Quantum information theory (20110401)

Lecturer: Jens Eisert

Chapter 6: Elements of quantum Shannon theory

Contents

Chapter 6

Elements of quantum Shannon theory

6.1 Definitions of channel capacities

The notion of a channel capacity captures what rate of information can be transmitted via a given communication channel, let this be quantum or classical, given a set of further rules of how this communication is supposed to happen. In the context of quantum communication, naturally, quantum channels are in the focus of attention. We will keep things relatively short, but still define the main quantities and state a couple of key results. There are also a couple of striking insights that we will briefly comment upon. We will also use this as an excuse to properly define asymptotic rates of quantum protocols, including a definition of the distillable entanglement that can be seen as an addendum to the previous chapter. Actually, historically, quantum Shannon theory was the first sub-field of quantum information theory, when it was still thought that quantum effects are a limitation to communication tasks, and not that they can be used to the users' advantage. It is still a field that is actively explored, mostly from the perspective of mathematical physics.

6.1.1 Diamond norms

But first things first. When considering channel capacities for quantum channels, we first need to know in what sense we can approximate a quantum channel. A starting point is the *trace-norm distance* for quantum states: It is defined for two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ as

$$D(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\mathrm{tr}|\rho - \sigma|, \tag{6.1}$$

where $|\cdot|$ denotes the operator absolute value (so the sum of the singular values of the argument). A moment of thought reveals (by invoking the above Kraus theorem) that it quantifies the statistical distinguishability of ρ from σ. This distance hence

operationally captures how different ρ is from σ. More specifically,

$$p := \frac{1}{2}(1 + D(\rho, \sigma)) \tag{6.2}$$

is the maximal success probability when trying to distinguish ρ and σ via a single-shot measurement. Let us now move to meaningful distance measures for quantum channels. One might think a good distance measure for two quantum channels T and S on $\mathcal{H} = \mathbb{C}^d$ (i.e., completely positive, trace-preserving maps) is

$$\|T - S\|_1 = \sup_{\|A\|_1=1} \|T(A) - S(A)\|_1. \tag{6.3}$$

However, there is a problem with this definition: The norms of $\|T \otimes \mathrm{id}_n\|_1$ may increase with n, even though the channel does not even act non-trivially on the second tensor factor. For this reason, one defines the *diamond norm* distance (the factor $1/2$ has no significance) as follows.

Diamond norm: For two channels T and S, the diamond norm is defined as

$$\|T - S\|_\diamond = \sup_n \sup_{\|A\|_1=1} \|(T \otimes \mathrm{id}_n)(A) - (S \otimes \mathrm{id}_n)(A)\|_1. \tag{6.4}$$

Given the clumsy definition as an unbounded supremum, one might be tempted to think that this norm cannot be computed. In fact, it can, even efficiently: It turns out to be the solution, once again, of a semi-definite problem.

6.1.2 Capacities as asymptotic transmission rates

The notion of a capacity asks at what rate information can be transmitted. Here, "rate" refers to an asymptotic rate, invoking a communication channel more than once. This makes a lot of sense, and classical channel capacities are also defined in such a fashion. This is in fact very much reminiscent of what we had in mind in the previous chapter when discussing entanglement manipulation. We define the capacity as the number of invocations of a quantum channel to "get something through", optimized over the respective encodings and decodings.

More specifically and precisely, we can define capacities of quantum channels T based on the quantity

$$\Delta(S, T) = \inf_{D,E} \|S - D \circ T \circ E\|_\diamond, \tag{6.5}$$

as the infimum over encoding channels E and decoding channels D. After all, we are only interested in the optimal encoding and decoding. S here is the identity channel over a certain algebra. S is seen as representing a word of the kind of message that is supposed to be sent, whereas T stands for a single invocation of the channel.

However, when defining a capacity, we are less interested in single invocations, but rather in many invocations and long messages. This refers to the situation of considering $T^{\otimes n}$ and $S^{\otimes n}$ for large n, but encodings and decodings over many channels of this type. A more precise and specific definition of a capacity looks as follows.

Capacity: Let S and T be quantum channels. Then a number $c \geq 0$ is called an "achievable rate" for T with respect to S, if for any sequences n_α, m_α of integers with $m_\alpha \to \infty$ and

$$\limsup_\alpha \left(\frac{n_\alpha}{m_\alpha} \right) < c \tag{6.6}$$

we have

$$\lim_\alpha \Delta(S^{\otimes n_\alpha}, T^{\otimes m_\alpha}) = 0. \tag{6.7}$$

The supremum of all achievable rates is called the *capacity* of T with respect to S and is denoted by $C(S,T)$.

As it is written, the definition makes a lot of sense. It should also be clear, however, that the definition alone equips us very little with tools to actually compute any quantum channel capacity defined in this fashion. Fortunately, a number of results simplifying this expression or providing bounds are known.

6.1.3 Addendum: The distillable entanglement

Now that we have familiarized ourselves with such asymptotic definitions we can come back to the distillable entanglement to have the previous chapter not appear too technically involved. The distillable entanglement is defined as

$$E_D(\rho) = \sup \left\{ \frac{m}{n} : \lim_{n\to\infty} \| T(|\psi\rangle\langle\psi|^{\otimes n}) - |\Omega\rangle\langle\Omega| \|_1 = 0 \right\} \tag{6.8}$$

for an LOCC (local operations with classical communication) protocol T. So this is the yield of the ratio of m output copies to n in an LOCC protocol. But for any finite number, the output does not need to be perfect. It should only be "asymptotically perfect" for a large number of input copies.

6.2 Properties of channel capacities

6.2.1 Classical information capacity and additivity problems

The classical information capacity C_c is defined as the rate at which classical bits can be sent via a quantum channel. The quantum capacity C_q in turn is the rate at which quantum bits, qubits, can be transmitted. If the one-bit system is defined as $\mathcal{C}_2$ and the one qubit system as $\mathcal{M}_2$, then they are

$$C_c(T) = C(\mathcal{C}_2, T), \tag{6.9}$$
$$C_q(T) = C(\mathcal{M}_2, T). \tag{6.10}$$

It is clear that

$$C_q(T) \leq C_c(T) \tag{6.11}$$

for any quantum channel T, as quantum channels can be used to send classical bits. But it may be true that some noisy channels only allow for the transmission of classical information, but no coherent quantum information. These capacities are notoriously difficult to compute. However, stringent bounds can be found, and formulae do exist. The most famous result is the expression of the classical information capacity In order to state this, we need to define the *von-Neumann entropy* of a quantum state. It is the quantum analogon of the Shannon entropy and given by

$$S(\rho) = -\mathrm{tr}(\rho \log \rho), \tag{6.12}$$

in terms of a matrix logarithm (that is computed on the spectrum of ρ as a matrix function). The *classical information capacity* is then found to be the following expression.

Classical information capacity: The single-shot classical capacity is given by

$$C_{c,1}(T) := \max_{\{p_i,\rho_i\}} \left(S(\sum_i p_i T(\rho_i)) - \sum_i p_i S(T(\rho_i)) \right), \tag{6.13}$$

where $\{p_i\}$ is a probability distribution and $\{\rho_i\}$ is a set of quantum states. The actual *classical information capacity* is then regularized as

$$C_c(T) := \sup_n \frac{1}{n} C_{c,1}(T^{\otimes n}). \tag{6.14}$$

This expression seems puzzling: How can it be an advantage to send information coherently over many channels, so why is not simply $C_c(T) = C_{c,1}(T)$? It turns out that it does help. It was an open problem for a long time whether or not the classical information capacity was *additive*. Using ingenious ideas of random coding, it was shown in large dimension to be beneficial to use entangled inputs, even though only classical information is to be transmitted. In fact, the additivity was a long-standing puzzle and open question in the field: It could be shown that many additivity questions in quantum information theory were equivalent until it was finally settled. The most puzzling of these (equivalent) formulations was the one of the *additivity of the minimum output entropy* of a quantum channel. This is the smallest von-Neumann entropy one can attain by a suitable input to the quantum channel,

$$S_{\min}(T) := \inf_{\rho \in \mathcal{S}(\mathcal{H})} S(T(\rho)). \tag{6.15}$$

Here, it seems particularly counter-intuitive that

$$S_{\min}(T \otimes T) \neq 2 S_{\min}(T) \tag{6.16}$$

in general. In fact, the latter publication made use of the proven equivalences and proved a counterexample for the additivity of the minimum output entropy of a quantum channel, maybe the most puzzling of the known additivity problems. It was done using random quantum coding, and it is still an open problem in the field to provide a constructive counterexample.

6.2.2 Quantum capacity and super-activation

How about the quantum capacity? For this, we define the *coherent information* as

$$C_{q,1}(T) := \sup_{\rho} \left(S(\rho_B) - S(\rho_E) \right), \tag{6.17}$$

where in a Stinespring dilation we write the channel as $T(\rho) = U(\rho \otimes \omega)U^\dagger$ and consider the output state as a bi-partite state over B and E. The genuine quantum capacity is again seen as an asymptotic limit.

Quantum information capacity: The quantum capacity is given by

$$C_q(T) := \sup_n \frac{1}{n} C_{q,1}(T^{\otimes n}). \tag{6.18}$$

Again, it is known that the "regularization" on the right hand side is needed. This renders the quantum capacity a quantity that in practice cannot be computed, but quite easily meaningfully bounded. A most intriguing phenomenon is that of super-activation.

Super-activation: There exists quantum channels T_1 and T_2 both of which featuring a vanishing quantum information capacity, i.e.,

$$C_q(T_1) = C_q(T_2) = 0. \tag{6.19}$$

Still, both channels used jointly have a strictly positive quantum information capacity so that

$$C_q(T_1 \otimes T_2) > 0. \tag{6.20}$$

That is to say, one can combine two quantum channels that have precisely zero quantum capacity. But used jointly, one can transmit faithfully quantum information. This is a highly counter-intuitive effect in quantum communication that received significant attention at the time.