Order p quantum Wasserstein distances from couplings

Emily Beatty, Daniel Stilck França

arXiv:2402.16477

ENS de Lyon // INRIA

Porquerolles, 10 June 2024

Outline

Classical optimal transport context

Classically, optimal transport costs describe the minimum cost needed to transport mass from one place to another.
Given:

- Measurable space $\mathcal{X}$
- Cost function $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$
- Probability measures μ, ν on $\mathcal{X}$

we define a coupling as a measure π on $\mathcal{X} \times \mathcal{X}$ whose marginals are μ and ν respectively, so for projections P_x and P_y we have $(P_x)_\star \pi = \mu$, $(P_y)_\star \pi = \nu$. Then for points x and y, $\pi(\mathrm{d}x, \mathrm{d}y)$ is interpreted as the amount of mass transported from point x to point y.

Classical optimal transport context

Classically, optimal transport costs describe the minimum cost needed to transport mass from one place to another.
Given:

- Measurable space $\mathcal{X}$
- Cost function $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$
- Probability measures μ, ν on $\mathcal{X}$

we define a coupling as a measure π on $\mathcal{X} \times \mathcal{X}$ whose marginals are μ and ν respectively, so for projections P_x and P_y we have $(P_x)_\star\pi = \mu$, $(P_y)_\star\pi = \nu$. Then for points x and y, $\pi(\mathrm{d}x, \mathrm{d}y)$ is interpreted as the amount of mass transported from point x to point y.
The transport cost of coupling π is then

$$\int_{\mathcal{X}\times\mathcal{X}} c(x, y)\pi(\mathrm{d}x, \mathrm{d}y). \tag{1}$$

Classical optimal transport context

Classically, optimal transport costs describe the minimum cost needed to transport mass from one place to another.
Given:

- Measurable space $\mathcal{X}$
- Cost function $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$
- Probability measures μ, ν on $\mathcal{X}$

we define a coupling as a measure π on $\mathcal{X} \times \mathcal{X}$ whose marginals are μ and ν respectively, so for projections P_x and P_y we have $(P_x)_\star \pi = \mu$, $(P_y)_\star \pi = \nu$. Then for points x and y, $\pi(\mathrm{d}x, \mathrm{d}y)$ is interpreted as the amount of mass transported from point x to point y.
The transport cost of coupling π is then

$$\int_{\mathcal{X} \times \mathcal{X}} c(x, y) \pi(\mathrm{d}x, \mathrm{d}y). \tag{1}$$

This gives the optimal transport cost,

$$\inf_{\pi} \int_{\mathcal{X} \times \mathcal{X}} c(x, y) \pi(\mathrm{d}x, \mathrm{d}y) \tag{2}$$

where the infimum is over couplings π of μ and ν.

Classical optimal transport context

When $\mathcal{X}$ is equipped with metric d, we can take $c(x, y) = d(x, y)^p$ and recover the classical Wasserstein distances:

Definition (Classical Wasserstein distance)

The Wasserstein distance $\mathcal{W}_p$ of order p between μ, ν on metric space $(\mathcal{X}, d)$ is given by

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\pi} \int_{\mathcal{X} \times \mathcal{X}} d(x, y)^p \pi(\mathrm{d}x, \mathrm{d}y) \right)^{1/p}. \tag{3}$$

Motivation

Many generalisations to the quantum setting have been proposed, all of which either do not satisfy basic properties, or only exist for specific circumstances.

Motivation

Many generalisations to the quantum setting have been proposed, all of which either do not satisfy basic properties, or only exist for specific circumstances. We note the property of the classical distances: that for all orders p, the Wasserstein distance between point masses agrees with the underlying distance:

$$\forall p, \mathcal{W}_p(\delta_x, \delta_y) = d(x, y). \tag{4}$$

Motivation

Many generalisations to the quantum setting have been proposed, all of which either do not satisfy basic properties, or only exist for specific circumstances. We note the property of the classical distances: that for all orders p, the Wasserstein distance between point masses agrees with the underlying distance:

$$\forall p, \mathcal{W}_p(\delta_x, \delta_y) = d(x, y). \tag{4}$$

Generalising this, given a Hilbert space $\mathcal{H}$ with underlying metric d on the space of pure states $\mathbb{P}\mathcal{H}$, we would expect any general definition of a quantum Wasserstein distance W_p^d to satisfy

$$W_p^d(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = d(|\psi\rangle, |\varphi\rangle) \tag{5}$$

for pure states $|\psi\rangle$, $|\varphi\rangle$.

Definitions

Definition (Quantum transport plan)

A quantum transport plan between states ρ, σ is a finite set of triples $Q = \{(q_j, |\psi_j\rangle, |\varphi_j\rangle)\}_j$ such that

$$\sum_j q_j |\psi_j\rangle\langle\psi_j| = \rho, \qquad \sum_j q_j |\varphi_j\rangle\langle\varphi_j| = \sigma. \tag{6}$$

The plan Q corresponds to separable coupling $\tau = \sum_j q_j |\psi_j\rangle\langle\psi_j| \otimes |\varphi_j\rangle\langle\varphi_j|$.

Definitions

Definition (Quantum transport plan)

A quantum transport plan between states ρ, σ is a finite set of triples $Q = \{(q_j, |\psi_j\rangle, |\varphi_j\rangle)\}_j$ such that

$$\sum_j q_j |\psi_j\rangle\langle\psi_j| = \rho, \qquad \sum_j q_j |\varphi_j\rangle\langle\varphi_j| = \sigma. \tag{6}$$

The plan Q corresponds to separable coupling $\tau = \sum_j q_j |\psi_j\rangle\langle\psi_j| \otimes |\varphi_j\rangle\langle\varphi_j|$.

Definition (Quantum transport cost)

For a quantum transport plan $Q = \{(q_j, |\psi_j\rangle, |\varphi_j\rangle)\}_j$ between ρ and σ, the transport cost of Q is

$$T_p^d(Q) = \sum_j q_j d(|\psi_j\rangle, |\varphi_j\rangle)^p. \tag{7}$$

Definitions

Optimising over Q allows us to define the quantum Wasserstein distance of order p with respect to d.

Definition (Quantum Wasserstein distance)

The p^{th}-order quantum Wasserstein distance between ρ and σ is defined as

$$W_p^d(\rho, \sigma) = \left(\inf_Q T_p^d(Q) \right)^{1/p} = \left(\inf_Q \sum_j q_j d(|\psi_j\rangle, |\varphi_j\rangle)^p \right)^{1/p} \tag{8}$$

where the infimum is over all quantum transport plans Q between ρ and σ.

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

Properties

1. **Faithfulness:** $W_p^d(\rho, \sigma) \geq 0$ with equality iff $\rho = \sigma$.

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

Properties

1. **Faithfulness:** $W_p^d(\rho, \sigma) \geq 0$ with equality iff $\rho = \sigma$.
2. **Agreement with** d**:** for pure states $|\psi\rangle, |\varphi\rangle$, we have $W_p^d(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = d(|\psi\rangle, |\varphi\rangle)$.

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

Properties

1. **Faithfulness:** $W_p^d(\rho, \sigma) \geq 0$ with equality iff $\rho = \sigma$.
2. **Agreement with d:** for pure states $|\psi\rangle, |\varphi\rangle$, we have $W_p^d(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = d(|\psi\rangle, |\varphi\rangle)$.
3. **Data processing for mixed unitary channels:** for unitaries U_i which are symmetries of d, the channel $T(\cdot) = \sum_i p_i U_i \cdot U_i^\dagger$ satisfies data processing.

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

Properties

1. **Faithfulness:** $W_p^d(\rho,\sigma) \geq 0$ with equality iff $\rho = \sigma$.
2. **Agreement with d:** for pure states $|\psi\rangle, |\varphi\rangle$, we have $W_p^d(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = d(|\psi\rangle, |\varphi\rangle)$.
3. **Data processing for mixed unitary channels:** for unitaries U_i which are symmetries of d, the channel $T(\cdot) = \sum_i p_i U_i \cdot U_i^\dagger$ satisfies data processing.
4. **Hierarchy in p:** For $p_1 < p_2$, we have $W_{p_1}^d \leq W_{p_2}^d$.

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

Properties

1. **Faithfulness:** $W_p^d(\rho, \sigma) \geq 0$ with equality iff $\rho = \sigma$.
2. **Agreement with d:** for pure states $|\psi\rangle, |\varphi\rangle$, we have $W_p^d(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = d(|\psi\rangle, |\varphi\rangle)$.
3. **Data processing for mixed unitary channels:** for unitaries U_i which are symmetries of d, the channel $T(\cdot) = \sum_i p_i U_i \cdot U_i^\dagger$ satisfies data processing.
4. **Hierarchy in p:** For $p_1 < p_2$, we have $W_{p_1}^d \leq W_{p_2}^d$.
5. **Continuity:** W_p^d is uniformly continuous.

General properties

Subject to gentle continuity conditions on d, the W_p^d distances satisfy the following properties:

Properties

1. **Faithfulness:** $W_p^d(\rho, \sigma) \geq 0$ with equality iff $\rho = \sigma$.
2. **Agreement with d:** for pure states $|\psi\rangle, |\varphi\rangle$, we have $W_p^d(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = d(|\psi\rangle, |\varphi\rangle)$.
3. **Data processing for mixed unitary channels:** for unitaries U_i which are symmetries of d, the channel $T(\cdot) = \sum_i p_i U_i \cdot U_i^\dagger$ satisfies data processing.
4. **Hierarchy in p:** For $p_1 < p_2$, we have $W_{p_1}^d \leq W_{p_2}^d$.
5. **Continuity:** W_p^d is uniformly continuous.
6. **Infimum achieved at polynomial sizes:** For Hilbert space $\mathcal{H}$ of dimension D, the infimum in the definition of W_p^d is achieved for a transport plan Q with at most $2D^2$ elements.

Dual setting

We can dualise this quantity as follows:

$$L_d(O) = \sup_{\rho \neq \sigma} \frac{\mathrm{Tr}[O(\rho - \sigma)]}{W_1^d(\rho, \sigma)} \tag{9}$$

which is a norm.

Dual setting

We can dualise this quantity as follows:

$$L_d(O) = \sup_{\rho \neq \sigma} \frac{\mathrm{Tr}[O(\rho - \sigma)]}{W_1^d(\rho, \sigma)} \tag{9}$$

which is a norm. In an analogy with the Kantorovich-Rubenstein theorem, this turns out to be the Lipschitz constant of the operator O on $\mathbb{P}\mathcal{H}$:

$$L_d(O) = \sup_{\psi \neq \varphi} \frac{\mathrm{Tr}[O(|\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|)]}{d(|\psi\rangle, |\varphi\rangle)}. \tag{10}$$

Dual setting

We can dualise this quantity as follows:

$$L_d(O) = \sup_{\rho \neq \sigma} \frac{\mathrm{Tr}[O(\rho - \sigma)]}{W_1^d(\rho, \sigma)} \tag{9}$$

which is a norm. In an analogy with the Kantorovich-Rubenstein theorem, this turns out to be the Lipschitz constant of the operator O on $\mathbb{P}\mathcal{H}$:

$$L_d(O) = \sup_{\psi \neq \varphi} \frac{\mathrm{Tr}[O(|\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|)]}{d(|\psi\rangle, |\varphi\rangle)}. \tag{10}$$

We can then define closely related norm

Definition (dual norm)

$$\|\rho - \sigma\|_{DW_1^d} = \sup_{L_d(O) \leq 1} \mathrm{Tr}[O(\rho - \sigma)] \tag{11}$$

which is indeed a norm.

Dual setting

We can dualise this quantity as follows:

$$L_d(O) = \sup_{\rho \neq \sigma} \frac{\mathrm{Tr}[O(\rho - \sigma)]}{W_1^d(\rho, \sigma)} \tag{9}$$

which is a norm. In an analogy with the Kantorovich-Rubenstein theorem, this turns out to be the Lipschitz constant of the operator O on $\mathbb{P}\mathcal{H}$:

$$L_d(O) = \sup_{\psi \neq \varphi} \frac{\mathrm{Tr}[O(|\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|)]}{d(|\psi\rangle, |\varphi\rangle)}. \tag{10}$$

We can then define closely related norm

Definition (dual norm)

$$\|\rho - \sigma\|_{DW_1^d} = \sup_{L_d(O) \leq 1} \mathrm{Tr}[O(\rho - \sigma)] \tag{11}$$

which is indeed a norm. We have $W_1^d(\rho, \sigma) \geq \|\rho - \sigma\|_{DW_1^d}$ and for d induced by a norm $\|\cdot\|_d$,

$$W_1^d(\rho, \sigma) \geq \|\rho - \sigma\|_{DW_1^d} \geq \|\rho - \sigma\|_d\,. \tag{12}$$

Specific examples

- ***d* from the trace distance:** $d(|\psi\rangle, |\varphi\rangle) = \frac{1}{2} \left\| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \right\|_1$. Then $\frac{1}{2} \|\rho - \sigma\|_1 = \|\rho - \sigma\|_{DW_1^1}$.

[1] G. de Palma et al., The Quantum Wasserstein Distance of Order 1, IEEE Transactions on Information Theory 67(10), Oct 2021

[2] M.A. Nielsen, A geometric approach to quantum circuit lower bounds, Quantum Information and Computation 6(3), May 2006

Specific examples

- ***d* from the trace distance:** $d(|\psi\rangle, |\varphi\rangle) = \frac{1}{2}\left\| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \right\|_1$. Then $\frac{1}{2}\|\rho - \sigma\|_1 = \|\rho - \sigma\|_{DW_1^1}$.
- ***d* from the W_1^H norm[1]:** $d(|\psi\rangle, |\varphi\rangle) = \frac{1}{2}\left\| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \right\|_{W_1^H}$. Notated as W_p^H. Then $2\|\rho - \sigma\|_{W_1^H} \geq \|\rho - \sigma\|_{DW_1^H} \geq \|\rho - \sigma\|_{W_1^H}$.

[1] G. de Palma et al., The Quantum Wasserstein Distance of Order 1, IEEE Transactions on Information Theory 67(10), Oct 2021

[2] M.A. Nielsen, A geometric approach to quantum circuit lower bounds, Quantum Information and Computation 6(3), May 2006

Specific examples

- d **from the trace distance:** $d(|\psi\rangle, |\varphi\rangle) = \frac{1}{2} \| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \|_1$. Then $\frac{1}{2}\|\rho - \sigma\|_1 = \|\rho - \sigma\|_{DW_1^1}$.
- d **from the W_1^H norm**[1]**:** $d(|\psi\rangle, |\varphi\rangle) = \frac{1}{2} \| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \|_{W_1^H}$. Notated as W_p^H. Then $2\|\rho - \sigma\|_{W_1^H} \geq \|\rho - \sigma\|_{DW_1^H} \geq \|\rho - \sigma\|_{W_1^H}$.
- d **from the Nielsen complexity geometry**[2]**:** notated as W_p^C.

[1] G. de Palma et al., The Quantum Wasserstein Distance of Order 1, IEEE Transactions on Information Theory 67(10), Oct 2021

[2] M.A. Nielsen, A geometric approach to quantum circuit lower bounds, Quantum Information and Computation 6(3), May 2006

Values on random states

We use model $\rho = \mathrm{Tr}_{\mathcal{A}}[|\Phi\rangle\langle\Phi|]$ where $|\Phi\rangle \sim \mu_{\mathrm{Haar}}$ on $\mathcal{H} \otimes \mathcal{A}$. Let $\dim \mathcal{H} = d^n$, $\log_d(\dim \mathcal{A}) = m$ and $c = m/n$.

Values on random states

We use model $\rho = \mathrm{Tr}_{\mathcal{A}}[|\Phi\rangle\langle\Phi|]$ where $|\Phi\rangle \sim \mu_{\mathrm{Haar}}$ on $\mathcal{H} \otimes \mathcal{A}$. Let $\dim \mathcal{H} = d^n$, $\log_d(\dim \mathcal{A}) = m$ and $c = m/n$.

Random states with large auxiliary systems

$$\mathbb{P}\left[W_1^d(\rho, \sigma) \geq \beta d^{-(c-3)n/2} \mathrm{diam}_d(\mathbb{P}\mathcal{H})\right] \leq \frac{1}{\beta^2}. \tag{13}$$

For $W_{p=1}^H$, this is significant when $c > 3$, and W_1^C when $c > 9$.

Values on random states

We use model $\rho = \mathrm{Tr}_{\mathcal{A}}[|\Phi\rangle\langle\Phi|]$ where $|\Phi\rangle \sim \mu_{\mathrm{Haar}}$ on $\mathcal{H} \otimes \mathcal{A}$. Let $\dim \mathcal{H} = d^n$, $\log_d(\dim \mathcal{A}) = m$ and $c = m/n$.

Random states with large auxiliary systems

$$\mathbb{P}\left[W_1^d(\rho, \sigma) \geq \beta d^{-(c-3)n/2} \mathrm{diam}_d(\mathbb{P}\mathcal{H})\right] \leq \frac{1}{\beta^2}. \tag{13}$$

For $W_{p=1}^H$, this is significant when $c > 3$, and W_1^C when $c > 9$.

Random states with small auxiliary systems

$$\text{When } c < 1, \qquad \mathbb{E}_{\rho,\sigma}\left[W_{p=1}^H(\rho, \sigma)\right] \geq \lambda_c n \tag{14}$$

where λ_c satisfies $(1-c)\log d = h_2(\lambda) + \lambda \log(d^2 - 1)$ for h_2 the binary entropy.

Values on random states

We use model $\rho = \mathrm{Tr}_{\mathcal{A}}[|\Phi\rangle\langle\Phi|]$ where $|\Phi\rangle \sim \mu_{\text{Haar}}$ on $\mathcal{H} \otimes \mathcal{A}$. Let $\dim \mathcal{H} = d^n$, $\log_d(\dim \mathcal{A}) = m$ and $c = m/n$.

Random states with large auxiliary systems

$$\mathbb{P}\left[W_1^d(\rho, \sigma) \geq \beta d^{-(c-3)n/2} \mathrm{diam}_d(\mathbb{P}\mathcal{H})\right] \leq \frac{1}{\beta^2}. \tag{13}$$

For $W_{p=1}^H$, this is significant when $c > 3$, and W_1^C when $c > 9$.

Random states with small auxiliary systems

$$\text{When } c < 1, \qquad \mathbb{E}_{\rho,\sigma}\left[W_{p=1}^H(\rho, \sigma)\right] \geq \lambda_c n \tag{14}$$

where λ_c satisfies $(1-c)\log d = h_2(\lambda) + \lambda \log(d^2 - 1)$ for h_2 the binary entropy. When $c = 0$,

$$\mathbb{P}_{|\varphi\rangle}\left[d_C(|\psi\rangle, |\varphi\rangle) \leq \epsilon^{2/3} n^{-1} \kappa \left(\frac{2^{(1-\delta)n}}{poly(n, \log \epsilon^{-1})}\right)^{1/3}\right] \leq e^{-\Omega(2^n \log((2\epsilon)^{-1}))}. \tag{15}$$

Moments of classical-quantum (c-q) sources

$W_p^d(\rho, \sigma)$ can be expressed as a sharp lower bound for the p^{th} moment between the output of c-q sources.

Moments of classical-quantum (c-q) sources

$W_p^d(\rho, \sigma)$ can be expressed as a sharp lower bound for the p^{th} moment between the output of c-q sources.

- Let R and S be controlled by X, $\mathbb{P}(X = i) = q_i$.

Moments of classical-quantum (c-q) sources

$W_p^d(\rho, \sigma)$ can be expressed as a sharp lower bound for the p^{th} moment between the output of c-q sources.

- Let R and S be controlled by X, $\mathbb{P}(X = i) = q_i$.
- On input i let R output $|\psi_i\rangle$ and S output $|\varphi_i\rangle$.

Moments of classical-quantum (c-q) sources

$W_p^d(\rho, \sigma)$ can be expressed as a sharp lower bound for the p^{th} moment between the output of c-q sources.

- Let R and S be controlled by X, $\mathbb{P}(X = i) = q_i$.
- On input i let R output $|\psi_i\rangle$ and S output $|\varphi_i\rangle$.
- We define the expected outputs ρ, σ of R and S as $\rho = \sum q_i |\psi_i\rangle\langle\psi_i|$, $\sigma = \sum q_i |\varphi_i\rangle\langle\varphi_i|$ respectively.

Moments of c-q sources

$$\mathbb{E}[d(R, S)^p] \geq W_p^d(\rho, \sigma)^p. \tag{16}$$

Moments of classical-quantum (c-q) sources

$W_p^d(\rho, \sigma)$ can be expressed as a sharp lower bound for the p^{th} moment between the output of c-q sources.

- Let R and S be controlled by X, $\mathbb{P}(X = i) = q_i$.
- On input i let R output $|\psi_i\rangle$ and S output $|\varphi_i\rangle$.
- We define the expected outputs ρ, σ of R and S as $\rho = \sum q_i |\psi_i\rangle\langle\psi_i|$, $\sigma = \sum q_i |\varphi_i\rangle\langle\varphi_i|$ respectively.

Moments of c-q sources

$$\mathbb{E}[d(R,S)^p] \geq W_p^d(\rho, \sigma)^p. \tag{16}$$

- This bound is sharp in the sense that there exist sources R', S' with expected outputs ρ, σ respectively such that $\mathbb{E}[d(R', S')^p] = W_p^d(\rho, \sigma)^p$.

For the W_p^C distance in particular, this means that the p^{th} moment of the complexity of converting one c-q source into another post-output is sharply lower bounded by their W_p^C distance.

Hypercontractivity

Hypercontractivity[3] can be used in the classical setting to quantify the noise of an operation.

[3] A. Bonami, Étude des coefficients de Fourier des fonctions de $L^p(G)$, 1970.

Hypercontractivity

Hypercontractivity[3] can be used in the classical setting to quantify the noise of an operation. Indeed, for $\alpha \in [0,1]$ let the standard Boolean noise operator T_α on functions $f : \{-1,1\}^n \to \{-1,1\}$ is defined as follows

$$(T_\alpha f)(x) = \mathbb{E}_y[f(y)] \tag{17}$$

where bit y_i is x_i with probability $1/2 + \alpha/2$ and $-x_i$ otherwise.
Then we have

Classical hypercontractivity theorem

Given $\alpha \leq \sqrt{\frac{p-1}{q-1}}$, we have

$$\|T_\alpha f\|_q \leq \|f\|_p \,. \tag{18}$$

[3] A. Bonami, Étude des coefficients de Fourier des fonctions de $L^p(G)$, 1970.

Hypercontractivity

Hypercontractivity[3] can be used in the classical setting to quantify the noise of an operation. Indeed, for $\alpha \in [0,1]$ let the standard Boolean noise operator T_α on functions $f : \{-1,1\}^n \to \{-1,1\}$ is defined as follows

$$(T_\alpha f)(x) = \mathbb{E}_y[f(y)] \tag{17}$$

where bit y_i is x_i with probability $1/2 + \alpha/2$ and $-x_i$ otherwise.
Then we have

Classical hypercontractivity theorem

Given $\alpha \leq \sqrt{\frac{p-1}{q-1}}$, we have

$$\|T_\alpha f\|_q \leq \|f\|_p . \tag{18}$$

In this way, for a general operator U the ratio $\frac{\|f\|_p}{\|U_\alpha f\|_q}$ can quantify the noise in f.

[3] A. Bonami, Étude des coefficients de Fourier des fonctions de $L^p(G)$, 1970.

Hypercontractivity

In the hierarchy $p_1 < p_2$, we know that $\frac{W_{p_1}(\rho,\sigma)}{W_{p_2}(\rho,\sigma)} < 1$. We have the standard noise channels $\mathcal{N} = \mathcal{R}_{x,\delta}, \mathcal{S}_\delta$

$$\mathcal{R}_{x,\delta}(\rho) = (1-\delta)\rho + \delta|x\rangle\langle x| \qquad \mathcal{S}_\delta(\rho) = (1-\delta)\rho + \delta\mathbb{I}/D. \tag{19}$$

Hypercontractivity

In the hierarchy $p_1 < p_2$, we know that $\frac{W_{p_1}(\rho,\sigma)}{W_{p_2}(\rho,\sigma)} < 1$. We have the standard noise channels $\mathcal{N} = \mathcal{R}_{x,\delta}, \mathcal{S}_\delta$

$$\mathcal{R}_{x,\delta}(\rho) = (1-\delta)\rho + \delta|x\rangle\langle x| \qquad \mathcal{S}_\delta(\rho) = (1-\delta)\rho + \delta\mathbb{I}/D. \tag{19}$$

Hypercontractivity in W_p^d

Let $1 \leq p_1 < p_2$, and suppose $W_{p_1}^d(\rho,\sigma) = M$. Then for $1-\delta \leq (M/diam_d(\mathbb{P}\mathcal{H}))^{p_2-p_1}$, the channels $\mathcal{N} = \mathcal{R}_{x,\delta}$ and $\mathcal{N} = \mathcal{S}_x$ have

$$\frac{W_{p_1}(\rho,\sigma)}{W_{p_2}(\mathcal{N}(\rho),\mathcal{N}(\sigma))} > 1. \tag{20}$$

Hypercontractivity

In the hierarchy $p_1 < p_2$, we know that $\frac{W_{p_1}(\rho,\sigma)}{W_{p_2}(\rho,\sigma)} < 1$. We have the standard noise channels $\mathcal{N} = \mathcal{R}_{x,\delta}, \mathcal{S}_\delta$

$$\mathcal{R}_{x,\delta}(\rho) = (1-\delta)\rho + \delta|x\rangle\langle x| \qquad \mathcal{S}_\delta(\rho) = (1-\delta)\rho + \delta\mathbb{I}/D. \tag{19}$$

Hypercontractivity in W_p^d

Let $1 \leq p_1 < p_2$, and suppose $W_{p_1}^d(\rho,\sigma) = M$. Then for $1-\delta \leq (M/diam_d(\mathbb{P}\mathcal{H}))^{p_2-p_1}$, the channels $\mathcal{N} = \mathcal{R}_{x,\delta}$ and $\mathcal{N} = \mathcal{S}_x$ have

$$\frac{W_{p_1}(\rho,\sigma)}{W_{p_2}(\mathcal{N}(\rho),\mathcal{N}(\sigma))} > 1. \tag{20}$$

For general $\mathcal{N}$, this suggests that the ratio above can be interpreted as quantifying the amount of noise in the channel.

Conclusion

- We propose a novel definition of the quantum Wasserstein distance for any underlying metric d on $\mathbb{P}\mathcal{H}$ and any order p, which (so far as we are aware) is the first to exhibit such flexibility.
- We show it has many desirable properties of a good quantum Wasserstein distance.
- We exhibit some applications to c-q sources and the noise of channels, which are only possible thanks to the full flexibility of the definition in both the order p and the distance d.
- Further avenues: triangle inequality, Markov chains, other underlying metrics d, ...

Further avenues

Further avenues

- What does this look like for **other underlying metrics** d on $\mathbb{P}\mathcal{H}$?

Further avenues

- What does this look like for **other underlying metrics** d on $\mathbb{P}\mathcal{H}$?
- How can we approach the **triangle inequality**?

Further avenues

- What does this look like for **other underlying metrics** d on $\mathbb{P}\mathcal{H}$?
- How can we approach the **triangle inequality**?
- How can we recover **other applications** of classical Wasserstein distances in the quantum setting (Markov chains, concentration inequalities etc.)?

Further avenues

- What does this look like for **other underlying metrics** d on $\mathbb{P}\mathcal{H}$?
- How can we approach the **triangle inequality**?
- How can we recover **other applications** of classical Wasserstein distances in the quantum setting (Markov chains, concentration inequalities etc.)?
- How can we **lower bound** this distance W_p?

Thanks for your attention :)