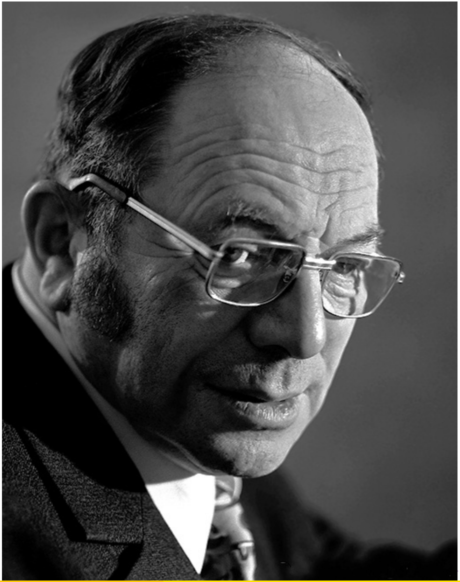
Optimal transport by quantum channels: non-quadratic problems, metric properties, and isometries

Dániel Virosztek
MTA RI 'Momentum' Optimal Transport and Quantum Information Geometry Research Group

WS II: Dynamics of Density Operators, UCLA IPAM, April 28, 2025

Gaspard Monge and Leonid Kantorovich

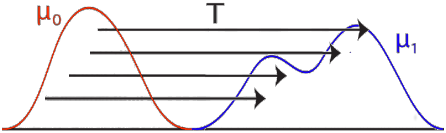


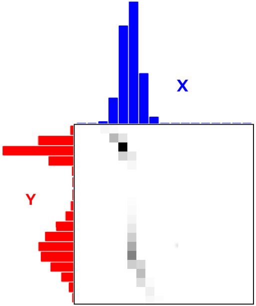
The classical (Monge-Kantorovich) optimal transport problem

- one has to transport goods (bread) from producers (bakeries) to costumers (convenience stores)
- the distribution of producers is described by a Borel probability measure μ on the underlying (Polish) metric space X: $\mathrm{d}\mu(x) \approx$ the *capacity of production* at x
- the distribution of costumers is described by another Borel probability measure ν: $\mathrm{d}\nu(y) \approx$ the *demand* at y

The classical (Monge-Kantorovich) optimal transport problem

- goal: find a **transport plan** from μ to ν that minimizes the cost
- what is a transport plan?
    - **Monge**: it is a measurable map from $T : X \to X$ such that $T_{\#}\mu = \nu$
    - **Kantorovich**: it is a probability measure π on $X \times X$ with $\pi_{|1} = \mu, \pi_{|2} = \nu$ and $\mathrm{d}\pi(x, y) \approx$ the amount of goods to be transferred from x to y





The classical (Monge-Kantorovich) optimal transport problem

- what is the transport cost?
- denote by $c(x, y)$ the cost of transporting a unit of goods from x to y (c is ≥ 0, l.s.c., typically symmetric, etc.)
- then the cost of the transport plan π is

$$C_\pi = \int_{X^2} c(x, y) \mathrm{d}\pi(x, y) = \langle c, \pi \rangle$$

- and the minimal cost of the transport $\mu \to \nu$ given the cost function $c(x, y)$ is

$$W_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} C_\pi$$

  where $\Pi(\mu, \nu) = \left\{ \pi \in \mathcal{P}\left(X^2\right) \middle| \pi_{|1} = \mu, \pi_{|2} = \nu \right\}$
- this is a linear optimization task on a convex domain no matter what $c(x, y)$ is!

The classical (Monge-Kantorovich) optimal transport problem

- what is $c(x, y)$?
- in many cases (from economics/real life) $c(x, y)$ is a *subadditive* function of the distance, say, $c(x, y) = d(x, y)^p$ for some $0 < p < 1$

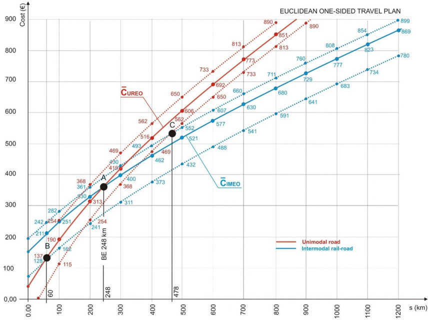


Figure: Zgonc, B., Tekavcic, M., Jakcic, M., The impact of distance on mode choice in freight transport. *Eur. Transp. Res. Rev.* **11** (2019), 10.

What if you are short on fuel?

... and you want to minimize the **time-average of the kinetic energy**?[1]

$$A[\rho, v] = \int_0^1 \int_X \rho_t(x) \left\| v_t(x) \right\|^2 \mathrm{d}x \mathrm{d}t,$$

where $\{\rho_t, v_t\}_{t \in [0,1]}$ is a weak solution of the linear transport equation $\frac{\partial \rho_t}{\partial t} + \nabla_x \cdot (\rho_t v_t) = 0$ with initial and final conditions $\rho_0 = \mu, \rho_1 = \nu$

[1] J.-D. Benamou, Y. Breiner, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), 375–393.

Dynamical interpretation and connections to fluid mechanics

- in general, which $(\rho_t)_{t\in[0,1]}$ flow minimizes the kinetic energy?
- answered by Benamou and Breiner in 2000:
  - consider the (static) OT problem with the quadratic cost $c(x,y)=d(x,y)^2$
  - denote by π the optimal transport plan
  - let $\gamma_{(x,y)} : [0,1] \ni t \mapsto \gamma_{(x,y)}(t) \in X$ be the unique geodesic connecting x and y
  - let $e(t) : (x,y) \mapsto \gamma_{(x,y)}(t)$ be the evaluation at time t
  - the minimizer is the **displacement interpolation** given by

$$\rho_t = e(t)_{\#}\pi$$

  - and the minimal energy is

$$\int_{X^2} \left(\frac{d(x,y)}{1}\right)^2 \mathrm{d}\pi(x,y) = d_{W_2}^2(\mu,\nu)$$

Wasserstein spaces and OT for probabilists

- for a Polish space (X, ρ) and a parameter $0 < p < \infty$ the *p-Wasserstein space* is

$$\mathcal{W}_p(X) = \left\{ \mu \in \mathcal{P}(X) \,\middle|\, \int_X \rho(x, \hat{x})^p \, \mathrm{d}\mu(x) < \infty \text{ for some } \hat{x} \in X \right\}$$

endowed with the *p-Wasserstein distance*

$$d_{\mathcal{W}_p}(\mu, \nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{X^2} \rho(x, y)^p \, \mathrm{d}\pi(x, y) \right)^{\min\left\{\frac{1}{p}, 1\right\}}$$

- the probabilistic interpretation of the OT problem:

$$\mathbb{E}\, c(X, Y) \to \min \qquad \texttt{subject to} \qquad \mathrm{law}(X) = \mu,\, \mathrm{law}(Y) = \nu$$

- example: if $c(x, y) = \|x - y\|^2$, then $\mathbb{E}\, \|X - Y\|^2 \to \min$ is equivalent to $\mathbb{E}\, \langle X, Y \rangle \to \max$

Basics of non-commutative optimal transport

- several different approaches (a far-from-complete list):
  - Biane and Voiculescu (free probability)
  - Carlen and Maas (dynamical interpretation)
  - Caglioti, Golse, Mouhot, and Paul (static interpretation)
  - De Palma and Trevisan (quantum channels)
  - Życzkowski and Słomczyński (semi-classical approach)
- most relevant approaches for us are that of Caglioti-Golse-Mouhot-Paul[2] and De Palma-Trevisan[3]

[2] F. Golse, C. Mouhot and T. Paul, *On the mean-field and classical limits of quantum mechanics,* Commun. Math. Phys., **343** (2016), 165–205.

[3] G. De Palma and D. Trevisan, *Quantum optimal transport with quantum channels*, Ann. Henri Poincaré **22** (2021), 3199–3234.

Basics of non-commutative optimal transport

- classical mechanics: the state of a particle moving in $\mathbb{R}^d$ is described by a probability measure μ on the phase space $\mathbb{R}^d \times \mathbb{R}^d$ which is the collection of all possible values of the position and momentum variables $q, p \in \mathbb{R}^d$
- so the classical quadratic OT distance of the states $\mu, \nu \in \mathcal{P}\left(\mathbb{R}^{2d}\right)$ is

$$d_{\mathcal{W}_2}^2(\mu,\nu) = \inf_{\mathrm{law}(Q_1,P_1)=\mu,\, \mathrm{law}(Q_2,P_2)=\nu} \left\{ \mathbb{E}\left\| (Q_1,P_1) - (Q_2,P_2) \right\|^2 \right\}$$

- quantum mechanics: a state of the same system is described by a wave function $\psi \in L^2\left(\mathbb{R}^d\right)$ of unit norm, or more generally, by a normalized, positive, trace-class operator ρ on $\mathcal{H} = L^2\left(\mathbb{R}^d\right)$
- measurable physical quantities $\leftrightarrow$ (possibly unbounded) self-adjoint operators on $\mathcal{H}$

Basics of non-commutative optimal transport

- Born's rule: when measuring an observable quantity $A = A^* \in \mathrm{Lin}(\mathcal{H})$ on a quantum system being in the state $\rho \in \mathcal{S}(\mathcal{H})$, the probability of the outcome lying in an interval $[a, b] \subset \mathbb{R}$ is $\mathrm{tr}_{\mathcal{H}}(\rho E_A([a, b]))$, where E_A is the spectral measure of A — the unique POVM satisfying $A = \int_{\mathbb{R}} \lambda \mathrm{d} E_A(\lambda)$
- a quantum state *encapsulates several classical probability distributions*, each corresponding to a physical quantity we are interested in
- let $\mathcal{A} = \{A_1, \dots, A_K\}$ be a finite collection of observable quantities, let us fix the initial state ρ and the final state ω
- let X_k (resp. Y_k) denote the random variable obtained by measuring A_k in ρ (resp. ω), that is, $\mathbb{P}(X_k \in [a, b]) = \mathrm{tr}_{\mathcal{H}}(\rho E_k([a, b]))$ (resp. $\mathbb{P}(Y_k \in [a, b]) = \mathrm{tr}_{\mathcal{H}}(\omega E_k([a, b]))$)

QOT via quantum channels

- transport plans by quantum channels?
- the classical correspondence between Markov maps and couplings: if $\xi_1, \xi_2, \ldots$ is a discrete-time Markov process (time-homogeneous) on a finite state space driven by the Markov kernel $K(x, y) := \mathbb{P}(\xi_{n+1} = y | \xi_n = x)$ and $\mu = \mathrm{law}(\xi_n)$, then π defined by $\pi(x, y) := \mu(\{x\}) K(x, y)$ is a coupling of $\mu = \mathrm{law}(\xi_n)$ and $\nu := \mathrm{law}(\xi_{n+1})$
- the idea of De Palma and Trevisan: take the *Choi-Jamiolkowski* isomorphism of $\mathcal{H} \to \mathcal{H}$ quantum channels and states on $\mathcal{H} \otimes \mathcal{H}^*$ given by

$$\mathrm{CH}(\mathcal{H}, \mathcal{H}) \ni \Phi \mapsto \frac{1}{N} \sum_{j,k=1}^{N} \Phi\left(|e_j\rangle \langle e_k|\right) \otimes |f_j\rangle \langle f_k| \in \mathcal{S}\left(\mathcal{H} \otimes \mathcal{H}^*\right)$$

QOT via quantum channels

- now take the canonical purification of $\mathcal{S}(\mathcal{H}) \ni \rho = \sum_{j=1}^{N} \lambda_j |e_j\rangle \langle e_j|$ on $\mathcal{H} \otimes \mathcal{H}^*$ which is

$$||\sqrt{\rho}\rangle\rangle = \sum_{j=1}^{N} \sqrt{\lambda_j} e_j \otimes f_j \leftrightarrow$$

$$\leftrightarrow ||\sqrt{\rho}\rangle\rangle \langle\langle\sqrt{\rho}|| = \sum_{j,k=1}^{N} \sqrt{\lambda_j \lambda_k} \left(|e_j\rangle \langle e_k|\right) \otimes \left(|f_j\rangle \langle f_k|\right)$$

- ... and $\Pi_\Phi := \left(\Phi \otimes \mathrm{Id}_{\mathcal{B}(\mathcal{H}^*)}\right) \left(\left|\left|\sqrt{\rho}\right\rangle\right\rangle \left\langle\left\langle\sqrt{\rho}\right|\right|\right)$
- straightforward calculation shows that $\mathrm{tr}_{\mathcal{H}^*} \Pi_\Phi = \Phi(\rho)$ and $\mathrm{tr}_{\mathcal{H}} \Pi_\Phi = \rho^T$
- the other direction: given a quantum coupling $\pi \in \mathcal{C}(\rho, \omega)$ the map

$$X \mapsto \mathrm{tr}_{\mathcal{H}^*} \left(\left(I_{\mathcal{H}} \otimes \left(\rho^{-\frac{1}{2}} X \rho^{-\frac{1}{2}} \right)^T \right) \cdot \pi \right)$$

is a quantum channel that sends ρ to ω

QOT via quantum channels

- the set of couplings of ρ and ω is defined by

$$\mathcal{C}(\rho,\omega)=\left\{\Pi\in\mathcal{S}(\mathcal{H}\otimes\mathcal{H}^*)\,\middle|\,\mathrm{tr}_{\mathcal{H}^*}\Pi=\omega,\ \mathrm{tr}_{\mathcal{H}}\Pi=\rho^T\right\}$$

- that is, a coupling of ρ and ω is a state Π on $\mathcal{H}\otimes\mathcal{H}^*$ such that

$$\mathrm{tr}_{\mathcal{H}\otimes\mathcal{H}^*}[(A\otimes I_{\mathcal{H}^*})\Pi]=\mathrm{tr}_{\mathcal{H}}[\omega A]$$

  and

$$\mathrm{tr}_{\mathcal{H}\otimes\mathcal{H}^*}\left[\left(I_{\mathcal{H}}\otimes B^T\right)\Pi\right]=\mathrm{tr}_{\mathcal{H}^*}[\rho^T B^T]=\mathrm{tr}_{\mathcal{H}}[\rho B]$$

  for all bounded $A,B\in\mathrm{Lin}(\mathcal{H})^{sa}$

- compare to: $\pi\in\mathcal{P}(X^2)$ is a coupling of $\mu\in\mathcal{P}(X)$ and $\nu\in\mathcal{P}(X)$ iff $\iint_{X^2}f(x)\mathrm{d}\pi(x,y)=\int_X f(x)\mathrm{d}\mu(x)$ and $\iint_{X^2}g(y)\mathrm{d}\pi(x,y)=\int_X g(y)\mathrm{d}\nu(y)$ for every $f,g\in C_b(X)$

QOT via quantum channels

- Let $\Pi \in \mathcal{C}(\rho, \omega)$ and Y_k and X_k denote the real random variables we obtain by measuring A_k on the $\mathcal{H}$ part and A_k^T on the $\mathcal{H}^*$ of the kth copy, respectively
- by Born's rule on quantum measurement, the (infinitesimal) probabilities describing the possible outcomes of the measurements are given for every k by

$$\mathrm{d}\mathbb{P}_{(\Pi)}^{(\mathcal{A})}(X_k = x_k, Y_k = y_k) = \mathrm{tr}_{\mathcal{H}\otimes\mathcal{H}^*}\left[\Pi\left(\mathrm{d}E_k(y_k) \otimes \mathrm{d}E_k^T(x_k)\right)\right]$$

QOT via quantum channels

- by the independence of the measurements on different copies, the joint law $\mathbb{P}^{(\mathcal{A})}_{(\Pi)}$ of the $\mathbb{R}^K$-valued random vectors $X = (X_1, \ldots, X_K)$ and $Y = (Y_1, \ldots Y_K)$ is given by

$$\mathrm{d}\mathbb{P}^{(\mathcal{A})}_{(\Pi)}(x_1, \ldots, x_K, y_1, \ldots, y_K) = \prod_{k=1}^{K} \mathrm{tr}_{\mathcal{H}\otimes\mathcal{H}^*} \left[\Pi \left(\mathrm{d}E_k(y_k) \otimes \mathrm{d}E_k^T(x_k)\right)\right]$$

$$= \mathrm{tr}_{(\mathcal{H}\otimes\mathcal{H}^*)^{\otimes K}} \left[\Pi^{\otimes K} \left(\mathrm{d}E_1(y_1) \otimes \mathrm{d}E_1^T(x_1) \otimes \cdots \otimes \mathrm{d}E_K(y_K) \otimes \mathrm{d}E_K^T(x_K)\right)\right]$$

- given a non-negative classical transport cost $c : \mathbb{R}^K \times \mathbb{R}^K \to \mathbb{R}_+$, our goal is to minimize

$$\mathcal{S}(\mathcal{H} \otimes \mathcal{H}^*) \ni \Pi \mapsto \mathbb{E}^{(\mathcal{A})}_{(\Pi)} [c(X, Y)]$$

$$= \iint_{\mathbb{R}^K \times \mathbb{R}^K} c(x_1, \ldots, x_K, y_1, \ldots, y_K) \mathrm{d}\mathbb{P}^{(\mathcal{A})}_{(\Pi)}(x_1, \ldots, x_K, y_1, \ldots, y_K)$$

QOT via quantum channels

- we[4] define the positive and possibly unbounded quantum cost operator $C_c^{(\mathcal{A})}$ by

$$C_c^{(\mathcal{A})} := \iint_{\mathbb{R}^K \times \mathbb{R}^K} c(x_1, \dots, x_K, y_1, \dots, y_K) \times$$

$$\times \mathrm{d}E_1(y_1) \otimes \mathrm{d}E_1^T(x_1) \otimes \dots \otimes \mathrm{d}E_K(y_K) \otimes \mathrm{d}E_K^T(x_K)$$

- and propose the following quantum optimal transport problem:

$$\text{minimize } \Pi \mapsto \mathrm{tr}_{(\mathcal{H} \otimes \mathcal{H}^*)^{\otimes K}} \left[\Pi^{\otimes K} C_c^{(\mathcal{A})} \right]$$

where Π runs over the set of all couplings of ρ and ω

[4] Bunth-Pitrik-Titkos-Virosztek, Wasserstein distances and divergences of order p by quantum channels, arXiv:2501.08066

Quantum Wasserstein distances and divergences of order p

Proposition (Existence of optimal plans)

Let $c: \mathbb{R}^K \times \mathbb{R}^K \to \mathbb{R}$ be a non-negative and lower semi-continuous function, let $\mathcal{A}$ be a finite collection of observables on $\mathcal{H}$, and $\rho, \omega \in \mathcal{S}(\mathcal{H})$ given marginals. Then there exists an optimal solution $\Pi_0 \in \mathcal{C}(\rho, \omega)$ of the optimization problem

$$\textit{minimize } \Pi \mapsto \operatorname{tr}_{(\mathcal{H} \otimes \mathcal{H}^*)^{\otimes K}} \left[\Pi^{\otimes K} C_c^{(\mathcal{A})} \right]$$

- proof: a compactness/tightness argument like in the classical case
- important special case: $c(x_1, \ldots, x_K, y_1, \ldots, y_K) = \left(\sum_{k=1}^K |x_k - y_k|^q \right)^{\frac{p}{q}}$ where $p > 0$ and $q \geq 1$

Quantum Wasserstein distances and divergences of order p

- in this case,

$$C_{p,q}^{(\mathcal{A})} = \iint_{\mathbb{R}^K \times \mathbb{R}^K} \left(\sum_{k=1}^{K} |x_k - y_k|^q \right)^{\frac{p}{q}} \bigotimes_{k=1}^{K} \mathrm{d}E_k(y_k) \otimes \mathrm{d}E_k^T(x_k)$$

$$= \left(\sum_{k=1}^{K} \left(\left| A_k \otimes I_{\mathcal{H}}^T - I_{\mathcal{H}} \otimes A_k^T \right|^q \right)^{(k)} \right)^{\frac{p}{q}}$$

- even more special: $q = p \to$ for $p > 0$ we define

$$C_{\mathcal{A},p} := \sum_{k=1}^{K} \iint_{\mathbb{R}^2} |x - y|^p \, \mathrm{d}E_k(y) \otimes \mathrm{d}E_k^T(x) = \sum_{k=1}^{K} \left| A_k \otimes I^T - I \otimes A_k^T \right|^p$$

- the p-Wasserstein distance of ρ and ω w.r.t. $\mathcal{A} = \{A_1, \dots, A_K\}$ is

$$D_{\mathcal{A},p}(\rho, \omega) := \left(\inf_{\Pi \in \mathcal{C}(\rho,\omega)} \{ \mathrm{tr}_{\mathcal{H} \otimes \mathcal{H}^*} [\Pi C_{\mathcal{A},p}] \} \right)^{\min\left\{\frac{1}{p}, 1\right\}}$$

Quantum Wasserstein distances and divergences of order p

- the quadratic quantum Wasserstein *divergences* suggested by De Palma and Trevisan are defined by

$$d_{\mathcal{A},2}(\rho,\omega) := \sqrt{D^2_{\mathcal{A},2}(\rho,\omega) - \frac{1}{2}\left(D^2_{\mathcal{A},2}(\rho,\rho) + D^2_{\mathcal{A},2}(\omega,\omega)\right)}$$

  and conjectured to be genuine metrics on quantum state spaces
- therefore we define[5] the $(\mathcal{A}, p)$-Wasserstein *divergence* of ρ and ω by

$$d_{\mathcal{A},p}(\rho,\omega) = \left(D^{\max\{p,1\}}_{\mathcal{A},p}(\rho,\omega) - \frac{1}{2}\left(D^{\max\{p,1\}}_{\mathcal{A},p}(\rho,\rho) + D^{\max\{p,1\}}_{\mathcal{A},p}(\omega,\omega)\right)\right)^{\min\left\{\frac{1}{p},1\right\}}$$

[5] Bunth-Pitrik-Titkos-Virosztek, Wasserstein distances and divergences of order p by quantum channels, arXiv:2501.08066

Quantum Wasserstein divergences: back to quadratic

- quadratic cost operators are of the form

$$C = \sum_{j=1}^{k} \left(A_j \otimes I_{\mathcal{H}^*} - I_{\mathcal{H}} \otimes A_j^T \right)^2$$

- the corresponding quadratic quantum Wasserstein distance is defined by

$$D_C(\rho,\omega)^2 = \inf_{\Pi \in \mathcal{C}(\rho,\omega)} \left\{ \mathrm{tr}_{\mathcal{H}\otimes\mathcal{H}^*} (\Pi\, C) \right\} =$$

$$= \inf_{\Phi \in \mathrm{CPTP}(\rho,\omega)} \left\{ \sum_{j=1}^{k} \mathrm{tr}_{\mathcal{H}} \left((\rho+\omega) A_j^2 - 2\sqrt{\rho} A_j \sqrt{\rho} \Phi^\dagger (A_j) \right) \right\}$$

- the fact that the self-distance $D_C(\rho,\rho)$ is realized by the identity channel for every $\rho \in \mathcal{S}(\mathcal{H})$ together implies that $d_C^2(\rho,\omega) =$

$$= \inf_{\Phi \in \mathrm{CPTP}(\rho,\omega)} \left\{ \sum_{j=1}^{k} \mathrm{tr}_{\mathcal{H}} \left(\left(\rho^{1/2} A_j\right)^2 + \left(\omega^{1/2} A_j\right)^2 - 2\rho^{1/2} A_j \rho^{1/2} \Phi^\dagger (A_j) \right) \right\}$$

Triangle inequality for quantum Wasserstein divergences

Theorem (Bunth-Pitrik-Titkos-Virosztek, Phys. Rev. A, 2024)

The triangle inequality

$$d_C(\tau,\rho)+d_C(\rho,\omega)\geq d_C(\tau,\omega)$$

holds for any $\tau,\omega\in\mathcal{S}(\mathcal{H})$, *any* $\rho\in\mathcal{P}_1(\mathcal{H})$, *and any quadratic cost* C.

- since ρ is pure, we have

$$d_C^2(\tau,\rho)=\sum_{j=1}^N\left(\mathrm{tr}(\tau^{1/2}A_j)^2+\mathrm{tr}(\rho^{1/2}A_j)^2-2(\mathrm{tr}\rho A_j)(\mathrm{tr}\tau A_j)\right)$$

and

$$d_C^2(\rho,\omega)=\sum_{j=1}^N\left(\mathrm{tr}(\rho^{1/2}A_j)^2+\mathrm{tr}(\omega^{1/2}A_j)^2-2(\mathrm{tr}\omega A_j)(\mathrm{tr}\rho A_j)\right)$$

Triangle inequality for quantum Wasserstein divergences

- by relaxing the infimum in the definition of QOT distance to the tensor product coupling,

$$d_C^2(\tau,\omega) \leq \sum_{j=1}^{N} \left(\mathrm{tr}(\tau^{1/2} A_j)^2 + \mathrm{tr}(\omega^{1/2} A_j)^2 - 2(\mathrm{tr}\omega A_j)(\mathrm{tr}\tau A_j) \right)$$

- the triangle inequality is equivalent to

$$2 d_C(\tau,\rho) d_C(\rho,\omega) \geq d_C^2(\tau,\omega) - \left(d_C^2(\tau,\rho) + d_C^2(\rho,\omega) \right)$$

- if X and Y are self-adjoint, $X \geq 0$, and $\mathrm{tr}X = 1$, then by the Cauchy-Schwartz for the HS-inner product on $X^{1/2}$ and $X^{1/4} Y X^{1/4}$ we have $\mathrm{tr}(X^{1/2} Y)^2 \geq (\mathrm{tr}XY)^2$

Triangle inequality for quantum Wasserstein divergences

- hence we get the following upper bound for the RHS of

$$RHS \leq \sum_{j=1}^{N} \left(\mathrm{tr}(\tau^{1/2} A_j)^2 + \mathrm{tr}(\omega^{1/2} A_j)^2 - 2(\mathrm{tr}\omega A_j)(\mathrm{tr}\tau A_j) \right) -$$

$$- \sum_{j=1}^{N} \left(\mathrm{tr}(\tau^{1/2} A_j)^2 + \mathrm{tr}(\rho^{1/2} A_j)^2 - 2(\mathrm{tr}\rho A_j)(\mathrm{tr}\tau A_j) \right) -$$

$$- \sum_{j=1}^{N} \left(\mathrm{tr}(\rho^{1/2} A_j)^2 + \mathrm{tr}(\omega^{1/2} A_j)^2 - 2(\mathrm{tr}\omega A_j)(\mathrm{tr}\rho A_j) \right) \leq$$

$$\leq \sum_{j=1}^{N} 2(\mathrm{tr}\rho A_j - \mathrm{tr}\omega A_j)(\mathrm{tr}\tau A_j - \mathrm{tr}\rho A_j) \tag{1}$$

Triangle inequality for quantum Wasserstein divergences

- now a Cauchy-Schwartz for the Euclidean space $\mathbb{R}^N$ tells us that

$$\sum_{j=1}^{N} 2(\mathrm{tr}\rho A_j - \mathrm{tr}\omega A_j)(\mathrm{tr}\tau A_j - \mathrm{tr}\rho A_j) \leq$$

$$\leq 2\left(\sum_{j=1}^{N}(\mathrm{tr}\rho A_j - \mathrm{tr}\omega A_j)^2\right)^{1/2}\left(\sum_{k=1}^{N}(\mathrm{tr}\tau A_k - \mathrm{tr}\rho A_k)^2\right)^{1/2} \leq$$

$$2\left(\sum_{j=1}^{N}\left(\mathrm{tr}(\rho^{1/2}A_j)^2 + \mathrm{tr}(\omega^{1/2}A_j)^2 - 2(\mathrm{tr}\omega A_j)(\mathrm{tr}\rho A_j)\right)\right)^{1/2} \times$$

$$\times\left(\sum_{k=1}^{N}\left(\mathrm{tr}(\tau^{1/2}A_k)^2 + \mathrm{tr}(\rho^{1/2}A_k)^2 - 2(\mathrm{tr}\rho A_k)(\mathrm{tr}\tau A_k)\right)\right)^{1/2} =$$

$$= 2d_C(\rho,\omega)d_C(\tau,\rho)$$

where we used again the HS-Cauchy-Schwartz $\mathrm{tr}(X^{1/2}A_j)^2 \geq (\mathrm{tr}XA_j)^2$ for $X = \rho, \tau, \omega$

A Wigner-type result for qubits

Definition (Quantum Wasserstein isometry)

A map $\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is called a *quantum Wasserstein isometry* with respect to the cost operator C if $D_C(\Phi(\rho), \Phi(\omega)) = D_C(\rho, \omega)$ for all $\rho, \omega \in \mathcal{S}(\mathcal{H})$.

- the Pauli operators are given by

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

- we consider the cost operator which is symmetric in the sense that it involves all the Pauli operators

$$C_{sym} := \sum_{j=1}^{3} \left(\sigma_j \otimes I_{\mathcal{H}^*} - I_{\mathcal{H}} \otimes \sigma_j^T \right)^2 = \begin{bmatrix} 4 & 0 & 0 & -4 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ -4 & 0 & 0 & 4 \end{bmatrix}$$

A Wigner-type result for qubits

Theorem (Gehér-Pitrik-Titkos-Virosztek, J. Math. Anal. Appl. (2023))

Let $\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ be a quantum Wasserstein isometry with respect to the cost operator C_{sym}. That is, assume that

$$D_{sym}\left(\Phi\left(\rho\right), \Phi\left(\omega\right)\right) = D_{sym}\left(\rho, \omega\right) \qquad \left(\rho, \omega \in \mathcal{S}\left(\mathcal{H}\right)\right).$$

Then there exist a unitary or anti-unitary operator U acting on $\mathcal{H} = \mathbb{C}^2$ such that

$$\Phi\left(\rho\right) = U\rho U^* \qquad \left(\rho \in \mathcal{S}\left(\mathcal{H}\right)\right).$$

Conversely, any map of the above form is a quantum Wasserstein isometry with respect to C_{sym}.
In other words, the isometry group of the quantum Wasserstein space defined by the cost operator C_{sym} coincides with the orthogonal group O(3) *by the Bloch representation.*

Isometries for the clock and shift cost

- we turn to the case when the cost operator involves the qubit "clock" and "shift" operators
- these are intimately related to the finite dimensional approximations of the position and momentum operators in quantum mechanics
- as the qubit "clock" operator is σ_3 and the "shift" is σ_1, let us define the corresponding cost operator C_{xz} by

$$C_{xz} := \sum_{j=1,3} \left(\sigma_j \otimes I_{\mathcal{H}^*} - I_{\mathcal{H}} \otimes \sigma_j^T \right)^2 =$$

$$= 4I - 2 \sum_{j=1,3} \sigma_j \otimes \sigma_j^T = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & 6 & -2 & 0 \\ 0 & -2 & 6 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$$

Isometries for the clock and shift cost

Theorem (Gehér-Pitrik-Titkos-Virosztek, J. Math. Anal. Appl. (2023))

If $\psi \in \mathcal{G} \times \mathcal{K} \cong \mathrm{O}(2) \times \mathcal{C}_2$ *and* $\xi \in \mathcal{F}_{\{-1,1\}}^{\left(\mathcal{P}_1(\mathcal{H}) \setminus \mathcal{P}_1^{\mathbb{R}}(\mathcal{H})\right)}$, *then the map* $\Phi = \psi \circ \xi$ *belongs to the semigroup* $\mathrm{Isom}\left(\mathcal{W}_2^{(xz)}\left(\mathcal{S}\left(\mathcal{H}\right)\right)\right)$. *On the other hand, if* $\Phi \in \mathrm{Isom}\left(\mathcal{W}_2^{(xz)}\left(\mathcal{S}\left(\mathcal{H}\right)\right)\right)$, *then there exists a unique* $\psi \in \mathcal{G} \cong \mathrm{O}(2)$ *and a unique* $\xi \in \mathcal{F}_{\{-1,1\}}^{\left(\mathcal{S}(\mathcal{H}) \setminus \mathcal{S}^{\mathbb{R}}(\mathcal{H})\right)}$ *such that* $\Phi = \psi \circ \xi$. *In other words,*

$$\left(\mathrm{O}(2) \times \mathcal{C}_2\right) \ltimes_{\varphi_1} \mathcal{F}_{\{-1,1\}}^{\left(\mathcal{P}_1(\mathcal{H}) \setminus \mathcal{P}_1^{\mathbb{R}}(\mathcal{H})\right)} \subseteq \mathrm{Isom}\left(\mathcal{W}_2^{(xz)}\left(\mathcal{S}\left(\mathcal{H}\right)\right)\right)$$

and

$$\mathrm{Isom}\left(\mathcal{W}_2^{(xz)}\left(\mathcal{S}\left(\mathcal{H}\right)\right)\right) \subseteq \mathrm{O}(2) \ltimes_{\varphi_2} \mathcal{F}_{\{-1,1\}}^{\left(\mathcal{S}(\mathcal{H}) \setminus \mathcal{S}^{\mathbb{R}}(\mathcal{H})\right)}.$$

Isometries for a single-observable transport cost

Theorem (Simon-Virosztek, Linear Algebra Appl. (2025))

Let D_z denote the quantum Wasserstein distance defined by the cost operator $C_z = \left(\sigma_z \otimes I_{\mathcal{H}^*} - I \otimes \sigma_z^T\right)$, and let $\Phi : \mathcal{S}(\mathbb{C}^2) \to \mathcal{S}(\mathbb{C}^2)$ be a map. Then the following are equivalent.

1. *The map Φ is a quantum Wasserstein isometry with respect to D_z, that is, $D_z(\Phi(\rho), \Phi(\omega)) = D_z(\rho, \omega)$ for all $\rho, \omega \in \mathcal{S}(\mathbb{C}^2)$.*
2. *The map Φ leaves the Euclidean length of the Bloch vector of a state invariant, that is, $\left|\mathrm{b}_{\Phi(\rho)}\right| = \left|\mathrm{b}_\rho\right|$ for all $\rho \in \mathcal{S}(\mathbb{C}^2)$, and either $\left(\mathrm{b}_{\Phi(\rho)}\right)_3 = (\mathrm{b}_\rho)_3$ for all $\rho \in \mathcal{S}(\mathbb{C}^2)$, or $\left(\mathrm{b}_{\Phi(\rho)}\right)_3 = -(\mathrm{b}_\rho)_3$ for all $\rho \in \mathcal{S}(\mathbb{C}^2)$.*

Future work

- metric properties of quadratic QW divergences: remove the assumption of one of the states being pure
- p-Wasserstein divergences:
  - proper definition
  - metric properties
- data processing inequality for quantum Wasserstein divergences
- isometries of quantum state spaces with respect to QW distances and divergences

Collaborators

Gergely Bunth

József Pitrik

György Gehér

Tamás Titkos

The definiteness of quantum Wasserstein divergences

- the key ingredient is the monotonicity version of Lieb's concavity theorem implying that

$$\mathrm{tr}_{\mathcal{H}}\left(\omega^{1/2}A_j\omega^{1/2}A_j\right)=\mathrm{tr}_{\mathcal{H}}\left(\Phi(\rho)^{1/2}A_j\Phi(\rho)^{1/2}A_j\right)\geq$$

$$\geq\mathrm{tr}_{\mathcal{H}}\left(\rho^{1/2}\Phi^{\dagger}(A_j)\rho^{1/2}\Phi^{\dagger}(A_j)\right)$$

- consequently,

$$\mathrm{tr}_{\mathcal{H}}\left(\left(\rho^{1/2}A_j\right)^2+\left(\omega^{1/2}A_j\right)^2-2\rho^{1/2}A_j\rho^{1/2}\Phi^{\dagger}\left(A_j\right)\right)\geq$$

$$\geq\left\|\rho^{1/4}\left(A_j-\Phi^{\dagger}(A_j)\right)\rho^{1/4}\right\|_{HS}^2$$

QOT via quantum couplings

- quantum couplings are defined as

$$\mathcal{C}(\rho, \omega) = \{\pi \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H}) \,|\, \mathrm{tr}_2 \pi = \rho,\ \mathrm{tr}_1 \pi = \omega\},$$

- cost operators

$$C = \sum_{j=1}^{M} (A_j \otimes I - I \otimes A_j)^2$$

where $A_j \in \mathcal{L}^{sa}(\mathcal{H})$.

- optimal transport cost:

$$D_C^2(\rho, \omega) = \inf_{\pi \in \mathcal{C}(\rho, \omega)} \mathrm{tr}\, \pi C$$

QOT via quantum couplings

- note that even if $\mathcal{H} = L^2(\mathbb{R})$, $M = 1$, and $A_1 = Q$ (position op.), the quantum optimal transport problem is *essentially different* from the classical one: for pure states $|\varphi\rangle\langle\varphi|$ and $|\zeta\rangle\langle\zeta|$ we have

$$D_Q^2\left(|\varphi\rangle\langle\varphi|, |\zeta\rangle\langle\zeta|\right) = \int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 |\varphi(x)|^2 |\zeta(y)|^2 \,\mathrm{d}y\mathrm{d}x$$

which is the quadratic transport cost of the classical probability measures $\widetilde{\rho}(\mathrm{d}x) = |\varphi(x)|^2 \mathrm{d}x$ and $\widetilde{\omega}(\mathrm{d}y) = |\zeta(y)|^2 \mathrm{d}y$ with the *trivial transport plan* $\widetilde{\rho} \otimes \widetilde{\omega}$ which is typically far from being optimal