

# Rényi divergences in quantum information theory

Anna Jenčová

Mathematical Institute, Slovak Academy of Sciences  
Bratislava

Joint work with Fumio Hiai

Smolenice, 15. 10. 2024



# What is a divergence?

- A "dissimilarity measure" on probability distributions:

For probability distributions  $p, q$

$$D(p\|q) \equiv \text{how different } p \text{ is from } q.$$

- A **contrast functional**:

$$D(p\|q) \geq 0, \quad D(p\|q) = 0 \iff p = q.$$

- Not a metric (not necessarily symmetric)
- Other properties?

## Axiomatic approach (A. Rényi, 1961)

Let  $p = (p_1, \dots, p_m)$ ,  $q = (q_1, \dots, q_m)$ ,  $p_i \geq 0$ ,  $q_i > 0$

A divergence  $D$  should satisfy the **postulates**:

- **invariance** under permutations:  $D(\pi(p) \parallel \pi(q)) = D(p \parallel q)$
- **continuity**
- **additivity**:  $D(p_1 \otimes p_2 \parallel q_1 \otimes q_2) = D(p_1 \parallel q_1) + D(p_2 \parallel q_2)$
- **generalized mean**: for a continuous, strictly increasing real function  $g$

$$D(p_1 \oplus p_2 \parallel q_1 \oplus q_2) = g^{-1} \left( \frac{g(D(p_1 \parallel q_1)) + g(D(p_2 \parallel q_2))}{2} \right)$$

- **order relations**:

$$p_i \leq q_i, \forall i \implies D(p \parallel q) \geq 0, \quad p_i \geq q_i, \forall i \implies D(p \parallel q) \leq 0$$

- **normalization**:  $D(\{1\} \parallel \{1/2\}) = 1$

# Rényi divergences

There is a **unique** family of divergences  $\{D_\alpha\}_{\alpha>0}$ , satisfying the Rényi postulates:

$$D_\alpha(p\|q) = \frac{1}{\alpha - 1} \log \left( \sum_k p_k^\alpha q_k^{1-\alpha} \right), \quad 1 \neq \alpha > 0$$

$$D_1(p\|q) = \lim_{\alpha \rightarrow 1} D_\alpha(p\|q) = \sum_k p_k \log \left( \frac{p_k}{q_k} \right)$$

- Fundamental quantities in information theory
- For  $\alpha = 1$ , we get the **Kullback-Leibler divergence** (relative entropy,  $I$ -divergence)

# Example: Asymptotic hypothesis testing

Testing simple hypothesis  $H_0 = p$  against simple alternative  $H_1 = q$ :

- A test statistic:  $T$
- Two kinds of error probabilities:

$\alpha(T)$  - rejecting true       $\beta(T)$  - accepting false

- Cannot minimize both errors simultaneously
- i.i.d. repetitions: a sequence of tests  $\{T_n\}$
- We can obtain  $\alpha(T_n) \rightarrow 0$ ,  $\beta(T_n) \rightarrow 0$  exponentially
- Rate of the convergence?

# Stein's lemma and error exponents

## Direct domain

$$\exists \{T_n\}: \beta(T_n) \approx e^{-Rn}$$

and

$$\alpha(T_n) \rightarrow 0$$

optimal exponent (largest):

$$\alpha(T_n) \approx e^{-R'n}$$

## Converse domain

$$\forall \{T_n\}, \beta(T_n) \approx e^{-Rn}$$

$\Downarrow$

$$\alpha(T_n) \rightarrow 1$$

optimal exponent (smallest):

$$\alpha(T_n) \approx 1 - e^{-R'n}$$

$$D_1(p||q)$$

Trade-off between  $R$  and  $R'$ :

- Direct domain -  $D_\alpha(p||q)$ ,  $\alpha \in (0, 1)$
- Converse domain -  $D_\alpha(p||q)$ ,  $\alpha > 1$

# A basic property: DPI and sufficient statistics

**Data processing inequality:** For a transformation

$T : \{1, \dots, m\} \rightarrow \{1, \dots, k\}$ , with  $p^T, q^T$  induced distributions

$$D_\alpha(p^T \| q^T) \leq D_\alpha(p \| q)$$

- Any reasonable divergence should satisfy DPI!

**Kullback-Leibler-Csiszár Theorem:** If  $\text{supp}(p) \subseteq \text{supp}(q)$ ,  $\alpha > 0$

$D_\alpha(p^T \| q^T) = D_\alpha(p \| q) \iff T$  is a **sufficient statistic** for  $\{p, q\}$  :

- conditional expectations  $E_p[\cdot | T] = E_q[\cdot | T]$
- $T$  contains all information needed to distinguish  $p$  from  $q$ .

# Quantum divergences

Quantum information theory:

- quantum states instead of probability measures
- simplest case: density matrices

$$\rho \in M_n(\mathbb{C}), \rho \geq 0, \operatorname{Tr}[\rho] = 1$$

- general case: normal states of a von Neumann algebra
  - covers most of interesting situations
  - powerful technical tools

Quantum divergences: dissimilarity measures for quantum states



# Postulates for quantum divergences?

- Postulates similar to Rényi (Müller-Lennert et al, 2013)
- In the **classical case** (commuting density matrices) we get the unique family of Rényi divergences  $\{D_\alpha\}_{\alpha>0}$
- In quantum case: **no unique solution**

# Quantum DPI

Quantum channel: a linear map  $\Phi : M_n(\mathbb{C}) \rightarrow M_m(\mathbb{C})$

- **completely positive**:  $\text{id}_k : M_k(\mathbb{C}) \rightarrow M_k(\mathbb{C})$  identity map

$\Phi \otimes \text{id}_k$  is positive for any  $k \geq 1$

- **trace-preserving**:  $\text{Tr} [\Phi(\rho)] = \text{Tr} [\rho]$

Equivalently:  $\Phi \otimes \text{id}_k$  maps states to states, for all  $k$ .

**Data processing inequality** for quantum divergences:

$$D(\Phi(\rho) \parallel \Phi(\sigma)) \leq D(\rho \parallel \sigma)$$

for any quantum channel  $\Phi$  and any pair of states  $\rho, \sigma$ .

# An important quantum divergence

Quantum relative entropy (Umegaki, 1962)

$$S(\rho||\sigma) = \text{Tr} [\rho (\log(\rho) - \log(\sigma))]$$

- satisfies postulates, DPI (Lindblad, 1975)
- fundamental in quantum information theory
- **operational interpretations**: quantum communication, quantum Stein's lemma (Petz & Hiai, 1991)
- related to many important quantities
- entanglement measures, uncertainty relations

# Quantum Rényi divergences

Petz-type (standard) quantum Rényi divergence: (Petz, 1985,1986)

$$D_{\alpha}(\rho\|\sigma) = \frac{1}{\alpha - 1} \log \text{Tr} [\rho^{\alpha} \sigma^{1-\alpha}] , \quad 1 \neq \alpha > 0$$

- satisfies postulates, DPI for  $\alpha \in (0, 2]$
- $\lim_{\alpha \rightarrow 1} D_{\alpha}(\rho\|\sigma) = S(\rho\|\sigma)$
- **operational interpretation** for  $\alpha \in (0, 1)$ : (Audenaert et al., 2008, Nagaoka, 2006)  
asymptotic hypothesis testing: error exponents, direct part

# Quantum Rényi divergences

Minimal (sandwiched) quantum Rényi divergence: (Müller-Lennert et al, 2013, Wilde et al, 2014)

$$\tilde{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha - 1} \log \text{Tr} \left[ \left( \sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right], \quad 1 \neq \alpha > 0$$

- satisfies postulates, DPI for  $\alpha \in [1/2, \infty)$  (Frank & Lieb, 2013)
- $\lim_{\alpha \rightarrow 1} \tilde{D}_\alpha(\rho\|\sigma) = S(\rho\|\sigma)$
- **operational interpretation** for  $\alpha > 1$ : (Mosonyi & Ogawa, 2015)  
asymptotic hypothesis testing: error exponents, converse part

# Quantum Rényi divergences

$\alpha - z$ -Rényi divergence: (Jaksic et al, 2011, Audenaert & Datta, 2015)

$$D_{\alpha,z}(\rho\|\sigma) = \frac{1}{\alpha - 1} \log \text{Tr} \left[ \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^z \right], \quad 1 \neq \alpha > 0, z > 0$$

- satisfies postulates, DPI if: (Zhang, 2020)

$$- \alpha \in (0, 1), \max\{\alpha, 1 - \alpha\} \leq z$$

$$- \alpha > 1, \max\{\frac{\alpha}{2}, \alpha - 1\} \leq z \leq \alpha$$

- $\lim_{\alpha \rightarrow 1} D_{\alpha,z}(\rho\|\sigma) = S(\rho\|\sigma), z > 1$
- Petz type:  $D_{\alpha,1}(\rho\|\sigma) = D_{\alpha}(\rho\|\sigma)$
- Minimal:  $D_{\alpha,\alpha}(\rho\|\sigma) = \tilde{D}_{\alpha}(\rho\|\sigma)$

# Quantum Rényi divergences

Maximal Rényi divergence: (Matsumoto, 2018)

$$D_{\alpha}^{max}(\rho\|\sigma) = \frac{1}{\alpha - 1} \log \text{Tr} \left[ \sigma \left( \sigma^{-1/2} \rho \sigma^{-1/2} \right)^{\alpha} \right], \quad 1 \neq \alpha > 0$$

- satisfies postulates, DPI if  $\alpha \in (0, 2]$
- Belavkin-Staszewski (maximal) relative entropy as limit

$$\lim_{\alpha \rightarrow 1} D_{\alpha}^{max}(\rho\|\sigma) = S_{BS}(\rho\|\sigma) := \text{Tr} \left[ \rho \log \left( \rho^{1/2} \sigma^{-1} \rho^{1/2} \right) \right]$$

# Extensions to von Neumann algebras

- In some infinite dimensional situations the previous definitions do not work.
- Useful also in e.g. QFT
- Technical problems: no density matrices (operators) in general, no matrix analysis tools...
- Other tools: modular theory, non-commutative  $L_p$ -spaces, complex interpolation



# Extensions to von Neumann algebras

- Relative entropy (Araki, 1976)
  - relative modular operator
- Petz-type Rényi divergences (Petz, 1985)
  - relative modular operator, operator convex functions
- Minimal Rényi divergences (Berta et al, 2018, AJ 2018, 2021)
  - weighted  $L_p$ -norms, interpolation
- $\alpha - z$ -Rényi divergences (Hiai & AJ, 2024)
  - weighted  $L_p$ -norms, variational formulas
- Maximal Rényi divergences (Hiai, 2019)
  - operator means, generalized connections

# Quantum Rényi divergences and $L_p$ -spaces

Rényi divergence:  $D_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1} \log Q_\alpha(\rho\|\sigma)$

- **Classical case:**  $Q$ -weighted  $L_p$ -norm

$$Q_\alpha(P\|Q) = \int (dP/dQ)^\alpha dQ = \|dP/dQ\|_{\alpha,Q}^\alpha$$

- **Quantum sandwiched case:**  $\sigma$ -weighted  $L_p$ -norm (AJ, 2018)

$$Q_\alpha(\rho\|\sigma) = \text{Tr} \left[ \left( \sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right] = \|\rho\|_{\alpha,\sigma}^\alpha,$$

- For  $\alpha > 1$ : complex interpolation norm (Kosaki, 1984)

$\|\sigma^{-1/2} \rho \sigma^{-1/2}\|$  (operator norm),  $\|\rho\|_1$  (trace norm)

- works in general von Neumann algebras

# $\alpha - z$ -Rényi divergences and $L_p$ -spaces

Variational formula: (Kato, 2024, Hiai & AJ, 2024)

- For  $\alpha \in (0, 1)$ ,  $p = \frac{z}{\alpha}$ ,  $r = \frac{z}{1-\alpha}$ :

$$Q_{\alpha,z}(\rho\|\sigma) = \inf_{a \text{ p.d.}} \left\{ \alpha \|\rho^{\frac{1}{2}} a \rho^{\frac{1}{2}}\|_{p,\rho}^p + (1-\alpha) \|\sigma^{\frac{1}{2}} a^{-1} \sigma^{\frac{1}{2}}\|_{r,\sigma}^r \right\}$$

- For  $\alpha > 1$ ,  $p = \frac{z}{\alpha}$ ,  $q = \frac{z}{\alpha-1}$ :

$$Q_{\alpha,z}(\rho\|\sigma) = \sup_{a \geq 0} \left\{ \alpha \|\rho^{\frac{1}{2}} a \rho^{\frac{1}{2}}\|_{p,\rho}^p - (\alpha-1) \|\sigma^{\frac{1}{2}} a \sigma^{\frac{1}{2}}\|_{q,\sigma}^q \right\}$$

- Connects to the weighted  $L_p$ -norms for all  $\alpha, z$
- Extends many results to von Neumann algebras

# Quantum sufficient statistics?

- Quantum statistics - quantum channels
- When is a channel  $\Phi$  **sufficient** w. r. to a set of states  $\mathcal{S}$ ?
- **Conditional expectations** do not exist in most situations

**Sufficient quantum channels:** (Petz, 1986)

A channel  $\Phi$  is sufficient with respect to  $\mathcal{S}$  if there is another channel  $\Psi$  such that

$$\Psi \circ \Phi(\rho) = \rho, \quad \rho \in \mathcal{S}.$$

- $\Phi$  is **reversible** on  $\mathcal{S}$ ,  $\Psi$  - **recovery map**
- sufficient statistics in classical case

# Sufficient quantum channels

Characterizations of sufficient quantum channels: (Petz, 1986, 1988)

- **Petz theorem**: if  $\text{supp } \rho \leq \text{supp } \sigma$  for all  $\rho \in \mathcal{S}$

$$S(\Phi(\rho) \parallel \Phi(\sigma)) = S(\rho \parallel \sigma), \quad \rho \in \mathcal{S}$$

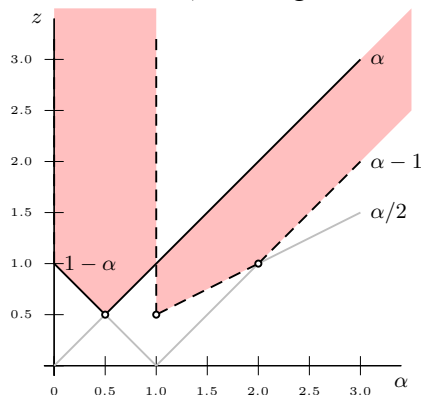
- There is a **universal recovery map**:  $\Phi_\sigma$  (Petz recovery map)

$$\Phi_\sigma \circ \Phi(\rho) = \rho, \quad \rho \in \mathcal{S}$$

- structure of the states  $\rho \in \mathcal{S}$ , strong conditions.
- For classical statistics:  $E_q[\cdot|T] = E_p[\cdot|T]$  is the Petz recovery.

# Quantum Rényi divergences and sufficient channels

Assume that  $\alpha, z$  belong to the following set:



Then  $\Phi$  is sufficient w.r. to  $\{\rho, \sigma\}$  if and only if (Hiai & AJ, 2024)

$$D_{\alpha, z}(\Phi(\rho) \parallel \Phi(\sigma)) = D_{\alpha, z}(\rho \parallel \sigma).$$

- holds in general von Neumann algebras

# Classical to quantum and in between

Classical	Classical/quantum	Quantum
discrete probability measures $p, q$	commuting density matrices $\rho, \sigma$	density matrices $\rho, \sigma$ in $M_n(\mathbb{C})$
probability measures $P, Q \ll \mu$ on a measure space $(X, \Omega, \mu)$	$L_\infty(X, \Omega, \mu)$ , densities $p, q \in L_1(X, \Omega, \mu)$	normal states $\rho, \sigma$ of a von Neumann algebra
$T : X \rightarrow Y$ statistic, Markov kernel $X \times Y \rightarrow [0, 1]$	Positive trace preserving map	Quantum channel $M_n(\mathbb{C}) \rightarrow M_m(\mathbb{C})$
transformation of probability measures	A Markov map $L_\infty(X, \Omega, \mu) \rightarrow L_\infty(Y, \Sigma, \nu)$	Unital normal cp map $\mathcal{M} \rightarrow \mathcal{N}$
conditional expectation $E_p[\cdot T]$	positive unital projection preserving $p$	Petz recovery map $\Phi_\sigma$