropy for random matrix models is equivalent to understanding the large-n behavior of these stochastic control problems.

Joint with Gangbo, Nam, Palmer: Systematic study of free stochastic control problems, which you will see later in this conference [15].

Ongoing/future work with various collaborators: Formulating types in a setting with a filtration and stochastic processes. Limits of the matrix versions.

References I

[1] Vadim Alekseev and Andreas Thom. On non-isomorphic universal sofic groups. Preprint, arXiv:2406.06741

[2] Itaï Ben Yaacov, Alexander Berenstein, C. Ward Henson, Alexander Usvyatsov. Model theory for metric structures. In *Model Theory with Applications to Algebra and Analysis, Vol. II*. 2008.

[3] Itaï Ben Yaacov. On theories of random variables. Israel J. Math., 194.2:957-1012 (2013).

[4] Philippe Biane and Dan-Virgil Voiculescu. A free probability analogue of the Wasserstein metric on the trace-state space. Geometric and Functional Analysis, 11:1125–1138, 2001.

[5] P. Biane, M. Capitaine, and A. Guionnet. Large deviation bounds for matrix Brownian motion. Inventiones Mathematicae, 152:433-459 (2003).

References II

[6] P. Biane and R. Speicher, Free diffusions, free entropy and free Fisher information, Ann. Inst. Henri Poincaré Probab. Statist. 37 (2001), 581–606.

[7] Michelle Boué and Paul Dupuis. A variational representation for certain functionals of Brownian motion. Ann. Prob. 26.4:1641-1659 (1998).

[8] A. Boutet de Monvel, L. Pastur, and M. Shcherbina. On the statistical mechanics approach in the random matrix theory: Integrated density of states. J. Stat. Phys. 79.3:585-611 (1995).

[9] Y. Dabrowski, A Laplace principle for Hermitian Brownian motion and free entropy I: the convex functional case, arXiv:1604.06420 (2017).

[10] Y. Dabrowski, A. Guionnet, and D. Shlyakhtenko, Free transport for convex potentials, New Zealand J. Math. 52 (2021), 251–359.

References III

[11] Ilijas Farah. Quantifier elimination in II_1 factors. To appear in Münster Journal of Math.

[12] Ilijas Farah, David Jekel, Jennifer Pi. Quantum expanders and quantifier reduction for tracial von Neumann algebras. Preprint, arXiv:2310.06197.

[13] Alessio Figalli and Alice Guionnet. Universality in several-matrix models via approximate transport maps. Acta Math., 217.1:81-176, 2016.

[14] Wilfrid Gangbo, David Jekel, Kyeongsik Nam, and Dimitri Shlyakhtenko. Duality for optimal couplings in free probability. Comm. Math. Phys., 2022.

[15] Wilfrid Gangbo, David Jekel, Kyeongsik Nam, and Aaron Z. Palmer. Viscosity solutions in non-commutative variables. Preprint, arXiv:2502.17329

References IV

[16] Isaac Goldbring and Bradd Hart. The universal theory of the hyperfinite II_1 factor is not computable, The Bulletin of Symbolic Logic. 2024;30(2):181-198. doi:10.1017/bsl.2024.7

[17] A. Guionnet and E. Maurel-Segala, Combinatorial aspects of random matrix models, ALEA, Lat. Amer. J. Probab. Math. Statist. 1 (2006), 241–279

[18] A. Guionnet and D. Shlyakhtenko, Free diffusions and matrix models with strictly convex interaction, Geom. Funct. Anal. 18 (2009), 1875-1916.

[19] Alice Guionnet and Dimitri Shlyakhtenko. Free monotone transport. Inventiones Mathematicae, 197(3):613–661, 2014.

[20] Uffe Haagerup and Magdalena Musat. An asymptotic property of factorizable completely positive maps and the Connes embedding problem. Comm. Math. Phys., 338(2):721–752, 2015.

[21] David Jekel, An elementary approach to free entropy theory for convex potentials, Anal. PDE 13 (2020), 2289-2374.

[22] David Jekel, Conditional expectation, entropy, and transport for convex Gibbs laws in free probability, International Mathematics Research Notices, 2022.6, pp. 4516-4619

[23] David Jekel. Free probability and model theory of tracial W^*-algebras. In Isaac Goldbring, editor, Model Theory of Operator Algebras, pages 215-267. DeGruyter, Berlin, Boston, 2023.

[24] David Jekel. Optimal transport for types and convex analysis for definable predicates in tracial W^*-algebras. J. Funct. Anal. 2024.

[25] David Jekel. Information geometry for types in the large-n limit of random matrices. Preprint, arXiv:2501.00703.

References VI

[26] David Jekel, Wuchen Li, and Dimitri Shlyakhtenko. Tracial non-commutative smooth functions and the free Wasserstein manifold. Dissertationes Mathematicae 580 (2022), pp. 1-150

[27] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, Henry Yuen. MIP*=RE. arXiv:2001.04383

[28] Naihuan Jing. Unitary and orthogonal equivalence of sets of matrices. Linear Algebra and its Applications, 481:235–242, 2015.

[29] Robert J. McCann. A convexity principle for interacting gases. Adv. Math., 128.1:153-179, 1997.

[30] Narutaka Ozawa. There is no separable universal II_1 factor. Proc. Amer. Math. Soc., 132:487–90, 2004.

[31] Claudio Procesi. The invariant theory of $n \times n$ matrices, Adv. Math. 19:306-381 (1976).

References VII

[32] Dimitri Shlyakhtenko and Terence Tao, with an appendix by David Jekel, Fractional free convolution powers. Indiana Univ. Math. J. 71.6 (2022), 2551-2594

[33] Stanisław J. Szarek, Metric entropy of homogeneous spaces, Quantum probability (Gdańsk, 1997), Banach Center Publ., vol. 43, Polish Acad. Sci. Inst. Math., Warsaw, 1998, pp. 395-410.

[34] Dan-Virgil Voiculescu, The analogues of entropy and Fisher's information in free probability, I, Comm. Math. Phys. 155.1:71-92 (1993).

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

References VIII

[36] 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 (1996), no. 1, 172-199.

[37] E.P. Wigner, On the distribution of the roots of certain symmetric matrices. Ann. of Math. 67 (1958), 325-327