One-Shot Quantum Information Theory I: Entropic Quantities

Nilanjana Datta

University of Cambridge,U.K.

In Quantum information theory, initially one evaluated:

- optimal rates of info-processing tasks, e.g.,
  - data compression,
  - transmission of information through a channel, etc.

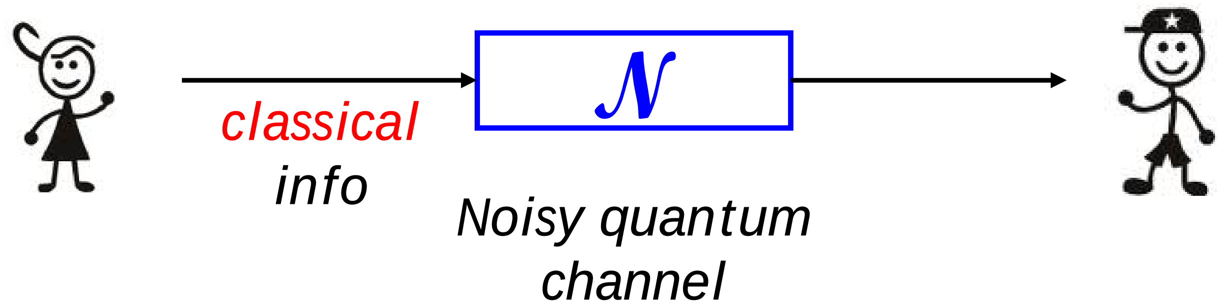
under the assumption of an *"asymptotic, memoryless setting"*

Assume:

- information sources & channels are memoryless
- They are available for asymptotically many uses

UNIVERSITY OF CAMBRIDGE

- **E.g. Transmission of classical information**



Optimal rate (of classical information transmission):

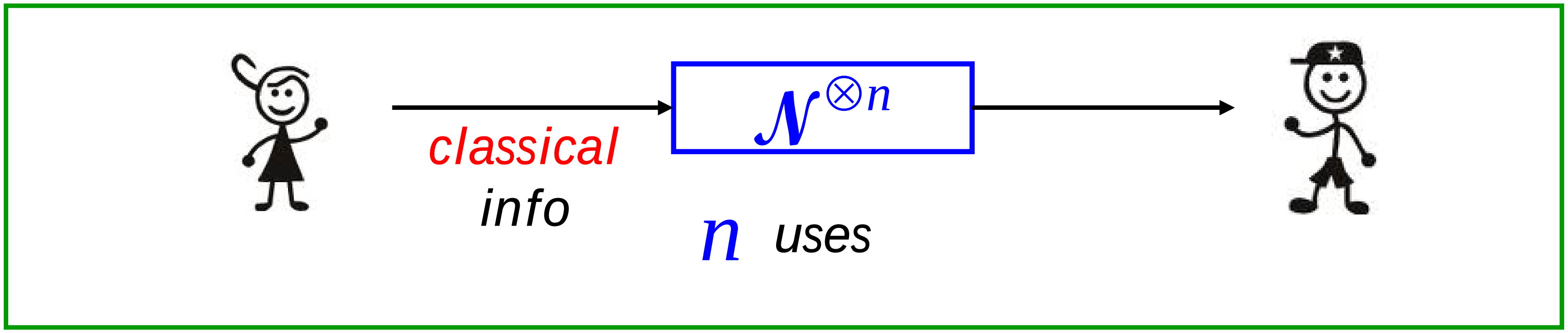
classical capacity

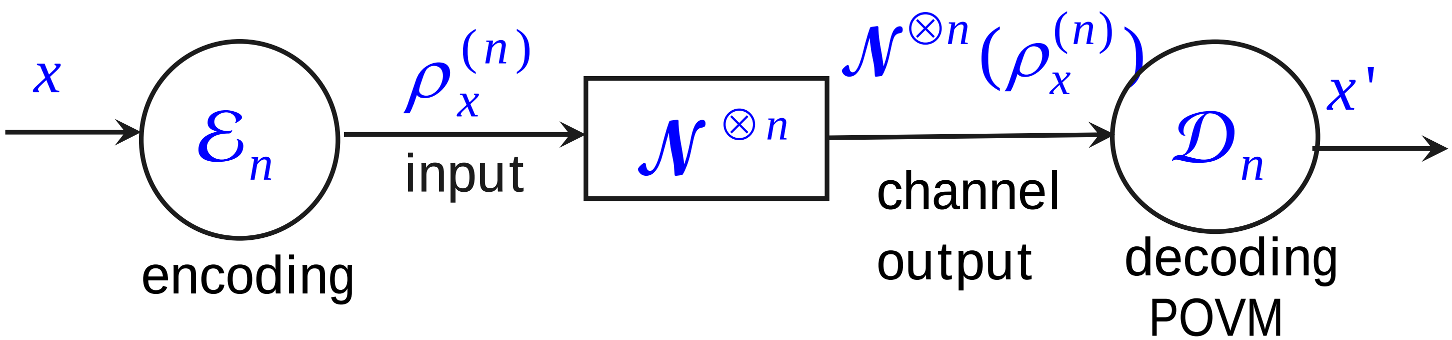
$C(\mathcal{N}) =$ maximum number of bits transmitted per use of $\mathcal{N}$

memoryless: there is no correlation in the noise acting on successive inputs

$\mathcal{N}^{\otimes n}$: n successive uses of the channel; independent *"asymptotic, memoryless setting"*

- *To evaluate* $C(\mathcal{N})$:





- One requires : **prob. of error** $p_e^{(n)} \to 0$ as $n \to \infty$

$C(\mathcal{N})$: *Optimal rate of reliable information transmission*

Entropic Quantities

Optimal rates of information-processing tasks in the

"asymptotic, memoryless setting"

- *Compression of Information*:

Memoryless quantum info. source $\{\rho, \mathcal{H}\}$ *[Schumacher]*

- Data compression limit: $S(\rho)$ *von Neumann entropy*

- *Info Transmission thro' a memoryless quantum channel* $\mathcal{N}$

- Classical capacity $C(\mathcal{N})$ *[Holevo, Schumacher, Westmoreland]*

  --given in terms of the Holevo capacity ;

- Quantum capacity $Q(\mathcal{N})$ *[Lloyd, Shor, Devetak]*

  --given in terms of the coherent information ;

UNIVERSITY OF CAMBRIDGE

These entropic quantities are all obtainable from a single parent quantity;

Quantum relative entropy: For $\rho, \sigma \geq 0; \quad \mathrm{Tr}\rho = 1$

$$D(\rho \| \sigma) := \mathrm{Tr}\left(\rho \log \rho\right) - \mathrm{Tr}\left(\rho \log \sigma\right)$$

$\mathrm{supp}\ \rho \subseteq \mathrm{supp}\ \sigma$

e.g. Data compression limit:

$$S(\rho) := -\mathrm{Tr}\left(\rho \log \rho\right) = -D(\rho \| I) \qquad (\sigma = I)$$

In real-world applications

"asymptotic memoryless setting" not necessarily valid

- **In practice:** information sources & channels are used a finite number of times;
- there are unavoidable correlations between successive uses *(memory effects)*

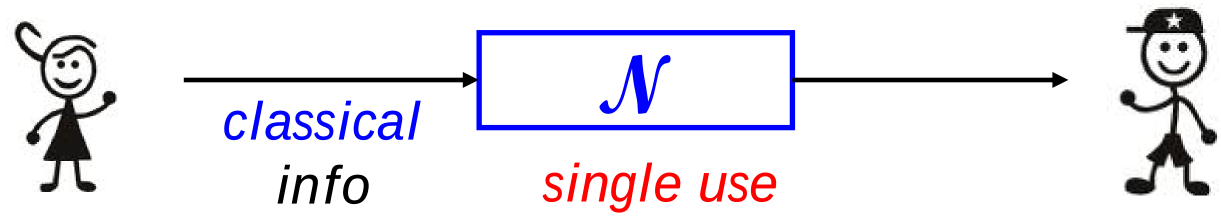
Hence it is important to evaluate optimal rates for a *finite number of uses (or even a single use)* *of an arbitrary source or channel*

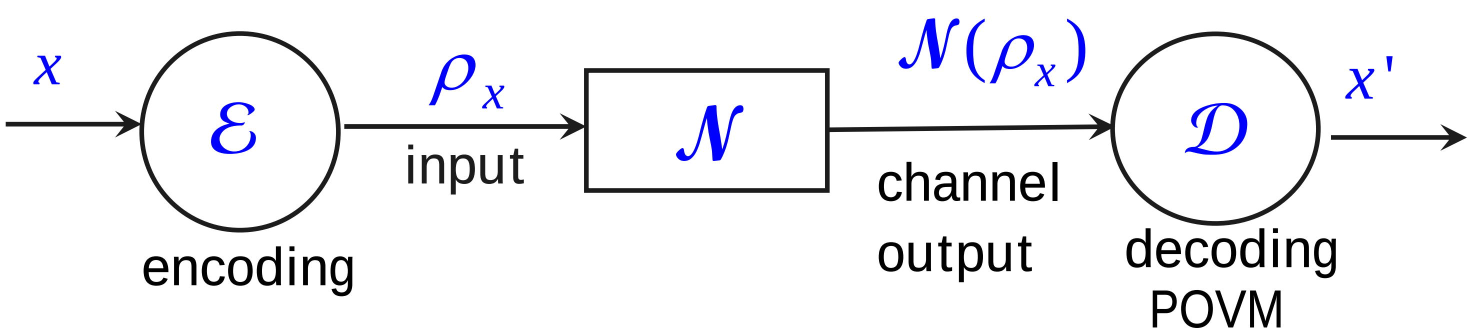
- Evaluation of corresponding optimal rates:

→ **One-shot information theory**

UNIVERSITY OF CAMBRIDGE

One-shot information theory





One-shot $\varepsilon -$ *error* *classical capacity* $:=$ *max. number of bits that can be transmitted on a single use of* $\mathcal{N}$

$$C^{(1)}_{\varepsilon}(\mathcal{N})$$

Prob. of error: $p_e \leq \varepsilon$ for some $\varepsilon > 0,$

UNIVERSITY OF CAMBRIDGE

In the one-shot setting too...

- Capacities, data compression limit etc. are

-- given in terms of entropic quantities

Min-/max-/0- entropies (R.Renner)

- Obtainable from certain (generalized) relative entropies

Parent quantities for optimal 'rates' in the one-shot setting

$D_{\max}(\rho\|\sigma)$	$D_0(\rho\|\sigma)$	$D_{\min}(\rho\|\sigma)$
Max-relative entropy	*0-relative Renyi entropy*	*Min-relative entropy*

"super-parent": $\tilde{D}_\alpha(\rho\|\sigma)$

Quantum Renyi Divergence

(sandwiched Renyi relative entropy)

[Wilde et al; Muller-Lennert et al]

UNIVERSITY OF CAMBRIDGE

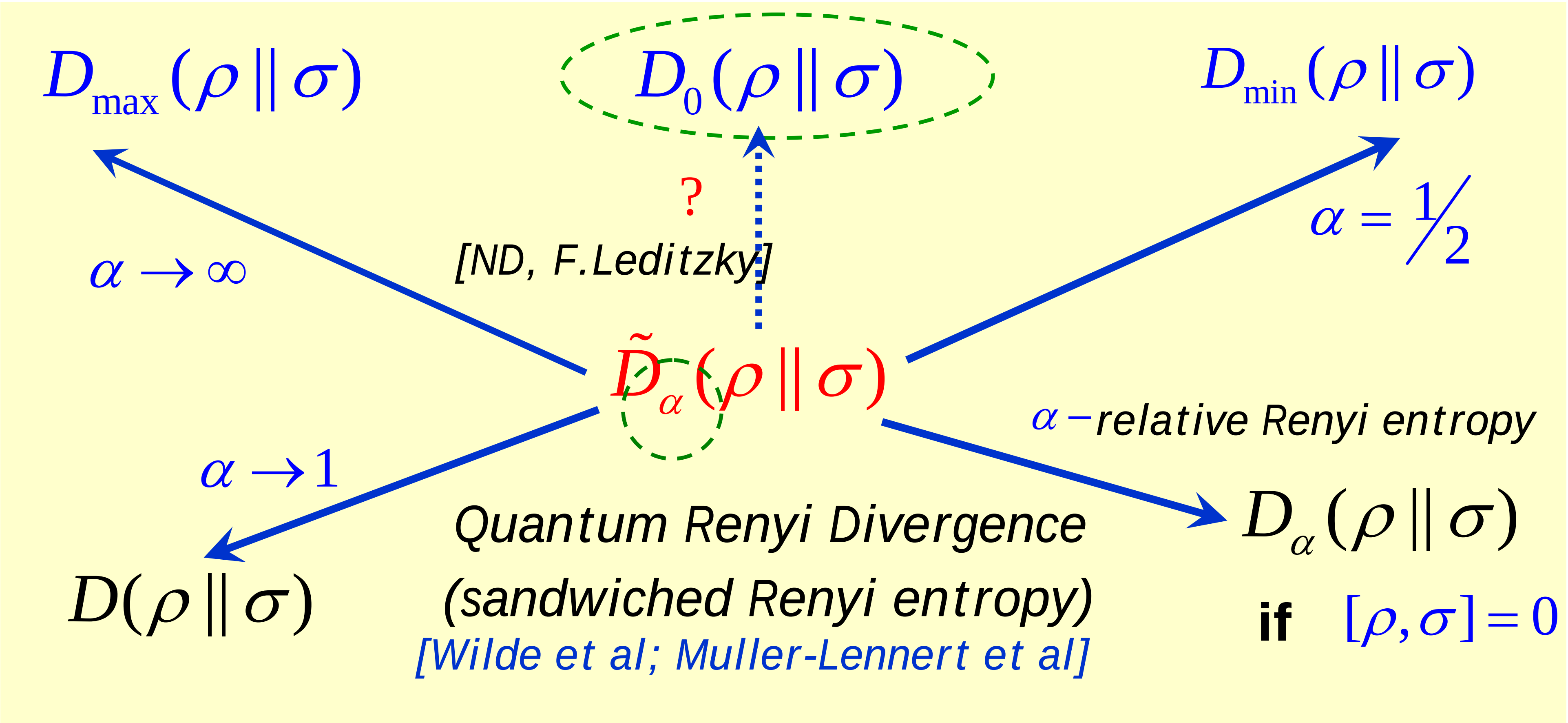
In the one-shot setting too...

- Capacities, data compression limit etc. are

-- given in terms of entropic quantities

Min-/max- entropies (R.Renner)

- Obtainable from certain (generalized) relative entropies



UNIVERSITY OF CAMBRIDGE

Outline

- *Mathematical Tool:*
  - *Semidefinite programming*

- *Definitions of generalized relative entropies:*

$$D_{\max}(\rho\|\sigma), D_0(\rho\|\sigma), D_{\min}(\rho\|\sigma)$$

- *Properties & operational significances of them*

- *Their "smoothed" versions*

- *Their children:* *the min-, max- and 0-entropies*

- *The "super-parent" : Quantum Renyi Divergence* $\tilde{D}_\alpha(\rho\|\sigma)$

- *Relationship between* $\tilde{D}_\alpha(\rho\|\sigma)$ & $D_0(\rho\|\sigma)$

UNIVERSITY OF CAMBRIDGE

Notations & Definitions

$$A \longleftrightarrow \mathcal{H}_A$$

quantum system *Hilbert space*

$\mathcal{B}(\mathcal{H})$: algebra of linear operators acting on $\mathcal{H}$

$\mathcal{P}(\mathcal{H})$: set of positive operators……

$\mathcal{D}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$: set of density matrices (states)

- **Linear maps:** If $\Lambda : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ $\left(\Lambda : A \to B\right)$

  its **adjoint map**: $\Lambda^* : B \to A$

  defined through $\mathrm{Tr}\left(\omega_B \Lambda(\rho_A)\right) = \mathrm{Tr}\left(\Lambda^*(\omega_B)\rho_A\right)$

- Quantum operations (quantum channels) : linear CPTP map

Mathematical Tool

Semi-definite programming (SDP)

- *A well-established form of convex optimization*
- *The objective function is linear in an input constrained to a semi-definite cone*
- *Efficient algorithms have been devised for its solution*

UNIVERSITY OF CAMBRIDGE

Mathematical Tool

(2) Semi-definite programming (SDP)

(formulation:Watrous)

$$(\Lambda, A, B); \quad A, B \in \mathcal{P}(\mathcal{H}),$$

$$\Lambda : \mathcal{P}(\mathcal{H}_A) \to \mathcal{P}(\mathcal{H}_B)$$ positivity-preserving map

- **Primal problem**

minimize $\mathrm{Tr}(AX)$

subject to $\Lambda(X) \geq B;$

$X \geq 0;$

- **Dual problem**

maximize $\mathrm{Tr}(BY)$

subject to $\Lambda^*(Y) \leq A;$

$Y \geq 0;$

Optimal solutions: $\alpha = \beta$

IF *Slater's duality condition holds.*

Outline

- *Mathematical Toolkit:*
  - *Semidefinite programming*

- *Definitions of generalized relative entropies:*

$$D_{\max}(\rho \| \sigma), D_0(\rho \| \sigma), D_{\min}(\rho \| \sigma)$$

UNIVERSITY OF CAMBRIDGE

Definitions of generalized relative entropies

$\rho \in \mathcal{D}(\mathcal{H}),\ \sigma \in \mathcal{P}(\mathcal{H});\ \text{supp}\,\rho \subseteq \text{supp}\,\sigma;$

- *Max-relative entropy [ND]*

$$D_{\max}(\rho \| \sigma) := \inf\left\{\gamma : \rho \le 2^{\gamma}\sigma\right\}$$

$$\sigma^{-1/2}\rho\sigma^{-1/2} \le 2^{\gamma} I$$

$$= \log\left(\lambda_{\max}(\sigma^{-1/2}\rho\sigma^{-1/2})\right)$$

- *Min-relative entropy [Dupuis et al]*

$$D_{\min}(\rho \| \sigma) := -2\log \|\sqrt{\rho}\sqrt{\sigma}\|_1$$

$$= -2\log\ F(\rho,\sigma)$$

fidelity

UNIVERSITY OF CAMBRIDGE

Definitions of generalized relative entropies

contd.

$\rho \in \mathcal{D}(\mathcal{H}),\ \sigma \in \mathcal{P}(\mathcal{H});\ \ \text{supp}\,\rho \subseteq \text{supp}\,\sigma;$

- *0-relative Renyi entropy*

$$D_0(\rho \| \sigma) := -\log\left(\text{Tr}\,(\pi_\rho \sigma)\right)$$

where π_ρ *denotes the projector onto* $\text{supp}\,\rho$

- α *-relative Renyi entropy* $(\alpha \neq 1)$

$$D_\alpha(\rho \| \sigma) := \frac{1}{\alpha - 1}\log\ \text{Tr}\,(\rho^\alpha \sigma^{1-\alpha})$$

$$\lim_{\alpha \to 0^+} D_\alpha(\rho \| \sigma) = D_0(\rho \| \sigma)$$

UNIVERSITY OF CAMBRIDGE

Properties of generalized relative entropies

for $* = \max,\ 0,\ \min$

- Positivity: If $\rho, \sigma \in \mathcal{D}(\mathcal{H})$,

$$D_*(\rho \| \sigma) \geq 0$$

just as $D(\rho \| \sigma)$

- Data-processing inequality:

$$D_*(\Lambda(\rho) \| \Lambda(\sigma)) \leq D_*(\rho \| \sigma)$$

for any CPTP map Λ

- Invariance under joint unitaries:

$$D_*(U \rho U^\dagger \| U \sigma U^\dagger) = D_*(\rho \| \sigma)$$

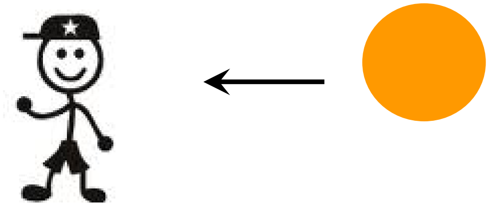
for any unitary operator U

- *Interestingly,*

$$D_0(\rho \| \sigma) \leq D_{\min}(\rho \| \sigma) \leq D(\rho \| \sigma) \leq D_{\max}(\rho \| \sigma)$$

Operational interpretation of the max-relative entropy

- *Multiple state discrimination problem:*



Bob

a quantum system

& told

	its state	with prob.
	ρ_1	$1/m$
	ρ_2	⋮
	⋮	
	ρ_m	$1/m$

- He does measurements to infer the state: **POVM**

$$\{E_1,..,E_m\}:\ 0 \le E_i \le I;\ \sum_{i=1}^{m} E_i = I$$

- *His optimal average success probability:*

$$p^*_{succ} := \max_{\{E_1,..,E_m\}} \frac{1}{m} \sum_{i=1}^{m} \mathrm{Tr}\left(E_i \rho_i\right)$$

UNIVERSITY OF CAMBRIDGE

- *Theorem 3* *[M.Mosonyi & ND]:*

The optimal average success probability in this multiple state discrimination problem is given by:

$$p^*_{succ} = \frac{1}{m} \min_{\sigma} \max_{1 \le i \le m} 2^{D_{\max}(\rho_i \| \sigma)}$$

UNIVERSITY OF CAMBRIDGE

Sketch of proof:

- Let $\left\{|i\rangle\right\}_{i=1}^{m}$: basis in $\mathbf{C}^m$
- Let $\tilde{\rho} := \frac{1}{m}\sum_{i=1}^{m}|i\rangle\langle i|\otimes\rho_i$

$\in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H})$

$$p^*_{succ} := \max_{\{E_1,..,E_m\}} \frac{1}{m}\sum_{i=1}^{m}\mathrm{Tr}\left(E_i\rho_i\right)$$

$$= \max_{\{E_i\}_{i=1}^{m}:POVM} \mathrm{Tr}\left[\tilde{\rho}\left(\sum_{i=1}^{m}|i\rangle\langle i|\otimes E_i\right)\right]$$

$$Y = \sum_{i=1}^{m}|i\rangle\langle i|\otimes E_i \in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H});$$

$$\mathrm{Tr}_{\mathbf{C}^m} Y = \sum_i E_i = I_{\mathcal{H}};$$

$$= \max_{\substack{Y\in\mathcal{P}(\mathbf{C}^m\otimes\mathcal{H});\\ \mathrm{Tr}_{\mathbf{C}^m}Y = I_{\mathcal{H}}}} \mathrm{Tr}\left(\tilde{\rho}Y\right)$$

UNIVERSITY OF CAMBRIDGE

Sketch of proof:

- Let $\{|i\rangle\}_{i=1}^{m}$: basis in $\mathbf{C}^m$
- Let $\tilde{\rho} := \frac{1}{m}\sum_{i=1}^{m} |i\rangle\langle i| \otimes \rho_i$ $\in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H})$

$$p^{*}_{succ} := \max_{\{E_1,\ldots,E_m\}} \frac{1}{m}\sum_{i=1}^{m} \mathrm{Tr}\left(E_i \rho_i\right)$$

$$= \max_{\{E_i\}_{i=1}^{m}:POVM} \mathrm{Tr}\left[\tilde{\rho}\left(\sum_{i=1}^{m} |i\rangle\langle i| \otimes E_i\right)\right]$$

$$Y = \sum_{i=1}^{m} |i\rangle\langle i| \otimes E_i \in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H});$$

$$\mathrm{Tr}_{\mathbf{C}^m} Y = \sum_i E_i = I_{\mathcal{H}};$$

$$= \max_{\substack{Y \in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H}); \\ \mathrm{Tr}_{\mathbf{C}^m} Y \le I_{\mathcal{H}}}} \mathrm{Tr}\left(\tilde{\rho} Y\right)$$

UNIVERSITY OF CAMBRIDGE

Sketch of proof:

- Let $\{|i\rangle\}_{i=1}^{m}$: basis in $\mathbf{C}^m$
- Let $\tilde{\rho} := \frac{1}{m}\sum_{i=1}^{m} |i\rangle\langle i| \otimes \rho_i$

$\in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H})$

$$p^*_{succ} := \max_{\{E_1,\ldots,E_m\}} \frac{1}{m}\sum_{i=1}^{m} \mathrm{Tr}\left(E_i \rho_i\right)$$

$$= \max_{\{E_i\}_{i=1}^{m}:POVM} \mathrm{Tr}\left[\tilde{\rho}\left(\sum_{i=1}^{m} |i\rangle\langle i| \otimes E_i\right)\right]$$

$$Y = \sum_{i=1}^{m} |i\rangle\langle i| \otimes E_i \in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H});$$

$$\mathrm{Tr}_{\mathbf{C}^m} Y = \sum_i E_i = I_{\mathcal{H}};$$

$$= \max_{\substack{Y \in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H});\\ \mathrm{Tr}_{\mathbf{C}^m} Y \le I_{\mathcal{H}}}} \mathrm{Tr}\left(\tilde{\rho} Y\right)$$

- *SDP*
- *duality condition holds*

[Koenig, Renner, Schaffner]

UNIVERSITY OF CAMBRIDGE

- **Primal problem**

minimize $\mathrm{Tr}(AX)$

subject to $\Lambda(X) \geq B;$

$X \geq 0;$

- **Dual problem**

maximize $\mathrm{Tr}(BY)$

subject to $\Lambda^{*}(Y) \leq A;$

$Y \geq 0;$

minimize $\mathrm{Tr}\, X$

subject to $I_{\mathbf{C}^m} \otimes X \geq \tilde{\rho}$

$X \geq 0;$

$$= \max_{\substack{Y \in \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H});\\ \mathrm{Tr}_{\mathbf{C}^m} Y \leq I_{\mathcal{H}}}} \mathrm{Tr}\left(\tilde{\rho} Y\right)$$

$$\mathrm{Tr}(AX) = \mathrm{Tr}(I_{\mathcal{H}} X) = \mathrm{Tr}\, X$$

$B = \tilde{\rho};\quad A = I_{\mathcal{H}};\quad \Lambda^{*} = \mathrm{Tr}_{\mathbf{C}^m}$

$$\Lambda^{*} = \mathrm{Tr}_{\mathbf{C}^m} : \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H}) \to \mathcal{P}(\mathcal{H})$$

$$\Rightarrow \Lambda : \mathcal{P}(\mathcal{H}) \to \mathcal{P}(\mathbf{C}^m \otimes \mathcal{H});\quad \Lambda(X) = I_{\mathbf{C}^m} \otimes X$$

UNIVERSITY OF CAMBRIDGE

Sketch of proof contd:

$$p^*_{succ} = \min \left\{ \mathrm{Tr}\, X : X \geq 0,\ \tilde{\rho} \leq I_{\mathbf{C}^m} \otimes X \right\}; \tilde{\rho} := \frac{1}{m}\sum_{i=1}^{m} |i\rangle\langle i| \otimes \rho_i$$

$$= \min \left\{ \mathrm{Tr}\, X : X \geq 0,\ \rho_i \leq mX\ \ \forall i = 1, 2, .., m \right\}$$

$$\left[\ldots = m\mathrm{Tr}X \Rightarrow \mathrm{Tr}X = \frac{\mathrm{Tr}\tilde{X}}{m} \right]$$

$$\ldots \forall i = 1, 2, .., m \}$$

$$\left[\ldots \tilde{X}; \sigma = \frac{\tilde{X}}{\mathrm{Tr}\tilde{X}};\ \tilde{X} = \lambda\sigma \right]$$

$$\ldots \sigma\ \forall i = 1, 2, .., m \}$$

$$\left[\min \left\{ \lambda : \rho_i \leq \lambda\sigma \right\} = 2^{D_{\max}(\rho_i \| \sigma)} \right]$$

$$\tilde{\rho} \leq I_{\mathbf{C}^m} \otimes X$$

$$\frac{1}{m}\sum_{i=1}^{m} |i\rangle\langle i| \otimes \rho_i \leq \sum_{i=1}^{m} |i\rangle\langle i| \otimes X$$

$$\Rightarrow \frac{\rho_i}{m} \leq X\ \ \forall i$$

$$= \frac{1}{m} \min_{\sigma} \max_{1 \leq i \leq M} 2^{D_{\max}(\rho_i \| \sigma)}$$

$\blacksquare$

UNIVERSITY OF CAMBRIDGE

Operational interpretation of $D_0(\rho \| \sigma) := -\log\left(\mathrm{Tr}\,(\pi_\rho \sigma)\right)$

- *Quantum binary hypothesis testing:*

Bob receives a state ρ *(null hypothesis)* or σ *(alternative hypothesis)*

- He does a measurement to infer which state it is

POVM $A\,[\rho]$ & $(I - A)\,[\sigma]$

-

Possible errors	*inference*	*actual state*
Type I	σ	ρ
Type II	ρ	σ

- *Error probabilities*

$$\alpha = \mathrm{Tr}((I - A)\rho) \quad \textit{Type I}$$

$$\beta = \mathrm{Tr}(A\sigma) \quad \textit{Type II}$$

UNIVERSITY OF CAMBRIDGE

- *Suppose* (POVM element) $A = \pi_\rho$

Prob(Type I error)

$$\alpha = \mathrm{Tr}((I - A)\rho)$$
$$= 0$$

Prob(Type II error)

$$\beta = \mathrm{Tr}(A\sigma)$$
$$= \mathrm{Tr}(\pi_\rho \sigma)$$

Bob never infers the state to be σ *when it is* ρ

BUT

$$D_0(\rho \| \sigma) := -\log \mathrm{Tr}\, \pi_\rho \sigma$$

Hence $\beta\big|_{\alpha=0} = 2^{-D_0(\rho\|\sigma)}$

= Prob(Type II error | Type I error = zero)

- *Suppose* (POVM element) $A = \pi_\rho$

Prob(Type I error)

$$\alpha = \mathrm{Tr}((I - A)\rho)$$
$$= 0$$

Prob(Type II error)

$$\beta = \mathrm{Tr}(A\sigma)$$
$$= \mathrm{Tr}(\pi_\rho \sigma)$$

Bob never infers the state to be σ *when it is* ρ

BUT

$$D_0(\rho \| \sigma) := -\log \mathrm{Tr}\, \pi_\rho \sigma$$

In fact, *min Prob(Type II error | Type I error = zero)*

$$\beta^* \big|_{\alpha=0} = 2^{-D_0(\rho\|\sigma)}$$

UNIVERSITY OF CAMBRIDGE

Smoothed relative entropies

- What i ... allows

non-zero ... α ?

D_0 ... $\rho\sigma$

$\alpha = \mathrm{Tr}$ ... $\mathrm{Tr}\,(\rho) = 1$

$$\beta^*|_{\alpha=0} = \min_{\substack{0\le A\le I \\ \mathrm{Tr}\,(A\rho)=1}} \mathrm{Tr}\,(A\sigma)$$

$$\beta^*|_{\alpha\le\varepsilon} = \min_{\substack{0\le A\le I \\ \mathrm{Tr}\,(A\rho)\ge 1-\varepsilon}} \mathrm{Tr}\,(A\sigma)$$

$\therefore$ For $\alpha \le \varepsilon$ choose A such that $\mathrm{Tr}\,(A\rho) \ge 1-\varepsilon$

$$D_0^\varepsilon(\rho\,\|\,\sigma) = -\log \beta^*|_{\alpha\le\varepsilon} = \max_{\substack{0\le A\le I \\ \mathrm{Tr}\,(A\rho)\ge 1-\varepsilon}} \big(-\log(\mathrm{Tr}\,(A\rho))\big)$$

Hypothesis testing relative entropy
[Wang & Renner]

$$\equiv D_H^\varepsilon(\rho\,\|\,\sigma)$$

UNIVERSITY OF CAMBRIDGE

Compare **operational significances** of $D_H^{\varepsilon}(\rho\|\sigma)$ & $D(\rho\|\sigma)$

$D(\rho\|\sigma)$ arises in asymptotic binary hypothesis testing

- Suppose Bob is given many (n) identical copies of the state

- He receives $\rho^{\otimes n}$ or $\sigma^{\otimes n}$

Bob's POVM

$$\{A_n, (I - A_n)\}$$

$\beta^{*(n)}|_{\alpha(n)\leq\varepsilon} :=$ *Minimum type II error when type I error* $\leq \varepsilon$

$$\forall \varepsilon \in [0,1): \quad \lim_{n\to\infty}\left\{-\frac{1}{n}\log \beta^{*(n)}|_{\alpha(n)\leq\varepsilon}\right\} = D(\rho\|\sigma)$$

[Quantum Stein's Lemma]

UNIVERSITY OF CAMBRIDGE

Operational interpretations in binary hypothesis testing

$D_H^{\varepsilon}(\rho \| \sigma)$

One-shot setting;

Single copy of the state:

$$= -\log \beta^* |_{\alpha \le \varepsilon}$$

$D(\rho \| \sigma)$

Asymptotic memoryless setting;

Multiple copies of the state:

$$= \lim_{n \to \infty} \left\{ -\frac{1}{n} \log \beta^{*(n)} |_{\alpha(n) \le \varepsilon} \right\}$$

$\forall \varepsilon \in [0,1):$

(Bob receives identical copies of the state: $\rho^{\otimes n}$ *or* $\sigma^{\otimes n}$ *)*

UNIVERSITY OF CAMBRIDGE

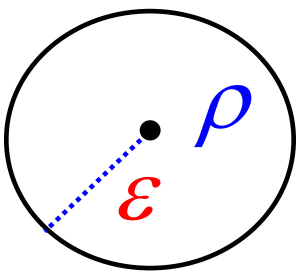
Smooth max-relative entropy

$\forall \varepsilon \geq 0.$

$$D_{\max}^{\varepsilon}(\rho \| \sigma) := \min_{\bar{\rho} \in B^{\varepsilon}(\rho)} D_{\max}(\bar{\rho} \| \sigma)$$

$$B^{\varepsilon}(\rho) := \{\bar{\rho} \geq 0, \operatorname{Tr}\bar{\rho} = 1 : F(\rho, \bar{\rho}) \geq 1 - \varepsilon\}$$

fidelity



$2^{D_{\max}^{\varepsilon}(\rho\|\sigma)}$ & $2^{-D_{H}^{\varepsilon}(\rho\|\sigma)}$ can both be formulated as SDPs

UNIVERSITY OF CAMBRIDGE

Outline

- *Mathematical Toolkit:*
  - *Semidefinite programming*
- *Definitions of generalized relative entropies:*

$$D_{\max}(\rho \| \sigma), D_0(\rho \| \sigma), D_{\min}(\rho \| \sigma)$$

- *Properties & operational significances of them*
- *Their children: the min-, max- and 0-entropies*

UNIVERSITY OF CAMBRIDGE

$D_{\max}(\rho\|\sigma), D_0(\rho\|\sigma)$ & $D_{\min}(\rho\|\sigma)$

as parent quantities for other entropies

Just as:

von Neumann entropy

$$S(\rho) = -D(\rho\|I)$$

$$(\sigma = I)$$

$$H_{\min}(\rho) := -D_{\max}(\rho\|I) = -\log\lambda_{\max}(\rho)$$

$$H_0(\rho) := -D_0(\rho\|I) = \log \operatorname{rank}(\rho)$$

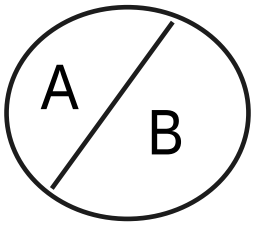
$$H_{\max}(\rho) := -D_{\min}(\rho\|I) = 2\log\left(\mathrm{Tr}\sqrt{\rho}\right)$$

[Renner]

UNIVERSITY OF CAMBRIDGE

Other min- & max- entropies

For a bipartite state ρ_{AB} :



Conditional entropy

$$S(A|B) = S(\rho_{AB}) - S(\rho_B) = \max_{\sigma_B} \{-D(\rho_{AB} \| I_A \otimes \sigma_B)\}$$

Conditional min-entropy

$$H_{\min}(A|B)_\rho := \max_{\sigma_B} \{-D_{\max}(\rho_{AB} \| I_A \otimes \sigma_B)\}$$

Max-conditional entropy

$$H_{\max}(A|B)_\rho := \max_{\sigma_B} \{-D_{\min}(\rho_{AB} \| I_A \otimes \sigma_B)\}$$

0-conditional entropy

$$H_0(A|B)_\rho := \max_{\sigma_B} \{-D_0(\rho_{AB} \| I_A \otimes \sigma_B)\}$$

UNIVERSITY OF CAMBRIDGE

- They have interesting mathematical properties:

- e.g. **Duality relation:** *[Koenig, Renner, Schaffner]:*

For any purification ρ_{ABC} of a bipartite state ρ_{AB} :

$$H_{\max}(A|B)_\rho = -H_{\min}(A|C)_\rho$$

(just as for the von Neumann entropy): $$H(A|B)_\rho = -H(A|C)_\rho$$

-- and -- interesting operational interpretations:

UNIVERSITY OF CAMBRIDGE

Operational interpretations

- *Conditional min-entropy* $\sim$

  maximum achievable singlet fraction

- *Conditional max-entropy* $\sim$

  [Koenig, Renner, Schaffner]

  decoupling accuracy

- *Conditional 0-entropy* $\sim$

  one-shot entanglement cost under LOCC

  [F.Buscemi, ND]

UNIVERSITY OF CAMBRIDGE

Operational interpretations contd.

- *Conditional min-entropy* $\sim$ *Max. achievable singlet fraction*

$$|\Phi_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i_A\rangle |i_B\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B :$$

max. entangled state

$$\Phi_{AB} = |\Phi_{AB}\rangle\langle\Phi_{AB}|$$

[Koenig, Renner, Schaffner]

$$2^{-H_{\min}(A|B)_\rho} = d \max_{\Lambda_B : CPTP} F\left(\left((\mathrm{id}_A \otimes \Lambda_B)\rho_{AB}\right), \Phi_{AB}\right)^2$$

fidelity

Given the bipartite state ρ_{AB}, *it is the maximum overlap with the singlet state* Φ_{AB}, *that can be achieved by local quantum operations* Λ_B *on the subsystem* B.

UNIVERSITY OF CAMBRIDGE

Operational interpretations contd.

- *Conditional max-entropy* $\sim$ *Decoupling accuracy*

Distance of ρ_{AB}, *from a product state* $\tau_A \otimes \sigma_B$

no correlations; *decoupled*

$\tau_A = \frac{I}{d_A}$ *completely mixed state on* $\mathcal{H}_A$

[Koenig, Renner, Schaffner]

$$2^{H_{\max}(A|B)_\rho} = d_A \max_{\sigma_B} F\left(\rho_{AB}, \tau_A \otimes \sigma_B\right)^2$$

fidelity

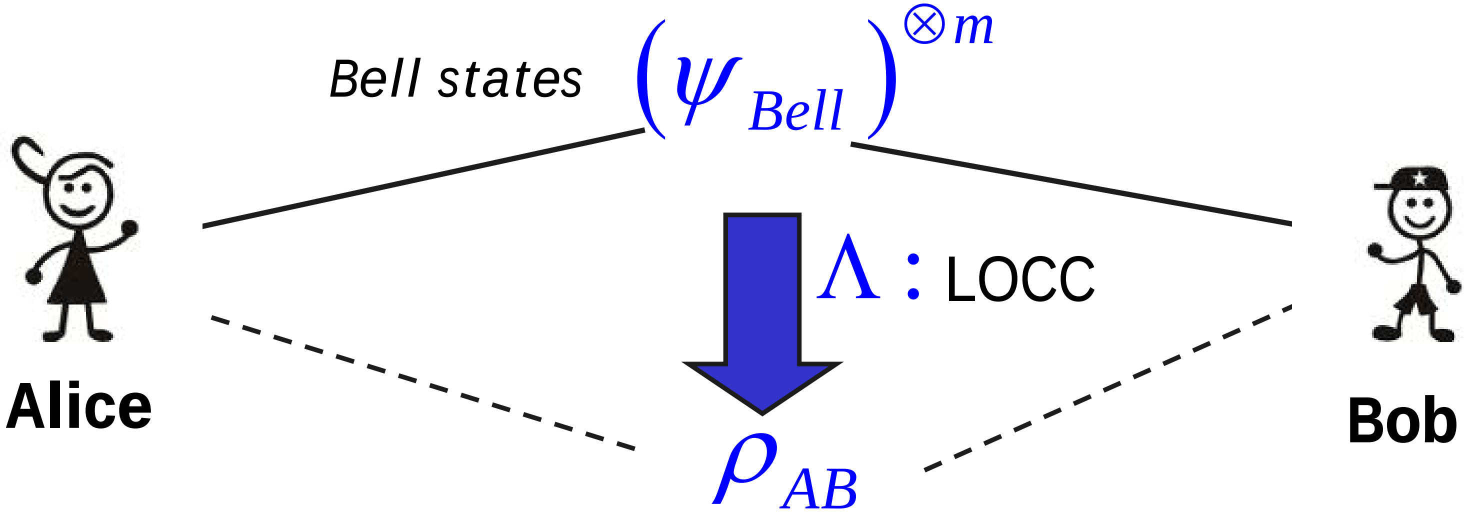
From the cryptographic point of view:

How random A *appears from the point of view of an adversary who has access to* B.

Operational interpretations contd.

- *Conditional 0-entropy* $\sim$ *one-shot entanglement cost*

One-shot Entanglement Dilution



One-shot entanglement cost

$$E_C^{(1)}(\rho_{AB}) := \min m$$

= *minimum number of Bell states needed to prepare a single copy of* ρ_{AB} *via LOCC*

Operational interpretations contd.

- *Theorem [F.Buscemi & ND]: One-shot perfect entanglement cost of a bipartite state* ρ_{AB} *under LOCC:*

$$E_C^{(1)}(\rho_{AB}) = \min_{\mathcal{E}} H_0(A|R)_{\rho^{\mathcal{E}}}$$

conditional 0-entropy

Pure-state ensembles:

$$\mathcal{E} = \left\{ p_i, \left|\psi_{AB}^i\right\rangle \right\}_i ; \quad \rho_{AB} = \sum_i p_i \left|\psi_{AB}^i\right\rangle\left\langle\psi_{AB}^i\right|$$

and $$\rho_{RAB}^{\mathcal{E}} = \sum_i p_i \left|i_R\right\rangle\left\langle i_R\right| \otimes \left|\psi_{AB}^i\right\rangle\left\langle\psi_{AB}^i\right|$$

classical-quantum state

$$\rho_{RA}^{\mathcal{E}} = \mathrm{Tr}_B \rho_{RAB}^{\mathcal{E}},$$

UNIVERSITY OF CAMBRIDGE

Outline

- *Mathematical Toolkit:*
  - *Semidefinite programming*

- *Definitions of generalized relative entropies:*

$$D_{\max}(\rho \| \sigma), D_0(\rho \| \sigma), D_{\min}(\rho \| \sigma)$$

- *Properties & operational significances of them*

and their children: the min-, max- and 0-entropies

- *Their "smoothed" versions*

- *The "super-parent" : Quantum Renyi Divergence* $\tilde{D}_{\alpha}(\rho \| \sigma)$

UNIVERSITY OF CAMBRIDGE

$D_{\max}(\rho \| \sigma)$ — *Max-relative entropy*

$D_0(\rho \| \sigma)$ — *0-relative Renyi entropy*

$D_{\min}(\rho \| \sigma)$ — *Min-relative entropy*

"super-parent": $\tilde{D}_\alpha(\rho \| \sigma)$

Quantum Renyi Divergence

(sandwiched Renyi entropy)

[Wilde et al; Muller-Lennert et al]

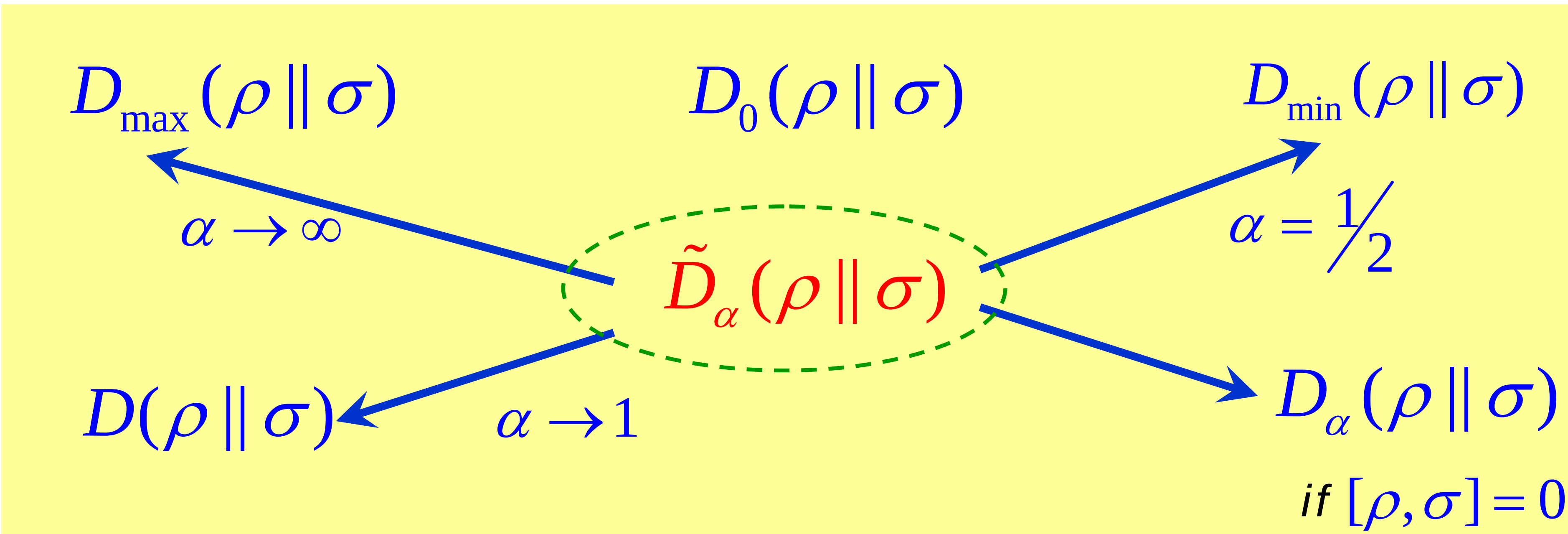
UNIVERSITY OF CAMBRIDGE

"super-parent":

Quantum Renyi Divergence

(sandwiched Renyi entropy)

[Wilde et al; Muller-Lennert et al]



Quantum Renyi Divergence

$$\rho \in \mathcal{D}(\mathcal{H}), \;\; \sigma \in \mathcal{P}(\mathcal{H}); \;\; \text{supp}\, \rho \subseteq \text{supp}\, \sigma;$$

For $\alpha \in (0,1) \cup (1,\infty):$

$$\tilde{D}_\alpha(\rho \| \sigma) := \frac{1}{\alpha - 1} \log\left(\text{Tr}\left((\sigma^\beta \rho \sigma^\beta)^\alpha\right)\right);$$

where $\beta = \frac{1-\alpha}{2\alpha}$ $\;\Rightarrow 2\beta\alpha = 1 - \alpha$

- **Note: If** $[\rho, \sigma] = 0$

$$\text{Tr}\left((\sigma^\beta \rho \sigma^\beta)^\alpha\right) = \text{Tr}\left((\rho \sigma^{2\beta})^\alpha\right) = \text{Tr}\left(\rho^\alpha \sigma^{2\beta\alpha}\right) = \text{Tr}\left(\rho^\alpha \sigma^{1-\alpha}\right)$$

$$\tilde{D}_\alpha(\rho \| \sigma) = \frac{1}{\alpha - 1} \log\, \text{Tr}\,(\rho^\alpha \sigma^{1-\alpha}) \;\; = D_\alpha(\rho \| \sigma)$$

α *-relative Renyi entropy*

UNIVERSITY OF CAMBRIDGE

Quantum Renyi Divergence

$\rho \in \mathcal{D}(\mathcal{H}),\ \sigma \in \mathcal{P}(\mathcal{H});\ \text{supp}\,\rho \subseteq \text{supp}\,\sigma;$

For $\alpha \in (0,1) \cup (1,\infty):$

$$\tilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\Big(\mathrm{Tr}\big((\sigma^\beta \rho \sigma^\beta)^\alpha\big)\Big);$$

where $\beta = \frac{1-\alpha}{2\alpha}$ $\Rightarrow 2\beta\alpha = 1-\alpha$

Non-commutative generalization of $D_\alpha(\rho\|\sigma)$

T $\sigma^{1-\alpha})$

$$\tilde{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\log\,\mathrm{Tr}\,(\rho^\alpha \sigma^{1-\alpha}) = D_\alpha(\rho\|\sigma)$$

α *-relative Renyi entropy*

UNIVERSITY OF CAMBRIDGE

Two properties of Quantum Renyi Divergence

(1) *Data-processing inequality; For* $\frac{1}{2} \le \alpha \le \infty,$

[Frank & Lieb; Beigi 2013]

$$\tilde{D}_\alpha(\Lambda(\rho) \| \Lambda(\sigma)) \le \tilde{D}_\alpha(\rho \| \sigma)$$

$\Lambda :$ *CPTP map*

$\therefore$ *holds also for* $\underset{\alpha = \frac{1}{2}}{D_{\min}(\rho \| \sigma)}, \; \underset{\alpha \to 1}{D(\rho \| \sigma)}, \; \underset{\alpha \to \infty}{D_{\max}(\rho \| \sigma)}$

UNIVERSITY OF CAMBRIDGE

Two properties of Quantum Renyi Divergence

(1) *Monotonicity under CPTP maps* Λ : *For* $\frac{1}{2} \le \alpha \le \infty,$

[Frank & Lieb; Beigi 2013]

$$\tilde{D}_\alpha(\Lambda(\rho) \| \Lambda(\sigma)) \le \tilde{D}_\alpha(\rho \| \sigma)$$

$\therefore$ *holds also for* $D_{\min}(\rho \| \sigma),\ D(\rho \| \sigma), D_{\max}(\rho \| \sigma)$

(2) *Joint convexity:* *For* $\frac{1}{2} \le \alpha \le 1,$ *[Frank & Lieb]*

$$\tilde{D}_\alpha\left(\sum_i p_i \rho_i \,\|\, \sum_i p_i \sigma_i\right) \le \sum_i p_i \tilde{D}_\alpha(\rho_i \| \sigma_i)$$

$\therefore$ *holds also for* $D_{\min}(\rho \| \sigma)\ ,\ D(\rho \| \sigma)$

- *Note:* $D_{\max}(\rho \| \sigma)$ *is quasi-convex:* *[ND]*

$$D_{\max}\left(\sum_i p_i \rho_i \,\|\, \sum_i p_i \sigma_i\right) \le \max_{1 \le i \le n} D_{\max}(\rho_i \| \sigma_i)$$

- *What about* $D_0(\rho \| \sigma)$?

UNIVERSITY OF CAMBRIDGE

Outline

- *Mathematical Toolkit:*
  - *Semidefinite programming*
- *Definitions of generalized relative entropies:*

$$D_{\max}(\rho \| \sigma), D_0(\rho \| \sigma), D_{\min}(\rho \| \sigma)$$

- *Properties & operational significances of them*
- *Their children: the min-, max- and 0-entropies*
- *Their "smoothed" versions*
- *The "super-parent" : Quantum Renyi Divergence* $\tilde{D}_\alpha(\rho \| \sigma)$
- *Relationship between* $\tilde{D}_\alpha(\rho \| \sigma)$ & $D_0(\rho \| \sigma)$ *& its implication*

UNIVERSITY OF CAMBRIDGE

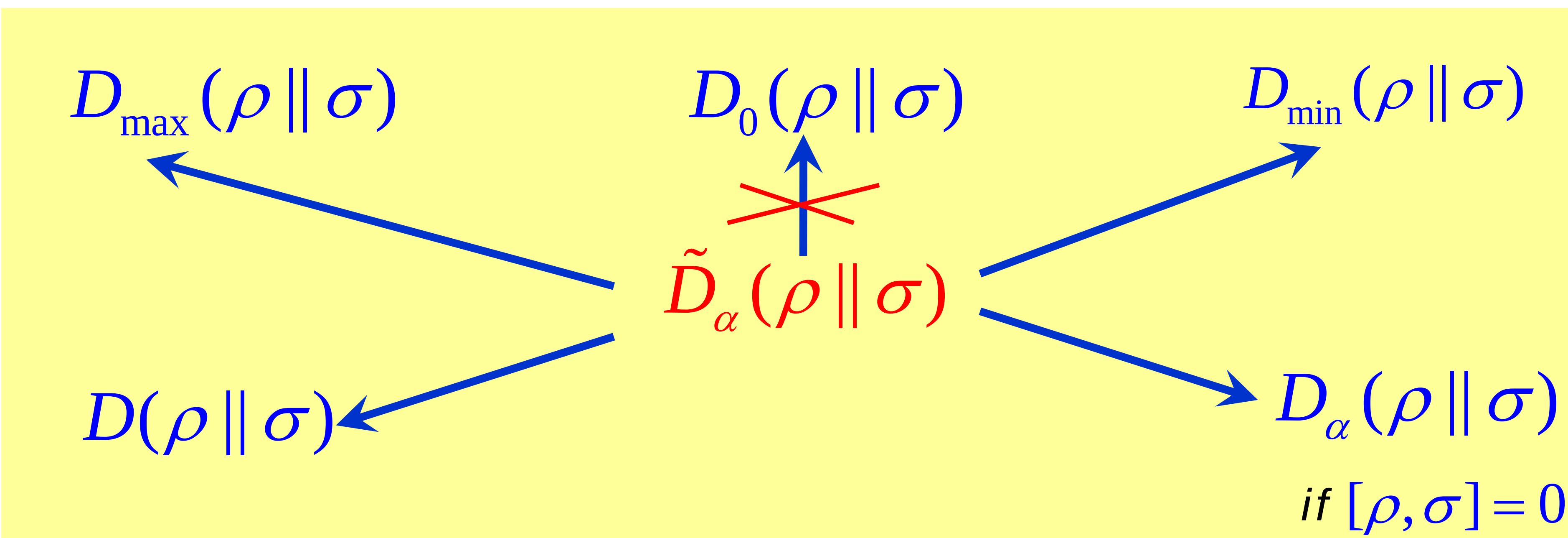
- *Theorem:* For $\rho, \sigma \geq 0, \mathrm{Tr}\,\rho = 1;$ *if* $\mathrm{supp}\,\rho = \mathrm{supp}\,\sigma;$

[ND, F.Leditzky] *then* $\lim_{\alpha \to 0} \tilde{D}_\alpha(\rho \| \sigma) = D_0(\rho \| \sigma) = 0 ....(1)$

If $\mathrm{supp}\,\rho \subset \mathrm{supp}\,\sigma,$ *then* (1) *does not necessarily hold*

$\tilde{D}_\alpha(\rho \| \sigma):$ *Quantum Renyi* Divergence

$D_0(\rho \| \sigma):$ *0-relative Renyi entropy*



UNIVERSITY OF CAMBRIDGE

- *Theorem:* For $\rho, \sigma \geq 0, \mathrm{Tr}\,\rho = 1;$ *if* $\mathrm{supp}\,\rho = \mathrm{supp}\,\sigma;$

[ND, F.Leditzky] *then* $\lim_{\alpha \to 0} \tilde{D}_\alpha(\rho \| \sigma) = D_0(\rho \| \sigma) = 0 .... (1)$

If $\mathrm{supp}\,\rho \subset \mathrm{supp}\,\sigma,$ *then* (1) *does not necessarily hold*

- *Proof (key steps):*

(1) If $\mathrm{supp}\,\rho \subseteq \mathrm{supp}\,\sigma: \ \tilde{D}_\alpha(\rho \| \sigma) \leq D_\alpha(\rho \| \sigma) \ \forall \alpha > 0.$

$$\Rightarrow \lim_{\alpha \to 0} \tilde{D}_\alpha(\rho \| \sigma) \leq \lim_{\alpha \to 0} D_\alpha(\rho \| \sigma) = D_0(\rho \| \sigma) ..... (a)$$

(Araki-Lieb-Thirring inequality)

(2) If $\mathrm{supp}\,\rho = \mathrm{supp}\,\sigma: \ \lim_{\alpha \to 0} \tilde{D}_\alpha(\rho \| \sigma) \geq D_0(\rho \| \sigma) ... (b)$

$(a) \,\&\, (b) \Rightarrow (1)$ *if* $\mathrm{supp}\,\rho = \mathrm{supp}\,\sigma;$

(a variant of the Pinching lemma)

UNIVERSITY OF CAMBRIDGE

- **Proof of the fact:** *If* $\text{supp}\,\rho \subset \text{supp}\,\sigma$, *then*

$$\lim_{\alpha\to 0} \tilde{D}_\alpha(\rho\|\sigma) = D_0(\rho\|\sigma)$$ *does not necessarily hold*

A simple counterexample:

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad \sigma = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix} \qquad c \in (0,1).$$

$$\rho, \sigma \geq 0, \quad [\rho, \sigma] \neq 0.$$

$$D_0(\rho\|\sigma) = 0$$

$$\lim_{\alpha\to 0} \tilde{D}_\alpha(\rho\|\sigma) = -\log(1+c) < 0;$$

■

UNIVERSITY OF CAMBRIDGE

Summary

- Generalized relative entropies:

$$D_{\max}(\rho\|\sigma), D_0(\rho\|\sigma), D_{\min}(\rho\|\sigma)$$

- *Properties & some operational significances*

$D_{\max}(\rho\|\sigma)$: *in multiple state discrimination,*

$D_0(\rho\|\sigma)$: *in binary hypothesis testing*

- *Their "smoothed" versions;* $D_0^{\varepsilon}(\rho\|\sigma) \equiv D_H^{\varepsilon}(\rho\|\sigma)$
- *Their children: the min-, max- and 0-entropies*
- *Operational significances of conditional entropies*
- *The "super-parent" : Quantum Renyi Divergence* $\tilde{D}_{\alpha}(\rho\|\sigma)$
- *Relationship between* $\tilde{D}_{\alpha}(\rho\|\sigma)$ & $D_0(\rho\|\sigma)$

UNIVERSITY OF CAMBRIDGE

Thank you!

- *Thanks also to:*

F.Buscemi, F.Brandao, M-H.Hsieh, F.Leditzky, M.Mosonyi, R.Renner, T.Rudolph,