Quantum OT: quantum channels and qubits
Tutorial 1

Dario Trevisan

Università di Pisa
dario.trevisan@unipi.it

March 12, 2025

Plan

Plan

Classical vs Quantum: a dictionary

Classical	Quantum
E (discrete)	H Hilbert space
$e \in E$	$\|\psi\rangle \in H$
$A \subseteq E$	$V < H$ (closed subspace)
$f : E \to \mathbb{C}$	$A : D(A) \subseteq H \to H$ linear operator
bounded	bounded operator
real-valued	self-adjoint
non-negative (psd)	non-negative
$\|f\|^2$	$\|A\|^2 = A^\dagger A$
$\sum_{x \in E} f(x)$	$\mathrm{Tr}[A]$
$\ell^p(E) = \{ f : \sum_{x \in E} \|f(x)\|^p < \infty \}$	$\ell^p(H) = \{ A : \mathrm{Tr}[(A^\dagger A)^{p/2}] < \infty \}$

Classical vs Quantum: a dictionary

Classical	Quantum
E (discrete)	H Hilbert space
$e \in E$	$\vert\psi\rangle \in H$
$A \subseteq E$	$V < H$ (closed subspace)
$f : E \to \mathbb{C}$	$A : D(A) \subseteq H \to H$ linear operator
bounded	bounded operator
real-valued	self-adjoint
non-negative (psd)	non-negative
$\vert f\vert^2$	$\vert A\vert^2 = A^\dagger A$
$\sum_{x \in E} f(x)$	$\mathrm{Tr}[A]$
$\ell^p(E) = \{f : \sum_{x\in E} \vert f(x)\vert^p < \infty\}$	$\ell^p(H) = \{A : \mathrm{Tr}[(A^\dagger A)^{p/2}] < \infty\}$

Classical vs Quantum: a dictionary

probability densities p	quantum states $\rho \in \mathcal{S}(H)$
δ_x	$\lvert\psi\rangle\langle\psi\rvert$ (pure state)
Markov operator (transition kernel)	CPTP operator
Product $E \times F$	Product space $H \otimes K$
Partial sum $\sum_x f(x, y)$	Partial trace $\mathrm{Tr}_H[A]$
$\sum_{(x,y)} f(x, y) = \sum_x \sum_y f(x, y)$	$\mathrm{Tr}[A] = \mathrm{Tr}_K[\mathrm{Tr}_H[A]]$
Marginal $p_E(x) = \sum_y p(x, y)$	Marginal $\rho_H = \mathrm{Tr}_K[\rho]$
$S(p) = -\sum_x p(x) \ln p(x)$	$S(\rho) = -\mathrm{Tr}[\rho \ln \rho]$
$D(p\|\|q) = \sum_x p(x) \ln(p(x)/q(x))$	$D(\rho\|\|\sigma) = \mathrm{Tr}[\rho(\ln \rho - \ln \sigma)]$

Classical vs Quantum: a dictionary

Classical	Quantum
probability densities p	quantum states $\rho \in \mathcal{S}(H)$
δ_x	$\|\psi\rangle\langle\psi\|$ (pure state)
Markov operator (transition kernel)	CPTP operator
Product $E \times F$	Product space $H \otimes K$
Partial sum $\sum_x f(x,y)$	Partial trace $\mathrm{Tr}_H[A]$
$\sum_{(x,y)} f(x,y) = \sum_x \sum_y f(x,y)$	$\mathrm{Tr}[A] = \mathrm{Tr}_K[\mathrm{Tr}_H[A]]$
Marginal $p_E(x) = \sum_y p(x,y)$	Marginal $\rho_H = \mathrm{Tr}_K[\rho]$
$S(p) = -\sum_x p(x) \ln p(x)$	$S(\rho) = -\mathrm{Tr}[\rho \ln \rho]$
$D(p\|\|q) = \sum_x p(x) \ln(p(x)/q(x))$	$D(\rho\|\|\sigma) = \mathrm{Tr}[\rho(\ln \rho - \ln \sigma)]$

States and density matrices

Given a state π on a (finite dimensional) space H, pick basis $(|\psi_i\rangle)_i$ and write its density matrix $(\pi_{i,j})_{ij}$:

$$\pi = \sum_{i,j} \pi_{i,j} |\psi_i\rangle\langle\psi_j|.$$

Then:

- $(\pi_{i,j})_{ij}$ is Hermitian positive semidefinite
- The diagonal $(\pi_{i,i})_i$ is a classical probability density
- By the spectral theorem one can always diagonalize

$$\pi = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|$$

but the basis $(|\varphi_i\rangle)_i$ depends on π.

Partial trace and density matrices

Given a state π on the product space $H \otimes K$, pick basis $(|\psi_i\rangle)_i$, $(|\varphi_j\rangle)_j$, and write its density matrix $(\pi_{ij,k\ell})_{ijk\ell}$:

$$\pi = \sum_{i,j,k,\ell} \pi_{ij,k\ell} |\psi_i\rangle \otimes |\phi_j\rangle\langle\psi_k| \otimes \langle\phi_\ell|.$$

Then

$$\mathrm{Tr}_H[\pi] = \sum_{j,\ell} \left(\sum_i \pi_{ij,i\ell} \right) |\phi_j\rangle\langle\phi_\ell|,$$

$$\mathrm{Tr}_K[\pi] = \sum_{i,k} \left(\sum_j \pi_{ij,kj} \right) |\psi_i\rangle\langle\psi_\ell|,$$

Well-defied (do not depend on the basis).

Partial trace and density matrices

Given a state π on the product space $H \otimes K$, pick basis $(|\psi_i\rangle)_i$, $(|\varphi_j\rangle)_j$, and write its density matrix $(\pi_{ij,k\ell})_{ijk\ell}$:

$$\pi = \sum_{i,j,k,\ell} \pi_{ij,k\ell} |\psi_i\rangle \otimes |\phi_j\rangle\langle\psi_k| \otimes \langle\phi_\ell|.$$

Then

$$\mathrm{Tr}_H[\pi] = \sum_{j,\ell} \left(\sum_i \pi_{ij,i\ell} \right) |\phi_j\rangle\langle\phi_\ell|,$$

$$\mathrm{Tr}_K[\pi] = \sum_{i,k} \left(\sum_j \pi_{ij,kj} \right) |\psi_i\rangle\langle\psi_\ell|,$$

Well-defied (do not depend on the basis).

Partial trace and density matrices

Given a state π on the product space $H \otimes K$, pick basis $(|\psi_i\rangle)_i$, $(|\varphi_j\rangle)_j$, and write its density matrix $(\pi_{ij,k\ell})_{ijk\ell}$:

$$\pi = \sum_{i,j,k,\ell} \pi_{ij,k\ell} |\psi_i\rangle \otimes |\phi_j\rangle\langle\psi_k| \otimes \langle\phi_\ell|.$$

Then

$$\mathrm{Tr}_H[\pi] = \sum_{j,\ell} \left(\sum_i \pi_{ij,i\ell} \right) |\phi_j\rangle\langle\phi_\ell|,$$

$$\mathrm{Tr}_K[\pi] = \sum_{i,k} \left(\sum_j \pi_{ij,kj} \right) |\psi_i\rangle\langle\psi_\ell|,$$

Well-defied (do not depend on the basis).

Exercises

On Hilbert spaces H, K, $p, q \in [0,1]$, orthonormal $(|\psi_i\rangle)_i \in H$, $(|\varphi_i)_i\rangle \in K$,

1. Define $\rho_p = (1-p)|\psi_0\rangle\langle\psi_0| + p|\psi_1\rangle\langle\psi_1|$. Show that $\rho_p \in \mathcal{S}(H)$ and compute $S(\rho_p)$.
2. Define $|\psi_p\rangle = \sqrt{(1-p)}|\psi_0\rangle + \sqrt{p}|\psi_1\rangle$ and compute $S(|\psi_p\rangle\langle\psi_p|)$.
3. Consider the (Bell) state

$$|\Phi^+\rangle := \frac{1}{\sqrt{2}}|\psi_0\rangle \otimes |\varphi_0\rangle + \frac{1}{\sqrt{2}}|\psi_1\rangle \otimes |\varphi_1\rangle$$

Compute

$$\mathrm{Tr}_H[|\Phi^+\rangle\langle\Phi^+|], \quad \mathrm{Tr}_K[|\Phi^+\rangle\langle\Phi^+|].$$

Qubits systems

A quantum analogue of $\{0,1\}^n$. Let

$$H = (\mathbb{C}^2)^{\otimes n}.$$

The standard (computational) basis

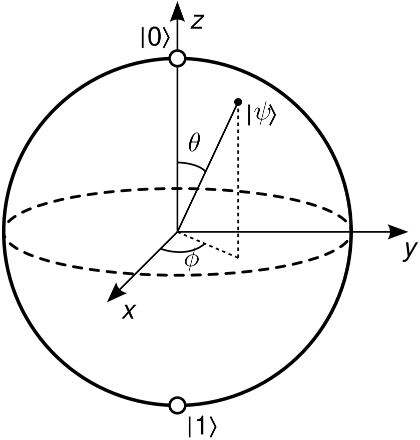
$$\{|s\rangle\}$$

with $s \in \{0,1\}^n$, e.g. for $n = 1$

$$\{|0\rangle, |1\rangle\},$$

for $n = 2$,

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}.$$



The Bloch sphere representation of a (pure) state $|\psi\rangle \in \mathbb{C}^2$.

Quantum states on systems of qubits

On n-qubits σ, $\rho \in \mathcal{S}(\mathbb{C}^2)^{\otimes n}$,

- States are $2^n \times 2^n$ complex matrices (Hermitian, positive semi-definite, with unit trace).
- Pure states are rank-one matrices, corresponding to unit norm vectors

$$\sigma = |\psi\rangle\langle\psi|$$

  but not necessarily $\psi \in \{0, 1\}^n$.
- Classical probabilities $(p(s))_{s\in\{0,1\}^n}$ correspond to diagonal states:

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

Quantum computing

(Most) quantum computing architectures are based on systems of n qubits (ideally $n \gg 1$) with sequence of unitary operations (gates) acting on a small subset of them (e.g., up to 3 at the time). Examples:

- Singe qubit gates:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

- Two-qubits gates:

$$CX = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad SWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Three-qubits gates: Toffoli gate (CCX)

Exercise: Show that the above are unitary operators.

Quantum computing

(Most) quantum computing architectures are based on systems of n qubits (ideally $n \gg 1$) with sequence of unitary operations (gates) acting on a small subset of them (e.g., up to 3 at the time). Examples:

- Singe qubit gates:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

- Two-qubits gates:

$$CX = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad SWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Three-qubits gates: Toffoli gate (CCX)

Exercise: Show that the above are unitary operators.

Postulates of Quantum Mechanics

1. A wave function $|\psi_0\rangle \in H$ in a closed (isolated) quantum system evolves according to Schrödinger's equation

$$\partial_t|\psi_t\rangle = iA|\psi_t\rangle$$

for a self-adjoint A. Integration gives

$$|\psi_t\rangle = U_t|\psi_0\rangle, \quad \text{with } U_t = e^{itA} \text{ unitary.}$$

2. A measurement on a system is described by an orthonormal basis $(|\phi_v\rangle)_v \subseteq H$. If the state is $|\psi\rangle \in H$, the outcome has value v with probability (Born's rule)

$$p(v) = |\langle\phi_v|\psi\rangle|^2.$$

After the measurement, the system is in state $|\phi_v\rangle$ with probability $p(v)$, i.e., the state is mixed:

$$\sum_v p(v)|\phi_v\rangle\langle\phi_v|.$$

Postulates of Quantum Mechanics

1. A wave function $|\psi_0\rangle \in H$ in a closed (isolated) quantum system evolves according to Schr̈odinger's equation

$$\partial_t |\psi_t\rangle = iA|\psi_t\rangle$$

for a self-adjoint A. Integration gives

$$|\psi_t\rangle = U_t|\psi_0\rangle, \quad \text{with } U_t = e^{itA} \text{ unitary.}$$

2. A measurement on a system is described by an orthonormal basis $(|\phi_v\rangle)_v \subseteq H$. If the state is $|\psi\rangle \in H$, the outcome has value v with probability (Born's rule)

$$p(v) = |\langle\phi_v|\psi\rangle|^2.$$

After the measurement, the system is in state $|\phi_v\rangle$ with probability $p(v)$, i.e., the state is mixed:

$$\sum_v p(v)|\phi_v\rangle\langle\phi_v|.$$

When applied directly to mixed states $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$, we obtain the two transformations:

1. For the unitary evolution of closed systems:

$$\rho \mapsto U_t \rho U_t^\dagger,$$

2. For the measurement $(|\phi_v\rangle)_v \subseteq H$:

$$\rho \mapsto \sum_v |\phi_v\rangle \left(\langle\phi_v|\rho|\phi_v\rangle\right) \langle\phi_v|.$$

Can one interpolate between these two, describing open quantum systems which interact with an external environment?

Notice that both transformations are linear and map states into states.

When applied directly to mixed states $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$, we obtain the two transformations:

1. For the unitary evolution of closed systems:

$$\rho \mapsto U_t \rho U_t^\dagger,$$

2. For the measurement $(|\phi_v\rangle)_v \subseteq H$:

$$\rho \mapsto \sum_v |\phi_v\rangle \left(\langle\phi_v|\rho|\phi_v\rangle\right) \langle\phi_v|.$$

Can one interpolate between these two, describing open quantum systems which interact with an external environment?

Notice that both transformations are linear and map states into states.

When applied directly to mixed states $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$, we obtain the two transformations:

1. For the unitary evolution of closed systems:

$$\rho \mapsto U_t \rho U_t^\dagger,$$

2. For the measurement $(|\phi_v\rangle)_v \subseteq H$:

$$\rho \mapsto \sum_v |\phi_v\rangle \left(\langle\phi_v|\rho|\phi_v\rangle\right) \langle\phi_v|.$$

Can one interpolate between these two, describing open quantum systems which interact with an external environment?

Notice that both transformations are linear and map states into states.

When applied directly to mixed states $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$, we obtain the two transformations:

1. For the unitary evolution of closed systems:

$$\rho \mapsto U_t \rho U_t^\dagger,$$

2. For the measurement $(|\phi_v\rangle)_v \subseteq H$:

$$\rho \mapsto \sum_v |\phi_v\rangle \left(\langle\phi_v|\rho|\phi_v\rangle\right) \langle\phi_v|.$$

Can one interpolate between these two, describing open quantum systems which interact with an external environment?

Notice that both transformations are linear and map states into states.

Quantum channels

We need a further property for a linear transformation of a quantum system H, $\Phi : \rho \mapsto \Phi(\rho)$ which maps states into states to be a quantum channel.

Complete positivity: enlarging H to any system $H \otimes K$ and acting with $\Phi \otimes \mathbb{I}_K$ still maps joint states into states.

Quantum channels Φ therefore are defined as maps that are

- linear: $\Phi(\lambda\rho + \sigma) = \lambda\Phi(\rho) + \Phi(\sigma)$,
- completely positive: $\Phi \otimes \mathbb{I}_K(\rho)$ is positive (semidefinite) whenever ρ is so,
- trace preserving: $\mathrm{Tr}[\Phi(\rho)] = \mathrm{Tr}[\rho]$ (so that states are mapped into states).

Exercise: Show that both unitary evolutions of closed systems and measurements induce quantum channels.

Quantum channels

We need a further property for a linear transformation of a quantum system H, $\Phi : \rho \mapsto \Phi(\rho)$ which maps states into states to be a quantum channel.

Complete positivity: enlarging H to any system $H \otimes K$ and acting with $\Phi \otimes \mathbb{I}_K$ still maps joint states into states.

Quantum channels Φ therefore are defined as maps that are

- linear: $\Phi(\lambda\rho + \sigma) = \lambda\Phi(\rho) + \Phi(\sigma)$,
- completely positive: $\Phi \otimes \mathbb{I}_K(\rho)$ is positive (semidefinite) whenever ρ is so,
- trace preserving: $\mathrm{Tr}[\Phi(\rho)] = \mathrm{Tr}[\rho]$ (so that states are mapped into states).

Exercise: Show that both unitary evolutions of closed systems and measurements induce quantum channels.

Quantum channels

We need a further property for a linear transformation of a quantum system H, $\Phi : \rho \mapsto \Phi(\rho)$ which maps states into states to be a quantum channel.

Complete positivity: enlarging H to any system $H \otimes K$ and acting with $\Phi \otimes \mathbb{I}_K$ still maps joint states into states.

Quantum channels Φ therefore are defined as maps that are

- linear: $\Phi(\lambda\rho + \sigma) = \lambda\Phi(\rho) + \Phi(\sigma)$,
- completely positive: $\Phi \otimes \mathbb{I}_K(\rho)$ is positive (semidefinite) whenever ρ is so,
- trace preserving: $\mathrm{Tr}[\Phi(\rho)] = \mathrm{Tr}[\rho]$ (so that states are mapped into states).

Exercise: Show that both unitary evolutions of closed systems and measurements induce quantum channels.

Quantum channels

We need a further property for a linear transformation of a quantum system H, $\Phi : \rho \mapsto \Phi(\rho)$ which maps states into states to be a quantum channel.

Complete positivity: enlarging H to any system $H \otimes K$ and acting with $\Phi \otimes \mathbb{I}_K$ still maps joint states into states.

Quantum channels Φ therefore are defined as maps that are

- linear: $\Phi(\lambda\rho + \sigma) = \lambda\Phi(\rho) + \Phi(\sigma)$,
- completely positive: $\Phi \otimes \mathbb{I}_K(\rho)$ is positive (semidefinite) whenever ρ is so,
- trace preserving: $\mathrm{Tr}[\Phi(\rho)] = \mathrm{Tr}[\rho]$ (so that states are mapped into states).

Exercise: Show that both unitary evolutions of closed systems and measurements induce quantum channels.

Representations of quantum channels

These are useful ways to write (and think of) a channel Φ:

- (Kraus) There exists linear operators $(B_k)_k$ such that

$$\sum_k B_k^\dagger B_k = \mathbb{I}$$

  and representing Φ as

$$\Phi(\rho) = \sum_k B_k \rho B_k^\dagger.$$

- (Stinespring) There exists an auxiliary quantum system K, a state $|0_K\rangle \in K$ and unitary U acting on $H \otimes K$ such that, for every ρ.

$$\Phi(\rho) = \mathrm{Tr}_K[U(\rho \otimes |0_K\rangle\langle 0_K|)\, U^\dagger].$$

  One can choose K as (a copy of) H.

Exercise: describe the two representations for unitary evolutions of closed systems and measurements.

Representations of quantum channels

These are useful ways to write (and think of) a channel Φ:

- (Kraus) There exists linear operators $(B_k)_k$ such that

  $$\sum_k B_k^\dagger B_k = \mathbb{I}$$

  and representing Φ as

  $$\Phi(\rho) = \sum_k B_k \rho B_k^\dagger.$$

- (Stinespring) There exists an auxiliary quantum system K, a state $|0_K\rangle \in K$ and unitary U acting on $H \otimes K$ such that, for every ρ.

  $$\Phi(\rho) = \mathrm{Tr}_K[U\,(\rho \otimes |0_K\rangle\langle 0_K|)\,U^\dagger].$$

  One can choose K as (a copy of) H.

Exercise: describe the two representations for unitary evolutions of closed systems and measurements.

Representations of quantum channels

These are useful ways to write (and think of) a channel Φ:

- (Kraus) There exists linear operators $(B_k)_k$ such that

  $$\sum_k B_k^\dagger B_k = \mathbb{I}$$

  and representing Φ as

  $$\Phi(\rho) = \sum_k B_k \rho B_k^\dagger.$$

- (Stinespring) There exists an auxiliary quantum system K, a state $|0_K\rangle \in K$ and unitary U acting on $H \otimes K$ such that, for every ρ.

  $$\Phi(\rho) = \mathrm{Tr}_K[U\,(\rho \otimes |0_K\rangle\langle 0_K|)\,U^\dagger].$$

  One can choose K as (a copy of) H.

Exercise: describe the two representations for unitary evolutions of closed systems and measurements.

Plan

Monge's Optimal Transport problem

Monge (1781): *sur la théorie des déblais et des remblais*.

How to transport soil during a construction with minimal expenses?

A discrete formulation: given a

- cost $c(x, y)$ of moving unit of soil from position x to position y, e.g.

$$c(x, y) = |x - y|^p,$$

- Source distribution of soil $\sigma = (\sigma(x_i))_i$
- Target distribution (dump) $\rho = (\rho(y_j))_j$

Find $T : \{x_i\} \to \{y_j\}$ that moves σ into ρ with minimal transport cost

$$\sum_i c(x_i, T(x_i))\sigma(x_i).$$

Monge's Optimal Transport problem

Monge (1781): *sur la théorie des déblais et des remblais*.

How to transport soil during a construction with minimal expenses?

A discrete formulation: given a

- cost $c(x, y)$ of moving unit of soil from position x to position y, e.g.

$$c(x, y) = |x - y|^p,$$

- Source distribution of soil $\sigma = (\sigma(x_i))_i$
- Target distribution (dump) $\rho = (\rho(y_j))_j$

Find $T : \{x_i\} \to \{y_j\}$ that moves σ into ρ with minimal transport cost

$$\sum_i c(x_i, T(x_i))\sigma(x_i).$$

Monge's Optimal Transport problem

Monge (1781): *sur la théorie des déblais et des remblais*.

How to transport soil during a construction with minimal expenses?

A discrete formulation: given a

- cost $c(x, y)$ of moving unit of soil from position x to position y, e.g.

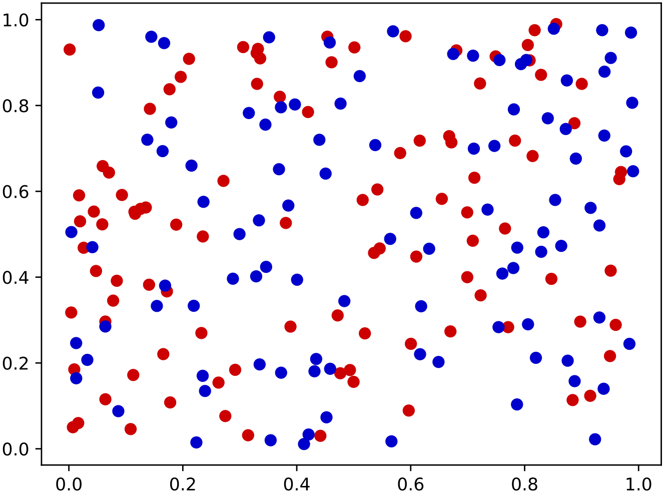
$$c(x, y) = |x - y|^p,$$

- Source distribution of soil $\sigma = (\sigma(x_i))_i$
- Target distribution (dump) $\rho = (\rho(y_j))_j$

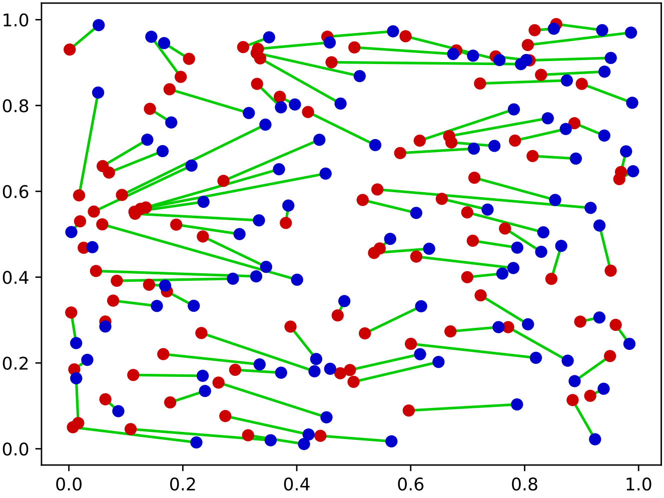
Find $T : \{x_i\} \to \{y_j\}$ that moves σ into ρ with minimal transport cost

$$\sum_i c(x_i, T(x_i))\sigma(x_i).$$

A random instance of OT in the plane



A random instance of OT in the plane



Kantorovich and linear programming

Relax a map T with a coupling

$$\pi(x_i, y_j) \geq 0$$

such that

$$\sum_j \pi(x_i, y_j) = \sigma(x_i), \quad \sum_i \pi(x_i, y_j) = \rho(y_j).$$

The problem becomes linear optimization with linear constraints:

$$\min_\pi \sum_i \sum_j c(x_i, y_j)\pi(x_i, y_j)$$

that can be solved via simplex algorithm.

Kantovorich and linear programming

Relax a map T with a coupling

$$\pi(x_i, y_j) \geq 0$$

such that

$$\sum_j \pi(x_i, y_j) = \sigma(x_i), \quad \sum_i \pi(x_i, y_j) = \rho(y_j).$$

The problem becomes linear optimization with linear constraints:

$$\min_\pi \sum_i \sum_j c(x_i, y_j)\pi(x_i, y_j)$$

that can be solved via simplex algorithm.

Wasserstein distance of order 1

If $c(x, y) = d(x, y)$ is a distance, then

$$W_1(\sigma, \rho) = \min_{\pi} \sum_i \sum_j d(x_i, y_j)\pi(x_i, y_j)$$

defines a distance between (discrete) probability densities.

Called Wasserstein distance of order 1 (aka Kantorovich-Rubinstein distance, or Earth Mover's distance).

Exercise: $(x, y) \mapsto d(x, y)^p$, with $p \in (0, 1)$, is a distance. What is

$$\lim_{p \to 0^+} W_{d^p}(\sigma, \rho)?$$

For $p > 1$, one defines a distance via

$$W_p(\sigma, \rho) = \left(\min_T \sum_i \sum_j d(x_i, y_j)^p T(x_i, y_j) \right)^{1/p}.$$

Wasserstein distance of order 1

If $c(x,y) = d(x,y)$ is a distance, then

$$W_1(\sigma,\rho) = \min_{\pi} \sum_i \sum_j d(x_i,y_j)\pi(x_i,y_j)$$

defines a distance between (discrete) probability densities.

Called Wasserstein distance of order 1 (aka Kantorovich-Rubinstein distance, or Earth Mover's distance).

Exercise: $(x,y) \mapsto d(x,y)^p$, with $p \in (0,1)$, is a distance. What is

$$\lim_{p \to 0^+} W_{d^p}(\sigma,\rho)?$$

For $p > 1$, one defines a distance via

$$W_p(\sigma,\rho) = \left(\min_T \sum_i \sum_j d(x_i,y_j)^p T(x_i,y_j) \right)^{1/p}.$$

Wasserstein distance of order 1

If $c(x, y) = d(x, y)$ is a distance, then

$$W_1(\sigma, \rho) = \min_{\pi} \sum_i \sum_j d(x_i, y_j)\pi(x_i, y_j)$$

defines a distance between (discrete) probability densities.

Called Wasserstein distance of order 1 (aka Kantorovich-Rubinstein distance, or Earth Mover's distance).

Exercise: $(x, y) \mapsto d(x, y)^p$, with $p \in (0, 1)$, is a distance. What is

$$\lim_{p \to 0^+} W_{d^p}(\sigma, \rho)?$$

For $p > 1$, one defines a distance via

$$W_p(\sigma, \rho) = \left(\min_T \sum_i \sum_j d(x_i, y_j)^p T(x_i, y_j) \right)^{1/p}.$$

Wasserstein distance of order 1

If $c(x, y) = d(x, y)$ is a distance, then

$$W_1(\sigma, \rho) = \min_{\pi} \sum_i \sum_j d(x_i, y_j)\pi(x_i, y_j)$$

defines a distance between (discrete) probability densities.

Called Wasserstein distance of order 1 (aka Kantorovich-Rubinstein distance, or Earth Mover's distance).

Exercise: $(x, y) \mapsto d(x, y)^p$, with $p \in (0, 1)$, is a distance. What is

$$\lim_{p \to 0^+} W_{d^p}(\sigma, \rho)?$$

For $p > 1$, one defines a distance via

$$W_p(\sigma, \rho) = \left(\min_T \sum_i \sum_j d(x_i, y_j)^p T(x_i, y_j) \right)^{1/p}.$$

Duality

- Kantorovich provided also the dual formulation for W_1:

$$W_1(\sigma, \rho) = \max_f \left\{ \sum_i f(x_i)\sigma(x_i) - \sum_j f(y_j)\rho(y_j) \ : |f(x) - f(y)| \leq d(x, y) \right\}.$$

  Inequality $\geq$ is trivial, the other follows from $\min\max$ theorems.

- It gives a definition of W_1 without transport plans, using only d-Lipschitz functions f.
- The formula also yields that $W_1(\sigma, \rho) = \|\sigma - \rho\|_{W_1}$ is induced by a norm.
- For general $c(x, y)$ duality uses conjugate functions (f, g):

$$f(x) - g(y) \leq c(x, y).$$

Duality

- Kantorovich provided also the dual formulation for W_1:

$$W_1(\sigma, \rho) = \max_f \left\{ \sum_i f(x_i)\sigma(x_i) - \sum_j f(y_j)\rho(y_j) \ : |f(x) - f(y)| \leq d(x, y) \right\}.$$

  Inequality $\geq$ is trivial, the other follows from $\min \max$ theorems.
- It gives a definition of W_1 without transport plans, using only d-Lipschitz functions f.
- The formula also yields that $W_1(\sigma, \rho) = \|\sigma - \rho\|_{W_1}$ is induced by a norm.
- For general $c(x, y)$ duality uses conjugate functions (f, g):

$$f(x) - g(y) \leq c(x, y).$$

Benamou-Brenier formula

- From a probabilist perspective

$$\min_{\pi} \sum_i \sum_j c(x_i, y_j)\pi(x_i, y_j) = \min_{X\sim\sigma, Y\sim\rho} \mathbb{E}\left[c(X, Y)\right]$$

- On $\mathbb{R}^d$ and $c(x, y) = |x - y|^p$, we have

$$|x - y|^p \leq \int_0^1 |\dot{x}_t|^p dt$$

along any (smooth) path $(x_t)_t$ with $x_0 = x$, $x_1 = y$.
- For any stochastic process $(X_t)_t$ with $X_0 \sim X$, $X_1 \sim Y$:

$$\mathbb{E}\left[|X - Y|^p\right] \leq \mathbb{E}\left[\int_0^1 |\dot{X}_t|^p dt\right]$$

- Benamou and Brenier (1999) proved that

$$W_p^p(\sigma, \rho) = \min_{(X_t)_t : X_0\sim\sigma, X_1\sim\rho} \mathbb{E}\left[\int_0^1 |\dot{X}_t|^p dt\right]$$

Other "distances" between probabilities

1. Total Variation distance

$$\|\rho - \sigma\|_{TV} = \frac{1}{2}\sum_x |\sigma(x) - \rho(x)|$$

2. Hellinger distance

$$H^2(\sigma, \rho) = \frac{1}{2}\sum_x |\sqrt{\sigma(x)} - \sqrt{\rho(x)}|^2 = 1 - \sum_x \sqrt{\sigma(x)}\sqrt{\rho(x)}$$

3. Kullback-Leibler divergence

$$D_{KL}(\sigma\|\rho) = \sum_x \sigma(x) \ln\left(\sigma(x)/\rho(x)\right).$$

No use of geometry of the space, i.e. the distance $d(x, y)$ between positions.

Exercise: Show that TV is Wasserstein distance w.r.t. $d(x, y) = 1_{\{x \neq y\}}$.

Other “distances” between probabilities

1. Total Variation distance

$$\|\rho - \sigma\|_{TV} = \frac{1}{2}\sum_x |\sigma(x) - \rho(x)|$$

2. Hellinger distance

$$H^2(\sigma, \rho) = \frac{1}{2}\sum_x |\sqrt{\sigma(x)} - \sqrt{\rho(x)}|^2 = 1 - \sum_x \sqrt{\sigma(x)}\sqrt{\rho(x)}$$

3. Kullback-Leibler divergence

$$D_{KL}(\sigma\|\rho) = \sum_x \sigma(x)\ln\left(\sigma(x)/\rho(x)\right).$$

No use of geometry of the space, i.e. the distance $d(x, y)$ between positions.

Exercise: Show that TV is Wasserstein distance w.r.t. $d(x, y) = 1_{\{x \neq y\}}$.

Other "distances" between probabilities

1. Total Variation distance

$$\|\rho - \sigma\|_{TV} = \frac{1}{2}\sum_x |\sigma(x) - \rho(x)|$$

2. Hellinger distance

$$H^2(\sigma, \rho) = \frac{1}{2}\sum_x |\sqrt{\sigma(x)} - \sqrt{\rho(x)}|^2 = 1 - \sum_x \sqrt{\sigma(x)}\sqrt{\rho(x)}$$

3. Kullback-Leibler divergence

$$D_{KL}(\sigma\|\rho) = \sum_x \sigma(x)\ln\left(\sigma(x)/\rho(x)\right).$$

No use of geometry of the space, i.e. the distance $d(x, y)$ between positions.

Exercise: Show that TV is Wasserstein distance w.r.t. $d(x, y) = 1_{\{x \neq y\}}$.

Plan

An overview of Quantum OT

Classical distances between probabilities have quantum analogues:

- Total variation $\to$ Trace distance $\frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\mathrm{Tr}[|\rho - \sigma|]$
- Hellinger distance $\to$ Fidelity $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$.
- Kullback-Leibler divergence $\to$ Relative entropy $S(\rho\|\sigma)$.

As their classical counterparts:

+ Quite general, easy to compute or approximate
- Not adapted to specific geometry, i.e., unitarily invariant:

$$d(\rho, \sigma) = d(U\rho U^\dagger, U\sigma U^\dagger).$$

What about Quantum Optimal Transport distances?

An overview of Quantum OT

Classical distances between probabilities have quantum analogues:

- Total variation $\to$ Trace distance $\frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\operatorname{Tr}[|\rho - \sigma|]$
- Hellinger distance $\to$ Fidelity $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$.
- Kullback-Leibler divergence $\to$ Relative entropy $S(\rho\|\sigma)$.

As their classical counterparts:

\+ Quite general, easy to compute or approximate

\- Not adapted to specific geometry, i.e., unitarily invariant:

$$d(\rho, \sigma) = d(U\rho U^\dagger, U\sigma U^\dagger).$$

What about Quantum Optimal Transport distances?

An overview of Quantum OT

Classical distances between probabilities have quantum analogues:

- Total variation $\rightarrow$ Trace distance $\frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\operatorname{Tr}[|\rho - \sigma|]$
- Hellinger distance $\rightarrow$ Fidelity $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$.
- Kullback-Leibler divergence $\rightarrow$ Relative entropy $S(\rho\|\sigma)$.

As their classical counterparts:

\+ Quite general, easy to compute or approximate

\- Not adapted to specific geometry, i.e., unitarily invariant:

$$d(\rho, \sigma) = d(U\rho U^\dagger, U\sigma U^\dagger).$$

What about Quantum Optimal Transport distances?

Quantum OT theories: a timeline

- 1992 - Connes/Lott:
  spectral distance in non-commutative geometry
- 1997 - Zyczkowski/Slomczynski:
  Wasserstein distance of Husimi distributions
- 2012 - Maas/Carlen:
  quantum analogue of Benamou-Brenier formula
- 2013 - Agredo:
  1-Wasserstein extending any distance on basis vectors
- 2016 - Golse/Mouhot/Paul:
  quantum Kantorovich problem (plans)
- 2019 - De Palma/T.:
  quantum optimal transport with channels (couplings)
- 2020 - De Palma/Marvian/T./Lloyd:
  Wasserstein distance on qubits (Hamming)

A classification

Classically equivalent definitions of $W_1(\rho, \sigma)$:

Monge-Kantorovich	Dual Kantorovich	Benamou-Brenier
$\min_{\pi} \int c(x,y)\pi(x,y)$	$\max_{(f,g)} \int f d\rho - \int g d\sigma$	$\min_{(X_t)} \mathbb{E} \int_0^1 \lvert\dot{X}_t\rvert^p dt$
• Zyczkowski/ Slomczynski • Golse/Mouhot/Paul • De Palma/T.	• Connes/Lott • Agredo • De Palma/Marvian /T./Lloyd	• Maas/Carlen

A classification

Classically equivalent definitions of $W_1(\rho, \sigma)$:

Monge-Kantorovich	Dual Kantorovich	Benamou-Brenier
$\min_{\pi} \int c(x,y)\pi(x,y)$	$\max_{(f,g)} \int f d\rho - \int g d\sigma$	$\min_{(X_t)} \mathbb{E} \int_0^1 \lvert \dot{X}_t \rvert^p dt$
• Zyczkowski/ Slomczynski • Golse/Mouhot/Paul • De Palma/T.	• Connes/Lott • Agredo • De Palma/Marvian /T./Lloyd	• Maas/Carlen

A classification

Classically equivalent definitions of $W_1(\rho, \sigma)$:

Monge-Kantorovich	Dual Kantorovich	Benamou-Brenier
$\min_{\pi} \int c(x,y)\pi(x,y)$	$\max_{(f,g)} \int f d\rho - \int g d\sigma$	$\min_{(X_t)} \mathbb{E} \int_0^1 \lvert \dot{X}_t \rvert^p dt$
Zyczkowski/ Slomczynski Golse/Mouhot/Paul De Palma/T.	Connes/Lott Agredo De Palma/Marvian /T./Lloyd	Maas/Carlen

A classification

Classically equivalent definitions of $W_1(\rho, \sigma)$:

Monge-Kantorovich	Dual Kantorovich	Benamou-Brenier
$\min_{\pi} \int c(x,y)\pi(x,y)$	$\max_{(f,g)} \int f d\rho - \int g d\sigma$	$\min_{(X_t)} \mathbb{E} \int_0^1 \lvert\dot{X}_t\rvert^p dt$
• Zyczkowski/ Slomczynski • Golse/Mouhot/Paul • De Palma/T.	• Connes/Lott • Agredo • De Palma/Marvian /T./Lloyd	• Maas/Carlen

A classification

Classically equivalent definitions of $W_1(\rho, \sigma)$:

Monge-Kantorovich	Dual Kantorovich	Benamou-Brenier
$\min_{\pi} \int c(x,y)\pi(x,y)$	$\max_{(f,g)} \int f d\rho - \int g d\sigma$	$\min_{(X_t)} \mathbb{E} \int_0^1 \lvert\dot{X}_t\rvert^p dt$
• Zyczkowski/ Slomczynski • Golse/Mouhot/Paul • De Palma/T.	• Connes/Lott • Agredo • De Palma/Marvian /T./Lloyd	• Maas/Carlen

The GMPC approach

Consider a quantum system H and states $\sigma, \rho \in \mathcal{S}(H)$.
Following Kantorovich, we need

- Couplings:any $\Pi \in \mathcal{S}(H \otimes H)$ with

$$\mathrm{Tr}_1 \Pi = \rho, \quad \mathrm{Tr}_2 \Pi = \sigma$$

- Transportation cost: any observable C on $H \otimes H$.

Example: fix $R_1, \ldots, R_d$ "quadratures" on H, set

$$C = \sum_{i=1}^{d} (R_i \otimes \mathbb{I} - \mathbb{I} \otimes R_1)^2.$$

The GMPC transportation problem is

$$\inf_{\Pi} \mathrm{Tr}[C\Pi].$$

Problems

The GMPC transportation problem is

$$\inf_{\Pi} \mathrm{Tr}[C\Pi].$$

1. Show that the inf is attained (assuming H finite dimensional, or C bounded). Are minimizers unique?
2. Describe the set of couplings if σ (or ρ) is a pure state.
3. (*) On the n-qubit system, assume that ρ, σ and C are diagonal in the computational basis (hence identified with classical functions and probabilities).
   Is the GMPC transportation cost the same as the classical transportation?

Plan

Peyré, G., and Cuturi, M.
Computational Optimal Transport
https://arxiv.org/abs/1803.00567

Vershik, A.M.
Long history of the Monge-Kantorovich transportation problem.
The Mathematical Intelligencer, 2013.

Villani, C.
Optimal transport: old and new.
Springer, 2009.

De Palma, G., and Trevisan, D.
Quantum optimal transport with quantum channels.
Annales Henri Poincaré. Springer International Publishing, 2021.

De Palma, G., Marvian, M., Trevisan, D., and Lloyd, S.
The quantum Wasserstein distance of order 1
IEEE Transactions on Information Theory, 2021.

Connes, A., and Lott, J.
The metric aspect of noncommutative geometry
New symmetry principles in quantum field theory, 1992

Zyczkowski, K. and Slomczynski, W.
The Monge distance between quantum states.
Journal of Physics A: Mathematical and General, 1998.

Agredo, J.
A Wasserstein-type distance to measure deviation from equilibrium of quantum Markov semigroups
Open Systems & Information Dynamics, 2013.

Carlen, E.A. and Maas, J.
An analog of the 2-Wasserstein metric in non-commutative probability under which the Fermionic Fokker–Planck equation is gradient flow for the entropy
Communications in mathematical physics, 2014.

Golse, F., Mouhot, C., and Paul, T.
On the mean field and classical limits of quantum mechanics
Communications in Mathematical Physics, 2016