Quantum OT: quantum channels and qubits
Tutorial 3

Dario Trevisan

Università di Pisa
dario.trevisan@unipi.it

March 13, 2025

Plan

1. OT on the Hamming cube
2. W_1 on qubits
3. Quantum Lipschitz constant
4. W_1 continuity of Shannon Entropy
5. Gaussian concentration inequalities
6. Bibliography

Plan

The Hamming cube

The discrete cube $\{0,1\}^n$ (i.e., bit strings of length n) can be endowed with

1. the Euclidean distance $|s-t|_E = \sqrt{\sum_{i=1}^n |s_i - t_i|^2}$, or
2. the Hamming distance $|s-t|_H = \sum_{i=1}^n |s_i - t_i| = \sum_{i=1}^n 1_{\{s_i \neq t_i\}}$.

We have in fact $|s-t|_E = |s-t|_H^{1/2}$, hence they give different geometries for large n.

We collect some features of the Wasserstein distance, w.r.t. the Hamming distance, to be later extended to quantum systems of n-qubits.

Given two probabilities σ, ρ over $\{0,1\}^n$, write

$$W_1(\sigma,\rho) = \min_T \sum_{s,t\in\{0,1\}^n} |t-s|_H T(s,t),$$

where T are transport plans between σ and ρ.

Example: let σ be uniform on $\{0,1\}^n$, $\sigma(s) = \frac{1}{2^n}$, while let $\rho(\mathbf{0}) = 1$. Then,

$$W_1(\sigma,\rho) = \frac{1}{2^n} \sum_{s\in\{0,1\}^n} \sum_{i=1}^n |s_i| = \frac{n}{2}.$$

The dual formulation is

$$W_1(\sigma,\rho) = \max_f \left\{ \sum_{s\in\{0,1\}^n} f(s)(\rho(s)-\sigma(s)) \right\}$$

where $f : \{0,1\}^n \to \mathbb{R}$ are Lipschitz with respect to the Hamming distance, i.e.,

$$|f(s)-f(t)| \leq |s-t|_H = \sum_{i=1}^n 1_{\{s_i\neq t_i\}}.$$

Behaviour with respect to local transformations

- If σ and ρ have same $(n-1)$-marginals, e.g.,

$$\sigma(0, s) + \sigma(1, s) = \rho(0, s) + \rho(1, s) \quad \text{for every } s \in \{0, 1\}^{n-1},$$

  then

$$W_1(\sigma, \rho) \leq 1.$$

- More generally, if σ and ρ have the same $(n-k)$-marginals, then

$$W_1(\sigma, \rho) \leq k.$$

- If σ is the law of a random variable X and ρ is the law of $Q(X)$, where $Q : \{0, 1\}^n \to \{0, 1\}^n$ acts only on k positions (e.g., Q is a transformation of the first k positions), then

$$W_1(\sigma, \rho) \leq k.$$

Comparison with Total Variation distance

The following inequalities hold for any σ, ρ on $\{0,1\}^n$:

$$\|\sigma - \rho\|_{TV} \leq W_1(\sigma, \rho) \leq n\|\sigma - \rho\|_{TV}.$$

Indeed TV is Wasserstein w.r.t distance $1_{\{s \neq t\}}$ (recall the exercise):

$$|s - t|_{TV} \leq |s - t|_H \leq n|t - s|_{TV}.$$

Plan

Definition of quantum W_1

On the system of n qubits $\mathcal{H}_n = \left(\mathbb{C}^2\right)^{\otimes n}$, we define the distance W_1 essentially by postulating that a state ρ is at distance ≤ 1 from a state σ if there exists $i = 1, \ldots, n$ such that

$$\mathrm{Tr}_i[\sigma] = \mathrm{Tr}_i[\rho].$$

The distance $W_1(\rho, \sigma) = \|\rho - \sigma\|_{W_1}$ will be induced by a norm on the subspace of self-adjoint linear operators on $(\mathbb{C}^2)^{\otimes n}$ with null trace.

We define the unit ball (centred at 0) as the convex envelope

$$\mathcal{B}_n = \left\{ \sum_{i=1}^n p_i \left(\rho^{(i)} - \sigma^{(i)}\right) : p_i \geq 0,\ \sum_{i=1}^n p_i = 1,\ \rho^{(i)}, \sigma^{(i)} \text{states}, \mathrm{Tr}_i[\rho^{(i)}] = \mathrm{Tr}_i[\sigma^{(i)}] \right\}$$

and the norm via the Minkowski functional:

$$\|X\|_{W_1} = \min\left(t \geq 0 : X \in t\,\mathcal{B}_n\right).$$

We obtain the explicit definition for the distance

$$\|\sigma-\rho\|_{W_1} = \min\left\{ \sum_{i=1}^{n} c_i : \sigma - \rho = \sum_{i=1}^{n} c_i(\sigma^{(i)} - \rho^{(i)}), c_i \geq 0, \mathrm{Tr}_i[\sigma^{(i)}] = \mathrm{Tr}_i[\rho^{(i)}] \right\}.$$

W_1 is invariant with respect to

- permutations of the qubits,
- unitary operations U acting on a single qubit (e.g., Hadamard gate):

$$\|\sigma - \rho\|_{W_1} = \|U\sigma U^\dagger - U\rho U^\dagger\|_{W_1}.$$

Exercise:

1. Show that (10) fails for $U = CX$ gate (acting on $n \geq 2$ qubits).

Simple properties

- for states σ, ρ on m qubits and τ on n qubits:

$$\|\sigma \otimes \tau - \rho \otimes \tau\|_{W_1} \leq \|\sigma - \rho\|_{W_1}.$$

- Consider states σ, ρ over $m+n$ qubits with reduced density operators

$$\sigma_{1,\ldots,m}, \quad \rho_{1,\ldots,n}, \quad \sigma_{m+1,\ldots,m+n}, \quad \rho_{m+1,\ldots,m+n}.$$

  Then,

$$\|\sigma - \rho\|_{W_1} \geq \|\sigma_{1,\ldots,m} - \rho_{1,\ldots,m}\|_{W_1} + \|\sigma_{m+1,\ldots,m+n} - \rho_{m+1,\ldots,m+n}\|_{W_1}$$

- Equality holds if

$$\sigma = \sigma_{1,\ldots,m} \otimes \sigma_{m+1,\ldots,m+n} \quad \text{and} \quad \rho = \rho_{1,\ldots,m} \otimes \rho_{m+1,\ldots,m+n}.$$

Relation with trace distance

We have the inequalities

$$\frac{1}{2}\|\sigma - \rho\|_1 \leq \|\sigma - \rho\|_{W_1} \leq \frac{n}{2}\|\sigma - \rho\|_1.$$

Exercise:

1. if there exists i such that $\mathrm{Tr}_i[\rho] = \mathrm{Tr}_i[\sigma]$ then

$$\|\sigma - \rho\|_{W_1} = \frac{1}{2}\|\sigma - \rho\|_1.$$

Recovery of classical distance

For diagonal states

$$\sigma = \sum_{s \in \{0,1\}^n} p(s) |s\rangle\langle s|, \qquad \rho = \sum_{s \in \{0,1\}^n} q(s) |s\rangle\langle s|$$

we recover the Wasserstein distance of order 1 w.r.t. Hamming distance:

$$\|\rho - \sigma\|_{W_1} = W_1(p, q).$$

- Tensorization for product states gives

$$\| |x\rangle\langle x| - |y\rangle\langle y| \|_{W_1} = \sum_{i=1}^{n} 1_{\{x_i \neq y_i\}} = |x - y|_H.$$

- Given a (classical) transport plan π, between p and q, the triangle inequality yields

$$\|\rho - \sigma\|_{W_1} = \| \sum_{x,y} \pi(x,y) \left(|x\rangle\langle x| - |y\rangle\langle y| \right) \|_{W_1} \leq \sum_{x,y} \pi(x,y) |x - y|_H.$$

- This proves inequality $\leq$, i.e., quantum W_1 is cheaper.

To prove that it is not strictly cheaper:

- Given $c_i \geq 0$, states $\rho^{(i)}$, $\sigma^{(i)}$ such that

$$\sigma - \rho = \sum_{i=1}^{n} c_i(\sigma^{(i)} - \rho^{(i)}),$$

  discard the off-diagonals of $\rho^{(i)}$ and $\sigma^{(i)}$, yielding probabilities $q^{(i)}$, $p^{(i)}$ with

$$p - q = \sum_{i=1}^{n} c_i(p^{(i)} - q^{(i)}).$$

- The $n-1$ marginals of $p^{(i)}$ and $q^{(i)}$ (discarding the bit i) coincide, hence

$$W_1(p^{(i)}, q^{(i)}) \leq 1.$$

- Since also the classical W_1 is induced by a norm, by triangle inequality

$$W_1(p, q) \leq \sum_{i=1}^{n} c_i W_1(p^{(i)}, q^{(i)}) \leq \sum_{i=1}^{n} c_i.$$

Remarks

A“Monge problem” without transport plans:

$$\|\rho-\sigma\|_{W_1} = \min\left\{ \sum_{i=1}^n c_i : \sigma - \rho = \sum_{i=1}^n c_i(\sigma^{(i)} - \rho^{(i)}), c_i \geq 0, \mathrm{Tr}_i[\sigma^{(i)}] = \mathrm{Tr}_i[\rho^{(i)}] \right\}.$$

- Computational aspects: Since $\rho^{(i)}$ and $\sigma^{(i)}$ are states, computing W_1 is a positive semidefinite problem. The computational cost grows however with the number of qubits (the dimension is 2^n).
- Generalizations: We actually cover the case of systems $H = (\mathbb{C}^d)^{\otimes n}$, possibly one can generalize to $d \to \infty$. We can also consider the limit $n \to \infty$ (spin systems).
- We search for a dual formulation

$$\|\rho - \sigma\|_{W_1} = \max_H \{\mathrm{Tr}[H(\rho - \sigma)] : \|H\|_L \leq 1\},$$

with $\|H\|_L$ a quantum notion of Lipschitz constant of H.

Remarks

A"Monge problem" without transport plans:

$$\|\rho - \sigma\|_{W_1} = \min \left\{ \sum_{i=1}^n c_i : \sigma - \rho = \sum_{i=1}^n c_i (\sigma^{(i)} - \rho^{(i)}), c_i \geq 0, \mathrm{Tr}_i[\sigma^{(i)}] = \mathrm{Tr}_i[\rho^{(i)}] \right\}.$$

- Computational aspects: Since $\rho^{(i)}$ and $\sigma^{(i)}$ are states, computing W_1 is a positive semidefinite problem. The computational cost grows however with the number of qubits (the dimension is 2^n).

- Generalizations: We actually cover the case of systems $H = (\mathbb{C}^d)^{\otimes n}$, possibly one can generalize to $d \to \infty$. We can also consider the limit $n \to \infty$ (spin systems).

- We search for a dual formulation

$$\|\rho - \sigma\|_{W_1} = \max_H \{ \mathrm{Tr}[H(\rho - \sigma)] : \|H\|_L \leq 1 \},$$

with $\|H\|_L$ a quantum notion of Lipschitz constant of H.

Remarks

A"Monge problem" without transport plans:

$$\|\rho - \sigma\|_{W_1} = \min \left\{ \sum_{i=1}^{n} c_i : \sigma - \rho = \sum_{i=1}^{n} c_i(\sigma^{(i)} - \rho^{(i)}), c_i \geq 0, \mathrm{Tr}_i[\sigma^{(i)}] = \mathrm{Tr}_i[\rho^{(i)}] \right\}.$$

- Computational aspects: Since $\rho^{(i)}$ and $\sigma^{(i)}$ are states, computing W_1 is a positive semidefinite problem. The computational cost grows however with the number of qubits (the dimension is 2^n).
- Generalizations: We actually cover the case of systems $H = (\mathbb{C}^d)^{\otimes n}$, possibly one can generalize to $d \to \infty$. We can also consider the limit $n \to \infty$ (spin systems).
- We search for a dual formulation

$$\|\rho - \sigma\|_{W_1} = \max_H \{\mathrm{Tr}[H(\rho - \sigma)] : \|H\|_L \leq 1\},$$

with $\|H\|_L$ a quantum notion of Lipschitz constant of H.

Remarks

A“Monge problem” without transport plans:

$$\|\rho-\sigma\|_{W_1} = \min\left\{\sum_{i=1}^{n} c_i : \sigma - \rho = \sum_{i=1}^{n} c_i(\sigma^{(i)} - \rho^{(i)}), c_i \geq 0, \mathrm{Tr}_i[\sigma^{(i)}] = \mathrm{Tr}_i[\rho^{(i)}]\right\}.$$

- Computational aspects: Since $\rho^{(i)}$ and $\sigma^{(i)}$ are states, computing W_1 is a positive semidefinite problem. The computational cost grows however with the number of qubits (the dimension is 2^n).
- Generalizations: We actually cover the case of systems $H = (\mathbb{C}^d)^{\otimes n}$, possibly one can generalize to $d \to \infty$. We can also consider the limit $n \to \infty$ (spin systems).
- We search for a dual formulation

$$\|\rho - \sigma\|_{W_1} = \max_H \{\mathrm{Tr}[H(\rho - \sigma)] \, : \, \|H\|_L \leq 1\},$$

with $\|H\|_L$ a quantum notion of Lipschitz constant of H.

Plan

The classical case

In the classical case, if $f : \{0,1\}^n \to \mathbb{R}$

$$\|f\|_L = \max_{x \neq y} \frac{|f(y) - f(x)|}{|x - y|_H}.$$

We obtain functions with bounded differences:

$$|f(x) - f(y)| \leq \|f\|_L k$$

if the strings $x, y \in \{0,1\}^n$ differ in k positions.

Exercise: to compute $\|f\|_L$ it is sufficient to bound $|f(x) - f(y)|$ for strings x, y differing in 1 position (argue along a “discrete geodesic”).

Using Kantorovich duality for W_1, we can also write

$$\left| \sum_x f(x)(p(x) - q(x)) \right| \leq \|f\|_L W_1(p, q).$$

The classical case

In the classical case, if $f : \{0,1\}^n \to \mathbb{R}$

$$\|f\|_L = \max_{x \neq y} \frac{|f(y) - f(x)|}{|x - y|_H}.$$

We obtain functions with bounded differences:

$$|f(x) - f(y)| \leq \|f\|_L k$$

if the strings $x, y \in \{0,1\}^n$ differ in k positions.

Exercise: to compute $\|f\|_L$ it is sufficient to bound $|f(x) - f(y)|$ for strings x, y differing in 1 position (argue along a “discrete geodesic”).

Using Kantorovich duality for W_1, we can also write

$$\left| \sum_x f(x)(p(x) - q(x)) \right| \leq \|f\|_L W_1(p, q).$$

The classical case

In the classical case, if $f : \{0,1\}^n \to \mathbb{R}$

$$\|f\|_L = \max_{x \neq y} \frac{|f(y) - f(x)|}{|x - y|_H}.$$

We obtain functions with bounded differences:

$$|f(x) - f(y)| \leq \|f\|_L k$$

if the strings $x, y \in \{0,1\}^n$ differ in k positions.

Exercise: to compute $\|f\|_L$ it is sufficient to bound $|f(x) - f(y)|$ for strings x, y differing in 1 position (argue along a "discrete geodesic").

Using Kantorovich duality for W_1, we can also write

$$\left| \sum_x f(x)(p(x) - q(x)) \right| \leq \|f\|_L W_1(p,q).$$

The classical case

In the classical case, if $f : \{0,1\}^n \to \mathbb{R}$

$$\|f\|_L = \max_{x \neq y} \frac{|f(y) - f(x)|}{|x - y|_H}.$$

We obtain functions with bounded differences:

$$|f(x) - f(y)| \leq \|f\|_L k$$

if the strings $x, y \in \{0,1\}^n$ differ in k positions.

Exercise: to compute $\|f\|_L$ it is sufficient to bound $|f(x) - f(y)|$ for strings x, y differing in 1 position (argue along a “discrete geodesic”).

Using Kantorovich duality for W_1, we can also write

$$\left| \sum_x f(x)(p(x) - q(x)) \right| \leq \|f\|_L W_1(p, q).$$

Definition of quantum Lipschitz constant

We define $\|H\|_L$ as the dual norm to W_1, for any self-adjoint operator H on $(\mathbb{C}^2)^{\otimes n}$:

$$\|H\|_L = \max_{\rho,\sigma} \{\mathrm{Tr}[H(\rho - \sigma)] \, : \, \|\rho - \sigma\|_{W_1} \leq 1\} \, .$$

Since we work in finite dimensional spaces, duality is straightforward:

$$\|\rho - \sigma\|_{W_1} = \max_{H} \{\mathrm{Tr}[H(\rho - \sigma)] \, : \, \|H\|_L \leq 1\} \, .$$

But we also have an operational definition of H:

$$\|H\|_L = 2 \max_{i=1,\ldots,n} \min_{H^{(i)}} \left\| H - \mathbb{I}^{(i)} \otimes H^{(i)} \right\|_{\infty} \, ,$$

where $\mathbb{I}^{(i)}$ is the identity on the i-th qubit and $H^{(i)}$ does not act on the i-th qubit, and $\| \cdot \|_\infty$ denotes the operator norm (dual to $\| \cdot \|_1$).

Definition of quantum Lipschitz constant

We define $\|H\|_L$ as the dual norm to W_1, for any self-adjoint operator H on $(\mathbb{C}^2)^{\otimes n}$:

$$\|H\|_L = \max_{\rho,\sigma} \left\{ \mathrm{Tr}[H(\rho - \sigma)] \, : \, \|\rho - \sigma\|_{W_1} \leq 1 \right\}.$$

Since we work in finite dimensional spaces, duality is straightforward:

$$\|\rho - \sigma\|_{W_1} = \max_{H} \left\{ \mathrm{Tr}[H(\rho - \sigma)] \, : \, \|H\|_L \leq 1 \right\}.$$

But we also have an operational definition of H:

$$\|H\|_L = 2 \max_{i=1,\dots,n} \min_{H^{(i)}} \left\| H - \mathbb{I}^{(i)} \otimes H^{(i)} \right\|_\infty ,$$

where $\mathbb{I}^{(i)}$ is the identity on the i-th qubit and $H^{(i)}$ does not act on the i-th qubit, and $\|\cdot\|_\infty$ denotes the operator norm (dual to $\|\cdot\|_1$).

Definition of quantum Lipschitz constant

We define $\|H\|_L$ as the dual norm to W_1, for any self-adjoint operator H on $(\mathbb{C}^2)^{\otimes n}$:

$$\|H\|_L = \max_{\rho,\sigma} \{\mathrm{Tr}[H(\rho - \sigma)] \,:\, \|\rho - \sigma\|_{W_1} \leq 1\}\,.$$

Since we work in finite dimensional spaces, duality is straightforward:

$$\|\rho - \sigma\|_{W_1} = \max_{H} \{\mathrm{Tr}[H(\rho - \sigma)] \,:\, \|H\|_L \leq 1\}\,.$$

But we also have an operational definition of H:

$$\|H\|_L = 2 \max_{i=1,\dots,n} \min_{H^{(i)}} \left\| H - \mathbb{I}^{(i)} \otimes H^{(i)} \right\|_\infty ,$$

where $\mathbb{I}^{(i)}$ is the identity on the i-th qubit and $H^{(i)}$ does not act on the i-th qubit, and $\|\cdot\|_\infty$ denotes the operator norm (dual to $\|\cdot\|_1$).

Write

$$\|H\|_L' = 2 \max_{i=1,\ldots,n} \min_{H^{(i)}} \left\| H - \mathbb{I}^{(i)} \otimes H^{(i)} \right\|_\infty .$$

We prove $\|H\|_L \leq \|H\|_L'$.

1. If $\mathrm{Tr}_i[\rho^{(i)}] = \mathrm{Tr}_i[\sigma^{(i)}]$, then

$$\mathrm{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho^{(i)} - \sigma^{(i)})] = \mathrm{Tr}[H^{(i)} \mathrm{Tr}_i[\rho^{(i)} - \sigma^{(i)}]] = 0.$$

2. It follows that if $\rho - \sigma = \sum_{i=1}^n c_i(\rho^{(i)} - \sigma^{(i)})$,

$$\mathrm{Tr}[H(\rho - \sigma)] = \sum_{i=1}^n c_i \mathrm{Tr}[H(\rho^{(i)} - \sigma^{(i)})] = \sum_{i=1}^n c_i \mathrm{Tr}[(H - \mathbb{I}^{(i)} \otimes H^{(i)})(\rho^{(i)} - \sigma^{(i)})]$$

3. Using $\mathrm{Tr}[AB] \leq \|A\|_\infty \|B\|_1$, we bound

$$\mathrm{Tr}[H(\rho - \sigma)] \leq \sum_{i=1}^n c_i \|H\|_L'.$$

4. Assuming $\|\rho - \sigma\|_{W_1} \leq 1$ gives the inequality.

Write

$$\|H\|_L' = 2 \max_{i=1,\ldots,n} \min_{H^{(i)}} \left\| H - \mathbb{I}^{(i)} \otimes H^{(i)} \right\|_\infty .$$

We prove $\|H\|_L \leq \|H\|_L'$.

1. If $\mathrm{Tr}_i[\rho^{(i)}] = \mathrm{Tr}_i[\sigma^{(i)}]$, then

$$\mathrm{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho^{(i)} - \sigma^{(i)})] = \mathrm{Tr}[H^{(i)}\,\mathrm{Tr}_i[\rho^{(i)} - \sigma^{(i)}]] = 0.$$

2. It follows that if $\rho - \sigma = \sum_{i=1}^n c_i(\rho^{(i)} - \sigma^{(i)})$,

$$\mathrm{Tr}[H(\rho-\sigma)] = \sum_{i=1}^n c_i\,\mathrm{Tr}[H(\rho^{(i)}-\sigma^{(i)})] = \sum_{i=1}^n c_i\,\mathrm{Tr}[(H-\mathbb{I}^{(i)}\otimes H^{(i)})(\rho^{(i)}-\sigma^{(i)})]$$

3. Using $\mathrm{Tr}[AB] \leq \|A\|_\infty \|B\|_1$, we bound

$$\mathrm{Tr}[H(\rho-\sigma)] \leq \sum_{i=1}^n c_i \|H\|_L'.$$

4. Assuming $\|\rho - \sigma\|_{W_1} \leq 1$ gives the inequality.

Write

$$\|H\|_L' = 2 \max_{i=1,\ldots,n} \min_{H^{(i)}} \left\| H - \mathbb{I}^{(i)} \otimes H^{(i)} \right\|_\infty .$$

We prove $\|H\|_L \leq \|H\|_L'$.

1. If $\mathrm{Tr}_i[\rho^{(i)}] = \mathrm{Tr}_i[\sigma^{(i)}]$, then

$$\mathrm{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho^{(i)} - \sigma^{(i)})] = \mathrm{Tr}[H^{(i)}\,\mathrm{Tr}_i[\rho^{(i)} - \sigma^{(i)}]] = 0.$$

2. It follows that if $\rho - \sigma = \sum_{i=1}^n c_i(\rho^{(i)} - \sigma^{(i)})$,

$$\mathrm{Tr}[H(\rho-\sigma)] = \sum_{i=1}^n c_i\,\mathrm{Tr}[H(\rho^{(i)}-\sigma^{(i)})] = \sum_{i=1}^n c_i\,\mathrm{Tr}[(H-\mathbb{I}^{(i)}\otimes H^{(i)})(\rho^{(i)}-\sigma^{(i)})]$$

3. Using $\mathrm{Tr}[AB] \leq \|A\|_\infty \|B\|_1$, we bound

$$\mathrm{Tr}[H(\rho - \sigma)] \leq \sum_{i=1}^n c_i \|H\|_L'.$$

4. Assuming $\|\rho - \sigma\|_{W_1} \leq 1$ gives the inequality.

To prove $\|H\|'_L \leq \|H\|_L$ we use that, for self-adjoint A,

$$2 \min_{c \in \mathbb{R}} \|A - c\mathbb{I}\|_\infty = \max_{\rho,\sigma} \operatorname{Tr}[A(\rho - \sigma)].$$

with ρ, σ states: prove it as an exercise.

1. Fix $i = 1, \ldots, n$ and use $\min\max = \max\min$ to obtain

$$\begin{aligned}
\min_{H^{(i)}} \|H - \mathbb{I}^{(i)} \otimes H^{(i)}\|_\infty &= \min_{H^{(i)}} \min_{c \in \mathbb{R}} \|(H - \mathbb{I}^{(i)} \otimes H^{(i)}) - c\mathbb{I}\|_\infty \\
&= \min_{H^{(i)}} \max_{\rho,\sigma} 2 \operatorname{Tr}[(H - \mathbb{I}^{(i)} \otimes H^{(i)})(\rho - \sigma)] \\
&= \max_{\rho,\sigma} \min_{H^{(i)}} 2 \operatorname{Tr}[(H - \mathbb{I}^{(i)} \otimes H^{(i)})(\rho - \sigma)]
\end{aligned}$$

2. Reduce to ρ, σ such that $\operatorname{Tr}_i[\rho] = \operatorname{Tr}_i[\sigma]$, otherwise one can choose $H^{(i)}$ so that

$$\operatorname{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho - \sigma)] = \operatorname{Tr}[H^{(i)} \operatorname{Tr}_i[\rho - \sigma]] \to -\infty$$

3. Therefore $\operatorname{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho - \sigma)] = 0$, and

$$\min_{H^{(i)}} \|H - \mathbb{I}^{(i)} \otimes H^{(i)}\|_\infty \leq \max_{\rho,\sigma} \operatorname{Tr}[H(\rho - \sigma)] \leq \|H\|_L \max_{\rho,\sigma} \|\rho - \sigma\|_{W_1}$$

4. Finally use that $\|\rho - \sigma\|_{W_1} \leq 1$ since $\operatorname{Tr}_i[\rho] = \operatorname{Tr}_i[\sigma]$.

To prove $\|H\|_L' \leq \|H\|_L$ we use that, for self-adjoint A,

$$2 \min_{c\in\mathbb{R}} \|A - c\mathbb{I}\|_\infty = \max_{\rho,\sigma} \operatorname{Tr}[A(\rho-\sigma)].$$

with ρ, σ states: prove it as an exercise.

1. Fix $i = 1, \ldots, n$ and use $\min\max = \max\min$ to obtain

$$\begin{aligned}\min_{H^{(i)}} \|H - \mathbb{I}^{(i)} \otimes H^{(i)}\|_\infty &= \min_{H^{(i)}} \min_{c\in\mathbb{R}} \|(H - \mathbb{I}^{(i)} \otimes H^{(i)}) - c\mathbb{I}\|_\infty \\ &= \min_{H^{(i)}} \max_{\rho,\sigma} 2 \operatorname{Tr}[(H - \mathbb{I}^{(i)} \otimes H^{(i)})(\rho-\sigma)] \\ &= \max_{\rho,\sigma} \min_{H^{(i)}} 2 \operatorname{Tr}[(H - \mathbb{I}^{(i)} \otimes H^{(i)})(\rho-\sigma)]\end{aligned}$$

2. Reduce to ρ, σ such that $\operatorname{Tr}_i[\rho] = \operatorname{Tr}_i[\sigma]$, otherwise one can choose $H^{(i)}$ so that

$$\operatorname{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho-\sigma)] = \operatorname{Tr}[H^{(i)} \operatorname{Tr}_i[\rho-\sigma]] \to -\infty$$

3. Therefore $\operatorname{Tr}[\mathbb{I}^{(i)} \otimes H^{(i)}(\rho-\sigma)] = 0$, and

$$\min_{H^{(i)}} \|H - \mathbb{I}^{(i)} \otimes H^{(i)}\|_\infty \leq \max_{\rho,\sigma} \operatorname{Tr}[H(\rho-\sigma)] \leq \|H\|_L \max_{\rho,\sigma} \|\rho-\sigma\|_{W_1}$$

4. Finally use that $\|\rho-\sigma\|_{W_1} \leq 1$ since $\operatorname{Tr}_i[\rho] = \operatorname{Tr}_i[\sigma]$.

Remarks

- It always holds $\|H\|_L \le 2 \min_{c\in\mathbb{R}} \|H - c\mathbb{I}\|_\infty \le 2\|H\|_\infty$.
- If $H = \sum_{\mathcal{I}\subseteq[n]} H_{\mathcal{I}}$ is a sum of local operators, i.e., $H_{\mathcal{I}}$ acts only on qubits in the subset $\mathcal{I}$, then

$$\|H\|_L \le 2 \max_i \left\| \sum_{\mathcal{I}\subseteq[n]: i\in\mathcal{I}} H_{\mathcal{I}} \right\|_\infty ,$$

simply by taking for $i = 1, \ldots, n$

$$\mathbb{I}^{(i)} \otimes H^{(i)} = \sum_{i\notin\mathcal{I}} H_{\mathcal{I}}.$$

Exercises:

1. If $H = \sum_x f(x)|x\rangle\langle x|$ is diagonal, then $\|H\|_L = \|f\|_L$.
2. Prove that $\|H - \mathbb{I}\,\mathrm{Tr}[H]/2^n\|_\infty \le n\|H\|_L$.

Remarks

- It always holds $\|H\|_L \leq 2\min_{c\in\mathbb{R}}\|H - c\mathbb{I}\|_\infty \leq 2\|H\|_\infty$.
- If $H = \sum_{\mathcal{I}\subseteq[n]} H_\mathcal{I}$ is a sum of local operators, i.e., $H_\mathcal{I}$ acts only on qubits in the subset $\mathcal{I}$, then

$$\|H\|_L \leq 2\max_i \left\| \sum_{\mathcal{I}\subseteq[n]:i\in\mathcal{I}} H_\mathcal{I} \right\|_\infty ,$$

simply by taking for $i = 1, \ldots, n$

$$\mathbb{I}^{(i)} \otimes H^{(i)} = \sum_{i\notin\mathcal{I}} H_\mathcal{I}.$$

Exercises:

1. If $H = \sum_x f(x)|x\rangle\langle x|$ is diagonal, then $\|H\|_L = \|f\|_L$.
2. Prove that $\|H - \mathbb{I}\,\mathrm{Tr}[H]/2^n\|_\infty \leq n\|H\|_L$.

Plan

Continuity of entropy

Reminder: Shannon's entropy $S(p) = -\sum_x p(x) \ln p(x)$ and Von Neumann entropy $S(\rho) = -\operatorname{Tr}[\rho \ln \rho]$.
Continuity of the entropy has applications in information theory:

$$\text{if } q \approx p, \text{ is } S(q) \approx S(p)?$$

On the cube $\{0,1\}^n$, Polyanski and Wu proved

$$|S(p) - S(q)| \leq n h_2\left(\frac{W_1(p,q)}{n}\right).$$

with $h_2(x) = -(1-x)\ln(1-x) - x \ln x$ the (binary) entropy of a Bernoulli.

Example: if X is uniform and Y is close to X, i.e., $W_1(X,Y) \ll n$, then

$$S(Y) \approx n.$$

Continuity of entropy

Reminder: Shannon's entropy $S(p) = -\sum_x p(x) \ln p(x)$ and Von Neumann entropy $S(\rho) = -\operatorname{Tr}[\rho \ln \rho]$.
Continuity of the entropy has applications in information theory:

if $q \approx p$, is $S(q) \approx S(p)$?

On the cube $\{0,1\}^n$, Polyanski and Wu proved

$$|S(p) - S(q)| \leq n h_2\left(\frac{W_1(p,q)}{n}\right).$$

with $h_2(x) = -(1-x)\ln(1-x) - x\ln x$ the (binary) entropy of a Bernoulli.

Example: if X is uniform and Y is close to X, i.e., $W_1(X,Y) \ll n$, then

$$S(Y) \approx n.$$

Continuity of entropy

Reminder: Shannon's entropy $S(p) = -\sum_x p(x) \ln p(x)$ and Von Neumann entropy $S(\rho) = -\operatorname{Tr}[\rho \ln \rho]$.
Continuity of the entropy has applications in information theory:

if $q \approx p$, is $S(q) \approx S(p)$?

On the cube $\{0,1\}^n$, Polyanski and Wu proved

$$|S(p) - S(q)| \leq n h_2 \left(\frac{W_1(p,q)}{n} \right).$$

with $h_2(x) = -(1-x)\ln(1-x) - x \ln x$ the (binary) entropy of a Bernoulli.

Example: if X is uniform and Y is close to X, i.e., $W_1(X,Y) \ll n$, then

$$S(Y) \approx n.$$

Quantum case

For states $\rho, \sigma \in \mathcal{S}(\mathbb{C}^d)$, Fannes-Audenaert proved

$$|S(\rho) - S(\sigma)| \leq h_2\left(\frac{1}{2}\|\rho - \sigma\|_1\right) + \frac{1}{2}\|\rho - \sigma\|_1 \ln(d-1).$$

Consider $d = 2^n$. Then

1. it is exactly Polyanskiy-Wu for $n = 1$,
2. for $n \gg 1$, useful only if $\|\rho - \sigma\|_1/2 \ll 1$.

Using the quantum Wasserstein we prove: for $\rho, \sigma \in \mathcal{S}((\mathbb{C}^2)^{\otimes n})$:

$$|S(\rho) - S(\sigma)| \leq nh_2\left(\frac{\|\rho - \sigma\|_{W_1}}{n}\right) + \|\rho - \sigma\|_{W_1} \ln(3).$$

Quantum case

For states $\rho, \sigma \in \mathcal{S}(\mathbb{C}^d)$, Fannes-Audenaert proved

$$|S(\rho) - S(\sigma)| \leq h_2\left(\frac{1}{2}\|\rho - \sigma\|_1\right) + \frac{1}{2}\|\rho - \sigma\|_1 \ln(d-1).$$

Consider $d = 2^n$. Then

1. it is exactly Polyanskiy-Wu for $n = 1$,
2. for $n \gg 1$, useful only if $\|\rho - \sigma\|_1/2 \ll 1$.

Using the quantum Wasserstein we prove: for $\rho, \sigma \in \mathcal{S}((\mathbb{C}^2)^{\otimes n})$:

$$|S(\rho) - S(\sigma)| \leq nh_2\left(\frac{\|\rho - \sigma\|_{W_1}}{n}\right) + \|\rho - \sigma\|_{W_1} \ln(3).$$

Quantum case

For states $\rho, \sigma \in \mathcal{S}(\mathbb{C}^d)$, Fannes-Audenaert proved

$$|S(\rho) - S(\sigma)| \leq h_2\left(\frac{1}{2}\|\rho - \sigma\|_1\right) + \frac{1}{2}\|\rho - \sigma\|_1 \ln(d-1).$$

Consider $d = 2^n$. Then

1. it is exactly Polyanskiy-Wu for $n = 1$,
2. for $n \gg 1$, useful only if $\|\rho - \sigma\|_1/2 \ll 1$.

Using the quantum Wasserstein we prove: for $\rho, \sigma \in \mathcal{S}((\mathbb{C}^2)^{\otimes n})$:

$$|S(\rho) - S(\sigma)| \leq n h_2\left(\frac{\|\rho - \sigma\|_{W_1}}{n}\right) + \|\rho - \sigma\|_{W_1} \ln(3).$$

Quantum case

For states $\rho, \sigma \in \mathcal{S}(\mathbb{C}^d)$, Fannes-Audenaert proved

$$|S(\rho) - S(\sigma)| \leq h_2\left(\frac{1}{2}\|\rho - \sigma\|_1\right) + \frac{1}{2}\|\rho - \sigma\|_1 \ln(d-1).$$

Consider $d = 2^n$. Then

1. it is exactly Polyanskiy-Wu for $n = 1$,
2. for $n \gg 1$, useful only if $\|\rho - \sigma\|_1/2 \ll 1$.

Using the quantum Wasserstein we prove: for $\rho, \sigma \in \mathcal{S}((\mathbb{C}^2)^{\otimes n})$:

$$|S(\rho) - S(\sigma)| \leq n h_2\left(\frac{\|\rho - \sigma\|_{W_1}}{n}\right) + \|\rho - \sigma\|_{W_1} \ln(3).$$

Plan

Since the n-qubits system has W_1 diameter $\leq n$, all the eigenvalues of an observable H belong to the interval

$$[\mathrm{Tr}[H]/2^n - n\|H\|_L, \mathrm{Tr}[H]/2^n + n\|H\|_L].$$

In fact most eigenvalues belong to much smaller intervals of length $\approx \sqrt{n}\|H\|_L$.

We aim to establish the following concentration inequality: for every $\delta > 0$,

$$\dim\left(H \geq \left(\mathrm{Tr}[H]/2^n + \delta\sqrt{n}\,\|H\|_L/2\right)\mathbb{I}\right) \leq 2^n \exp(-\delta^2/2).$$

Since $\dim((\mathbb{C}^2)^{\otimes n}) = 2^n$, it yields that the relative distribution of eigenvalues is (roughly) concentrated as a Gaussian with mean $\mathrm{Tr}[H]/2^n$ and standard deviation $\sqrt{n}\|H\|_L/2$.

Since the n-qubits system has W_1 diameter $\leq n$, all the eigenvalues of an observable H belong to the interval

$$[\mathrm{Tr}[H]/2^n - n\|H\|_L, \mathrm{Tr}[H]/2^n + n\|H\|_L].$$

In fact most eigenvalues belong to much smaller intervals of length $\approx \sqrt{n}\|H\|_L$.

We aim to establish the following concentration inequality: for every $\delta > 0$,

$$\dim\left(H \geq \left(\mathrm{Tr}[H]/2^n + \delta\sqrt{n}\,\|H\|_L/2\right)\mathbb{I}\right) \leq 2^n \exp(-\delta^2/2).$$

Since $\dim((\mathbb{C}^2)^{\otimes n}) = 2^n$, it yields that the relative distribution of eigenvalues is (roughly) concentrated as a Gaussian with mean $\mathrm{Tr}[H]/2^n$ and standard deviation $\sqrt{n}\|H\|_L/2$.

Since the n-qubits system has W_1 diameter $\leq n$, all the eigenvalues of an observable H belong to the interval

$$[\mathrm{Tr}[H]/2^n - n\|H\|_L, \mathrm{Tr}[H]/2^n + n\|H\|_L].$$

In fact most eigenvalues belong to much smaller intervals of length $\approx \sqrt{n}\|H\|_L$.

We aim to establish the following concentration inequality: for every $\delta > 0$,

$$\dim\left(H \geq \left(\mathrm{Tr}[H]/2^n + \delta\sqrt{n}\,\|H\|_L\,/2\right)\mathbb{I}\right) \leq 2^n \exp(-\delta^2/2).$$

Since $\dim((\mathbb{C}^2)^{\otimes n}) = 2^n$, it yields that the relative distribution of eigenvalues is (roughly) concentrated as a Gaussian with mean $\mathrm{Tr}[H]/2^n$ and standard deviation $\sqrt{n}\|H\|_L/2$.

Since the n-qubits system has W_1 diameter $\leq n$, all the eigenvalues of an observable H belong to the interval

$$[\mathrm{Tr}[H]/2^n - n\|H\|_L, \mathrm{Tr}[H]/2^n + n\|H\|_L].$$

In fact most eigenvalues belong to much smaller intervals of length $\approx \sqrt{n}\|H\|_L$.

We aim to establish the following concentration inequality: for every $\delta > 0$,

$$\dim\left(H \geq \left(\mathrm{Tr}[H]/2^n + \delta\sqrt{n}\,\|H\|_L\,/2\right)\mathbb{I}\right) \leq 2^n \exp(-\delta^2/2).$$

Since $\dim((\mathbb{C}^2)^{\otimes n}) = 2^n$, it yields that the relative distribution of eigenvalues is (roughly) concentrated as a Gaussian with mean $\mathrm{Tr}[H]/2^n$ and standard deviation $\sqrt{n}\|H\|_L/2$.

Transport-entropy inequalities

A classical tool to establish concentration are inequalities of the form

$$W_1(p, q) \leq \sqrt{\frac{n}{2} D(p \| q)}\,. \tag{1}$$

for probabilities p, q, and

$$D(p\|q) = \sum_x p(x) \ln(p(x)/q(x))$$

is the relative entropy (or Kullback-Leibler divergence).

- If (1) holds for a given q and for every p, then duality yields a form of Gaussian concentration.
- Relevant cases (due to K. Marton) are: q product state or q Markov state (i.e., the law of a Markov chain) under some assumptions.
- We establish the quantum analogue of (1) for product states on qubits. For non-product states there are partial results.

Transport-entropy inequalities

A classical tool to establish concentration are inequalities of the form

$$W_1(p, q) \leq \sqrt{\frac{n}{2} D(p \| q)}\,. \tag{1}$$

for probabilities p, q, and

$$D(p\|q) = \sum_x p(x) \ln(p(x)/q(x))$$

is the relative entropy (or Kullback-Leibler divergence).

- If (1) holds for a given q and for every p, then duality yields a form of Gaussian concentration.
- Relevant cases (due to K. Marton) are: q product state or q Markov state (i.e., the law of a Markov chain) under some assumptions.
- We establish the quantum analogue of (1) for product states on qubits. For non-product states there are partial results.

Quantum Marton's inequality

For any $\rho, \sigma \in \mathcal{S}((\mathbb{C}^2)^{\otimes n})$, with

$$\sigma = \sigma_1 \otimes \ldots \otimes \sigma_n$$

product state, the following inequality holds:

$$\|\rho - \sigma\|_{W_1} \leq \sqrt{\frac{n}{2} S(\rho\|\sigma)},$$

where the relative entropy is defined as

$$S(\rho\|\sigma) = \mathrm{Tr}[\rho\,(\ln(\rho) - \ln(\sigma))].$$

Quantum Pinsker's inequality

If $n = 1$, then any σ is product and W_1 is the trace distance, hence the inequality becomes

$$\|\rho - \sigma\|_1 \leq \sqrt{2\, S(\rho\|\sigma)}.$$

This is known as Pinsker's inequality in the quantum case (of a single qubit).

Sketch of proof:

1. Establish the classical Pinsker's inequality for classical probabilities r, s on a two point set, e.g. $\{-,+\}$.
2. Introduce the orthogonal projectors $\Pi_+ = 1_{\{\rho-\sigma\geq 0\}}$, $\Pi_- = 1_{\{\rho-\sigma<0\}}$ and probabilities on $\{-,+\}$

   $$r_\pm = \mathrm{Tr}[\Pi_\pm \rho], \quad s_\pm = \mathrm{Tr}[\Pi_\pm \sigma].$$

   So that

   $$\|\rho - \sigma\|_1 = |r_+ - s_+| + |r_- - s_-| = \|r - s\|_1 \leq \sqrt{2S(r\|s)}.$$

3. The pair $\{\Pi_+, \Pi_-\}$ gives a measurement (Helstrom). Conclude by monotonicity of the relative entropy:

   $$S(r\|s) \leq S(\rho\|\sigma).$$

Quantum Pinsker's inequality

If $n = 1$, then any σ is product and W_1 is the trace distance, hence the inequality becomes

$$\|\rho - \sigma\|_1 \leq \sqrt{2\, S(\rho\|\sigma)}.$$

This is known as Pinsker's inequality in the quantum case (of a single qubit).

Sketch of proof:

1. Establish the classical Pinsker's inequality for classical probabilities r, s on a two point set, e.g. $\{-, +\}$.
2. Introduce the orthogonal projectors $\Pi_+ = 1_{\{\rho-\sigma\geq 0\}}$, $\Pi_- = 1_{\{\rho-\sigma<0\}}$ and probabilities on $\{-, +\}$

   $$r_\pm = \mathrm{Tr}[\Pi_\pm \rho], \quad s_\pm = \mathrm{Tr}[\Pi_\pm \sigma].$$

   So that

   $$\|\rho - \sigma\|_1 = |r_+ - s_+| + |r_- - s_-| = \|r - s\|_1 \leq \sqrt{2S(r\|s)}.$$

3. The pair $\{\Pi_+, \Pi_-\}$ gives a measurement (Helstrom). Conclude by monotonicity of the relative entropy:

   $$S(r\|s) \leq S(\rho\|\sigma).$$

Quantum Pinsker's inequality

If $n = 1$, then any σ is product and W_1 is the trace distance, hence the inequality becomes

$$\|\rho - \sigma\|_1 \leq \sqrt{2\, S(\rho\|\sigma)}.$$

This is known as Pinsker's inequality in the quantum case (of a single qubit).

Sketch of proof:

1. Establish the classical Pinsker's inequality for classical probabilities r, s on a two point set, e.g. $\{-, +\}$.
2. Introduce the orthogonal projectors $\Pi_+ = 1_{\{\rho-\sigma\geq 0\}}$, $\Pi_- = 1_{\{\rho-\sigma<0\}}$ and probabilities on $\{-, +\}$

   $$r_\pm = \mathrm{Tr}[\Pi_\pm \rho], \quad s_\pm = \mathrm{Tr}[\Pi_\pm \sigma].$$

   So that

   $$\|\rho - \sigma\|_1 = |r_+ - s_+| + |r_- - s_-| = \|r - s\|_1 \leq \sqrt{2S(r\|s)}.$$
3. The pair $\{\Pi_+, \Pi_-\}$ gives a measurement (Helstrom). Conclude by monotonicity of the relative entropy:

   $$S(r\|s) \leq S(\rho\|\sigma).$$

Proof of quantum Marton's inequality

1. Write
$$\rho - \sigma = \sum_{i=1}^{n} \left(\rho_{1\ldots i} \otimes \sigma_{i+1\ldots n} - \rho_{1\ldots i-1} \otimes \sigma_{i\ldots n} \right)$$
2. Apply Pinsker's inequality for every $i = 1, \ldots, n$:
$$\left\| \rho_{1\ldots i} \otimes \sigma_{i+1\ldots n} - \rho_{1\ldots i-1} \otimes \sigma_{i\ldots n} \right\|_1 \leq \sqrt{2\, S\left(\rho_{1\ldots i} \otimes \sigma_{i+1\ldots n} \| \rho_{1\ldots i-1} \otimes \sigma_{i\ldots n} \right)}$$
3. Summing upon i and using concavity of $\sqrt{\cdot}$,
$$\|\rho - \sigma\|_{W_1} \leq \sqrt{\frac{n}{2} \sum_{i=1}^{n} S\left(\rho_{1\ldots i} \otimes \sigma_{i+1\ldots n} \| \rho_{1\ldots i-1} \otimes \sigma_{i\ldots n} \right)}$$
4. Conclude by the identity
$$\begin{aligned} S\left(\rho_{1\ldots i} \otimes \sigma_{i+1\ldots n} \| \rho_{1\ldots i-1} \otimes \sigma_{i\ldots n} \right) &= S\left(\rho_{1\ldots i} \| \rho_{1\ldots i-1} \otimes \sigma_i \right) \\ &= -S(\rho_{1\ldots i}) + S(\rho_{1\ldots i-1}) - \mathrm{Tr}\left[\rho_i \ln \sigma_i \right] \end{aligned}$$

and telescopic summation.

Proof of quantum Marton's inequality

1. Write
$$\rho - \sigma = \sum_{i=1}^{n} \left(\rho_{1\dots i} \otimes \sigma_{i+1\dots n} - \rho_{1\dots i-1} \otimes \sigma_{i\dots n}\right)$$
2. Apply Pinsker's inequality for every $i = 1, \dots, n$:
$$\left\|\rho_{1\dots i} \otimes \sigma_{i+1\dots n} - \rho_{1\dots i-1} \otimes \sigma_{i\dots n}\right\|_1 \le \sqrt{2\, S\left(\rho_{1\dots i} \otimes \sigma_{i+1\dots n} \| \rho_{1\dots i-1} \otimes \sigma_{i\dots n}\right)}$$
3. Summing upon i and using concavity of $\sqrt{\cdot}$,
$$\|\rho - \sigma\|_{W_1} \le \sqrt{\frac{n}{2} \sum_{i=1}^{n} S\left(\rho_{1\dots i} \otimes \sigma_{i+1\dots n} \| \rho_{1\dots i-1} \otimes \sigma_{i\dots n}\right)}$$
4. Conclude by the identity
$$\begin{aligned} S\left(\rho_{1\dots i} \otimes \sigma_{i+1\dots n} \| \rho_{1\dots i-1} \otimes \sigma_{i\dots n}\right) &= S\left(\rho_{1\dots i} \| \rho_{1\dots i-1} \otimes \sigma_i\right) \\ &= -S(\rho_{1\dots i}) + S(\rho_{1\dots i-1}) - \mathrm{Tr}\left[\rho_i \ln \sigma_i\right] \end{aligned}$$

   and telescopic summation.

Proof of Gaussian concentration inequality

1. Assume without loss of generality $\mathrm{Tr}[H] = 0$ and $\|H\|_L \leq 1$. We prove,

$$\mathrm{Tr}[e^{tH}] \leq 2^n \exp(nt^2/8), \quad \text{for every } t > 0,$$

so that concentration follows by Markov inequality.

2. Choose $\sigma = \mathbb{I}/2^n$ (maximally mixed state), which is a product state. Duality and Marton's inequality give

$$t\,\mathrm{Tr}[H\rho] = t\,\mathrm{Tr}[H(\rho - \sigma)] \leq t\|\rho - \sigma\|_{W_1} \leq t\sqrt{\frac{n}{2}S(\rho\|\sigma)} \leq \frac{nt^2}{8} + S(\rho\|\sigma).$$

3. Set $\rho = e^{tH}/\,\mathrm{Tr}[e^{tH}]$ so that $\ln\rho = tH - \ln\mathrm{Tr}[e^{tH}]$ and

$$S(\rho\|\sigma) = \mathrm{Tr}[\rho\ln\rho] + \ln 2^n = t\,\mathrm{Tr}[H\rho] - \ln\mathrm{Tr}[e^{tH}] + \ln 2^n.$$

4. Conclude that

$$\ln\mathrm{Tr}[e^{tH}] \leq \frac{nt^2}{8} + \ln 2^n.$$

Proof of Gaussian concentration inequality

1. Assume without loss of generality $\mathrm{Tr}[H] = 0$ and $\|H\|_L \leq 1$. We prove,

$$\mathrm{Tr}[e^{tH}] \leq 2^n \exp(nt^2/8), \quad \text{for every } t > 0,$$

   so that concentration follows by Markov inequality.

2. Choose $\sigma = \mathbb{I}/2^n$ (maximally mixed state), which is a product state. Duality and Marton's inequality give

$$t\,\mathrm{Tr}[H\rho] = t\,\mathrm{Tr}[H(\rho - \sigma)] \leq t\|\rho - \sigma\|_{W_1} \leq t\sqrt{\frac{n}{2} S(\rho\|\sigma)} \leq \frac{nt^2}{8} + S(\rho\|\sigma).$$

3. Set $\rho = e^{tH} / \mathrm{Tr}[e^{tH}]$ so that $\ln \rho = tH - \ln \mathrm{Tr}[e^{tH}]$ and

$$S(\rho\|\sigma) = \mathrm{Tr}[\rho \ln \rho] + \ln 2^n = t\,\mathrm{Tr}[H\rho] - \ln \mathrm{Tr}[e^{tH}] + \ln 2^n.$$

4. Conclude that

$$\ln \mathrm{Tr}[e^{tH}] \leq \frac{nt^2}{8} + \ln 2^n.$$

Further applications/problems

- Quantum generative models: qWGAN (Kiani et al.)
- Quantum matching problem (Rouzé/França)
- Quantum Ricci curvature (Gao/Rouzé)

Problems:

1. Efficient computation/approximation
2. Geometric inequalities on quantum Hamming cube
3. Mixing times for semigroups

Further applications/problems

- Quantum generative models: qWGAN (Kiani et al.)
- Quantum matching problem (Rouzé/França)
- Quantum Ricci curvature (Gao/Rouzé)

Problems:

1. Efficient computation/approximation
2. Geometric inequalities on quantum Hamming cube
3. Mixing times for semigroups

Plan

Nielsen, M. A., and Chuang, I.
Quantum computation and quantum information
Cambridge University Press, 2010.

Polyanskiy, Y. and Wu, Y.
Wasserstein continuity of entropy and outer bounds for interference channels
IEEE Transactions on Information Theory, 2016.

Marton, K.
Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration
The Annals of Probability, 1996.

Gozlan, N., and Léonard, C.
Transport inequalities. A survey
Markov Processes and Related Fields,, 2010.

De Palma, G., and Rouzé, C.
Quantum concentration inequalities
Annales Henri Poincaré, 2022.

De Palma, G., and Trevisan, D.
The Wasserstein distance of order 1 for quantum spin systems on infinite lattices
Annales Henri Poincaré. Vol. 24. No. 12.