1 Relation between maximum and average error probability

Show that for any \( M \)-code \((E, D)\) for a channel \( \{W_x\}_{x \in \mathcal{X}} \), there exists an \([M/2]\)-code \((E', D')\) for the same channel with \( p_{\text{err, max}}(E', D') \leq 2p_{\text{err}}(E, D) \).

2 Channel capacities

1. For a classical-quantum channel \( \{W_x\}_{x \in \mathcal{X}} \), we defined the function \( f(n) = \sup_{P_{X^n}} I(X^n : B^n) \), where \( \rho_{X^n B^n} = \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) |x^n\rangle \langle x^n| \otimes W_{x^n} \) and \( W_{x^n} = W_{x_1} \otimes W_{x_2} \cdots \otimes W_{x_n} \). Show that the function \( f \) is superadditive.

2. Let \( p \in [0, 1] \) and consider the channel \( W \) with input alphabet \( \mathcal{X} = \{0, 1\} \) defined by \( W_0 = (1 - p)|0\rangle\langle 0| + p|1\rangle\langle 1| \) and \( W_1 = p|0\rangle\langle 0| + (1 - p)|1\rangle\langle 1| \) (this is called the binary symmetric channel). Compute the capacity \( C(W) \).

3. Let \( e \in [0, 1] \) and consider the channel \( W \) with input alphabet \( \mathcal{X} = \{0, 1\} \) defined by \( W_0 = (1 - e)|0\rangle\langle 0| + e|E\rangle\langle E| \) and \( W_1 = (1 - e)|1\rangle\langle 1| + e|E\rangle\langle E| \), where \( |E\rangle \) is orthogonal to both \( |0\rangle \) and \( |1\rangle \) (this is called the binary erasure channel). Compute the capacity \( C(W) \).

4. Prove that the von Neumann entropy is strictly concave, i.e., for any distribution \( \{p(x)\} \) with support \( S \) we have \( H(\sum_{x \in S} p(x)W_x) \geq \sum_{x \in S} p(x)H(W_x) \) with equality if and only if \( W_x = W \) for all \( x \in S \). (You may want to use a question from Sheet 3 for this). Conclude that a classical-quantum channel \( \{W_x\}_{x \in \mathcal{X}} \) satisfies \( C(W) = 0 \) if and only if there exists \( W \) such that \( W_x = W \) for all \( x \in \mathcal{X} \).

3 Properties of the Holevo information

Recall that the Holevo information \( \chi(\{p(x), \sigma_A^x\}) \) of an ensemble \( \{p(x), \sigma_A^x\} \) is defined as \( I(X : A) \), where \( \rho_{XA} = \sum_x p(x)|x\rangle \langle x| \otimes \sigma_A^x \).

1. Show that \( \chi(\{p(x), \sigma_A^x\}) = H(\sum_x p(x)\sigma_A^x) - \sum_x p(x)H(\sigma_A^x) \)

2. Show that \( \chi \) is concave in \( \{p(x)\} \) (for fixed \( \{\sigma_A^x\} \)) and convex in \( \{\sigma_A^x\} \) (for fixed \( \{p(x)\} \)).

3. The Holevo information of a channel \( \mathcal{W} : L(A) \rightarrow L(B) \) is defined as the supremum over all ensembles \( \{p(x), \sigma_A^x\} \) of the Holevo information \( \chi(\{p(x), \mathcal{W}(\sigma_A^x)\}) \). Show that we may restrict the optimization to ensembles where \( \sigma_A^x \) are all pure states.