Role of Entropies in Quantum Communication

LECTURE II

Nilanjana Datta
University of Cambridge,U.K.

In the last lecture we saw that:

In Quantum information theory, initially one evaluated:

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

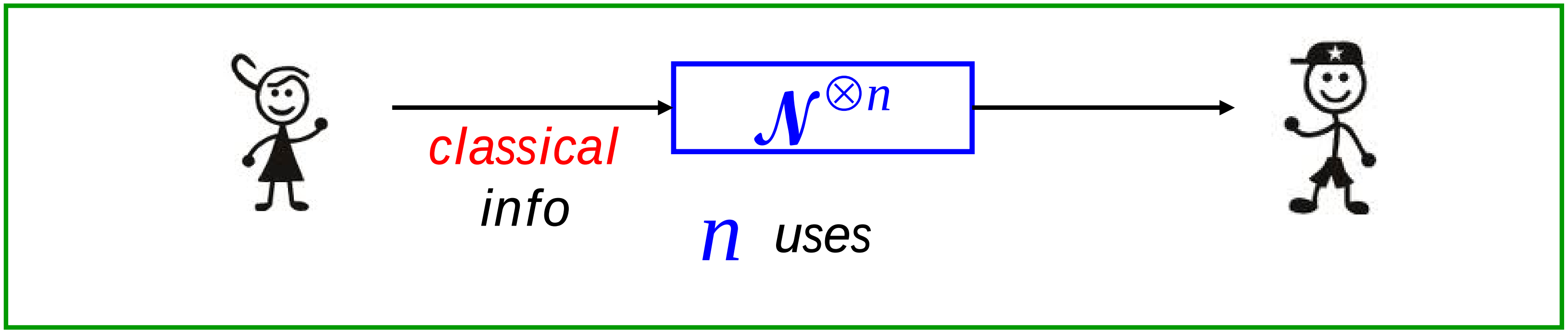
under the assumption of an *"asymptotic, memoryless setting"*

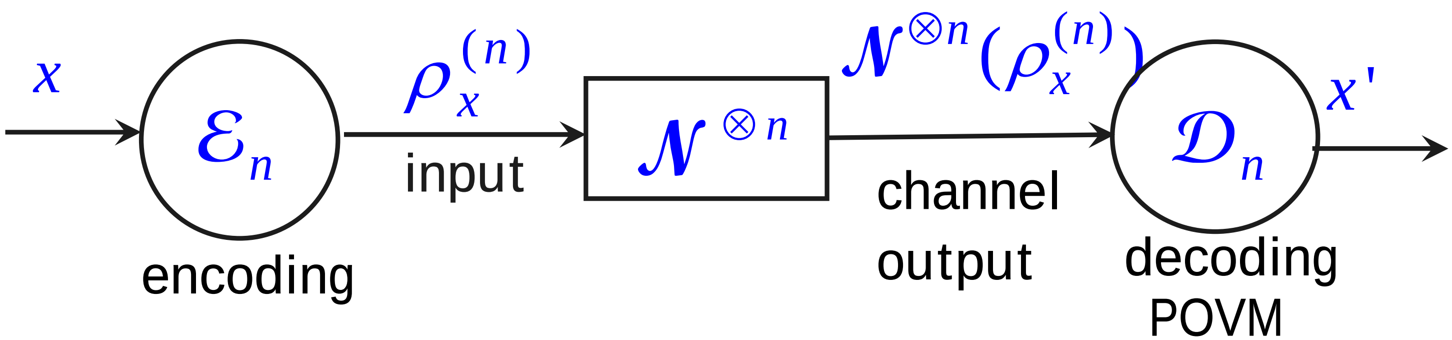
Assume:

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

"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 $\rho, \sigma \geq 0; \quad \mathrm{Tr}\rho = 1$

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

e.g. Data compression limit:

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

e.g. Holevo quantity:

$$\chi\left(\{p_x, \rho_x\}\right) = \sum_x p_x D(\rho_x \| \rho); \; \rho = \sum_x p_x \rho_x$$

etc.

acts as a parent quantity for optimal rates in the "asymptotic, memoryless setting"

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)*

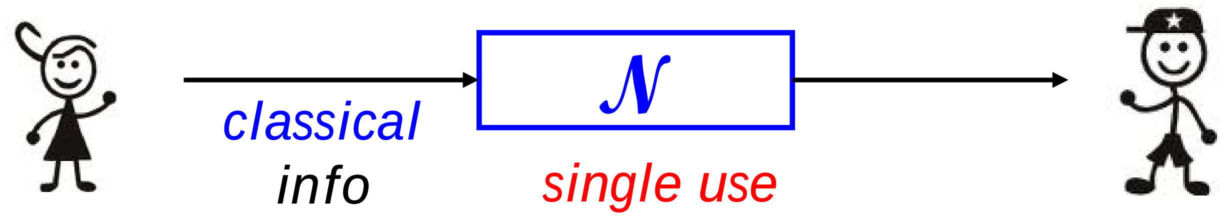
Hence it is important to evaluate optimal rates for *finite number of uses* *(or even a single use)* *of an arbitrary source or channel*

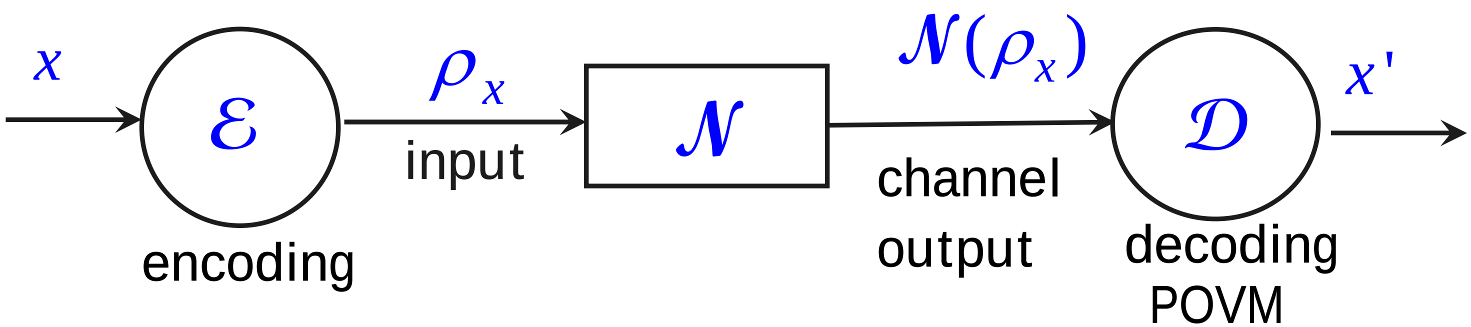
- Evaluation of corresponding optimal 'rates':

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* $\mathcal{N}$

$$C_{\varepsilon}^{(1)}(\mathcal{N})$$

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

- Obtainable from certain (generalized) relative entropies

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

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

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

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

- **Rest of this lecture:**

Part I

Entropies relevant in One-Shot Information Theory

Part II

These entropies as operational quantities in One-Shot Information Theory

UNIVERSITY OF CAMBRIDGE

Part I

Entropies relevant in One-Shot Information Theory

Outline

- *Notations & Definitions*
- *Tool: Decoupling*
- *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*

UNIVERSITY OF CAMBRIDGE

Notations & Definitions

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

(finite-dimensional)

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

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

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

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

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

- Quantum operations (quantum channels) : linear CPTP map

Λ is CPTP if and only if Λ^* is CPUM

completely positive unital map: $\Lambda^*(I) = I$

UNIVERSITY OF CAMBRIDGE

Notations & Definitions

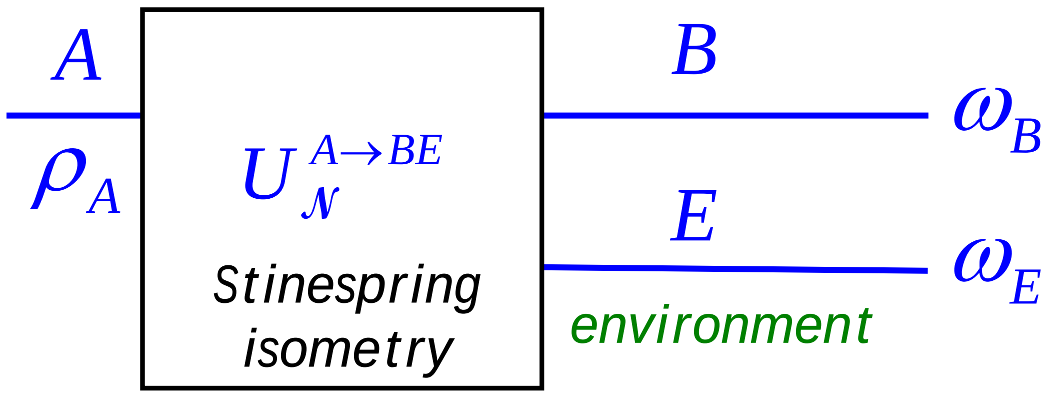
- **Quantum channel :** $\mathcal{N}^{A\to B}$.

- **Stinespring isometry** of $\mathcal{N}$: $U_{\mathcal{N}}^{A\to BE}$

$$\omega_B := \mathcal{N}^{A\to B}(\rho_A) = \mathrm{Tr}_E U_{\mathcal{N}}^{A\to BE}(\rho_A)$$

- **Complementary channel**: $\tilde{\mathcal{N}}^{A\to E}$,

$$\omega_E := \tilde{\mathcal{N}}^{A\to E}(\rho_A) = \mathrm{Tr}_B U_{\mathcal{N}}^{A\to BE}(\rho_A)$$



UNIVERSITY OF CAMBRIDGE

Notations & Definitions

- A figure of merit in quantum communication tasks:

- **Fidelity:** *For* $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $F(\rho,\sigma) := \| \sqrt{\rho}\sqrt{\sigma} \|_1$

$$F(\rho,\sigma) = F(\sigma,\rho); \quad 0 \le F(\rho,\sigma) \le 1$$

For 2 pure states ψ, ϕ: $F(\psi,\phi) = |\langle \psi | \phi \rangle|$

$$F(\psi,\rho) = \sqrt{\mathrm{Tr}(\rho\psi)}; \quad \therefore F^2(\psi,\rho) = \mathrm{Tr}(\rho\psi) = \langle \psi | \rho | \psi \rangle$$

- **Uhlmann's Theorem:**

$$F(\rho,\sigma) = \max_{\psi_\rho, \psi_\sigma} |\langle \psi_\rho | \psi_\sigma \rangle|,$$ ψ_ρ, ψ_σ : *purifications of* ρ, σ.

$$F(\rho,\sigma) \le F(\Lambda(\rho), \Lambda(\sigma)) \quad \forall \Lambda \text{ CPTP}$$

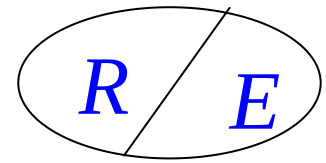
Mathematical Tool

Decoupling: -- a central concept in quantum info theory

- Has wide-ranging applications:
  - transmission of quantum information
  - other protocols, e.g. state merging, coherent state merging, ……

UNIVERSITY OF CAMBRIDGE

Mathematical Tool



Decoupling:

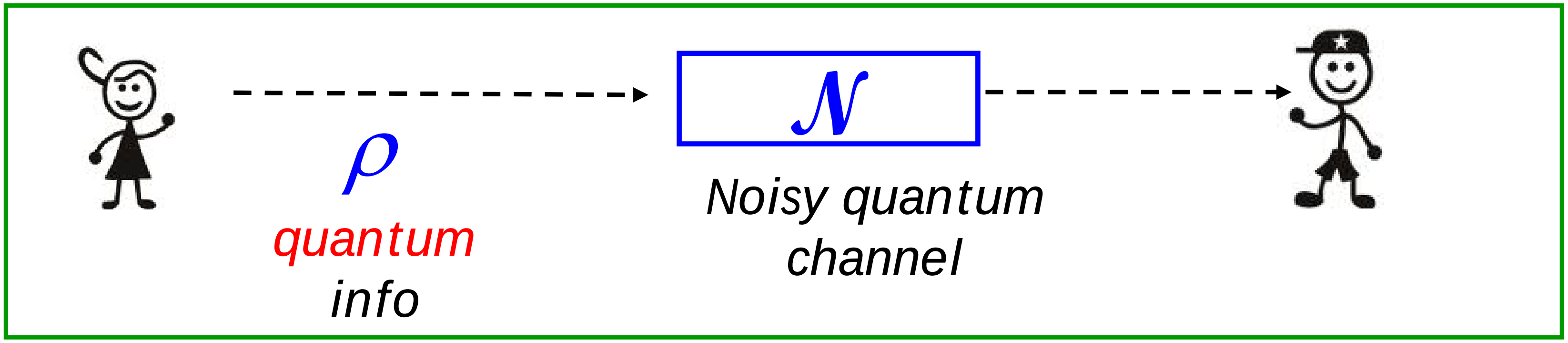
- Consider a composite system RE in a joint state ω_{RE}
- The subsystem R is decoupled (or uncorrelated) from E

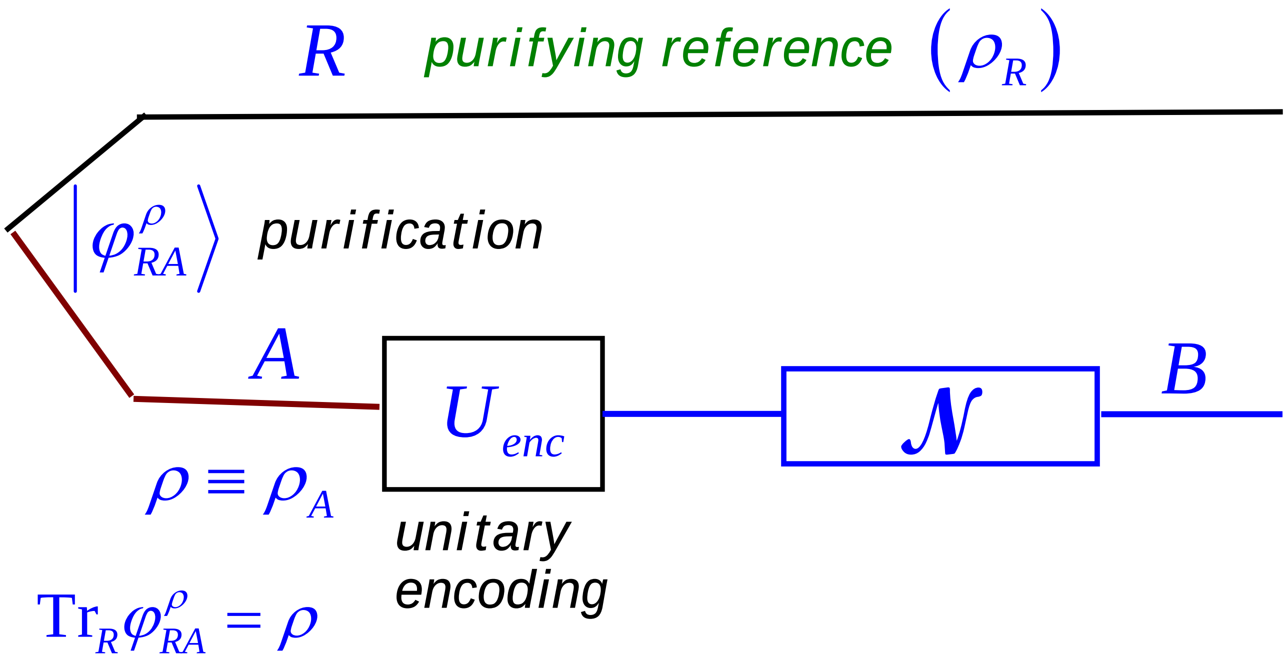
  if:

$$\omega_{RE} = \rho_R \otimes \sigma_E$$

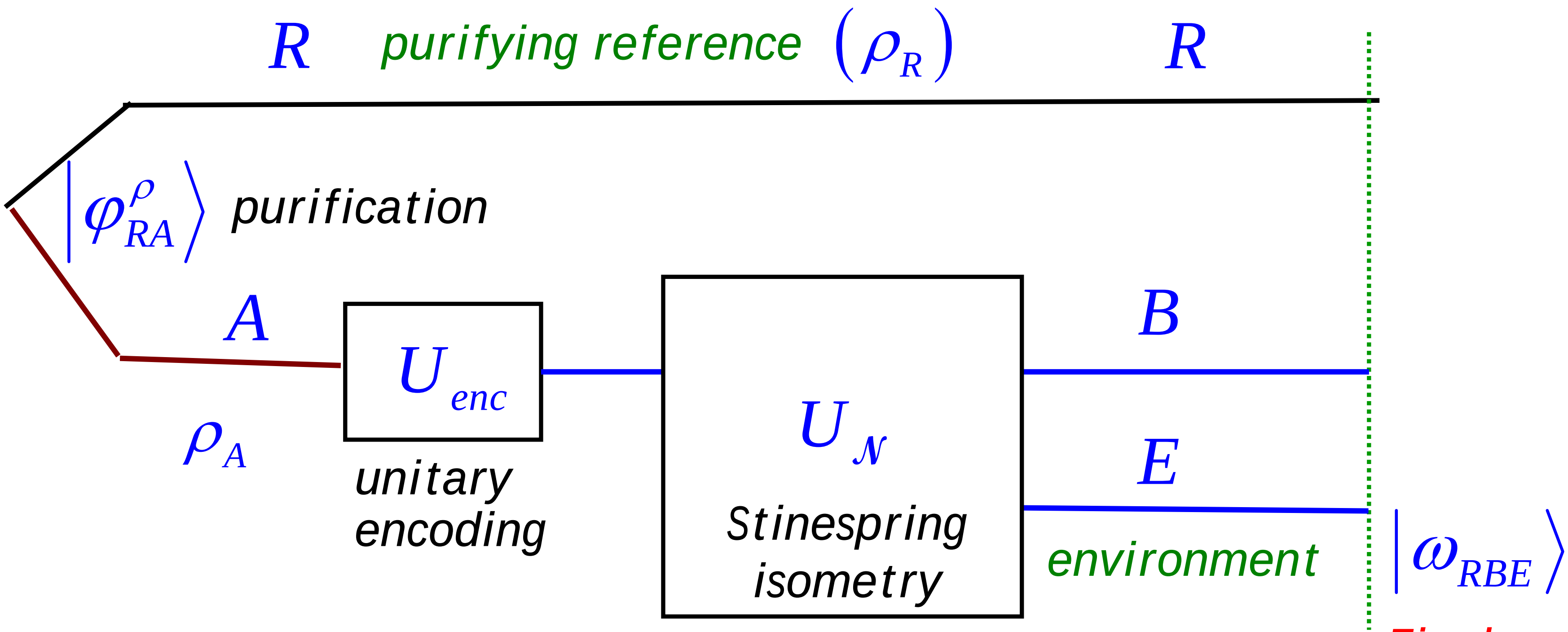
- The outcome of any measurement on R is statistically independent of any measurement on E
- The system R does not give any information about system E

(I) Transmission of quantum information




R
purifying reference (ρ_R)
$|\varphi^{\rho}_{RA}\rangle$ purification
A
U_{enc}
$\mathcal{N}$
B
$\rho \equiv \rho_A$
unitary encoding
$\mathrm{Tr}_R \varphi^{\rho}_{RA} = \rho$

UNIVERSITY OF CAMBRIDGE

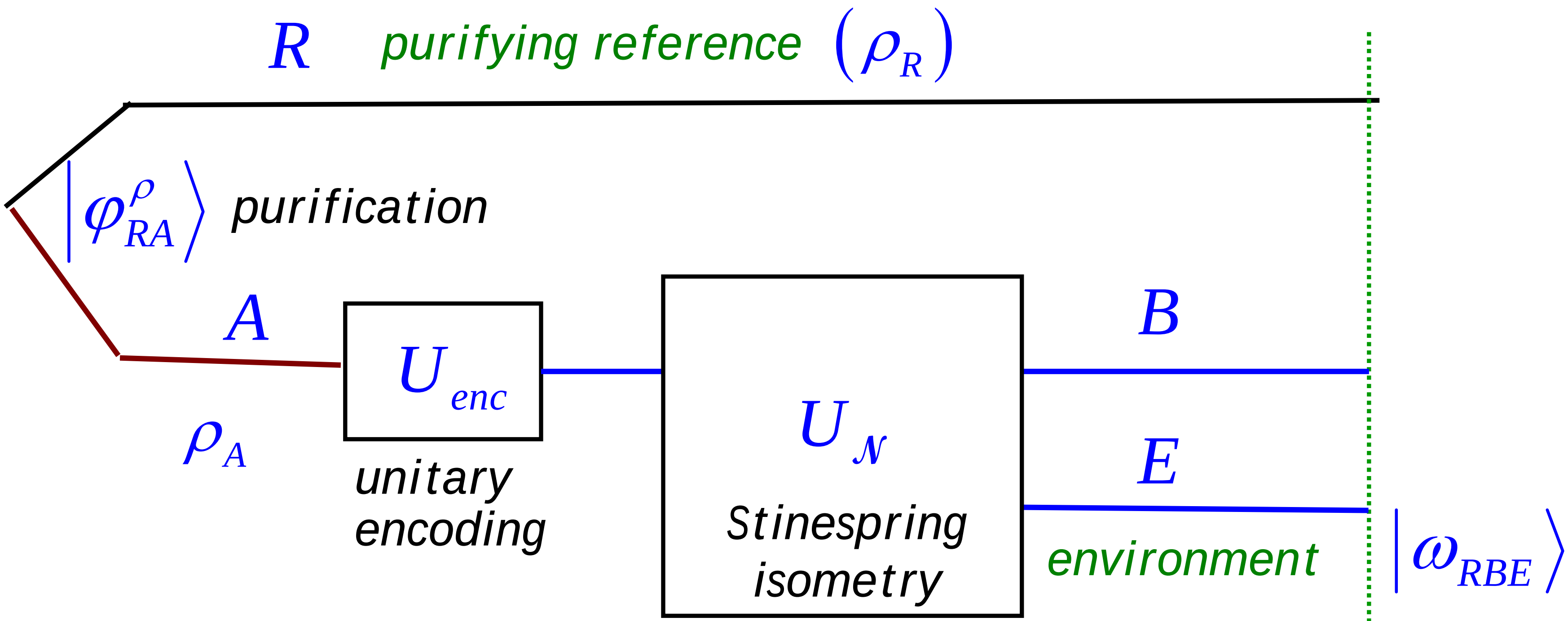


- *Let* U_{enc} *be such that* ω_{RE} *satisfies:*

$$\omega_{RE} = \rho_R \otimes \sigma_E$$

(decoupled)

for some state σ_E



Final pure state

- Let U_{enc} be such that ω_{RBE} satisfies:

$$\omega_{RE} = \rho_R \otimes \sigma_E$$

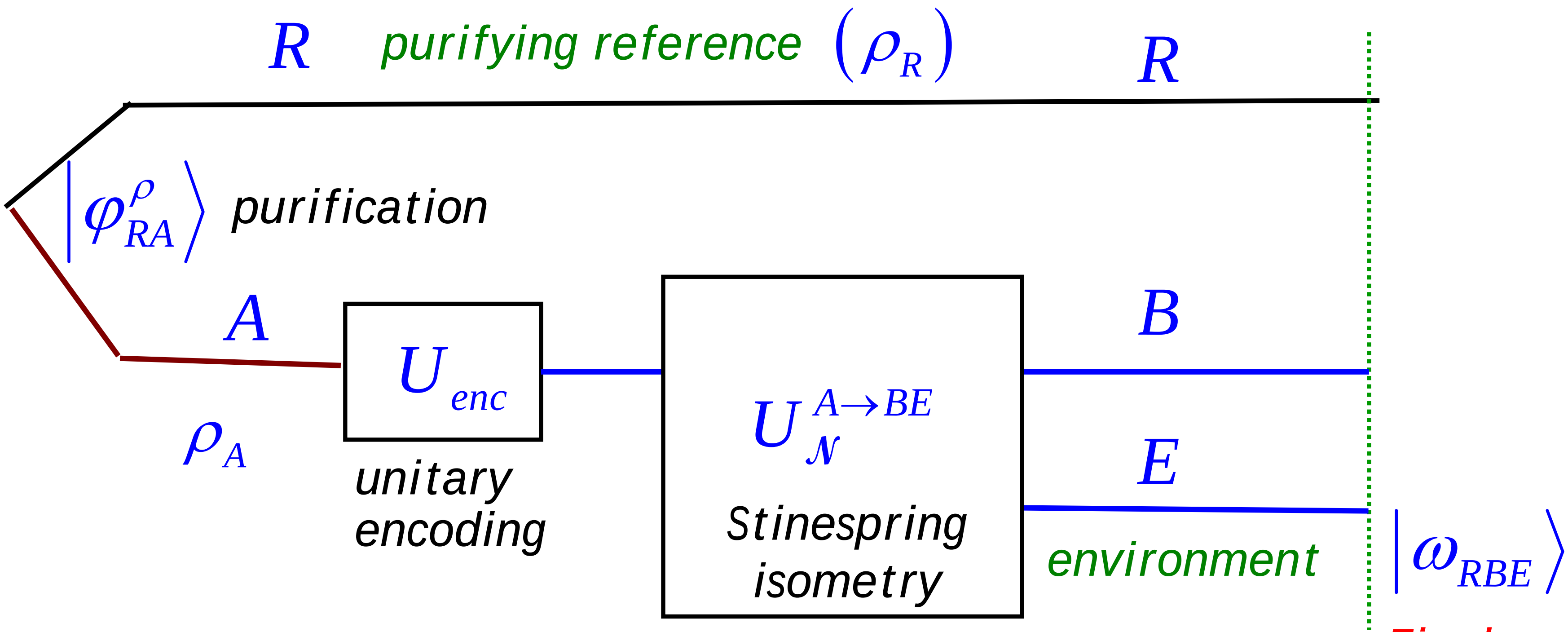
(decoupled)

for some state σ_E

purifications

$|\omega_{RBE}\rangle$ ⟷ $|\varphi^{\rho}_{RA}\rangle \otimes |\sigma_{EE'}\rangle$:

related by a partial isometry



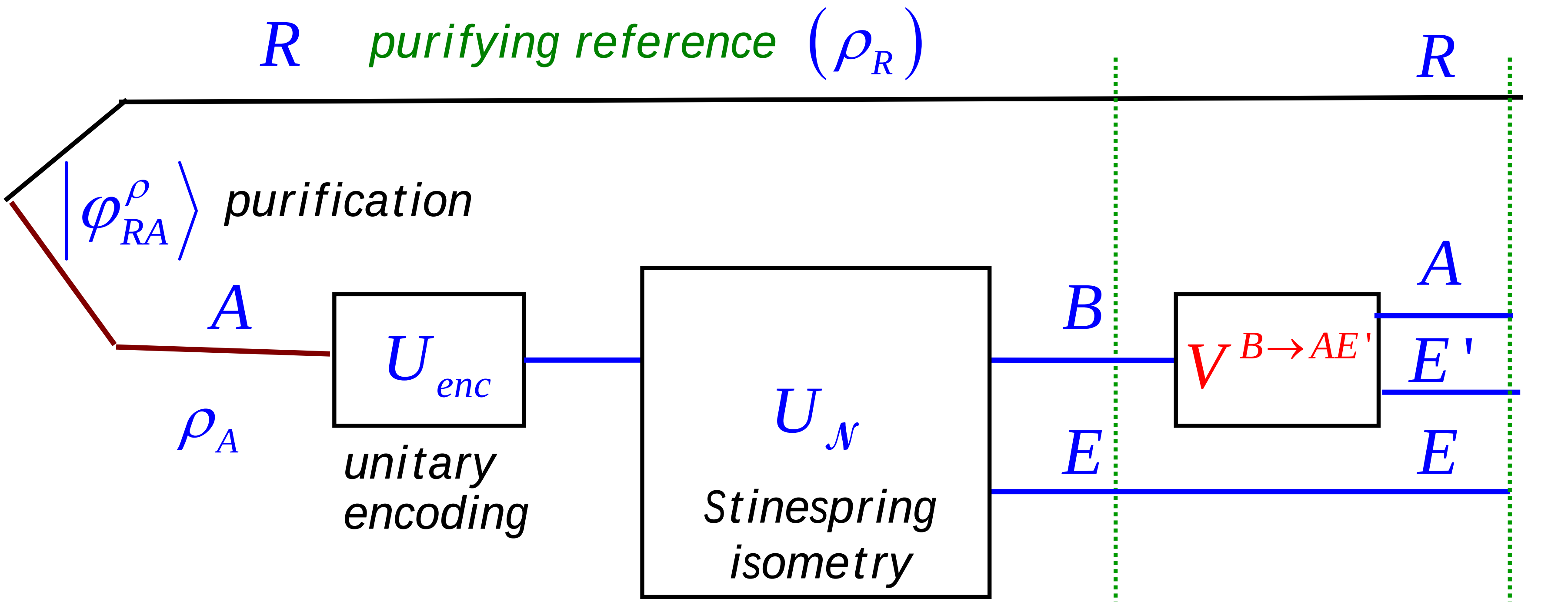
- Let U_{enc} be such that ω_{RBE} satisfies:

$$\omega_{RE} = \rho_R \otimes \sigma_E$$ (decoupled)

$\exists$ a partial isometry $V^{B\to AE'}$ such that

$$V^{B\to AE'}|\omega_{RBE}\rangle = |\varphi^{\rho}_{RA}\rangle \otimes |\sigma_{EE'}\rangle$$

This acts as Bob's decoding!

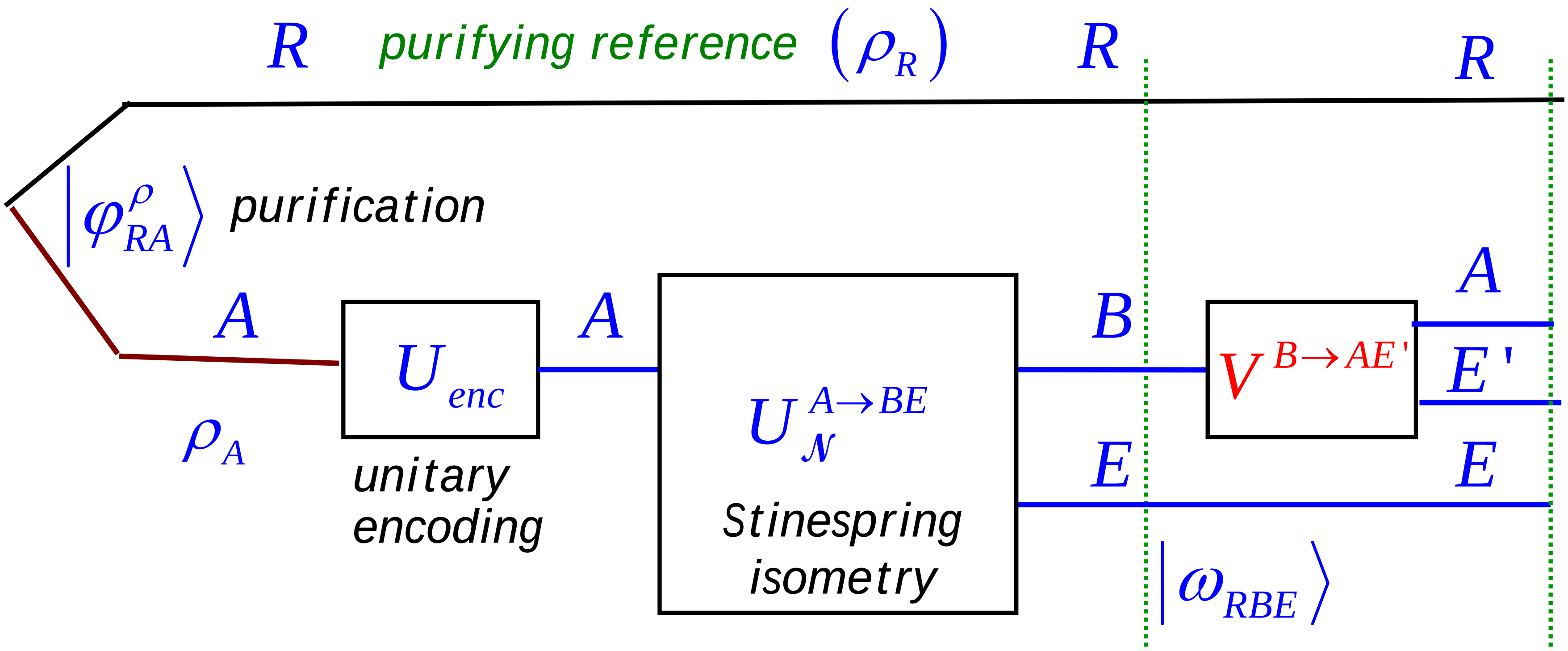


$$V^{B\to AE'}|\omega_{RBE}\rangle = |\varphi^{\rho}_{RA}\rangle \otimes |\sigma_{EE'}\rangle$$

- *Final state in Bob's possession:* $\mathrm{Tr}_{RE}\left(\varphi^{\rho}_{RA} \otimes \sigma_{EE'}\right) = \rho_A \otimes \sigma_{E'}$
- *Bob traces out over the system* E':

  $\mathrm{Tr}_{E'}\left(\rho_A \otimes \sigma_{E'}\right) = \rho_A$ *to recover Alice's message !*

UNIVERSITY OF CAMBRIDGE



Thus: If U_{enc} be such that ω_{RE} is decoupled:

$$\omega_{RE} = \rho_R \otimes \sigma_E$$

then Bob can recover Alice's message!

UNIVERSITY OF CAMBRIDGE

- *In fact, if* $\omega_{RE} \overset{\varepsilon}{\approx} \rho_R \otimes \sigma_E$ *(approximately decoupled)*

that is, $F\left(\omega_{RE}, \rho_R \otimes \sigma_E\right) \geq 1-\varepsilon$ *for some* $\varepsilon \geq 0$:

then $\exists$ *a decoder such that after decoding Bob has*

a state $\overset{\varepsilon}{\approx} \rho_A$ *(Alice's message)*

- *This follows from Uhlmann's theorem:*

Let $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A)$, *purifications* $\left|\varphi_{AR}^{\rho}\right\rangle, \left|\psi_{AR'}^{\sigma}\right\rangle$

$$F\left(\rho, \sigma\right) = \max_{V^{R\to R'}} \left|\left\langle \psi_{AR'}^{\sigma} \right| V^{R\to R'} \left| \varphi_{AR}^{\rho} \right\rangle\right|$$

$$1-\varepsilon \leq F\left(\omega_{RE}, \rho_R \otimes \sigma_E\right) = \max_{V^{B\to AE'}} \left|\left\langle \varphi_{RA}^{\rho} \otimes \sigma_{EE'} \right| V^{B\to AE'} \left| \omega_{RBE} \right\rangle\right|$$

UNIVERSITY OF CAMBRIDGE

$$1-\varepsilon \le F\left(\omega_{RE}, \rho_R \otimes \sigma_E\right) = \max_{V^{B\to AE'}} \left| \left\langle \varphi^{\rho}_{RA} \otimes \sigma_{EE'} \right| V^{B\to AE'} \left| \omega_{RBE} \right\rangle \right|$$

The optimizing partial isometry $V^{B\to AE'}$ *acts as Bob's decoding*

Bob ends up with a state $\overset{\varepsilon}{\approx} \mathrm{Tr}_{RE}\left(\varphi^{\rho}_{RA} \otimes \sigma_{EE'}\right) \overset{\varepsilon}{\approx} \rho_A \otimes \omega_{E'}$

And after doing a partial trace over E', *he ends up with*

a state $\overset{\varepsilon}{\approx} \rho_A$ *(Alice's message)*

i.e., Bob ends up with a state which is $\varepsilon -$*close*

to the quantum state that Alice sent

UNIVERSITY OF CAMBRIDGE

- *In a nutshell:*

For transmission of quantum information thro' a noisy channel $\mathcal{N}$ *in the one-shot setting (up to an error* ε *), require:*

$$\omega_{RE} \overset{\varepsilon}{\approx} \rho_R \otimes \sigma_E$$

(state before decoding)

i.e., the state of the reference system R *is (approxly.) decoupled from the state of the environment* E *of* $\mathcal{N}$.

Outline

- *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 2008]*

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

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

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

- *Min-relative entropy [Dupuis et al 2012]*

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

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

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

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

- *Proof:*

$$D_{\max}(\rho \| \sigma) := \inf\left\{\gamma : \rho \leq 2^{\gamma} \sigma\right\} \quad = \gamma_0$$

$$\rho \leq 2^{\gamma_0} \sigma, \qquad (2^{\gamma_0}\sigma - \rho) \geq 0, \quad \textit{Also} \quad \pi_\rho \geq 0$$

$$\mathrm{Tr}\,[\pi_\rho (2^{\gamma_0}\sigma - \rho)] \geq 0 \qquad \because A, B \geq 0 \Rightarrow \mathrm{Tr}\,(AB) \geq 0$$

$$2^{\gamma_0}\,\mathrm{Tr}\,[\pi_\rho \sigma] \geq \mathrm{Tr}\,[\pi_\rho \rho] = 1$$

$$\gamma_0 + \log\,[\mathrm{Tr}(\pi_\rho \sigma)] \geq 0$$

$$\gamma_0 \geq -\log\,[\mathrm{Tr}(\pi_\rho \sigma)]$$

$$D_{\max}(\rho \| \sigma) \quad \geq \quad D_0(\rho \| \sigma)$$

Properties of generalized relative entropies

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

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

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

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,*

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

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*

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

$\beta = \mathrm{Tr}(A\sigma)$ *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 $\beta = 2^{-D_0(\rho\|\sigma)}$ *when* $\alpha = 0$

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

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

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

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 if Bob has a single copy of the state but one allows non-zero but small value of the *Prob(Type I error)* α?

i.e., let $\alpha \leq \varepsilon$ *for some* $\varepsilon \geq 0.$

$$D_0(\rho \| \sigma) = -\log \beta^* |_{\alpha=0} = -\log \operatorname{Tr} \pi_\rho \sigma$$

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

$\alpha = \operatorname{Tr}((I - A)\rho);\ \alpha = 0$ for $A = \pi_\rho \therefore \operatorname{Tr}(A\rho) = 1$

$\therefore$ For $\alpha \leq \varepsilon$ choose A such that $\operatorname{Tr}(A\rho) \geq 1 - \varepsilon$

$$D_0^\varepsilon(\rho \| \sigma) = -\log \beta^* |_{\alpha \leq \varepsilon} = \max_{\substack{0 \leq A \leq I \\ \operatorname{Tr}(A\rho) \geq 1-\varepsilon}} \left(-\log(\operatorname{Tr}(A\rho))\right)$$

Hypothesis testing relative entropy [Wang & Renner]

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

UNIVERSITY OF CAMBRIDGE

$$D_0(\rho \| \sigma) = -\log \beta^* |_{\alpha=0} = -\log \ \mathrm{Tr}\ \pi_\rho \sigma$$

$$\beta^* |_{\alpha=0} = \mathrm{Tr}\ \pi_\rho \sigma$$

$$= \min_{\substack{0 \le A \le I \\ \mathrm{Tr}(A\rho) = 1}} \mathrm{Tr}(A\sigma) = 2^{-D_0(\rho\|\sigma)}$$

$$\beta^* |_{\alpha \le \varepsilon} = \min_{\substack{0 \le A \le I \\ \mathrm{Tr}(A\rho) \ge 1-\varepsilon}} \mathrm{Tr}(A\sigma) = 2^{-D_H^\varepsilon(\rho\|\sigma)}$$

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

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

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

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

$$\beta^{*(n)}\big|_{\alpha(n)\le\varepsilon} \approx 2^{-nD(\rho\|\sigma)}$$

[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:

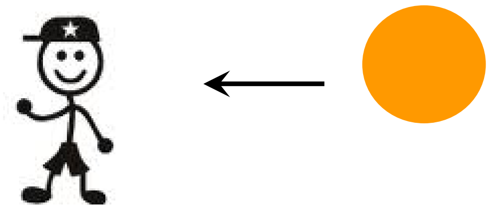
$$= \lim_{n \to \infty} \left\{ -\frac{1}{n} \log \beta^{*(n)} |_{\alpha(n) \le \varepsilon} \right\}$$

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

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

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:

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

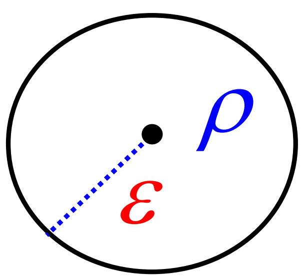
Smooth max-relative entropy

$\forall \varepsilon \geq 0.$

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

$$B^{\varepsilon}(\rho) := \left\{ \bar{\rho} \geq 0, \mathrm{Tr}\bar{\rho} = 1 : \sqrt{1 - F(\rho, \bar{\rho})} \leq \varepsilon \right\}$$

→ *fidelity*



UNIVERSITY OF CAMBRIDGE

Outline

- *Mathematical Tool:* *Decoupling*

- *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*

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

as parent quantities for other entropies

Just as:

von Neumann entropy

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

$$(\sigma = I)$$

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

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

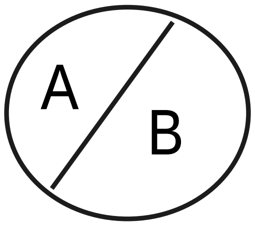
$$H_{\max}(\rho) := -D_{\min}(\rho \| I) = \log \| \sqrt{\rho} \|_1^2$$

[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): $S(A|B)_\rho = -S(A|C)_\rho$

-- 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 interpretation

- *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 **(MES)**

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

[Koenig, Renner, Schaffner]

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

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]

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

fidelity

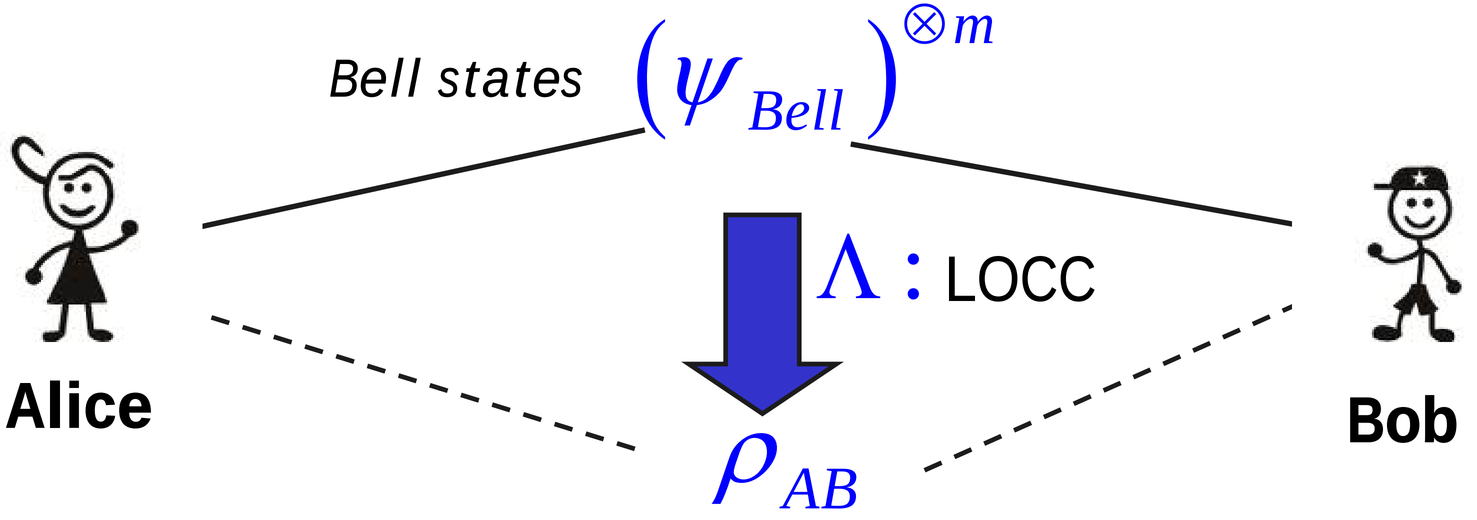
From the cryptographic point of view:

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

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

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

UNIVERSITY OF CAMBRIDGE

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

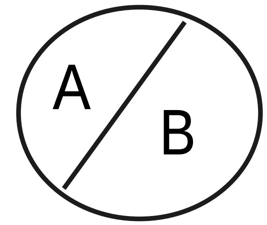
classical-quantum state

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

UNIVERSITY OF CAMBRIDGE

Other min- & max- entropies contd.

For a bipartite state ρ_{AB} :



- just as: *Mutual information*

$$I(A:B) = D(\rho_{AB} \,||\, \rho_A \otimes \rho_B) = \max_{\sigma_B} D(\rho_{AB} \,||\, \rho_A \otimes \sigma_B)$$

Max-mutual entropy

$$I_{\max}(A:B)_\rho := \max_{\sigma_B} D_{\max}(\rho_{AB} \,||\, \rho_A \otimes \sigma_B)$$

etc.

Smoothed entropies $\forall \varepsilon \geq 0.$

$$H^{\varepsilon}_{\min}(A|B)_\rho, H^{\varepsilon}_{\max}(A|B)_\rho, H^{\varepsilon}_{0}(A|B)_\rho, I^{\varepsilon}_{\max}(A:B)_\rho$$

UNIVERSITY OF CAMBRIDGE

PROOF OF:

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

Equivalently,

$$-H_{\min}(A|B)_\rho = \log\left(d \max_{\Lambda_B:CPTP} \mathrm{Tr}\left(((\mathrm{id}_A \otimes \Lambda_B)\rho_{AB})\Phi_{AB}\right)\right) \quad \text{....(a)}$$

Proof via SDP (=semidefinite programming)

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*

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: $\alpha \quad = \quad \beta$

IF *Slater's duality condition holds.*

UNIVERSITY OF CAMBRIDGE

PROOF OF:

$$-H_{\min}(A|B)_\rho = \log\left(d \max_{\Lambda_B:CPTP} \mathrm{Tr}\left(((\mathrm{id}_A \otimes \Lambda_B)\rho_{AB})\Phi_{AB}\right)\right)$$ ....(a)

Proof via SDP

- LHS of **(a)** $= \log\left(\min \mathrm{Tr}\tilde{\sigma}_B; (\mathrm{id}_A \otimes \tilde{\sigma}_B) \geq \rho_{AB}; \tilde{\sigma}_B \geq 0\right)$ ...(i)

- RHS of **(a)** $= \log\left(\min \mathrm{Tr}(\rho_{AB} Y_{AB}); \mathrm{Tr}_A Y_{AB} \leq I_B; Y_{AB} \geq 0\right)$ ...(ii)

(i)=(ii) *(details given in the lecture)*

UNIVERSITY OF CAMBRIDGE

Part II

Smooth entropies as operational quantities in One-Shot Information Theory

- *Consider quantum communication tasks in the the*

  one-shot setting

- *See how……*

  - *some of the smooth entropies that we discussed arise as operational quantities for these tasks.*

  - *the known results for the asymptotic memoryless setting can be obtained from these one-shot results.*

Smooth entropies

- *Relative entropies* $D_{\max}(\rho\|\sigma), D_0(\rho\|\sigma), D_{\min}(\rho\|\sigma)$

  -- their smoothed versions $D^{\varepsilon}_{\max}(\rho\|\sigma), D^{\varepsilon}_{H}(\rho\|\sigma).....$

- *Min-/max- entropies* $H_{\min}(A|B)_\rho, H_0(\rho), H_{\max}(A|B)_\rho$ etc.

  -- their smoothed versions $H^{\varepsilon}_{\min}(A|B)_\rho, H^{\varepsilon}_{\max}(A|B)_\rho,....$

(Smooth) Entropies: properties

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

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

$$H^\varepsilon_{\min}(A|B)_\rho := \max_{\bar\rho \in B^\varepsilon(\rho)} H_{\min}(A|B)_{\bar\rho};$$

$$H^\varepsilon_{\max}(A|B)_\rho := \min_{\bar\rho \in B^\varepsilon(\rho)} H_{\max}(A|B)_{\bar\rho};$$

- If $\rho_{RA} = \Phi^m_{RA}$; MES $|\Phi^m_{RA}\rangle = \frac{1}{\sqrt{m}}\sum_{i=1}^m |i\rangle|i\rangle$

$$H^\varepsilon_{\min}(A|R)_\rho \geq \quad H_{\min}(A|R)_\rho = -\log m = H_{\max}(A|R)_\rho$$

UNIVERSITY OF CAMBRIDGE

(Smooth) Entropies: properties

Duality of smoothed min- and max- entropies: *[Colbeck, Renner Tomamichel]*

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

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

Data-processing inequality:

- e.g. If $\tilde{\omega}_{RA} = (\mathrm{id}_R \otimes \Lambda^{B\to A})\omega_{RB}$

(quantum operation)

$$H^{\varepsilon}_{\max}(R|B)_{\omega} \le H^{\varepsilon}_{\max}(R|A)_{\tilde{\omega}}$$

One-shot to asymptotics

- *Relation between smooth entropies & quantum entropies*

$$\forall \varepsilon > 0: \quad \lim_{n\to\infty} \frac{1}{n} D^{\varepsilon}_{\max}(\rho^{\otimes n} \| \sigma^{\otimes n}) \quad = D(\rho \| \sigma)$$

[Audenaert, Mosonyi, Verstraete ; Tomamichel; ND &Renner]

QAEP

$$\forall \varepsilon > 0: \lim_{n\to\infty} \frac{1}{n} H^{\varepsilon}_{\min}(A|B)_{\rho_{AB}^{\otimes n}} = H(A|B)_{\rho}$$

$$\forall \varepsilon > 0: \lim_{n\to\infty} \frac{1}{n} H^{\varepsilon}_{\max}(A|B)_{\rho_{AB}^{\otimes n}} = H(A|B)_{\rho}$$

[Colbeck, Renner, Tomamichel; Tomamichel]

These results allow us to recover the results of the
"asymptotic memoryless setting"
from those of the "one-shot setting"