Lecture notes in QI

Part I: Foundations

L1 Kinematics: Qubits, Bloch sphere, quantum ensembles, mixing theorem

L2 Kinematics: Composite systems, Schmidt's theorem, reduced density operator, partial trace

L3 Entropy: Shannon and von Neumann entropy, relative entropy

L4 Dynamics: Completely positive maps, Stinespring dilation

L5 Dynamics: Measurements, master equations (Lindblad)

Part II: Entanglement and locality

L6 Entanglement: Separable and inseparable states, PPT
Entanglement measures: Concurrence, negativity

L7 Entanglement as a resource: Quantum teleportation, Entanglement of formation
Multipartite entanglement: Residual entanglement, GHZ and W states

L8 Quantum non-locality: Bell inequalities, Quantum key distribution - Ekert protocol (E91)

L9 Signalling: No-cloning, quantum copier

Part III: Distance measures

L10 Trace distance, measure of non-Markovianity

L11 Fidelity, Uhlmann's theorem

Reading:

- M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press)
- S. Stenholm and K.A. Suominen, *Quantum Approach to Informatics* (Wiley)
- J. Preskill, *Lecture notes in quantum computation* (Caltech) http://www.theory.caltech.edu/people/preskill/ph229/

Lecture 1

Unit of quantum information: qubit

- The classical **bit** encodes information in one of two distinguishable values, e.g., heads or tails of a coin. The explicit physical realization of the bit does not matter. These values are therefore denoted by the generic symbols 0 and 1.

  A string of n bits can encode 2^n symbols.

- The quantum-mechanical bit (**qubit**) encodes two distinguishable values in quantum-mechanical superpositions.

  Examples: Polarization states $|H\rangle$ and $|V\rangle$ of a photon; spin states $|\uparrow\rangle$ and $|\downarrow\rangle$ of a spin-$\frac{1}{2}$ particle; ground and excited states $|g\rangle$ and $|e\rangle$ of an atom/ion/molecule.

  Just as for the classical bit, the exact nature of the qubit does not matter so we introduce the generic notation $|0\rangle$ and $|1\rangle$ that defines the **computational basis** of the qubit state space.

  Several qubits can be combined to form states in a 2^n-dimensional Hilbert space, in which all 2^n computational basis states can be superposed.

- The qubit is an analog encoding (state preparation) but a digitial decoding (read out).

  The qubit state is geometrically represented by the **Bloch sphere**: a unit sphere with polar angles θ, φ defining the arbitrary qubit superposition $\cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$. Each single qubit state is in one-to-one correspondence with a point on Bloch sphere. Distinguishable read-outs are given by antipodal points on the Bloch sphere.

  Example: Spin-$\frac{1}{2}$ states can be associated with directions in ordinary 3D-space that provides an explicit interpretation of the Bloch sphere. Read out can be carried out in, e.g., a Stern-Gerlach setting: Different pairs of antipodal points correspond to different orientations of the SG-magnet.

Quantum ensembles: The density operator

- Conceptually, a mixed quantum state can be thought of as a source that produces sometimes the state ψ_1, sometimes the state ψ_2, and so on. The probabilities (relative frequencies) for these to happen are $p_1, p_2, \ldots$

- For such a source, what would be the average value of an observable $\mathcal{O}$? First, recall that if the system is described by the state $|\psi_k\rangle$, then the average value of $\mathcal{O}$ is given by the expectation value: $\langle\mathcal{O}\rangle_{\psi_k} = \langle\psi_k|\mathcal{O}|\psi_k\rangle$. It thus follows:

$$\langle\mathcal{O}\rangle_{\{\psi_k,p_k\}} = p_1\langle\psi_1|\mathcal{O}|\psi_1\rangle + p_2\langle\psi_2|\mathcal{O}|\psi_2\rangle + \ldots = \sum_{k=1}^{K} p_k\langle\psi_k|\mathcal{O}|\psi_k\rangle,$$

  where K is the number of different output states. Note that $|\psi_k\rangle$ need not be mutually orthogonal, which means that $K \neq \dim\mathcal{H}$ in general.

  To proceed, we need the following identity:

$$\langle\psi|\mathcal{O}|\psi\rangle = \sum_n \langle\psi|\mathcal{O}|n\rangle\langle n|\psi\rangle = \sum_n \langle n|\psi\rangle\langle\psi|\mathcal{O}|n\rangle = \mathrm{Tr}\left(|\psi\rangle\langle\psi|\mathcal{O}\right).$$

  By using the identity and linearity of trace, we find:

$$\langle\mathcal{O}\rangle_{\{\psi_k,p_k\}} \;=\; \sum_{k=1}^{K} p_k \mathrm{Tr}\Big(|\psi_k\rangle\langle\psi_k|\mathcal{O}\Big) = \mathrm{Tr}\left(\sum_{k=1}^{K} p_k|\psi_k\rangle\langle\psi_k|\mathcal{O}\right).$$

  The sum-over-k factor is independent of the observable, thus we may write

$$\langle\mathcal{O}\rangle_\rho \;=\; \mathrm{Tr}\left(\rho\mathcal{O}\right)$$

  with

$$\rho = \sum_{k=1}^{K} p_k|\psi_k\rangle\langle\psi_k|.$$

  ρ is a mathematical object, the **density operator**, that describes *all* the statistical information about the system in the case the source has the mixed characterization $\{\psi_k, p_k\}$.

- Properties of the density operator:

  – $\rho \geq 0$.

– $\mathrm{Tr}(\rho) = 1$.

– $\rho^2 \leq \rho$.

Special cases:

(i) ρ is a **pure** state if $p_k = \delta_{kk'}$ for some given k';

(ii) ρ is a **maximally mixed** state if $p_k = \frac{1}{d}$ for a qudit.

Note: ρ is a pure state iff $\rho^2 = \rho$.

- The density operator of a single qubit takes the form

$$\rho = p_0|\psi_0\rangle\langle\psi_0| + p_1|\psi_1\rangle\langle\psi_1|,$$

where $|\psi_0\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle$ and $|\psi_1\rangle = -e^{-i\varphi}\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle$. In terms of Pauli operators $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$:

$$\rho = \frac{1}{2}\left(\hat{1} + \mathbf{r}\cdot\boldsymbol{\sigma}\right),$$

where $\mathbf{r} = (p_0 - p_1)(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ is the **Bloch vector**.

Note that all states with $|\mathbf{r}| = 1$ are pure. These lie on the Bloch sphere. The interior $|\mathbf{r}| < 1$ contains all non-pure states. Pure and non-pure states form the **Bloch ball**. In particular, the origin $|\mathbf{r}| = 0$ of the Bloch ball is the maximally mixed qubit state $\rho = \frac{1}{2}\hat{1}$.

Mixing theorem

- Decomposition freedom: Given the mixture $\{\psi_k, p_k\}$ one can construct uniquely the density operator ρ, but the converse is not true unless the state is a pure state (contains a single term).

  Example: For a single qubit, consider an equal mixing of the non-orthogonal states $|\psi_\pm\rangle = \pm\frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle$. Geometrically this mixture can be viewed as the intersection point between a straight lie connecting these two points and the z axis in the Bloch ball. Thus, we may equally well represent it as an unequal mixture of the orthogonal states $|0\rangle$ and $|1\rangle$. [cf. Problem 10]

- Theorem: Consider the density operators $\rho = \sum_{k=1}^{K} p_k |e_k\rangle\langle e_k|$, $\langle e_k|e_l\rangle = \delta_{kl}$, and $\sigma = \sum_{l=1}^{L} q_l |\phi_l\rangle\langle\phi_l|$. Then, $\sigma = \rho$ iff there exists a unitary $L \times L$ matrix v such that

$$\sqrt{q_l}|\phi_l\rangle = \sum_{k=1}^{K} \sqrt{p_k}|e_k\rangle v_{kl}.$$

  Proof: [L. P. Hughston *et al.*, Phys. Lett. A **183**, 14 (1993)] If there exists a unitary v_{kl} such that $\sqrt{q_l}|\phi_l\rangle = \sum_k \sqrt{p_k}|e_k\rangle v_{kl}$, then

$$\begin{aligned}
\sigma &= \sum_l \sqrt{q_l}|\phi_l\rangle\langle\phi_l|\sqrt{q_l} = \sum_l \sum_k \sqrt{p_k}|e_k\rangle v_{kl} \sum_{k'} \sqrt{p_{k'}}\langle e_{k'}|v^*_{k'l} \\
&= \sum_{kk'} \sqrt{p_k p_{k'}}|e_k\rangle\langle e_{k'}| \sum_l v_{kl} v^\dagger_{lk'} = \sum_{kk'} \sqrt{p_k p_{k'}}|e_k\rangle\langle e_{k'}|\delta_{kk'} \\
&= \sum_k p_k |e_k\rangle\langle e_k| = \rho.
\end{aligned}$$

  Conversely, if $\sigma = \rho$ then the matrix $v_{kl} = \frac{\sqrt{q_l}}{\sqrt{p_k}}\langle e_k|\phi_l\rangle$ is unitary:

$$\sum_{l=1}^{L} v_{kl} v^\dagger_{lk'} = \sum_{l=1}^{L} \frac{q_l}{\sqrt{p_k p_{k'}}}\langle e_k|\phi_l\rangle\langle\phi_l|e_{k'}\rangle = \frac{1}{\sqrt{p_k p_{k'}}}\langle e_k|\rho|e_{k'}\rangle = \delta_{kk'}$$

  such that

$$\sum_{k=1}^{K} \sqrt{p_k}|e_k\rangle v_{kl} = \sum_{k=1}^{K} \sqrt{q_l}|e_k\rangle\langle e_k|\phi_l\rangle = \sqrt{q_l}|\phi_l\rangle$$

  □

  Note: If $L > K$ then only the $K \times K$ submatrix of v is used. The remaining $L - K$ rows are arbitrary up to the condition that v should be unitary.

Lecture 2

Composite systems

- A quantum system with two or more degrees of freedom that can be separately measured = **composite system**.

  Examples: Two distant atoms in an optical lattice, polarization + momentum of a photon, ...

- States of composite systems are given by **tensor products** $|k\rangle \otimes |l\rangle \otimes \ldots$, which span a vector space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \ldots$, whose dimension is:

$$\dim(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \ldots) = \dim(\mathcal{H}_A) \cdot \dim(\mathcal{H}_B) \cdot \ldots$$

  Example: For n qubits, $\dim(\mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n) = 2^n$.

  These vectors can be superposed:

$$|\Psi\rangle = \sum_{kl\ldots} a_{kl\ldots} |k\rangle \otimes |l\rangle \otimes \ldots$$

  If N degrees of freedom, the system is said to be ***N*-partite**. Simplifying notations:

$$|k\rangle \otimes |l\rangle \otimes \ldots = |k\rangle|l\rangle \ldots = |kl\ldots\rangle.$$

- **Bipartite** case: $N = 2$, for which we may write:

$$|\Psi^{AB}\rangle = \sum_{k=1}^{n_A} \sum_{l=1}^{n_B} a_{kl} |k\rangle \otimes |l\rangle,$$

  where $n_A = \dim \mathcal{H}_A$ and $n_B = \dim \mathcal{H}_B$.

  Schmidt's theorem: Pure bipartite states can always be written as:

$$|\Psi^{AB}\rangle = \sum_{m=1}^{\min\{n_A, n_B\}} \sqrt{d_m} |A_m\rangle \otimes |B_m\rangle,$$

  where $\langle A_m | A_n \rangle = \delta_{mn}$ and $\langle B_m | B_n \rangle = \delta_{mn}$.

  Mathematical note: Let $\boldsymbol{a}$ be an $n_A \times n_B$ matrix. Then there exists unitary $n_A \times n_A$ and $n_B \times n_B$ matrices $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively, and

a diagonal $n_A \times n_B$ matrix $\boldsymbol{d} \geq 0$, defining the **singular value decomposition** (SVD) $\boldsymbol{a} = \boldsymbol{udv}$. The non-negative entries of $\boldsymbol{d}$ are the **singular values** of $\boldsymbol{a}$.

Proof: View a_{kl} as a $n_A \times n_B$ matrix $\boldsymbol{a}$ and set $\boldsymbol{d}_{mm} = \sqrt{d_m}$, where $m = 1, \ldots, \min\{n_A, n_B\}]$. SVD of $\boldsymbol{a}$ reads

$$a_{kl} = \sum_{m=1}^{\min\{n_A,n_B\}} u_{km}\sqrt{d_m}v_{ml},$$

which implies:

$$|\Psi^{AB}\rangle = \sum_{k=1}^{n_A}\sum_{l=1}^{n_B}\sum_{m=1}^{\min\{n_A,n_B\}} u_{km}\sqrt{d_m}v_{ml}|k\rangle \otimes |l\rangle.$$

Let

$$\begin{aligned} |A_m\rangle &= \sum_{k=1}^{n_A} |k\rangle u_{km}, \\ |B_m\rangle &= \sum_{l=1}^{n_B} v_{ml}|l\rangle, \end{aligned}$$

which yields

$$|\Psi^{AB}\rangle = \sum_{m=1}^{\min\{n_A,n_B\}} \sqrt{d_m}|A_m\rangle \otimes |B_m\rangle.$$

It remains to prove orthogonality:

$$\begin{aligned} \langle A_m|A_n\rangle &= \sum_{k,k'=1}^{n_A} u^*_{km}u_{k'n}\langle k|k'\rangle = \sum_{k=1}^{n_A} u^*_{km}u_{kn} \\ &= \sum_{k=1}^{n_A} u^\dagger_{mk}u_{kn} = \delta_{mn}, \\ \langle B_m|B_n\rangle &= \sum_{l,l'=1}^{n_B} v^*_{ml}v_{nl'}\langle l|l'\rangle = \sum_{l=1}^{n_b} v^*_{ml}v_{nl} \\ &= \sum_{l=1}^{n_B} v_{nl}v^\dagger_{lm} = \delta_{nm}. \end{aligned}$$

□

Note: Schmidt decompositions do not exist for more than two subsystems ($N \geq 3$) except for special states.

Example: The three-qubit W-state $|W\rangle = |001\rangle + |010\rangle + |100\rangle$ cannot be turned into Schmidt form $\sum_k \sqrt{d_k}|A_k\rangle \otimes |B_k\rangle \otimes |C_k\rangle$; the GHZ state $|\mathrm{GHZ}\rangle = |000\rangle + |111\rangle$ is explicitly on Schmidt form.

Reduced density operators, partial trace

- Consider bipartite system $A \otimes (BC\ldots)$ prepared in the pure state $|\Psi^{A(BC\ldots)}\rangle$. Suppose we measure an arbitrary $\mathcal{O}$ on A. What is the expectation value? We use the Schmidt form:

$$\begin{aligned}
\langle\mathcal{O}\rangle_{\Psi^{A(BC\ldots)}} &= \sum_{kl}\sqrt{d_k d_l}\langle A_k|\mathcal{O}|A_l\rangle\langle(BC\ldots)_k|(BC\ldots)_l\rangle \\
&= \sum_{kl}\sqrt{d_k d_l}\langle A_k|\mathcal{O}|A_l\rangle\delta_{kl} \\
&= \sum_{k} d_k\langle A_k|\mathcal{O}|A_k\rangle \\
&= \mathrm{Tr}\left(\sum_k d_k|A_k\rangle\langle A_k|\mathcal{O}\right) \\
&= \mathrm{Tr}\left(\rho_A\mathcal{O}\right).
\end{aligned}$$

  Here,

$$\rho_A = \sum_k d_k|A_k\rangle\langle A_k|$$

  is the **reduced density operator**. Since ρ_A determines all expectation values for measurements of observables pertaining to subsystem A, it is the state of subsystem A.

  The reduced density operator can be found by computing the **partial trace**:

$$\rho_A = \mathrm{Tr}_{BC\ldots}\rho_{ABC\ldots} = \sum_n\langle(BC\ldots)_n|\rho_{ABC\ldots}|(BC\ldots)_n\rangle.$$

  Note: ρ_A is a non-pure state unless all d_k except one vanish. This is an important feature underlying the theory of entanglement.

Lecture 3

Shannon entropy

- Entropy measures the information-carrying capacity of a source. It quantifies how much information is gained *on average* when we learn the value of a random variable.
- The **information gain** for an outcome m_k occuring with probability p_k, is defined as $\log(1/p_k) = -\log p_k$ motivated by:

  1. A very common event (high probability p_k) does not carry much information;
  2. 'log' ensures extensivity: $\log(a_1 a_2) = \log a_1 + \log a_2$

  Averaging the information gain leads to the **Shannon entropy**:

$$H(\{p_k\}) = -\sum_k p_k \log p_k$$

  given the probability distribution $\{p_k\}$. The extensive property can be checked by considering an uncorrelated joint probability distribution $p_k q_l$, with marginals $\{p_k\}$ and $\{q_l\}$ such that $\sum_k p_k = 1$ and $\sum_l q_l = 1$, which yields

$$\begin{aligned} H(\{p_k q_l\}) &= -\sum_{kl} p_k q_l \left(\log p_k + \log q_l\right) \\ &= -\sum_l q_l \sum_k p_k \log p_k - \sum_k p_k \sum_l q_l \log q_l \\ &= H(\{p_k\}) + H(\{q_l\}). \end{aligned}$$

  By convention, we take $\log \equiv \log_2$ throughout, since the maximum capacity for a single bit (two outcomes with $p_0 + p_1 = 1$)

$$\max_{p_0,p_1} H(p_0, p_1) = \max_{p_0,p_1} \Big(-p_0 \log p_0 - p_1 \log p_1 \Big) = 1,$$

  which occurs for $p_0 = p_1 = \frac{1}{2}$ (e.g., a 'fair' coin). Thus, with this convention, a fair binary system carries one **bit** of information.

  Note: $p_k \log p_k \to 0$ when $p_k \to 0$. Thus, we define $0 \log 0 \equiv 0$.

von Neumann and relative entropy

- The information-carrying capacity of a quantum source is fully contained in the density operator produced by the source. This ρ describes two types of randomness:

  (i) The randomness in the preparation procedure, as captured by the 'classical' probabilities $\{p_k\}$.

  (ii) The 'intrinsic' randomness in the quantum states $\{|\psi_k\rangle\}$, as captured by Born's probability rule (see Lecture 5).

  To take care of both these types of randomness, one uses the **von Neumann entropy**:

  $$S(\rho) = -\mathrm{Tr}\Big(\rho \log \rho\Big).$$

  This quantity is the Shannon entropy of the eigenvalues of ρ. In particular, it follows that $S(\rho)$ vanishes for pure states, since such states have eigenvalues $p_k = \delta_{kl}$. But $S(\rho)$ is more subtle than this due to the non-commuting nature of quantum probability distributions (density operators) related to the intrinsic 'randomness' in the quantum states $\{|\psi_k\rangle\}$.

- To examine the consequences of the non-commuting feature of density operators, the **relative entropy** is a useful quantity. For two arbitrary density operators ρ_1 and ρ_2 of a quantum system, the relative entropy is defined as

  $$S(\rho_1 \parallel \rho_2) = \mathrm{Tr}\left(\rho_1 \log \rho_1\right) - \mathrm{Tr}\left(\rho_1 \log \rho_2\right).$$

  Its usefulness derives from **Klein's inequality**:

  $$S(\rho_1 \parallel \rho_2) \geq 0$$

  with equality iff $\rho_1 = \rho_2$.

  Roughly, this inequality can be understood from the operator version of the inequality $\log(p/q) \leq p/q - 1 \Rightarrow q(\log q - \log p) \geq q - p$, which reads

  $$A(\log A - \log B) \geq A - B,$$

  by putting $A = \rho_1$, $B = \rho_2$, and taking the trace.

To see the effect of the non-commutative nature of density operators, let us consider the 'classical' relative entropy

$$S(\{p_k\} \parallel \{q_k\}) = \sum_k p_k \log p_k - \sum_k p_k \log q_k$$

between two probability distributions $\{p_k\}$ and $\{q_k\}$. Clearly, $S(\{p_k\}, \{q_k\}) = 0$ if $q_k = p_k$, $\forall k$. Let us compare this with the corresponding quantum-mechanical case of two density operators ρ and σ having the same set of eigenvalues, i.e., with eigendecompositions

$$\begin{aligned} \rho &= \sum_k p_k |\phi_k\rangle\langle\phi_k|, \\ \sigma &= \sum_k p_k |\psi_k\rangle\langle\psi_k|. \end{aligned}$$

We find:

$$S(\rho \parallel \sigma) = \sum_k p_k \log p_k - \sum_{kl} |\langle\phi_k\psi_l|^2 \, p_k \log p_l \neq 0.$$

Thus, although the 'classical' probabilities are the same, the relative entropy may be nonzero due to the difference in eigenbasis of ρ and σ, or, equivalently that ρ and σ may be non-commuting.

- One can derive several important results from Klein's inequality, illustrating the non-commuting feature of density operators.

  Example: For a bipartite state ρ^{AB} and its marginals $\rho_A = \mathrm{Tr}_B \rho^{AB}$ and $\rho_B = \mathrm{Tr}_A \rho^{AB}$, we have $S(\rho^{AB}) \leq S(\rho_A) + S(\rho_B)$ (sub-additivity):

$$\begin{aligned} 0 &\leq S(\rho^{AB} \parallel \rho_A \otimes \rho_B) \\ &= \mathrm{Tr}(\rho^{AB} \log \rho^{AB}) - \mathrm{Tr}\left(\rho^{AB} \log [\rho_A \otimes \rho_B]\right) \\ &= \mathrm{Tr}(\rho^{AB} \log \rho^{AB}) - \mathrm{Tr}(\rho^{AB} \log \rho_A \otimes \hat{1}_B) - \mathrm{Tr}(\rho^{AB} \hat{1}_A \otimes \log \rho_B) \\ &= \mathrm{Tr}(\rho^{AB} \log \rho^{AB}) - \mathrm{Tr}(\rho_A \log \rho_A) - \mathrm{Tr}(\rho_B \log \rho_B) \\ &= -S(\rho^{AB}) + S(\rho_A) + S(\rho_B). \end{aligned}$$

  Thus, the information in the subsystems is "more random" than in the full system.

  Note: The right-hand side $S(\rho_A) + S(\rho_B) - S(\rho^{AB})$ is called the **quantum mutual information** that measures correlations in bipartite quantum states. It vanishes iff $\rho^{AB} = \rho^A \otimes \rho^B$, i.e., when the two subsystems are uncorrelated.

Lecture 4

Completely positive maps

- Generalizes Schrödinger evolution $\rho \mapsto U(t,0)\rho U^\dagger(t,0)$ valid for a **closed** quantum system.
- A completely positive map (CPM) $\rho \mapsto \mathcal{E}(\rho)$ relies on the following physically motivated axioms:

  1. $\mathrm{Tr}\left[\mathcal{E}(\rho)\right]$ is the probability that $\mathcal{E}$ occurs when ρ is the initial state. Thus, $0 \leq \mathrm{Tr}\mathcal{E}(\rho) \leq 1$. With upper equality for all ρ, the map is said to be **trace-preserving**, or CPTP (completely positive trace preserving).
  2. $\mathcal{E}\Big(\sum_k p_k\rho_k\Big) = \sum_k p_k\mathcal{E}(\rho_k)$ (linearity).
  3. $A \geq 0 \Rightarrow \mathcal{E}(A) \geq 0$ (positivity) and $\tilde{A} \geq 0 \Rightarrow (\mathcal{E}\otimes\mathcal{I})(\tilde{A}) \geq 0$ for *any* extension of the system (**complete** positivity).

  The third axiom ensures that remote systems cannot influence the physical nature of the map $\mathcal{E}$.

  The difference between a positive and a completely positive map can be illustrated by the following example:

  Example: Transposition

  $$\Big(|A_k\rangle\langle A_l|\Big)^T = |A_l\rangle\langle A_k|$$

  is a positive but not completely positive map. To see this, consider the extension:

  $$\Big(|A_k\rangle\langle A_l| \otimes |B_m\rangle\langle B_n|\Big)^{T_A} = |A_l\rangle\langle A_k| \otimes |B_m\rangle\langle B_n|$$

  called **partial transposition**. This map does in general not preserve positivity. Such (non-physical) operations play an important role in entanglement theory (see Lecture 6).

- For practical applications, the following theorem is central:

  Theorem: (Without proof) $\mathcal{E}$ is a CPM iff there exists a set of linear operators $\{E_k\}$ such that $\sum_k E_k^\dagger E_k \leq \hat{1}$ (with equality for CPTP) in terms of which

  $$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$$

for all ρ of the system.

This way to write $\mathcal{E}$ is called the **operator-sum** or **Kraus representation**. The operators $\{E_k\}$ are called **Kraus operators**.

Example: **Amplitude damping** of a single qubit.

$$\begin{aligned}\rho \mapsto \mathcal{E}_{ad}(\rho) &= \left(|0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|\right)\rho\left(|0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|\right) \\ &\quad +\gamma|0\rangle\langle 1|\rho|1\rangle\langle 0|, \quad 0 \leq \gamma \leq 1.\end{aligned}$$

This models **decay** of a system: when γ increases towards 1, the density operator tends to $|0\rangle\langle 0|$ for all input ρ. For $\gamma = 1$, the output ensemble is described by the 'ground state' $|0\rangle$. Kraus operators: $E_0 = |0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|$ and $E_1 = \sqrt{\gamma}|0\rangle\langle 1|$.

Example: Let $\{\Pi_k = |\psi_k\rangle\langle\psi_k|\}$ be mutually orthogonal projectors spanning the Hilbert space of the system. The CPTP

$$\rho \mapsto \sum_k \Pi_k \rho \Pi_k = \sum_k |\psi_k\rangle\rho_{kk}\langle\psi_k| \equiv \rho_D,$$

with $\rho_{kk} = \langle\psi_k|\rho|\psi_k\rangle$, is called a **nonselective projective measurement** with the following important property: it can only increase the randomness (as measured by the von Neumann entropy) in the system. This can be seen from Klein's inequality:

$$\begin{aligned}0 \leq S(\rho \parallel \rho_D) &= \mathrm{Tr}(\rho\log\rho) - \sum_k \mathrm{Tr}(\rho|\psi_k\rangle\langle\psi_k|\log\rho_{kk}) \\ &= -S(\rho) - \sum_k \rho_{kk}\log\rho_{kk} = -S(\rho) + S(\rho_D) \\ &\Rightarrow \quad S(\rho) \leq S(\rho_D).\end{aligned}$$

Stinespring dilation

- Physical model for a CPM: Consider a system s and its **environment** e, formally defined as any extension of s such that $s+e$ is a closed system. Let $\{|\epsilon_n\rangle\}_{n\geq 0}$ be an orthonormal basis of $\mathcal{H}_e$ and assume the input $s+e$ state is:

$$\varrho_{se} = \rho \otimes |\epsilon_0\rangle\langle\epsilon_0|.$$

Let U_{se} be the time evolution operator ($s+e$ is a closed system) acting on $\mathcal{H}_s \otimes \mathcal{H}_e$. Thus,

$$\varrho_{se} \mapsto U_{se}\varrho_{se}U_{se}^\dagger = U_{se}\rho \otimes |\epsilon_0\rangle\langle\epsilon_0|U_{se}^\dagger.$$

The transformation of ρ is obtained by tracing over the environment:

$$\rho \mapsto \sum_n \langle\epsilon_n|U_{se}|\epsilon_0\rangle\rho\langle\epsilon_0|U_{se}^\dagger|\epsilon_n\rangle.$$

This construction is called **Stinespring dilation** and defines the operator-sum representation of a CPM with Kraus operators:

$$E_n \equiv \langle\epsilon_n|U_{se}|\epsilon_0\rangle.$$

In fact, it is a CPTP:

$$\begin{aligned}\sum_n E_n^\dagger E_n &= \sum_n \Big(\langle\epsilon_n|U_{se}|\epsilon_0\rangle\Big)^\dagger \langle\epsilon_n|U_{se}|\epsilon_0\rangle \\ &= \sum_n \langle\epsilon_0|U_{se}^\dagger|\epsilon_n\rangle\langle\epsilon_n|U_{se}|\epsilon_0\rangle \\ &= \langle\epsilon_0|U_{se}^\dagger U_{se}|\epsilon_0\rangle = \langle\epsilon_0|\epsilon_0\rangle \cdot \hat{1} = \hat{1}.\end{aligned}$$

- The preceding model has an interesting implication: Since trace is basis independent the operator-sum representation cannot be unique.

To see this, consider the basis transformation:

$$|\epsilon_n\rangle \mapsto |f_n\rangle = \sum_m |\epsilon_m\rangle u_{mn}^\dagger,$$

u_{mn} being unitary. Since this is a basis transformation, the Kraus operators $F_n = \langle f_n|U_{se}|\epsilon_0\rangle$ define the same CPTP. F_n can be related to E_n:

$$F_n = \langle f_n|U_{se}|\epsilon_0\rangle = \sum_m u_{nm}\langle\epsilon_m|U_{se}|\epsilon_0\rangle = \sum_m u_{nm}E_m.$$

Note the similarity with the mixing theorem.

Lecture 5

Measurements

- In the process of measurement, a measuring device interacts with the system and outcomes m_k are read out. This can be viewed as a CPTP with the additional interpretations:

  (i) The Kraus operators represent different measurement outcomes:

  $$E_k \leftrightarrow m_k.$$

  (ii) **Born's probability interpretation**: Given system in state ρ, the probability for the outcome m_k is:

  $$P(m_k|\rho) = \mathrm{Tr}\left(E_k^\dagger E_k \rho\right).$$

  CPTP implies that $\sum_k P(m_k|\rho) = 1$. The positive operators $\{\Pi_k\} = \{E_k^\dagger E_k\}$ form a **positive operator valued measure (POVM)**.

  (iii) Updating probability distribution given outcome m_k:

  $$\rho \mapsto E_k \rho E_k^\dagger.$$

- *Important special case*: **Projective measurements** defined as $E_k = |k\rangle\langle k|$. Here, $\Pi_k = E_k$.

- Theorem: Non-orthogonal states cannot be reliably distinguished.

  Proof: We prove the theorem by *reductio ad absurdum*: the assumption that two non-orthogonal state $|\psi\rangle$ and $|\phi\rangle$ can be disinguished leads to a contradiction with quantum mechanics.

  Suppose such a measurement is possible. Then there must exist Π_1 and Π_2 such that $\langle\psi|\Pi_1|\psi\rangle = 1$ and $\langle\phi|\Pi_2|\phi\rangle = 1$. Since probabilities sum up to one, we have

  $$\langle\psi|\Pi_2|\psi\rangle = \| \sqrt{\Pi_2}|\psi\rangle \|^2 = 0,$$

  which implies $\sqrt{\Pi_2}|\psi\rangle = 0$. We may write $|\phi\rangle = a|\psi\rangle + b|\psi^\perp\rangle$, where $|a|^2 + |b|^2 = 1$ and $|b| < 1$ (non-orthogonality assumption). Thus, $\sqrt{\Pi_2}|\phi\rangle = b\sqrt{\Pi_2}|\psi^\perp\rangle$, which implies

  $$\langle\phi|\Pi_2|\phi\rangle = |b|^2\langle\psi^\perp|\Pi_2|\psi^\perp\rangle \leq |b|^2 < 1,$$

  which contradicts the assumption $\langle\phi|\Pi_2|\phi\rangle = 1$.

  □

Lindblad equation

- *Conceptual framework*: At each (small) time step, the system undergoes a CPTP. This means that the environment does not memorize its interaction with the system. This is the essence of the **Markov approximation**.

- *Derivation*: Let us consider evolution $\rho_t, t \in [0, \tau]$, and divide τ in N small time steps such that $t_{k+1} - t_k = \tau/N \equiv \delta t$. Under the Markovian assumption, we may write the evolution as a sequence of CPTPs:

$$\rho_0 \mapsto \mathcal{E}_{k\delta t,(k-1)\delta t} \circ \ldots \circ \mathcal{E}_{\delta t,0}(\rho_0).$$

For notational simplicity we assume time-translational symmetry:

$$\mathcal{E}_{t+\delta t,t} \equiv \mathcal{E}_{\delta t}.$$

We can turn the evolution into a differential equation for ρ_t. Consider the mapping:

$$\rho_t \mapsto \rho_{t+\delta t} \quad = \quad \mathcal{E}_{\delta t}(\rho_t) = \sum_{n\geq 0} E_n(\delta t)\rho_t E_n^\dagger(\delta t).$$

Since $\mathcal{E}_{\delta t\to 0} = \mathcal{I}$, we make the ansatz

$$\begin{aligned} E_0(\delta t) &= \hat{1} - i(H - iG)\delta t = \hat{1} - iH\delta t - G\delta t, \\ E_{n\geq 1}(\delta t) &= \sqrt{\delta t} L_n, \end{aligned}$$

H and G being Hermitian and L_n arbitrary. G can be found from the normalization associated with the CPTP property:

$$\hat{1} = E_0^\dagger(\delta t)E_0(\delta t) + \sum_{n\geq 1} E_n^\dagger(\delta t)E_n(\delta t) = \hat{1} - 2G\delta t + \delta t \sum_{n\geq 1} L_n^\dagger L_n + O\left(\delta t^2\right),$$

which implies

$$G = \frac{1}{2}\sum_{n\geq 1} L_n^\dagger L_n.$$

Insert this and the ansatz into the mapping $\rho_t \mapsto \rho_{t+\delta t}$, one finds

$$\dot{\rho}_t = \lim_{\delta t\to 0} \frac{1}{\delta t}\Big(\rho_{t+\delta t} - \rho_t\Big) = -i\left[H, \rho_t\right] + \sum_{n\geq 1}\left(L_n \rho_t L_n^\dagger - \frac{1}{2}\{L_n^\dagger L_n, \rho_t\}\right).$$

This is the **Lindblad equation**. L_n are **Lindblad operators**. [G. Lindblad, Comm. Math. Phys. **48**, 119 (1976)]

- *Example*: Single qubit phase damping. Let $L = \sqrt{\gamma}\sigma_z$ and $H = 0$ ('wide open' system). We find:

$$\dot{\rho}_t = \gamma\Big(\sigma_z\rho_t\sigma_z - \frac{1}{2}\{\sigma_z\sigma_z, \rho_t\}\Big) = \gamma\Big(\sigma_z\rho_t\sigma_z - \rho_t\Big).$$

Let $\rho = \frac{1}{2}\big(\hat{1} + \mathbf{r}_t \cdot \boldsymbol{\sigma}\Big)$, yielding:

$$\dot{\mathbf{r}}_t = -2\gamma\,(x_t, y_t, 0)$$

with solution:

$$x_t = x_0 e^{-2\gamma t}; y_t = y_0 e^{-2\gamma t}; z_t = z_0.$$

Lecture 6

Definition of quantum entanglement

- Entanglement describes the separability of quantum systems. It refers to the preparation procedure: If a bipartite state ρ^{AB} can be prepared by using local operations (CPMs) on a product state $\rho^A \otimes \rho^B$, then ρ^{AB} is **separable**. If this is *not* possible, then ρ^{AB} is **entangled**.

  For instance, if Alice and Bob perform coordinated local CPMs:

  $$\{\mathcal{E}_k \otimes \mathcal{F}_k, p_k\},$$

  on the product input state $\rho^A \otimes \rho^B$, then

  $$\rho^A \otimes \rho^B \mapsto \sum_k p_k \mathcal{E}_k(\rho^A) \otimes \mathcal{F}_k(\rho^B) \equiv \sum_k p_k \rho_k^A \otimes \rho_k^B.$$

  Thus, we arrive at the conclusion that a state ρ^{AB} is separable if there exists a decomposition such that

  $$\rho^{AB} = \sum_k p_k \rho_k^A \otimes \rho_k^B.$$

  Otherwise ρ^{AB} is said to be entangled. Entanglement theory focuses on the converse issue, to detect and characterize the entanglement content of a given ρ^{AB}.

- *Special case*: The pure bipartite state $|\Psi^{AB}\rangle$ is entangled if it is impossible to write it as a product state.

  One may check this by using the Schmidt form: If two or more of the Schmidt coefficients d_k are non-zero, then $|\Psi^{AB}\rangle$ is entangled. Equivalently: if the marginal state $\rho^A = \mathrm{Tr}_B |\Psi^{AB}\rangle\langle\Psi^{AB}|$ is non-pure, then $|\Psi^{AB}\rangle$ is entangled.

  The distribution of the Schmidt coefficients d_k provides information about the entanglement in the state. If $d_k = \delta_{kk'}$, then $|\Psi^{AB}\rangle$ is not entangled (product state); if $d_k = d^{-1}$, $d = \min\{n_A, n_B\}$, then $|\Psi^{AB}\rangle$ is said to be **maximally entangled**.

Detecting quantum entanglement: PPT

- Given a bipartite state ρ^{AB} how do we determine whether it is entangled or not? If it is a pure state it is simple: Compute one of the reduced density operators and check whether it is pure or not. However, if ρ^{AB} is non-pure, the task looks formidable since we have to search through an uncountably number of decompositions.

- Entanglement detection is based on the fact that some operations are positive but not completely positive. This can be used to formulate a precise theorem that underlies entanglement detection:

  Theorem: [M. Horodecki *et al.* PLA **223**, 1 (1996)] A mixed bipartite state ρ^{AB} is separable iff for *every* positive map $\mathcal{A}$ on one of the subsystems, the output $\mathcal{A} \otimes \mathcal{I}_B(\rho^{AB}) \geq 0$.

- For $2 \otimes 2$ and $2 \otimes 3$ systems, the above theorem reduces to a useful test (positive partial transpose, **PPT**):

  $$\rho^{AB} \text{ separable} \Leftrightarrow \left(\rho^{AB}\right)^{T_A} \geq 0.$$

  Note that for other dimensionalities there can be entangled states with $\left(\rho^{AB}\right)^{T_A} \geq 0$.

Quantifying quantum entanglement

- *Idea*: Given two bipartite states ρ^{AB} and $\tilde{\rho}^{AB}$, which one is 'most' entangled? That is, we wish to find physically proper **entanglement measures** $\mu\left(\rho^{AB}\right) \geq 0$ that quantify the *amount* of entanglement in a given state. By 'physically proper' essentially means that $\mu\left(\rho^{AB}\right)$ vanishes for separable states and is non-increasing under LOCC. We describe two important proper entanglement measures:

  - PPT suggests **negativity** [G. Vidal and R. F. Werner, PRA **65**, 032314 (2002)]:

    $$\mathcal{N}_e(\rho^{AB}) = \frac{\|\left(\rho^{AB}\right)^{T_A}\| - 1}{2},$$

    with the trace norm $\|A\| = \mathrm{Tr}\sqrt{A^\dagger A}$.

  - For qubit pairs, an entanglement measure can be based on the universal flip operation

    $$\Theta\left(\alpha|0\rangle + \beta|1\rangle\right) = -\beta^*|0\rangle + \alpha^*|1\rangle,$$

    which takes any input qubit state into an orthogonal state. Note that

    $$\Theta = -i\sigma_y\mathcal{K},$$

    where $\mathcal{K}$ is complex conjugation with respect to the computational basis $\{|0\rangle, |1\rangle\}$.

    To see how Θ can be used to quantify entanglement, we first consider a pure two-qubit state $|\Psi^{AB}\rangle$. Acting with the universal flip on each qubit, yields:

    $$\begin{aligned} |\Psi^{AB}\rangle &= \sqrt{d_0}|A_0\rangle \otimes |B_0\rangle + \sqrt{d_1}|A_1\rangle \otimes |B_1\rangle \\ &\mapsto |\tilde{\Psi}^{AB}\rangle = \sqrt{d_0}|A_1\rangle \otimes |B_1\rangle + \sqrt{d_1}|A_0\rangle \otimes |B_0\rangle. \end{aligned}$$

    The distinguishability of the input and output states, as measured by the overlap, defines the **concurrence** $C(\Psi^{AB})$ of the state:

    $$C(\Psi^{AB}) = \left|\langle\tilde{\Psi}^{AB}|\Psi^{AB}\rangle\right| = 2\sqrt{d_0 d_1}.$$

    Since $d_0 + d_1 = 1$, we find $C(\Psi^{AB}) = 2\sqrt{d_0(1-d_0)}$, from which we deduce that $0 \leq C(\Psi^{AB}) \leq 1$. The lower bound is attained for

product states $d_0(1-d_0)=0$; the upper bound is attained for the maximally entangled states $d_0=\frac{1}{2}$.

Concurrence can be generalized to mixed two-qubit states ρ^{AB} according to the following procedure [W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998)]:

- Define the qubit-flipped state $\tilde{\rho}^{AB}=\sigma_y\otimes\sigma_y\mathcal{K}\left(\rho^{AB}\right)\sigma_y\otimes\sigma_y$.
- Compute the eigenvalues λ_k of $\rho^{AB}\tilde{\rho}^{AB}$. These eigenvalues turn out to be real and positive and can be ordered $\lambda_1\geq\ldots\geq\lambda_4$.
- Concurrence of ρ^{AB} takes the form:

$$C(\rho^{AB})=\max\left\{0,\sqrt{\lambda_1}-\sum_{k=2}^{4}\sqrt{\lambda_k}\right\}.$$

One can show that $C(\rho^{AB})=0$ iff there exists a separable decomposition of the density operator. We shall see later (lecture 7) that $C(\rho^{AB})$ has an operational meaning.

Lecture 7

Application: Quantum teleportation

- The key point with entanglement is that it can be used as a *resource* for communicating quantum information between distant parties.

- Important communication scheme: **quantum teleportation**.

  [Theoretical proposal: C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993). Experiment: D. Bouwmeester *et al.* Nature **390**, 575 (1997).]

  Alice wishes to communicate an *unknown* qubit state $|\phi\rangle$ to Bob. They share a maximally entangled state, say $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, where the first (second) entry belongs to Alice (Bob). Thus, the full initial state reads $|\Gamma\rangle = |\phi\rangle \otimes |\Phi_+\rangle$.

  Given $|\Gamma\rangle$, Alice and Bob can teleport $|\phi\rangle$ by using LOCC:

  - LO (Alice): Alice performs a **Bell measurement** meaning a projective measurement onto the four mutually orthogonal maximally entangled **Bell states**:

$$
\begin{aligned}
|\Phi_\pm\rangle &= \frac{1}{\sqrt{2}}\Big(|00\rangle \pm |11\rangle\Big), \\
|\Psi_\pm\rangle &= \frac{1}{\sqrt{2}}\Big(|01\rangle \pm |10\rangle\Big).
\end{aligned}
$$

  - CC (Alice and Bob): For each measurement outcome, Alice sends two classical bits xy of information to Bob:

    if Φ_+ is obtained then she sends $xy = 00$
    if Φ_- is obtained then she sends $xy = 01$
    if Ψ_+ is obtained then she sends $xy = 10$
    if Ψ_- is obtained then she sends $xy = 11$

  - LO (Bob): For each communicated bit pair, Bob performs a unitary map U_{xy} on his part of the original Bell pair:

    if $xy = 00$ is received then he performs $U_{00} = \hat{1}$
    if $xy = 01$ is received then he performs $U_{01} = \sigma_z$
    if $xy = 10$ is received then he performs $U_{10} = \sigma_x$
    if $xy = 11$ is received then he performs $U_{11} = \sigma_y$

    In all four cases, Bob's qubit ends up in the state $|\phi\rangle$ thereby completing the teleportation.

Entanglement of formation

- Entanglement can be viewed as a *resource* that can be consumed (used) to perform certain tasks (e.g., teleporting unknown quantum states). It is natural to ask how to quantify this resource.

  Entanglement of formation E_F is an entropy-based entanglement measure that captures the feature that the reduced state of a pure entangled state is mixed (non-pure). In the two-qubit case, it turns out that E_F has a relation to concurrence.

- Entanglement of formation, pure state case. Let $|\psi^{AB}\rangle$ be a bipartite state. We define:

  $$E_F(\psi^{AB}) = S(\rho_A)$$

  with $\rho_A = \mathrm{Tr}_B|\psi^{AB}\rangle\langle\psi^{AB}|$ the reduced density operator.

  Properties:

  - E_F vanishes for product states, i.e., states of the form $|\psi^{AB}\rangle = |\psi^A\rangle \otimes |\psi^B\rangle$: The reduced states of $|\psi^A\rangle \otimes |\psi^B\rangle$ are pure, i.e., $\rho_A = |\psi^A\rangle\langle\psi^A|$ and $\rho_B = |\psi^B\rangle\langle\psi^B|$, which implies $E_F(\psi^{AB}) = 0$ since the von Neumann entropy of a pure state vanishes.
  - For a maximally entangled state, we find

    $$\rho_A = \mathrm{Tr}_B\left(|\mathrm{MES}\rangle\langle\mathrm{MES}|\right) = \frac{1}{d}\sum_{m=1}^{d}|A_m\rangle\langle A_m|.$$

    This implies

    $$E_F(\mathrm{MES}) = -\sum_{m=1}^{d}\frac{1}{d}\log\frac{1}{d} = \log d,$$

    which is the maximal value for qudits. For qubit-pairs ($d = 2$) we find: $E_F(\mathrm{MES}) = 1$. One says that a maximally entangled qubit state represents one unit of entanglement, i.e., one **e-bit**.
  - For a qubit-pair state with Schmidt coefficients d_0 and d_1:

    $$E_F(\psi^{AB}) = -d_0\log d_0 - d_1\log d_1 = H(d_0, d_1 = 1 - d_0) \equiv h(d_0),$$

    $h(d_0) = -d_0\log d_0 - (1-d_0)\log(1-d_0)$ being the one-bit Shannon entropy. From normalization $d_0 + d_1 = 1$, one finds:

    $$d_0 = \frac{1 + \sqrt{1 - [C(\psi^{AB})]^2}}{2},$$

where $C(\psi^{AB}) = 2\sqrt{d_0 d_1}$ is the concurrence of ψ^{AB} and we have assumed that $d_0 \geq \frac{1}{2} \geq d_1$ (without loss of generality). We thus obtain:

$$E_F(\psi^{AB}) = h\left(\frac{1+\sqrt{1-[C(\psi^{AB})]^2}}{2}\right),$$

which demonstrates a direct relation between concurrence and entanglement of formation.

- Entanglement of formation, general case. Consider the bipartite state ρ^{AB}. We define:

$$E_F(\rho^{AB}) = \min_{\{p_k, \psi_k^{AB}\}} \sum_k p_k E_F(\psi_k^{AB}),$$

where minimum is taken over all decompositions of ρ^{AB} into pure states.

  - This kind of extension of a quantity defined on pure states to mixed states is called a **convex roof** construction.
  - For qubit pairs one can show [W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998)]:

$$E_F(\rho^{AB}) = h\left(\frac{1+\sqrt{1-[C(\rho^{AB})]^2}}{2}\right).$$

    This provides an operational meaning to concurrence for arbitrary two-qubit states.

Multipartite entanglement

- How can one characterize and quantify multipartite entanglement? What is the meaning of *genuine* multipartite entanglement?

- We limit to pure states. We may write a general multipartite state as

$$|\Psi^{ABC\ldots}\rangle = \sum_{n_A n_B n_C \ldots} a_{n_A n_B n_C \ldots} |n_A n_B n_C \ldots\rangle,$$

  where $|n_A n_B n_C \ldots\rangle$ span the Hilbert space of the full system. One can characterize entanglement in terms of:

  - bipartite partitioning of the full state $|\Psi^{A(BC\ldots)}\rangle$, $|\Psi^{B(AC\ldots)}\rangle$, $|\Psi^{(AB)(C\ldots)}\rangle$ etc;
  - bipartite mixed state entanglement $\rho^{AB} = \mathrm{Tr}_{C\ldots}|\Psi^{ABC\ldots}\rangle\langle\Psi^{ABC\ldots}|$;
  - SL-invariant polynomials: Let $H_n = (\mathbb{C})^{\otimes n}$ and $G = SL(2,\mathbb{C})^{\otimes n}$. $I : H_n \mapsto \mathbb{C}$ is SL-invariant if $I(g\Psi^{AB\ldots}) = I(\Psi^{AB\ldots})$ for all $g \in G$. $SL(2,\mathbb{C})$ is the group of complex 2×2 matrices with unit determinant.

- The simplest multipartite case: three qubits prepared in $|\Psi^{ABC}\rangle$.
  Consider for instance the three-qubit state $|W^{ABC}\rangle = \frac{1}{\sqrt{3}}\big(|001\rangle + |010\rangle + |100\rangle\big)$ and the state $|\mathrm{GHZ}^{\mathrm{ABC}}\rangle = \frac{1}{\sqrt{2}}\big(|000\rangle + |111\rangle\big)$. One finds:

$$\begin{aligned}
|W^{A(BC)}\rangle &= \sqrt{\frac{2}{3}}|0\rangle \otimes |\Psi_+\rangle + \sqrt{\frac{1}{3}}|1\rangle \otimes |00\rangle \to C\left(W^{A(BC)}\right) = \frac{2\sqrt{2}}{3}, \\
|\mathrm{GHZ}^{\mathrm{A(BC)}}\rangle &= \sqrt{\frac{1}{2}}|0\rangle \otimes |00\rangle + \sqrt{\frac{1}{2}}|1\rangle \otimes |11\rangle \to C_{\mathrm{GHZ;A(BC)}} = 1, \\
\rho_W^{AB} &= \frac{2}{3}|\Psi_+\rangle\langle\Psi_+| + \frac{1}{3}|00\rangle\langle 00| \to C\left(\rho_W^{AB}\right) = \frac{2}{3}, \\
\rho_{\mathrm{GHZ}}^{AB} &= \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| \to C\left(\rho_{\mathrm{GHZ}}^{AB}\right) = 0,
\end{aligned}$$

  These numbers give apparently conflicting views on the entanglement in these two states.

- **Residual entanglement** τ_{ABC} is a way to combine the above concurrences so as to provide a measure of genuine multipartite entanglement

in a pure tripartite state Ψ^{ABC}. It is defined as [V. Coffman, J. Kundu, W.K. Wootters Phys. Rev. A **61**, 052306 (2000)]:

$$\tau\left(\Psi^{ABC}\right) = C^2_{A(BC)} - C^2_{AB} - C^2_{AC}.$$

It has the following properties:

- $\tau_{ABC} \geq 0$.
- τ_{ABC} is invariant under permutation of A, B, and C.
- τ_{ABC} is invariant under $SL(2,\mathbb{C})^{\otimes 3}$; transformations that belong to the group of complex 2×2 matrices $\mathbf{A}$ with $\det \mathbf{A} = 1$.
- τ_{ABC} defines a SL-invariant polynomial: For $|\Psi^{ABC}\rangle = |0\rangle|\psi_0\rangle + |1\rangle|\psi_1\rangle$, the function

$$I(\Psi^{ABC}) = \det \begin{pmatrix} \langle\tilde{\psi}_0|\psi_0\rangle & \langle\tilde{\psi}_0|\psi_1\rangle \\ \langle\tilde{\psi}_1|\psi_0\rangle & \langle\tilde{\psi}_1|\psi_1\rangle \end{pmatrix}$$

  is an SL-invariant degree four polynomial in the expansion coefficients $\langle klm|\Psi^{ABC}\rangle$ for which $\tau_{ABC} = |I(\Psi^{ABC})|$.

- One finds: $\tau\left(W^{ABC}\right) = 0$ and $\tau\left(\mathrm{GHZ}^{\mathrm{ABC}}\right) = 1$. W^{ABC} has bipartite entanglement ($C\left(\rho_W^{AB}\right) \neq 0$) but no tripartite entanglement ($\tau\left(W^{ABC}\right) = 0$); $\mathrm{GHZ}^{\mathrm{ABC}}$ has no bipartite entanglement ($C\left(\rho_W^{AB}\right) = 0$) but has tripartite entanglement ($\tau\left(W^{ABC}\right) \neq 0$).

- There are two types of pure state entanglement for three qubits: W and GHZ: single copies of states of these classes cannot be converted into each other by means of local invertible transformations (SLOCC).

- There is no straightforward extension of these considerations to more than three qubits. Each number of qubits has its special characters and must therefore be treated separately.

Lecture 8

Bell's inequalities

- The **Bell inequalities** [J. S. Bell, Physics **1**, 195 (1964)] put restrictions on the strength of classical correlations. We focus on the Clauser-Horne-Shimony-Holt (**CHSH**) inequality [J. F. Clauser *et al.*, Phys. Rev. Lett. **23**, 880 (1969)], which provides a bound on the correlations between two pairs of two-valued classical variables $a, a' = \pm 1$ and $b, b' = \pm 1$. It reads:

$$|\langle ab\rangle + \langle ab'\rangle + \langle a'b\rangle - \langle a'b'\rangle| \leq 2.$$

- *Derivation*: The assumption $a, a', b, b' = \pm 1$ implies the identity:

$$a(b+b') + a'(b-b') = \pm 2.$$

  Averaging gives:

$$-2 \leq \langle a(b+b') + a'(b-b')\rangle \leq 2 \Rightarrow |\langle a(b+b') + a'(b-b')\rangle| \leq 2.$$

- We now show that quantum mechanics violate CHSH. Suppose we have two qubits prepared in a state ρ^{AB} and distributed on Alice (A) and Bob (B). Alice performs local projective measurements of the observables $\mathbf{a}\cdot\boldsymbol{\sigma}, \mathbf{a}'\cdot\boldsymbol{\sigma}$ and Bob similarly of $\mathbf{b}\cdot\boldsymbol{\sigma}, \mathbf{b}'\cdot\boldsymbol{\sigma}$; $\mathbf{a}, \ldots, \mathbf{b}'$ being unit vectors so that each observable has two possible outcomes ± 1. Define the **Bell operator**:

$$\begin{aligned}\mathcal{B}_{\rm CHSH} &= (\mathbf{a}\cdot\boldsymbol{\sigma})\otimes(\mathbf{b}\cdot\boldsymbol{\sigma}) + (\mathbf{a}\cdot\boldsymbol{\sigma})\otimes(\mathbf{b}'\cdot\boldsymbol{\sigma}) \\ &\quad +(\mathbf{a}'\cdot\boldsymbol{\sigma})\otimes(\mathbf{b}\cdot\boldsymbol{\sigma}) - (\mathbf{a}'\cdot\boldsymbol{\sigma})\otimes(\mathbf{b}'\cdot\boldsymbol{\sigma}),\end{aligned}$$

  which clearly has the same structure as in the CHSH inequality. In other words, the validity of the CHSH inequality in quantum mechanics can be tested by evaluating the expectation value (average) of $\mathcal{B}_{\rm CHSH}$ for ρ^{AB}.

  For $\rho^{AB} = |\Psi_-\rangle\langle\Psi_-|$, one finds:

$$\begin{aligned}\left|\langle\mathcal{B}_{\rm CHSH}\rangle_{\Psi_-}\right| &= |{\rm Tr}\left(|\Psi_-\rangle\langle\Psi_-|\mathcal{B}_{\rm CHSH}\right)| \\ &= |\langle\Psi_-|\mathcal{B}_{\rm CHSH}|\Psi_-\rangle| \\ &= |\mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{b}' + \mathbf{a}'\cdot\mathbf{b} - \mathbf{a}'\cdot\mathbf{b}'|\,.\end{aligned}$$

By putting $\mathbf{a}, \ldots, \mathbf{b}'$ in a plane, making angles $\varphi_a, \ldots, \varphi_{b'}$ relative to some fixed direction in this plane, we find:

$$\begin{aligned} \left|\langle \mathcal{B}_{\mathrm{CHSH}} \rangle_{\Psi_-}\right| &= \big|\cos\left(\varphi_a - \varphi_b\right) + \cos\left(\varphi_a - \varphi_{b'}\right) \\ &\quad + \cos\left(\varphi_{a'} - \varphi_b\right) - \cos\left(\varphi_{a'} - \varphi_{b'}\right)\big| \,. \end{aligned}$$

By choosing $\varphi_a = 0$, $\varphi_b = \frac{\pi}{4}$, $\varphi_{a'} = \frac{\pi}{2}$, and $\varphi_{b'} = -\frac{\pi}{4}$, we find:

$$\left|\langle \mathcal{B}_{\mathrm{CHSH}} \rangle_{\Psi_-}\right| = \left|\cos\frac{\pi}{4} + \cos\frac{\pi}{4} + \cos\frac{\pi}{4} - \cos\frac{3\pi}{4}\right| = 2\sqrt{2} > 2.$$

Thus, quantum mechanics violates the CHSH inequality. In fact, one can show that $2\sqrt{2}$ is the largest possible violation of CHSH (**Cirelson bound**).

- The reason why we cannot expect the CHSH inequality to hold in quantum mechanics is that although the outcomes of each of the pairwise qubit measurements are ± 1, we cannot expect the eigenvalues of $\mathcal{B}_{\mathrm{CHSH}}$ to be ± 2. This would only hold if the four terms in $\mathcal{B}_{\mathrm{CHSH}}$ had the same eigenvectors, i.e., that they were mutually commuting. Clearly this is not the case for the above chosen $\varphi_a, \varphi_{a'}, \varphi_b, \varphi_{b'}$. Thus, the assumption $a(b + b') + a'(b - b') = \pm 2$ is not *a priori* valid, basically due to the complementarity feature of observables in quantum mechanics. The violation of the CHSH inequality demonstrates that the underlying assumption behind the CHSH inequality underestimates the potential strength of correlations in quantum mechanics.

- *Note*: ρ^{AB} violates CHSH $\Rightarrow$ ρ^{AB} is entangled. The converse does not hold: there are entangled states that do not violate CHSH. Thus, violation of CHSH is a stronger correlation criterion than entanglement.

Application: Quantum key distribution, E91 protocol

- Idea: Use maximally entangled particles for **quantum key distribution** and test eavesdropping by means of Bell's inequalities. Security is guaranteed if the CHSH inequality is violated.
- Proposed in [A. Ekert Phys. Rev. Lett. **67**, 661 (1991)]. First *application* of quantum entanglement.
- Alice and Bob wish to share a secret key that can be used to encode and decode messages.
  - Alice and Bob share a large set of Bell singlet states $|\Psi_-\rangle$.
  - Alice perform measurements randomly in directions $\mathbf{a}, \mathbf{a}', \mathbf{a}''$ and Bob measures randomly in directions $\mathbf{b}, \mathbf{b}', \mathbf{b}''$ on each Bell pair and record the outcomes. All measurement directions lie in the same plane and define angles $\varphi_a = 0, \varphi_{a'} = \frac{\pi}{2}, \varphi_{a''} = \frac{\pi}{4}$ and $\varphi_b = \frac{\pi}{4}, \varphi_{b'} = -\frac{\pi}{4}, \varphi_{b''} = 0$ relative a common reference direction.
  - Alice and Bob announce their measurement directions and sort them into one group ($\mathcal{G}_1$) where they happen to choose different directions and one ($\mathcal{G}_2$) where they happen to choose the same direction. $\mathcal{G}_2$ consists of the pairs $(\varphi_a, \varphi_{b''})$ and $(\varphi_{a''}, \varphi_b)$.
  - Alice and Bob reveal the outcomes but only from $\mathcal{G}_1$. They test the CHSH inequality from the set $(\varphi_a, \varphi_b), (\varphi_a, \varphi_{b'}), (\varphi_{a'}, \varphi_b), (\varphi_{a'}, \varphi_{b'})$, which as we have seen above saturates the Cirelson bound $2\sqrt{2}$ in the ideal case. On the other hand, if an eavesdropper has acquired any useful information about the key, then the CHSH cannot be violated. Thus,
    (i) CHSH inequality violated $\Rightarrow$ no eavesdropper.
    (ii) CHSH inequality holds $\Rightarrow$ absence of an eavesdropper cannot be guaranteed.
  - In case (i), group $\mathcal{G}_2$ can be used to generate a common key since there is perfect anti-correlation of outcomes when Alice and Bob choose the same direction ($\langle\Psi_-|\mathbf{a}\cdot\boldsymbol{\sigma}\otimes\mathbf{b}''\cdot\boldsymbol{\sigma}|\Psi_-\rangle = -\mathbf{a}\cdot\mathbf{b}'' = -1$). The security of the key is guaranteed by the violation of the CHSC inequality (i.e. by the laws of physics) and not by the complexity of some computational task as in ordinary classical cryptography (such as factorization, as in the RSA scheme).

Lecture 9

Quantum cloning

- Cloning can be used for superluminal signalling.

  Suppose Alice and Bob share a single copy of the Bell state $|\Phi_+\rangle$. Alice encodes a bit of information into a filtering measurement of either σ_z (bit value 0) or σ_x (bit value 1). Bob task is to tell which bit value (σ_z or σ_x) Alice had chosen. Is this possible?

  We show that it would indeed be possible if cloning was possible.

  Suppose Alice choose σ_z and she obtains $|0\rangle$. This changes Bob's state to $|0\rangle$, which is cloned into $|\psi\rangle = |0\rangle^{\otimes n}$. Suppose Alice instead chooses σ_x and she obtains $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. This changes Bob's state to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, which is cloned into $|\phi\rangle = \frac{1}{2^{n/2}}(|0\rangle + |1\rangle)^{\otimes n}$. Bob can asymptotically distinguish these two cloned states:

  $$|\langle\psi|\phi\rangle|^2 = \frac{1}{2^n},$$

  as this overlap tends to zero when the number of copies n tends to infinity.

  Now, the problem with this is that the above cloning scheme would allow for superluminal signalling since Alice and Bob could be very far apart. This is to say that quantum states cannot be cloned:

- Theorem: (No-cloning) Non-orthogonal quantum states cannot be cloned.

  Proof: We prove the theorem by *reductio ad absurdum*: the assumption that a quantum cloning device of non-orthogonal states exists leads to a contradiction with quantum mechanics.

  Suppose we have a quantum device with three slots, the data and the target slots, as well as an auxiliary system whose role is to make the data+target+auxiliary a closed system. The data slot starts out in the *unknown* pure state $|\psi\rangle$. The target slot is prepared in some standard pure state $|0\rangle$ and the auxiliary system in $|a\rangle$.

  Assume now the device can perform perfect cloning. This means that there should exist a unitary U (since data+target+auxiliary is a closed system) such that

  $$U(|\psi\rangle \otimes |0\rangle \otimes |a\rangle) = |\psi\rangle \otimes |\psi\rangle \otimes |a_\psi\rangle.$$

Since U is by assumption a perfect cloner, it should also clone any other state $|\phi\rangle$ in the same way, i.e.:

$$U(|\phi\rangle \otimes |0\rangle \otimes |a\rangle) = |\phi\rangle \otimes |\phi\rangle \otimes |a_\phi\rangle.$$

We assume that $|\phi\rangle \neq |\psi\rangle$ and non-orthogonal, i.e., $0 < |\langle\phi|\psi\rangle| < 1$. By taking the scalar products of the left- and right-hand sides of the cloning transformations, we find:

$$\begin{aligned}(\langle\phi| \otimes \langle 0| \otimes \langle a|)\, U^\dagger U\, (|\psi\rangle \otimes |0\rangle \otimes |a\rangle) &= \langle\phi|\psi\rangle\langle 0|0\rangle\langle a|a\rangle = \langle\phi|\psi\rangle \\ = (\langle\phi| \otimes \langle\phi| \otimes \langle a_\phi|)\, (|\psi\rangle \otimes |\psi\rangle \otimes |a_\psi\rangle) &= (\langle\phi|\psi\rangle)^2\, \langle a_\phi|a_\psi\rangle.\end{aligned}$$

Since $|\langle\phi|\psi\rangle| > 0 \Rightarrow \langle\phi|\psi\rangle \neq 0$, which implies:

$$\langle\phi|\psi\rangle\langle a_\phi|a_\psi\rangle = 1 \Rightarrow |\langle\phi|\psi\rangle\langle a_\phi|a_\psi\rangle| = 1. \tag{1}$$

Since $|\langle\phi|\psi\rangle| < 1$, we obtain:

$$|\langle\phi|\psi\rangle\langle a_\phi|a_\psi\rangle| < |\langle a_\phi|a_\psi\rangle| \leq 1,$$

which contradicts Eq. (1).

□

- *Note*: Quantum teleportation is *not* cloning, since Alice copy of the unknown state $|\phi\rangle$ is lost (*consumed*) after her Bell measurement.

Application: Quantum copying

- Even if perfect cloning is impossible, one may ask *how well* an unknown state can be copied. More precisely, we wish to construct a scheme that copies all pure input states of a single qubit with the same and largest possible quality (fidelity). This is the idea of **optimal cloning** or **quantum copying**.

  Suppose we wish to copy the unknown state $|\phi\rangle = a|0\rangle + b|1\rangle$ of a data qubit as well as possible on a target qubit. The copying quality should be independent of $|\phi\rangle$. This can be done with the help of an extra auxiliary qubit as follows:

  - Implement a unitary transformation U acting on all three qubits such that:

$$\begin{aligned} U : |000\rangle &\mapsto \sqrt{\frac{2}{3}}|000\rangle - \frac{1}{\sqrt{6}}\Big(|011\rangle + |101\rangle\Big), \\ U : |100\rangle &\mapsto -\sqrt{\frac{2}{3}}|111\rangle + \frac{1}{\sqrt{6}}\Big(|010\rangle + |100\rangle\Big) \end{aligned}$$

    with entries from left to right corresponding to the data, target, and auxiliary qubit, respectively.

  - By linearity, we find:

$$\begin{aligned} U : |\phi 00\rangle \mapsto |\Psi\rangle \;=\; & \sqrt{\frac{2}{3}}\Big(a|000\rangle - b|111\rangle\Big) - \frac{a}{\sqrt{6}}\Big(|011\rangle + |101\rangle\Big), \\ & + \frac{b}{\sqrt{6}}\Big(|010\rangle + |100\rangle\Big). \end{aligned}$$

  - The output state is unchanged when permuting the data and target qubits; thus, the reduced states ρ_{data} and ρ_{target} are identical. In other words, we have obtained two copies of the same state. We find:

$$\rho_{\text{data}} = \rho_{\text{target}} = \frac{5}{6}|\phi\rangle\langle\phi| + \frac{1}{6}\Big(\hat{1} - |\phi\rangle\langle\phi|\Big) \equiv \rho.$$

  - The quality of the cloning is measured by the (squared) **fidelity**:

$$\mathcal{F}^2(|\phi\rangle, \rho) = \langle\phi|\rho|\phi\rangle = \frac{5}{6}$$

    Thus, the quality of the cloning procedure is about 83 % independent of the input state $|\phi\rangle$. It can be shown that this is the optimal fidelity for a single qubit.

Lecture 10

Distance measures between probability distributions

- What does it mean to say that two items of information are *similar*? And how is the information *preserved* in some process? These two questions can be tackled by introducing distance measures between quantum probability distributions as represented by density operators.
- Let us first consider the 'classical' case. Let $\{p_k\} = (p_1, p_2, \ldots, p_K)$ and $\{q_k\} = (q_1, q_2, \ldots, q_K)$ be two probability distributions over the same index k. How similar are they? Two measures that answer this question:
  - The l_1 distance:
    $$D(\{p_k\}, \{q_k\}) = \frac{1}{2} \sum_{k=1}^{K} \big|p_k - q_k\big|,$$
    which vanishes iff $p_k = q_k$.
  - The fidelity:
    $$F(\{p_k\}, \{q_k\}) = \sum_k \sqrt{p_k q_k},$$
    which takes its maximal value $F_{\max} = 1$ iff $p_k = q_k$.

Trace distance

- The quantum analog of the l_1 distance is the **trace distance**:

$$D(\rho_1, \rho_2) = \frac{1}{2}\mathrm{Tr}\left|\rho_1 - \rho_2\right|.$$

- $D(\rho_1, \rho_2)$ satisfies all criteria for being a proper distance:

  (i) *Non-negative*: $D(\rho_1, \rho_2) \geq 0$. $\left|\rho_1 - \rho_2\right|$ has eigenvalues $\lambda_k \geq 0$, which implies $\mathrm{Tr}\left|\rho_1 - \rho_2\right| = \sum_k \lambda_k \geq 0$.

  (ii) *Non-degenerate*: $D(\rho_1, \rho_2) = 0$ implies $\rho_2 = \rho_1$. Since $D(\rho_1, \rho_2) = \frac{1}{2}\sum_k \lambda_k$ with $\lambda_k \geq 0$, $D(\rho_1, \rho_2) = 0$ can only happen if all $\lambda_k = 0$, which implies $|\rho_1 - \rho_2| = 0$.

  (iii) *Symmetric*: $D(\rho_1, \rho_2) = D(\rho_2, \rho_1)$. Follows directly from definition.

  (iv) *Triangle inequality*:

$$D(\rho_1, \rho_2) \leq D(\rho_1, \rho_3) + D(\rho_3, \rho_2)$$

  for any density operators ρ_1, ρ_2, ρ_3 of the system.

  Proof: $\rho_1 - \rho_2$ is an Hermitian operator with spectrum on $[-1, 1]$. We may thus divide it into a positive and a negative part:

$$\begin{aligned} \rho_1 - \rho_2 &= V_{12}^+ - V_{12}^- \text{ with } V_{12}^{\pm} \geq 0 \\ &\Rightarrow \left|\rho_1 - \rho_2\right| = V_{12}^+ + V_{12}^- \end{aligned}$$

  and introduce the projection Π_{12} onto V_{12}^+, i.e., $V_{12}^+ = \Pi_{12}(\rho_1 - \rho_2)$ and $V_{12}^- = (\Pi_{12} - \hat{1})(\rho_1 - \rho_2)$. By taking the trace we find:

$$\mathrm{Tr}(\rho_1 - \rho_2) = 0 = \mathrm{Tr}(V_{12}^- - V_{12}^+) \Rightarrow \mathrm{Tr}V_{12}^- = \mathrm{Tr}V_{12}^+.$$

  We obtain:

$$\begin{aligned} D(\rho_1, \rho_2) &= \frac{1}{2}\mathrm{Tr}\left|\rho_1 - \rho_2\right| = \frac{1}{2}\mathrm{Tr}\left(V_{12}^+ + V_{12}^-\right) = \mathrm{Tr}V_{12}^+ \\ &= \mathrm{Tr}\left[\Pi_{12}(\rho_1 - \rho_2)\right]. \end{aligned}$$

  Since trace is linear, we may write:

$$D(\rho_1, \rho_2) = \mathrm{Tr}\left[\Pi_{12}(\rho_1 - \rho_3)\right] + \mathrm{Tr}\left[\Pi_{12}(\rho_3 - \rho_2)\right].$$

Now,

$$\begin{aligned}\mathrm{Tr}\left[\Pi_{12}(\rho_1-\rho_3)\right] &= \mathrm{Tr}\left[\Pi_{12}(V_{13}^{+}-V_{13}^{-})\right] \\ &= \mathrm{Tr}\left[\Pi_{12}V_{13}^{+}\right]-\mathrm{Tr}\left[\Pi_{12}V_{13}^{-}\right] \\ &\leq \mathrm{Tr}\left[\Pi_{12}V_{13}^{+}\right]\leq \mathrm{Tr}V_{13}^{+}=D(\rho_1,\rho_3),\end{aligned}$$

where we have used $\mathrm{Tr}(\Pi A)\leq \mathrm{Tr}(A)$ for any $A\geq 0$. We conclude:

$$D(\rho_1,\rho_2) \leq D(\rho_1,\rho_3)+D(\rho_3,\rho_2).$$

□

- Physical interpretation of trace distance: We have found that

$$[\Pi(\rho_1-\rho_2)]\leq \mathrm{Tr}D(\rho_1,\rho_2)$$

with equality if $\Pi=\Pi_{12}$. The left-hand side may be interpreted as the probability difference for obtaining the POVM element $\Pi\geq 0$, given the state ρ_1 or the state ρ_2. Trace distance is thus the maximal value of this probability difference.

Application: Measure of non-Markovianity

- Theorem: Trace distance is *contractive* under CPTP, i.e.,

$$D(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)) \leq D(\rho_1, \rho_2)$$

for any CPTP $\rho \mapsto \mathcal{E}(\rho)$.

Proof: Same technique as when proving the triangle inequality above and by using that $\mathrm{Tr}\mathcal{E}(\rho) = \mathrm{Tr}\rho$ for CPTP.

- The contractive nature of trace distance under CPTP can be used to quantify the **non-Markovianity** of a given evolution $\rho \mapsto \mathcal{E}_t(\rho)$ [H.-P. Breuer *et al.*, Phys. Rev. Lett. **103**, 042103 (2010)]. The idea is based on the observation that any Markovian evolution can be viewed as a composite map of near-unity CPTPs, from which follows, according to the theorem above, that trace distance cannot increase when a system undergoes Markovian evolution. Thus, if the trace distance would increase for some time interval, there must be a non-Markovian component involved in the equations of motion of the system.

- This observation suggests the following **measure of non-Markovianity**:

$$\mathcal{N}(\mathcal{E}_t) = \max_{\rho_1,\rho_2} \int_{\delta(t;\rho_1,\rho_2)>0} \delta(t;\rho_1,\rho_2)dt,$$

where

$$\delta(t;\rho_1,\rho_2) = \frac{d}{dt} D\left[\mathcal{E}_t(\rho_1), \mathcal{E}_t(\rho_2)\right].$$

- To find the optimal pair is in general difficult, but it can be proved that they must satisfy the following two conditions:

  - The should have orthogonal support, i.e., the subspaces spanned by nonzero eigenvalues should be orthogonal.
  - They should both have at least one zero eigenvalue.

  For a qubit: The optimal states should be orthogonal and pure.

Lecture 11

Fidelity

- The fidelity for two quantum states ρ_1 and ρ_2 is defined as:

$$F(\rho_1, \rho_2) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right),$$

where $\sqrt{A}$ is defined as $\left(\sqrt{A}\right)^2 = A$. One can prove that $0 \leq F(\rho_1, \rho_2) \leq 1$ with equality in the first inequality iff ρ_1 and ρ_2 have orthogonal support and equality in the second inequality iff $\rho_2 = \rho_1$.

- Fidelity is symmetric: $F(\rho_2, \rho_1) = F(\rho_1, \rho_2)$. To see this, we need to introduce the concept of **purification**.

Consider a density operator ρ acting on Hilbert space $\mathcal{H}$. A pure state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}_a$ is a purification of ρ if

$$\rho = \mathrm{Tr}_a |\psi\rangle\langle\psi|.$$

Theorem: (Uhlmann) [A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976)]

$$F(\rho_1, \rho_2) = \max_{|\psi_1\rangle, |\psi_2\rangle} |\langle\psi_1|\psi_2\rangle|,$$

where the maximization is over all purifications $|\psi_1\rangle$ and $|\psi_2\rangle$ of ρ_1 and ρ_2, respectively.

Proof: The proof uses some results in linear algebra: for A arbitrary and V unitary, we have:

$$\begin{aligned} |\mathrm{Tr}(AV)| &= |\mathrm{Tr}(|A|UV)| = \left|\mathrm{Tr}(\sqrt{|A|}\sqrt{|A|}UV)\right| \\ &\leq \sqrt{\mathrm{Tr}|A|\mathrm{Tr}(V^\dagger U^\dagger |A| VU)} = \mathrm{Tr}|A| \end{aligned} \tag{2}$$

by using the right polar decomposition $A = |A|U$ and Cauchy-Schwarz for the Hilbert-Schmidt inner product $(A, B) = \mathrm{Tr}(A^\dagger B)$, i.e., $|(A, B)| \leq \sqrt{|(A, A)| \cdot |(B, B)|}$; for $|\mu\rangle = \sum_k |kk\rangle \in \mathcal{H} \otimes \mathcal{H}_a$, we have:

$$\begin{aligned} \langle\mu|(A \otimes B)|\mu\rangle &= \sum_l \langle ll|A \otimes B \sum_k |kk\rangle = \sum_{kl} \langle l|A|k\rangle\langle l|B|k\rangle \\ &= \sum_{kl} \langle k|A^T|l\rangle\langle l|B|k\rangle = \sum_k \langle k|A^T B|k\rangle \\ &= \mathrm{Tr}(A^T B) \end{aligned} \tag{3}$$

We may write arbitrary purifications of ρ_1 and ρ_2 as:

$$\begin{aligned} |\psi_1\rangle &= (U \otimes \sqrt{\rho_1} U_a)|\mu\rangle, \\ |\psi_2\rangle &= (V \otimes \sqrt{\rho_2} V_a)|\mu\rangle. \end{aligned}$$

Thus, we find:

$$\begin{aligned} |\langle\psi_1|\psi_2\rangle| &= \left|\langle\mu|U^\dagger V \otimes U_a\sqrt{\rho_1}\sqrt{\rho_2}V_a|\mu\rangle\right| \\ &= \left|\mathrm{Tr}\left(V^T U^* U_a \sqrt{\rho_1}\sqrt{\rho_2} V_a\right)\right| = \left|\mathrm{Tr}\left(V_a V^T U^* U_a \sqrt{\rho_1}\sqrt{\rho_2}\right)\right| \end{aligned}$$

by using Eq. (3) and cyclicity of trace. Now, by using Eq. (2), we find:

$$|\langle\psi_1|\psi_2\rangle| \leq \mathrm{Tr}\,|\sqrt{\rho_1}\sqrt{\rho_2}| = \mathrm{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}.$$

Equality is obtained by setting $U = U_a = V = \hat{1}$ and $V_a = W^\dagger$, where $\sqrt{\rho_1}\sqrt{\rho_2} = \left|\sqrt{\rho_1}\sqrt{\rho_2}\right| W$. This last point further shows that

$$F(\rho_1, \rho_2) = \max_{|\psi_2\rangle} |\langle\psi_1|\psi_2\rangle| \,.$$

□

- Uhlmann's theorem provides a physical interpretation of fidelity as the worst case measure of distinguishability (orthogonality) of purifications of the states ρ_1 and ρ_2.

- Fidelity is *not* a proper distance since it does not satisfy the triangle inequality. However, fidelity can be used to introduce a proper distance measure in terms of the **angle** between states:

$$\mathcal{A}(\rho_1, \rho_2) = \arccos F(\rho_1, \rho_2).$$

- Monotonicity of fidelity:

$$F\left(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)\right) \geq F\left(\rho_1, \rho_2\right)$$

for any CPTP map $\mathcal{E}$.

- *Application*: Quality Q of ideal quantum gates $|\psi\rangle \mapsto U|\psi\rangle$ for quantum computation, under the factual evolution $|\psi\rangle\langle\psi| \mapsto \mathcal{E}(|\psi\rangle\langle\psi|)$, is naturally measured by the 'worst-case' fidelity:

$$\begin{aligned} Q &\equiv \min_{|\psi\rangle} F\left(U|\psi\rangle\langle\psi|U^\dagger, \mathcal{E}(|\psi\rangle\langle\psi|)\right) \\ &= \min_{|\psi\rangle} \sqrt{\langle\psi|U^\dagger \mathcal{E}(|\psi\rangle\langle\psi|) U|\psi\rangle}. \end{aligned}$$