Quantum Information Theory

Lecture 20

1 HSW Theorem revisited

In the last lecture we saw that the product state classical capacity of a memoryless quantum channel $\Lambda$ is given by the Holevo-Schumacher-Westmoreland (HSW) Theorem [2, 1]:

**Theorem 1 (HSW)**: The product state classical capacity of a memoryless quantum channel $\Lambda$ is given by

$$C^{(1)}(\Lambda) = \chi^*(\Lambda),$$

where $\chi^*(\Lambda)$ is the Holevo capacity (or Holevo information) of the channel and is defined as follows:

$$\chi^*(\Lambda) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, \Lambda(\rho_x)\}),$$

where

$$\chi(\{p_x, \Lambda(\rho_x)\}) := S(\Lambda(\sum_x p_x \rho_x)) - \sum_x p_x S(\Lambda(\rho_x)).$$

The maximum in (1) is taken over all ensembles $\{p_x, \rho_x\}$ of possible input states $\rho_x$ of the channel\(^1\), with $p_x \geq 0$, $\sum_x p_x = 1$.

The fact that the maximization in (2) can be restricted to ensembles of pure states, can be seen as follows:

$$\chi^*(\Lambda) := \max_{\{p_x, \rho_x\}} S(\Lambda(\sum_x p_x \rho_x)) - \sum_x p_x S(\Lambda(\rho_x))$$

Note that the first term on the RHS of (3) does not depend on the details of the ensemble $\{p_x, \rho_x\}$ but only on the average state $\rho := \sum_x p_x \rho_x$. Hence if $\mathcal{H}_A$ is the input Hilbert space of $\Lambda$, we can write

$$\chi^*(\Lambda) := \max_{\rho \in \mathcal{D}(\mathcal{H})} S(\Lambda(\rho)) - \min_{\{p_x, \rho_x\}: \sum_x p_x \rho_x = \rho} \sum_x p_x S(\Lambda(\rho_x))$$

\(^1\)The maximum in (2) is potentially over an unbounded set. However, it is known that one can restrict the maximisation to pure state ensembles, containing at most $d_A^2$ elements, where $d_A$ is the dimension of the input Hilbert space of the channel.
Look at the second term on the RHS of (4). If $\rho_x$ is a mixed state, it can be written in terms of an ensembles of pure states, e.g. consider its eigenvalue decomposition

$$\rho_x = \sum_j q_j^x |\psi_j^x\rangle \langle \psi_j^x|.$$ 

Then by concavity of the von Neumann entropy we have that

$$S(\rho_x) \geq \sum_j q_j^x S(\Lambda(|\psi_j^x\rangle \langle \psi_j^x|)),$$

that is, $\sum_x p_x S(\Lambda(\rho_x))$ can only be decreased if $\rho_x$’s are chosen to be pure. Hence it suffices to consider pure state ensembles.

The following problem is given in Example Sheet 4 and will be used in the next section.

**Exercise 1:** Use the Holevo bound to justify that by transmitting $n$ qubits to Bob, Alice can send at most $n$ bits of classical information to him.

## 1 Proof of the converse of the HSW theorem

The HSW Theorem tells us that Alice can reliably transmit classical information to Bob, through a memoryless quantum channel $\Lambda$ using product state inputs, at any rate below $\chi^*(\Lambda)$. In contrast, any rate $R > \chi^*(\Lambda)$ is not achievable. In this section we prove the latter (also called the converse part of the HSW theorem). That is, we prove that if Alice attempts to send classical messages to Bob at a rate larger than $\chi^*(\Lambda)$, then

$$p_{\text{max}}^{(n)} \not\to 0 \text{ as } n \to \infty,$$

where $n$ denotes the number of successive uses of the memoryless channel $\Lambda$. That is, $\lim_{n \to \infty} p_{\text{max}}^{(n)} > 0$ for such rates.

Suppose Alice wants to send a message $M \in \mathcal{M}$ to Bob. Consider the messages to be uniformly distributed.

**Remark:** We can assume the messages to be uniformly distributed, without loss of generality, because our criterion for reliability of classical information transmission is given in terms of the maximum probability of error $p_{\text{max}}^{(n)}$ (i.e., we consider information transmission to be reliable if $p_{\text{max}}^{(n)} \to 0$ as $n \to \infty$), and $p_{\text{max}}^{(n)}$ does not depend on the distribution of the messages.

She encodes the message $M$ into the quantum state

$$\rho_M^{(n)} = \rho_1^{(n)} \otimes \cdots \otimes \rho_n^{(n)} \in \mathcal{D}(\mathcal{H}^\otimes n),$$

( where $\mathcal{H}$ is the input Hilbert space of the channel $\Lambda$) and sends if to Bob through $n$ uses of the channel (i.e. through $\Lambda^\otimes n$). The state that Bob receives is

$$\Lambda^\otimes n (\sigma_M^{(n)}) = \sigma_1^{(n)} \otimes \cdots \otimes \sigma_n^{(n)} \in \mathcal{D}(\mathcal{H}^\otimes n),$$
where $\sigma^j_M = \Lambda(\rho^j_M)$ for $j = 1, 2, \ldots, n$ and $\mathcal{H}_B$ is the output Hilbert space of the channel $\Lambda$. Let Bob’s measurement be given by a POVM $\{E^{(n)}_M\}$, with each $E^{(n)}_M \in \mathcal{B}(\mathcal{H}_B^\otimes n)$. If $\tilde{M}$ denotes Bob’s inference of Alice’s message, based on the outcome of his measurement, then

$$P(\tilde{M} \neq M) = 1 - \text{tr}(\sigma^{(n)}_M E^{(n)}_M)$$

(5)

If the rate of the encoding-decoding scheme is $R := (\log |\mathcal{M}|)/n$, then $R \leq \log d_B$, where $d_B = \dim \mathcal{H}_B$. This follows from Exercise 1 given above\(^2\). Let $q = P(\tilde{M} \neq M)$ and $h(q)$ denote the corresponding binary entropy. Then by (generalized) Fano’s inequality we have

$$H(M|\tilde{M}) \leq h(q) + q \log (|\mathcal{M}| - 1) \leq h(q) + q \log |\mathcal{M}| \leq h(q) + q n \log d_B$$

(6)

This can be rewritten in terms of the mutual information as follows:

$$qn \log d_B \geq H(M) - I(M : \tilde{M}) - h(q).$$

(7)

By the Holevo bound

$$I(M : \tilde{M}) \leq \chi\left(\left\{\frac{1}{|\mathcal{M}|} \sigma^{(n)}_M\right\}\right).$$

(8)

Hence, defining

$$\bar{\sigma}^j := \frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} \sigma^j_M,$$

for each $j \in \{1, 2, \ldots, n\}$, and using subadditivity of the von Neumann entropy, and its additivity under tensor products, we have

$$I(M : \tilde{M}) \leq S\left(\frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} \sigma^{(n)}_M\right) - \frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} S(\sigma^{(n)}_M)$$

$$\leq \sum_{j=1}^n S(\bar{\sigma}^j) - \frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} S(\sigma^1_M \otimes \sigma^2_M \ldots \otimes \sigma^n_M)$$

$$= \sum_{j=1}^n [S(\bar{\sigma}^j) - \frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} S(\sigma^j_M)]$$

$$\leq \sum_{j=1}^n \chi\left(\left\{\frac{1}{|\mathcal{M}|}, \sigma^j_M\right\}\right)$$

$$\leq n \chi^*(\Lambda)$$

(10)

\(^2nR = \text{the number of classical bits sent, while the maximum possible number of qubits that Bob receives is } \log d_B^\otimes n.\)
Substituting (10) into (7) we obtain
\[ qn \log d_B \geq H(M) - n\chi^*(\Lambda) - h(q) \]
\[ = n(R - \chi^*(\Lambda)) - h(q), \]  \hfill (11)
where the last line follows from the fact that since the messages \( M \) are uniformly distributed over the set \( \mathcal{M} \),
\[ H(M) = \log |\mathcal{M}| = \log 2^{nR} = nR. \]
Hence,
\[ p_{\text{max}}^{(n)} \geq q \geq \frac{(R - \chi^*(\Lambda))}{\log d_B} - \frac{h(q)}{n \log d_B}, \]  \hfill (12)
and hence
\[ \lim_{n \to \infty} p_{\text{max}}^{(n)} \geq \frac{(R - \chi^*(\Lambda))}{\log d_B} > 0, \]  \hfill (13)
for \( R > \chi^*(\Lambda) \). This concludes the proof.

To any ensemble of states \( \{p_x, \rho_x\} \) one can associate a classical-quantum state (or c-q state)
\[ \rho_{XA} = \sum_x p_x |x\rangle \langle x| \otimes \rho_x. \]  \hfill (14)
Here \( \rho_x \in \mathcal{D}(\mathcal{H}_A) \) and \( |x\rangle \in \mathcal{H}_X \). Consider the memoryless quantum channel \( \Lambda : \mathcal{D}(\mathcal{H}_A) \to \mathcal{D}(\mathcal{H}_B) \). Its Holevo capacity can equivalently be expressed as:
\[ \chi^*(\Lambda) = \max_{\rho_{XA}} I(X : B|\tilde{\rho}), \]  \hfill (15)
where \( \rho_{XA} \) is the state defined by (14) and
\[ \tilde{\rho}_{XB} = \sum_x p_x |x\rangle \langle x| \otimes \Lambda(\rho_x). \]  \hfill (16)
This is because
\[ \chi(\{p_x, \rho_x\}) = I(X : A)_{\tilde{\rho}}. \]  \hfill (17)
Since, \( \tilde{\rho}_{XB} = (\text{id}_X \otimes \Lambda)\rho_{XA} \), we have that
\[ \chi(\{p_x, \Lambda(\rho_x)\}) = I(X : B)_{\tilde{\rho}}. \]

Exercise 2: Prove that the Holevo capacity is superadditive,
\[ \chi^*(\Lambda_1 \otimes \Lambda_2) \geq \chi^*(\Lambda_1) + \chi^*(\Lambda_2), \]  \hfill (18)
where \( \Lambda_1, \Lambda_2 \) denote quantum channels.
Exercise 3: Use the HSW theorem to find the product state capacity of the qubit depolarizing channel, $\Lambda$, defined by

$$\Lambda(\rho) = p \rho + (1 - p) \frac{I}{2}.$$ 

An interesting consequence of the HSW theorem is the following lemma.

**Lemma 1** Any arbitrary quantum channel $\Lambda$ can be used to transmit classical information, provided the channel is not simply a constant, i.e., $\Lambda(\rho)$ is not identical for all $\rho$.

**Proof:** This can be seen as follows. If $\Lambda$ is not a constant, then there exist pure states $|\psi\rangle$ and $|\phi\rangle$ such that

$$\Lambda(|\psi\rangle\langle\psi|) \neq \Lambda(|\phi\rangle\langle\phi|).$$

(19)

Consider the ensemble

$$\{p_1 = p_2 = 1/2, \rho_1 = |\psi\rangle\langle\psi|, \rho_2 = |\phi\rangle\langle\phi|\}.$$ 

From (19) and the concavity of the von Neumann entropy it follows that $\chi(\{p_i, \Lambda(\rho_i)\}) > 0$. Hence, $\chi^*(\Lambda) > 0$, which in turn implies that the quantum channel $\Lambda$ can transmit classical information if the latter is encoded into product states which are then sent through the channel.

The HSW Theorem naturally leads to the following question:

**Q** Can one increase the classical capacity of a quantum channel by using entangled input states?

This question is related to an important conjecture, namely, the *additivity conjecture of the Holevo capacity*, which is as follows:

For any two quantum channels, $\Lambda_1$ and $\Lambda_2$:

$$\chi^*(\Lambda_1 \otimes \Lambda_2) = \chi^*(\Lambda_1) + \chi^*(\Lambda_2),$$

(20)

The product channel $\Lambda_1 \otimes \Lambda_2$ for $\Lambda_1 = \Lambda_2 = \Lambda$, denotes two successive uses of $\Lambda$. Let us see how the conjecture (20) is related to the question (Q) above.

It can be shown that the capacity, $C_{\text{classical}}(\Lambda)$, of a quantum channel $\Lambda$ to transmit classical information, in the case in which inputs are not restricted to be product states, is given by the regularised Holevo capacity:

$$C_{\text{classical}}(\Lambda) = \lim_{n \to \infty} \frac{1}{n} \chi^*(\Lambda^\otimes n).$$

(21)
The superadditivity of the Holevo capacity implies that \( C_{\text{classical}}(\Lambda) \geq \chi^*(\Lambda) \), as expected. However if the Holevo capacity is additive then \( \chi^*(\Lambda^\otimes n) = n \chi^*(\Lambda) \), which implies that \( C_{\text{classical}}(\Lambda) = \chi^*(\Lambda) \), i.e., the unrestricted classical capacity is equal to the Holevo capacity of the channel (i.e. its product state capacity), which is a fixed quantity characteristic of the channel. Hence, if the Holevo capacity is additive then the capacity of the quantum channel to transmit classical information cannot be increased by using entangled inputs.

The additivity of the Holevo capacity was proved for various different channels but whether the conjecture was true “globally”, i.e., for all quantum channels, had remained an open question until 2008, when M.Hastings [5] gave a counterexample to it. Hence we now know that the classical capacity of a quantum channel can be increased by using entangled inputs.

References


