**2224-1**

# School on New Trends in Quantum Dynamics and Quantum Entanglement

*14 - 18 February 2011*

# Quantum Channels with Memory

Nilanjana DATTA
*DPMMS, Mathematical Sciences*
*Univ. of Cambridge*
*U.K.*

# Memory Effects in Quantum Channels

Nilanjana Datta
University of Cambridge,U.K.

- One of the most common and essential tasks of everyday life:

  *transmission of information*

- Examples of **classical communications** channels:

  *telephones/mobile phones/computers*

- A **quantum communications channel** : one which

incorporates intrinsically quantum-mechanical effects

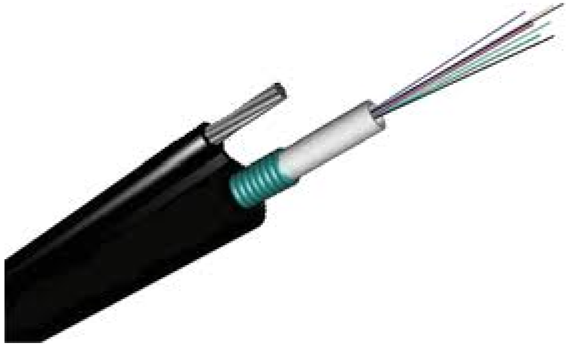
- Example of a **quantum communications channel** :

  -- an optical fibre

  - input to the channel :

    a photon in some

    quantum-mechanical state



- In Quantum Information Theory, information is carried by (or embodied in) *physical states of quantum-mechanical systems*:
- e.g. polarization states of a photon, spin states of electrons
- State space : Hilbert space $H$ associated with the system,

(finite-dimensional Hilbert spaces)

e.g. $H \simeq \mathbb{C}^2$ *Qubit space*

$B(H)$: algebra of linear operators acting on $H$

- States: density matrices, $\rho\ ; \rho \geq 0,\ \mathrm{Tr}\,\rho = 1$

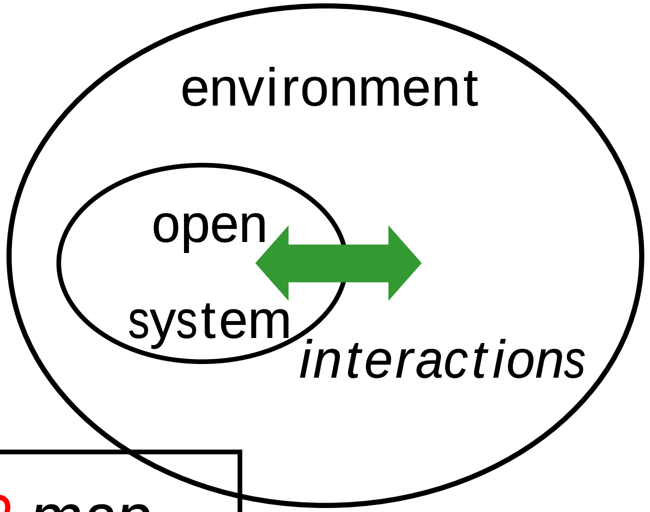
$D(H) \subset B(H)$: set of all density matrices,

# Quantum Channels

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

  *linear completely-positive,*

  *trace preserving (CPTP) map*

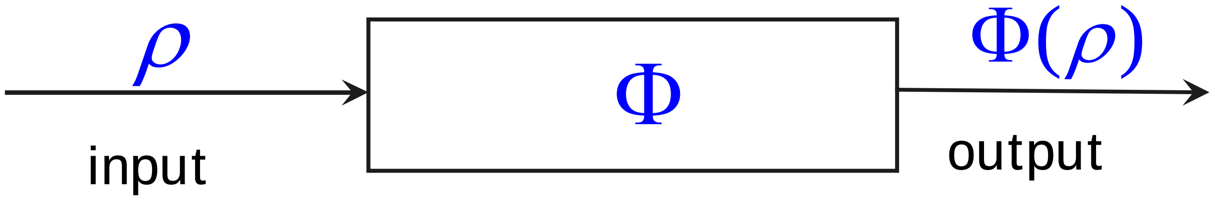

environment

open system

*interactions*

e.g. transmission through a quantum communications channel

*Quantum channel* *: a linear CPTP map*


$\rho$

input

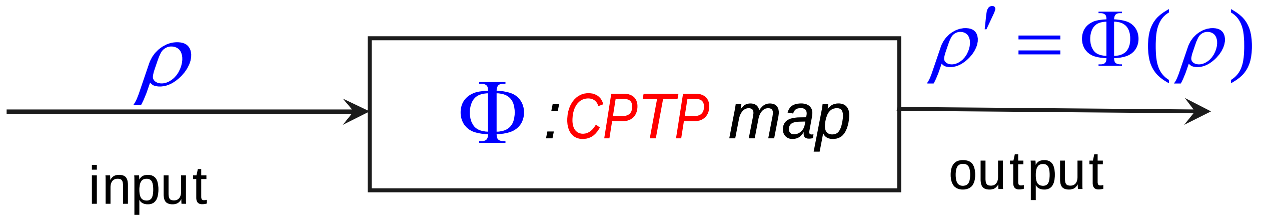
$\Phi$

$\Phi(\rho)$

output

$$\Phi : D\,(H_A) \rightarrow D\,(H_B)$$

$H_A, H_B$ : Hilbert spaces of the input and output systems

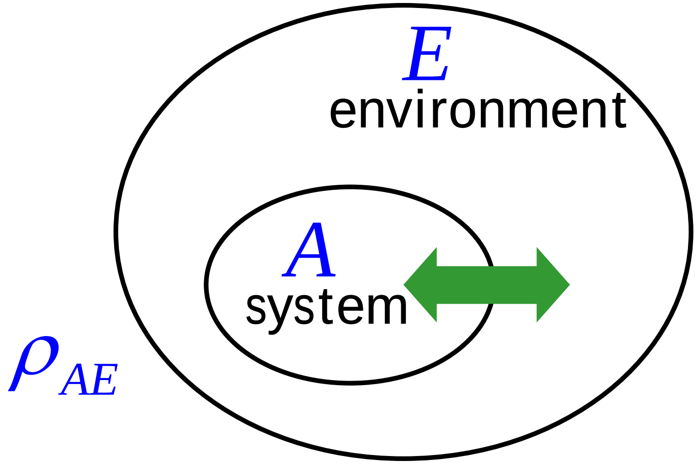


- *Trace preserving (TP):* $\mathrm{Tr}\,\rho' = \mathrm{Tr}\rho = 1$

- *Positive:*

$$\rho' = \Phi(\rho) \geq 0$$

- *Completely positive (CP):*



$$\Phi : D\ (H_A) \rightarrow D\ (H_B)$$

$(\Phi \otimes id_E)(\rho_{AE})$ = *an allowed state of the composite system* $\in D\ (H_B \otimes H_E)$

$$\boxed{(\Phi \otimes id)(\rho_{AE}) \geq 0}$$

# “Church of the larger Hilbert Space”

$$\begin{tikzcd}
\rho \otimes |0_E\rangle\langle 0_E| \arrow[rr, "U"] & & \sigma_{AE} := U\left(\rho \otimes |0_E\rangle\langle 0_E|\right)U^{\dagger} \arrow[d, "\mathrm{Tr}_E \sigma_{AE}"] \\
\rho \arrow[u] \arrow[rr, "\Phi"] & & \Phi(\rho)
\end{tikzcd}$$

system + ancilla
(environment)

$H_A \otimes H_E$

system

$H_A$

$$\Phi(\rho) = \mathrm{Tr}_E\left(U\left(\rho \otimes |0_E\rangle\langle 0_E|\right)U^{\dagger}\right)$$

- *Stinespring’s Dilation Theorem*

# Kraus Representation Theorem:

A quantum channel $\Phi : D\,(H_A) \rightarrow D\,(H_B)$

can be represented as follows:

$$\boxed{\Phi(\rho) = \sum_{i=1}^{M} A_i \rho A_i^{\dagger}}$$

$\{A_i\}_{i=1}^{M}$ : a finite set of linear operators acting on the Hilbert space $H_A$ of the system, satisfying

$$\boxed{\sum_{i=1}^{M} A_i^{\dagger} A_i = \mathrm{I}_A}$$

*completeness relation*

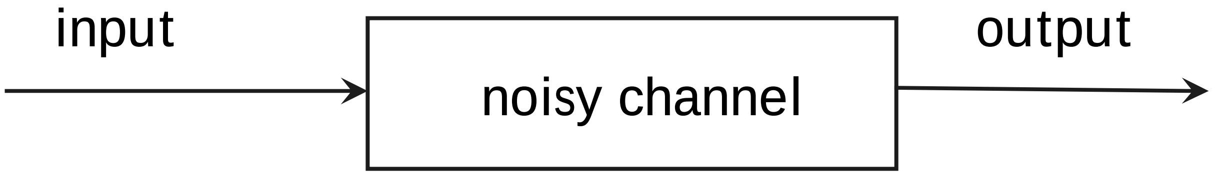
*Kraus operators*

*Identity operator*

- The biggest hurdle in the path of efficient transmission of information:

-- Presence of noise in communications channels.

- Noise distorts the information sent through the channel.



output $\neq$ input

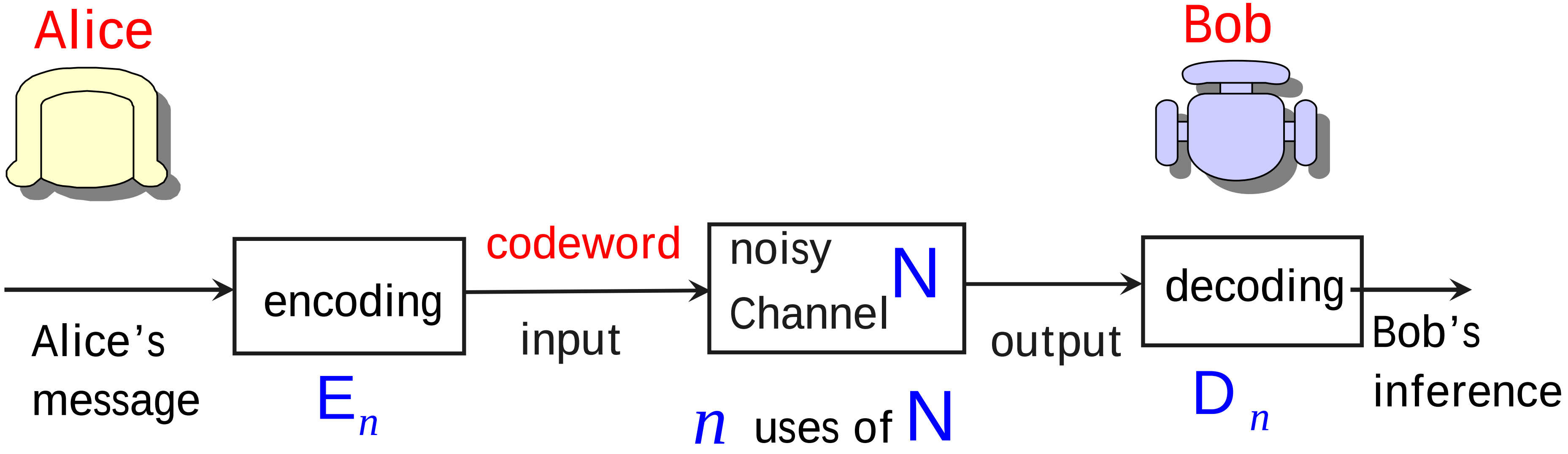
- To combat the effects of noise: use error-correcting codes

To overcome the effects of noise:

instead of transmitting the original messages,

-- the sender encodes her messages into suitable codewords

-- these codewords are then sent through (multiple uses of) the channel



Alice

Bob

Alice's message

encoding

$E_n$

codeword

input

noisy Channel $N$

$n$ uses of $N$

output

decoding

$D_n$

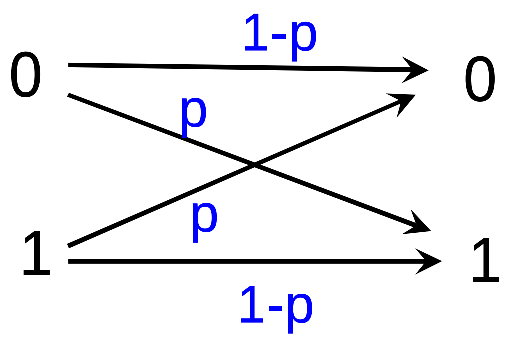
Bob's inference

- *Error-correcting code:* $C_n := (E_n, D_n)$:

- The idea behind the encoding:
  - To introduce redundancy in the message so that upon decoding, Bob can retrieve the original message with a low probability of error:
  - The amount of redundancy which needs to be added – depends on the noise in the channel

# Example

- Memoryless binary symmetric channel (m.b.s.c.)



- it transmits single bits
- effect of the noise: to flip the bit with probability p

*Repetition Code*

- Encoding: 0 ⟶ 000

  1 ⟶ 111

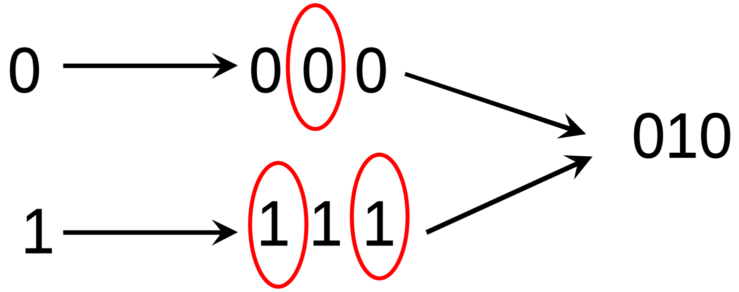
codewords

- the 3 bits are sent through 3 successive uses of the m.b.s.c.
- Suppose $000 \xrightarrow[\text{m.b.s.c.}]{} 010$ (Bob receives)

  codeword

- Decoding : (*majority voting*) 010 ⟶ 0 (Bob infers)

- Probability of error for the m.b.s.c. :
  - without encoding = p
  - with encoding = *Prob (2 or more bits flipped) := q*



0 → 000

1 → 111

010

- *Prove: q < p if p < 1/2*

*-- in this case encoding helps!*

- (Encoding – Decoding) : Repetition Code.

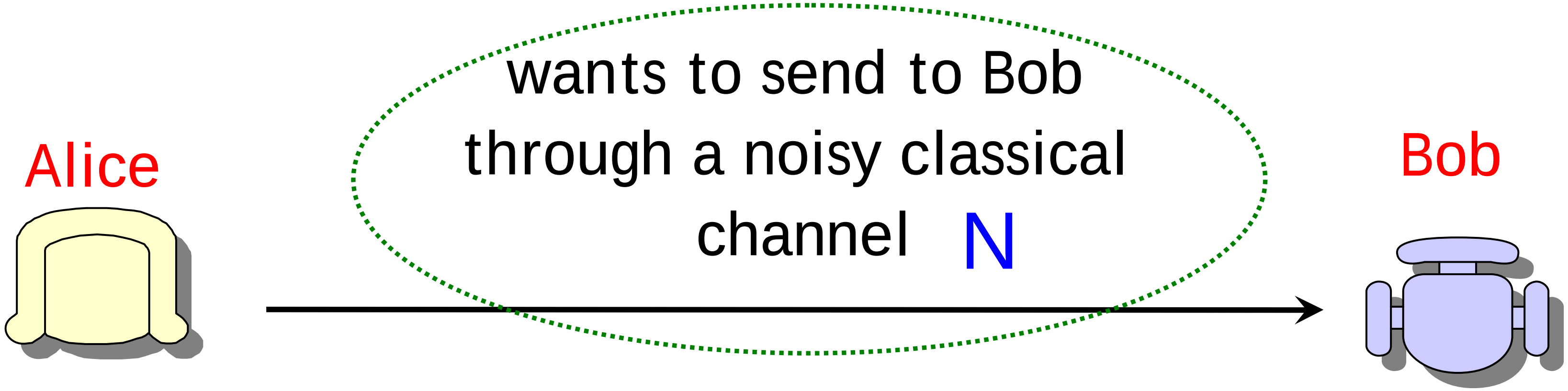
- Information transmission is said to be reliable if:
  -- the probability of error in decoding the output
  vanishes asymptotically in the number of uses of the channel

- Aim: to achieve reliable information transmission
  whilst optimizing the rate

  - the maximum amount of information that can be sent
    per use of the channel

- The optimal rate of reliable info transmission: capacity

# Transmission of info through a classical channel

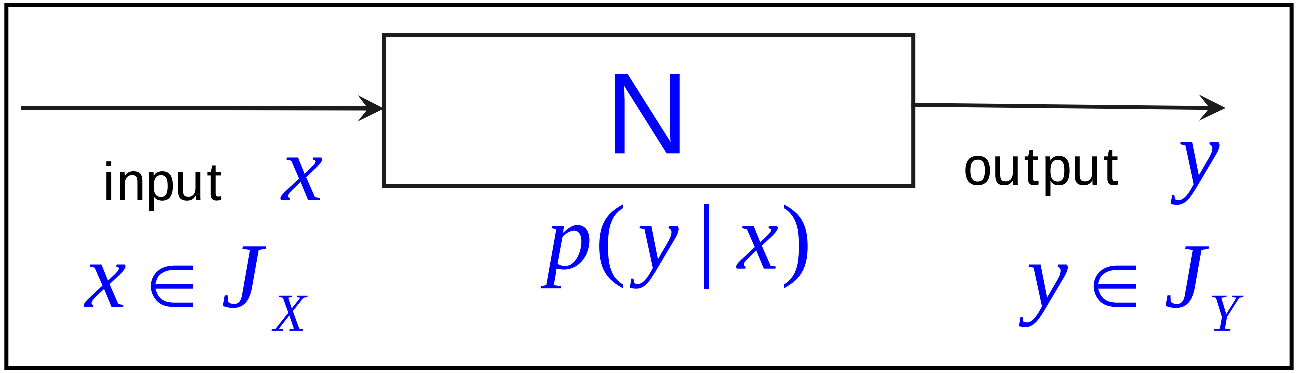


Alice

wants to send to Bob through a noisy classical channel $\mathcal{N}$

Bob

$\mathcal{M}_n =$ a set of classical messages



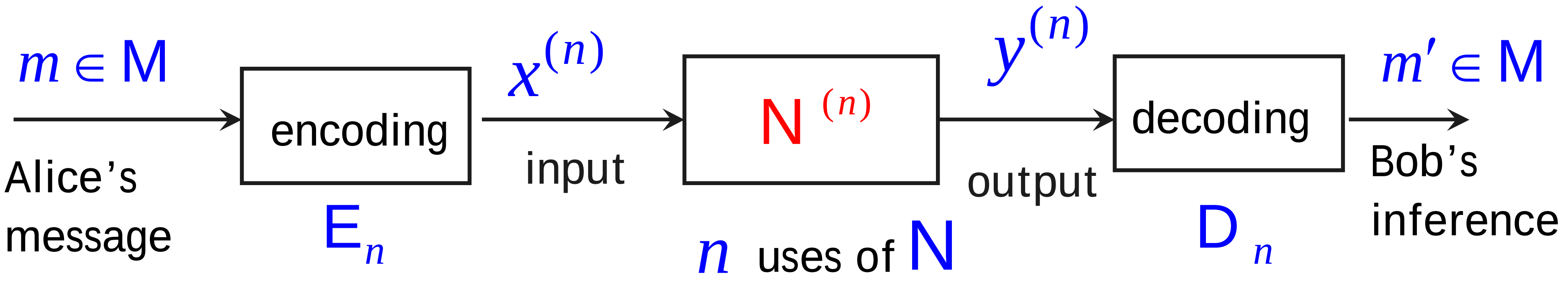
Let :

$$J_X = J_Y = \{0,1\}^n$$

- To overcome the effects of noise:

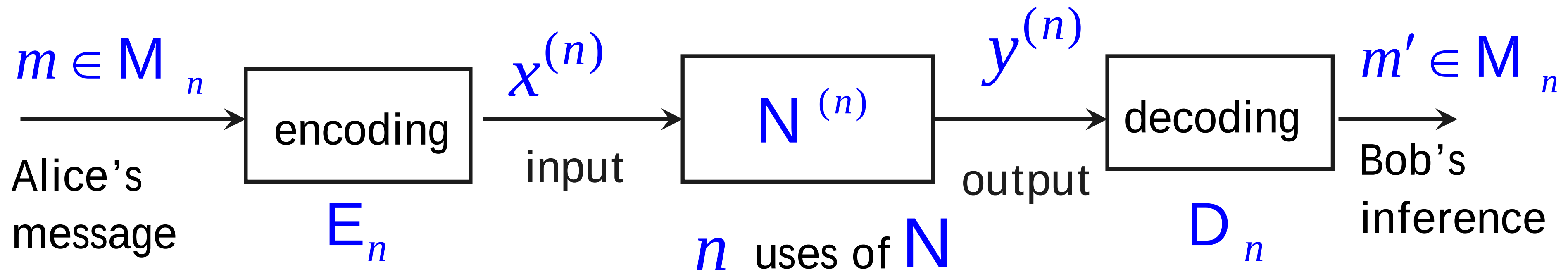
- Alice encodes: messages $\longrightarrow$ codewords ;

codewords $\longrightarrow$ ( n uses of) the channel



- *codeword:* $x^{(n)} = (x_1, x_2, \ldots, x_n);\ x_i \in \{0,1\}^n$
- *encoding:* $\mathrm{E}_n : \mathrm{M} \mapsto \{0,1\}^n$
- *output:* $y^{(n)} = (y_1, y_2, \ldots, y_n);\ \mathrm{y}_i \in \{0,1\}^n$
- *decoding:* $\mathrm{D}_n : \{0,1\}^n \mapsto \mathrm{M}$
- *Error-correcting code:* $\mathrm{C}_n := (\mathrm{E}_n, \mathrm{D}_n):$

$$\boxed{\begin{array}{c} \mathrm{N}^{(n)}: \\ p(y^{(n)} \mid x^{(n)}) \end{array}}$$



$m \in \mathcal{M}_n$ → encoding → $x^{(n)}$ → $\mathcal{N}^{(n)}$ → $y^{(n)}$ → decoding → $m' \in \mathcal{M}_n$

Alice's message

$\mathcal{E}_n$

input

$n$ uses of $\mathcal{N}$

output

$\mathcal{D}_n$

Bob's inference

- *If* $m' \neq m$ *then an error occurs!*
- *Information transmission is reliable:*

  *Prob. of error* $\rightarrow 0$ *as* $n \rightarrow \infty$

- *Rate of info transmission* = number of bits of message transmitted per use of the channel $= \dfrac{\text{size of message (in bits)}}{\text{size of codeword (in bits)}}$

- *Aim: achieve reliable transmission whilst optimizing the rate*
- *Capacity: optimal rate of reliable information transmission*

UNIVERSITY OF
CAMBRIDGE

- Shannon in his Noisy Channel Coding Theorem:

-- obtained an explicit expression for the capacity of a memoryless classical channel

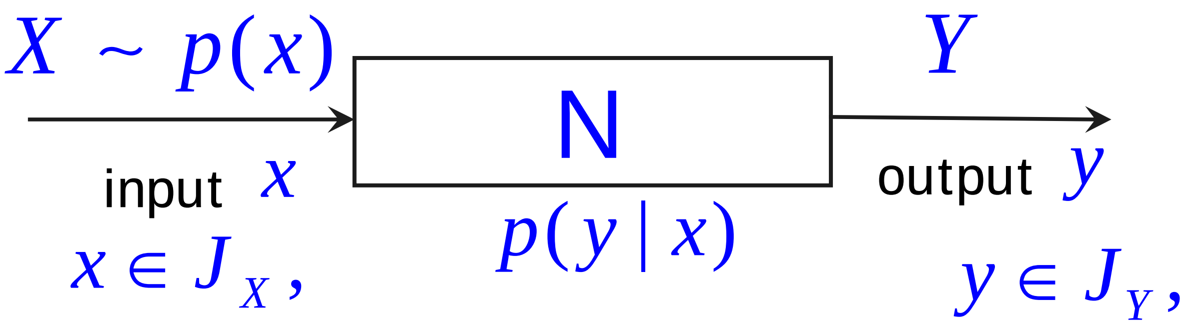
$$p(y^{(n)} \mid x^{(n)}) = \prod_{i=1}^{n} p(y_i \mid x_i)$$

**Memoryless (classical or quantum) channels**

- action of each use of the channel is identical and it is independent for different uses

-- i.e., the noise affecting states transmitted through the channel on successive uses is assumed to be uncorrelated.

- Classical memoryless channel: a schematic representation



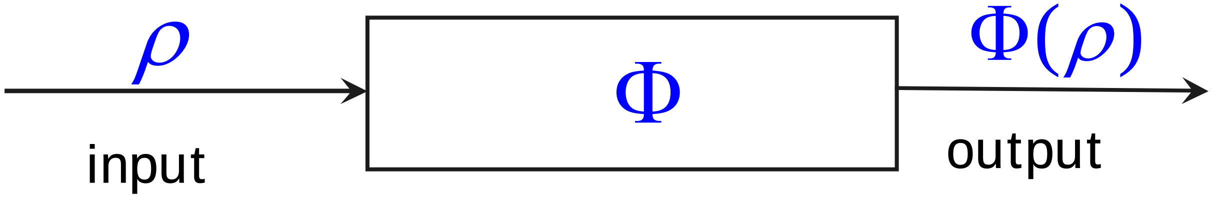
- channel: a set of conditional probs. $\{p(y|x)\}$

- Capacity $\boxed{C(\mathrm{N}) = \max_{\{p(x)\}} I(X:Y)}$

*input distributions* *mutual information*

$$I(X:Y) = H(X) + H(Y) - H(X,Y)$$

*Shannon Entropy* $H(X) = -\sum_{x} p(x) \log p(x)$



$$\boxed{\Phi(\rho) \neq \rho}$$

- A classical channel has a unique capacity

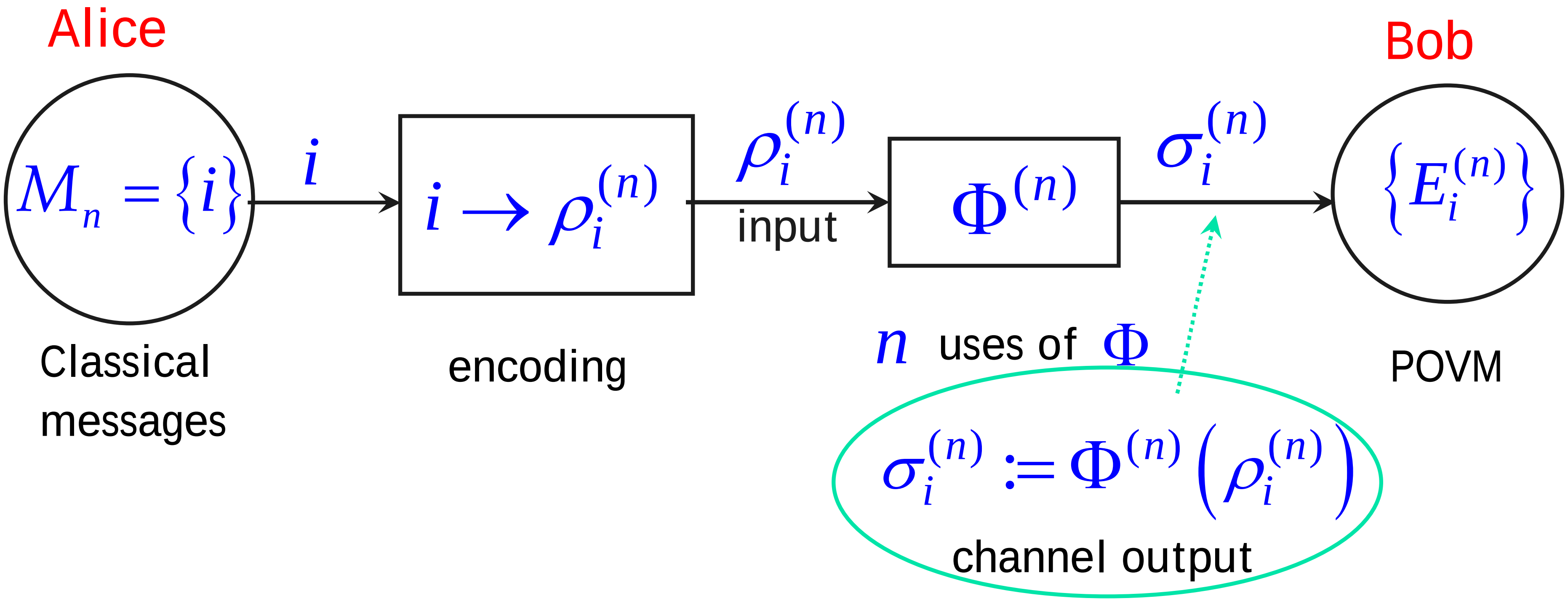
BUT

a quantum channel has various different capacities

-- This is due to the greater flexibility in the use of a quantum channel

- 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

# Transmission of Classical Info through a quantum channel



Alice

$M_n = \{i\}$

Classical messages

$i$

$i \to \rho_i^{(n)}$

encoding

$\rho_i^{(n)}$

input

$\Phi^{(n)}$

$n$ uses of $\Phi$

$\sigma_i^{(n)}$

$\sigma_i^{(n)} := \Phi^{(n)}\left(\rho_i^{(n)}\right)$

channel output

Bob

$\left\{E_i^{(n)}\right\}$

POVM

- Probability(Bob infers $i$ correctly)= $\mathrm{Tr}\left(E_i^{(n)}\sigma_i^{(n)}\right)$
- Average probability of error:

$$\boxed{p_{av}^{(n)} = \frac{1}{|M_n|}\sum_{i\in M_n}\left[1-\mathrm{Tr}\left(E_i^{(n)}\sigma_i^{(n)}\right)\right]}$$

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

- In this case, any $R \leq \liminf_{n\to\infty} \frac{\log|M_n|}{n}$ is said to be an achievable rate

$|M_n| =$ number of messages in $M_n$

$$\frac{\text{number of bits of classical message}}{\text{number of uses of } \Phi}$$

- $C(\Phi):$ Classical capacity of $\Phi$
  = the maximum amount of classical info (in bits) that can be reliably transmitted per use of

$$\boxed{C(\Phi) = \sup R}$$

--the supremum taken over all achievable rates

- If Alice restricts her codewords to product states, i.e., if

$$i \to \rho_i^{(n)} = \rho_{i_1} \otimes \rho_{i_2} \otimes \ldots\ldots \otimes \rho_{i_n}$$

- And Bob does a collective measurement (POVM) on

$\sigma_i^{(n)} := \Phi^{(n)}\left(\rho_i^{(n)}\right)$ : the output of $n$ uses of the channel

- Capacity : product state capacity $C^{(1)}(\Phi)$

# Memoryless (classical and quantum ) channels

- action of each use of the channel is identical and it is independent for different uses

$\Phi$

-- i.e., the noise affecting states transmitted through the channel on successive uses is assumed to be uncorrelated.

- Let $\Phi^{(n)}$ : $n$ successive uses of a quantum channel $\Phi$
- For a memoryless channel: $\boxed{\Phi^{(n)} = \Phi^{\otimes n}}$

# Multiple uses of a memoryless channel

- Consider a memoryless channel defined by

$$\Phi(\rho) = \sum_{i=1}^{M} A_i \rho A_i^\dagger \qquad \forall\ \rho \in \mathcal{D}(\mathcal{H}\ )$$

- Then the output of $n$ uses of the channel is given by

$$\Phi^{(n)}\left(\rho^{(n)}\right) \equiv \Phi^{\otimes n}\left(\rho^{(n)}\right),\ \ \forall\ \rho^{(n)} \in \mathcal{D}(\mathcal{H}^{\ \otimes n})$$

- where

$$\boxed{\Phi^{(\otimes n)}\left(\rho^{(n)}\right) = \sum_{k_1,\ldots,k_n}^{M} \left(A_{k_1} \otimes \ldots. \otimes A_{k_n}\right)\rho^{(n)}\left(A_{k_1}^\dagger \otimes \ldots. \otimes A_{k_n}^\dagger\right)}$$

Transmission of classical info through a memoryless quantum channel

$$\boxed{\Phi^{(n)} = \Phi^{\otimes n}}$$

- For product state inputs: $i \to \rho_i^{(n)} = \rho_{i_1} \otimes \rho_{i_2} \otimes ...... \otimes \rho_{i_n}$
- Outputs = product states

$$\sigma_i^{(n)} := \Phi^{\otimes n}\left(\rho_i^{(n)}\right) = \Phi(\rho_{i_1}) \otimes \Phi(\rho_{i_2}) \otimes ...... \otimes \Phi(\rho_{i_n})$$

- Product State Capacity $C^{(1)}(\Phi)$ given by the
  - Holevo-Schumacher-Westmoreland (HSW) Theorem

- HSW Theorem

$$\boxed{C^{(1)}(\Phi) = \max_{\{p_i,\rho_i\}} \chi\left(\{p_{i,}\Phi(\rho_i)\}\right) = \chi^*(\Phi)}$$

*Holevo Capacity*

- where

$$\boxed{\chi\left(\{p_{i,}\Phi(\rho_i)\}\right) = S\left(\sum_i p_i\Phi(\rho_i)\right) - \sum_i p_i S\left(\Phi(\rho_i)\right)}$$

$$S(\sigma) = -\mathrm{tr}\sigma\log\sigma$$ : von Neumann entropy

- Holevo $\chi -$ quantity: Let $\sigma_i := \Phi(\rho_i)$

$$\boxed{\chi\left(\{p_{i,}\sigma_i\}\right) = S\left(\sum_i p_i\sigma_i\right) - \sum_i p_i S(\sigma_i)}$$

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

$$\boxed{\chi\big(\{p_{i,}\sigma_i\}\big) := S\Big(\sum_i p_i\sigma_i\Big) - \sum_i p_i S(\sigma_i)}$$

- Holevo Bound

The maximum amount of classical info that Alice can send to Bob using $\{p_i, \sigma_i\}$ is

$$\leq \chi\big(\{p_{i,}\sigma_i\}\big)$$

$$\boxed{\chi\big(\{p_{i,}\sigma_i\}\big) \to S(\sigma) \quad \text{where} \quad \sigma := \sum_i p_i\sigma_i}$$

if the $\sigma_i$ are pure $\because S(\sigma_i) = 0$

- HSW Theorem

$$\boxed{C^{(1)}(\Phi) = \max_{\{p_i,\rho_i\}} \chi\left(\{p_{i,}\Phi(\rho_i)\}\right) = \chi^*(\Phi)}$$

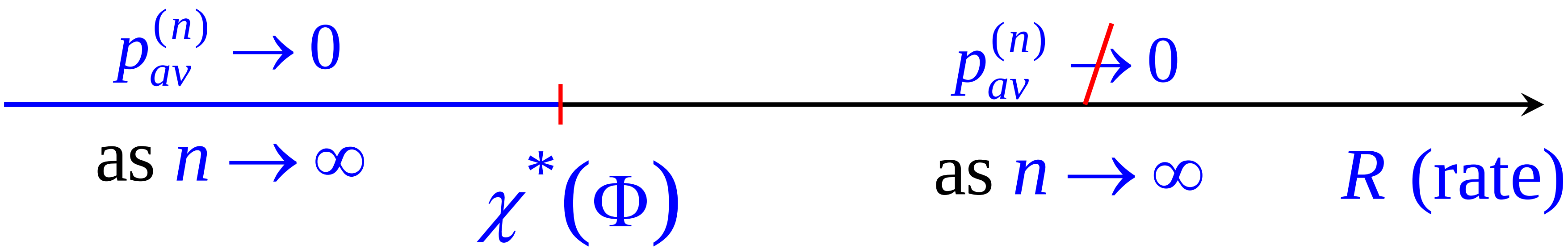
*Holevo Capacity*

- tells us that the Holevo bound can be achieved --

IF Alice uses product state inputs

& Bob does a collective measurement

- Optimal signal ensemble



- Classical capacity of a memoryless channel $\Phi$ :

(without the restriction of inputs being product states) :

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

*regularised Holevo capacity*

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

- Classical capacity of a memoryless channel

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

(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:

$$\boxed{\chi^*(\Phi_1 \otimes \Phi_2) = \chi^*(\Phi_1) + \chi^*(\Phi_2)} \Rightarrow \chi^*(\Phi^{\otimes n}) = n\chi^*(\Phi)$$

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

- IF the Holevo capacity is additive for a memoryless quantum channel then using entangled inputs would not increase its classical capacity

- An interesting question:

Could entangled inputs increase the classical capacity of quantum channels with memory ?

- For **real-world communications channels**, the assumption :

  *noise is uncorrelated between successive uses of a channel*

  cannot be justified!

Hence, memory effects need to be taken into account

→ *quantum channels with memory*

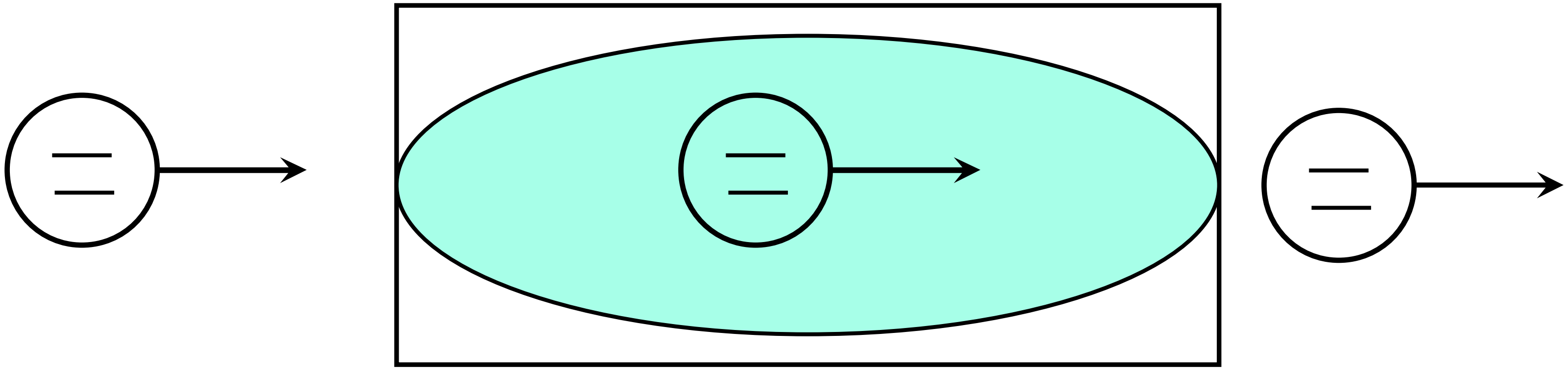
- There are various examples of quantum channels with memory :

  e.g. (1) one-atom maser or micromaser

  (2) spin-chain

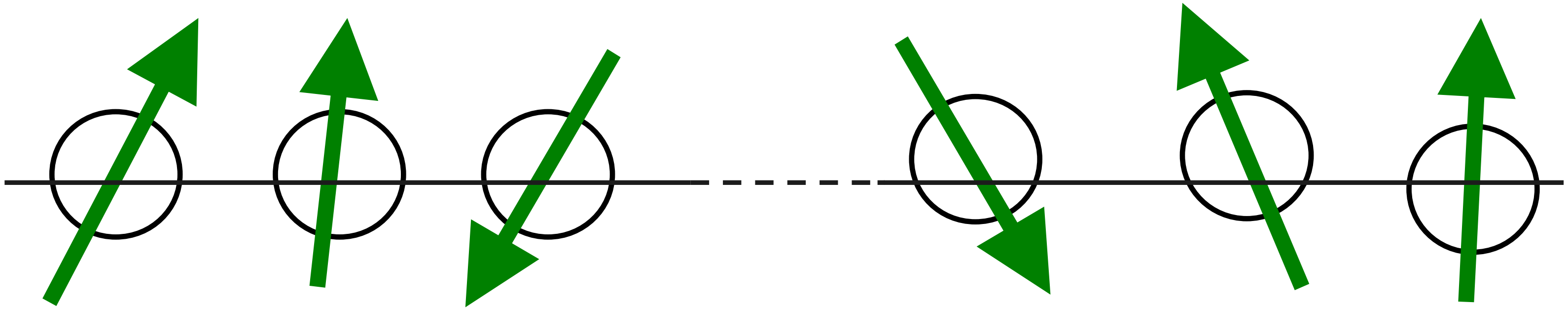
UNIVERSITY OF CAMBRIDGE

- *(1) One-atom maser or micromaser*



- A stream of two-level atoms injected into an optical cavity.
- States of these input two-level atoms: signal states
- The atoms interact with the photons in the cavity
- If these photons have sufficiently long lifetimes, then the atoms entering the cavity feel the effect of the preceding atoms
- This introduces correlations between consecutive signal states

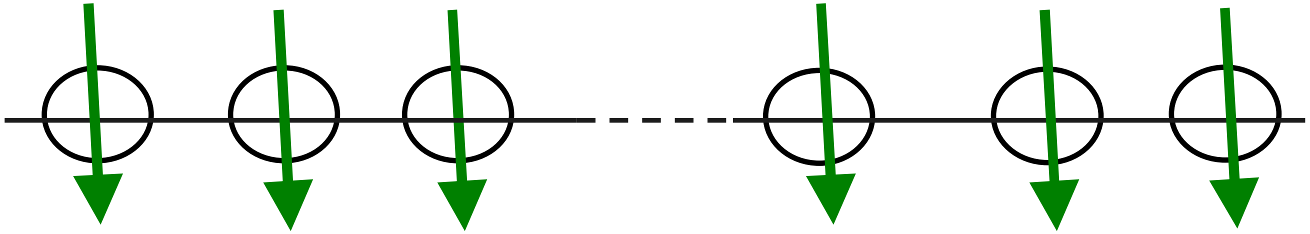
- *(2) State transfer across a spin chain*



- a spin chain : governed by a suitable Hamiltonian
- Spins at one end of the chain are prepared (by Alice) in the state which is required to be transmitted
- The spin chain is allowed to evolve for a specific amount of time under the action of the Hamiltonian; causing state to propagate
- The state is then retrieved from a set of spins at the other end of the spin chain – thus state transfer is achieved !

UNIVERSITY OF CAMBRIDGE

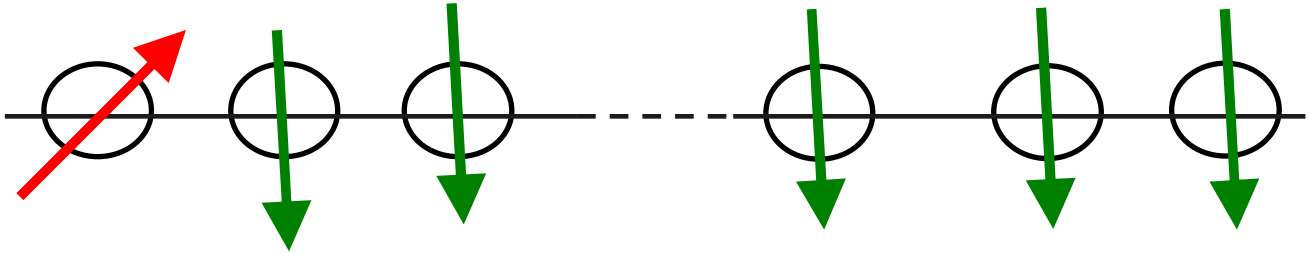
## - *(2) State transfer across a spin chain*



- When considered as a model for quantum communication:

assume : a reset of the spin chain occurs after the transmission of each signal

- e.g. by applying an external magnetic field



Alice

➡ memoryless quantum channel

- A continuous operation without reset might lead to higher transmission rates

➡ quantum channel with memory

# Exercises

(1) Use the HSW theorem to prove that any quantum channel can be used to transmit classical information, as long as it is not a constant.

(2) Use the HSW theorem to evaluate the product state capacity of a qubit depolarizing channel.

$$\Phi_{dep}(\rho) = (1-p)\rho + \frac{p}{3}\left[\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z\right]$$

UNIVERSITY OF CAMBRIDGE

# Memory Effects in Quantum Channels

## LECTURE II

## Forgetful Channels

UNIVERSITY OF
CAMBRIDGE

# Lecture I - revisited

- IF Alice sends classical info through a quantum channel $\Phi$
- using product state inputs
- & Bob does a collective measurement
- Then capacity : Product-state capacity

*HSW Theorem*

$$C^{(1)}(\Phi) = \max_{\{p_i,\rho_i\}} \chi\left(\{p_{i,}\Phi(\rho_i)\}\right) = \chi^*(\Phi)$$

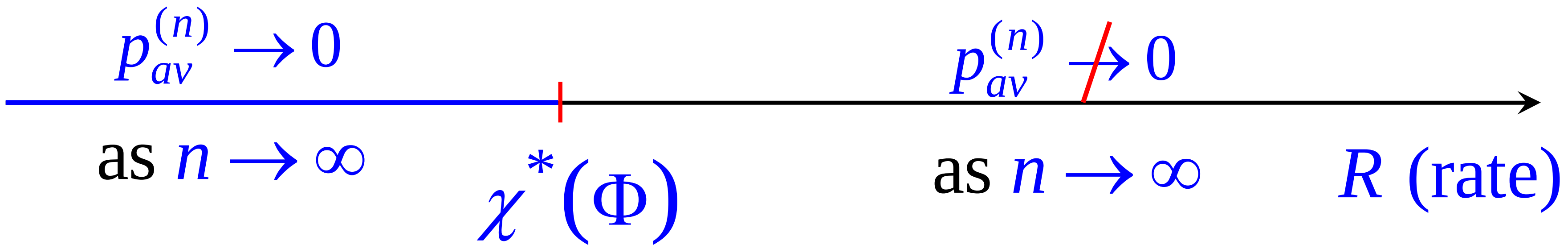
- Holevo $\chi$ – quantity

*Holevo Capacity*

$$\boxed{C^{(1)}(\Phi) = \max_{\{p_i,\rho_i\}} \chi\left(\{p_i,\Phi(\rho_i)\}\right) = \chi^*(\Phi)}$$

*Holevo Capacity*

- Optimal signal ensemble



$p_{av}^{(n)} \to 0$

as $n \to \infty$

$\chi^*(\Phi)$

$p_{av}^{(n)} \not\to 0$

as $n \to \infty$

$R$ (rate)

- Classical capacity of a memoryless channel

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

*Regularised Holevo capacity*

## *Quantum channels with memory* = *quantum memory channels*

- **Strategy :** (i) start with a simple example of a memoryless quantum channel

  (ii) using it, construct a quantum memory channel

Qubit depolarizing channel $\Phi_{dep} : B(H) \to B(H)$; $H \simeq \mathbb{C}^2$

$$\Phi_{dep}(\rho) = (1-p)\rho + \frac{p}{3}\left[\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z\right] \quad \forall\ \rho \in D(H)$$

$$= \sum_{i=1}^{4} p_i \sigma_i \rho \sigma_i \qquad p_1 = (1-p)\ ; \quad p_2 = p_3 = p_4 = \frac{p}{3}$$

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};\ \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\ \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix};\ \sigma_4 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\sigma_x \qquad\qquad \sigma_y \qquad\qquad \sigma_z$$

- Consider n successive uses of $\boxed{\Phi_{dep}(\rho) = \sum_{i=1}^{4} p_i \sigma_i \rho \sigma_i}$
  - since the channel is memoryless $\Phi_{dep}^{(n)} = \Phi_{dep}^{\otimes n}$
  - input : $\rho^{(n)} \in D(H^{\otimes n})$ ; the output is given by

$$\boxed{\Phi_{dep}^{\otimes n}(\rho^{(n)}) = \sum_{i=1}^{4} p_{i_1} p_{i_2} \ldots . p_{i_n} \left(\sigma_{i_1} \otimes \sigma_{i_2} \otimes \ldots \otimes \sigma_{i_n}\right)\left(\rho^{(n)}\right)\left(\sigma_{i_1} \otimes \sigma_{i_2} \otimes \ldots \otimes \sigma_{i_n}\right)}$$

- Let $p_{i_1 i_2 \ldots\ldots i_n}$ = joint prob. of the n successive qubits being acted on by $\sigma_{i_1}, \sigma_{i_2}, \ldots., \sigma_{i_n}$ resply.

$$\boxed{p_{i_1 i_2 \ldots\ldots i_n} = p_{i_1} p_{i_2} \ldots . p_{i_n}}$$

- this is in keeping with the notion that the noise acts independently on each successive use

- **Next:**
  - we consider an interesting generalization of this model
  - which yields a model of a quantum memory channel
- This generalization involves a :

discrete-time Markov Chain

# MARKOV CHAINS

- simplest mathematical models for random phenomena evolving in time

- It is a random process with the characteristic property that it retains no memory of where it has been in the past

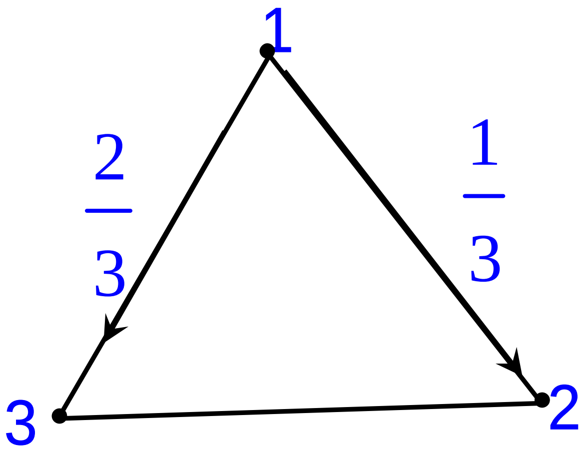
- so only the current state of the process can influence where it goes next.

Discrete-time Markov Chain:

- time is discrete
- the instants of time are labelled by $n \in Z^{+} = \{0,1,2,...\}$

# An Example

- Consider a fly hopping on the vertices of a triangle



$I = \{1,2,3\}$ = state space of the MC

- Suppose the fly hops
  - clockwise with prob. $\frac{1}{3}$
  - anticlockwise with prob. $\frac{2}{3}$

- So where it hops next depends only on where it is now

$q_{12} =$ Prob( hops to 2 in next step | it is at 1) $= \frac{2}{3}$

$q_{13} =$ Prob ( hops to 3 in next step | it is at 1) $= \frac{1}{3}$ etc.

$q_{ij} = P($next state is j | current state is i$)$ ; transition probability

$q_{ij},\ i, j = 1, 2, 3$ are the elements of a matrix $Q$

$Q$ : Transition matrix $q_{ij} = Q_{ij}$

It is a stochastic matrix $\sum_j q_{ij} = 1$

n-step transition probability

$$q_{ij}^{(n)} = P(\text{state after n steps is j} \mid \text{current state is i}) = \left(Q^n\right)_{ij}$$

A distribution on the state space $I$ is given by

$\lambda = \{\lambda_i\}_{i \in I}$ ; $\lambda_i \geq 0, \sum_{i \in I} \lambda_i = 1$ *(probabilities)*

- Invariant distribution

$$\lambda = \{\lambda_i\}_{i \in I} \;; \qquad \boxed{\lambda = \lambda Q \;;} \qquad \lambda_j = \sum_j \lambda_i q_{ij} \;;$$

- many of the long-term properties of a MC depend on its invariant distribution

- A Markov Chain is defined by a sequence of random variables

$X_0, X_1, ... X_n;\ (X_n)_{n \geq 0}$ ; each $X_i$ takes values in $I$

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i \;, ... X_n = i_n) = P(X_{n+1} = i_{n+1} \mid X_n = i_n)$$

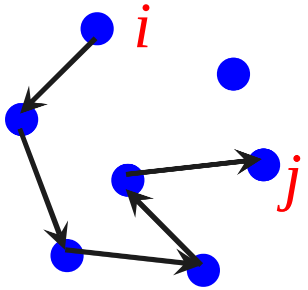
*Markov Property*

$P(X_{n+1} = j \mid X_n = i) = q_{ij}$ = *transition probability*

# Some properties of a Markov Chain

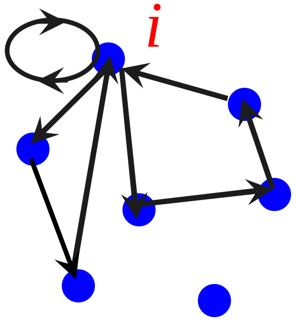
- Irreducibility:

-- a Markov Chain is said to be irreducible if it is possible to go from any state to any other state in the chain



- Aperiodicity :

-- a Markov Chain is said to be aperiodic if the return time to any state in the chain is not periodic (or if it has a period = 1) i.e., return can occur at irregular times



- n uses of a memoryless depolarising channel :

$$\Phi_{dep}^{\otimes n}(\rho^{(n)}) = \sum_{i=1}^{4} p_{i_1 i_2 .....i_n} \left(\sigma_{i_1} \otimes \sigma_{i_2} \otimes ... \otimes \sigma_{i_n}\right)\left(\rho^{(n)}\right)\left(\sigma_{i_1} \otimes \sigma_{i_2} \otimes ... \otimes \sigma_{i_n}\right)$$

$$\boxed{p_{i_1 i_2 .....i_n} = p_{i_1} p_{i_2} ....p_{i_n}}$$

- Now consider the case in which: $\boxed{p_{i_1 i_2 .....i_n} = \gamma_{i_1} q_{i_1 i_2} ....q_{i_{n-1} i_n}}$

$q_{ij}, \; i, j = 1, 2, 3, 4:$ the elements of the transition matrix $Q$

of a discrete-time Markov Chain with state space $I = \{1, 2, 3, 4\}$

- Note : the states $1, 2, 3, 4,$ label the matrices

$\sigma_1, \sigma_2, \sigma_3, \sigma_4$ of the depolarizing channel

- In this case : the output after n uses

$$\Phi^{(n)}_{dep}(\rho^{(n)}) = \sum_{i_1,..,i_n} \gamma_{i_1} q_{i_1 i_2} .. q_{i_{n-1} i_n} \left(\sigma_{i_1} \otimes .. \otimes \sigma_{i_n}\right)\left(\rho^{(n)}\right)\left(\sigma_{i_1} \otimes .. \otimes \sigma_{i_n}\right)$$

$$q_{i_{k-1} i_k} = P(k^{th} \text{qubit acted on by } \sigma_{i_k} \mid (k-1)^{th} \text{qubit acted on by } \sigma_{i_{k-1}})$$

$$\gamma_{i_1} = P\left(\text{the } 1^{st} \text{ qubit acted on by } \sigma_{i_k}\right)$$

- the noise acting on the $k^{th}$ qubit depends on the noise acting on the $(k-1)^{th}$ qubit

- **Note :** the noise acting on successive qubits is correlated

model of a quantum memory channel – with Markovian correlated noise

Macchiavello & Palma : -- introduced this model

-- studied the transmission of classical information through 2 successive uses of this quantum memory channel with

$$\boxed{q_{ij} = (1\text{-}\mu)\gamma_j + \mu\delta_{ij} \; ; 0 \leq \mu \leq 1} \qquad i, j = 1, 2, 3, 4 :$$

- with prob. $\mu$ the 2 qubits are acted on identically
- with prob. $(1-\mu)$ the action of the channel on the 2 qubits is uncorrelated

$\mu$ : the degree of memory of the channel

$\mu = 0$ : uncorrelated noise ●

$\mu = 1$ : fully correlated noise (successive actions identical)

Macchiavello & Palma : -- showed that

above a certain threshold value of the parameter $\mu$

entangled inputs increase the Holevo $\chi -$quantity

for 2 successive uses of the channel

- This suggests that above this value of $\mu$:

One might be able to transmit a higher amount of classical information through this channel by using entangled input states

- They did not, however, compute the capacity of the channel

- **Next:**
- we consider a more general quantum memory channel

  with Markovian correlated noise

  (of which the above model is a special case)
- and study its capacities

- The model is constructed from

  ---- a finite set of memoryless quantum channels:

$$\left\{\phi_1, \phi_2, \ldots, \phi_M\right\}$$

*qubit channels*

$$\forall i = 1, 2, .., M, \quad \phi_i : B(H\ ) \to B(H\ ); \quad H\ \simeq \mathrm{C}^2$$

# Quantum Channel with Markovian Correlated Noise

- $n$ uses of the channel

$$\boxed{\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,\dots,i_n=1}^{M} \gamma_{i_1} q_{i_1 i_2} \dots . q_{i_{n-1} i_n} \left(\phi_{i_1} \otimes \dots \otimes \phi_{i_n}\right)\left(\rho^{(n)}\right)}$$

$q_{ij}$ elements of the transition matrix of a discrete-time Markov chain with finite state space $\mathrm{I} = \{1, 2, \dots, M\}$

$\{\gamma_i\}$ = invariant distribution

- For each $i \in \mathrm{I}$, $\phi_i$ CPT map on $B(H)$:
- On each use of the channel, one of the given set of CPTP maps $\{\phi_1, \phi_2, \dots, \phi_M\}$ acts on the qubit

Depending on the nature of the Markov Chain the channel

Either : (1) forgetful or (2) not-forgetful

(1) Forgetful channel : a channel in which the correlation in the noise dies out with time

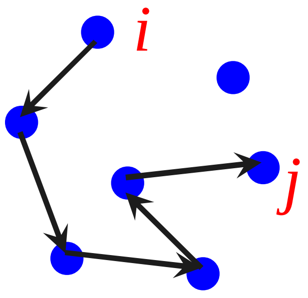
(2) not-forgetful channel : a channel with long-term memory

(Q) When is the quantum memory channel with Markovian correlated noise forgetful?

UNIVERSITY OF CAMBRIDGE

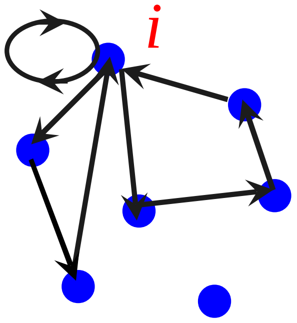
## (A) If the underlying Markov chain is

(1) irreducible



**AND**

(2) aperiodic



- In this case the Markov Chain has a unique invariant distribution and it satisfies the property called

"convergence to equilibrium"

$$i, j \in I, \qquad q_{ij}^{(n)} \xrightarrow[n\to\infty]{} \gamma_j$$

no dependence on $i$

n-step transition probability

$\{\gamma_j\}_{j\in I}$ invariant distribution of the Markov chain

If

# Quantum Channel with Markovian Correlated Noise

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,..,i_n=1}^{M} \gamma_{i_1} q_{i_1 i_2} \ldots q_{i_{n-1} i_n} \left(\phi_{i_1} \otimes \ldots \otimes \phi_{i_n}\right)\left(\rho^{(n)}\right)$$

*satisfies* “convergence to equilibrium”

$$i, j \in \mathrm{I}, \qquad q_{ij}^{(n)} \xrightarrow[n\to\infty]{} \gamma_j$$

- $\Rightarrow$ For $n$ large enough, the prob. that the $n^{th}$ qubit sent through the channel is acted upon by the memoryless channel $\phi_j$ does not depend on which memoryless channel $\phi_i$ acted on the first qubit.

## In this case (of a forgetful channel) :

- The classical capacity of the channel is given by a formula which is very similar to that of a memoryless channel

- For a memoryless channel

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

$$= \lim_{n\to\infty} \frac{1}{n} \max_{\{p_i,\rho_i^{(n)}\}} \chi\left(\{p_i, \Phi^{\otimes n}(\rho_i^{(n)})\}\right)$$

*regularised Holevo capacity*

- For our forgetful channel

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{\{p_i,\rho_i^{(n)}\}} \chi\left(\{p_i, \Phi^{(n)}(\rho_i^{(n)})\}\right)$$

- The reason behind getting such a similar result:

can be explained by a simple double-blocking argument

- We shall consider this argument in a more general setting.

Why ?

(I) Forgetful channels form an important subclass of ALL quantum channels with memory – (not only those with Markovian correlated noise)

(II) For forgetful channels, expressions for each of the different capacities are similar to the corrs. capacity formulas for memoryless channels ---

-- and can be understood by a double-blocking argument

# General model for quantum channels with memory

- Thus far : we have studied only a small class of quantum memory channels – those in which the memory is

(i) classical and (ii) governed by an underlying Markov Chain

- Bowen & Mancini : introduced a more general model for quantum memory channels in which the memory could even be quantum.

- Kretschmann & Werner : studied this model exhaustively in the Heisenberg picture

-- they were the first to evaluate capacities of forgetful channels.

- In this model : a forgetful channel is one in which :

The effect of the initializing memory dies away with time

- **Recall:** for the Markovian correlated noise model

*condition for forgetfulness*

"convergence to equilibrium"

$$i, j \in \mathrm{I}, \qquad q_{ij}^{(n)} \xrightarrow[n \to \infty]{} \gamma_j$$

- it ensures that the initializing memory dies out asymptotically

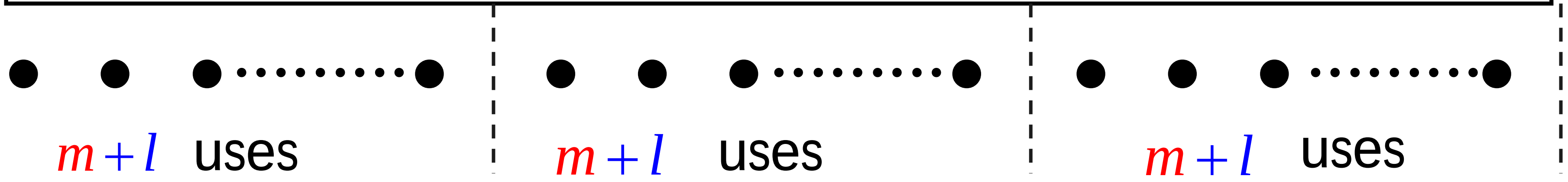
- It is easy to evaluate the capacities of forgetful channel by reducing them to a memoryless setting via a double-blocking argument

# The double-blocking argument

- Consider a strictly forgetful channel $\Phi$
- one in which : the effect of the initializing memory dies away after a finite number of uses (say, m uses)

- **e.g.** transmission of info over a quantum spin chain which is reset after every third use (m = 3).

- For processing of long messages (signal states) we group the successive uses of the channel in blocks of length (m+ l)

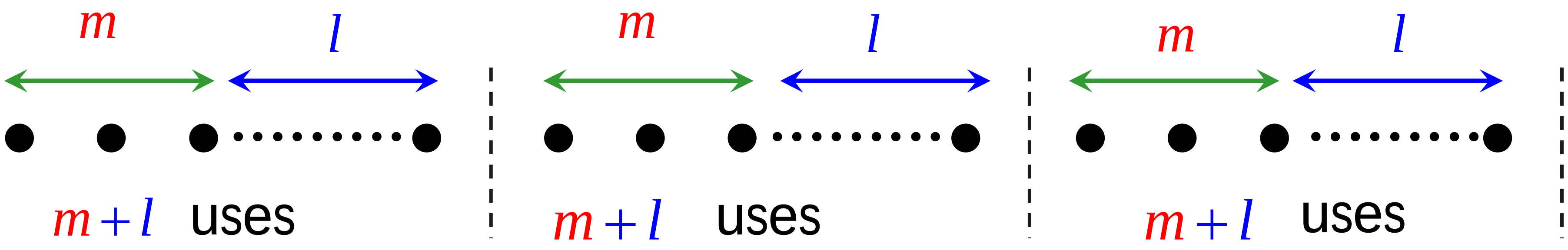


$m+l$ uses

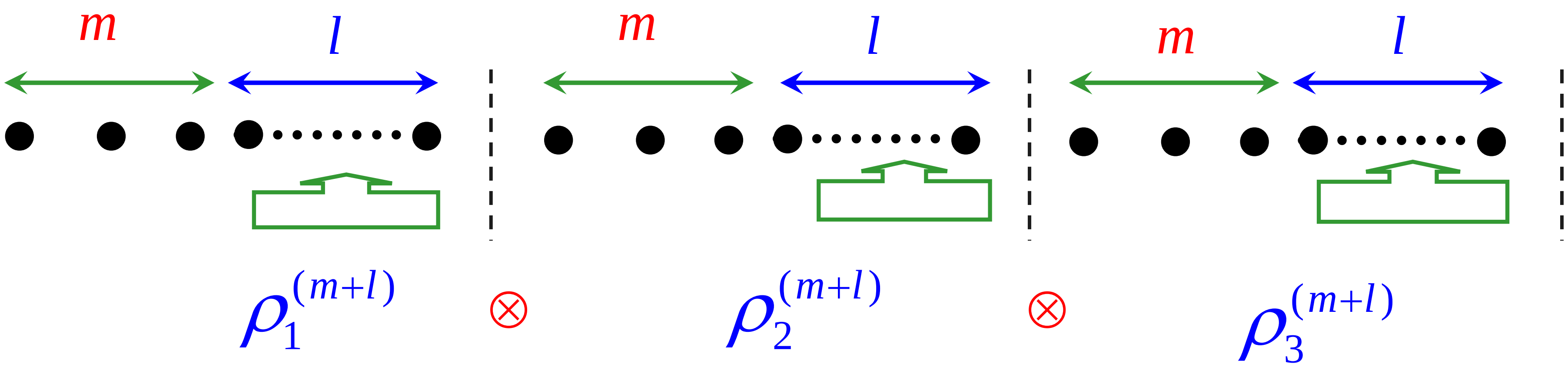
$m+l$ uses

$m+l$ uses

# Strictly forgetful channel



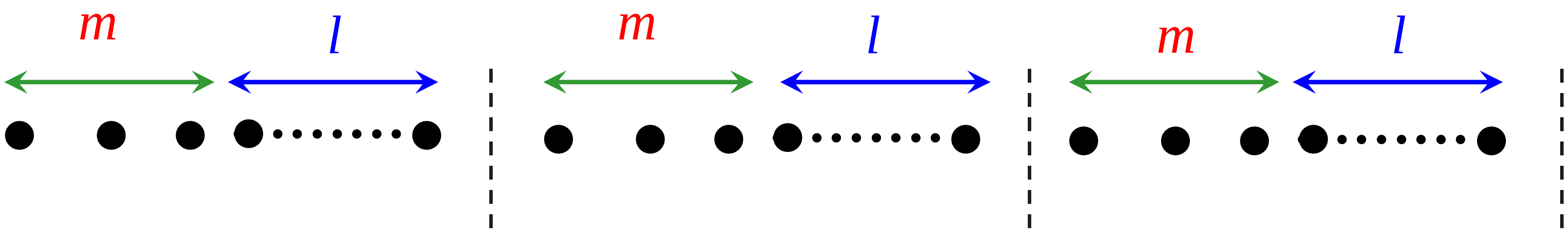
# Strictly forgetful channel



- ignore the outputs of the first $m$ channels of each such block
- actual encoding is done for the remaining $l$ blocks
- Eventually let $l \rightarrow \infty$
- If we restrict inputs to products states of block length $m+l$

$$\text{input} = \rho_1^{(m+l)} \otimes \rho_2^{(m+l)} \otimes \ldots\ldots$$

# Strictly forgetful channel



$$\rho_1^{(m+l)} \otimes \rho_2^{(m+l)} \otimes \rho_3^{(m+l)}$$

- due to the strict forgetfulness of the channel:

  -- the (relevant part of the) output state factorizes

- The whole set-up corrs. to a memoryless channel acting

  on a larger Hilbert space $\Phi^{(m+l)} \approx$ *memoryless channel*

- Problem $\xrightarrow[\textit{double-blocking}]{}$ *memoryless setting*

  for which we know the classical capacity

The same double-blocking argument can be applied to channels which are forgetful (and not just strictly forgetful)

## Classical Capacity

- For a memoryless channel

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \chi^{*}\left(\Phi^{\otimes n}\right)$$
$$= \lim_{n\to\infty} \frac{1}{n} \max_{\{p_i,\rho_i^{(n)}\}} \chi\left(\{p_{i,}\Phi^{\otimes n}(\rho_i^{(n)})\}\right)$$

*regularised Holevo capacity*

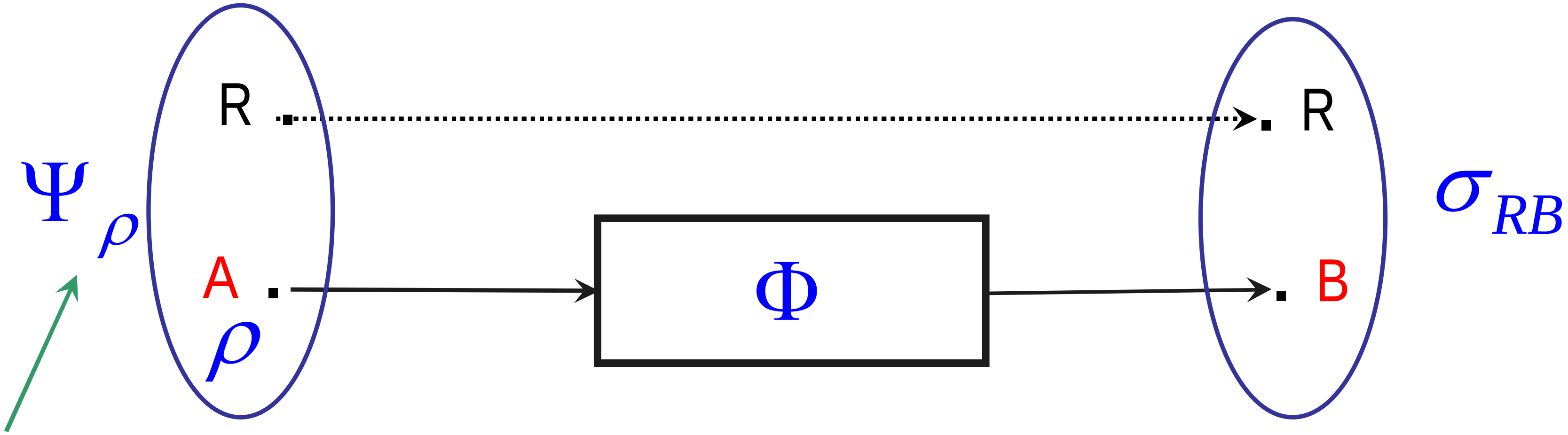
- For forgetful channels

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{\{p_i,\rho_i^{(n)}\}} \chi\left(\{p_{i,}\Phi^{(n)}(\rho_i^{(n)})\}\right)$$

- *Lloyd, Shor & Devetak:* LSD Theorem

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho^{(n)}} I_c\left(\rho^{(n)}, \Phi^{\otimes n}\right)$$

coherent information



purification of $\rho$

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

*coherent information*

$$I_c\left(\rho, \Phi\right) = -S(\sigma_{RB}) + S(\sigma_B) \quad = -S(R\,|\,B)_\sigma$$

# Quantum Capacity

- For a memoryless channel

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho^{(n)}} I_c\left(\rho^{(n)}, \Phi^{\otimes n}\right)$$

*Regularised*

*Coherent information*

- For forgetful channels

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho^{(n)}} I_c\left(\rho^{(n)}, \Phi^{(n)}\right)$$

# LECTURE III

A channel with long-term memory

(not-forgetful)

*Coding Theorem for a Class of Quantum Channels with Long-Term Memory,*

*ND and Tony Dorlas,*

*J. Phys. A: Math. Theor. 40, 8147-8164 (2007).*

# A channel with long-term memory

- The correlation in the noise does not die out with time
- evaluating their capacities is a more challenging task
- Simplest example:

*convex combinations of a finite number of memoryless channels*

$$\{\phi_1, \phi_2, \ldots, \phi_M\}$$

$$\forall i = 1, 2, .., M, \quad \phi_i : D(H_A) \to D(H_B);$$

$$\boxed{\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \phi_i^{\otimes n}\left(\rho^{(n)}\right)}$$

- $n$ uses of the channel: $\gamma_i > 0 \ \forall i = 1, 2, .., M, \ \sum_{i=1}^{M} \gamma_i = 1$

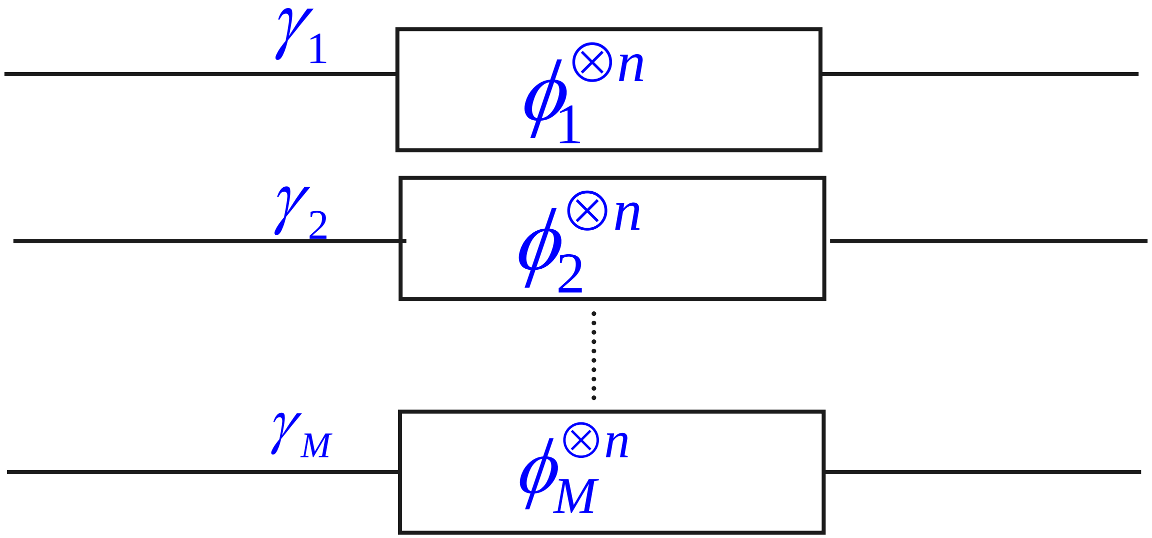
# A channel with long-term memory

$$\boxed{\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \phi_i^{\otimes n}\left(\rho^{(n)}\right)}$$

$$H_A, H_B \simeq \mathbb{C}^2$$

$$\gamma_i > 0 \ \ \forall i = 1, 2, .., M, \ \sum_{i=1}^{M} \gamma_i = 1$$

- The channel has $M$ memoryless branches



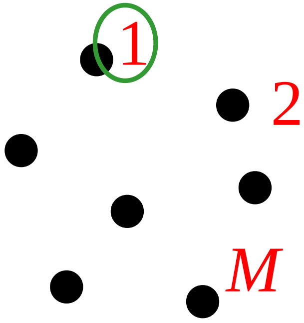
- Comparing this channel:

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \phi_i^{\otimes n}\left(\rho^{(n)}\right) \quad \text{..........(a)}$$

- with the Markovian correlated noise model:

$$\boxed{\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,..,i_n=1}^{M} \gamma_{i_1} q_{i_1 i_2} \ldots. q_{i_{n-1} i_n} \left(\phi_{i_1} \otimes \ldots \otimes \phi_{i_n}\right)\left(\rho^{(n)}\right) \quad \text{.............(b)}}$$

- We note that (a) is a special case of (b):
  - The Markov Chain has $M$ states



$$\left\{\phi_1, \phi_2, \ldots, \phi_M\right\}$$

states of the MC

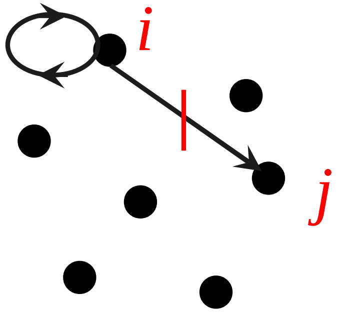
- Comparing this channel:

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \phi_i^{\otimes n}\left(\rho^{(n)}\right) \qquad \text{..........(a)}$$

- with the Markovian correlated noise model:

$$\boxed{\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,..,i_n=1}^{M} \gamma_{i_1} q_{i_1 i_2} \ldots q_{i_{n-1} i_n} \left(\phi_{i_1} \otimes \ldots \otimes \phi_{i_n}\right)\left(\rho^{(n)}\right) \qquad \text{.............(b)}}$$

- We note that (a) is a special case of (b):
  - The Markov Chain has $M$ states
  - $q_{ij} = \delta_{ij}$
  - aperiodic but not irreducible



- ~~Convergence to equilibrium~~ : so it is not forgetful

- Macchiavello and Palma considered:

$$q_{ij} = (1-\mu)\gamma_j + \mu\delta_{ij} \;;\; 0 \le \mu \le 1$$

- Our choice $q_{ij} = \delta_{ij}$ corresponds to $\mu = 1$

(fully correlated noise -- successive actions identical)

Let us evaluate: the product state capacity of the channel

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \phi_i^{\otimes n}\left(\rho^{(n)}\right) \qquad \text{..........(a)}$$

- i.e., the classical capacity under the restriction of

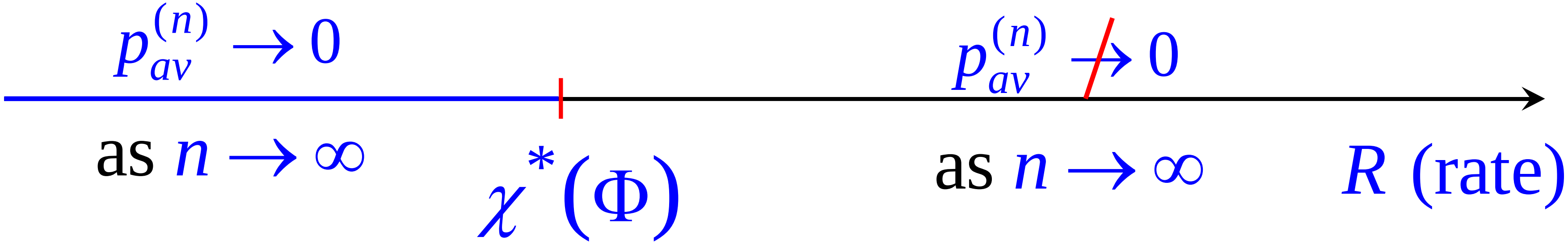
$$C^{(1)}(\Phi)$$

product-state inputs

Let us start by making a naïve guess:

- Recall : for a memoryless channel $\phi$ : *[HSW Theorem]*

$$\boxed{C^{(1)}(\phi) = \sup_{\{p_j,\rho_j\}} \chi\left(\{p_j,\phi(\rho_j)\}\right) \quad = \chi^*(\phi) \text{ .........(A)}}$$



$p_{av}^{(n)} \to 0$

as $n \to \infty$

$\chi^*(\Phi)$

$p_{av}^{(n)} \not\to 0$

as $n \to \infty$

$R$ (rate)

- Any $R \leq \chi^*(\Phi)$ is achievable.

[*HSW Theorem*]

For a memoryless channel $\phi$:

$$\boxed{C^{(1)}(\phi) = \sup_{\{p_j,\rho_j\}} \chi\left(\{p_j,\phi(\rho_j)\}\right) \quad = \chi^*(\phi) \;\ldots\ldots\ldots\text{(A)}}$$

- So in this case, because the channel has $M$ memoryless branches, one might naively expect:

$$\boxed{\begin{aligned} C^{(1)}(\Phi) &= \min_{1\le i\le M} \chi^*(\phi_i) \\ &= \min_{1\le i\le M} \max_{\{p_j,\rho_j\}} \chi\left(\{p_j,\phi_i(\rho_j)\}\right) \;\ldots\ldots\ldots\ldots\ldots\ldots\text{(B)} \end{aligned}}$$

BUT

- (B) is NOT TRUE ; min ⟷ max

- **Theorem:**

*ND & Dorlas*

The product-state capacity of the long-term memory channel

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \phi_i^{\otimes n}\left(\rho^{(n)}\right) \quad \text{..........(a)}$$

is given by

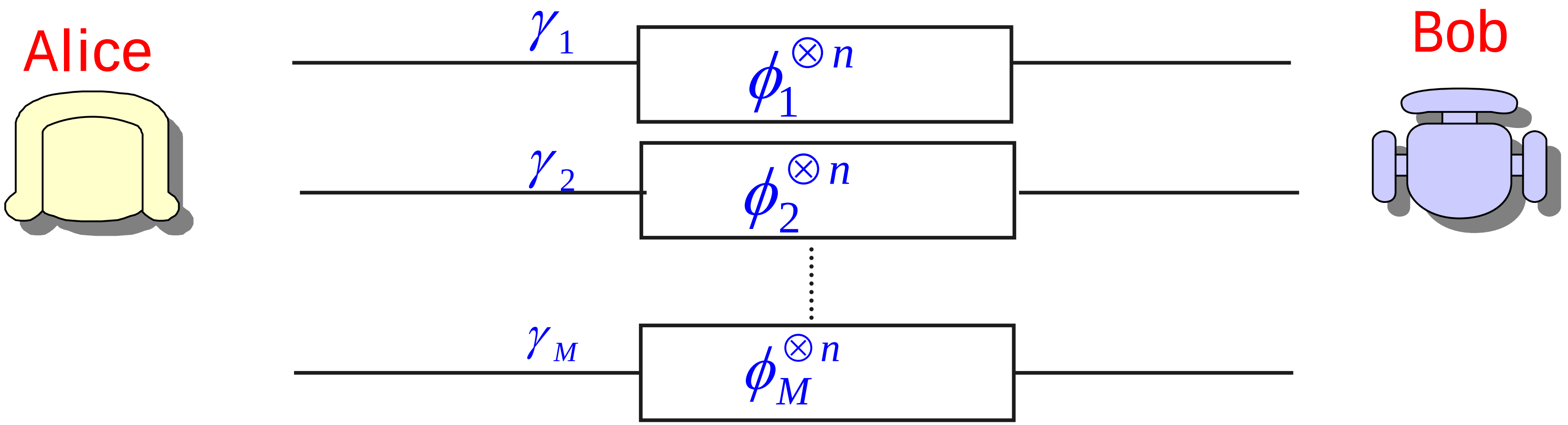
$$\boxed{C^{(1)}(\Phi) = \max_{\{p_j,\rho_j\}} \min_{1\le i\le M} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)}$$

- whereas are guess was:

$$\boxed{C^{(1)}(\Phi) = \min_{1\le i\le M} \max_{\{p_j,\rho_j\}} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)}$$

Why ?

# Sketch of the proof



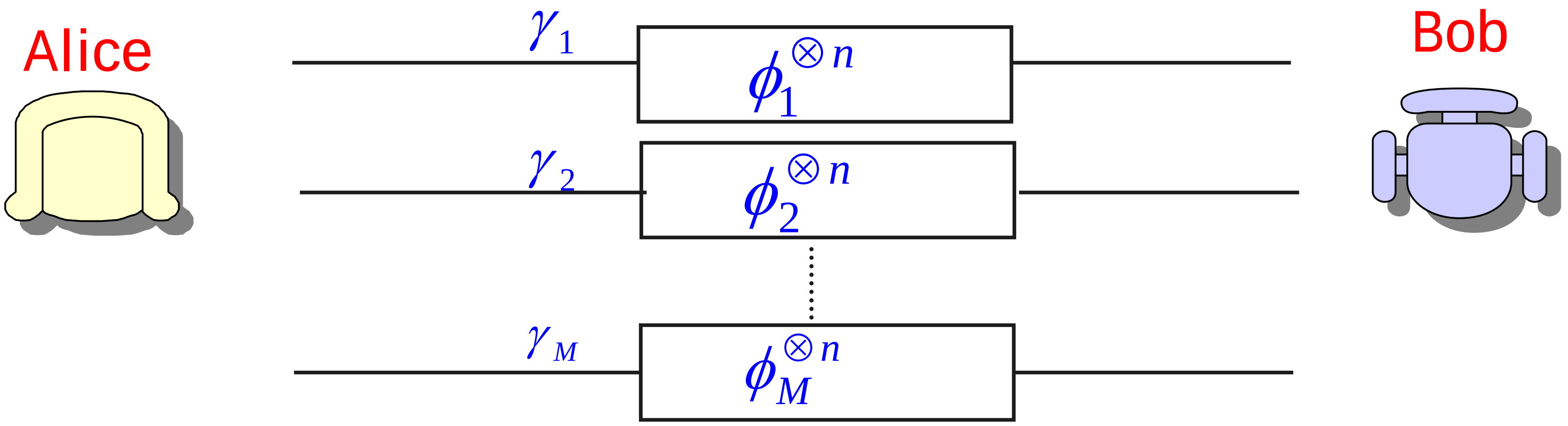
(Q) Is there any way in which Bob can find out which of the $M$ memoryless branches the qubits have been sent through?

i.e., Can Bob distinguish between the outputs of the different memoryless branches ?

- If so, then at least from his point of view:

problem → memoryless channel

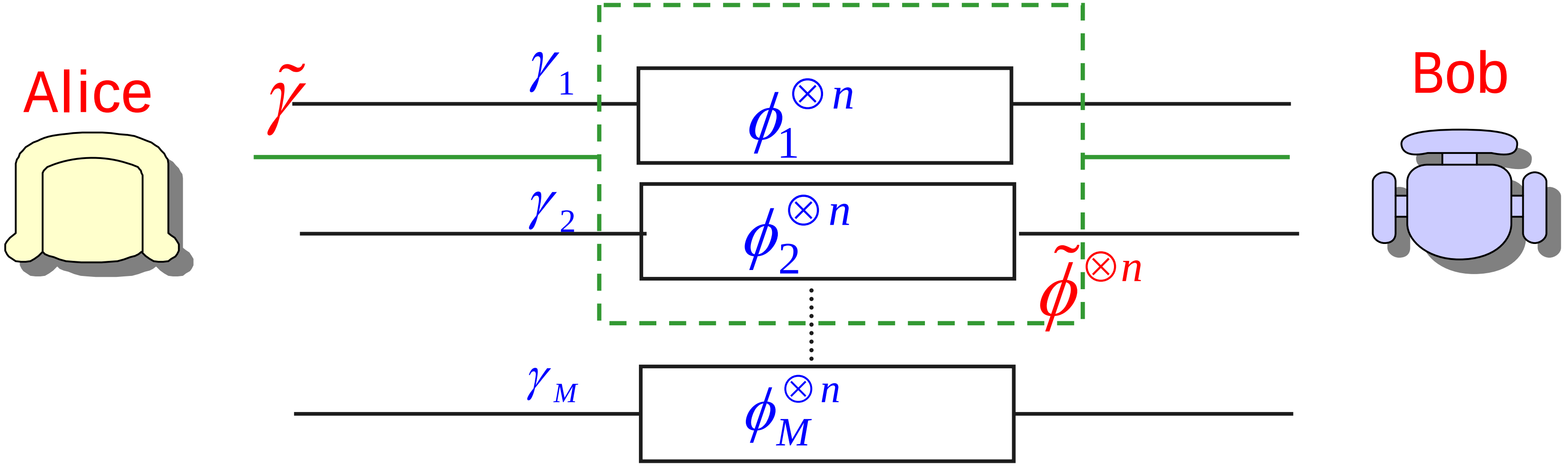
# Sketch of the proof contd.



(A) Yes - provided

- **Alice** adds a preamble to her codewords &
- **Bob** does a collective measurement on the qubits he receives

# Sketch of the proof contd.

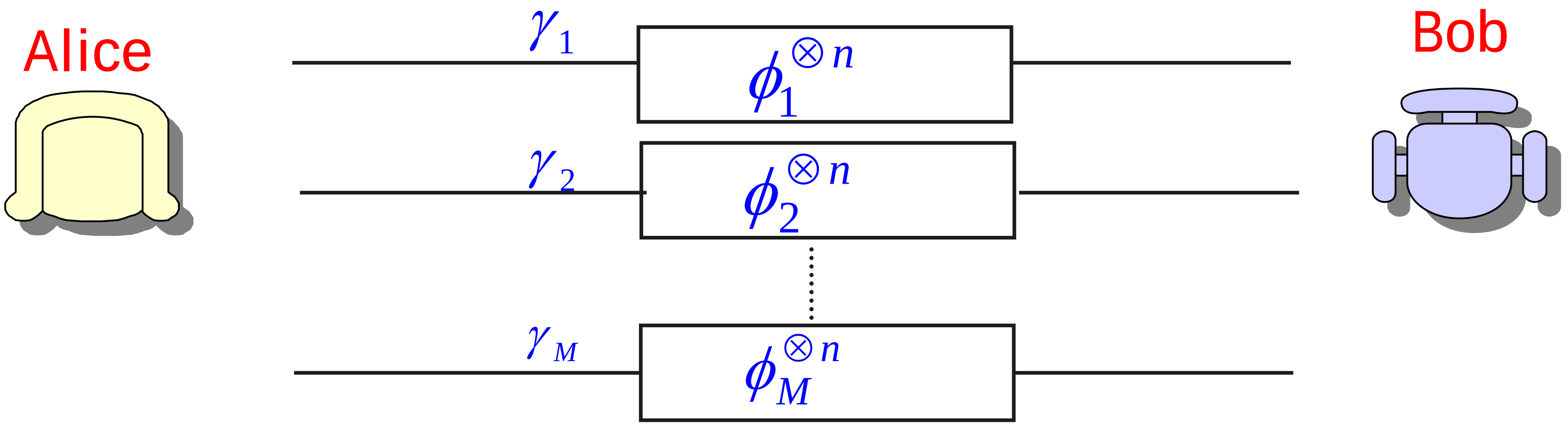


Assume: $\phi_i,\ i = 1, 2, \ldots, M$ are all different

- **Else** we do not need to distinguish between all of them
- & we can introduce a compound prob. for each set of identical branches.

- e.g. If $\phi_1 = \phi_2 = \tilde{\phi}$ prob. $\tilde{\gamma} := \gamma_1 + \gamma_2$

# Sketch of the proof contd.



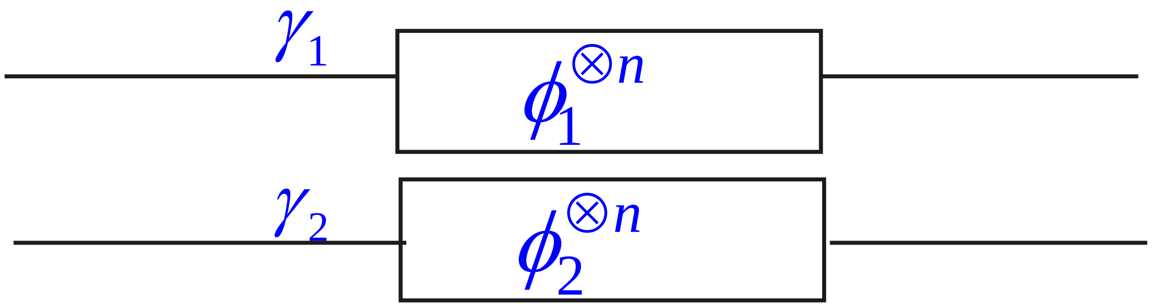
$\phi_i, \; i = 1, 2, \ldots, M$ are all different

- For each pair $\phi_i, \phi_j : \; 1 \le i, j \le M,$

-- there exists states $\omega^{(ij)},$ such that

$$\boxed{\phi_i\left(\omega^{(ij)}\right) \neq \phi_j\left(\omega^{(ij)}\right),}$$

- For simplicity consider $M = 2$



(Q) Can Bob distinguish the outputs of these 2 branches ?

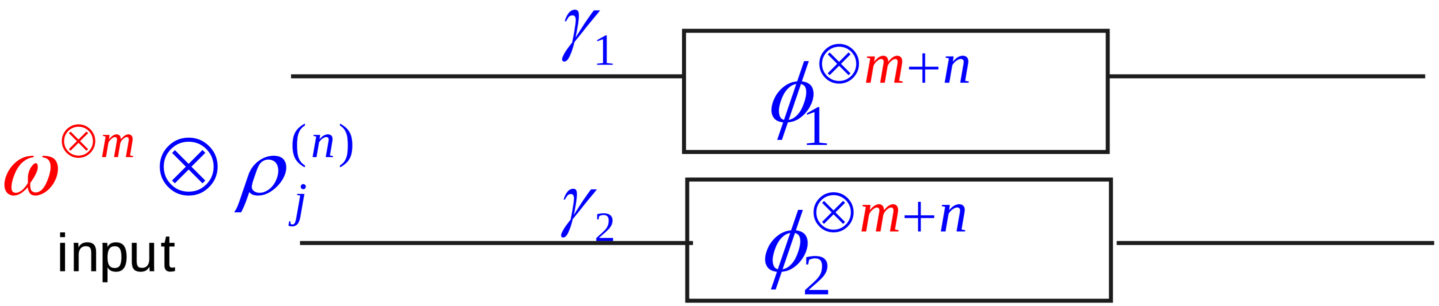
Let $\omega$ be a state such that $\phi_1(\omega) \neq \phi_2(\omega)$

- To allow Bob to distinguish between the 2 branches,

Alice adds a preamble to the input state $\rho_i^{(n)} \longmapsfrom i$

codeword

message



- Instead of encoding $j \mapsto \rho_j^{(n)}$
- She encodes $j \mapsto \omega^{\otimes m} \otimes \rho_j^{(n)}$ $(m+n)-$ qubit state

-- where $\omega$ is a state such that $\phi_1(\omega) \neq \phi_2(\omega)$

- Bob receives the state $\sigma_i := \phi_i(\omega)$

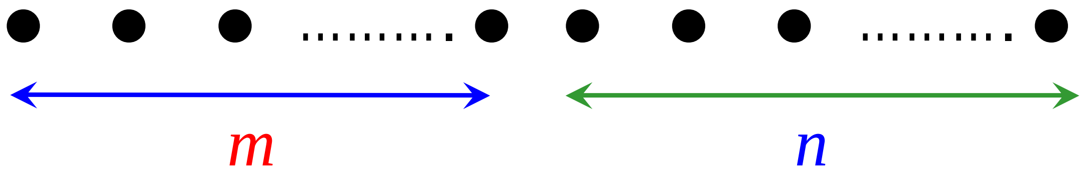
$$\phi_i^{\otimes m+n}\left(\omega^{\otimes m} \otimes \rho_j^{(n)}\right) \text{ with prob. } \gamma_i$$

$$= \left(\phi_i(\omega)\right)^{\otimes m} \otimes \phi_i^{\otimes n}\left(\rho_j^{(n)}\right) = \boxed{\sigma_i^{\otimes m} \otimes \phi_i^{\otimes n}\left(\rho_j^{(n)}\right)}$$

Alice encodes $j \to \omega^{\otimes m} \otimes \rho_j^{(n)}$

$\omega^{\otimes m}$ → **preamble**

$(m+n)-$ uses



$$\omega^{\otimes m} \otimes \rho_j^{(n)} \xrightarrow{\text{input}} \boxed{\phi_i^{(m+n)}} \xrightarrow{\text{output}} \phi_i^{(m+n)}\left(\omega^{\otimes m} \otimes \rho_j^{(n)}\right)$$

with prob. $\gamma_i$

- **Bob gets** the state $\sigma_i^{\otimes m} \otimes \phi_i^{\otimes n}(\rho_j^{(n)})$ with probability $\gamma_i$

- Let us focus on the output of the first $m$ qubits state of the first $m$ qubits that Bob receives

$$\sigma_1^{\otimes m} = [\phi_1(\omega)]^{\otimes m}$$ with probability $\gamma_1$

$$\sigma_2^{\otimes m} = [\phi_2(\omega)]^{\otimes m}$$ with probability $\gamma_2$

(Q) Can Bob do a measurement to distinguish between $\sigma_1^{\otimes m}$ & $\sigma_2^{\otimes m}$ ?

(A) Yes. Consider the operator:

$$\boxed{A^{(m)} = \gamma_1 \sigma_1^{\otimes m} - \gamma_2 \sigma_2^{\otimes m}}$$

- Let $\Pi_1^{(m)}$ : orthogonal projection onto the non-negative eigenspace of $A^{(m)}$

  and

$$\Pi_2^{(m)} = 1^{(m)} - \Pi_1^{(m)}$$

- Let Bob does a projective measurement (a la Helstrom) described by the operators $\Pi_1^{(m)}$ & $\Pi_2^{(m)}$ on the state $\sigma_j^{\otimes m}$ that he receives:

$$\sigma_j^{\otimes m}, j = 1, 2$$ with probs. $\gamma_1$ & $\gamma_2$ resply.

- For $m$ large enough, by using Helstrom's strategy,

Bob can indeed distinguish between $\sigma_1^{\otimes m}$ & $\sigma_2^{\otimes m}$

-- with arbitrarily low probability of error.

- Thus he can determine which memoryless branch the

qubits have come through!

- Bob determines which branch the input has come through
- $\therefore$ from Bob's point of view : problem reduces to

decoding codewords sent through a memoryless channel

- So now he can do the appropriate decoding operation

on the remaining output state to infer Alice's message

Measurement outcome $i$ $\Rightarrow$ remaining output state $\equiv$ $\phi_i^{\otimes n}(\rho_j^{(n)})$

codeword corrs. to Alice's message $i$

- This idea can be generalized to distinguish between all $M$ branches.

BUT

- Alice does not know what $i$ is (no feedback)

- Now one can understand why:

$$\boxed{C^{(1)}(\Phi) = \max_{\{p_j,\rho_j\}} \min_{1\le i\le M} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)}$$ ............(A)

and $$\boxed{C^{(1)}(\Phi) \neq \min_{1\le i\le M} \max_{\{p_j,\rho_j\}} \chi\left(\{p_j,\Phi_i(\rho_j)\}\right)}$$ ............(B)

- For a memoryless channel $\phi_i$

$$C^{(1)}(\phi_i) = \max_{\{p_j,\rho_j\}} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)$$

- The input ensemble for which the max is achieved $=$ optimal signal ensemble

- IF Alice knew $i$ apriori then she could encode her messages using the optimal signal ensemble for $\phi_i$ & obtain (B)

BUT Alice does NOT know $i$ apriori.

- For $\phi_i$ , for any given input ensemble $\{p_j, \rho_j\}$

Max. amount of classical info that can be sent through $\phi_i$ $\Big\} = \chi\left(\{p_j, \phi_i(\rho_j)\}\right)$

- In our channel there are $M$ memoryless branches:

$\therefore$ Max. amount of classical info that can be sent through it (for any given input ensemble $\{p_j, \rho_j\}$ ):

$$= \min_{1 \le i \le M} \chi\left(\{p_j, \phi_i(\rho_j)\}\right)$$

- & this $\Longrightarrow$ - any rate $R \le \max_{\{p_j, \rho_j\}} \min_{1 \le i \le M} \chi\left(\{p_j, \phi_i(\rho_j)\}\right)$

is achievable

- & this $\Longrightarrow$ $C^{(1)}(\Phi) \ge \max_{\{p_j, \rho_j\}} \min_{1 \le i \le M} \chi\left(\{p_j, \phi_i(\rho_j)\}\right)$

Theorem: $C^{(1)}(\Phi) = \max_{\{p_j,\rho_j\}} \min_{1\le i\le M} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)$

- We have proved

*Direct part (achievability)*

$$C^{(1)}(\Phi) \ge \max_{\{p_j,\rho_j\}} \min_{1\le i\le M} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)$$

- We also need to prove that : any rate

$R \ge \max_{\{p_j,\rho_j\}} \min_{1\le i\le M} \chi\left(\{p_j,\phi_i(\rho_j)\}\right)$ is not achievable

*Weak Converse*

- *Ingredients needed to prove the Weak Converse:*
  - Holevo bound
  - Subadditivity of the von Neumann entropy
  - Fano’s inequality

- Recall: The quantum channel with Markovian correlated noise is forgetful IF the Markov Chain is

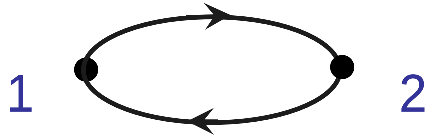
  (1) irreducible and (2) aperiodic

- The "not-forgetful" channel that we considered was

  aperiodic but not irreducuble

- Another example of a "not-forgetful" channel is one for which the Markov Chain is : irreducible but not aperiodic (i.e., memory governed by a periodic Markov Chain)

- *E.g. 2-state Markov Chain :*



- Transition Matrix

$$Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

*irreducible, periodic (period=2).*

- 2 states of the Markov Chain corrs. to 2 single qubit channels $\phi_1, \phi_2$ which act alternatively on successive inputs

$$\boxed{\Phi^{(n)}(\rho^{(n)}) = \frac{1}{2}[\underbrace{\phi_1 \otimes \phi_2 \otimes \phi_1 \otimes ....}_{n \text{ times}} + \underbrace{\phi_2 \otimes \phi_1 \otimes \phi_2 \otimes ....}_{n \text{ times}}]\left(\rho^{(n)}\right)}$$

- In this case,

$$\boxed{C^{(1)}(\phi_i) = \sup_{\{p_j,\rho_j\}} \frac{1}{2}\sum_{i=1}^{2} \chi\left(\{p_j,\phi_i(\rho_j)\}\right) = \frac{1}{2}\sum_{i=1}^{2} \chi^*(\phi_i)}$$

= average of the Holevo capacities of the individual channels

- Similarly once can consider a channel where the underlying Markov Chain has a period $L > 2$

- ND and Tony Dorlas, Coding Theorem for a Class of Quantum Channels with Long-Term Memory, J. Phys. A: Math. Theor. 40, 8147-8164 (2007).

- ND, T.C.Dorlas and Y.M. Suhov, Entanglement Assisted Classical Capacity of a Class of Quantum Channels with Long-Term Memory,' Quantum Information Processing, 7, no. 6, 251--262 (2008).

- I._Bjelakovic and H. Boche, On quantum capacity of compound channels, arXiv:0808.1007.

- ND and Tony Dorlas,

  Classical capacity of quantum channels with general Markovian correlated noise,

  J.Stat.Phys. 134, pp 1173-1195, (2009).

- I._Bjelakovic and H. Boche,

  Entanglement transmission capacity of compound channels,

  arXiv:0904.3011.

……………………………….. etc.

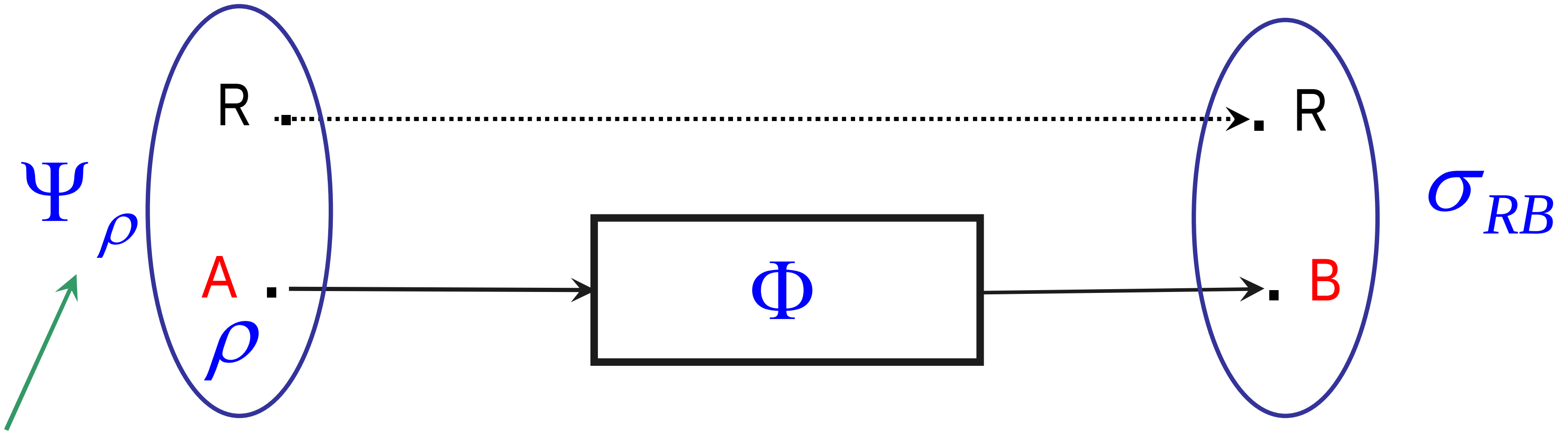
# Information Spectrum Method

- S.Verdu and T.S.Han, IEEE Trans. Inf. Theory 40, pp. 1147-1157 (1994);
- T.S.Han, Information-Spectrum Methods in Information Theory (Springer-Verlag, 2002);
- H.Nagaoka and M.Hayashi, IEEE Trans. Inf. Theory {\bf 53}, 534--549 (2007)
- F.Buscemi and N.Datta, The quantum capacity of channels with arbitrarily correlated noise, IEEE Trans. inf. Th. 56, 1447-1460 (2010),

- *Lloyd, Shor & Devetak:* LSD Theorem

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho} I_c\left(\rho, \Phi^{\otimes n}\right)$$

coherent information



purification of $\rho$

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

*coherent information*

$$I_c\left(\rho, \Phi\right) = -S(\sigma_{RB}) + S(\sigma_B) \quad = -S(R\,|\,B)_{\sigma}$$