Master QLMN Antoine Browaeys
Physics of Quantum Information antoine.browaeys@institutoptique.fr

Lecture 4: Entanglement (II)

October 3rd 2025

In Lecture 3, we have studied the basic properties of entangled, i.e. non separable, states. We have in particular seen that entanglement implies abandoning local realism in the sense that the subparts of the system do not have defined properties before a measurement. Moreover, entangled states feature non-local correlations between their subparts. These correlations are stronger than any classical correlations, as demonstrated by the violation of the Bell's inequalities. We have also seen that these quantum correlations could be used as a resource (teleportation, quantum computation, quantum metrology...).

In this lecture, we will address the difficult question of how to quantify entanglement. To see why this question arises, consider the entangled two-qubit states:

$$\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \quad \text{and} \quad \sqrt{1-\epsilon^2}\,|01\rangle + \epsilon\,|10\rangle \ . \tag{1}$$

Intuitively, the second one seems closer to a separable state than the first state when $\epsilon \ll 1$, but can we make this intuition quantitative ? Also, at a fundamental level, a system is generally not in a pure state (see below) but rather described by a density operator: how can we define entanglement in this case ? Although some criteria exist, as we will see, the quantification of entanglement remains an outstanding problem in the general case and a very intense line of research experimentally and theoretically.

1 Density operator description of a subsystem

Consider a bipartite system made of two parts A and B. Each part can be two individual particles or describe a large number of constituents or degrees of freedom. Entangled states of the $(A+B)$ system are states which can not be written as tensor product, i.e.:

$$|\psi_{AB}\rangle \neq |\psi_A\rangle \otimes |\psi_B\rangle \ . \tag{2}$$

This means that we can not assign a state vector to any of the subparts. How can we then describe each subpart ? The answer makes use of the density operator.

1.1 Review of density operator: statistical approach

The traditional introduction to the density operator relies on the imperfect knowledge of the state prepared in an experiment. Assume that each realization of the experiment produces a state $|\psi_k\rangle$ with a probability π_k. The $|\psi_k\rangle$'s do not need to be orthogonal, but are normalized. The average value of an observable $\hat{A}$ over many realizations of the experiment is

$$\langle \hat{A} \rangle = \sum_k \pi_k \langle \psi_k | \hat{A} | \psi_k \rangle \ . \tag{3}$$

Introducing the closure relation involving the basis $\{|u\rangle_n\}$, $\hat{\mathbb{1}} = \sum_n |u_n\rangle\langle u_n|$, yields:

$$\langle \hat{A} \rangle = \sum_n \langle u_n | \left(\sum_k \pi_k |\psi_k\rangle\langle\psi_k| \right) \hat{A} |u_n\rangle = \mathrm{Tr}[\hat{\rho}\hat{A}] \ , \tag{4}$$

with the density operator $\hat{\rho} = \sum_k \pi_k |\psi_k\rangle\langle\psi_k|$. The main properties of the density operator are (check them): (i) $\hat{\rho}^\dagger = \hat{\rho}$ (hermitian) ; (ii) $\mathrm{Tr}[\hat{\rho}] = 1$. Furthermore, we impose (iii) the positivity of $\hat{\rho}$, i.e. that all the eigenvalues are non-negative so that they can be interpreted as probabilities (also called population). An important quantity is the *purity*, defined as $\mathrm{Tr}[\hat{\rho}^2]$: a *pure state* $|\psi\rangle$ has a density operator $\hat{\rho} = |\psi\rangle\langle\psi|$, and $\mathrm{Tr}[\hat{\rho}^2] = 1$ (it is a projector). Any state for which $\mathrm{Tr}[\hat{\rho}^2] < 1$ is called a *mixed state*.

In this *statistical* approach, the density operator is simply a convenient tool to describe the uncertainty in the knowledge of the state of the system. In an experiment aiming at preparing a target state $|\psi_\mathrm{t}\rangle$, the density operator "prepared by the experiment" is:

$$\hat{\rho}_\mathrm{exp} = (1-\epsilon)|\psi_\mathrm{t}\rangle\langle\psi_\mathrm{t}| + \epsilon\hat{\rho}_\mathrm{junk} \ , \tag{5}$$

where $\epsilon \ll 1$ if the experiment is performed correctly. Here $\hat{\rho}_\mathrm{junk}$ acts in the subspace orthogonal to $|\psi_\mathrm{t}\rangle$.

Density matrix for a qubit. Consider first a pure qubit state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. The matrix associated to the density operator is

$$\hat{\rho} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} \ . \tag{6}$$

More generally, for any two-level system, being it in a pure or mixed states, the matrix density has the form

$$\hat{\rho} = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \ . \tag{7}$$

The diagonal coefficients ρ_{00} and ρ_{11} are called *populations*, while the off-diagonal ones $\rho_{01} = \rho_{10}^*$ are named *coherences* for reasons that will be clarified below.

The positivity of $\hat{\rho}$ implies that $|\rho_{01}|^2 \leq \rho_{00}\rho_{11}$, and the equality holds for a pure state. The Bloch vector $\mathbf{u} = \langle\boldsymbol{\sigma}\rangle = \mathrm{Tr}[\hat{\rho}\,\boldsymbol{\sigma}]$ associated to this two-level system is such that

$$\hat{\rho} = \frac{1 + \mathbf{u}\cdot\boldsymbol{\sigma}}{2} \quad \text{with} \quad \mathbf{u} = (2\mathrm{Re}[\rho_{10}], 2\mathrm{Im}[\rho_{10}], \rho_{00} - \rho_{11}) \ . \tag{8}$$

(see Exercice A.1.1). Hence for a mixed state, the Bloch vector lies inside the Bloch sphere of unit radius. The purity is $\mathrm{Tr}[\hat{\rho}^2] = (1 + |\mathbf{u}|^2)/2$.

Pure versus mixed states. We will illustrate this distinction on the case of a qubit. Consider the two density matrices:

$$\hat{\rho}_\mathrm{pure} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} \quad \text{and} \quad \hat{\rho}_\mathrm{mixed} = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} \ . \tag{9}$$

Both describe a situation where the probability to measure $|0\rangle$ is $|\alpha|^2$. However, the mixed states has to be interpreted in the following way: each time you repeat the

experiment the system is either in state $|0\rangle$ or $|1\rangle$, in the same way you get head or tail when tossing a coin. Instead, the pure case corresponds to a situation where, each time, the experiment prepares $\alpha|0\rangle + \beta|1\rangle$. As a an analogy to understand the difference between mixed and pure states, consider single photons prepared either in $|h\rangle$ or $|v\rangle$ respectively, with probabilities $|\alpha|^2$ and $|\beta|^2$. If you place a polarizer at 45° of the vertical direction, the photons are transmitted with a probability $(|\alpha|^2 + |\beta|^2)/2 = 1/2$. If now each photon is prepared in a state $\alpha|h\rangle + \beta|v\rangle$, they are transmitted with a probability $|\alpha + \beta|^2/2 = 1/2 + \mathrm{Re}[\alpha^*\beta]$: the last term is the interference one, related to the coherence $\alpha^*\beta = \rho_{10}$. This ability to yield interferences is the main distinction between pure and mixed states.

1.2 Density operator of a subpart of a system

We come back to the problem of how to describe the subpart of a bipartite system $(A+B)$, characterized by the *pure* state $|\psi_{AB}\rangle$ associate to the density operator $\hat{\rho}_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$. The Hilbert states are $\mathcal{E}_A$ and $\mathcal{E}_B$ (dimensions $d_{A,B}$) with basis $\{|u_n\rangle\}$ and $\{|v_n\rangle\}$. The average value of an observable $\hat{A} \otimes \hat{\mathbb{1}}_B$ acting on the subpart A is

$$\begin{aligned}
\langle \hat{A} \rangle = \mathrm{Tr}[\hat{\rho}_{AB}\hat{A} \otimes \hat{\mathbb{1}}_B] &= \sum_{m=1}^{d_A}\sum_{n=1}^{d_B} \langle u_m| \langle v_n| \hat{\rho}_{AB}\hat{A} \otimes \hat{\mathbb{1}}_B |u_m\rangle |v_n\rangle & (10)\\
&= \sum_{m=1}^{d_A} \langle u_m| \left[\sum_{n=1}^{d_B} \langle v_n| \hat{\rho}_{AB} |v_n\rangle\right] \hat{A} |u_m\rangle & (11)\\
&= \mathrm{Tr}[\hat{\rho}_A \hat{A}] \ , & (12)
\end{aligned}$$

with $\hat{\rho}_A = \mathrm{Tr}_B[\hat{\rho}_{AB}]$. This operator, acting in the Hilbert space $\mathcal{E}_A$ associated to A, results from the *partial trace* over the states associated to B. It fulfills the properties (i)-(iii) of a density operator, and allows one to calculate the properties of the subsystem A. It is called the *reduced density operator*. Hence the density operator introduced here as a partial trace is not *statistical*, i.e. resulting from an imperfect knowledge over the system: on the contrary, it contains *all* the information necessary to describe the *sub*-system A, in the same way the state vector of a *isolated* system fully characterizes it. We now understand why this is the quantity of choice when describing open quantum systems, i.e. a situation where a quantum system is coupled to another external system called a reservoir, a bath, a measuring device...: the system and the reservoir can be entangled and thus the system must be described by a reduced density operator.

Examples of reduced density operators. Take the four Bell states $|\psi\rangle$ involving two qubits A and B. Calling $\hat{\rho} = |\psi\rangle\langle\psi|$, the density operators of each of the qubits are

$$\hat{\rho}_A = {}_B\langle 0|\hat{\rho}|0\rangle_B + {}_B\langle 1|\hat{\rho}|1\rangle_B = \frac{1}{2}|0\rangle\langle 0|_A + \frac{1}{2}|1\rangle\langle 1|_A = \frac{1}{2}\hat{\mathbb{1}}_A \ , \qquad (13)$$

with a similar expression for $\hat{\rho}_B$. The purities are $\mathrm{Tr}[\hat{\rho}_A^2] = \mathrm{Tr}[\hat{\rho}_B^2] = 1/2$. Hence for the maximally entangled Bell states (in a sense that we will make quantitative below), the density operator is maximally mixed, i.e. the matrix is proportional to the identity operator. This gives a criteria to decide whether a bipartite state features entanglement:

trace over the states of one of the subsystem; if the reduced density operator is maximally mixed (proportional to $\hat{\mathbb{1}}$), the state is maximally entangled.

To see that this indeed is plausible, consider now the entangled state

$$|\psi\rangle = \sqrt{1-\epsilon^2}\,|01\rangle + \epsilon\,|10\rangle \ , \tag{14}$$

which intuitively should not be as entangled as the Bell's states when $\epsilon \ll 1$. Show that the reduced density operators are

$$\hat{\rho}_A = (1-\epsilon^2)\,|0\rangle\langle 0| + \epsilon^2\,|1\rangle\langle 1| \quad \text{and} \quad \hat{\rho}_B = \epsilon^2\,|0\rangle\langle 0| + (1-\epsilon^2)\,|1\rangle\langle 1| \ . \tag{15}$$

As one of the coefficient of the matrix is closer to 1 than the other, it means that we have more information on the subsystems than for the Bell's states.

2 State tomography

We now describe how to measure the density operator experimentally. The procedure for measuring a quantum state is called *state tomography*.

Tomography of a pure qubit. Consider first a qubit state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. Performing the tomography of the state means measuring $|\alpha|$ (from which $|\beta|$ follows as $|\alpha|^2 + |\beta|^2 = 1$), and $\arg[\alpha] - \arg[\beta]$. In terms of Bloch vector, this is equivalent to measuring the θ and φ angles. Start by measuring in the z-basis, *i.e.* in the $\{|0\rangle, |1\rangle\}$ basis. Then: $p_0 = |\langle 0|\psi\rangle|^2 = |\alpha|^2$, and $p_1 = |\langle 1|\psi\rangle|^2 = |\beta|^2$. In practice, you repeat N times the same experiment with the system prepared in $|\psi\rangle$ (you could also measure just once on an ensemble containing N systems if you cannot prepare individual systems). Measure the number of times $N_{0,1}$ you obtain 0 or 1. Then $N_0/N \to p_0$ and $N_1/N \to p_1$ when $N \gg 1$.

Now, to get $\arg[\alpha] - \arg[\beta]$, one has to measure in a different basis, for example $|\pm\rangle_x = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Indeed, $|{}_x\langle \pm|\psi\rangle|^2 = \frac{1}{2} \pm \mathrm{Re}[\alpha^*\beta]$, and measuring, for example, the probability $p_{0,x}$ allows one to measure $\cos(\arg[\alpha^*\beta])$.

In the case of a Stern and Gerlach experiment, measuring in the x-basis means rotating the magnet such that the gradient of magnetic field is aligned along x. It is however not always possible to "rotate the measuring apparatus". Take for example a two-level atom: we discussed in Lecture 2 that one can measure if an atom is in state $|0\rangle$ or $|1\rangle$ by scattering light on it. If the atoms scatters, it means that the atom is, say, in $|0\rangle$, while if is does not it is in state $|1\rangle$. Hence, this way of distinguishing the two states allows you to measure only in the z basis. If you cannot rotate the measuring apparatus, you may be able to rotate the state of the system: $|+\rangle_x = \mathcal{R}_y(\frac{\pi}{2})\,|0\rangle$ (see notation Lecture 1 eq. 19). As ${}_x\langle +|\psi\rangle = \langle 0|\,\mathcal{R}_y(-\frac{\pi}{2})\,|\psi\rangle$, one could first rotate the system, and then measure in the z-basis. This gives of course yields the same result: $|{}_x\langle \pm|\psi\rangle|^2 = |\alpha + \beta|^2/2$. In practice such a rotation can be performed using microwaves or light coupling the two states, as seen in Lecture 2.

Tomography of a general density matrix for a qubit. As shown above the density operator is $\hat{\rho} = (1 + \mathbf{u}\cdot\boldsymbol{\sigma})/2$, with $\mathbf{u} = (\mathrm{Re}[\rho_{10}], \mathrm{Im}[\rho_{10}], \rho_{00} - \rho_{11})$. Now, for a *mixed*

state, contrarily to the pure state case, we need to measure three quantities, which are the three components of the non-unit Bloch vector:

$$\hat{\rho} = \frac{1}{2}(1 + u_x\hat{\sigma}_x + u_y\hat{\sigma}_y + u_z\hat{\sigma}_z) \ . \tag{16}$$

The component u_z is directly related to the populations in the $\{|0\rangle, |1\rangle\}$ basis. Hence, repeat the experiment on N identical systems and count $N_{0,1}$. Then $N_0/N \to \rho_{00}$ and $N_1/N \to \rho_{11}$ when $N \gg 1$, thus yielding u_z. To measure u_x and u_y, one needs to measure in the basis where the σ_x or σ_y are diagonal. For example, ${}_x\langle +|\hat{\rho}|+\rangle_x = 1/2 + \mathrm{Re}[\rho_{01}]$. As above, it may not be possible to rotate the apparatus, and one could rotate the state instead: ${}_x\langle +|\hat{\rho}|+\rangle_x = \langle 0|\, \mathcal{R}_y(-\frac{\pi}{2})\hat{\rho}\mathcal{R}_y(\frac{\pi}{2})\, |0\rangle$. One then measures the observable $\hat{\rho}' = \mathcal{R}_y^{-1}\hat{\rho}\mathcal{R}_y$ in the z-basis. Remember (or check using Eq.(18) in Lecture 1) that $\mathcal{R}_y^{-1}\sigma_x\mathcal{R}_y = \sigma_z$, $\mathcal{R}_y^{-1}\sigma_y\mathcal{R}_y = \sigma_y$ and $\mathcal{R}_y^{-1}\sigma_z\mathcal{R}_y = -\sigma_x$, thus

$$\hat{\rho}' = \frac{1}{2}(1 + u_x\hat{\sigma}_z + u_y\hat{\sigma}_y - u_z\hat{\sigma}_x) \ . \tag{17}$$

The u_x coefficient is now on the diagonal in the z-basis, and can be read out. The idea of this rotation is always the same: bring the off-diagonal terms on the diagonal in the basis you can measure, or said differently, transform coherences into populations.

Tomography of a two-qubit density matrix. Generalizing the approach above, a two-qubit density operator $\hat{\rho}$ can be decomposed on $\sigma_i \otimes \sigma_j$, with $i, j = x, y, z$:

$$\hat{\rho} = \sum_{i=0,x,y,z} \sum_{j=0,x,y,z} \lambda_{ij}\, \hat{\sigma}_i \otimes \hat{\sigma}_j \quad \text{with} \quad \lambda_{ij} = \mathrm{Tr}[\hat{\rho}\, \hat{\sigma}_i \otimes \hat{\sigma}_j]/4 = \langle \hat{\sigma}_i \otimes \hat{\sigma}_j \rangle/4 \ . \tag{18}$$

As for the single qubit case, usually, one can only measure terms on the diagonal in the $\mathcal{B} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ basis. One thus applies a rotation prior to the measurement in this basis to bring a particular off-diagonal element onto the diagonal. For example, to measure the coefficient $\lambda_{x,y}$, apply a $\pi/2$-rotation around y on the first qubit and a $\pi/2$-rotation around x to transform the $\sigma_x \otimes \sigma_y$ into a $\sigma_z \otimes \sigma_z$ operator diagonal in $\mathcal{B}$. An example of experimental results on a pair of ions is shown in Fig. 1.

Tomography of a general density matrix for N qubits. Extending the state tomography beyond $N = 8 - 10$ qubits is extremely difficult, if not impossible. Indeed the number of matrix elements one has to measure is $N_\mathrm{m} \sim (2^N)^2/$ For $N = 8$ (the record so far), $N_\mathrm{m} \approx 65500$. But remember that to measure a population or coherence with a good precision, one needs to repeat typically ~ 1000 times the same experiment, preparing the same state as well as possible. If the experiment operates at about 100 Hz (already a high frequency), the full tomography requires ~ 200 hours.... For $N = 9$, this would be 800 hours ! Full state tomography is thus restricted to small quantum systems.

3 Schmidt decomposition

We now derive a general entanglement criteria valid for bipartite systems in a pure state. It will be useful to define the entanglement entropies in the next section.

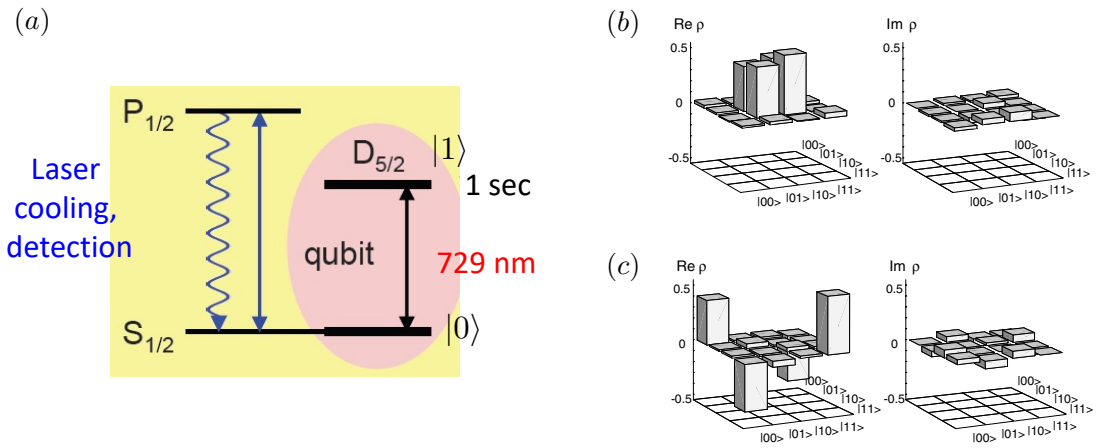

Figure 1: *(a) State tomography on two Ca^+ ions encoding two qubit states $|0\rangle$ and $|1\rangle$ separated by a transition at 729 nm. (b) Results for the state: $(|01\rangle + |10\rangle)/\sqrt{2}$. (c) State: $(|00\rangle - |11\rangle)/\sqrt{2}$. From C.F. Roos et al., Phys. Rev. Lett.* **92**, *220402 (2004).*

Consider the general decomposition of a state $|\psi_{AB}\rangle$ of a bipartite system $A + B$:

$$|\psi_{AB}\rangle = \sum_{n=1}^{\dim\mathcal{E}_A} \sum_{m=1}^{\dim\mathcal{E}_B} c_{nm} |u_n\rangle \otimes |v_m\rangle \ , \tag{19}$$

with $\{|u_n\rangle\}$ and $\{|v_n\rangle\}$ two orthogonal bases of $\mathcal{E}_A$ and $\mathcal{E}_B$. The Schmidt decomposition states that you can find for any *bipartite* system in a *pure* state a decomposition:

$$|\psi_{AB}\rangle = \sum_{n=1}^{r} \sqrt{\lambda_n} |u_n\rangle \otimes |\chi_n\rangle \ , \tag{20}$$

with $\{|u_n\rangle\}$ and $\{|\chi_n\rangle\}$ two *orthogonal* basis of $\mathcal{E}_A$ and $\mathcal{E}_B$, and $r \leq \dim\mathcal{E}_A$ the Schmidt rank. The important difference between the decompositions (19) and (20) is the fact that the sum runs over *only one* index n. This immediately gives a criteria for entanglement: if $r = 1$, the state $|\psi_{AB}\rangle$ is separable, and if $r > 1$, it is entangled. At this stage this is nothing more than a binary criteria stating if the state is entangled or not.

To demonstrate the Schmidt decomposition, take for $\{|u_n\rangle\}$ a basis in which the reduced density matrix of A is diagonal:

$$\hat{\rho}_A = \sum_{n=1}^{\dim\mathcal{E}_A} \lambda_n |u_n\rangle \langle u_n| \ . \tag{21}$$

Rewriting the decomposition (19), we get

$$|\psi_{AB}\rangle = \sum_{n=1}^{\dim\mathcal{E}_A} |u_n\rangle |\phi_n\rangle \quad \text{with} \quad |\phi_n\rangle = \sum_{m=1}^{\dim\mathcal{E}_B} c_{nm} |v_m\rangle \ . \tag{22}$$

The states $|\phi_n\rangle$ are not normalized and not necessarily orthogonal with each other at this stage. The density matrix of the $A + B$ system is

$$\hat{\rho} = |\psi_{AB}\rangle \langle\psi_{AB}| = \sum_n \sum_{n'} |u_n\rangle \langle u_{n'}| \otimes |\phi_n\rangle \langle\phi_{n'}| \ . \tag{23}$$

The reduced density matrix of A is thus:

$$\hat{\rho}_A = \mathrm{Tr}_B[\hat{\rho}] = \sum_{m=1}^{\dim \mathcal{E}_B} \langle v_m | \left[\sum_n \sum_{n'} |u_n\rangle \langle u_{n'}| \otimes |\phi_n\rangle \langle \phi_{n'}| \right] |v_m\rangle \quad (24)$$

$$= \sum_n \sum_{n'} |u_n\rangle \langle u_{n'}| \langle \phi_n | \phi_{n'} \rangle \ , \quad (25)$$

having used $\hat{\mathbb{1}} = \sum_m |v_m\rangle\langle v_m|$. Comparing to Eq. (21), we get $\langle \phi_n | \phi_{n'} \rangle = \lambda_n \delta_{nn'}$, meaning that the $|\phi_n\rangle$'s are in fact orthogonal. Introducing the normalized states $|\chi_n\rangle = |\phi_n\rangle / \sqrt{\lambda_n}$ leads to the (not unique !) Schmidt decomposition.

The Schmidt decomposition has several consequences. Firstly the reduced density operators of the subparts A and B have the same form (same eigenvalues, but act in different Hilbert space):

$$\hat{\rho}_A = \sum_{n=1}^{r} \lambda_n |u_n\rangle \langle u_n| \quad \text{and} \quad \hat{\rho}_B = \sum_{n=1}^{r} \lambda_n |\chi_n\rangle \langle \chi_n| \ . \quad (26)$$

Their purity is thus the same. Besides, Eq. (26) gives us the method to calculate the Schmidt decomposition of a bipartite state: trace over one of the subpart to get $\hat{\rho}_A$ or $\hat{\rho}_B$, expressed in whatever convenient basis, and diagonalize the reduced density matrices to get the λ_n's and the eigenstates $|u_n\rangle$ and $|\chi_n\rangle$. Secondly, any local rotation, *i.e.* a unitary operation acting on each subsystem independently, conserves the decomposition: it can therefore neither entangle nor disentangle a state. To check this, assume the Hamiltonian of $A+B$ is $H = H_A + H_B$, with $[H_A, H_B] = 0$. The associated evolution operator is $U(t) = U_A(t) \otimes U_B(t)$, with $U_{A,B}(t) = \exp[-iH_{A,B}t/\hbar]$. Hence, starting from a Schmidt form $|\psi_{AB}\rangle = \sum_{n=1}^{r} \sqrt{\lambda_n} |u_n\rangle \otimes |\chi_n\rangle$, one gets $U(t)|\psi_{AB}\rangle = \sum_{n=1}^{r} \sqrt{\lambda_n} [U_A(t)|u_n\rangle] \otimes [U_B(t)|\chi_n\rangle]$. The decomposition is thus preserved. So, once again, to change the entanglement (create or destroy), the Hamiltonian must contain a term H_{AB} acting on the two parts of the systems at the same time. Examples of such hamiltonians are:

1. the spin orbit coupling: $H_{AB} = A\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$, or the Ising Hamiltonian $H_{AB} = J\hat{S}_A^z \hat{S}_B^z$;
2. The Jaynes-Cummings Hamiltonian: $H_{AB} = g(\hat{\sigma}_+ \hat{a} + \hat{\sigma}_- \hat{a}^+)$.

Lastly, any mixed state described by the density operator $\hat{\rho}_A = \sum_{n=1}^{r} \lambda_n |u_n\rangle \langle u_n|$ can be "purified" by considering it as the partial trace of a pure state $|\psi_{AB}\rangle = \sum_{n=1}^{r} \sqrt{\lambda_n} |u_n\rangle \otimes |\chi_n\rangle$ of larger system.

Schmidt decomposition and many-body physics. As soon as quantum particles interact, entanglement between them is created and the question naturally arises as how much entanglement. It turns out that there exist general statements about the amount of entanglement in ensembles of quantum particles. To illustrate this, let us consider the Transverse Field Ising model (TFI): it describes an ensemble of spin 1/2 particles with an interaction of the form $\hat{\sigma}_i^z \hat{\sigma}_j^z$ between pairs (i,j), placed in a magnetic field B perpendicular to the z-axis. For a chain of N spins with nearest neighbor interactions, the Hamiltonian is:

$$H = -J \sum_n \hat{\sigma}_n^z \hat{\sigma}_{n+1}^z + B\hat{\sigma}_n^x \quad (J > 0, B > 0) \ . \quad (27)$$

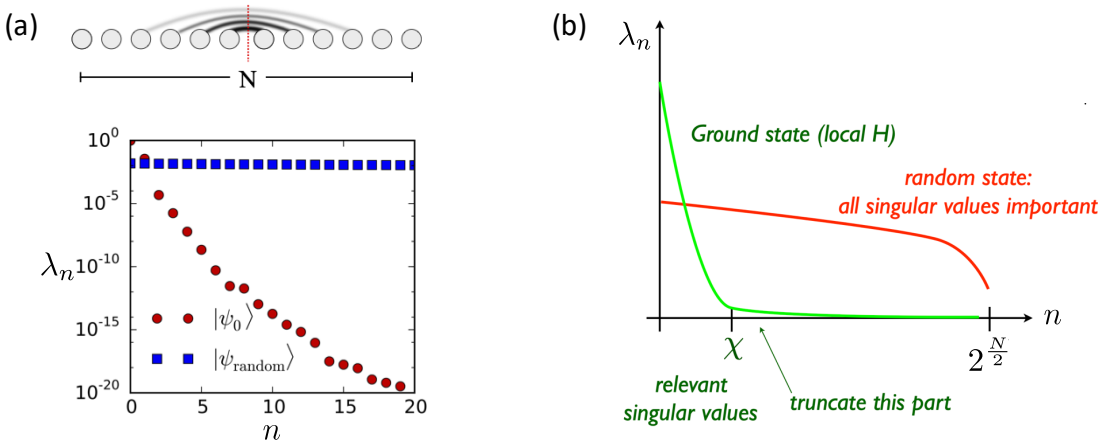

Figure 2: *(a) A chain of N spins prepared in a state $|\psi\rangle$ is split in two equal parts and one calculates the Schmidt numbers λ_n. For the ground state $|\psi_0\rangle$, only a few have significant values, while for any random excited states they all contribute. (b) Generic figure showing that the ground state of a local Hamiltonian features non-negligible Schmidt numbers up to a bound dimension χ. Figures adapted from [https://scipost.org/SciPostPhysLectNotes.5]*

This model is used as an idealization of quantum magnets in condensed matter. It is easy to calculate the ground state in limiting cases: if $B \gg J$, the $\hat{\sigma}^x$ term dominates and the ground state is $(|\uparrow\rangle + |\downarrow\rangle)^{\otimes N}$; if $J \gg B$, the interaction term is minimized by the ferromagnetic states $|\uparrow,\uparrow,\uparrow ...\rangle$ and $|\downarrow,\downarrow,\downarrow ...\rangle$. In the general case $B \sim J$, the non-commutation between the two parts of the Hamiltonian, *i.e.* $[\hat{\sigma}^z_n \hat{\sigma}^z_{n+1}, \hat{\sigma}^x_n] \neq 0$, makes it impossible to diagonalize them independently and one needs to resort to a full diagonalization, which becomes challenging beyond a few tens of spins.

It was realized in the early 2000's that the ground-states of many-body systems with local interactions are actually quite special in that they do not feature a lot of entanglement. This can be quantified mode precisely using the Schmidt decomposition: split the chain into two parts A and B containing $N/2$ spins each and apply the decomposition to the ground-state:

$$|\psi_{\text{gs}}\rangle = \sum_{n=1}^{r} \sqrt{\lambda_n}\, |u_n\rangle_A\, |\chi_n\rangle_B \ . \tag{28}$$

A priori, the rank is such that $r \lesssim N/2$. However it is much better than that: only a few of the Schmidt numbers λ_n are non negligible. This is illustrated in Fig. 2(a) where we show the λ_n's for the ground state of the TFI model for a chain of 16 atoms and $B = 1.5J$. By contrast, had we taken randomly any excited states, we would have found that all the λ_n's have nearly the same value. It turns out that this is a general property of the ground states of local Hamiltonians (the demonstration is very hard and way beyond what we discuss in this course), as shown in Fig. 2(b). The upper bound in the sum is called the bound dimension χ: it represents the largest Schmidt number that contribute significantly to the sum at a given level of approximation ϵ defined such

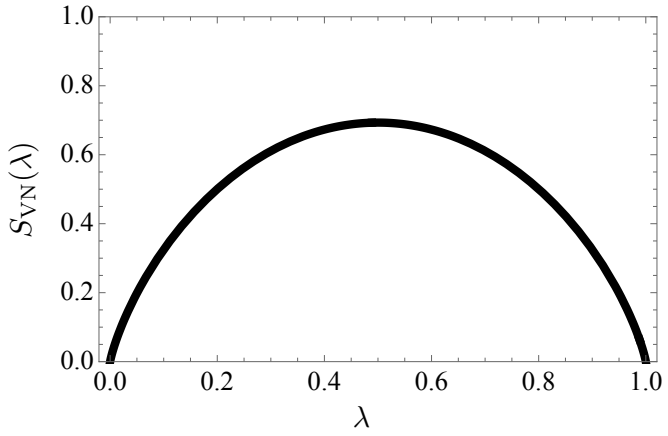

Figure 3: *Von Neumann entropy for a system of two qubits.*

that:

$$\| |\psi_{\rm gs}\rangle - \sum_{n=1}^{\chi} \sqrt{\lambda_n}\, |u_n\rangle_A\, |\chi_n\rangle_B \| < \epsilon \ . \tag{29}$$

This is a considerable simplification for numerical simulations: instead of having to diagonalize matrices of size $2^N \times 2^N$ to find the ground state, it is enough to handle matrices of dimension $\chi \times \chi$ (see also at the end of next section). This remark is at the heart of modern numerical methods used to calculate the properties of many-body systems, such as tensor networks and Matrix-Product states.

4 Entanglement entropies and their measurement

We have seen in Section 1.2 two examples of entangled states that do not have the same reduced density matrix: the Bell state (13) features $\hat{\rho}_A = \hat{\mathbb{1}}/2$, while the state (14) has $\hat{\rho}_A = (1-\epsilon^2)\,|0\rangle\langle 0| + \epsilon^2\,|1\rangle\langle 1|$. The Bell state is the one for which we have the smallest amount of information on each subpart: the two states $|0\rangle$ and $|1\rangle$ have exactly the same probability, contrarily to the second one. This information content can be quantified by introducing *entanglement entropies*.

Remember first that in statistical physics, one introduces the Shannon entropy as:

$$S_{\rm Shannon} = -\sum_l p_l \ln p_l \ , \tag{30}$$

where the p_l's are the probability to find a system in the micro-state l. One fundamental assumption of statistical physics is the fact that for an *isolated* system, all the micro-states realizing a given macrostate have the same probability $p_l = 1/\Omega$, with Ω the total number of microstates that lead to this macrostate. This is equivalent to saying that the entropy of the isolated system is maximal under the constraint $\sum_l p_l = 1$. The fact that all microstates have the same probability indicates that we have the least amount of information possible about the microstate the system is in.

Von Neumann entropy. The extension of this idea in quantum physics leads to the Von Neumann entropy for a system described by a density operator $\hat{\rho}$:

$$S_{\mathrm{VN}}(\hat{\rho}) = -\mathrm{Tr}[\hat{\rho}\ln\hat{\rho}] \ . \tag{31}$$

For a bipartite system in a pure state, the reduced density operators of A and B have the same form thanks to the Schmidt decomposition (see Eq. 26). Thus

$$S_{\mathrm{VN}}(\hat{\rho}_A) = S_{\mathrm{VN}}(\hat{\rho}_B) = -\mathrm{Tr}[\hat{\rho}_A\ln\hat{\rho}_A] = -\sum_{n=1}^{r}\lambda_n\ln\lambda_n \ , \tag{32}$$

which is nothing but the Shannon entropy with $p_l = \lambda_l$, the eigenvalues of the reduced density matrices. We then find that for a subsystem in a pure state, *i.e.* $|\psi_{AB}\rangle$ separable, $r = 1$, and $\lambda = 1$, $S_{\mathrm{VN}}(\hat{\rho}_A) = 0$: the von Neumann entropy of a pure state is 0, which indicates again that we have perfect knowledge of the state the system is in.

Let us consider the case of two entangled qubits. The Schmidt decomposition applies with two eigenvalues λ_1 and $\lambda_2 = 1 - \lambda_1$ (remember that $\mathrm{Tr}[\hat{\rho}_A] = \lambda_1 + \lambda_2 = 1$). The evolution of $S_{\mathrm{VN}}(\lambda_1)$ is shown in Fig. 3. We do observe that the entropy is maximum for $\lambda_1 = 1/2$, and is 0 for $\lambda_1 = 0, 1$. Hence we now have a quantitative criteria to measure the degree of entanglement for a bipartite system: the larger the entanglement entropy, the more entangled the state is (for this particular bipartition).

Rényi entropies. Experimentally however, the von Neumann entropy is not a convenient quantity: we have seen above that the experimental reconstruction of the density matrix $\hat{\rho}_{\mathrm{exp}}$ of a system of more than $\sim 8-10$ particles (or subparts) is too challenging. We can therefore not calculate S_{VN} from $\hat{\rho}_{\mathrm{exp}}$, and besides no one has so far been able to devise an experiment to measure directly S_{VN} without the precise knowledge of $\hat{\rho}_{\mathrm{exp}}$. It turns out however that one can define other entropies that are experimentally measurable. Those are the Rényi entropies defined as:

$$S_{\mathrm{R}}^{(\alpha)}(\hat{\rho}) = \frac{1}{1-\alpha}\ln(\mathrm{Tr}[\hat{\rho}^{\alpha}]) \quad \text{with} \quad \alpha \geq 1 \ . \tag{33}$$

One can prove that $S_{\mathrm{R}}^{(\alpha)}(\hat{\rho}) \to S_{\mathrm{VN}}(\hat{\rho})$ for $\alpha \to 1$ (see exercise A.2). The most interesting one experimentally is the second-order Rényi entropy $S_{\mathrm{R}}^{(2)}(\hat{\rho})$, which is directly related to the purity of $\hat{\rho}$, $\mathrm{Tr}[\hat{\rho}^2]$.

By now, several experiments have measured Rényi entropies using either atoms or superconducting qubits. They used mainly two types of methods which we illustrate below. A first experiment by R. Islam *et al.* [Nature **528**, 77 (2015)] explored the entanglement between 4 atoms placed in an optical lattices. When the lattice is very deep the atoms are well isolated in each site, and independent from each other, hence should not be entangled. On the contrary, when the lattice amplitude is weak, the atoms tunnel between sites and become entangled: this is not obvious and we take this as a fact. The idea to demonstrate the presence of entanglement is the following: divide the system in two parts A and B and measure the purity $\mathrm{Tr}[\hat{\rho}_A^2]$. If entanglement is present in the system, we should have $\mathrm{Tr}[\hat{\rho}_A^2] < \mathrm{Tr}[\hat{\rho}_{AB}^2]$, or equivalently $S_{\mathrm{R}}^{(2)}(\hat{\rho}_A) > S_{\mathrm{R}}^{(2)}(\hat{\rho}_{AB})$. The measurement of the purity is done in the following way: prepare two copies 1 and 2 of the

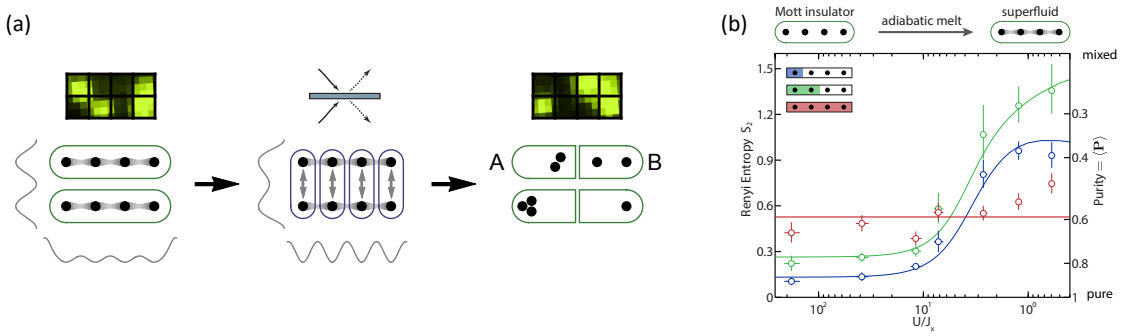

Figure 4: *Measurement of the purity of 4 atoms in an optical lattice. (a) One prepares two copies of the system and perform a swap operation akin to a two-photon interference, followed by the measurement of the parity of the number of atoms on each side. (b) Results of the experiment for different partitions of the 4 atoms. From [Nature* **528**, *77 (2015)].*

system, described by density operators $\hat{\rho}_{1,2}$. Weakening the potential barrier between the two copies allows the tunneling between sites (see Fig. 4a). A two-atom interference akin to the Hong-Ou-Mandel effect between photons occurs, that swaps the atoms between copy 1 and 2. It then turns out (and this is far from being obvious) that the average parity in the number of atoms in each well after this is directly $\mathrm{Tr}[\hat{\rho}_1\hat{\rho}_2] = \mathrm{Tr}[\hat{\rho}_1^2]$ when the two copies are identical. Figure 4(b) shows the result of the experiment: the Rényi entropy of a sub-part increases when the atoms are more delocalized along the 4 sites, as expected.

The second method demonstrated for the first time by T. Brydges *et al.* [Science **364**, 260 (2019)] does not require two copies of the same system. The idea is to apply on the system of N qubits a set of random unitary operations (i.e. rotations on the Bloch sphere) $u_1 \otimes u_2 \otimes ... \otimes u_N$. For a given set of unitaries, you repeat the experiment many times to get the statistics of the outcomes of the experiment, and then repeat for a new set of unitaries. It turns out (and this is very hard to show) that the correlations between results is related to the purity of the density matrix. To see that it is plausible, consider a single qubit, apply random rotation followed by a measurement in the z-basis to obtain $\langle \sigma_z \rangle$. For a pure state the Bloch vector following any random rotation is always on the Bloch sphere and the distribution of results is uniform between $[-1, 1]$ (Fig. 5a). If the state is mixed, the vector has a length smaller than 1, resulting in a narrower distribution of the results. Thus, measuring the width of the distribution tells about the purity $\mathrm{Tr}[\hat{\rho}^2]$. The experiment was performed using a chain of up to 20 entangled ions, each encoding a qubit. The authors measured the purity of a sub-system of N_A ions. An example of results for $N = 10$ ions is shown in Fig. 5(b): the maximum entropy (hence lowest purity) is obtained for $N_A = N/2$.

Entropy and Area Laws. We have seen in Sec. 3 that the Schmidt rank of the ground state of a many-body system of N spins-1/2 is usually bounded by $\chi \ll 2^{N/2}$. On the contrary, for most excited states taken randomly, $\chi \sim 2^{N/2}$. Hence, in the ground

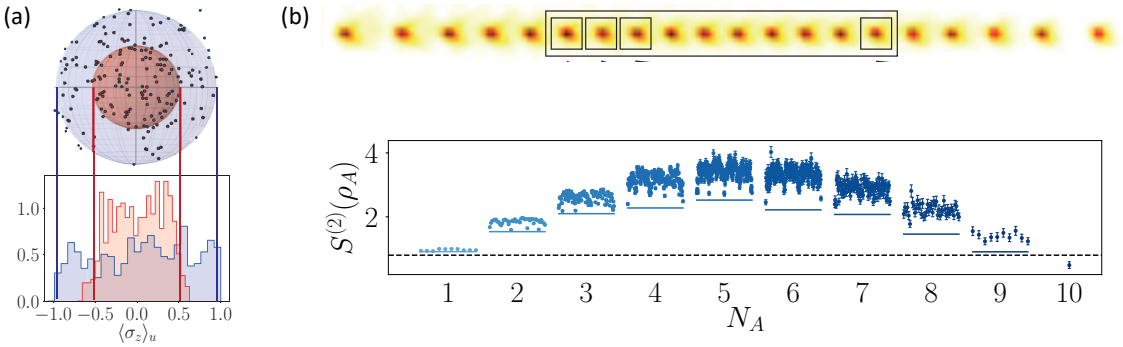

Figure 5: *Measurement of the purity by application of random unitaries. (a) Example on a single pure or mixed qubit. (b) Results for 10 ions. From [Science* **364***, 260 (2019)].*

state, the entanglement entropy is such that:

$$S_A \approx -\sum_{n=1}^{\chi} \lambda_n \ln \lambda_n \lesssim \frac{1}{\chi} \ln \chi \times \chi = \ln \chi \ , \tag{34}$$

assuming $\lambda_n \sim 1/\chi$. Hence the entanglement entropy is related to the bound dimension by $S_A \sim \ln \chi$. In thermodynamics, we are used to the entropy being extensive, *i.e.* proportional to the volume of the system or the number of particles N. Things are more subtle with entanglement entropies that can be non extensive. In general, for any random states but the ground state, S_A fulfills volume laws: $S_A \sim N/2 \ln 2$ for N qubits (bipartition). However, in the ground state it obeys area laws: S_A is proportional to the "area" between the two sub-parts A and B of the system. For example, in a chain the "area" of the separation between the two subpart is 1. In a 2-dimensional square array, it would be the perimeter of the boundary between the two subparts, hence L (see Fig. 6). This is the recognition that most of the entanglement is located at the boundary between the two sub-systems. Numerically, again, this is very useful: it means that in 2D, $S_A = \alpha L \sim \alpha\sqrt{N}$, hence the bound dimension is $\chi \sim e^{\alpha\sqrt{N}}$. This is the size of the matrix one has to diagonalize. In 1D it is even better: S_A, and thus χ are constant.

5 Multipartite entanglement and mixed entangled states

So far we have described mainly bipartite entanglement. This concept naturally applies to systems with two degrees of freedom (e.g. spin and orbital angular momentum, two modes of a field, two spins...). Entanglement can be generalized to multipartite cases following the same procedure as for the bipartite case. Take a multipartite system consisting of N subparts $1, 2, 3, ...N$ and described by a pure state $|\psi_N\rangle$. The state is separable if $|\psi_N\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes |\phi_3\rangle \otimes ... \otimes |\phi_N\rangle$. If not separable, it is entangled. A *biseparable* state can be decomposed in two separable parts containing n and $N - n$ subpart: $|\psi_N\rangle = |\phi_n\rangle \otimes |\phi_N - n\rangle$, but each of the subpart can be entangled. A *fully entangled* state is a state which is not biseparable with respect to any bipartition of the system.

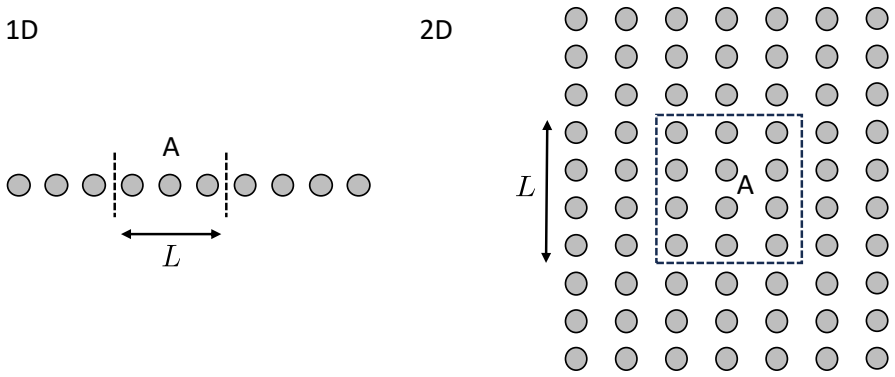

Figure 6: *Illustration of the area laws for the entanglement entropy, in 1 and 2 dimensions.*

As a matter of fact, most systems that we deal with experimentally *are not* pure states, but rather mixed states. This raises the question of how to define entanglement in this case. A bipartite system is separable when, for a realization k that has a probability p_k, $\hat{\rho}_{AB} = \hat{\rho}_A^{(k)} \otimes \hat{\rho}_B^{(k)}$. Summing over all the possible realizations, one gets

$$\text{Separable}: \quad \hat{\rho}_{AB} = \sum_k p_k \, \hat{\rho}_A^{(k)} \otimes \hat{\rho}_B^{(k)} \tag{35}$$

$$\text{Entangled}: \quad \hat{\rho}_{AB} \neq \sum_k p_k \, \hat{\rho}_A^{(k)} \otimes \hat{\rho}_B^{(k)} \tag{36}$$

All these rather abstract definitions (of little practical use) suggest us that defining entanglement is an extremely challenging problem, and as a matter of fact we are still at the beginning of attempting to classify entanglement in multipartite systems.

Examples of multipartite states and classes of entanglement. Here, we revisit briefly some ideas already introduced in Lecture 3, now making use of the density operator. Consider the two states containing N qubits, already seen in Lecture 3:

$$|\mathrm{GHZ}_N\rangle = \frac{1}{\sqrt{2}}(|0,0,0,...\rangle + |1,1,1...\rangle) \text{ (Greenberger-Horne-Zeilinger state)} , \tag{37}$$

$$|\mathrm{W}_N\rangle = \frac{1}{\sqrt{N}}(|1,0,0,...\rangle + |0,1,0,...\rangle + ... + |0,0,...,1\rangle \text{ (Werner state)} . \tag{38}$$

Both are entangled but belong to different entanglement class. The GHZ state has however a major drawback: it is very sensitive to atom losses. Suppose that you lose one of the N atoms of the state. As you don't know in which state is the lost particle, we have to trace over its two possible states to calculate the reduced density operator of the $N-1$ remaining particles. We obtain

$$\hat{\rho}_{N-1}^{\mathrm{GHZ}} = \mathrm{Tr}_{0,1}[\hat{\rho}_N^{\mathrm{GHZ}}] = \frac{1}{2} |0\rangle^{\otimes N-1} \langle 0|^{\otimes N-1} + \frac{1}{2} |1\rangle^{\otimes N-1} \langle 1|^{\otimes N-1} \quad . \tag{39}$$

The operator is diagonal: we have lost the quantum coherences, and the $N-1$ atoms are now in a statistical mixture. The W state is not very useful for metrology, but is

robust with respect to the loss of particles. Indeed (show it...):

$$\hat{\rho}^{\rm W}_{N-1} = {\rm Tr}_{0,1}[\hat{\rho}^{\rm W}_N] = \frac{N-1}{N} \left|W_{N-1}\right\rangle \left\langle W_{N-1}\right| + \frac{1}{N} \left|000...\right\rangle \left\langle 000...\right| \ . \tag{40}$$

The state remains nearly pure for large N.

Experimentally, GHZ states of up to 24 trapped ions have been prepared and characterized [Phys. Rev. X Quantum **2**, 020343 (2021)]. Werner states of up to 200 atoms have been reported [Phys. Rev. X **5**, 031015 (2015)], but the characterization was not direct and relied on assumptions on the system.

6 Entanglement witnesses

As probably obvious from the sections above, if preparing entangled states in the lab is already a challenge, characterizing them is the experimentalist's nightmare. The main problem is to devise experimentally measurable criteria.

Bell test. For two qubits, we have seen that a violation of the Bell's inequalities characterizes entanglement. Hence, if one prepares a two-qubit state, subjects it to a Bell test (i.e. measure the S parameter introduced by John Bell) and obtain a violation, then the state is entangled. However, it is a very strong test: a state can be entangled without violating Bell's inequalities. This would for example be the case if the system prepared in an experiment is described by the density matrix: $\hat{\rho}_{\rm exp} = (1-\epsilon)|\psi_{\rm B}\rangle\langle\psi_{\rm B}| + \epsilon\hat{\rho}_{\rm junk}$, with $|\psi_{\rm B}\rangle$ of the Bell states. This states leads to a violation of Bell's inequality if $\epsilon < 1 - 1/\sqrt{2} \approx 0.3$.

Fidelity. When the number of qubits grows, measuring the correlations or the density matrix is too hard, and one has to resort to more global (but less precise...) criteria. The fidelity is one of them. Assume that you want to prepare a target state $|\psi_{\rm t}\rangle$, and that you actually produce on the experiment a state $|\psi_{\rm exp}\rangle$. The fidelity $\mathcal{F}$ is defined as the square of the overlap between what you prepare and what you actually want to prepare: $\mathcal{F} = |\langle\psi_{\rm exp}|\psi_{\rm t}\rangle|^2$. If what you get in the experiment is a density matrix $\hat{\rho}_{\rm exp}$, then $\mathcal{F} = \langle\psi_{\rm t}|\,\hat{\rho}_{\rm exp}\,|\psi_{\rm t}\rangle$. It turns out that the fidelity is the easiest quantity to measure experimentally (see homework later). A criterion you will derive in the homework states that $\mathcal{F} > 1/2 \Rightarrow$ entanglement. For a two-qubit state to violate Bell's inequality, you need $\mathcal{F} > 1/\sqrt{2} \approx 0.7$. However be careful that $\mathcal{F} \leq 1/2$ does *not* imply the absence of entanglement ! Assume for example you want to prepare the state $|\phi_+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$, and that your experiment generates instead $|\psi_+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$: $\mathcal{F} = 0$ and yet the state prepared is entangled !

Entanglement witness. A witness operator $\hat{\mathcal{W}}$ is defined such that $\langle\hat{\mathcal{W}}\rangle = {\rm Tr}[\hat{\rho}_{\rm exp}\hat{\mathcal{W}}] < 0$ if $\hat{\rho}_{\rm exp}$ is entangled and $\langle\hat{\mathcal{W}}\rangle \geq 0$ if $\hat{\rho}_{\rm exp}$ is separable (the choice of the sign is a convention). To be useful $\langle\hat{\mathcal{W}}\rangle$ has to be measurable in a lab. As a first example, take $\hat{\mathcal{W}} = 1 - 2\left|\psi_{\rm t}\right\rangle\left\langle\psi_{\rm t}\right|$. Then $\langle\hat{\mathcal{W}}\rangle = 1 - 2\mathcal{F}$ is related to the measurement of the fidelity. As another example, consider a chain of spin-1/2 particles for which the hamiltonian $H = -J\sum_n \hat{\mathbf{S}}_n \cdot \hat{\mathbf{S}}_{n+1}$ ($J > 0$). If you suppose that the N pair of atoms in the system are all in a separable state, then $\langle\hat{\mathbf{S}}_n \cdot \hat{\mathbf{S}}_{n+1}\rangle = \langle\hat{\mathbf{S}}_n\rangle \cdot \langle\hat{\mathbf{S}}_{n+1}\rangle$, and $|\langle H\rangle| \leq NJ/4$. If on

the contrary the pairs are, say, in the entangled state $(|01\rangle - |10\rangle)/\sqrt{2}$, then one finds $\langle H\rangle = 3NJ/4$ (use $\hat{\mathbf{S}}_n \cdot \hat{\mathbf{S}}_{n+1} = (\hat{\mathbf{S}}^2 - \hat{\mathbf{S}}_n^2 - \hat{\mathbf{S}}_{n+1}^2)/2$). Thus, if you take as a witness $\langle \hat{\mathcal{W}}\rangle = 1 - 2H/J$, you are able to tell is there is entanglement in the system. The advantage of the witness method lies in the fact that it only requires measuring global observables, like here $\langle H\rangle$.

A Problem set for Lecture 4

A.1 Some derivations...

A.1.1 Density matrix of a qubit in a mixed state.

Take for the density matrix the form of Eq. (7).

1. Give the components of the Bloch vector $\mathbf{u}(\theta, \varphi)$ as a function of the coefficients of $\hat{\rho}$.
2. Show that $|\rho_{01}|^2 \leq \rho_{00}\rho_{11}$. In which case is it an equality ?
3. Show that $|\mathbf{u}| \leq 1$.

A.1.2 Partial traces and reduced density operators

1. Calculate the reduced density operators $\hat{\rho}_\pm$ associated to the qubit states $|\varphi_\pm\rangle = (|00\rangle + |01\rangle + |10\rangle \pm |11\rangle)/2$. Which of the two states is entangled?
2. Calculate the reduced density operators associated to the state $(|00\rangle + |01\rangle + |11\rangle)/\sqrt{3}$.
3. Demonstrate that $\lambda_{ij} = \langle \hat{\sigma}_i \otimes \hat{\sigma}_j\rangle = \mathrm{Tr}[\hat{\rho}\,\hat{\sigma}_i \otimes \hat{\sigma}_j]/4$ in equation (18).
4. Demonstrate equations (39) and (40).

A.2 Entropies

1. Show that $S_{\mathrm{VN}}(\hat{\rho}_A) = -\sum_{n=1}^{r} \lambda_n \ln \lambda_n$ (notations of the lecture).
2. Consider the Rényi entropy. Show that $S_{\mathrm{R}}^{(\alpha)}(\hat{\rho}) = \frac{1}{1-\alpha} \ln[\sum_k p_k^\alpha]$, with p_k the eigenvalues of $\hat{\rho}$, such that $\sum_k p_k = 1$.
3. We want to show that $S_{\mathrm{R}}^{(\alpha)}(\hat{\rho}) \to S_{\mathrm{VN}}(\hat{\rho})$ for $\alpha \to 1$. To do so write, that $\sum_k p_k^\alpha = \sum_k p_k \exp[(\alpha - 1)\ln p_k]$ and Taylor expand the exponential for $\alpha \to 1$ to recover the expression of the Shannon entropy.
4. Use the von Neumann entropy to decide which of the two states is the most entangled: $|\psi_1\rangle = \sqrt{2/7}\,|00\rangle + \sqrt{5/7}\,|11\rangle$ or $|\psi_2\rangle = \sqrt{2/3}\,|01\rangle + \sqrt{1/3}\,|10\rangle$?
5. Same question as before using the second-order Rényi entropy.