Remark: By requiring high entanglement fidelity $F_C(c)$ for a good code, one demand for $C$ was to preserve "pure-state entanglement" (in $|\psi_{AR}\rangle$) well. In this sense, the rate $R(C)$ should be seen as an "entanglement bit (e-bit)" rate.

But Schumacher's compression theorem also holds for other accuracy notions, in particular for the "ensemble average fidelity"

$$F(C) = \sum_{x_1, \ldots, x_n} p(x_1) \ldots p(x_n) \mathbb{F}(|x_1\rangle\langle x_1| \otimes \cdots \otimes |x_n\rangle\langle x_n|),$$

quantifying how well $C$ preserves a pure product state $|x_1\rangle\otimes \cdots \otimes |x_n\rangle \in \mathcal{H}^n$ drawn at random from the pure state ensemble $\{p(x), |x\rangle\}$ in an i.i.d. fashion. This fidelity notion supports better the idea of a "gubit rate" $R(C)$.

The optimal compression rate in this case is

$$S\left(\sum_x p(x) |x\rangle\langle x|\right).$$

In Exercise 5.1 it is shown that

$$S\left(\sum_x p(x) |x\rangle\langle x|\right) \leq H\left(\{p(x)\}\right),$$

showing that we usually need fewer qubits to compress the state $|x_1\rangle\otimes \cdots \otimes |x_n\rangle$ than we need classical bits to store the message $(x_1, \ldots, x_n)$. 
II.3 - Holevo-Schumacher-Westmoreland Theorem

Quantum preliminaries

- A map $T : B(C^d) \to B(C^d)$ is called positive if $T(A) \geq 0$ for all $A \geq 0$
  
  (linear!)

  positive semidefinite order

  Example: $T = \text{id} : B(C^d) \to B(C^d)$ identity map
  
  $T = \Theta : B(C^d) \to B(C^d)$ transposition map (see Exercise 4.2)

- In quantum mechanics (tensor products!), we want more:
  
  $T : B(C^d) \to B(C^d)$ is called completely positive (CP) if $(T \otimes \text{id}_n) : B(C^d \otimes C^d) \to B(C^d \otimes C^d)$ is positive $\forall n \in \mathbb{N}$

  map $T$ acts only on one subsystem

  Schmidt decomposition

  Fact: $T \in \text{CP} \iff T \otimes \text{id}_n$ is positive

  $\iff (T \otimes \text{id}_n)(1_{d_1} \otimes \rho_{d_1}) \geq 0$

  where $1_{d_1} = \sum_{i=1}^{d_1} |i\rangle \langle i|$ "maximally entangled state" (normalized)
Ex. • \( \Theta : \mathcal{B}(\mathcal{C}^d) \rightarrow \mathcal{B}(\mathcal{C}^1) \) not CP (exercise 5.4)

• \( \text{id} : \mathcal{B}(\mathcal{C}^d) \rightarrow \mathcal{B}(\mathcal{C}^d) \), \( \tau : \mathcal{B}(\mathcal{C}^d) \rightarrow \mathcal{C} = \mathcal{B}(\mathcal{C}) \) are CP

• for any fixed \( K \in \mathcal{C}^{d_2 \times d_1} \), the map

\[
T : \mathcal{B}(\mathcal{C}^{d_2}) \rightarrow \mathcal{B}(\mathcal{C}^{d_2}), \; X \mapsto T(X) = K \otimes K^* \quad (\ast)
\]

is CP: \( \text{let } Y \in \mathcal{B}(\mathcal{C}^{d_2 \otimes d_2}) \) and \( Y \geq 0 \)

\[
\Rightarrow (T \circ \text{id}_n)(Y) = (K \otimes 1_n)Y(K \otimes 1_n)^* \geq 0.
\]

The maps (\ast) are in fact the 'basic' CP maps:

\[
\text{Fact (Kraus representation): } T : \mathcal{B}(\mathcal{C}^{d_2}) \rightarrow \mathcal{B}(\mathcal{C}^{d_2}) \; \text{is CP}
\]

if and only if it can be written as

\[
T(X) = \sum K_i X K_i^* \quad \text{with } K_i \in \mathcal{C}^{d_2 \times d_1}
\]

The sum needs at most \( d^2 d^2 \) terms.

• \( T : \mathcal{B}(\mathcal{C}^{d_2}) \rightarrow \mathcal{B}(\mathcal{C}^{d_2}) \) is called trace-preserving (TP)

if \( \forall X \in \mathcal{B}(\mathcal{C}^{d_2}) : \; \text{tr}[T(X)] = \text{tr}[X] \)

\( \iff \forall \text{ state } \rho \in \mathcal{B}(\mathcal{C}^{d_2}) : \; \text{tr}[T(\rho)] = \text{tr}[\rho] = 1 \)

A linear map \( T : \mathcal{B}(\mathcal{C}^{d_2}) \rightarrow \mathcal{B}(\mathcal{C}^{d_2}) \) that is CP and TP

is called a quantum channel.

A quantum channel describes the most general (time) evolution (that succeeds always) allowed by the laws of quantum physics.

This is due to the following theorem and since the only deterministic quantum operations are: state preparation, unitary evolution, and neglecting subsystems (but not conditioning on measurement outcome).
Fact (Stinespring representation)

\[ T : \mathcal{B}(\mathbb{C}^{d_1}) \to \mathcal{B}(\mathbb{C}^{d_3}) \text{ is a quantum channel if and only if there exists a unitary } U : \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_3} \to \mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3} \text{ such that } \]

\[ T(X) = \text{Tr}_{d_4} \left[ U (X \otimes 1_{d_2}) U^+ \right] \forall X \in \mathcal{B}(\mathbb{C}^{d_1}). \]

It is always possible to choose \( d_2 = d_1 d_2 \), \( d_4 = d_1^2 \).

The (de-) compression maps for quantum data compression were allowed to be such general quantum operations.

Now we use quantum channels to model noise, e.g., in the transmission of a quantum system from Alice to Bob, or in a noisy quantum memory.

Ex.: (1) partially depolarizing channel \( T : \mathcal{B}(\mathbb{C}^{d_1}) \to \mathcal{B}(\mathbb{C}^{d_1}) \):

\[ T(\rho) = (1 - p) \rho + p \frac{1_{d_1}}{d_1} \quad \text{for } p \in [0, 1] \]

with probability \( p \), the state is replaced by the totally mixed state.

(2) dephasing channel: let \( \{ |x\} \) be a basis in \( \mathbb{C}^{d_1} \):

\[ T(\rho) = \sum_{x=x}^{d_1} |x\rangle \langle x| \rho |x\rangle \langle x| \quad \text{as reduces } \rho \text{ to its diagonal elements ('diagonal basis')} \]

partial dephasing channel: \( (1 - p) \text{id} + p T \) for \( p \in [0, 1] \).