Role of Entropies in Quantum Communication

LECTURE I

Nilanjana Datta
University of Cambridge,U.K.

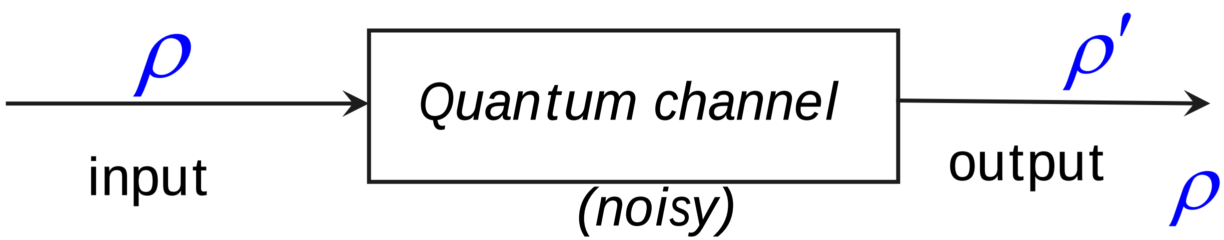
See lecture notes on: http://www.qi.damtp.cam.ac.uk/node/223

"Quantum communication is the art of transferring a quantum state from one place to another." [Gisin]

- quantum states encode information – classical or quantum;
- quantum communication allows transmission of information

The main hurdle in the path of quantum communication:

- Presence of noise in the quantum channel
- Disturbs the quantum state sent through the quantum channel
- Distorts the information encoded in the state



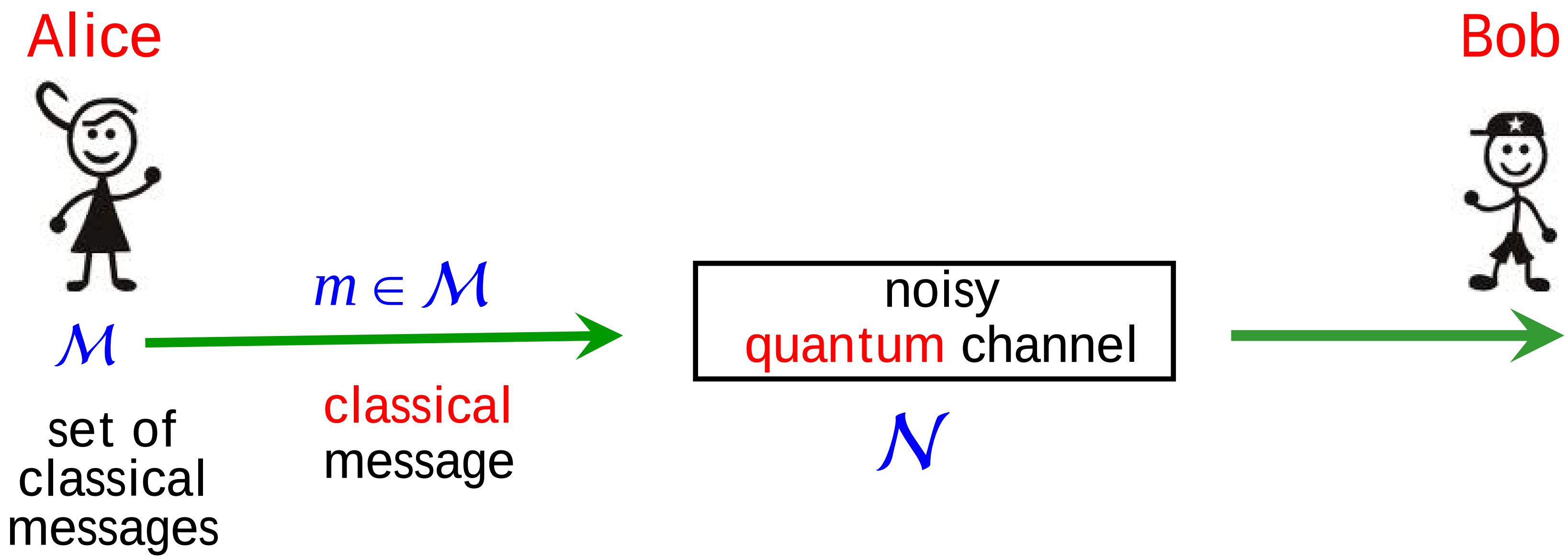
$$\rho' \neq \rho$$

UNIVERSITY OF CAMBRIDGE

To overcome the effects of noise

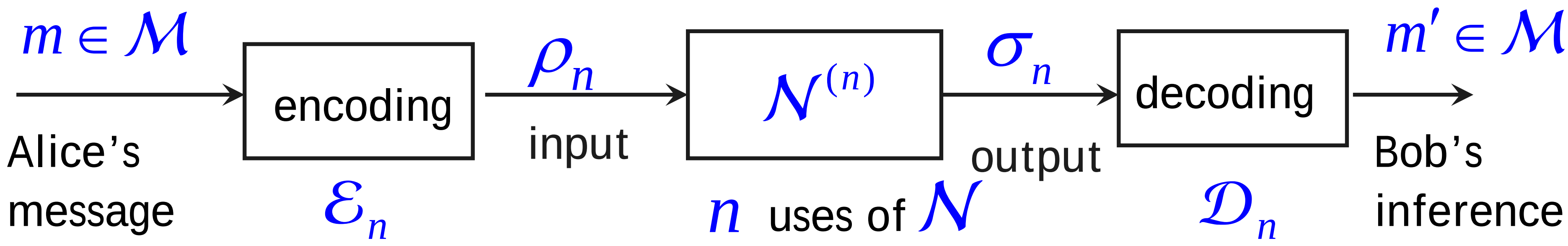
use quantum error-correcting codes

e.g. consider the case of classical information transmission;



- Alice encodes her messages into suitable quantum states (codewords)
- she sends these codewords through (multiple uses of) the channel

UNIVERSITY OF CAMBRIDGE



- *If* $m' \neq m$ *then an error occurs!*
- *Reliability: e.g. if Probability of error* $\to 0$ *as* $n \to \infty$

- *Rate of info transmission* = number of bits of message transmitted per use of the channel

- *Aim: achieve reliable transmission whilst maximizing the rate*
- *There is a fundamental limit on the rate of reliable info transmission (depends on the channel)*

(a property of the channel)
Classical capacity of the quantum channel $\mathcal{N}$ $:=$ *max. rate of reliable transmission of classical info through* $\mathcal{N}$

- An important class of problems in QIT concerning the **transmission of information** through quantum channels:

  evaluating the capacities of a quantum channel

Another essential task in QIT :

- Efficient **storage of information** emitted by a quantum info source:

  This involves reliable compression of quantum info

  i.e. Quantum Data Compression

  - Why do we need to compress information?
  - What is meant by "reliable"?
  - What is a quantum info source?

It can also be viewed as a quantum communication task

UNIVERSITY OF CAMBRIDGE

- Quantum info source: characterized by an ensemble

$\mathcal{E} = \{p_i, |\psi_i\rangle\}$ of pure states $|\psi_i\rangle \in \mathcal{H}$

source ensemble & a priori probs p_i

$|\psi_i\rangle$: signal emitted with prob. p_i

In general $\langle \psi_i | \psi_j \rangle \neq \delta_{ij}$

source state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$

- Equivalently the source is characterized by $\{\rho, \mathcal{H}\}$

UNIVERSITY OF CAMBRIDGE

Quantum Data Compression

- Storage setting:

quantum info source — ψ_i signal → $\mathcal{E}$ encoding → $\mathcal{E}(\psi_i)$ compressed signal → $\mathcal{D}$ decoding → recovered signals

$$\psi_i \equiv |\psi_i\rangle\langle\psi_i| \in \mathcal{D}(\mathcal{H}); \qquad \mathcal{E}(\psi_i) \in \mathcal{D}(\mathcal{H}_c)$$

set of density matrices acting on $\mathcal{H}$

$$\dim \mathcal{H}_c < \dim \mathcal{H}$$

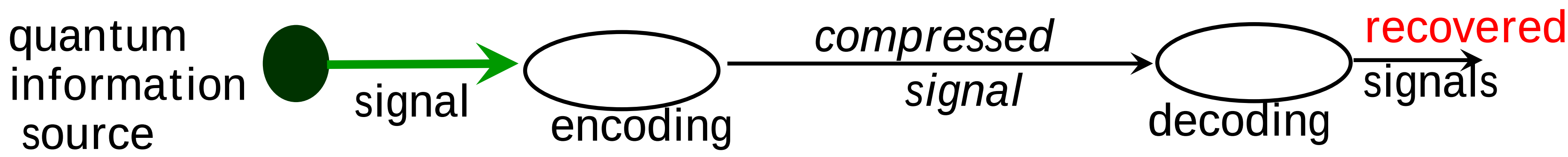
$\log(\dim \mathcal{H}_c)$: number of qubits needed to compress the signals

Quantity of interest: minimum number of qubits needed to compress the signals

Quantum Data Compression as quantum communication

- Storage setting:

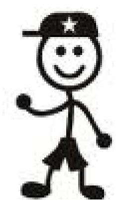
minimum number of qubits Alice needs to compress the signals

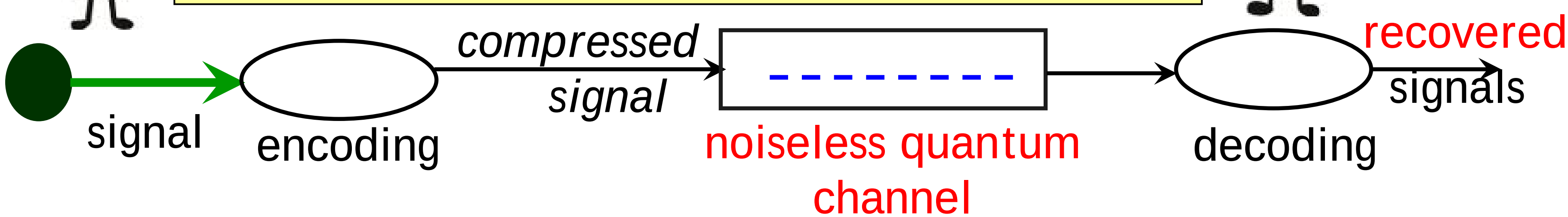


An Equivalent Scenario for Quantum Data Compression

- Communication setting:

minimum number of qubits that Alice needs to send to Bob through the noiseless channel

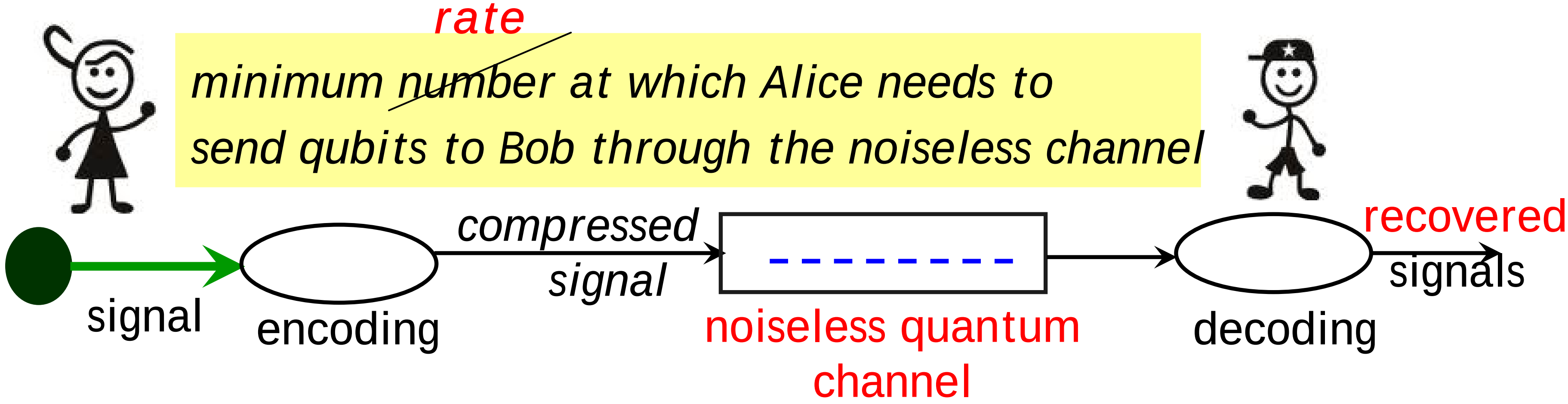




Quantum Data Compression as quantum communication

Usually one uses the source a multiple number (n) of times

- **Communication setting:**



UNIVERSITY OF CAMBRIDGE

The fundamental operational quantities:

optimal rates of info-processing tasks

- **Info transmission:** maximum rate of reliable info transmission through a noisy quantum channel

  capacity *(of the channel)*

- **Storage of information (data compression) :**

  minimum rate of reliable info transmission through a noiseless quantum channel

  data compression limit

  (of the source)

- **Aim:** to evaluate these optimal rates :

i.e. find mathematical expressions for them in terms of

entropic quantities

UNIVERSITY OF CAMBRIDGE

These optimal rates were initially evaluated under the assumption of an:

"asymptotic, memoryless setting"

- info sources & channels are assumed to be memoryless
- they are used an infinite number of times:

  (asymptotic limit) $n \to \infty$

- one requires that the error incurred vanishes in this limit

e.g. Memoryless channel

n successive uses : $\boxed{\mathcal{N}^{(n)} = \mathcal{N}^{\otimes n}}$

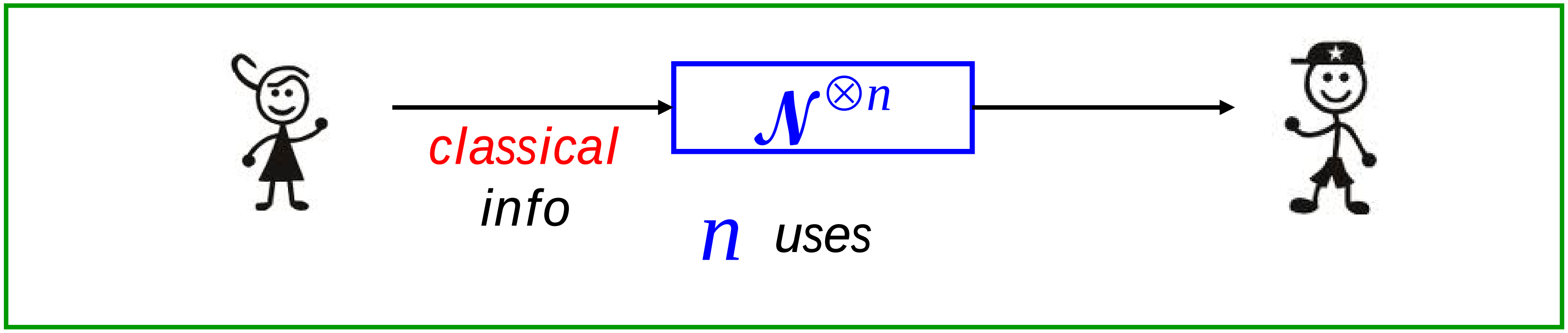
- action of each use of the channel : identical & independent for different uses

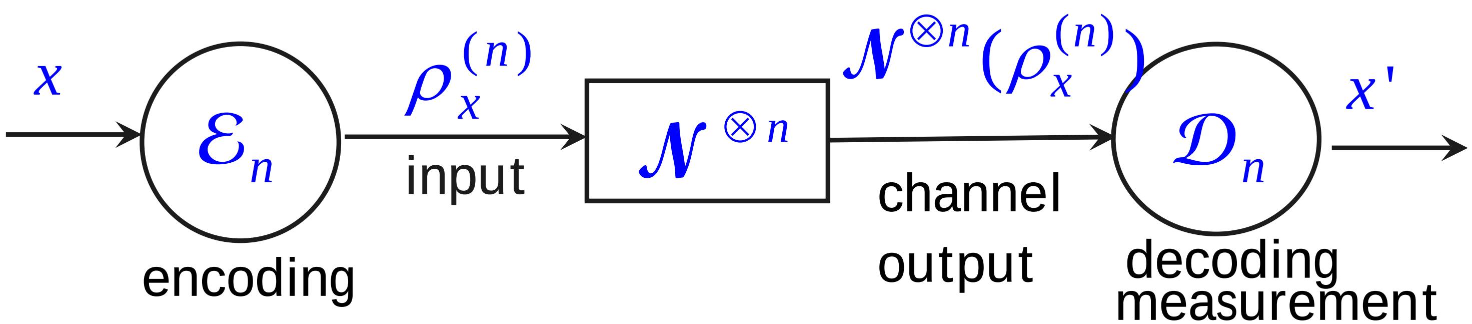
-- the noise affecting successive input states uncorrelated.

UNIVERSITY OF CAMBRIDGE

"asymptotic, memoryless setting"

- e.g. *To evaluate* $C(\mathcal{N})$: *classical capacity of a noisy quantum channel* $\mathcal{N}$





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

Outline of the rest of the lecture

- Recall some standard definitions
- Define the relevant entropic quantities
- Description of the info-processing tasks in more detail
- Statement of results expressing

  optimal rates in terms of entropic quantities
- Sketch of some proofs

- **Standard Definitions**

UNIVERSITY OF CAMBRIDGE

Notations & Definitions

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

quantum system *Hilbert space*

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

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

set of density matrices (states)

$$\mathcal{D}(\mathcal{H}) = \{\rho \in \mathcal{L}(\mathcal{H}) : \rho \geq 0,\ \mathrm{Tr}\,\rho = 1\}$$

- Spectral decomposition:

$$\lambda_i \geq 0, \quad \sum_{i=1}^{d} \lambda_i = 1$$

$$\rho = \sum_{i=1}^{d} \lambda_i |\varphi_i\rangle\langle\varphi_i|;$$

eigenvalues eigenvectors

$\{\lambda_i\}_{i=1}^{d}$: probability distribution

UNIVERSITY OF CAMBRIDGE

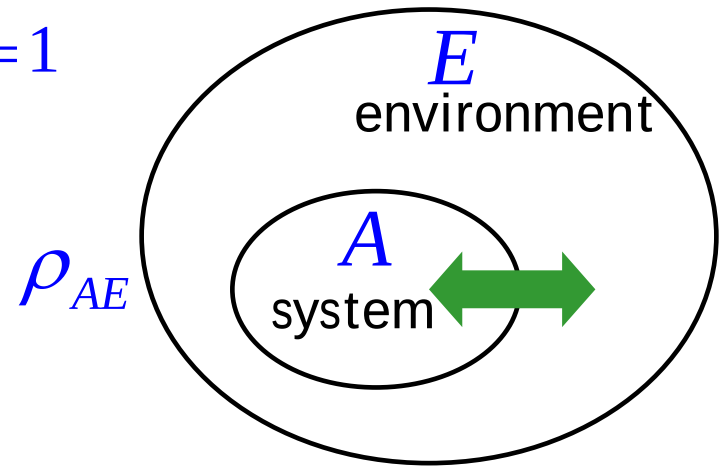
Quantum Operations or Quantum Channels

- Any allowed physical process that a quantum system can undergo is described by a :

linear completely-positive, trace preserving (CPTP) map

$$\rho \xrightarrow{\text{input}} \boxed{\mathcal{N} \text{ : CPTP map}} \xrightarrow{\text{output}} \rho' = \mathcal{N}(\rho)$$

- *Trace-preserving (TP):* $\mathrm{Tr}\,\rho' = \mathrm{Tr}\rho = 1$
- *Positive:* $\rho' = \mathcal{N}(\rho) \geq 0$
- *Completely positive (CP):*



$$\mathcal{N} : \mathcal{D}(\mathcal{H}_A) \to \mathcal{D}(\mathcal{H}_B)$$

$(\mathcal{N} \otimes id_E)(\rho_{AE})$ = *an allowed state of the composite system* $\in \mathcal{D}(\mathcal{H}_B \otimes \mathcal{H}_E)$

$$\boxed{(\mathcal{N} \otimes id_E)(\rho_{AE}) \geq 0}$$

UNIVERSITY OF CAMBRIDGE

Generalized measurements – POVM:

A quantum measurement is described by a **POVM**

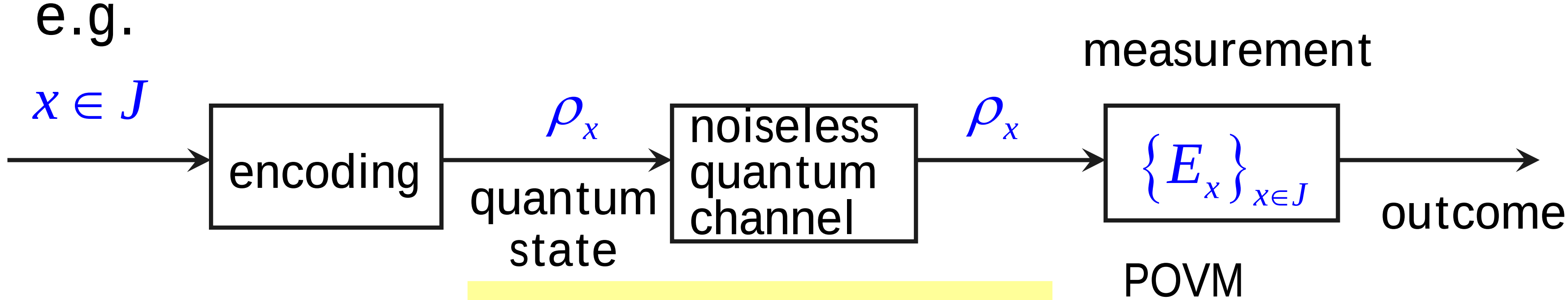
$E = \{E_i\};$ (finite set) $\quad E_i \geq 0, \ \sum_i E_i = I$

If the system is in a state ρ before the measurement,

Then, probability of getting the i^{th} outcome is:

$$p_i = \mathrm{Tr}(E_i \rho)$$

- e.g.



$$p(y \mid x) = \mathrm{Tr}(E_y \rho_x)$$

UNIVERSITY OF CAMBRIDGE

Purification

Any mixed state

$$\rho_A \in \mathcal{H}_A \longleftrightarrow |\psi_{AR}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_R;$$

A pure state

$$\rho_A = \mathrm{Tr}_R |\Psi_{AR}\rangle\langle\Psi_{AR}|;$$

purifying reference system

Schmidt decomposition: Any pure bipartite state

$$|\psi_{AR}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_R;$$

$$|\psi_{AR}\rangle = \sum_{i=1}^{d} \lambda_i |i_A\rangle |i_B\rangle; \qquad \lambda_i \geq 0, \sum_{i=1}^{d} \lambda_i^2 = 1$$

Consequences: Reduced states,

$$\rho_A := \mathrm{Tr}_R |\psi_{AR}\rangle\langle\psi_{AR}|, \qquad \rho_R := \mathrm{Tr}_A |\psi_{AR}\rangle\langle\psi_{AR}|$$

have identical non-zero eigenvalues

- **Entropic Quantities**

UNIVERSITY OF CAMBRIDGE

Von Neumann entropy

of a state ρ:

$$\boxed{S(\rho) := -\mathrm{Tr}\,(\rho \log \rho)}$$

$\log \equiv \log_2$

Spectral decomposition: $\rho = \sum_{i=1}^{d} \lambda_i |\varphi_i\rangle\langle\varphi_i|;$

Shannon entropy

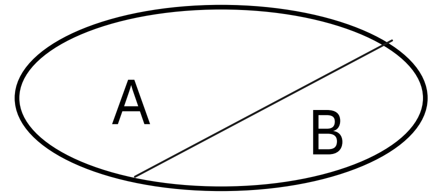
$$S(\rho) := -\mathrm{Tr}\,(\rho \log \rho) = -\sum_{i=1}^{d} \lambda_i \log \lambda_i = H(\{\lambda_i\})$$

- $S(\rho) = 0$ if and only if ρ is a pure state: $\rho = |\Psi\rangle\langle\Psi|$
- $S(\rho) \geq 0;\quad S(\rho) \leq \log d;$ *where* $d = \dim \mathcal{H}$
- $S(\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2)$
- If ρ_A, ρ_B are reduced states of a pure state ψ_{AB} then,

$$S(\rho_A) = S(\rho_B)$$

Other Entropies

For a bipartite system in a state ρ_{AB} :



- **Joint entropy:**

$$\boxed{S(\rho_{AB}) = -\mathrm{Tr}(\rho_{AB} \log \rho_{AB})}$$

- **Conditional entropy:**

$$\boxed{S(A|B)_{\rho} := S(\rho_{AB}) - S(\rho_B)}$$

$$\rho_B = \mathrm{Tr}_A \rho_{AB}$$

reduced state

- **Quantum mutual information:**

$$\boxed{I(A:B)_{\rho} := S(\rho_A) + S(\rho_B) - S(\rho_{AB});}$$

- Quantum Relative Entropy

of ρ w.r.t. σ, $\rho \geq 0$, $\mathrm{Tr}\,\rho = 1$, $\sigma \geq 0$:

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

well-defined if $\mathrm{supp}\,\rho \subseteq \mathrm{supp}\,\sigma$

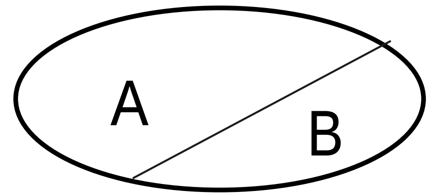
- It acts as a parent quantity for the *von Neumann entropy*:

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

UNIVERSITY OF CAMBRIDGE

- It also acts as a parent quantity for other entropies:

e.g. for a bipartite state ρ_{AB} :



- *Conditional entropy*

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

- *Mutual information*

$$\rho_B = \mathrm{Tr}_A \, \rho_{AB}$$

$$\boxed{I(A:B) := S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = D(\rho_{AB} \| \rho_A \otimes \rho_B)}$$

UNIVERSITY OF CAMBRIDGE

Some Properties of $D(\rho \| \sigma)$

$$D(\rho \| \sigma) \geq 0 \qquad \rho, \sigma \text{ states}$$

$$= 0 \text{ if \& only if } \rho = \sigma \qquad \text{..........(1)}$$

"distance"

~~symmetric~~
~~triangle inequality~~

Data-processing inequality:

i.e. monotonicity under a quantum operation (CPTP map)

$$D(\Lambda(\rho) \| \Lambda(\sigma)) \leq D(\rho \| \sigma) \qquad \text{.........(2)}$$

This is a fundamental property ;

quantum operations never increase mutual information

Many properties of other entropies can be proved using (1) & (2)

- e.g. *If* $\sigma_{AB'} = (\mathrm{id}_A \otimes \Lambda_{B \to B'})\rho_{AB}$ *then* $I(A:B')_\sigma \leq I(A:B)_\rho$

UNIVERSITY OF CAMBRIDGE

Further Properties of $D(\rho \| \sigma)$

- Joint convexity:

For two mixtures of states $\rho = \sum_{i=1}^{n} p_i \rho_i$ & $\sigma = \sum_{i=1}^{n} p_i \sigma_i$

$$\boxed{D\left(\sum_k p_k \rho_k \,\|\, \sum_k p_k \sigma_k\right) \le \sum_k p_k D(\rho_k \,\|\, \sigma_k) \text{....(a)}}$$

- Invariance under joint unitaries

$$\boxed{D(U\rho U^\dagger \,\|\, U\sigma U^\dagger) = D(\rho \,\|\, \sigma) \text{....(b)}}$$

Implications for the von Neumann entropy:

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

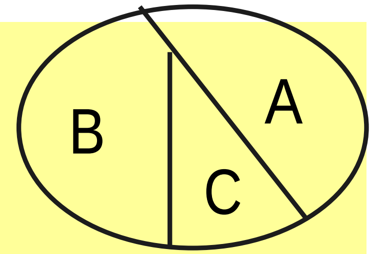
(a) $\Rightarrow$ *Concavity:* $S\left(\sum_i p_i \rho_i\right) \ge \sum_i p_i S(\rho_i)$

(b) $\Rightarrow$ *Invariance under unitaries:* $S(U\rho U^\dagger) = S(\rho)$

UNIVERSITY OF CAMBRIDGE

Properties of quantum entropies contd.

- Strong subadditivity: ρ_{ABC} tripartite state



$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$$

Lieb & Ruskai '73

Consequences of strong subadditivity:

- Conditioning reduces entropy

$$S(A|BC)_\rho \leq S(A|B)_\rho$$

- Discarding quantum systems never increases mutual information

$$I(A:B)_\rho \leq I(A:BC)_\rho$$

- **Description of the info-processing tasks in more detail**

-- in the "asymptotic, memoryless setting"

Quantum Data Compression

Quantum Info source → *signals*

signals (*pure states*) $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_r\rangle \in \mathcal{H}$

with probabilities $p_1, p_2, \ldots, p_r$ $\langle \psi_i | \psi_j \rangle \neq \delta_{ij}$

- Then source characterized by: $\{\rho, \mathcal{H}\}$

$$\boxed{\rho = \sum_{i=1}^{r} p_i |\psi_i\rangle\langle\psi_i|}$$

density matrix

- Memoryless quantum information source

State of n copies of the source: $\rho_n = \rho^{\otimes n}$

no correlation

Quantum data compression

- Evaluated in the asymptotic limit $n \to \infty$

  $n =$ number of copies/uses of the source

- emits signals $|\psi_1^{(n)}\rangle, |\psi_2^{(n)}\rangle, \ldots, |\psi_m^{(n)}\rangle \in \mathcal{H}^{\otimes n}$
- with probs. $p_1^{(n)}, p_2^{(n)}, \ldots, p_m^{(n)}$

  $\langle \psi_i^{(n)} | \psi_j^{(n)} \rangle \neq \delta_{ij}$ in general

- Source State : $\rho_n = \sum_{i=1}^{m} p_i^{(n)} |\psi_i^{(n)}\rangle\langle\psi_i^{(n)}|$

Compression-Decompression Scheme

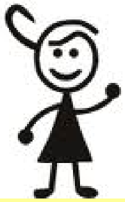
- Encoding: $\mathcal{E}_n : \underbrace{|\psi_i^{(n)}\rangle\langle\psi_i^{(n)}|}_{\text{signal}} \to \underbrace{\sigma_i^{(n)}}_{\text{compressed state}} \in \mathcal{D}(\mathcal{H}_c^{\,n})$

  $\mathcal{H}_c^{\,n}$ → compressed Hilbert space

- Decoding: $\mathcal{D}_n : \sigma_i^{(n)} \to \underbrace{\omega_i^{(n)}}_{\text{recovered signal}} \in \mathcal{D}(\mathcal{H}^{\otimes n})$

UNIVERSITY OF CAMBRIDGE

Quantum Data Compression or Fixed length quantum source coding



n copies of a quantum information source ρ_n

$|\psi_i^{(n)}\rangle\langle\psi_i^{(n)}|$ signal $\in \mathcal{D}(\mathcal{H}^{\otimes n})$ with prob. $p_i^{(n)}$

$\xrightarrow{} \mathcal{E}_n \dashrightarrow \sigma_i^{(n)} \dashrightarrow \mathcal{D}_n \xrightarrow{} \omega_i^{(n)}$

compressed state in $\mathcal{H}_c^n$

recovered state $\in \mathcal{D}(\mathcal{H}^{\otimes n})$

$M_n := \dim \mathcal{H}_c^n < \dim \mathcal{H}^{\otimes n}$

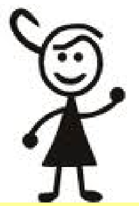
- Ensemble average fidelity: figure of merit used for determining reliability

$$\bar{F}_n = \sum p_i^{(n)} \left\langle \psi_i^{(n)} \right| \mathcal{D}_n \circ \mathcal{E}_n \left(\left| \psi_i^{(n)} \right\rangle \left\langle \psi_i^{(n)} \right| \right) \left| \psi_i^{(n)} \right\rangle$$

- The compression-decompression scheme is reliable if $\boxed{\bar{F}_n \to 1 \text{ as } n \to \infty}$

UNIVERSITY OF CAMBRIDGE

Quantum Data Compression



n copies of a quantum information source ρ_n

signal $|\psi_i^{(n)}\rangle\langle\psi_i^{(n)}| \in \mathcal{B}(\mathcal{H}^{\otimes n})$ with prob. $p_i^{(n)}$

$$|\psi_i^{(n)}\rangle\langle\psi_i^{(n)}| \longrightarrow \mathcal{E}_n \overset{\sigma_i^{(n)}}{\dashrightarrow} \mathcal{D}_n \overset{\omega_i^{(n)}}{\longrightarrow}$$

compressed state in $\mathcal{H}_c^n$

$$M_n := \dim \mathcal{H}_c^n < \dim \mathcal{H}^{\otimes n}$$

recovered state $\in \mathcal{B}(\mathcal{H}^{\otimes n})$

$C_n := (\mathcal{E}_n, \mathcal{D}_n, M_n)$ defines a **code** of **rate** : $\dfrac{\log M_n}{n}$

- $\mathcal{E}_n, \mathcal{D}_n$ is a compression-decompression scheme of rate R

$$M_n := \dim \mathcal{H}_c^n = 2^{nR}$$

- Optimal rate of data compression: **Data compression limit**

$$R_{opt} := \inf \left\{ R \mid \exists a\ seq.\ of\ codes\ C_n := (\mathcal{E}_n, \mathcal{D}_n, 2^{nR})\ s.t.\ \bar{F}_n \to 1\ as\ n \to \infty \right\}$$

UNIVERSITY OF CAMBRIDGE

Schumacher's Theorem : Quantum Data Compression

Suppose $\{\rho, \mathcal{H}\}$ is an *memoryless, quantum information source*

$$\rho_n = \rho^{\otimes n}; \quad S(\rho):$$ von Neumann entropy

- Suppose $R > S(\rho)$: then there exists a reliable compression scheme of rate R for the source.
- If $R < S(\rho)$ then any compression scheme of rate R will not be reliable.

$$R_{opt} = S(\rho):$$

Proof follows from the Typical Subspace theorem

UNIVERSITY OF CAMBRIDGE

DIGRESSION

The notion of "typicality"

Typical sequences and Typical Subspace Theorem

UNIVERSITY OF CAMBRIDGE

Typical Sequences

- **Defn:** Consider a sequence of i.i.d. random variables:

$$U_1, U_2, \ldots U_n; \;\; p(u) \; ; \; u \in J$$

For any $\varepsilon > 0,$ sequences $\underline{u} := (u_1, u_2, \ldots u_n) \in J^n$ for which

$$2^{-n(H(U)+\varepsilon)} \leq p(u_1, u_2, \ldots u_n) \leq 2^{-n(H(U)-\varepsilon)},$$

where $H(U) = -\sum p(u) \log p(u);$ *Shannon entropy*

are called $\varepsilon -$ typical sequences

$T_\varepsilon^{(n)} := \varepsilon -$ typical set = set of $\varepsilon -$ typical sequences

- **Note:** Typical sequences are almost equiprobable

$$\forall \; \underline{u} \in T_\varepsilon^{(n)}, \;\; p(\underline{u}) \approx 2^{-nH(U)}$$

$$\forall\ \underline{u} \in T_{\varepsilon}^{(n)}, \quad p(\underline{u}) \approx 2^{-nH(U)}$$

$U_1, U_2, \ldots U_n;$

$p(u)\ ;\ u \in J$

(Q) Does this agree with our intuitive notion of typical sequences?

(A) Yes! For an i.i.d. sequence : $U_1, U_2, \ldots U_n;\ \ U_i \sim p(u)\ ;\ u \in J$

A typical sequence $\underline{u} := (u_1, u_2, \ldots u_n)$ of length n,

is one which contains approx. $np(u)$ copies of u, $\forall\ u \in J$

- Probability of such a sequence is approximately given by

$$\approx \prod_{u \in J} p(u)^{np(u)} = \prod_{u \in J} 2^{np(u)\log p(u)} = 2^{\sum_{u\in J} p(u)\log p(u)}$$

$$= 2^{-nH(U)}$$

UNIVERSITY OF CAMBRIDGE

Properties of the Typical Set $T_\varepsilon^{(n)}$

- Let $\left|T_\varepsilon^{(n)}\right|$: number of typical sequences

  $P\left(T_\varepsilon^{(n)}\right)$: probability of the typical set

- **Typical Sequence Theorem:** Fix $\varepsilon > 0,$ then $\forall \delta > 0,$ and n large enough,
  - $$P\left(T_\varepsilon^{(n)}\right) > 1 - \delta$$
  - $$(1-\delta)2^{n(H(U)-\varepsilon)} \le \left|T_\varepsilon^{(n)}\right| \le 2^{n(H(U)+\varepsilon)}$$

$\Rightarrow J^n = T_\varepsilon^{(n)} \bigcup A_\varepsilon^{(n)}$

atypical set

(disjoint union)

- *sequences in the atypical set rarely occur*

  $$P\left(A_\varepsilon^{(n)}\right) \le \delta$$

- *typical sequences are almost equiprobable*

UNIVERSITY OF CAMBRIDGE

Memoryless quantum information source

state of n *copies of the source*
$$\rho_n = \sum_{i=1}^{m} p_i^{(n)} \left|\psi_i^{(n)}\right\rangle\left\langle\psi_i^{(n)}\right| = \rho^{\otimes n};$$

$\left|\psi_i^{(n)}\right\rangle$:*signal emitted with prob.* $p_i^{(n)}$; $\quad \left\langle\psi_i^{(n)}\middle|\psi_j^{(n)}\right\rangle \neq \delta_{ij}$

$$\rho \in \mathcal{H}, \dim \mathcal{H} = d \qquad \therefore \rho_n = \rho^{\otimes n} \in \mathcal{H}^{\otimes n}$$

Spectral decompositions:

$$\rho = \sum_{j=1}^{d} q_j \left|\varphi_j\right\rangle\left\langle\varphi_j\right|; \qquad \rho_n = \sum_{\underline{k}} \lambda_{\underline{k}}^{(n)} \left|\Psi_{\underline{k}}^{(n)}\right\rangle\left\langle\Psi_{\underline{k}}^{(n)}\right|$$

eigenstates

$$\therefore \rho_n = \rho^{\otimes n} \Rightarrow \quad \left|\Psi_{\underline{k}}^{(n)}\right\rangle = \left|\varphi_{k_1}\right\rangle \otimes \left|\varphi_{k_2}\right\rangle \otimes \ldots \left|\varphi_{k_n}\right\rangle$$

$$\lambda_{\underline{k}}^{(n)} = q_{k_1} q_{k_2} \ldots\ldots q_{k_n}$$

Identification of the label k *as a sequence of classical indices*

$$\underline{k} \equiv k = (k_1, k_2, \ldots\ldots, k_n)$$

- sum over all possible sequences

$$\rho_n \equiv \rho^{\otimes n} = \sum_{\underline{k}} \lambda_{\underline{k}}^{(n)} \left|\Psi_{\underline{k}}^{(n)}\right\rangle\left\langle\Psi_{\underline{k}}^{(n)}\right|$$

$$\underline{k} \equiv (k_1, k_2, \ldots\ldots, k_n):$$
$$k_i \in \{1, 2, .., d\}; \quad d = \dim \mathcal{H}$$

$$\lambda_{\underline{k}}^{(n)} = q_{k_1} q_{k_2} \ldots\ldots q_{k_n}$$

von Neumann entropy
$$S(\rho_n) = S(\rho^{\otimes n}) = nS(\rho) = nH(\{q_k\})$$

Probability: $p(\underline{k}) \equiv \lambda_{\underline{k}}^{(n)} = q_{k_1} q_{k_2} \ldots\ldots q_{k_n}$

$\forall \varepsilon > 0,$ a sequence $\underline{k} \equiv (k_1, k_2, \ldots\ldots, k_n)$ is $\varepsilon -$typical if:

$$2^{-n(H(\{q_k\})+\varepsilon)} \le p(\underline{k}) \le 2^{-n(H(\{q_k\})-\varepsilon)},$$

$$2^{-n(S(\rho)+\varepsilon)} \le p(\underline{k}) \le 2^{-n(S(\rho)-\varepsilon)},$$

$T_\varepsilon^{(n)} :=$ $\varepsilon -$ typical set

eigenvalues $\lambda_{\underline{k}}^{(n)}$ $\longleftrightarrow$ sequences $\underline{k}$

eigenvectors $\left|\Psi_{\underline{k}}^{(n)}\right\rangle$

$\longrightarrow$ $\mathcal{T}_\varepsilon^{(n)} :=$ $\varepsilon -$ typical subspace

UNIVERSITY OF CAMBRIDGE

$\varepsilon -$ typical subspace $\mathcal{T}_\varepsilon^{(n)} \subset \mathcal{H}^{\otimes n}$

- Subspace spanned by those eigenvectors

$$\left|\Psi_{\underline{k}}^{(n)}\right\rangle = \left|\varphi_{k_1}\right\rangle \otimes \left|\varphi_{k_2}\right\rangle \otimes \ldots \left|\varphi_{k_n}\right\rangle \quad \text{for which} \quad \underline{k} \in T_\varepsilon^{(n)}$$

- Let $P_\varepsilon^{(n)}$: orthogonal projection on to the typical subspace

Typical Sequence Theorem $\Longrightarrow$ **Typical Subspace Theorem**

Fix $\varepsilon > 0,$ then $\forall \delta > 0,$ and n large enough:

$P\left(T_\varepsilon^{(n)}\right) > 1-\delta$	$\mathrm{Tr}\left(P_\varepsilon^{(n)} \rho_n\right) > 1-\delta$
$(1-\delta)2^{n(H(\{q_k\})-\varepsilon)} \le \left\|T_\varepsilon^{(n)}\right\|$	$(1-\delta)2^{n(S(\rho)-\varepsilon)} \le \dim \mathcal{T}_\varepsilon^{(n)}$
$\le 2^{n(H(\{q_k\})+\varepsilon)}$	$\le 2^{n(S(\rho)+\varepsilon)}$

UNIVERSITY OF CAMBRIDGE

Schumacher's Theorem : Quantum Data Compression

- Suppose $R > S(\rho)$: then there exists a reliable compression scheme of rate R for the source.

- **Proof:**

  Compressed Hilbert space $\mathcal{H}_c^n; \ \dim \mathcal{H}_c^n = 2^{nR}$ $\qquad R > S(\rho)$

  - *Choose* $\varepsilon > 0,$ *such that* $R > S(\rho) + \varepsilon$

  Fix $\delta > 0,$ choose n large enough such that:

$$\mathrm{Tr}\left(P_\varepsilon^{(n)} \rho_n\right) > 1 - \delta; \qquad \dim \mathcal{T}_\varepsilon^{(n)} \leq 2^{n(S(\rho)+\varepsilon)} < 2^{nR} = \dim \mathcal{H}_c^n$$

$$\Rightarrow \ \mathcal{T}_\varepsilon^{(n)} \subset \mathcal{H}_c^n$$

UNIVERSITY OF CAMBRIDGE

Idea behind the compression scheme

$\left|\psi_i^{(n)}\right\rangle$:*signal emitted with prob.* $p_i^{(n)}$; $\quad \left\langle \psi_i^{(n)} \middle| \psi_j^{(n)} \right\rangle \neq \delta_{ij}$

$$\left|\psi_i^{(n)}\right\rangle = P_\varepsilon^{(n)} \left|\psi_i^{(n)}\right\rangle + \left(I - P_\varepsilon^{(n)}\right)\left|\psi_i^{(n)}\right\rangle$$

$\in \mathcal{T}_\varepsilon^{(n)}$

keep this part unchanged

$\notin \mathcal{T}_\varepsilon^{(n)}$

map this onto a fixed pure state

$\left|\phi_0^{(n)}\right\rangle \in \mathcal{T}_\varepsilon^{(n)}$

Compression scheme $\quad \mathcal{E}_n\left(\left|\psi_i^{(n)}\right\rangle\left\langle\psi_i^{(n)}\right|\right) = \tilde{\rho}_i^{(n)}$

$$\tilde{\rho}_i^{(n)} = \alpha_i^2 \left|\tilde{\psi}_i^{(n)}\right\rangle\left\langle\tilde{\psi}_i^{(n)}\right| + \beta_i^2 \left|\phi_0^{(n)}\right\rangle\left\langle\phi_0^{(n)}\right| \in \mathcal{D}\left(\mathcal{T}_\varepsilon^{(n)}\right)$$

$$\left|\tilde{\psi}_i^{(n)}\right\rangle \propto P_\varepsilon^{(n)}\left|\psi_i^{(n)}\right\rangle; \alpha_i^2 = \left\|P_\varepsilon^{(n)}\left|\psi_i^{(n)}\right\rangle\right\|^2; \beta_i^2 = \left\|\left(I - P_\varepsilon^{(n)}\right)\left|\psi_i^{(n)}\right\rangle\right\|^2$$

Decompression scheme $\quad \mathcal{D}_n\left(\tilde{\rho}_i^{(n)}\right) = \tilde{\rho}_i^{(n)} \oplus 0$

UNIVERSITY OF CAMBRIDGE

$$\tilde{\rho}_i^{(n)} = \alpha_i^2 \left|\tilde{\psi}_i^{(n)}\right\rangle\left\langle\tilde{\psi}_i^{(n)}\right| + \beta_i^2 \left|\phi_0^{(n)}\right\rangle\left\langle\phi_0^{(n)}\right| \in \mathcal{D}\left(\mathcal{T}_\varepsilon^{(n)}\right)$$

$$\alpha_i^2 = \left\| P_\varepsilon^{(n)} \left|\psi_i^{(n)}\right\rangle \right\|^2 = \left\langle \psi_i^{(n)} \right| P_\varepsilon^{(n)} \left| \psi_i^{(n)} \right\rangle$$

Ensemble average fidelity

$$\bar{F}_n = \sum_i p_i^{(n)} \left\langle \psi_i^{(n)} \right| \tilde{\rho}_i^{(n)} \left| \psi_i^{(n)} \right\rangle \geq 2\sum_i p_i^{(n)} \alpha_i^2 - 1$$

$$> 1 - 2\delta$$

$$\sum_i p_i^{(n)} \alpha_i^2 = \sum_i p_i^{(n)} \left\langle \psi_i^{(n)} \right| P_\varepsilon^{(n)} \left| \psi_i^{(n)} \right\rangle$$

$$= \mathrm{Tr}\left(P_\varepsilon^{(n)} \rho_n\right) > 1 - \delta;$$ *(by the Typical Subspace Theorem)*

$$\Rightarrow \quad \bar{F}_n \to 1 \text{ as } n \to \infty \quad \blacksquare$$

Schumacher's Theorem : Quantum Data Compression

Suppose $\{\rho, \mathcal{H}\}$ is an *memoryless, quantum information source*

$$\rho_n = \rho^{\otimes n}; \quad S(\rho):$$ von Neumann entropy

- Suppose $R > S(\rho)$: then there exists a reliable compression scheme of rate R for the source.

- If $R < S(\rho)$ then any compression scheme of rate R will not be reliable.

(See Cambridge lecture notes)

UNIVERSITY OF CAMBRIDGE

Schumacher proved (1995): for a memoryless source

$$\{\rho,\ \mathcal{H}\}$$

Data compression limit = $S(\rho)$: *von Neumann entropy of the source*

Transmission of information

Transmission of classical info through a noiseless quantum channel

$X \sim p(x);$

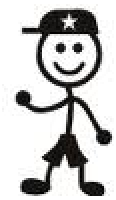
$x \in J$

x

classical info source

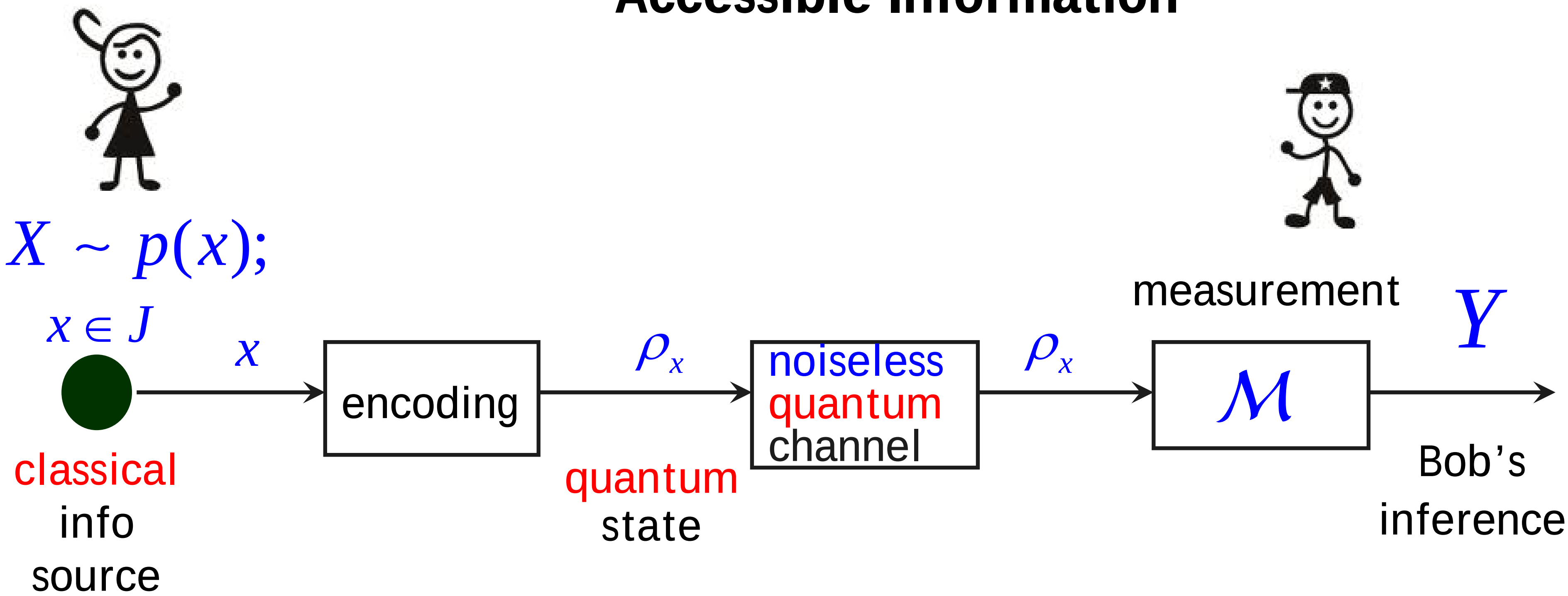
classical message

noiseless quantum channel



UNIVERSITY OF CAMBRIDGE

Accessible Information



- Bob receives the ensemble: $\mathcal{E} = \{p(x), \rho_x\}$

The maximum amount of info Bob can extract by doing any measurement

Accessible Information: $$I_{acc}(\mathcal{E}) = \max_{\mathcal{M}} I(X:Y)$$

(classical) mutual info

Holevo Bound

$$I_{acc}(\mathcal{E}) \leq \chi\left(\{p(x), \rho_x\}\right)$$

The maximum amount of info Alice can send to Bob using the ensemble $\mathcal{E} = \{p(x), \rho_x\}$

- Holevo $\chi-$quantity of an ensemble of states $\{p_i, \sigma_i\}$

$$\chi\left(\{p(x), \rho_x\}\right) := S\left(\sum_x p(x)\rho_x\right) - \sum_x p(x) S\left(\rho_x\right)$$

If the ρ_x are pure :

$\chi\left(\{p(x), \rho_x\}\right) = S(\rho)$; where $\rho := \sum_x p(x)\rho_x$

UNIVERSITY OF CAMBRIDGE

Holevo Bound: **sketch of proof**

Idea: Use strong subadditivity: need a tripartite system

Embed the classical r.v. X in a dummy quantum system $A;\ \mathcal{H}_A$

$\{|x\rangle : x \in J\}$: orthonormal basis in $\mathcal{H}_A$

- $A:$ a quantum register; keeps a record of the classical symbol x which Alice wants to send to Bob
- $Q:$ the quantum system in whose states ρ_x Alice encodes her messages
- $B:$ a quantum system representing Bob's measuring device; originally in some pure state $|0\rangle\langle 0|_B$

- Initial state: $$\rho_{AQB} = \Big(\sum_x p(x)|x\rangle\langle x|_A \otimes \rho_x\Big) \otimes |0\rangle\langle 0|_B$$

UNIVERSITY OF CAMBRIDGE

Holevo Bound: **sketch of proof**

- Initial state: $$\rho_{AQB} = \left(\sum_x p(x)|x\rangle\langle x|_A \otimes \rho_x\right) \otimes |0\rangle\langle 0|_B$$
- Bob's measurement $\mathcal{M}$: POVM $E = \{E_x\}_{x \in J}$;
- State after measurement $$\rho_{A'Q'B'} = \sum_{x,y} p(x)|x\rangle\langle x|_{A'} \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|_{B'}$$

  - $I(A:Q) = I(A:QB)$
  - $I(A:QB) \geq I(A':Q'B')$
  - $I(A':Q'B') \geq I(A':B')$

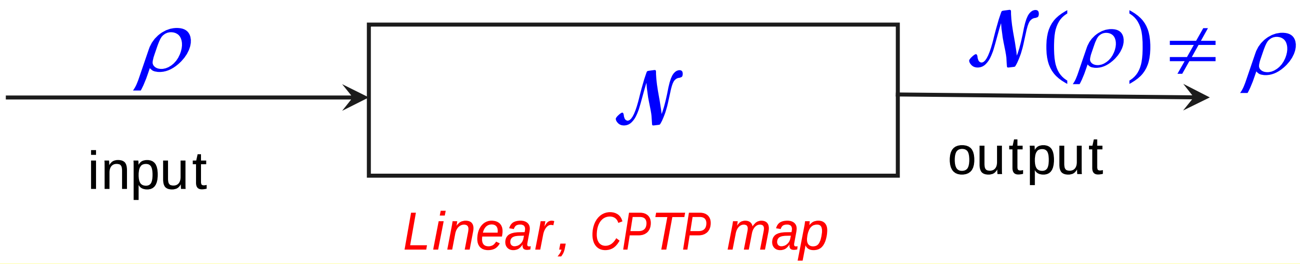
$$\longrightarrow I(A':B') \leq I(A:Q)$$

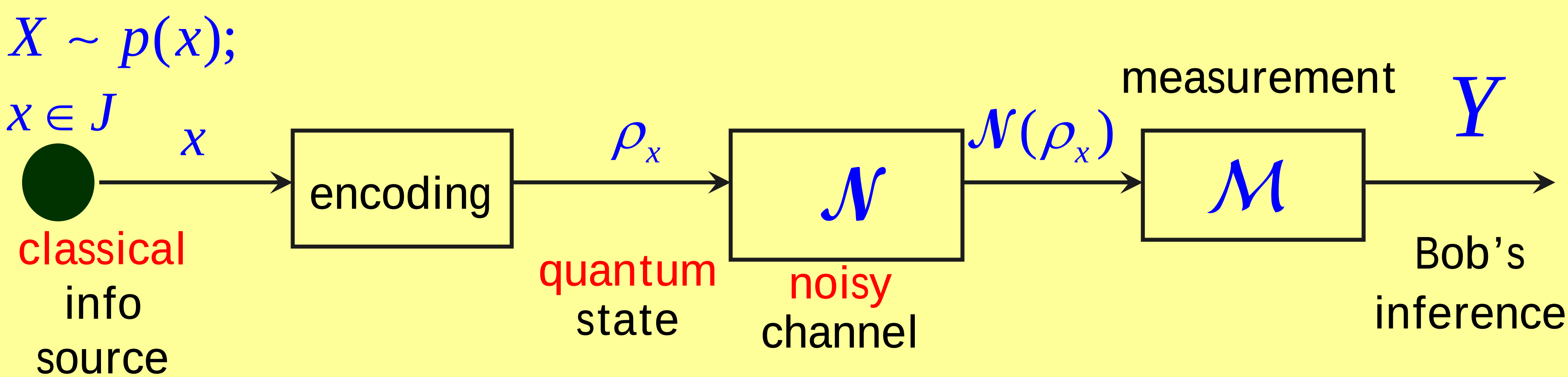
$$I(X:Y) \leq \chi\left(\{p(x),\rho_x\}\right) \quad \forall \mathcal{M}$$

$$I_{acc}(\mathcal{E}) = \max_{\mathcal{M}} I(X:Y) \leq \chi\left(\{p(x),\rho_x\}\right)$$ Holevo Bound

UNIVERSITY OF CAMBRIDGE

Transmission of classical info through noisy quantum channels





- Bob receives the ensemble: $\mathcal{E} = \{p(x), \mathcal{N}(\rho_x)\}$

Holevo bound $I_{acc}(\mathcal{E}) \leq \chi\left(\{p(x), \mathcal{N}(\rho_x)\}\right)$

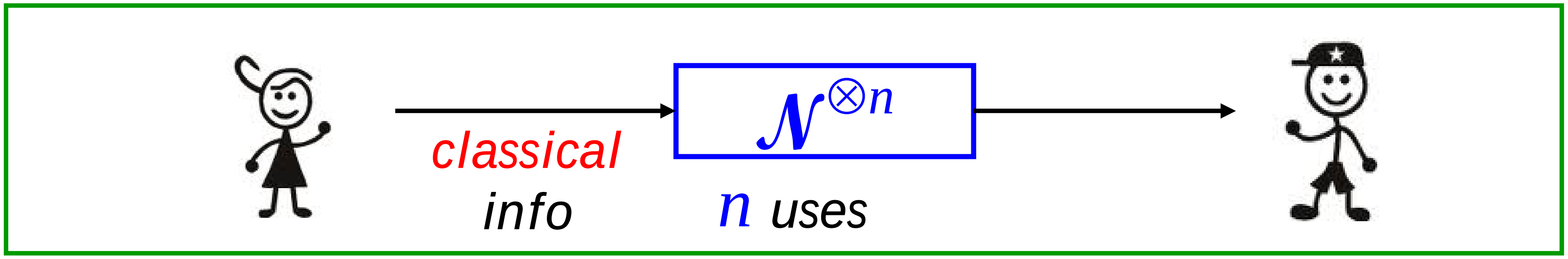
(Q) Is the Holevo bound achievable?

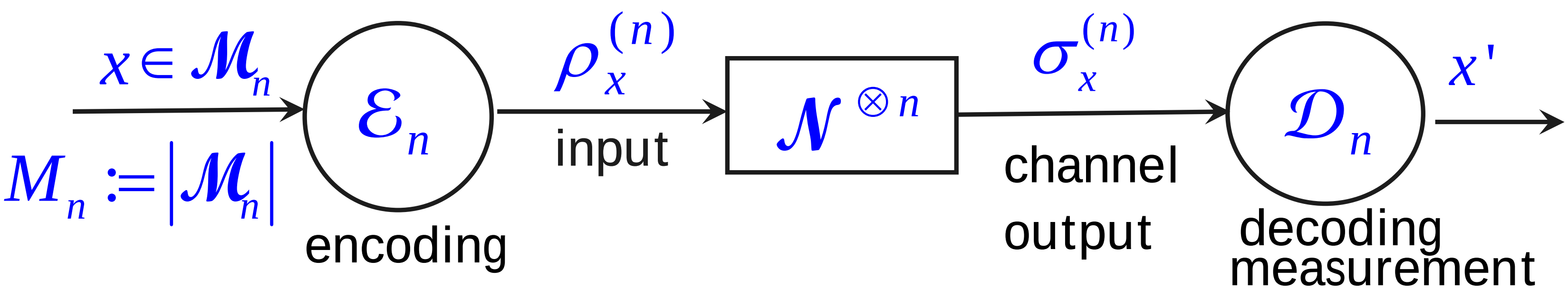
Yes, in the asymptotic, memoryless setting

UNIVERSITY OF CAMBRIDGE

"asymptotic, memoryless setting"

classical info transmission through a noisy quantum channel $\mathcal{N}$





- Measurement: POVM $\left\{E_x^{(n)}\right\}$ $\qquad \sigma_x^{(n)} = \mathcal{N}^{\otimes n}\left(\rho_x^{(n)}\right)$
- Probability (Bob infers x correctly)= $\mathrm{Tr}\left(E_x^{(n)}\sigma_x^{(n)}\right)$
- Average probability of error:

$$\boxed{p_{av}^{(n)} = \frac{1}{|\mathcal{M}_n|}\sum_{x\in\mathcal{M}_n}\left[1-\mathrm{Tr}\left(E_x^{(n)}\sigma_x^{(n)}\right)\right]}$$

UNIVERSITY OF CAMBRIDGE

If $p_{av}^{(n)} \to 0$ *as* $n \to \infty$ ……*(1)* *: information transmission is reliable*

If number of bits of message sent: $\log M_n \approx 2^{nR}$ *& (1) holds, then*

R*: an achievable rate* $R = \liminf_{n\to\infty} \frac{\log|M_n|}{n}$

Classical capacity of the quantum channel

$$C(\mathcal{N}) = \sup R$$

--the supremum taken over all achievable rates

UNIVERSITY OF CAMBRIDGE

A quantum channel has many capacities

- The different capacities depend on:
  - the nature of the transmitted information
    (classical or quantum)
  - the nature of the input states
    (entangled or product states)
  - the nature of the measurements done on the outputs
    (collective or individual)
  - the presence or absence of any additional resource
    (e.g. prior shared entanglement between Alice & Bob)
  - whether Alice & Bob are allowed to communicate
    classically with each other

- *Capacities evaluated in the "asymptotic memoryless setting"*

$$\Phi^{(n)} = \Phi^{\otimes n}; \quad n \to \infty$$

UNIVERSITY OF CAMBRIDGE

- If Alice restricts the inputs to product states, i.e., if

$$x \to \rho_x^{(n)} = \rho_{x_1} \otimes \rho_{x_2} \otimes \ldots\ldots \otimes \rho_{x_n}$$

- And Bob does a collective measurement (POVM) on

$$\sigma_x^{(n)} := \mathcal{N}^{\otimes n}\left(\rho_x^{(n)}\right)$$: the output of n uses of the channel

$$= \mathcal{N}(\rho_{x_1}) \otimes \mathcal{N}(\rho_{x_2}) \otimes \ldots\ldots \otimes \mathcal{N}(\rho_{x_n})$$

Capacity **: product state capacity** $C_p(\mathcal{N})$

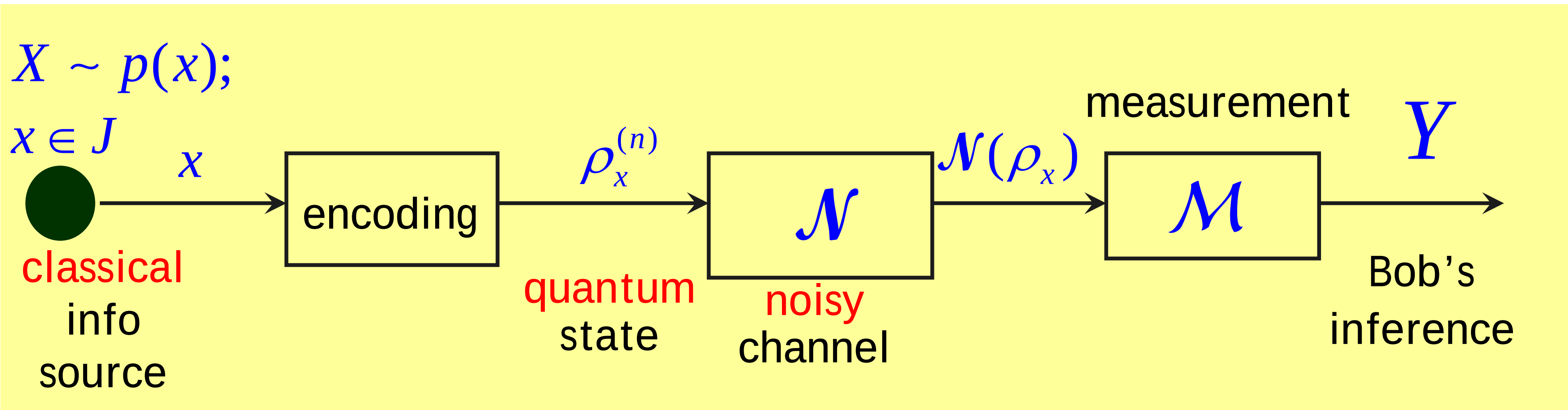
- **Holevo-Schumacher-Westmoreland (HSW) Theorem**

$$\boxed{C_p(\mathcal{N}) = \max_{\{p_i,\rho_i\}} \chi\left(\{p_i, \mathcal{N}(\rho_i)\}\right) = \chi^*(\mathcal{N})}$$

Holevo Capacity

- *Can be expressed as a relative entropy*

Classical info transmission through noisy quantum channels



- Bob receives the ensemble: $\mathcal{E} = \{p(x), \mathcal{N}(\rho_x)\}$

$$I_{acc}(\mathcal{E}) \leq \chi\left(\{p(x), \mathcal{N}(\rho_x)\}\right)$$

UNIVERSITY OF CAMBRIDGE

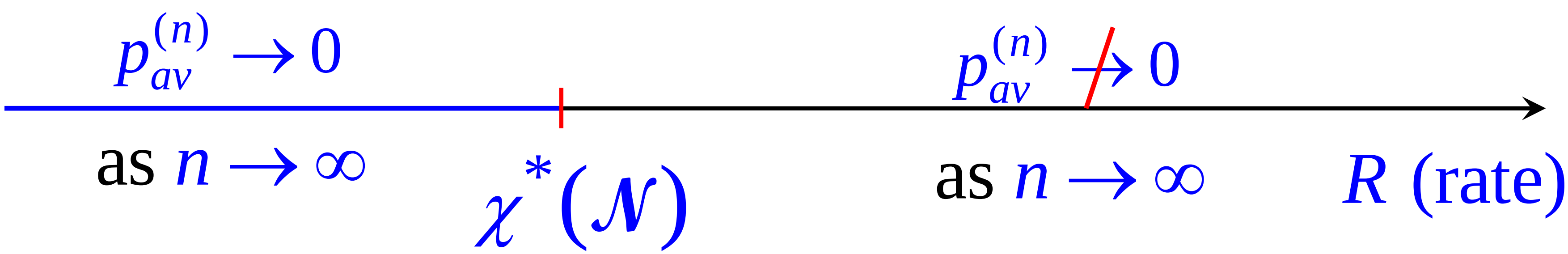
HSW Theorem

$$\boxed{C_p(\mathcal{N}) = \max_{\{p_i,\rho_i\}} \chi\left(\{p_i, \mathcal{N}(\rho_i)\}\right) = \chi^*(\mathcal{N})}$$

Holevo Capacity

Holevo bound can be achieved in the "asymptotic memoryless setting"

IF Alice uses product state inputs & Bob does a collective measurement



UNIVERSITY OF CAMBRIDGE

- **Classical capacity of a memoryless channel $\mathcal{N}$:**

(without the restriction of inputs being product states):

$$\boxed{C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \chi^*\left(\mathcal{N}^{\otimes n}\right)}$$

regularised Holevo capacity

$\chi^*\left(\mathcal{N}^{\otimes n}\right)$ *Holevo Capacity* of the block $\mathcal{N}^{\otimes n}$ of n channels

(This generalization is obtained by considering inputs which are product states over blocks of n channels but which may be entangled within each block)

(Q) Can the classical capacity of a memoryless quantum channel be increased by using entangled states as inputs ?

UNIVERSITY OF CAMBRIDGE

$$\boxed{C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \chi^*\left(\mathcal{N}^{\otimes n}\right)}$$

(Q) *Can the classical capacity of a memoryless quantum channel be increased by using entangled states as inputs ?*

- *This is related to the additivity conjecture of the Holevo capacity* :

$$\boxed{\chi^*(\mathcal{N}_1 \otimes \mathcal{N}_2) = \chi^*(\mathcal{N}_1) + \chi^*(\mathcal{N}_2)} \Rightarrow \chi^*\left(\mathcal{N}^{\otimes n}\right) = n\chi^*(\mathcal{N})$$

$$\Rightarrow \boxed{\begin{aligned} C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \chi^*\left(\mathcal{N}^{\otimes n}\right) = \lim_{n\to\infty} \frac{1}{\cancel{n}} \cancel{n} \chi^*(\mathcal{N}) &= \chi^*(\mathcal{N}) \\ &= C_p(\mathcal{N}) \end{aligned}}$$

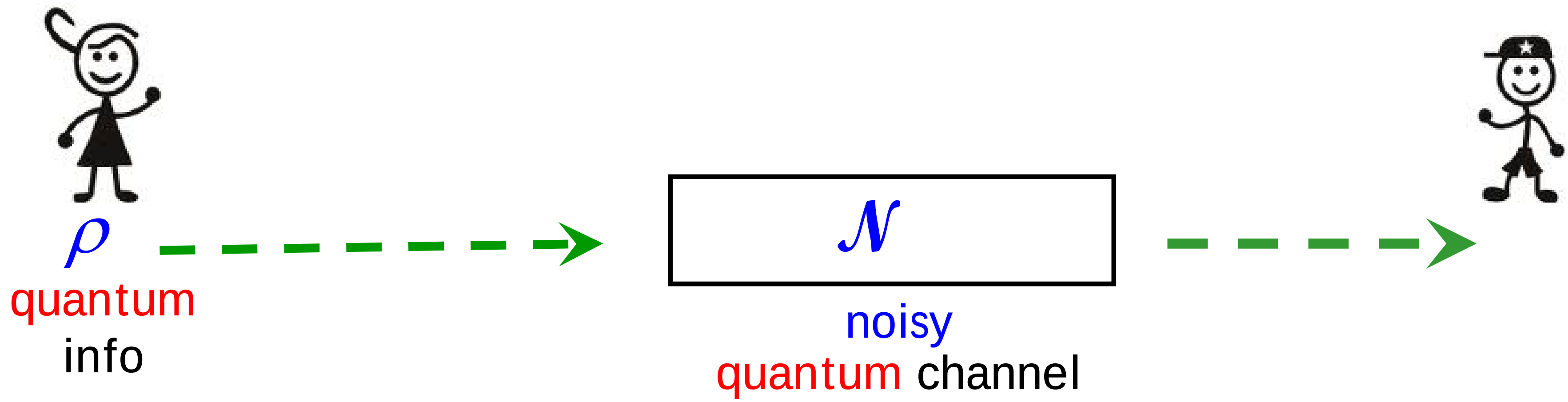
- *IF the Holevo capacity is additive then using entangled inputs would not increase its classical capacity!*

- **Additivity conjecture disproved by Matt Hastings 2008**

Using entangled inputs might help in transmitting classical information through a quantum channel

UNIVERSITY OF CAMBRIDGE

Transmission of Quantum Information



- Quantum capacity : *max. rate at which qubits can be transmitted reliably*

  $Q(\mathcal{N})$

- *Evaluated in the "asymptotic memoryless setting"*

(LSD theorem)

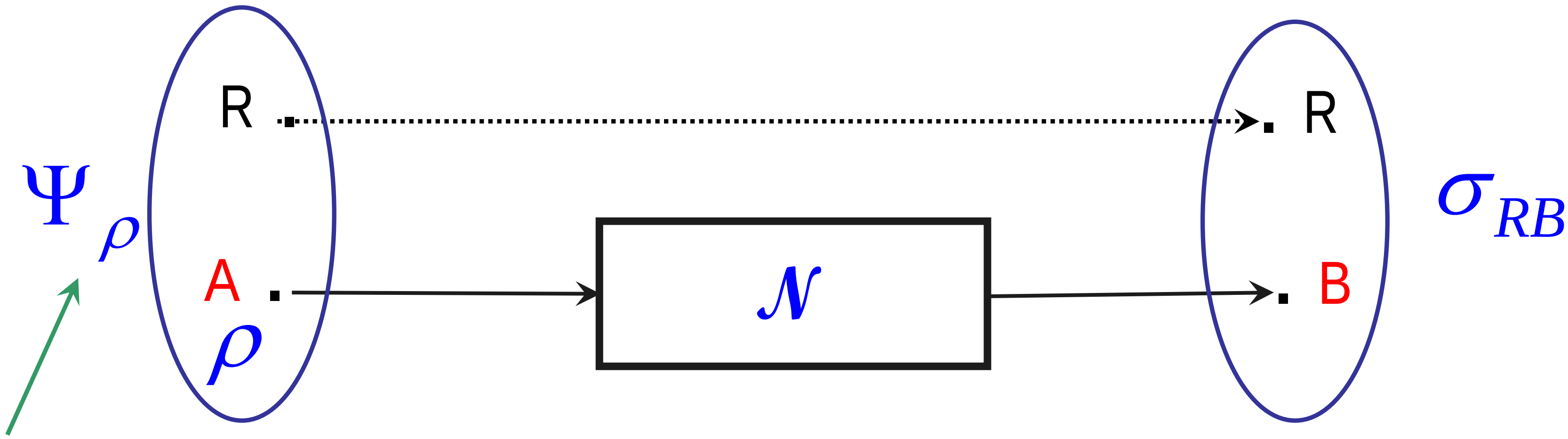
$$Q(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho^{(n)}} I_{coh}\left(\rho^{(n)}, \mathcal{N}^{\otimes n}\right)$$

coherent information

UNIVERSITY OF CAMBRIDGE

Quantum Capacity

- *Given in terms of the coherent information:* $I_{coh}(\rho, \mathcal{N})$



purification of ρ

$$\sigma_{RB} = (\mathrm{id}_R \otimes \mathcal{N})\Psi_\rho$$

coherent information

$$\boxed{I_{coh}(\rho, \mathcal{N}) = -S(\sigma_{RB}) + S(\sigma_B) \quad = -S(R\,|\,B)_\sigma}$$

Quantum Capacity

- For a memoryless channel

(LSD theorem)

$$Q(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho^{(n)}} I_{coh}\left(\rho^{(n)}, \mathcal{N}^{\otimes n}\right)$$

Regularised

Coherent information

In next lecture and example session:

- Discussion of degradable channels
- Proof of the fact that the coherent info is additive for degradable channels