Quantum information theory (20110401)

Lecturer: Jens Eisert

Chapter 4: Quantum channels

Contents

Chapter 4

Quantum channels

4.1 General quantum state transformations

What is the most general transformation of a quantum state that one can perform? One might wonder what this question is supposed to mean: We already know of unitary Schrödinger evolution generated by some Hamiltonian H. We also know of the measurement postulate that alters a state upon measurement. So what is the question supposed to mean? In fact, this change in mindset we have already encountered above, when we thought of unitary operations. Of course, one can interpret this a-posteriori as being generated by some Hamiltonian, but this is not so much the point. The emphasis here is on what can be done, what unitary state transformations are possible. It is the purpose of this chapter to bring this mindset to a completion, and to ask what kind of state transformations are generally possible in quantum mechanics. There is an abstract, mathematically minded approach to the question, introducing notions of complete positivity. Contrasting this, one can think of putting ingredients of unitary evolution and measurement together. Fortunately, these pictures turn out to be equivalent. Either way, this is given by the notion of a *quantum channel*. Given that we think here of the most general transformation, the connotation of an actual communication channel is perfectly accurate: We will see that natural communication channels (such as provided by fibres and so on) can be captured nicely in terms of quantum channels.

4.2 Complete positivity

Mathematically speaking, quantum channels capture the legitimate transformations that quantum states can undergo. They directly generalize the concept of a *stochastic matrix* that maps probability vectors onto probability vectors. A stochastic matrix is a matrix $P \in \mathbb{R}_+^{d\times d}$ with the property that

$$\sum_{k=1}^{d} P_{j,k} = 1, \tag{4.1}$$

so that probability vectors $p \in \mathbb{R}_+^d$ with $\sum_{j=1}^d p_j = 1$ are mapped to probability vectors.

Turning to the quantum world again: A quantum channel captures two things. First, it is the most general valid operation that one can perform on a quantum system. Second, it describes real communication channels as a special case. From the perspective of a mathematical characterization, what defines a quantum channel T? Surely, such channels must be linear maps, so that if T_1 and T_2 are quantum channels, then

$$T = \alpha T_1 + \beta T_2 \tag{4.2}$$

with $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$ is a quantum channel. A quantum channel T must also be a *positive map*, $T \geq 0$, mapping positive operators $\rho \geq 0$ onto positive operators, such that

$$\sigma = T(\rho) \geq 0 \tag{4.3}$$

must again be a valid positive operator. Interestingly, this turns out not to be enough: One needs a stronger form of positivity, referred to as *complete positivity*.

Complete positivity and quantum channels: Linear maps T on $\mathcal{H}$ are called completely positive iff

$$T \otimes \mathrm{id} \geq 0, \tag{4.4}$$

where $T \otimes \mathrm{id}$ is a linear map on $\mathcal{H} \otimes \mathcal{H}$ with $\mathcal{H} = \mathbb{C}^d$. Quantum channels are trace-preserving completely positive maps satisfying

$$\mathrm{tr}(T(\rho)) = 1 \tag{4.5}$$

for all ρ satisfying $\mathrm{tr}(\rho) = 1$.

It turns out that it is sufficient to take d having the same dimension as the dimension of the first tensor factor. Why is that? Because T could act on a part of a larger system, and then the operator $(T \otimes \mathrm{id})(\rho)$ must again be a valid quantum state. This is a feature of quantum mechanics absent in classical mechanics: Although the map acts only on a part of the system and "does nothing" to the second tensor factor, the joint map still needs to be a positive map. The best known example of a positive but not completely positive map is the *transposition* t, mapping

$$t : \rho \mapsto \rho^T = \rho^*. \tag{4.6}$$

Note that ^{T} denotes the element-wise transposition and * the element-wise complex conjugation. Hermitian conjugation will be denoted by †. Physically, this map reflects a *time reversal*. It is easy to see that this is a positive map, so whenever $\rho \geq 0$ then also $\rho^T \geq 0$. But *partial transposition* is not completely positive. Think of the quantum state of two qubits in $\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, given by

$$\rho = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \tag{4.7}$$

with eigenvalues $\{1,0,0,0\}$. Its partial transposition is then

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.8}$$

clearly not a positive matrix since it has eigenvalues $\left\{-\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\}$. In fact, for single qubits, this is basically already all there is for positive matrices. We do not offer a proof of this statement.

Structure theorem for positive maps on qubits: An arbitrary positive linear map T acting on $\mathbb{C}^2$ can be written as

$$T = \alpha T_1 \circ t + \beta T_2 \tag{4.9}$$

with $\alpha, \beta \geq 0$, t is the transposition and T_1, T_2 are completely positive linear maps.

4.2.1 Choi-Jamiolkowski isomorphism

More important is the following: The above definition of complete positivity does not give rise to a criterion that can be efficiently checked. Fortunately, the following statement provides such a criterion: It is necessary and sufficient for complete positivity to apply the linear map to a certain single reference state.

Criterion for complete positivity: A linear map T on $\mathcal{H}$ is completely positive iff

$$(T \otimes \mathrm{id})(\Omega) \geq 0 \tag{4.10}$$

where $\Omega \in \mathcal{H} \otimes \mathcal{H}$ is a maximally entangled state.

Proof: We will briefly prove this statement. We will need the following tiny Lemma for this: For any $\mathbb{C}^{d\times d} \ni P \geq 0$ and any $A \in \mathbb{C}^{d\times d}$, we have that

$$APA^\dagger \geq 0. \tag{4.11}$$

This is an immediate consequence of the fact that for every $|\psi\rangle \in \mathbb{C}^d$,

$$\langle\psi|APA^\dagger|\psi\rangle = (\langle\psi|A)P(A^\dagger|\psi\rangle) \geq 0. \tag{4.12}$$

Let us assume that Eq. (4.10) holds true. We will now show that

$$(T \otimes \mathrm{id})(\rho) \geq 0 \tag{4.13}$$

for all $\rho \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H})$. We make use of the spectral decomposition

$$\rho = \sum_{j=1}^{d^2} p_j |\psi_j\rangle\langle\psi_j|. \tag{4.14}$$

From linearity, we have

$$(T \otimes \mathrm{id})(\rho) = \sum_{j=1}^{d^2} p_j (T \otimes \mathrm{id})(|\psi_j\rangle\langle\psi_j|), \tag{4.15}$$

so it is sufficient to show the statement for single state vectors $|\psi\rangle \in \mathcal{H}$. Now every such state vector can be written as

$$|\psi\rangle = (\mathrm{id} \otimes X)|\Omega\rangle \tag{4.16}$$

for a suitable $X \in \mathbb{C}^{d\times d}$. But then we have

$$\begin{aligned}(T \otimes \mathrm{id})(|\psi\rangle\langle\psi|) = (T \otimes \mathrm{id}) \left((\mathrm{id} \otimes X)|\Omega\rangle\langle\Omega|(\mathrm{id} \otimes X^\dagger)\right) \\ (\mathrm{id} \otimes X)(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)(\mathrm{id} \otimes X^\dagger) \geq 0\end{aligned} \tag{4.17}$$

from which the statement follows. In fact, it turns out that $(T\otimes\mathrm{id})(|\Omega\rangle\langle\Omega|)$ completely specifies the channel.

Choi-Jamiolkowski isomorphism: Quantum channels as completely positive, trace preserving maps T on $\mathcal{H}$ are isomorphic to the quantum states

$$(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|). \tag{4.18}$$

The proof is left as an exercise. In fact, one direction of the proof of the isomorphism we have already elaborated upon. This may not be a particularly deep statement, but it has profound implications. Channels can be viewed as quantum states on a larger Hilbert space. That also comes along with the insight that the set of quantum channels is again a convex set. In fact, any kind of optimization of linear functionals over quantum channels can be cast into the form of a *convex optimization problem*. We will see that in fact *semi-definite programming* is at the heart of the optimization of many quantum protocols.[1] In fact, many optimal success probabilities of protocols can

[1] Semi-definite programming generalizes linear programming and is a form of a convex optimization problem for which the theory is very much developed, and for which interior point methods provide an efficient solution. They are optimization problems of the form, for vectors $c \in \mathbb{R}^d$ and matrices $F_0, \ldots, F_d \in \mathbb{R}^{D\times D}$

$$\text{minimize } c^T x, \tag{4.19}$$

$$\text{subject to } F_0 + \sum_{j=1}^{d} x_j F_j \geq 0. \tag{4.20}$$

The Lagrange dual is again a semi-definite problem of the form

$$\text{maximize } -\mathrm{tr}(ZF_0), \tag{4.21}$$

$$\text{subject to } \mathrm{tr}(ZF_j) = c_j \forall j = 1, \ldots, d, \tag{4.22}$$

$$Z \geq 0. \tag{4.23}$$

Any solution to the Lagrange dual provides a lower bound to any solution to the original, the primal, problem, which is a property most useful when using semi-definite programming in proofs.

readily be captured as semi-definite programs of this form. In a bigger picture, ideas of convex and non-convex programming feature strongly in quantum mechanics. In fact, the latter works provide hierarchies of semi-definite programs to decide the above separability problem, where each level of the hierarchy can be solved in polynomial time.

4.2.2 Kraus' theorem and Stinespring dilations

We have understood what a completely positive map is, but not how it can be parametrized and what specific form it takes. This is given by Kraus' theorem.

Kraus' theorem: A linear map T on $\mathcal{H}$ is completely positive and trace-preserving exactly if it can be written as

$$T(\rho) = \sum_{j=1}^{r} K_j \rho K_j^\dagger \tag{4.24}$$

satisfying

$$\sum_{j=1}^{r} K_j^\dagger K_j = \mathbb{1}. \tag{4.25}$$

The smallest number r that can be achieved in such a decomposition is called the Kraus rank.

We do not have the time to present the full proof of this. But we sketch the idea. One direction of the proof is trivial: T is linear by construction. Also, applying (4.32) to (4.10) immediately gives rise to a positive operator. The more technical direction is to show that such a form can always be achieved. The key steps are to start from the spectral decomposition

$$(T \otimes \mathrm{id})(\Omega) = \sum_i p_i |e_i\rangle\langle e_i|. \tag{4.26}$$

Now take an arbitrary state vector $|\psi\rangle \in \mathcal{H}$, and to extend it onto $\mathcal{H} \otimes \mathcal{H}$ as $|\psi^*\rangle \otimes |\psi\rangle$. One can then write

$$|\psi\rangle\langle\psi| = d\langle\psi^*|\Omega\rangle\langle\Omega|\psi^*\rangle = d\mathrm{tr}_A(|\psi^*\rangle\langle\psi^*| \otimes \mathbb{1})|\Omega\rangle\langle\Omega|). \tag{4.27}$$

Then applying T can be done on the second tensor factor. The Kraus operators are then defined by

$$K_j|\psi\rangle = \sqrt{dp_j}\langle\psi^*|e_j\rangle. \tag{4.28}$$

Note that the Kraus decomposition is not unique: Any set $\{l_k\}$ is again a set of Kraus operators if

$$l_k = \sum_i U_{k,i} K_i \tag{4.29}$$

for U being unitary is again a legitimate set of Kraus operators. It is also not difficult to see that the Kraus rank is exactly the standard rank of the Choi-Jamiolkowski isomorph $(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)$, an insight that is again left to the reader as an exercise.

Any channel can be seen as a unitary map in a larger vector space, a statement captured by *Stinespring's theorem*. We will spell it out in a slightly unusual and redundant form, yet one that is easier to communicate. This is a most important form: Its significance stems from the observation that unitary operations originate from time evolution in quantum mechanics, the most important quantum channel.

Hamiltonian evolution: The channel

$$\rho \mapsto U\rho U^\dagger \tag{4.30}$$

with U being a unitary on $\mathcal{H}$ captures *Hamiltonian time evolution* generated by a *Hamiltonian* $H = H^\dagger$ via $U = \exp(-itH)$. Such dynamics is referred to as *Schrödinger dynamics*.

In fact, most elementary courses on quantum mechanics elaborate on the consequences of such time evolution generated by meaningful Hamiltonians capturing important physical systems: The Schödinger equation is one of the key equations and one of the axioms of quantum mechanics. The point of the Stinespring dilation is now to see that any channel can be seen as such a unitary channel on a larger vector space.

Stinepring dilations: Any completely positive and trace-preserving map T on $\mathcal{H} = \mathbb{C}^d$ can be written as

$$T(\rho) = \mathrm{tr}_2(U(\rho \otimes \eta)U^\dagger) \tag{4.31}$$

where η is a quantum state on $\mathbb{C}^D$, U is a unitary defined on $\mathbb{C}^d \otimes \mathbb{C}^D$, and tr_2 is the partial trace with respect to the second tensor factor. D has at most be taken to be d.

4.2.3 Disturbance versus information gain

We briefly mention here that there is no way to attain any information about a system without changing its quantum state. This is elucidated at by means of the following statement. In fact, the labels of the Kraus' theorem exactly correspond to the labels in a von-Neumann measurement when the measurement postulate is applied to an auxiliary quantum system initially in η.

Generalized measurement: The Kraus decomposition can be realized as

$$K_j \rho K_j^\dagger = \mathrm{tr}_2((\mathbb{1} \otimes \pi_j) U (\rho \otimes \eta) U^\dagger) \tag{4.32}$$

where η is a quantum state on $\mathbb{C}^D$, U is a unitary defined on $\mathbb{C}^d \otimes \mathbb{C}^D$, tr_2 is the partial trace with respect to the second tensor factor, and $\pi_j = |\psi_j\rangle\langle\psi_j|$ are unit rank projections from the measurement postulate.

We will now go too much into detail here: But when captured in this form, it should be clear that the information gain (the knowledge obtained via the statistics of measurement outcomes) and the disturbance (the alteration of ρ to the state conditioned on measurement outcomes) are in a close relationship to one another.

4.3 Local operations and classical communication

We have so far learned what operations can be done, and what role Kraus operators play. In this section, we turn to the important problem of how this manifests itself in the composite setting where quantum operations can only be implemented locally. This is key to the understanding of distributed settings.

4.3.1 One-local operations

Let us consider a bi-partite system with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. These could be parts of a distributed quantum systems. Naturally, one can only perform local operations in each of the parts. The action may, however, be classically coordinated.

Local operations: Local operations in a bi-partite system with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ are operations that act as either

$$\rho \mapsto \sum_{j=1}^{J} (A_j \otimes \mathbb{1}) \rho (A_j \otimes \mathbb{1})^\dagger, \tag{4.33}$$

$$\rho \mapsto \sum_{j=1}^{J} (\mathbb{1} \otimes B_j) \rho (\mathbb{1} \otimes B_j)^\dagger, \tag{4.34}$$

where trace preservation requires that

$$\sum_{j=1}^{J} A_j^\dagger A_j = \mathbb{1}, \ \sum_{j=1}^{J} B_j^\dagger B_j = \mathbb{1}. \tag{4.35}$$

Let us imagine that the measurement is performed in Alice's laboratory. The label j can be viewed as a measurement outcome. This measurement outcome can be

communicated to Bob who would then make use of this label to implement a unitary operation that depends on this label.

One-local operations: One-local operations in a bi-partite system with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ with one-way communication from Alice to Bob are of the form

$$\rho \mapsto \sum_{j=1}^{J} (A_j \otimes U^{(j)}) \rho (A_j \otimes U^{(j)})^\dagger, \tag{4.36}$$

where all $\{U^{(j)} : j = 1, \ldots, J\}$ are unitary and where trace preservation requires

$$\sum_{j=1}^{J} A_j^\dagger A_j = \mathbb{1}. \tag{4.37}$$

4.3.2 General local quantum operations

Having said that, Bob may again implement a local operation himself, and communicate the results back to Alice.

Local operation with classical communication: General local quantum operations with classical communication (LOCC) in a bi-partite system with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ are sequences of protocols in which one party, say, Alice, performs a local quantum operation who transmits the label reflecting the outcome to Bob, who then again perform a local quantum operation in general dependent on the outcome of the form

$$\rho \mapsto \sum_{j=1}^{J} \sum_{k=1}^{K_j} (A_j \otimes B_k^{(j)}) \rho (A_j \otimes B_k^{(j)})^\dagger, \tag{4.38}$$

where trace preservation requires

$$\sum_{j=1}^{J} A_j^\dagger A_j = \mathbb{1}, \ \sum_{k=1}^{K_s} B_k^\dagger B_k = \mathbb{1} \tag{4.39}$$

for all $s = 1, \ldots, J$.

There is a large class of operations, one that we will not give a box, however: *Separable operations are of the form*

$$\rho \mapsto \sum_{j=1}^{J} (A_j \otimes B_j) \rho (A_j \otimes B_j)^\dagger, \tag{4.40}$$

again satisfying trace preservation. A number of interesting comments are in order:

- One can see that when manipulating pure quantum states, it is always sufficient to make use of *one-local* transformations. This is basically a consequence of the Schmidt decomposition.
- In general, one can reach more states by considering more than one round of local quantum operations and classical communication.
- In fact, one can prove that the reachable set of each round r is strictly smaller than the reachable set in round $r + 1$, for all r. This is an intriguing result. So one has to consider LOCC protocols with an arbitrary number of steps in order not to restrict generality.
- The above box also considers *separable operations*. It is rather obvious that every LOCC is also a separable operation: One has to suitably see the labels as super-labels that collect all the communication that is done. The converse is not true, however. In fact, since the label j is shared, they can in general not even physically be implemented. They still constitute a convenient outer approximation of the set of LOCC, which for the above mentioned reason is hard to handle. Hence, often separable operations are used as proxies for LOCC.