Free Entropy for Free Gibbs Laws Given by Convex Potentials

David A. Jekel

University of California, Los Angeles

University of Virginia Apr. 23, 2019

Motivation

We will discuss Voiculescu's *free entropy* of a *non-commutative law* μ *of an m-tuple of self-adjoint random variables*. This is an analogue in *free probability theory* of the *continuous entropy* of a probability measure $(\int -\rho \log \rho)$.

Motivation

We will discuss Voiculescu's *free entropy* of a *non-commutative law* μ *of an m-tuple of self-adjoint random variables*. This is an analogue in *free probability theory* of the *continuous entropy* of a probability measure ($\int -\rho \log \rho$).

Voiculescu defined two types of free entropy, $\chi(\mu)$ (*microstates*) and $\chi^*(\mu)$ (*non-microstates*). They both measure the "regularity" of the law μ.

Motivation

We will discuss Voiculescu's *free entropy* of a *non-commutative law* μ *of an m-tuple of self-adjoint random variables*. This is an analogue in *free probability theory* of the *continuous entropy* of a probability measure ($\int -\rho \log \rho$).

Voiculescu defined two types of free entropy, $\chi(\mu)$ (*microstates*) and $\chi^*(\mu)$ (*non-microstates*). They both measure the "regularity" of the law μ.

They are based on two different viewpoints for classical entropy: χ is based on the microstates interpretation of entropy and is defined by "counting" matrix approximations to μ, while χ^* is defined in terms of free Fisher information Φ^*, which describes how μ interacts with derivatives.

Motivation — von Neumann Algebras

Suppose that $X = (X_1, \dots, X_m)$ is a tuple of non-commutative self-adjoint random variables with law μ and $M = W^*(X_1, \dots, X_m)$.

Motivation — von Neumann Algebras

Suppose that $X = (X_1, \ldots, X_m)$ is a tuple of non-commutative self-adjoint random variables with law μ and $M = W^*(X_1, \ldots, X_m)$.

- Based on χ or χ^*, one defines various notions of free entropy dimension. This is something like the dimension of the support of a measure. (For instance, if χ is finite, then microstates free entropy dimension would be m, and the same for non-microstates versions.)

Motivation — von Neumann Algebras

Suppose that $X = (X_1, \dots, X_m)$ is a tuple of non-commutative self-adjoint random variables with law μ and $M = W^*(X_1, \dots, X_m)$.

- Based on χ or χ^*, one defines various notions of free entropy dimension. This is something like the dimension of the support of a measure. (For instance, if χ is finite, then microstates free entropy dimension would be m, and the same for non-microstates versions.)
- If $\delta_0(X) > 1$, then M has no Cartan subalgebras and no asymptotically central sequences [Voiculescu 1996]. In fact, if $\delta_0(X) > 0$, then M is prime [Ge 1998].

Motivation — von Neumann Algebras

Suppose that $X = (X_1, \dots, X_m)$ is a tuple of non-commutative self-adjoint random variables with law μ and $M = W^*(X_1, \dots, X_m)$.

- Based on χ or χ^*, one defines various notions of free entropy dimension. This is something like the dimension of the support of a measure. (For instance, if χ is finite, then microstates free entropy dimension would be m, and the same for non-microstates versions.)
- If $\delta_0(X) > 1$, then M has no Cartan subalgebras and no asymptotically central sequences [Voiculescu 1996]. In fact, if $\delta_0(X) > 0$, then M is prime [Ge 1998].
- If Φ^* is finite, then M has no asymptotically central sequences [Dabrowski 2010].

Motivation — von Neumann Algebras

Suppose that $X = (X_1, \dots, X_m)$ is a tuple of non-commutative self-adjoint random variables with law μ and $M = W^*(X_1, \dots, X_m)$.

- Based on χ or χ^*, one defines various notions of free entropy dimension. This is something like the dimension of the support of a measure. (For instance, if χ is finite, then microstates free entropy dimension would be m, and the same for non-microstates versions.)
- If $\delta_0(X) > 1$, then M has no Cartan subalgebras and no asymptotically central sequences [Voiculescu 1996]. In fact, if $\delta_0(X) > 0$, then M is prime [Ge 1998].
- If Φ^* is finite, then M has no asymptotically central sequences [Dabrowski 2010].
- $\delta^* = m$, then every non-constant self-adjoint polynomial in $X_1, \dots, X_m$ has diffuse spectral measure [Charlesworth-Shlyakhtenko 2016, Mai-Speicher-Weber 2017].

Motivation — von Neumann Algebras

Suppose that $X = (X_1, \dots, X_m)$ is a tuple of non-commutative self-adjoint random variables with law μ and $M = W^*(X_1, \dots, X_m)$.

- Based on χ or χ^*, one defines various notions of free entropy dimension. This is something like the dimension of the support of a measure. (For instance, if χ is finite, then microstates free entropy dimension would be m, and the same for non-microstates versions.)
- If $\delta_0(X) > 1$, then M has no Cartan subalgebras and no asymptotically central sequences [Voiculescu 1996]. In fact, if $\delta_0(X) > 0$, then M is prime [Ge 1998].
- If Φ^* is finite, then M has no asymptotically central sequences [Dabrowski 2010].
- $\delta^* = m$, then every non-constant self-adjoint polynomial in $X_1, \dots, X_m$ has diffuse spectral measure [Charlesworth-Shlyakhtenko 2016, Mai-Speicher-Weber 2017].

(List adapted from Charlesworth-Nelson 2019 "Free Stein Discrepancy.")

Motivation — von Neumann Algebras

- We hope that free entropy-related results will provide some sufficient or necessary conditions for M to be isomorphic to $L(\mathbb{F}_m)$, contain $L(\mathbb{F}_m)$, or be contained in $L(\mathbb{F}_m)$.

Motivation — von Neumann Algebras

- We hope that free entropy-related results will provide some sufficient or necessary conditions for M to be isomorphic to $L(\mathbb{F}_m)$, contain $L(\mathbb{F}_m)$, or be contained in $L(\mathbb{F}_m)$.
- Having χ, χ^*, or Φ^* finite does *not* imply that that M is a free group factor. Counterexamples are provided by $X + t^{1/2}S$ where $X = (X_1, \dots, X_m)$ generates a property (T) von Neumann algebra, S is a freely independent semicircular tuple, and t is sufficiently small.

Motivation — von Neumann Algebras

- We hope that free entropy-related results will provide some sufficient or necessary conditions for M to be isomorphic to $L(\mathbb{F}_m)$, contain $L(\mathbb{F}_m)$, or be contained in $L(\mathbb{F}_m)$.
- Having χ, χ^*, or Φ^* finite does *not* imply that that M is a free group factor. Counterexamples are provided by $X + t^{1/2}S$ where $X = (X_1, \dots, X_m)$ generates a property (T) von Neumann algebra, S is a freely independent semicircular tuple, and t is sufficiently small.
- Hayes has used a related notion of one-bounded free entropy to study one-bounded von Neumann algebras and maximal amenable subalgebras of free group factors.

Motivation — Random Matrix Theory

- We expect microstates free entropy to the be large N limit of normalized classical entropies of random matrix models that have good concentration, and free Fisher information to be the large N limit of normalized classical Fisher information.

Motivation — Random Matrix Theory

- We expect microstates free entropy to the be large N limit of normalized classical entropies of random matrix models that have good concentration, and free Fisher information to be the large N limit of normalized classical Fisher information.
- Microstates free entropy defines the rate function for a large deviation principle describing the Gaussian unitary ensemble (see Biane-Capitaine-Guionnet 2003).

Motivation — Random Matrix Theory

- We expect microstates free entropy to the be large N limit of normalized classical entropies of random matrix models that have good concentration, and free Fisher information to be the large N limit of normalized classical Fisher information.
- Microstates free entropy defines the rate function for a large deviation principle describing the Gaussian unitary ensemble (see Biane-Capitaine-Guionnet 2003).
- The results of this paper will be based on studying the asymptotic properties of random matrix models.

Motivation — Some Key Ideas

- We give more explicit statements of how free probability arises as the large N limit of classical probability, consistent with Voiculesu's original motivation.

Motivation — Some Key Ideas

- We give more explicit statements of how free probability arises as the large N limit of classical probability, consistent with Voiculesu's original motivation.
- Then we study the large N behavior of functions (e.g. solutions to PDE) related to the random matrix models and their entropy. We show that these functions are *asymptotically approximable by trace polynomials*.

Motivation — Some Key Ideas

- We give more explicit statements of how free probability arises as the large N limit of classical probability, consistent with Voiculesu's original motivation.
- Then we study the large N behavior of functions (e.g. solutions to PDE) related to the random matrix models and their entropy. We show that these functions are *asymptotically approximable by trace polynomials*.
- This means roughly that the behave asymptotically like a non-commutative function (e.g. NC polynomial rather than an entrywise function in the classical sense), and like the *same* non-commutative function for different values of N.

Motivation — Main Question

- Voiculescu: When does $\chi = \chi^*$?

Motivation — Main Question

- Voiculescu: When does $\chi = \chi^*$?
- Note χ being finite implies embeddability of M into $\mathcal{R}^\omega$, so $\chi = \chi^*$ always with no additional assumptions would solve Connes embedding. A more modest goal would be to prove $\chi = \chi^*$ this under the assumption that some nice enough matrix models exist.

Motivation — Main Question

- Voiculescu: When does $\chi = \chi^*$?
- Note χ being finite implies embeddability of M into $\mathcal{R}^\omega$, so $\chi = \chi^*$ always with no additional assumptions would solve Connes embedding. A more modest goal would be to prove $\chi = \chi^*$ this under the assumption that some nice enough matrix models exist.
- Biane-Capitaine-Guionet 2003 showed that $\chi \leq \chi^*$ always.

Motivation — Main Question

- Voiculescu: When does $\chi = \chi^*$?
- Note χ being finite implies embeddability of M into $\mathcal{R}^\omega$, so $\chi = \chi^*$ always with no additional assumptions would solve Connes embedding. A more modest goal would be to prove $\chi = \chi^*$ this under the assumption that some nice enough matrix models exist.
- Biane-Capitaine-Guionet 2003 showed that $\chi \leq \chi^*$ always.
- Dabrowski 2017 showed that $\chi = \chi^*$ for free Gibbs states given by nice enough convex potentials.

Motivation — Main Question

- Voiculescu: When does $\chi = \chi^*$?
- Note χ being finite implies embeddability of M into $\mathcal{R}^\omega$, so $\chi = \chi^*$ always with no additional assumptions would solve Connes embedding. A more modest goal would be to prove $\chi = \chi^*$ this under the assumption that some nice enough matrix models exist.
- Biane-Capitaine-Guionet 2003 showed that $\chi \leq \chi^*$ always.
- Dabrowski 2017 showed that $\chi = \chi^*$ for free Gibbs states given by nice enough convex potentials.
- The result of this paper is similar to Dabrowski's although our proof takes a PDE rather than SDE viewpoint.

What is non-commutative probability?

classical	non-commutative
$L^\infty(\Omega, P)$	W^*-algebra M
expectation E	trace τ
bdd. real rand. var. X	self-adjoint $X \in M$
law of X	spectral distribution of X w.r.t. τ

What is free probability?

We replace classical independence by *free independence*.

What is free probability?

We replace classical independence by *free independence*.

Definition by Example

For groups G_1 and G_2, the algebras $L(G_1)$ and $L(G_2)$ are freely independent in $(L(G_1 * G_2), \tau)$.

What is free probability?

We replace classical independence by *free independence*.

Definition by Example

For groups G_1 and G_2, the algebras $L(G_1)$ and $L(G_2)$ are freely independent in $(L(G_1 * G_2), \tau)$.

Free Central Limit Theorem: There's a free central limit theorem with normal distribution replaced by semicircular distribution.

What is free probability?

We replace classical independence by *free independence*.

Definition by Example

For groups G_1 and G_2, the algebras $L(G_1)$ and $L(G_2)$ are freely independent in $(L(G_1 * G_2), \tau)$.

Free Central Limit Theorem: There's a free central limit theorem with normal distribution replaced by semicircular distribution.

Free Convolution: If X and Y are classically independent, then $\mu_{X+Y} = \mu_X * \mu_Y$. If X and Y are freely independent, then $\mu_{X+Y} = \mu_X \boxplus \mu_Y$.

What is the law of a tuple?

Classically, the law of $X = (X_1, \ldots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

What is the law of a tuple?

Classically, the law of $X = (X_1, \ldots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

Assuming finite moments, this can be viewed as a map

$$\mu_X : \mathbb{C}[x_1, \ldots, x_m] \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto E[p(X_1, \ldots, X_m)].$$

What is the law of a tuple?

Classically, the law of $X = (X_1, \ldots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

Assuming finite moments, this can be viewed as a map

$$\mu_X : \mathbb{C}[x_1, \ldots, x_m] \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto E[p(X_1, \ldots, X_m)].$$

In the non-commutative case, the *law of* $X = (X_1, \ldots, X_m) \in M_{sa}^m$ is defined as the map

$$\mu_X : \mathbb{C}\langle x_1, \ldots, x_m \rangle \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto \tau[p(X_1, \ldots, X_m)],$$

What is the law of a tuple?

Classically, the law of $X = (X_1, \ldots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

Assuming finite moments, this can be viewed as a map

$$\mu_X : \mathbb{C}[x_1, \ldots, x_m] \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto E[p(X_1, \ldots, X_m)].$$

In the non-commutative case, the *law of* $X = (X_1, \ldots, X_m) \in M_{sa}^m$ is defined as the map

$$\mu_X : \mathbb{C}\langle x_1, \ldots, x_m \rangle \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto \tau[p(X_1, \ldots, X_m)],$$

The *moment topology* on laws is given by pointwise convergence on $\mathbb{C}\langle x_1, \ldots, x_m \rangle$.

Notation

τ_N is the normalized trace on $M_N(\mathbb{C})$.

$\|\cdot\|_2$ is the corresponding 2-norm, that is, for $x \in M_N(\mathbb{C})_{sa}^m$, we set $\|x\|_2^2 = \sum_{j=1}^m \tau_N(x_j^2)$.

$\|\cdot\|$ is the operator norm of a single matrix and $\|x\|_\infty$ denotes the maximum of the operator norms of $x_1, \dots, x_m$.

$\sigma_{N,t}$ denotes the law of m independent $N \times N$ GUE matrices which each have mean zero and variance t.

σ_t denotes the non-commutative law of m freely independent semicirculars which each have mean zero and variance t.

Asymptotic Approximation by Trace Polynomials

Trace Polynomials

Trace polynomials in $x_1, \dots, x_m$ are linear combinations of functions of the form $p_0 \tau(p_1) \dots \tau(p_n)$ where p_j is a non-commutative polynomial in $x_1, \dots, x_m$. For example,

$$\tau(x_1 x_2) x_1 + 3\tau(x_2^2)\tau(x_1 x_3) x_3 x_2 + 5\tau(x_3^2)$$

Trace Polynomials

Trace polynomials in $x_1, \ldots, x_m$ are linear combinations of functions of the form $p_0 \tau(p_1) \ldots \tau(p_n)$ where p_j is a non-commutative polynomial in $x_1, \ldots, x_m$. For example,

$$\tau(x_1 x_2) x_1 + 3\tau(x_2^2)\tau(x_1 x_3) x_3 x_2 + 5\tau(x_3^2)$$

If p is a trace polynomial, then p defines a function $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})$. We interpret τ as the normalized trace on $M_N(\mathbb{C})$ and evaluate p at the point x.

Trace Polynomials

Trace polynomials in $x_1, \dots, x_m$ are linear combinations of functions of the form $p_0 \tau(p_1) \dots \tau(p_n)$ where p_j is a non-commutative polynomial in $x_1, \dots, x_m$. For example,

$$\tau(x_1 x_2) x_1 + 3\tau(x_2^2)\tau(x_1 x_3) x_3 x_2 + 5\tau(x_3^2)$$

If p is a trace polynomial, then p defines a function $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})$. We interpret τ as the normalized trace on $M_N(\mathbb{C})$ and evaluate p at the point x.

More generally, if (M, τ) is a tracial von Neumann algebra, then p defines a map $M_{sa}^m \to M$.

Asymptotic Approximation by Trace Polynomials

Definition

A sequence of functions $\phi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ is *asymptotically approximable by trace polynomials* if for every $\epsilon > 0$ and $R > 0$, there exists an m-tuple of trace polynomials f such that

$$\limsup_{N \to \infty} \sup_{\substack{x \in M_N(\mathbb{C})_{sa}^m \\ \|x\|_\infty \leq R}} \|\phi_N(x) - f(x)\|_2 \leq \epsilon.$$

We make a similar definition for scalar-valued functions being approximated by scalar-valued trace polynomials.

Properties of AATP

Consider sequences of functions $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^n$ which are globally Lipschitz in $\|\cdot\|_2$.

Properties of AATP

Consider sequences of functions $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^n$ which are globally Lipschitz in $\|\cdot\|_2$.

- AATP is preserved under addition and scalar multiplication.

Properties of AATP

Consider sequences of functions $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^n$ which are globally Lipschitz in $\|\cdot\|_2$.

- AATP is preserved under addition and scalar multiplication.
- AATP is preserved under limits, that is, if $\{\phi_{N,k}\}_{N\in\mathbb{N}}$ converges to $\{\phi_N\}_{N\in\mathbb{N}}$ as $k \to \infty$ in a certain sense, then $\{\phi_N\}_{N\in\mathbb{N}}$ would inherit AATP.

Properties of AATP

Consider sequences of functions $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^n$ which are globally Lipschitz in $\|\cdot\|_2$.

- AATP is preserved under addition and scalar multiplication.
- AATP is preserved under limits, that is, if $\{\phi_{N,k}\}_{N \in \mathbb{N}}$ converges to $\{\phi_N\}_{N \in \mathbb{N}}$ as $k \to \infty$ in a certain sense, then $\{\phi_N\}_{N \in \mathbb{N}}$ would inherit AATP.
- AATP is preserved under convolution with the GUE law $\sigma_{t,N}$.

Properties of AATP

Consider sequences of functions $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^n$ which are globally Lipschitz in $\|\cdot\|_2$.

- AATP is preserved under addition and scalar multiplication.
- AATP is preserved under limits, that is, if $\{\phi_{N,k}\}_{N \in \mathbb{N}}$ converges to $\{\phi_N\}_{N \in \mathbb{N}}$ as $k \to \infty$ in a certain sense, then $\{\phi_N\}_{N \in \mathbb{N}}$ would inherit AATP.
- AATP is preserved under convolution with the GUE law $\sigma_{t,N}$.
- AATP is preserved under composition. This is easy to prove, though a little unexpected because the approximation occurs on an operator norm ball but the error is measured in $\|\cdot\|_2$.

Properties of AATP

Consider sequences of functions $M_N(\mathbb{C})^m_{sa} \to M_N(\mathbb{C})^n_{sa}$ which are globally Lipschitz in $\|\cdot\|_2$.

- AATP is preserved under addition and scalar multiplication.
- AATP is preserved under limits, that is, if $\{\phi_{N,k}\}_{N \in \mathbb{N}}$ converges to $\{\phi_N\}_{N \in \mathbb{N}}$ as $k \to \infty$ in a certain sense, then $\{\phi_N\}_{N \in \mathbb{N}}$ would inherit AATP.
- AATP is preserved under convolution with the GUE law $\sigma_{t,N}$.
- AATP is preserved under composition. This is easy to prove, though a little unexpected because the approximation occurs on an operator norm ball but the error is measured in $\|\cdot\|_2$.
- AATP is preserved under solving ODE. That is, if we have a vector field with AATP, then the flow along this vector field also has AATP.

Microstates Free Entropy χ

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

"Entropy measures regularity."

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

"Entropy measures regularity."

1. If μ is highly concentrated, then there is large negative entropy.

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

"Entropy measures regularity."

1. If μ is highly concentrated, then there is large negative entropy.
2. For mean zero and variance 1, the highest entropy is achieved by Gaussian.

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

"Entropy measures regularity."

1. If μ is highly concentrated, then there is large negative entropy.
2. For mean zero and variance 1, the highest entropy is achieved by Gaussian.
3. If you smooth μ out by convolution, the entropy increases.

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Classical case: Given a vector in $x = (x_1, \dots, x_m) \in (\mathbb{R}^N)^m$, let's define its *empirical distribution* as

$$\mu_x = \frac{1}{N} \sum_{j=1}^{N} \delta_{((x_1)_j, \dots, (x_m)_j)}.$$

Then $\{x : \mu_x \text{ is close to } \mu\}$ has measure approximately $\exp(-Nh(\mu))$.

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Classical case: Given a vector in $x = (x_1, \dots, x_m) \in (\mathbb{R}^N)^m$, let's define its *empirical distribution* as

$$\mu_x = \frac{1}{N} \sum_{j=1}^{N} \delta_{((x_1)_j, \dots, (x_m)_j)}.$$

Then $\{x : \mu_x \text{ is close to } \mu\}$ has measure approximately $\exp(-Nh(\mu))$. Thus, $h(\mu)$ can be expressed as

$$\inf_{\text{(nbhd's of } \mu)} \lim_{N \to \infty} \frac{1}{N} \log \operatorname{vol}\{x : \mu_x \text{ close to } \mu\}.$$

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Classical case: Given a vector in $x = (x_1, \ldots, x_m) \in (\mathbb{R}^N)^m$, let's define its *empirical distribution* as

$$\mu_x = \frac{1}{N} \sum_{j=1}^N \delta_{((x_1)_j, \ldots, (x_m)_j)}.$$

Then $\{x : \mu_x \text{ is close to } \mu\}$ has measure approximately $\exp(-Nh(\mu))$. Thus, $h(\mu)$ can be expressed as

$$\inf_{\text{(nbhd's of } \mu)} \lim_{N \to \infty} \frac{1}{N} \log \operatorname{vol}\{x : \mu_x \text{ close to } \mu\}.$$

Intuition: If μ is more regular and spread out, then there are more microstates because most choices of N vectors are "evenly distributed."

Microstates Free Entropy

Idea for free case: Replace $\mathbb{R}^N$ (self-adjoints in $L^\infty(\{1,\dots,N\})$) by $M_N(\mathbb{C})_{sa}$.

Microstates Free Entropy

Idea for free case: Replace $\mathbb{R}^N$ (self-adjoints in $L^\infty(\{1, \dots, N\})$) by $M_N(\mathbb{C})_{sa}$.

Given $(x_1, \dots, x_m) \in M_N(\mathbb{C})^m$, the *empirical distribution* μ_x is the non-commutative law of x w.r.t. normalized trace on $M_N(\mathbb{C})$. For a neighborhood $\mathcal{U}$ of μ in the moment topology and $R > 0$, define

$$\Gamma_{N,R}(\mathcal{U}) = \{x : \|x_j\| \leq R \text{ and } \mu \in \mathcal{U}\}.$$

Define

$$\chi(\mu) = \sup_{R>0} \inf_{\mathcal{U} \ni \mu} \limsup_{N \to \infty} \left(\frac{1}{N^2} \log \operatorname{vol} \Gamma_{N,R}(\mathcal{U}) + \frac{m}{2} \log N \right).$$

Microstates Free Entropy

Idea for free case: Replace $\mathbb{R}^N$ (self-adjoints in $L^\infty(\{1, \dots, N\})$) by $M_N(\mathbb{C})_{sa}$.

Given $(x_1, \dots, x_m) \in M_N(\mathbb{C})^m$, the *empirical distribution* μ_x is the non-commutative law of x w.r.t. normalized trace on $M_N(\mathbb{C})$. For a neighborhood $\mathcal{U}$ of μ in the moment topology and $R > 0$, define

$$\Gamma_{N,R}(\mathcal{U}) = \{x : \|x_j\| \leq R \text{ and } \mu \in \mathcal{U}\}.$$

Define

$$\chi(\mu) = \sup_{R>0} \inf_{\mathcal{U} \ni \mu} \limsup_{N \to \infty} \left(\frac{1}{N^2} \log \operatorname{vol} \Gamma_{N,R}(\mathcal{U}) + \frac{m}{2} \log N \right).$$

(Voiculescu) χ has properties similar to h.

Free Entropy as the Limit of Classical Entropy

Lemma

Suppose that $d\mu_N = e^{-N^2 V_N(x)}\,dx$*, where* $V_N : M_N(\mathbb{C})_{sa}^m \to \mathbb{R}$*. Suppose that* $|V_N(x)|$ *is bounded by a constant times* $1 + \|x\|^k$*, and that for some* R *we have* $\int_{\|x\|_\infty > R} (1 + \|x\|_\infty^k)\,d\mu_N(x) \to 0$ *as* $N \to \infty$*. Suppose that the law of* x *with respect to* τ_N *converges in probability to the non-commutative law* λ*. Then* $\chi(\lambda) = \limsup_{N\to\infty}(N^{-2}h(\mu_N) + (m/2)\log N)$*.*

Free Entropy as the Limit of Classical Entropy

The idea of the proof is as follows.

Free Entropy as the Limit of Classical Entropy

The idea of the proof is as follows.

- First, replace μ_N by μ_N truncated to an operator norm ball of radius R.

Free Entropy as the Limit of Classical Entropy

The idea of the proof is as follows.

- First, replace μ_N by μ_N truncated to an operator norm ball of radius R.
- For any given neighborhood $\mathcal{U}$ of λ, the measure μ_N will be concentrated on the microstate space $\Gamma_{N,R}(\mathcal{U})$.

Free Entropy as the Limit of Classical Entropy

The idea of the proof is as follows.

- First, replace μ_N by μ_N truncated to an operator norm ball of radius R.
- For any given neighborhood $\mathcal{U}$ of λ, the measure μ_N will be concentrated on the microstate space $\Gamma_{N,R}(\mathcal{U})$.
- $\{V_N\}$ can be approximated by a trace polynomial, which will be approximately constant on $\Gamma_{N,R}(\mathcal{U})$ if $\mathcal{U}$ is sufficiently small.

Free Entropy as the Limit of Classical Entropy

The idea of the proof is as follows.

- First, replace μ_N by μ_N truncated to an operator norm ball of radius R.
- For any given neighborhood $\mathcal{U}$ of λ, the measure μ_N will be concentrated on the microstate space $\Gamma_{N,R}(\mathcal{U})$.
- $\{V_N\}$ can be approximated by a trace polynomial, which will be approximately constant on $\Gamma_{N,R}(\mathcal{U})$ if $\mathcal{U}$ is sufficiently small.
- So the entropy of μ_N should be approximately the entropy of the uniform distribution on $\Gamma_{N,R}(\mathcal{U})$, which is the log volume.

Free Entropy as the Limit of Classical Entropy

The idea of the proof is as follows.

- First, replace μ_N by μ_N truncated to an operator norm ball of radius R.
- For any given neighborhood $\mathcal{U}$ of λ, the measure μ_N will be concentrated on the microstate space $\Gamma_{N,R}(\mathcal{U})$.
- $\{V_N\}$ can be approximated by a trace polynomial, which will be approximately constant on $\Gamma_{N,R}(\mathcal{U})$ if $\mathcal{U}$ is sufficiently small.
- So the entropy of μ_N should be approximately the entropy of the uniform distribution on $\Gamma_{N,R}(\mathcal{U})$, which is the log volume.
- Divide by N^2, add $(m/2)\log N$ and take the $\limsup$ as $N \to \infty$.

Non-microstates Free Entropy χ^*

Classical Fisher Information

Classical case: Let μ be a probability measure on $\mathbb{R}^m$ with density ρ. Let γ_t be the law of a Gaussian random vector with variance tI. Then

$$\frac{d}{dt} h(\mu * \gamma_t) = \int |\nabla \rho_t|^2 / \rho_t = \|\nabla \rho_t / \rho_t\|^2_{L^2(\mu * \gamma_t)}.$$

Classical Fisher Information

Classical case: Let μ be a probability measure on $\mathbb{R}^m$ with density ρ. Let γ_t be the law of a Gaussian random vector with variance tI. Then

$$\frac{d}{dt} h(\mu * \gamma_t) = \int |\nabla \rho_t|^2 / \rho_t = \|\nabla \rho_t / \rho_t\|_{L^2(\mu * \gamma_t)}^2.$$

The quantity $\|\nabla \rho_t / \rho_t\|_{L^2(\mu * \gamma_t)}^2$ is called the *Fisher information* of $\mu * \gamma_t$. The entropy can be recovered by integrating the Fisher information.

Classical Fisher Information

Classical case: Let μ be a probability measure on $\mathbb{R}^m$ with density ρ. Let γ_t be the law of a Gaussian random vector with variance tI. Then

$$\frac{d}{dt} h(\mu * \gamma_t) = \int |\nabla \rho_t|^2 / \rho_t = \|\nabla \rho_t / \rho_t\|_{L^2(\mu * \gamma_t)}^2.$$

The quantity $\|\nabla \rho_t / \rho_t\|_{L^2(\mu * \gamma_t)}^2$ is called the *Fisher information* of $\mu * \gamma_t$. The entropy can be recovered by integrating the Fisher information.

Intuition: The Fisher information measures the regularity of μ by looking at its derivatives.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Classical Fisher information is L^2 norm of the conjugate variable $\xi = (\nabla \rho / \rho)(X)$, which is characterized by an integration-by-parts formula $E[\xi f(X)] = E[\nabla f(X)]$.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Classical Fisher information is L^2 norm of the conjugate variable $\xi = (\nabla \rho / \rho)(X)$, which is characterized by an integration-by-parts formula $E[\xi f(X)] = E[\nabla f(X)]$.

Voiculescu used the free version $\tau[\xi_j f(X)] = \tau \otimes \tau[\mathcal{D}_{X_j} f(X)]$ to define the *free conjugate variables* and hence the *free Fisher information*.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Classical Fisher information is L^2 norm of the conjugate variable $\xi = (\nabla \rho / \rho)(X)$, which is characterized by an integration-by-parts formula $E[\xi f(X)] = E[\nabla f(X)]$.

Voiculescu used the free version $\tau[\xi_j f(X)] = \tau \otimes \tau[\mathcal{D}_{X_j} f(X)]$ to define the *free conjugate variables* and hence the *free Fisher information*.

$\chi^*(\mu)$ is defined by integrating the free Fisher information of $\mu \boxplus \sigma_t$, where σ_t is the law of a free semicircular family where each variable has mean zero and variance t.

Convergence of Fisher Information

In the case where $d\mu_N(x) = (1/Z_N)e^{-N^2 V_N(x)}\,dx$, the classical conjugate variables would be DV_N (up to normalization). So the normalized Fisher information would be $\int \|DV_N\|_2^2 \, d\mu_N$.

Convergence of Fisher Information

In the case where $d\mu_N(x) = (1/Z_N)e^{-N^2 V_N(x)}\,dx$, the classical conjugate variables would be DV_N (up to normalization). So the normalized Fisher information would be $\int \|DV_N\|_2^2\,d\mu_N$.

Lemma

Let μ_N be given by the potential V_N. Suppose that $\|DV_N(x)\|_2^2$ is bounded by a constant times $1 + \|x\|^k$, and that for some R we have $\int_{\|x\|_\infty > R}(1 + \|x\|_\infty^k)\,d\mu_N(x) \to 0$ as $N \to \infty$. Suppose that the law of x with respect to τ_N converges in probability to the non-commutative law λ. If $\{DV_N\}$ has AATP, then the (normalized) classical Fisher information converges to the free Fisher information (and the latter is finite).

Convergence of Fisher Information

Sketch of proof:

Convergence of Fisher Information

Sketch of proof:

- Suppose that f_k is a sequence of trace polynomials which as k increases provide better and better asymptotic approximations for DV_N.

Convergence of Fisher Information

Sketch of proof:

- Suppose that f_k is a sequence of trace polynomials which as k increases provide better and better asymptotic approximations for DV_N.
- Then f_k will converge in $L^2(\lambda)$ to some f.

Convergence of Fisher Information

Sketch of proof:

- Suppose that f_k is a sequence of trace polynomials which as k increases provide better and better asymptotic approximations for DV_N.
- Then f_k will converge in $L^2(\lambda)$ to some f.
- Also, f is a free conjugate variable for λ since the f_k's approximately satisfy the integration by parts formula.

Convergence of Fisher Information

Sketch of proof:

- Suppose that f_k is a sequence of trace polynomials which as k increases provide better and better asymptotic approximations for DV_N.
- Then f_k will converge in $L^2(\lambda)$ to some f.
- Also, f is a free conjugate variable for λ since the f_k's approximately satisfy the integration by parts formula.
- Then we check that $\|DV_N\|_{L^2(\mu_N)} \to \|f\|_{L^2(\lambda)}$.

Main Results and Strategy

The Upshot

The following conditions would be sufficient for $\chi(\lambda) = \chi^*(\lambda)$:

The Upshot

The following conditions would be sufficient for $\chi(\lambda) = \chi^*(\lambda)$:

- We have matrix models μ_N such that the laws of the random matrices given by μ converge in probability to the non-commutative law λ.

The Upshot

The following conditions would be sufficient for $\chi(\lambda) = \chi^*(\lambda)$:

- We have matrix models μ_N such that the laws of the random matrices given by μ converge in probability to the non-commutative law λ.
- V_N and DV_N do not grow too fast as $\|x\| \to \infty$. We have reasonable tail bounds on the probability of large operator norm under μ_N.

The Upshot

The following conditions would be sufficient for $\chi(\lambda) = \chi^*(\lambda)$:

- We have matrix models μ_N such that the laws of the random matrices given by μ converge in probability to the non-commutative law λ.
- V_N and DV_N do not grow too fast as $\|x\| \to \infty$. We have reasonable tail bounds on the probability of large operator norm under μ_N.
- V_N and DV_N are asymptotically approximable by trace polynomials.

The Upshot

The following conditions would be sufficient for $\chi(\lambda) = \chi^*(\lambda)$:

- We have matrix models μ_N such that the laws of the random matrices given by μ converge in probability to the non-commutative law λ.
- V_N and DV_N do not grow too fast as $\|x\| \to \infty$. We have reasonable tail bounds on the probability of large operator norm under μ_N.
- V_N and DV_N are asymptotically approximable by trace polynomials.
- The laws $\mu_N * \sigma_{t,N}$ satisfy all the same conditions.

The Upshot

The following conditions would be sufficient for $\chi(\lambda) = \chi^*(\lambda)$:

- We have matrix models μ_N such that the laws of the random matrices given by μ converge in probability to the non-commutative law λ.
- V_N and DV_N do not grow too fast as $\|x\| \to \infty$. We have reasonable tail bounds on the probability of large operator norm under μ_N.
- V_N and DV_N are asymptotically approximable by trace polynomials.
- The laws $\mu_N * \sigma_{t,N}$ satisfy all the same conditions.

Indeed, in the case, $\chi(\lambda)$ would be the $\limsup$ of the classical entropies. Since the classical Fisher information of $\mu_N * \sigma_{t,N}$ would converge to the free Fisher information of $\lambda \boxplus \sigma_t$, then the classical entropy would also converge to $\chi^*(\lambda)$.

The Upshot

When can we check these conditions are true?

The Upshot

When can we check these conditions are true?

- Operator norm tail bounds for μ_N would follow from exponential concentration for $\|\cdot\|_2$-Lipschitz functions (e.g. coming from the log-Sobolev inequality), provided that the expectation of μ_N is a multiple of the identity matrix [Guionnet and Maurel-Segala].

The Upshot

When can we check these conditions are true?

- Operator norm tail bounds for μ_N would follow from exponential concentration for $\|\cdot\|_2$-Lipschitz functions (e.g. coming from the log-Sobolev inequality), provided that the expectation of μ_N is a multiple of the identity matrix [Guionnet and Maurel-Segala].
- Given this concentration of measure, the convergence of the NC law in probability as $N \to \infty$ would be equivalent to convergence in expectation.

The Upshot

When can we check these conditions are true?

- Operator norm tail bounds for μ_N would follow from exponential concentration for $\|\cdot\|_2$-Lipschitz functions (e.g. coming from the log-Sobolev inequality), provided that the expectation of μ_N is a multiple of the identity matrix [Guionnet and Maurel-Segala].
- Given this concentration of measure, the convergence of the NC law in probability as $N \to \infty$ would be equivalent to convergence in expectation.
- The log-Sobolev inequality and exponential concentration are known to hold provided that V_N is uniformly convex ($HV_N \geq c$ for some $c > 0$ independent of N). [Bakry-Emery, Herbst, Ledoux, etc.]

The Upshot

- If $\{DV_N\}$ is asymptotically approximable by trace polynomials, then so is $\{V_N - V_N(0)\}$. You just integrate your approximating polynomial for DV_N along the straight-line path from 0 to x.

The Upshot

- If $\{DV_N\}$ is asymptotically approximable by trace polynomials, then so is $\{V_N - V_N(0)\}$. You just integrate your approximating polynomial for DV_N along the straight-line path from 0 to x.
- Concentration, convergence in expectation, and tail bounds are preserved under convolution by Gaussian. This is another lemma that is not too difficult.

Main Goals

Suppose we're given potentials $\{V_N\}$. Assume uniform convexity of V_N and that $\{DV_N\}$ is AATP. Then we want to show two claims:

Main Goals

Suppose we're given potentials $\{V_N\}$. Assume uniform convexity of V_N and that $\{DV_N\}$ is AATP. Then we want to show two claims:

Claim 1

If DV_N is AATP, then $\int \tau_N(p)\, d\mu_N$ converges as $N \to \infty$ for any non-commutative polynomial p.

Main Goals

Suppose we're given potentials $\{V_N\}$. Assume uniform convexity of V_N and that $\{DV_N\}$ is AATP. Then we want to show two claims:

Claim 1

If DV_N is AATP, then $\int \tau_N(p)\, d\mu_N$ converges as $N \to \infty$ for any non-commutative polynomial p.

Hence, there *is* some non-commutative law λ that arises as the large-N limit. We could take this as the definition of a free Gibbs state.

Main Goals

Suppose we're given potentials $\{V_N\}$. Assume uniform convexity of V_N and that $\{DV_N\}$ is AATP. Then we want to show two claims:

Claim 1

If DV_N is AATP, then $\int \tau_N(p)\, d\mu_N$ converges as $N \to \infty$ for any non-commutative polynomial p.

Hence, there *is* some non-commutative law λ that arises as the large-N limit. We could take this as the definition of a free Gibbs state.

In the special case where $V_N(x) = V(x) = \tau_N(p(x))$ for a fixed p that is a small or convex perturbation of quadratic, the existence and uniqueness of a NC law with conjugate variables $DV(x)$ was shown in works of Guionnet, Maurel-Segala, Shlyaktenko, Dabrowski. They also deduce convergence of certain random matrix models.

Main Goals

Claim 2

If DV_N is AATP, then the same holds for $DV_{N,t}$, where $V_{N,t}$ is the potential corresponding to $\mu_N * \sigma_{N,t}$.

Main Goals

Claim 2

If DV_N is AATP, then the same holds for $DV_{N,t}$, where $V_{N,t}$ is the potential corresponding to $\mu_N * \sigma_{N,t}$.

If we can prove this, then $\chi(\lambda) = \chi^*(\lambda)$. Also, it's equal to the limit of the normalized classical entropies.

Main Result

Theorem

Let $V_N(x) - (c/2)\|x\|_2^2$ is convex and $V_N(x) - (C/2)\|x\|_2^2$ is concave for some $0 < c < C$. Let $d\mu_N(x) = \frac{1}{Z_N} e^{-N^2 V_N(x)}\,dx$. Suppose $\{DV_N\}$ is AATP. Suppose that the expectation of μ_N is bounded in operator norm as $N \to \infty$. Then

1. $\mu(p) := \lim_{N\to\infty} \int \tau_N(p(x))\,d\mu_N(x)$ *exists for every non-commutative polynomial p.*
2. *The non-commutative law λ has finite free Fisher information and finite free entropy.*
3. $\chi(\lambda) = \chi^*(\lambda) = \lim_{N\to\infty}[N^{-2}h(\mu_N) + (m/2)\log N]$.
4. *The normalized Fisher information of $\mu_N * \sigma_{N,t}$ converges to the free Fisher information of $\mu \boxplus \sigma_t$ for every $t \geq 0$.*
5. *The free Fisher information is locally Lipschitz in t.*

Some of the Proof

Evolution of Potentials

Let's focus on the proof of Claim 2 (assuming Claim 1), since Claim 2 is harder and more interesting.

Evolution of Potentials

Let's focus on the proof of Claim 2 (assuming Claim 1), since Claim 2 is harder and more interesting.

Let $\mu_{N,t} = \mu_N * \sigma_{N,t}$ and let $V_{N,t}$ be the potential such that the density of $\mu_{N,t}$ is $(1/Z_N)e^{-N^2 V_{N,t}}$.

Evolution of Potentials

Let's focus on the proof of Claim 2 (assuming Claim 1), since Claim 2 is harder and more interesting.

Let $\mu_{N,t} = \mu_N * \sigma_{N,t}$ and let $V_{N,t}$ be the potential such that the density of $\mu_{N,t}$ is $(1/Z_N)e^{-N^2 V_{N,t}}$.

We know that the density of $\mu_{N,t}$ evolves according to the heat equation (with $(1/2N)\Delta$), but this does not immediately help us analyze $DV_{N,t}$ asymptotically because of the dimension-dependent factor of N^2 in the exponent.

Evolution of Potentials

Thus, we rewrite the equation in terms of $V_{N,t}$:

$$\partial_t V_{N,t} = \frac{1}{2N} \Delta V_{N,t} - \frac{1}{2} \|DV_{N,t}\|_2^2.$$

This is the normalization of the Laplacian that corresponds to convolution with GUE. So this is a dimension-independent equation for free probabilistic normalization.

Evolution of Potentials

Thus, we rewrite the equation in terms of $V_{N,t}$:

$$\partial_t V_{N,t} = \frac{1}{2N} \Delta V_{N,t} - \frac{1}{2} \| D V_{N,t} \|_2^2.$$

This is the normalization of the Laplacian that corresponds to convolution with GUE. So this is a dimension-independent equation for free probabilistic normalization.

Using PDE tools and the convexity assumptions, we will "build" an approximation to $V_{N,t}$ by taking V_N and applying nice explicit operations that preserve AATP (that is, AATP for the gradient of V rather than V itself).

Approximation of Solutions

As heuristic, recall that to solve the equation

$$\partial_t v = \frac{1}{2N} \Delta v,$$

we would use the Gaussian convolution semigroup $P_t v = v * \sigma_{N,t}$.

Approximation of Solutions

As heuristic, recall that to solve the equation

$$\partial_t v = \frac{1}{2N} \Delta v,$$

we would use the Gaussian convolution semigroup $P_t v = v * \sigma_{N,t}$.

To solve the equation

$$\partial_t v = -\frac{1}{2} \|Dv\|_2^2,$$

we would use the Hopf-Lax inf-convolution semigroup

$$Q_t v(x) = \inf_y \left[v(y) + \frac{1}{2t} \|x - y\|_2^2 \right].$$

(This is a well-known fact in PDE.)

Approximation of Solutions

The solution $V_{N,t}$ can be obtained by combining these operations together:

$$V_{N,t} = \lim_{k \to \infty} (P_{t/k} Q_{t/k})^k V_N.$$

The paper gives an elementary but technical argument for this, which we will not explain in detail.

Approximation of Solutions

The solution $V_{N,t}$ can be obtained by combining these operations together:

$$V_{N,t} = \lim_{k \to \infty} (P_{t/k} Q_{t/k})^k V_N.$$

The paper gives an elementary but technical argument for this, which we will not explain in detail. But note the following key points.

- The error estimates are dimension-independent. *This is the whole point!*

Approximation of Solutions

The solution $V_{N,t}$ can be obtained by combining these operations together:

$$V_{N,t} = \lim_{k \to \infty} (P_{t/k} Q_{t/k})^k V_N.$$

The paper gives an elementary but technical argument for this, which we will not explain in detail. But note the following key points.

- The error estimates are dimension-independent. *This is the whole point!*
- The regularizing properties of P_t are not used here because they disappear as $N \to \infty$.

Approximation of Solutions

The solution $V_{N,t}$ can be obtained by combining these operations together:

$$V_{N,t} = \lim_{k \to \infty} (P_{t/k} Q_{t/k})^k V_N.$$

The paper gives an elementary but technical argument for this, which we will not explain in detail. But note the following key points.

- The error estimates are dimension-independent. *This is the whole point!*
- The regularizing properties of P_t are not used here because they disappear as $N \to \infty$.
- It relies on the fact that P_t and Q_t preserve the space of functions with $0 \leq Hv \leq C$, and for such functions the gradient is automatically C-Lipchitz.

Approximation of Solutions

The solution $V_{N,t}$ can be obtained by combining these operations together:

$$V_{N,t} = \lim_{k \to \infty} (P_{t/k} Q_{t/k})^k V_N.$$

The paper gives an elementary but technical argument for this, which we will not explain in detail. But note the following key points.

- The error estimates are dimension-independent. *This is the whole point!*
- The regularizing properties of P_t are not used here because they disappear as $N \to \infty$.
- It relies on the fact that P_t and Q_t preserve the space of functions with $0 \leq Hv \leq C$, and for such functions the gradient is automatically C-Lipchitz.
- The proof goes by showing that the limit exists as k ranges over powers of 2, and the limit is a viscosity solution.

Inf-Convolution Preserves AATP

Lemma

Let $0 \leq Hu_N \leq C$. If $\{Du_N\}$ is AATP, then so is $\{D(Q_t u_N)\}$.

Inf-Convolution Preserves AATP

Lemma

Let $0 \leq Hu_N \leq C$. If $\{Du_N\}$ is AATP, then so is $\{D(Q_t u_N)\}$.

Proof.

The inf-convolution $Q_t u$ is differentiable and satisfies

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x))$$

(This is derived from the fact that the minimizer in the definition of $Q_t u$ has to be a critical point.)

Inf-Convolution Preserves AATP

Lemma

Let $0 \leq Hu_N \leq C$. *If* $\{Du_N\}$ *is AATP, then so is* $\{D(Q_t u_N)\}$.

Proof.

The inf-convolution $Q_t u$ is differentiable and satisfies

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x))$$

(This is derived from the fact that the minimizer in the definition of $Q_t u$ has to be a critical point.) Thus, $D(Q_t u)(x)$ is a fixed point of $y \mapsto Du(x - ty)$, which is a contraction mapping when $t < 1/C$.

Inf-Convolution Preserves AATP

Lemma

Let $0 \leq Hu_N \leq C$. *If* $\{Du_N\}$ *is AATP, then so is* $\{D(Q_t u_N)\}$.

Proof.

The inf-convolution $Q_t u$ is differentiable and satisfies

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x))$$

(This is derived from the fact that the minimizer in the definition of $Q_t u$ has to be a critical point.) Thus, $D(Q_t u)(x)$ is a fixed point of $y \mapsto Du(x - ty)$, which is a contraction mapping when $t < 1/C$. So $D(Q_t u)(x)$ can be obtained as $\lim_{n \to \infty} \phi_n(x)$ where $\phi_0(x) = x$ and $\phi_{n+1}(x) = Du(x - t\phi_n(x))$, and the rate of convergence is dimension-independent.

Inf-Convolution Preserves AATP

Lemma

Let $0 \leq Hu_N \leq C$. *If* $\{Du_N\}$ *is AATP, then so is* $\{D(Q_t u_N)\}$.

Proof.

The inf-convolution $Q_t u$ is differentiable and satisfies

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x))$$

(This is derived from the fact that the minimizer in the definition of $Q_t u$ has to be a critical point.) Thus, $D(Q_t u)(x)$ is a fixed point of $y \mapsto Du(x - ty)$, which is a contraction mapping when $t < 1/C$. So $D(Q_t u)(x)$ can be obtained as $\lim_{n \to \infty} \phi_n(x)$ where $\phi_0(x) = x$ and $\phi_{n+1}(x) = Du(x - t\phi_n(x))$, and the rate of convergence is dimension-independent. Since AATP is preserved by composition and limits, the claim holds for $t < 1/C$.

Inf-Convolution Preserves AATP

Lemma

Let $0 \leq Hu_N \leq C$. If $\{Du_N\}$ is AATP, then so is $\{D(Q_t u_N)\}$.

Proof.

The inf-convolution $Q_t u$ is differentiable and satisfies

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x))$$

(This is derived from the fact that the minimizer in the definition of $Q_t u$ has to be a critical point.) Thus, $D(Q_t u)(x)$ is a fixed point of $y \mapsto Du(x - ty)$, which is a contraction mapping when $t < 1/C$. So $D(Q_t u)(x)$ can be obtained as $\lim_{n \to \infty} \phi_n(x)$ where $\phi_0(x) = x$ and $\phi_{n+1}(x) = Du(x - t\phi_n(x))$, and the rate of convergence is dimension-independent. Since AATP is preserved by composition and limits, the claim holds for $t < 1/C$. But Q_t preserves the class of functions with $0 \leq Hu \leq C$ and Q_t is a semigroup, so the claim holds for all t. □