Free information geometry and the model theory of stochastic processes [1]

David Jekel
University of Copenhagen

Probabilistic Operator Algebras Seminar 2026-06-01

[1] Funded by the European Union's Marie Skłodowska-Curie Actions

Overview

Goal: Free analog of "heat evolution is the gradient flow of entropy on the Wasserstein space." This is the larger context behind the question of χ versus χ^*.

The correct definitions of entropy and Wasserstein distance need to take into account something about the ambient algebra. Entropy in the presence is the first example of this.

This motivated the development of model-theoretic free entropy [30], where the test functions involve suprema and infima (answers to optimization problems over auxiliary variables) in addition to the non-commutative law (moments).

Overview

Combination of techniques:

- Evolution variational inequality / theory of gradient flow in general metric spaces [40].
- Noncommutative Wasserstein distance [4, 17] in model-theoretic setting [31].
- Free microstate entropy [49] in model-theoretic setting [30].
- Noncommutative filtrations and stochastic optimization problems [18, 19].
- Interaction of continuous model theory with randomness [3, 20].

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.
- $\|X\|_2 = \mathrm{tr}_n(X^2)^{1/2}$ is the normalized Hilbert-Schmidt norm.
- $\langle X, Y\rangle_2 = \mathrm{tr}_n(X^* Y)$.
- $(\mathbb{M}_n)_{\mathrm{sa}}$ 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).

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 $\tau(f) = \int f \, dP$.

Multi-matrix models

Let f be a non-commutative $*$-polynomial in m variables, and let $V(\mathbf{x}) = \operatorname{Re}\operatorname{tr}(f(\mathbf{x}))$, which is defined for $\mathbf{x}$ in any tracial von Neumann algebra, including $\mathbb{M}_n$. Consider $\mu_V^{(n)} \in \mathcal{P}(\mathbb{M}_n^m)$ given by

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

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 $\operatorname{tr}_n(p(\mathbf{X}^{(n)}))$ converge almost surely for trace polynomials p?
- 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 $d_W(\mu_{V_0}^{(n)}, \mu_{V_1}^{(n)})$?

Multi-matrix models

Biane–Speicher [6], Guionnet–Shlyakhtenko [23]: We can analyze the $\mu_V^{(n)}$ as the large t limit distribution of a stochastic process satisfying

$$d\mathbf{X}_t^{(n)} = d\mathbf{Z}_t^{(n)} - \frac{1}{2}\nabla V(\mathbf{X}_t^{(n)})\,dt,$$

where $\mathbf{Z}_t^{(n)}$ is a matrix Brownian motion. In the large-n limit the traces of polynomials in $\mathbf{Z}_t^{(n)}$ converge to those of a free Brownian motion $\mathbf{z}_t$.

For convex V, the traces of polynomials in $\mathbf{X}^{(n)}$ converge in the large-n limit to some functional $\mu_V : \mathbb{C}\langle x_1, \dots, x_m\rangle \to \mathbb{C}$ (a *non-commutative law*).

$\mu_V(p) = \tau(p(\mathbf{x}))$ for some tuple $\mathbf{x}$ in a tracial von Neumann algebra (M, τ).

Multi-matrix models

Guionnet–Shlyakhtenko [24] (also [11, 34]): The von Neumann algebra $\mathrm{W}^*(\mathbf{x})$ is isomorphic to $L(\mathbb{F}_{2m})$, or it is isomorphic to the von Neumann algebra of a free circular family $\mathbf{z}$, and you can realize $\mathbf{x}$ as $\nabla\varphi(\mathbf{z})$ for some convex φ.

This is an analog of classical results from optimal transport theory.

Dabrowski [10], J. [28]: Two different definitions of free entropy agree for μ_V when V is convex.

Wasserstein information geometry framework

One of my main goals is to study the relationship between entropy and Wasserstein distance in the free setting. This is motivated by a lot of work in the classical case that establishes a unified framework of the *Wasserstein manifold*:

- The points of $\mathcal{P}^\infty(\mathbb{R}^d)$ are smooth positive probability measures (see Lafferty [38]).
- A tangent vector at μ is an equivalence class of C^1-paths in $\mathcal{P}^\infty(\mathbb{R}^d)$ through μ.
- The Riemannian metric is as follows. Suppose φ and $\psi : \mathbb{R}^d \to \mathbb{R}$ are smooth and consider the tangent vectors $\alpha_t = (\mathrm{id} + t\nabla\varphi)_*\mu$ and $\beta_t = (\mathrm{id} + t\nabla\psi)_*\mu$. Then

$$\langle \dot{\alpha}_0, \dot{\beta}_0 \rangle_g = \langle \nabla\varphi, \nabla\psi \rangle_{L^2(\mu)}.$$

Wasserstein information geometry framework

- **Theorem:** The Riemannian distance on $\mathcal{P}^\infty(\mathbb{R}^d)$ agrees with the Wasserstein distance

$$d_W(\mu, \nu) = \inf\{\|X - Y\|_{L^2} : X \sim \mu, Y \sim \nu\}.$$

- **Theorem:** Geodesics in $\mathcal{P}^\infty(\mathbb{R}^d)$ have the form $(\mathrm{id} + t\nabla\varphi)_*\mu$ where φ is convex.
- **Theorem (special case of McCann [39]):** Let $h(\mu) = -\int \rho \log \rho$. If μ_t is a geodesic in $\mathcal{P}^\infty(\mathbb{R}^d)$, then $t \mapsto h(\mu_t)$ is concave.
- **Theorem (special case of Otto [41]):** The gradient flow $\dot{\mu}_t = \operatorname{grad} h(\mu_t)$ on $\mathcal{P}^\infty(\mathbb{R}^d)$ agrees with the heat equation $\dot{\rho}_t = \Delta\rho_t$ for the density ρ_t.
- Inequalities such as the Talagrand inequality, log–Sobolev, etc. are statements are convexity of h and curvature of $\mathcal{P}^\infty(\mathbb{R}^d)$ (and for a Riemannian manifold M, they relate closely to the curvature of M itself).

Evolution variational inequality

The *evolution variational inequality* is a description of gradient flow that can be applied for abstract metric spaces.

Let φ be a real-valued function on (X, d) and $g : X \times [0, \infty) \to X$ a family of trajectories. They satisfy EVI_0'' [40] if

$$\frac{1}{2}[d(g(x,t),y)^2 - d(g(x,s),y)^2] \leq (t-s)(\varphi(g(x,t)) - \varphi(y)) \text{ for } s \leq t.$$

(sign convention here: upward gradient flow of concave function).

Motivating example: Suppose that $X = \mathbb{R}^m$ and φ is differentiable and $\partial_t g(x,t)$ exists. Then EVI implies that φ is concave and $\partial_t g(x,t) = \nabla \varphi(g(x,t))$.

Evolution variational inequality

Motivating example: Suppose that $X = \mathbb{R}^m$ and φ is differentiable and $\partial_t g(x,t)$ exists.

$$\frac{1}{2}[d(g(x,t+\varepsilon),y)^2 - d(g(x,t),y)^2] \leq \varepsilon(\varphi(g(x,t+\varepsilon)) - \varphi(y))$$
$$\left\langle \frac{g(x,t+\varepsilon)-g(x,t)}{\varepsilon}, \frac{g(x,t+\varepsilon)+g(x,t)}{2} - y \right\rangle \leq \varphi(g(x,t+\varepsilon)) - \varphi(y)$$
$$\langle \partial_t g(x,t), g(x,t) - y\rangle \leq \varphi(g(x,t)) - \varphi(y),$$

or

$$\varphi(y) \leq \varphi(g(x,t)) + \langle \partial_t g(x,t), y - g(x,t)\rangle.$$

So $\partial_t g(x,t)$ is a supergradient vector for φ at $g(x,t)$.

General theory of Daneri, Muratori, and Savaré [12, 40] shows that EVI implies something like geodesic concavity (convexity in their sign convention).

Summary of setup

The goal of this work is to define a framework for free entropy that has the correct properties of Wasserstein geometry.

To achieve this, we must expand our class of test functions beyond the traces of non-commutative polynomials.

We consider a class of test functions that is closed under taking suprema or infima with respect to some of the variables over the unit ball, as well as closed under substituting a free Brownian motion $\mathbf{z}_t$ (see next slide).

These are called *chronological formulas* because of the role of the time parameter for the filtration / Brownian motion.

Chronological formulas

Suppose we are given a *filtration* of von Neumann algebras $\mathcal{M} = (M_t)_{t \geq 0}$ where $M_s \subseteq M_t$ for $s \leq t$ and an associated *free Brownian motion* $\mathbf{z}_t = (z_{t,j})_{j=1}^m$ such that

- The real and imaginary parts of $\mathbf{z}_t$ are free semicircular operators.
- $\mathbf{z}_t \in M_t^m$.
- The increment $\mathbf{z}_t - \mathbf{z}_s$ is free from M_s for $s \leq t$.

Definition by example: A chronological formula is some expression like

$$\begin{aligned}&\varphi^{\mathcal{M}}(x_1, \dots, x_m) \\ &\quad = \sup_{y_1 \in B_1(M_{t_1})} \inf_{y_2 \in B_1(M_{t_2})} \operatorname{Re} \operatorname{tr}^M(p(x_1, \dots, x_m, y_1, y_2, \mathbf{z}_{s_1}, \dots, \mathbf{z}_{s_m})),\end{aligned}$$

where $t_1 \leq t_2$. There can be an arbitrary number of sup and inf's but the associated time parameters have to be in chronological order!

Setup

- We have a space $\mathcal{F}^0_{\text{chron},m}$ of *chronological formulas*.
- The chronological type of $\mathbf{x}$ is the map $\mathcal{F}^0_{\text{chron},m} \to \mathbb{R}$ given by $\varphi^{\mathcal{M}} \mapsto \varphi(\mathbf{x})$.
- For practical purposes, assume that $\mathbf{x}$ is chosen from Q_0^m, where $\mathcal{Q} = (Q_t)_{t \geq 0}$ is a filtration constructed from an ultraproduct of the matrix algebras (explained more later).
- Space of *chronological types* $\mathbb{S}(\mathrm{T}_{\mathcal{U}})$ associated to each ultrafilter $\mathcal{U}$ on $\mathbb{N}$.
- The *Wasserstein distance* $d_W(\mu, \nu)$ is the minimal distance $\|\mathbf{x} - \mathbf{y}\|_2$ where $\mathrm{tp}_{\text{chron}}(\mathbf{x}) = \mu$ and $\mathrm{tp}_{\text{chron}}(\mathbf{y}) = \nu$. We call $(\mathbf{x}, \mathbf{y})$ an *optimal coupling* of (μ, ν) if it achieves the infimum.
- We have a version of Voiculescu's free entropy for chronological types, $\chi^{\mathcal{U}}_{\text{chron}}(\mu)$ for $\mu \in \mathbb{S}(\mathrm{T}_{\mathcal{U}})$ (described more precisely later).

Main results

Theorem (Geodesic concavity) [33]

Let $t \mapsto \mu_t$ be a geodesic with respect to d_W. More explicitly, suppose $\mathbf{x}$ and $\mathbf{y}$ have types μ_0 and μ_1 with $\|\mathbf{x} - \mathbf{y}\|_2 = d_W(\mu_0, \mu_1)$, and let μ_t be the chronological type of $(1-t)\mathbf{x} + t\mathbf{y}$. Then $t \mapsto \chi^{\mathcal{U}}_{\text{chron}}(\mu_t)$ is concave.

Theorem (Evolution Variational Inequality) [33]

Let μ, $\sigma \in \mathbb{S}(\mathrm{T}_{\mathcal{U}})$. Let μ_t be the heat evolution of μ; namely, μ_t is the chronological type of $\mathbf{x} + \mathbf{z}_t$ where $\mathbf{x}$ has type μ and $\mathbf{z}_t$ is a non-commutative Brownian motion. Then we have the EVI

$$\frac{1}{2}\left[d_W(\mu_t, \sigma)^2 - d_W(\mu_s, \sigma)^2\right] \leq (t - s)(\chi^{\mathcal{U}}_{\text{chron}}(\mu_t) - \chi^{\mathcal{U}}_{\text{chron}}(\sigma))$$

for $0 \leq s \leq t < \infty$.

Auxiliary results

To discuss the proof, we need to relate the chronological types and entropy and Wasserstein distance, with the corresponding finite-dimensional matrix models.

- Each chronological formula $\varphi \in \mathcal{F}^0_{\text{chron},m}$ has an associated sequence of functions $\Lambda_\varphi^{(n)} : \mathbb{M}_n^m \to \mathbb{R}$.
- Given $\mu \in \mathbb{S}_{\text{chron}}(\mathrm{T}_\mathcal{U})$, a sequence of random matrix tuples $\mathbf{X}^{(n)}$ is said to be a *sequence of matrix models* for μ if

$$\lim_{n \to \mathcal{U}} \Lambda_\varphi^{(n)}(\mathbf{X}^{(n)}) = (\mu, \varphi)$$

for all $\varphi \in \mathcal{F}^0_{\text{chron},m}$.

Auxiliary results

We have the following characterization of $\chi^{\mathcal{U}}_{\text{chron}}$ as the largest possible limit of normalized classical entropy of matrix models for μ.

This can be taken as the definition of $\chi^{\mathcal{U}}_{\text{chron}}$ for purposes of this talk.

Proposition (classical-to-free variational principle); see [35]

For a matrix model $\mathbf{X}^{(n)}$, let $h^{(n)}(\mathbf{X}^{(n)}) = n^{-2} h(\mathbf{X}^{(n)}) + 2m \log n$. If $\mathbf{X}^{(n)}$ is a matrix model for μ satisfying reasonable tail bounds on $\mathbb{P}(\|\mathbf{X}^{(n)}\| \geq R)$ and $\mu^{(n)} = \text{law}(\mathbf{X}^{(n)})$, then

$$\chi^{\mathcal{U}}_{\text{chron}}(\mu) \geq \lim_{n \to \mathcal{U}} h^{(n)}(\mu^{(n)}).$$

Moreover, there exist matrix models for μ that achieve equality.

Auxiliary results

Theorem (Extension of matrix models for optimal couplings) [32, 33]

Fix types μ and ν. Let $(\mathbf{x}, \mathbf{y})$ be a pair of m-tuples which form an optimal coupling of chronological types μ and ν. Let $\mathbf{X}^{(n)}$ be a sequence of matrix models for μ. Then there exists a sequence of matrix models $\mathbf{Y}^{(n)}$ for ν such that

- $\|\mathbf{X}^{(n)} - \mathbf{Y}^{(n)}\|_{L^2}$ is equal to the classical Wasserstein distance between $\text{law}(\mathbf{X}^{(n)})$ and $\text{law}(\mathbf{Y}^{(n)})$.
- $(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})$ is a sequence of matrix models for the joint chronological type of $(\mathbf{x}, \mathbf{y})$.

Summary of objects and finite-dimensional analogs

chron. type setting	**random matrix setting**
m-tuple $\mathbf{x}$ from Q_0^m	random matrix m-tuple $\mathbf{X}^{(n)}$
chron. type $\mu \in \mathbb{S}_{\mathrm{chron},m}(\mathrm{T}_{\mathcal{U}})$	probability dist. $\nu^{(n)} \in \mathcal{P}(\mathbb{M}_n^m)$
free entropy $\chi_{\mathrm{chron}}^{\mathcal{U}}$	normalized entropy $h^{(n)}$
logical Wasserstein dist. d_W	classical Wasserstein dist. $d_{W,\mathrm{class}}$
chron. formula φ	f.d. approx. $\Lambda_\varphi^{(n)} : \mathbb{M}_n^m \to \mathbb{R}$.

Proof of EVI

By the semigroup property of the heat semigroup, assume WLOG that the earlier time $s = 0$.

Fix chronological types μ and σ. By the classical-to-free variational principle, fix matrix models $\mathbf{Y}^{(n)}$ such that

$$\lim_{n \to \mathcal{U}} h^{(n)}(\sigma^{(n)}) = \chi_{\text{chron}}^{\mathcal{U}}(\sigma).$$

Let $\sigma^{(n)} \in \mathcal{P}(\mathbb{M}_n^m)$ be its probability distribution. By the optimal coupling extension theorem, fix matrix models $\mathbf{X}^{(n)}$ for μ such that

$$\lim_{n \to \mathcal{U}} d_{W,\text{class}}(\mu^{(n)}, \sigma^{(n)}) = d_W(\mu, \sigma),$$

where $\mu^{(n)}$ is the law of $\mathbf{X}^{(n)}$.

Proof of EVI

Let $\mathbf{Z}_t^{(n)}$ be a matrix Brownian motion independent of $\mathbf{X}^{(n)}$. Then $\mathbf{X}^{(n)} + Z_t^{(n)}$ is a sequence of matrix models for μ_t (the heat semigroup will be defined so that this is true). Then

$$\lim_{n \to \mathcal{U}} h^{(n)}(\mu_t^{(n)}) \leq \chi_{\text{chron}}^{\mathcal{U}}(\mu_t).$$

and

$$\lim_{n \to \mathcal{U}} d_{W,\text{class}}(\mu_t^{(n)}, \sigma^{(n)}) \geq d_W(\mu_t, \sigma).$$

Proof of EVI

Apply the classical EVI:

$$\frac{1}{2}\left[d_{W,\text{class}}(\mu_t^{(n)},\sigma^{(n)})^2 - d_{W,\text{class}}(\mu_0^{(n)},\sigma^{(n)})^2\right] \leq t\left[h^{(n)}(\mu_t^{(n)}) - h^{(n)}(\sigma^{(n)})\right].$$

Then using each of the equalities and inequalities above, take the limit as $n \to \mathcal{U}$. We then obtain

$$\frac{1}{2}\left[d_W(\mu_t,\sigma)^2 - d_W(\mu_0,\sigma)^2\right] \leq t(\chi_{\text{chron}}^{\mathcal{U}}(\mu_t) - \chi_{\text{chron}}^{\mathcal{U}}(\sigma)),$$

which proves the EVI for $\chi_{\text{chron}}^{\mathcal{U}}$.

Remark: The fact that the formulas involve the filtration and Brownian motion is used here in order for the chron. type of $\mathbf{x} + \mathbf{z}_t$ to be determined from the chron. type of $\mathbf{x}$.

Monge–Kantorovich duality

Next, I will go into the proof of the Extension Theorem. To make this work, we need non-commutative optimal transport theory.

Theorem (MK duality for chronological types) [31, 33]

Let $\mu, \nu \in \mathbb{S}_{\text{chron},m}(\mathrm{T}_{\mathcal{U}})$. Then there exists chron. definable predicates φ, ψ (chron. definable predicates are the completion of the space of chron. formulas) such that

- $\varphi^{\mathcal{M}}$ and $\psi^{\mathcal{M}}$ are convex functions for all $\mathcal{M}$.
- $\varphi^{\mathcal{M}}(\mathbf{x}) + \psi^{\mathcal{M}}(\mathbf{y}) \geq \mathrm{Re}\langle \mathbf{x}, \mathbf{y} \rangle_2$ for all $\mathcal{M}$, $\mathbf{x}$, $\mathbf{y}$.
- Equality is achieved when $(\mathbf{x}, \mathbf{y})$ is an optimal coupling of (μ, ν).

Remark: The proof of this is not too hard. However, it really requires the ability to take partial suprema and infima of formulas. The naïve analog for plain NC laws is false (but this can be studied with a larger class of convex functions [17]).

Proof of Extension Theorem (with some lies)

Fix μ and ν and let φ and ψ be as in the MK duality. Suppose μ and ν are types of elements with operator norm $\leq R$.

Construct appropriate f.d. approx. $\Lambda_\varphi^{(n)}$ and $\Lambda_\psi^{(n)}$, which can be chosen to be convex.

Let $F^{(n)} : \mathbb{M}_n^m \to \mathbb{M}_n^m$ be a Borel function which for each $\mathbf{X}$ selects a maximizer $\mathbf{Y} = F^{(n)}(\mathbf{X})$ of

$$\mathrm{Re}\langle \mathbf{X}, \mathbf{Y} \rangle_2 - \Lambda_\psi^{(n)}(\mathbf{Y}) \text{ over } \|\mathbf{Y}\| \leq R,$$

which can be done by the Kuratowski–Ryll-Nardzewski selection theorem.

Given a random matrix model $\mathbf{X}^{(n)}$ for μ, let $\mathbf{Y}^{(n)} = F^{(n)}(\mathbf{X}^{(n)})$.

Problems in the naïve framework

In the usual setting of free probability, we don't work with model-theoretic formulas and types, but simply with non-commutative laws, or tracial states on the polynomial algebra $\mathbb{C}^*\langle x_1, \dots, x_m \rangle$.

Biane and Voiculescu's Wasserstein distance [4] considers pairs $(\mathbf{x}, \mathbf{y})$ from *any* tracial von Neumann algebra ($\mathcal{M}$ is allowed to vary).

Free entropy in this case can also be characterized using the limits of classical entropy of random matrix models such that the traces of non-commutative polynomials converge in probability to prescribed law [35].

Problems in the naïve framework

Proposition [32]

There exist some NC laws μ and ν such that there **do not** exist random matrix models $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$ such that

- $\mathrm{tr}(f(\mathbf{X}^{(n)})) \to \mu(f)$ and $\mathrm{tr}(f(\mathbf{Y}^{(n)})) \to \nu(f)$ for NC polynomials f.
- $h^{(n)}(\mathbf{Y}^{(n)}) \to \chi(\mathbf{Y})$.
- $\|\mathbf{X}^{(n)} - \mathbf{Y}^{(n)}\|_{L^2}$ converges to the (Connes-embeddable) Wasserstein distance of μ and ν.

As we'll see from the argument, the issue is more fundamental than merely about finite-dimensional approximations; the contradiction arises simply from the behavior of entropy under conditioning and change of variables.

In this discussion, we assume the tuples are self-adjoint.

Problems in the naïve framework

- S_1 and S_2 are freely independent semicirculars.
- S_3 is a semicircular in tensor position with S_1 and S_2.
  Group theorists: Think $F_2 \times \mathbb{Z}$.
- S_1', S_2', and S_3' are free semcirculars, free from S_1, S_2, and S_3.
  Group theorists: Think $(F_2 \times \mathbb{Z}) * F_3$
- $X_j = (1-\varepsilon)^{1/2} S_j + \varepsilon^{1/2} S_j'$ for $j = 1, 2, 3$.
- $\mathbf{Y} = (S_1', S_2', \varepsilon S_3')$.
- μ and ν are NC laws of $\mathbf{X}$ and $\mathbf{Y}$.
- $\chi(\nu) = (3/2)\log 2\pi e + (1/2)\log \varepsilon$.
- $1 - \varepsilon \leq d_W(\mu, \nu) \leq 1 - \varepsilon^{3/2}$. The upper bound is found by finding a coupling where $(X_1, X_2) = (Y_1, Y_2)$.

Problems in the naïve framework

Main idea: This coupling for $\mathbf{X}$ and $\mathbf{Y}$ is badly behaved because X_3 almost commutes with Y_1 and Y_2, but models for $\mathbf{Y}$ with high entropy should not allow anything to approximately commute with Y_1 and Y_2.

Suppose we have matrix models $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$. We give an upper bound on $h^{(n)}(\mathbf{Y}_1^{(n)}, \mathbf{Y}_2^{(n)} \mid \mathbf{X}_3^{(n)})$ showing that it is "too small."

By a standard bound for entropy in terms of variance,

$$h^{(n)}(i[Y_1^{(n)}, X_3^{(n)}], i[Y_2^{(n)}, X_3^{(n)}] \mid X_3^{(n)}) \leq \frac{1}{2} \sum_{j=1}^{2} \log \|[Y_j^{(n)}, X_3^{(n)}]\|_{L^2}^2 + \log(2\pi e)$$

Problems in the naïve framework

We use the change of variables formula for entropy:

$$h^{(n)}(i[Y_1^{(n)}, X_3^{(n)}], i[Y_2^{(n)}, X_3^{(n)}] \mid \mathbf{X}_3^{(n)}) \\ = h^{(n)}(Y_1^{(n)}, Y_2^{(n)} \mid X_3^{(n)}) + 2 \log \det |X_3^{(n)} \otimes 1 - 1 \otimes X_3^{(n)}|.$$

Since $X_3^{(n)}$ approximates a semicircular,

$$\log \det |X_3^{(n)} \otimes 1 - 1 \otimes X_3^{(n)}| \approx \int \log |s - t| \, d\sigma(s) \, d\sigma(t) \approx \text{constant},$$

where σ is the semicircular measure. This is the same as the free entropy of a single semicircular.

Problems in the naïve framework

Thus, the upper bound on the entropy of $i[Y_j^{(n)}, X_3^{(n)}]$ leads to an upper bound on the entropy of $Y_j^{(n)}$.

We also use that for $j = 1, 2$, we have

$$\begin{aligned} \|[Y_j^{(n)}, X_3^{(n)}]\|_2 &\leq \|[X_j^{(n)}, X_3^{(n)}]\|_2 + C\|X_j^{(n)} - Y_j^{(n)}\|_2 \\ &\leq C\varepsilon^{1/2} + C\|X_j^{(n)} - Y_j^{(n)}\|_2. \end{aligned}$$

After this, we apply the chain rule for conditioning $h(A, B) = h(A \mid B) + h(B)$ several times and estimate everything carefully . . .

Overall, if $\|X_j^{(n)} - Y_j^{(n)}\|_{L^2}$ is small enough to get close to the theoretical Wasserstein distance, we get an upper bound on $h^{(n)}(\mathbf{Y}^{(n)})$.

Problems in the naïve framework

Conclusion: Allowing couplings of non-commutative laws in arbitrary ways is simply wrong; it gives *too small* of a Wasserstein distance. We need to incorporate some information about how $\mathbf{x}$ and $\mathbf{y}$ sit inside the larger von Neumann algebra (using formulas with sup and inf).

Remark: This example also implies that the MK duality cannot hold for non-commutative laws in the naïve way, because otherwise, the proof of the Extension Theorem would work for non-commutative laws, causing a contradiction.

Chronological formulas and entropy: Motivation

In order to evaluate the entropy of random matrix models / large deviations rate function, we need to consider matrix integrals like the ones below; compare [5] and [19, §6].

Proposition (Corollary of Borell or Boué-Dupuis)

Let $\mathbf{Z}_t^{(n)}$ 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

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

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

Chronological formulas and entropy: motivation

Question: What is the large-n limit of the stochastic optimum, especially without assuming convexity of f?

Construction (Gangbo–J.–Nam–Palmer [18])

Let f be a trace polynomial, or more generally, a formula. Let M be a tracial von Neumann algebra, $(M_t)_{t\in[0,\infty)}$ a filtration, and compatible $\mathbf{z}_t$ m-variable free Brownian motion. Consider

$$u(\mathbf{x}) = \inf_{\alpha} \left[\frac{1}{2} \int_0^1 \|\alpha_t\|^2 \, dt + f\left(\mathbf{z}_1 + \int_0^1 \alpha_t \, dt \right) \right],$$

where $\mathbf{x} \in M_0^m$ and where α ranges over measurable maps $[0,1] \to L^2(M)$ such that $\alpha(t) \in L^2(M_t)$ for all t.

This is clearly an analog of the finite-dimensional problem, but which filtration and Brownian motion should we use?

Which filtration?

Which filtration and Brownian motion should we use?

Largest candidate filtration [18]: We consider all possible choices of filtration and Brownian motion, in any larger von Neumann algebra containing $\mathbf{x}$. This will make the infimum as small as possible.

Smallest candidate filtration: Fix M_0 generated by $\mathbf{x}$. Let M_t be the von Neumann algebra generated by $\mathbf{x}$ and the free Brownian motion up to time t.

If f is E-convex as defined in [17], these give the same answer. In general they do not, e.g. because of non-Connes-embeddable algebras.

Remark: This gives some indication of why the question of $\chi = \chi^*$ is easier for matrix models associated to a *convex* potential [10, 28].

Which filtration?

The work of Gao–J. [20] allows construction of a filtration as an ultraproduct of the random matrix models. Let $\mathcal{F}_t$ be a classical filtration on $(\Omega, \mathcal{F}, \mathbb{P})$, associated with a matrix Brownian motion $\mathbf{Z}_t^{(n)}$. Let

$$\mathcal{M} = \prod_{n \to \mathcal{U}} L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \otimes \mathbb{M}_n$$

$$\mathcal{M}_t = \prod_{n \to \mathcal{U}} L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}) \otimes \mathbb{M}_n$$

$$\mathbf{z}_t = [\mathbf{Z}_t^{(n)}]_{n \in \mathbb{N}}.$$

Let $\pi : \mathcal{M} \to \mathcal{Q}$ be a quotient by a maximal ideal, and let $\mathcal{Q}_t = \pi(\mathcal{M}_t)$. Thus, $\mathcal{Q}$ will have trivial center.

Claim: $\mathcal{Q}_t$ and $\pi(\mathbf{z}_t)$ are a useful choice of filtration and Brownian motion.

Which filtration?

Theorem (essentially Gao–J. [20])

- $\mathcal{Q}_t$ is a II_1 factor (trivial center).
- $\mathcal{Q}_t$ is elementarily equivalent to $\prod_{n \to \mathcal{U}} \mathbb{M}_n$ (hence under continuum hypothesis they are isomorphic).
- $(\pi(\mathbf{z}_t))_{t \in [0,\infty)}$ is a free Brownian motion compatible with $(\mathcal{Q}_t)_{t \in [0,\infty)}$.

Idea: In random matrix theory, we often use concentration of measure to obtain almost sure convergence. We could thus fix $\omega \in \Omega$ such that $[\mathbf{Z}_t^{(n)}(\omega)]_{n \in \mathbb{N}}$ is a free Brownian motion in $\prod_{n \to \mathcal{U}} \mathbb{M}_n$, but it is unclear what the filtration should be ...

Instead, in this construction, we take *first* take the ultraproduct and *then* derandomize by taking the quotient.

Stochastic optimization and chronological formulas

We approximate the optimization problem

$$\inf_{\alpha} \left[\frac{1}{2} \int_0^1 \|\alpha_t\|_2^2 \, dt + f \left(\mathbf{z}_1 + \int_0^1 \alpha_t \, dt \right) \right]$$

with a discretized version for $k \in \mathbb{N}$ and $R > 0$:

$$\inf_{\beta_1 \in B_R(M_{0/k})} \dots \inf_{\beta_k \in B_R(M_{(k-1)/k})} \left[\frac{1}{2k} \sum_{j=1}^{k} \|\beta_j\|_2^2 + f \left(\mathbf{z}_1 + \frac{1}{k} \sum_{j=1}^{k} \beta_j \right) \right].$$

The latter is a chronological formula! We can also make uniform estimates on how close the value is to the continuous-time and $R = \infty$ problem if we assume that f is $\|\cdot\|_1$-Lipschitz.

Idea: Chronological formulas contain enough information to evaluate the entropy, so using them as test functions produces a better behaved theory.

F.d. approximations of chronological formulas

For convenience, we consider *restricted* chronological formulas, which are actually formulas in the resolvents $(\mathrm{Re}(x_j) + i)^{-1}$ and $(\mathrm{Im}(x_j) + i)^{-1}$, in order to produce functions that are globally Lipschitz with respect to $\|\cdot\|_1$.

Proposition [33]

For every $\varphi \in \mathcal{F}^0_{\mathrm{chron}}$ with free variables among $x_1, \ldots, x_m$, and for every $n \in \mathbb{N}$, there exists $\Lambda^{(n)}_\varphi : \mathbb{M}_n^m \to \mathbb{R}$ bounded and $\|\cdot\|_1$-Lipschitz such that

$$\varphi^{\mathcal{Q}}(x_1, \ldots, x_m) = \pi([\Lambda^{(n)}_\varphi(X_1^{(n)}, \ldots, X_m^{(n)})]_{n \in \mathbb{N}}),$$

where

$$x_j = [X_j^{(n)}]_{n \in \mathbb{N}} \in \mathcal{M}_0, \qquad [\Lambda^{(n)}_\varphi(X_1^{(n)}, \ldots, X_m^{(n)})]_{n \in \mathbb{N}} \in Z(\mathcal{M}_0).$$

$\Lambda^{(n)}_\varphi$ are chosen independently of $\mathcal{U}$, π, and $(\mathcal{G}_t)_{t \geq 0}$.

F.d. approximations of chronological formulas

In a moment, I'll describe the construction of Λ_φ. I need the following notation:

If $\mathcal{M} = (M_t)_{t \geq 0}$ is any filtration, then the *shifted filtration* is $S_{t_0}\mathcal{M} = (M_{t+t_0})_{t \geq 0}$.

If φ is a chronological formula, then the shifted formula $S_{t_0}\varphi$ is given by shifting all the time indices to the right by t_0, which leads to

$$(S_t\varphi)^{\mathcal{M}}(\mathbf{x}) = \varphi^{S_t\mathcal{M}}(\mathbf{x}).$$

In the shift, we replace $\mathbf{z}_t$ by $\mathbf{z}_{t_0+t} - \mathbf{z}_{t_0}$.

F.d. approximations of chronological formulas

The f.d. approximations are constructed in the way you might expect, following the recursive construction of formulas with variables in M_0:

- If φ is a basic formula (i.e. a trace of a polynomial in $(\mathrm{Re}(x_j) + i)^{-1}$ and $(\mathrm{Im}(x_j) + i)^{-1}$), then $\Lambda_\varphi^{(n)} = \varphi^{\mathbb{M}_n}$.
- Suppose that $\varphi(\mathbf{x}) = \sup_{y \in B_R(M_0)} \psi(\mathbf{x}, y)$. Then
$$\Lambda_\varphi^{(n)}(\mathbf{X}) = \sup_{\mathbf{Y} \in B_R(\mathbb{M}_n)} \Lambda_\psi^{(n)}(\mathbf{X}, \mathbf{Y}).$$
- Suppose that $\varphi = F(\varphi_1, \dots, \varphi_k)$ where F is a continuous connective. Then
$$\Lambda_\varphi^{(n)} = F(\Lambda_{\varphi_1}^{(n)}, \dots, \Lambda_{\varphi_k}^{(n)}).$$
- Suppose that $\varphi(\mathbf{x}) = (S_t\psi)(\mathbf{x}, \mathbf{z}_t)$. Then
$$\Lambda_\varphi^{(n)}(\mathbf{X}) = \mathbb{E}\Lambda_\psi^{(n)}(\mathbf{X}, \mathbf{Z}_t^{(n)}).$$

F.d. approximations of chronological formulas

Remarks:

- The functions $\Lambda_\varphi^{(n)} : \mathbb{M}_n^m \to \mathbb{R}$ are deterministic, despite the fact that the formulas φ are defined in terms of a filtration.
- In the last item, we need to use concentration of measure to show that $\mathbb{E}\Lambda_\psi^{(n)}(\mathbf{X}, \mathbf{Z}_t^{(n)})$ is close to its expectation with high probability. This is what allows the reduction to deterministic $\Lambda_\varphi^{(n)}$'s.
- In order to apply the concentration argument, it is essential that the formulas can be generated by the operations above, which is only possible because the quantifiers are assumed to occur in chronological order.
- The choice of $\Lambda_\varphi^{(n)}$ is somewhat non-unique since the continuous connectives and the substitution of Brownian motion can be done in different orders, but asymptotically it is unique.

Definition of entropy

Let $\Lambda_\varphi^{(n)}$ be as above. Let $\mathbf{x}$ be a tuple from Q_0 for some filtration $\mathcal{Q}$ (for instance, this could be the matrix quotient filtration above).

Definition [33]

Let Φ be a finite subset of $\mathcal{F}^0_{\text{chron},m}$ and $\varepsilon > 0$. Define

$$\Gamma^{(n)}(\mathbf{x}; \Phi, \varepsilon) := \left\{ \mathbf{X} \in \mathbb{M}_n^m : \max_{\varphi \in \Phi} |\Lambda_\varphi^{(n)}(\mathbf{X}) - \varphi^{\mathcal{Q}}(\mathbf{x})| < \varepsilon \right\}.$$

Let $\sigma^{(n)} \in \mathcal{P}(\mathbb{M}_n^m)$ be the Ginibre Gaussian distribution. Define

$$\tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{x}; \Phi, \varepsilon) := \lim_{n \to \mathcal{U}} -\frac{1}{n^2} \log \sigma^{(n)}(\Gamma^{(n)}(\mathbf{x}; \Phi, \varepsilon))$$

and

$$\tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{x}) = \sup_{\Phi, \varepsilon} \tilde{\chi}^{\mathcal{U}}(\mathbf{x}; \Phi, \varepsilon)$$

Entropy and pressure

For $\varphi \in \mathcal{F}^0_{\mathrm{chron},m}$, let

$$\mathcal{P}^{(n)}(\varphi) := -\frac{1}{n^2} \log \frac{1}{C^{(n)}} \int_{\mathbb{M}_n^m} e^{-n^2(\frac{1}{2}\|\mathbf{X}\|_2^2 + \Lambda_\varphi^{(n)}(\mathbf{X}))}\, d\mathbf{X}$$
$$= \inf_\alpha \mathbb{E}\left[\frac{1}{2}\int_0^1 \|\alpha_t\|_2^2\, dt + \Lambda_\varphi^{(n)}\left(\mathbf{Z}_1^{(n)} + \int_0^1 \alpha_t\, dt\right)\right].$$

Let

$$\mathcal{P}^{\mathcal{U}}(\varphi) = \inf_\alpha \left[\frac{1}{2}\int_0^1 \|\alpha_t\|_2^2\, dt + \varphi^{\mathcal{Q}}\left(\mathbf{z}_1 + \int_0^1 \alpha_t\, dt\right)\right].$$

Entropy and pressure

Theorem [33]

1. We have $\lim_{n \to \mathcal{U}} \mathcal{P}^{(n)}(\varphi) = \mathcal{P}^{\mathcal{U}}(\varphi)$.
2. For a tuple $\mathbf{x}$ in Q_0 for a filtration $\mathcal{Q}$, we have

$$\tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{x}) = \sup_{\varphi \in \mathcal{F}^0_{\text{chron},m}} \left[\mathcal{P}(\varphi) - \varphi^{\mathcal{Q}}(\mathbf{x}) \right].$$

Item (1) is proved using the discrete approximation for the stochastic optimization problem, together with all the development of $\Lambda_\varphi^{(n)}$ and the matrix ultraproduct filtration.

Item (2) is proved by standard methods in large deviations. Compare Hiai's work on free pressure [26].

Conditional entropy

This entropy can be generalized to a conditional setting. Fix an $m + m'$-tuple $(\mathbf{x}, \mathbf{y})$. Fix deterministic matrix approximations $\mathbf{Y}^{(n)}$ for $\mathbf{y}$.

Definition [33]

Let Φ be a finite subset of $\mathcal{F}^0_{\mathrm{chron},m+m'}$ and $\varepsilon > 0$. Define

$$\Gamma^{(n)}(\mathbf{x} \mid \mathbf{Y}^{(n)} \rightsquigarrow \mathbf{y}; \Phi, \varepsilon) := \left\{ \mathbf{X} \in \mathbb{M}_n^m : \max_{\varphi \in \Phi} |\Lambda_\varphi^{(n)}(\mathbf{X}, \mathbf{Y}^{(n)}) - \varphi^{\mathcal{Q}}(\mathbf{x}, \mathbf{y})| < \varepsilon \right\}.$$

Let $\sigma^{(n)} \in \mathcal{P}(\mathbb{M}_n^m)$ be the Ginibre Gaussian distribution. Define

$$\tilde{\chi}^{\mathcal{U}}_{\mathrm{chron}}(\mathbf{x} \mid \mathbf{Y}^{(n)} \rightsquigarrow \mathbf{y}; \Phi, \varepsilon) := \lim_{n \to \mathcal{U}} -\frac{1}{n^2} \log \sigma^{(n)}(\Gamma^{(n)}(\mathbf{x} \mid \mathbf{Y}^{(n)} \rightsquigarrow \mathbf{y}; \Phi, \varepsilon))$$

and

$$\tilde{\chi}^{\mathcal{U}}_{\mathrm{chron}}(\mathbf{x} \mid \mathbf{Y}^{(n)} \rightsquigarrow \mathbf{y}) = \sup_{\Phi, \varepsilon} \tilde{\chi}^{\mathcal{U}}(\mathbf{x}; \Phi, \varepsilon)$$

Conditional entropy

In the conditional setting, we end up with pressure functionals that depend on $\mathbf{y}$, namely

$$\mathcal{P}^{(n)}(\varphi)(\mathbf{Y}^{(n)}), \qquad \mathcal{P}^{\mathcal{U}}(\varphi)(\mathbf{y}).$$

Since the pressure is given by a stochastic optimization problem which is a uniform limit of chronological formulas, we have that $\mathcal{P}^{(n)}(\varphi)(\mathbf{Y}^{(n)}) \to \mathcal{P}^{\mathcal{U}}(\varphi)(\mathbf{y})$.

Theorem [33]

$\tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{x} \mid \mathbf{Y}^{(n)} \rightsquigarrow \mathbf{y})$ is independent of the particular choice of matrix approximations $\mathbf{Y}^{(n)}$.

Remark: We cannot expect this to be true for plain free entropy if $\mathrm{W}^*(\mathbf{y})$ is not amenable, since there can be embeddings with different behaviors.

Conditional entropy

Because of the independence of the choice of matrix approximations, and the expression in terms of pressure, we obtain an actual chain rule for entropy and not just an inequality as in [45].

Theorem [33]

$$\tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{x}, \mathbf{y} \mid \mathbf{w}) = \tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{x} \mid \mathbf{y}, \mathbf{w}) + \tilde{\chi}^{\mathcal{U}}_{\text{chron}}(\mathbf{y} \mid \mathbf{w}).$$

Idea: Chronological formulas contain enough information to evaluate the entropy, so using them as test functions produces a better behaved theory.

Question: Does entropy depend on the ultrafilter?

We do not know whether the values of simple sentences in $\mathbb{M}_n$ converge as $n \to \infty$. For example, if f is a trace polynomial and

$$\varphi^{(n)} = \sup_{X \in B_1(\mathbb{M}_n)} \inf_{Y \in B_1(\mathbb{M}_n)} \operatorname{Re} \operatorname{tr}_n(f(X,Y)),$$

then we do not know if $\lim_{n \to \infty} \varphi^{(n)}$ exists.

- In the analogous situation for permutation groups equipped with the Hamming distance, the limit does not exist [1]!
- Peterson's work [43] implies that for some f, the limit $\lim_{n \to \mathcal{U}} \varphi^{(n)}$ does *not* agree with the corresponding variational formula in the free group von Neumann algebra $L(\mathbb{F}_m)$.

For GUE matrices to satisfy a large deviations principle as $n \to \infty$ is equivalent to the pressure $\mathcal{P}^{\mathcal{U}}(\varphi)$ being independent of the ultrafilter for all quantifier-free formulas φ.

Question: Is there a non-microstates approach?

This talk has focused on developing a theory that accurately describes the large-n behavior of entropy and Wasserstein distance for random matrix models, which is extremely difficult and subtle.

The weakness of the microstate framework is that it is largely unhelpful for studying non-Connes-embeddable von Neumann algebras.

The stochastic optimization framework does provide some hope of defining non-microstate versions, since it only requires a filtration and Brownian motion as input.

Problem: Define an entropy-like expression in terms of stochastic optimization problems, which also satisfies a change-of-variables formula.

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] Guillaume Cébron, Max Fathi, and Tobias Mai. A note on existence of free Stein kernels. Proceedings of the AMS 148: 1583-1594, 2020.

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

References III

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

[12] Sara Daneri and Giuseppe Savaré, Eulerian Calculus for the Displacement Convexity in the Wasserstein Distance, *SIAM Journal on Mathematical Analysis*, 40.3: 1104-1122 (2008).

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

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

[15] M. Fathi and B. Nelson. Free Stein kernels and an improvement of the free logarithmic Sobolev inequality. Adv. Math. 317 (2017), 193-223.

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

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

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

[19] Wilfrid Gangbo, David Jekel, Kyeongsik Nam, and Aaron Z. Palmer. Large-n limit of matrix control problems and non-commutative controls. arXiv:2511.22804

References V

[20] David Gao and David Jekel. Elementary equivalence and disintegration of tracial von Neumann algebras. Forum of Mathematics, Sigma 13: e105 (2025).

[21] 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

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

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

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

References VI

[25] 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.

[26] Fumio Hiai. Free Analog of Pressure and Its Legendre Transform. Commun. Math. Phys. 255:229-252 (2005).

[27] F. Hiai, D. Petz, and Y. Ueda, Free transportation cost inequalities via random matrix approximation, Probab. Theory Related Fields 130 (2004), 199-221.

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

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

References VII

[30] 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.

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

[32] David Jekel. Information geometry for types in the large-n limit of random matrices. *Commun. Math. Phys.* 406: art. 272 (2025).

[33] David Jekel. Free information geometry and the model theory of noncommutative stochastic processes. arXiv:2604.12212

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

References VIII

[35] David Jekel and Jennifer Pi. An elementary proof of the inequality $\chi \leq \chi^*$ for conditional free entropy. Doc. Math., 29(5):1085-1124, 2024.

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

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

[38] John D. Lafferty. The density manifold and configuration space quantization. Transactions of the American Mathematical Society, 305(2):699-741, 1988.

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

References IX

[40] Matteo Muratori and Giuseppe Savaré, Gradient flows and Evolution Variational Inequalities in metric spaces. I: Structural properties, *Journal of Functional Analysis* 278.4: 108347 (2020)

[41] Felix Otto. The geometry of dissipative evolution equations the porous medium equation. Communications in Partial Differential Equations, 26(1-2):101-174, 2001.

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

[43] Jesse Peterson. On ultraproduct approximations and property (T) factors. Preprint, arXiv:2605.16669

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

References X

[45] Dimitri Shlyakhtenko. A Microstates Approach to Relative Free Entropy, *Internat. J. Math.* 13.6: 605-623 (2002)

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

[47] 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.

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

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

References XI

[50] 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.

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