Unit 10: Quantum Entangles

Tzu-Chieh Wei
C. N. Yang Institute for Theoretical Physics and Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, NY 11794-3840, USA
(Dated: August 12, 2022)

In this unit, we discuss entanglement of quantum states, entanglement of formation and distillation, entanglement entropy, Schmidt decomposition, majorization, quantum Shannon theory, etc.

Learning outcomes: You'll be able to understand the basics of quantum information and entanglement theory.

I. INTRODUCTION

Quantum entangled states have correlations stronger than classical states. We have seen the use of two-particle singlet states to violate the Bell inequality, strongly supporting quantum mechanics over any local realistic classical theory. In this unit, we will give more quantitative definition for entanglement, both from the resource quantification perspective and from the Hilbert-space geometry one.

II. TWO-QUBIT ENTANGLEMENT

We have seen the following four Bell states, which enable many important quantum information processing tasks,

$$|\Phi^{\pm}\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle),\ |\Psi^{\pm}\rangle \equiv \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle).$$

What is essential is their entanglement. But how do we quantify entanglement for these quantum states? One idea is to use them as the unit, assigning to them the so-called *e-bit* and analyze how many e-bits can be distilled from other entangled states or how many e-bits are needed to create them. This is very intuitive and operational, but seems very challenging.

Another approach is to investigate how much these state differ from unentangled states. It is clear that these Bell states cannot be written as a product form: $|\text{Bell}\rangle \neq |\chi\rangle \otimes |\phi\rangle$. Is this entangled? $|\uparrow\uparrow\rangle - |\uparrow\downarrow\rangle - |\uparrow\downarrow\rangle + |\downarrow\downarrow\rangle$? What about $|\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle - |\uparrow\downarrow\rangle + |\downarrow\downarrow\rangle$?

Inseparability and concurrence. Let us take a two-qubit pure state example,

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle,$$

and seek the condition that it is entangled. If it is not entangled, it can be factorized into

$$(\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle.$$

We can identify the folowing

$$a \sim \alpha\gamma,\ b \sim \alpha\delta,\ c \sim \beta\gamma,\ d \sim \beta\delta,$$

which is equivalent to $ad - bc = 0$. Therefore, two-qubit pure states are not entangled, if and only if $ad - bc \neq 0$. Then, we can define a quantity called *concurrence* C for the two-qubit pure state $|\psi\rangle$,

$$C(\psi) \equiv 2|ad - bc|,$$

where the factor 2 is added such that $C(\text{Bell states}) = 1$, as one can easily verify. This definition is based inseparability, but is related to time-reversal of the state $|\psi\rangle$.

For a spin-1/2 or qubit, the time-reversal operation is defined as $\hat{T} = -i\sigma_y K$, where K denotes the complex conjugate. Thus, the detailed representation of $\hat{T}$ depends on the basis chosen. If we have two qubits, then the time reversal is thus

$$\hat{T}_2 \equiv -\sigma_y \otimes \sigma_y K = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} K.$$

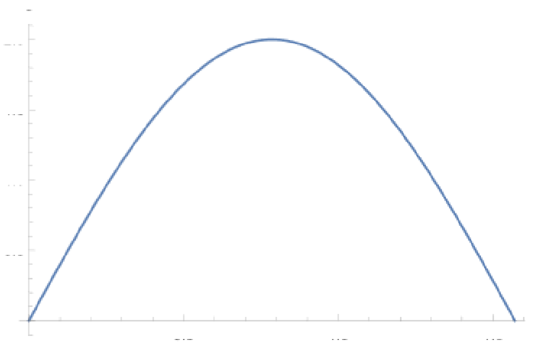

FIG. 1. Concurrence for $|\psi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$.

Thus the action of time-reversal on $|\psi\rangle$ leads to

$$|\tilde{\psi}\rangle = \hat{T}_2|\psi\rangle = d^*|00\rangle - c^*|01\rangle - b^*|10\rangle + a^*|11\rangle.$$

We have that

$$\langle\tilde{\psi}|\psi\rangle = 2(ad - bc),$$

and $C(\psi) = |\langle\tilde{\psi}|\psi\rangle|$, which is given via the overlap between the state and its time reversal partner. We shall see later that concurrence is related to the so-called entanglement of formation.

Example. Let us consider $|\psi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$. It is straightforward to see that $C = 2|\sin\theta\cos\theta| = |\sin(2\theta)|$, which we plot in Fig. 1. It achieves maximum at $\theta = \pi/4$, which is a Bell state.

We will postpone our discussions of mixed-state entanglement to later.

III. BIPARTITE PURE-STATE ENTANGLEMENT VIA SCHDMIT DECOMPOSITION AND ENTANGLEMENT ENTROPY

For any bipartite pure state (bipartite could arise by grouping subsystems into two parts), it can be written in the tensor-product basis $\{|i\rangle \otimes |j\rangle\}$ as

$$|\psi\rangle = \sum_{i=1}^{N_A}\sum_{j=1}^{N_B} \psi_{ij}|i,j\rangle.$$

The coefficient ψ_{ij} is an $N_A \times N_B$ matrix and can be singular-valude decomposed (SVD),

$$\psi_{ij} = \sum_{k=1}^{\max(N_A,N_B)} U_{ik}\sigma_k V^*_{jk}.$$

We can use this to write $|\psi\rangle$ in a more compact way,

$$|\psi\rangle = \sum_k \sigma_k(\sum_i U_{ik}|i\rangle) \otimes (\sum_j V^*_{jk}|j\rangle) = \sum_k \sigma_k|\tilde{k}\rangle_A \otimes |\tilde{k}\rangle_B,$$

which involves at most $\max(N_A, N_B)$ components. Note here that $\tilde{k}\rangle$'s are orthonormal, matrices U and V describe how to transform to this basis, and the singular values σ_k's satisfy that $\sum_k \sigma_k^2 = 1$. This decomposion is called Schmidt decomposition and works generally for bipartite pure states. However, it does not work beyond two qubits. For example, $|W\rangle = (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}$ does not have a Schmidt decomposition. It seems that 3 is the smallest number of components necessary to write down this W state.

Entanglement entropy. Given that the Schmidt coefficients $\{\sigma_k\}$ give rise to a probability distribution $\{\sigma_k^2\}$, we can define an entropy: $S_V = -\sum_k \sigma_k^2 \log_2(\sigma_k^2)$, where the base is usually taken as 2.

Let us consider an example: $|\psi\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |01\rangle + |10\rangle)$. From the previous section, we know that its concurrence $C = 2/3$.

$$\psi = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & 0 \end{pmatrix} = USV^\dagger,\ \mathrm{diag}(S) = \{\sigma_1, \sigma_2\} = \{\sqrt{\frac{3+\sqrt{5}}{6}}, \sqrt{\frac{3-\sqrt{5}}{6}}\},$$

and

$$U \approx \begin{pmatrix} 0.8507 & 0.5257 \\ 0.5257 & -0.8507 \end{pmatrix},\ V \approx \begin{pmatrix} 0.8507 & -0.5257 \\ 0.5257 & 0.8507 \end{pmatrix}.$$

We can also verify that $C = 2\sigma_1\sigma_2 = 2/3$ and that the entanglement entropy $S_V \approx 0.55$.
Partial trace. Once we have the Schmidt decomposition,

$$|\psi\rangle = \sum_k \sigma_k |\tilde{k}\rangle_A \otimes |\tilde{k}\rangle_B,$$

we can trace over the second party and obtain the reduced density matrix for the first party,

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \sum_k \langle\tilde{k}|_B \cdot |\psi\rangle\langle\psi| \cdot |\tilde{k}\rangle_B = \sum_k \sigma_k^2 |\tilde{k}\rangle_A\langle\tilde{k}|_A,$$

which is diagonal. The von Neumann entropy for a density matrix is defined in terms of its egienvalues,

$$S_V(\rho_A) = -\mathrm{Tr}(\rho_A \log \rho_A) = -\sum_k \sigma_k^2 \log(\sigma_k^2),$$

which is the entanglement entropy of $|\psi\rangle$. If ρ_A is pure, then $S_V(\rho_A) = 0$.

IV. BI-PARTITE MIXED STATE ENTANGLEMENT

In the case a bi-partite state ρ_{AB}, how do we determine whether it is entangled? It is generally a difficult problem. But we can at least define whether a state is entangled or not. Let us first define when a state is *not* entangled, which is sometimes called *separable*,

$$\rho_{\text{unentangled}} = \sum_i p_i \rho_i^A \otimes \rho_i^B,$$

i.e., if it can be decomposed into a probability distribution of unentangled states. Note that the above definition is equivalent to

$$\rho_{\text{unentangled}} = \sum_k q_k |\phi_i^A\rangle\langle\phi_i^A| \otimes |\phi_i^B\rangle\langle\phi_i^B|. \tag{1}$$

For example,

$$\frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$$

is unentangled. How about

$$\frac{1}{2}|\Phi^+\rangle\langle\Phi^+| + \frac{1}{2}|\Psi^+\rangle\langle\Psi^+|?$$

where $|\Phi^\pm\rangle$ and $|\Psi^\pm\rangle$ are Bell states. Note that unentangled states can be prepared between two parties, Alice and Bob, that are far apart. If they can communicate with each other, they can create e.g.

$$\rho_{\text{unentangled}} = \sum_k q_k |\phi_i^A\rangle\langle\phi_i^A| \otimes |\phi_i^B\rangle\langle\phi_i^B|,$$

by coordinating what states they create depending on some random number.

Once we have defined unentangled states, entangled states are thus those that are not unentangled. Let us see how a mixed state can arise. Consider two qubits interacting via the Heisenberg interaction,

$$\hat{H}_{12} = X_1X_2 + Y_1Y_2 + Z_1Z_2.$$

You will diagonalize this Hamiltonian and work out the four eigenstates and their eigenvalues. At a temperature T, the system is in a mixed state (the specific form below is usually called the Werner state [1])

$$\rho_{12} = \frac{e^{-\hat{H}_{12}/T}}{\mathrm{Tr}e^{-\hat{H}_{12}/T}} = r(T)|\Psi^-\rangle\langle\Psi^-| + \frac{1-r(T)}{4}I_{4\times4},$$

where

$$r(T) = \frac{e^{3/T} - e^{-1/T}}{e^{3/T} + 3e^{-1/T}}.$$

How do we know when this state is entangled? This will be an exercise for you.

V. POSITIVE-PARTIAL TRANSPOSE AND PERES-HORODECKI CRITERION FOR SEPARABILITY

If a bi-partite state cannot be written as a convex sum of direct products of density matrices then it is entangled. However, this definition does not offer a practical way of determining separability or entanglement.

Peres [2] proposed a very simple but useful criterion for separability. As a density matrix is Hermitian and positive semi-definite, its transpose is still a valid density matrix. If we take the transpose of the matrices $\{\rho_i^{\mathrm{B}}\}$'s in Eq. (1), the resulting matrix, denoted by $\rho_{\mathrm{s}}^{\mathrm{T_B}}$, still contains non-negative eigenvalues. The operation is called partial transpose and can be defined for any bi-partite state:

$$\rho = \sum_{i,j,k,l} \rho_{ij;kl}|e_i^{\mathrm{A}} \otimes e_j^{\mathrm{B}}\rangle\langle e_k^{\mathrm{A}} \otimes e_l^{\mathrm{B}}| \longrightarrow \rho^{\mathrm{T_B}} \equiv \sum_{i,j,k,l} \rho_{i\underline{l};k\underline{j}}|e_i^{\mathrm{A}} \otimes e_j^{\mathrm{B}}\rangle\langle e_k^{\mathrm{A}} \otimes e_l^{\mathrm{B}}|, \tag{2}$$

where $|e_i^{\mathrm{A}} \otimes e_j^{\mathrm{B}}\rangle \equiv |e_i^{\mathrm{A}}\rangle \otimes |e_j^{\mathrm{B}}\rangle$ is the product basis used to represent the density matrix, and the underscores are used to highlight the changes under the partial transpose. Thus we have that if the state is separable, its partially transposed matrix has non-negative eigenvalues (usually called PPT). Said equivalently, if the state is not PPT under the partial transpose, the state must be entangled [3]. This is the Peres-Horodecki positive partial transpose (PPT) criterion for separability [2, 4].

Let us examine the example of a singlet state $|\Psi^-\rangle$. When written in the form of density matrix in the basis $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$, it corresponds to the density matrix

$$|\Psi^-\rangle\langle\Psi^-| \longleftrightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3}$$

The partial transpose takes it to

$$\begin{pmatrix} 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \end{pmatrix}, \tag{4}$$

which has one negative eigenvalue, $-1/2$. Thus, via the PPT criterion we see that $|\Psi^-\rangle$ is entangled.

Exercise. Find out the range of r such that

$$\rho(r) = r|\Psi^-\rangle\langle\Psi^-| + \frac{1-r}{4}I_{4\times4}$$

is entangled.

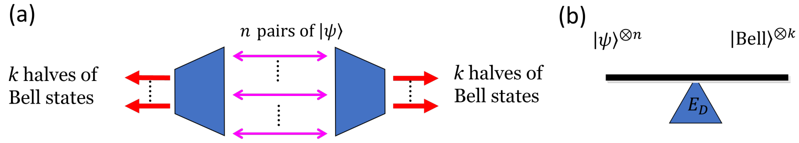

FIG. 2. The illustration of entanglement of distillation.

In general, this PPT criterion is necessary but not sufficient for establishing separability. However, it was shown by Horodecki and co-workers [4] that PPT is sufficient in the cases of $C^2 \otimes C^2$ (two-qubit) and $C^2 \otimes C^3$ (qubit-qutrit) systems: if two-qubit or qubit-qutrit states obey PPT, they are separable. On the other hand, if PPT is violated, the state is entangled. The extent to which a state violates PPT is manifested in the negative eigenvalues of the partially transposed density matrix, and can be used as a measure (not just an identifier) of entanglement; this measure is called the *negativity* [5, 6]. Following Życzkowski and co-workers [5], we define the negativity $\mathcal{N}$ to be twice the absolute value of the sum of the negative eigenvalues:

$$\mathcal{N}(\rho) = 2\max(0, -\lambda_{\mathrm{neg}}), \tag{5}$$

where λ_{neg} is the sum of the negative eigenvalues of $\rho^{\mathrm{T_B}}$ and the factor of two is a normalization chosen such that the singlet state $|\Psi^-\rangle$ has $\mathcal{N} = 1$.

VI. ENTANGLEMENT OF DISTILLATION

The notion of the entanglement of distillation was introduced by Bennett and co-workers [7, 8] to give an operational definition of the degree of entanglement. Suppose ρ represents the state of two particles possessed by two parties (usually referred to as Alice and Bob) separated by some distance. A way to envisage the degree of entanglement that ρ has is to ask how useful ρ is compared to a standard state, such as any of the four Bell states:

$$|\Psi^{\pm}\rangle \equiv \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \quad |\Phi^{\pm}\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle). \tag{6}$$

Here $\{|0\rangle, |1\rangle\}$ represents an orthonormal basis of a two-level system, for instance, the z-component of the spin of a spin-1/2 particle, or the polarization of a photon. More specifically, given n copies of the state ρ shared between Alice and Bob, how many pairs, say k, of Bell states can be obtained if each of Alice and Bob is allowed to (i) perform any local operations (including measurement) on the particles he or she possesses and (ii) share with the other party classical information, e.g., the outcome of some measurements. These operations are called local operations and classical communication (LOCC). The asymptotic limit

$$E_{\mathrm{D}}(\rho) \equiv \lim_{n\to\infty} (k/n), \tag{7}$$

is called the entanglement of distillation [7, 8] In it, k is the average number of Bell states taken over different possibilities (due to measurement) of an optimal procedure. E_{D} quantifies the entanglement as a resource, using Bell states as a standard ruler; see the illustration in Fig. 2.

A useful example. Let us illustrate the idea with an example. Suppose ρ is a pure state corresponding to the ket

$$|\psi_\theta\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle, \tag{8}$$

with $\theta \in [0, \pi/2]$. For $\theta = 0$ or $\pi/2$, it is not entangled. For the intermediate range of θ, the state is entangled, and maximally so at $\theta = \pi/4$. Suppose that two copies of non-maximally entangled $|\psi_\theta\rangle$ are shared between Alice and Bob:

$$\begin{aligned} |\psi_\theta\rangle_{12} \otimes |\psi_\theta\rangle_{34} &= (\cos\theta|0_1 0_2\rangle + \sin\theta|1_1 1_2\rangle) \otimes (\cos\theta|0_3 0_4\rangle + \sin\theta|1_3 1_4\rangle) \\ &= \cos^2\theta|0_1 0_2 0_3 0_4\rangle + \sqrt{2}\cos\theta\sin\theta\frac{1}{\sqrt{2}}(|0_1 0_2 1_3 1_4\rangle + |1_1 1_2 0_3 0_4\rangle) + \sin^2\theta|1_1 1_2 1_3 1_4\rangle, \end{aligned}$$

where Alice has particles 1 and 3, whereas Bob has 2 and 4. Their joint goal is to extract a Bell state under LOCC.

Both parties can perform any local operations allowed by quantum mechanics. A possible operation is to measure the number of his/her particles in state $|1\rangle$ (e.g., the z-component of total angular momentum). If Alice measures the

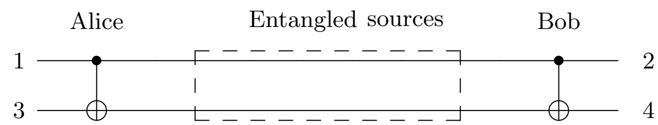

FIG. 3. Entanglement of distillation.

number of 1's of her particles to be 0 or 2 then the resulting state ($|0_10_20_30_4\rangle$ or $|1_11_21_31_4\rangle$) is unentangled. She needs to tell Bob to abort the operation, as there is now no entanglement to extract. But with probability $2\cos^2\theta\sin^2\theta$ she gets the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0_10_21_31_4\rangle + |1_11_20_30_4\rangle), \tag{9}$$

which is evidently entangled. But how do they establish from this a Bell state, say, $|\Phi^+\rangle$?

This time, Alice proceeds to perform a unitary transformation U on her particles and contacts Bob (which is when the classical communication takes place) and asks him to perform the same unitary transformation on *his* particles. Suppose that the unitary transformation they agree to perform is (also known as a CNOT operation)

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{10}$$

in the basis of $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. In particular, for it we have

$$U_{13}|0_11_3\rangle = |0_11_3\rangle,\ U_{13}|1_10_3\rangle = |1_11_3\rangle,\ U_{24}|0_21_4\rangle = |0_21_4\rangle,\ U_{24}|1_20_4\rangle = |1_21_4\rangle. \tag{11}$$

Then the joint state after the transformations becomes

$$\begin{aligned} U_{13}U_{24}|\psi\rangle &= U_{13}U_{24}\left(\frac{1}{\sqrt{2}}(|0_10_21_31_4\rangle + |1_11_20_30_4\rangle)\right) \\ &= \frac{1}{\sqrt{2}}(|0_10_21_31_4\rangle + |1_11_21_31_4\rangle) \\ &= \frac{1}{\sqrt{2}}(|0_10_2\rangle + |1_11_2\rangle) \otimes |1_31_4\rangle. \end{aligned} \tag{12}$$

As particles 3 and 4 are not entangled with 1 and 2, what now needs to be done is that Alice throws away her particle 3 and Bob throws away his particle 4. Finally, they have distilled one maximally entangled pair $\frac{1}{\sqrt{2}}(|0_10_2\rangle + |1_11_2\rangle)$ out of two non-maximally entangled pairs. The probability P of success is $2\cos^2\theta\sin^2\theta$, i.e., on average they can distill $k/n = \frac{1}{2}P = \cos^2\theta\sin^2\theta$ Bell pairs per initial pair.

A modification. In Fig. 3 we show a distillation scheme slightly modified from the two-pair example. In this modified scheme, Alice and Bob both perform the CNOT operation *before* the measurement. This transforms the initial state $|\psi_\theta\rangle_{12} \otimes |\psi_\theta\rangle_{34}$ as follows:

$$|\psi_\theta\rangle_{12} \otimes |\psi_\theta\rangle_{34} \to (\cos^2\theta|0_10_2\rangle + \sin^2\theta|1_11_2\rangle)|0_30_4\rangle + \sqrt{2}\cos\theta\sin\theta\frac{1}{\sqrt{2}}(|0_10_2\rangle + |1_11_2\rangle)|1_31_4\rangle. \tag{13}$$

If Alice and/or Bob then measures the third and/or fourth qubit, respectively, and the outcome is $|1\rangle$, they immediately obtain a Bell state shared between particles 1 and 2. If the outcome is $|0\rangle$, they get a slightly less entangled state, which they can store for a second trial of distillation. What we mean by this is that two pairs of the states

$$\frac{1}{\sqrt{\cos^4\theta + \sin^4\theta}}\big(\cos^2\theta|00\rangle + \sin^2\theta|11\rangle\big), \tag{14}$$

although less entangled than the original pairs of Eq. (8), is distillable. Thus, this modified scheme performs slightly better than the original two-pair scheme.

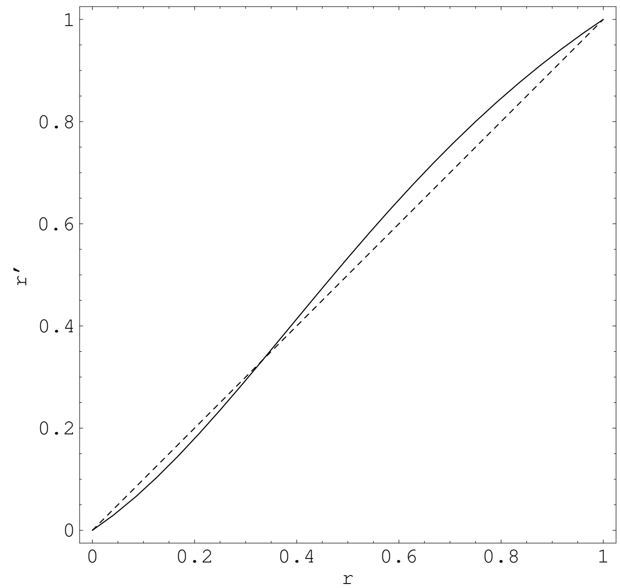

FIG. 4. Mixed state distillation for Werner state $\rho_{W+}(r)$. States above the dashed line (i.e., $r > 1/3$) can be distilled by the procedure.

General case. The above discussion involves Alice and Bob dealing with two pairs at a time. In fact, it can be extended to the case where they can manipulate n copies at a time [7]. The average number of Bell pairs per initial pair can be dervied to be

$$\overline{E}(n) = \frac{1}{n}\sum_{k=0}^{n} P(k)E_k = \frac{1}{n}\sum_{k=0}^{n} P(k)\log_2(C_k^n), \tag{15a}$$

$$P(k) \equiv (\cos^2\theta)^{n-k}(\sin^2\theta)^k C_k^n, \tag{15b}$$

where $C_k^n \equiv n!/[k!(n-k)!]$. As the number n of copies approaches infinity,

$$\lim_{n\to\infty} \overline{E}(n) \to E_{\rm D} = S(\rho_{\rm A}), \tag{16}$$

where $\rho_A \equiv {\rm Tr_B}|\psi_\theta\rangle\langle\psi_\theta|$, and $S(\rho) \equiv -{\rm Tr}\rho\log_2\rho$ is the von Neumann entropy of ρ. In the case of $|\psi_\theta\rangle$, its entanglement of distillation is $E_{\rm D} = h(\cos^2\theta)$, where $h(x) \equiv -x\log_2(x) - (1-x)\log_2(1-x)$, i.e., is the Shannon entropy. The result

$$E_{\rm D} = -{\rm Tr}\big(\rho_A \log_2(\rho_A)\big), \tag{17}$$

is valid for *any* bi-partite pure state.

Mixed states. For mixed entangled states, there are very few cases for which $E_{\rm D}$ is known. No general optimal distillation procedure is known for generic states. But a similar set-up to the one shown in Fig. 3 (except that the measurement is performed in a different basis) does provide a way (although not optimal) to distill very general two-qubit states. For example, Bennett and co-workers [8] have shown that after one step of the mixed-state distillation procedure, two initial pairs of the state (which is usually called the Werner state)

$$\rho_{\rm W+}(r) \equiv r|\Psi^+\rangle\langle\Psi^+| + \frac{1-r}{4}\mathbb{1}, \tag{18}$$

will be transformed into one pair with a new parameter (see Fig. 4)

$$r' = \frac{2r(1+2r)}{3(1+r^2)}. \tag{19}$$

Note that the larger the parameter r is, the higher entanglement the Werner state possesses. If $r' > r$, i.e., when $r > 1/3$ (which you should prove in your exercise) or equivalently the fidelity $F \equiv \langle\Psi^+|\rho_{\rm W}|\Psi^+\rangle > 1/2$, the entanglement

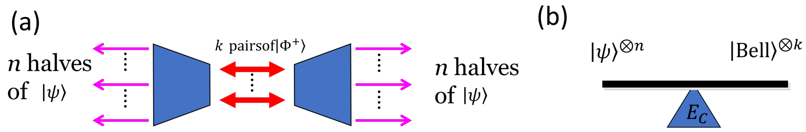

FIG. 5. The illustration of entanglement of creation/dilution.

is said to be increased. Horodecki and co-workers [9] further showed that any entangled two-qubit state can be transformed into a state $\rho_{\mathrm{W}+}(r)$ with $r > 1/3$, and hence can be distilled via the scheme of Bennett and co-workers (also known as the BBPSSW scheme) [8].

However, the procedure is not optimal for an arbitrary state ρ, and it is generally rather difficult to compute $E_{\mathrm{D}}(\rho)$. Nevertheless, if $E_{\mathrm{D}}(\rho) > 0$ we say that the state ρ is *distillable*. We remark that there is a connection between the PPT criterion and the distillability of a bi-partite state. Horodecki and co-workers [10] found that if a state has PPT then it cannot be distilled. But the converse is not generally true. In fact, there is still no simple criterion to determine whether or not a state is distillable.

VII. ENTANGLEMENT COST AND ENTANGLEMENT OF FORMATION

The distillation is a process for concentrating entanglement from a large number of pairs with less entanglement into a small number of pairs with more (and ultimately maximal) entanglement. On the other hand, we can consider the converse process, which is usually called *dilution*. Given k pairs of Bell states shared between Alice and Bob, how many pairs n of a given state ρ can be obtained by local operations (including adding unentangled particles) and classical communication? This is illustrated in Fig. 5. The goal is to maximize the number n of copies of the output state ρ. The optimal ratio defines the *entanglement cost* [11]:

$$E_{\mathrm{C}}(\rho) \equiv \lim_{k\to\infty} (k/n). \tag{20}$$

As with E_{D}, E_{C} is very difficult to calculate for general mixed states, and is only known for a very few special cases.

However, for pure states such as the state $|\psi_\theta\rangle$ discussed previously, $E_{\mathrm{C}} = -\mathrm{Tr}\big(\rho_A \log_2(\rho_A)\big)$, which equals E_{D}. The optimal way to realize this dilution process for the pure state is to utilize two techniques: (i) quantum teleportation, which we have introduced at the beginning and which simply says that a Bell state shared between two parties can be used to tranfer an unknown qubit state with certainty, and (ii) *quantum data compression* [12], which basically states that a large message consistive of say n qubits, with each qubit on average being described by a density matrix ρ_{A}, can be compressed into a possibly smaller number $k = nS(\rho_{\mathrm{A}}) \leq n$ of qubits; and one can faithfully recover the whole message, as long as n is large enough. We will discuss *quantum data compression* later.

Reversibility for pure states in the asymptotic limit. With these two tools in hand, Alice can first prepare n copies of $|\psi_\theta\rangle$ ($2n$ qubits in total) locally, compress the n qubits to k qubits that she will "send" to Bob, and teleport the compressed k qubits to Bob using the shared k Bell states. Bob then decompresses the k qubits back to the uncompressed n qubits, which belong to half of the n copies of the entangled state $|\psi_\theta\rangle$. Thus, Alice and Bob establish n pairs of $|\psi_\theta\rangle$. This describes the optimal procedure for the dilution process for a pure state.

The entanglement of distillation and entanglement cost are defined asymptotically, i.e., both processes involve an infinite number of copies of the initial states. For pure states, $E_{\mathrm{C}} = E_{\mathrm{D}}$ [7], which means that the two processes are reversible asymptotically. Yet, for mixed states, both quantities are very difficult to calculate. Nevertheless, it is expected that $E_{\mathrm{C}}(\rho) \geq E_{\mathrm{D}}(\rho)$, viz. that one can not distill more entanglement than is put in.

Entanglement of formation—an average quantity. However, as we now explain, there is a modification of E_{C}, obtained by averaging E_{C} over pure states, and it is called the *entanglement of formation* E_{F} [11, 13]. Any mixed state ρ can be decomposed into mixture of pure states $\{p_i, |\psi_i\rangle\langle\psi_i|\}$, although the decomposition is far from unique. To construct the mixed state via mixing pure states in this way will cost, on average, $\sum_i p_i E(|\psi_i\rangle\langle\psi_i|)$ pairs of Bell states. The entanglement of formation for a mixed state ρ is thus defined as the *minimal* average number of Bell states needed to realize an ensemble described by ρ, i.e.,

$$E_{\mathrm{F}}(\rho) \equiv \min_{\{p_i,\psi_i\}} \sum_i p_i\, E_{\mathrm{C}}(|\psi_i\rangle\langle\psi_i|), \tag{21}$$

where the minimization is taken over those probabilities $\{p_i\}$ and pure states $\{\psi_i\}$ that, taken together, reproduce the density matrix $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$. Such a construction is usually called a *convex hull* construction. Furthermore, the quantity $E_{\rm C}(|\psi_i\rangle\langle\psi_i|)$ is the entropy of entanglement of pure state $|\psi_i\rangle$, viz. the expression in the right-hand side of Eq. (17). However, $E_{\rm F}$ is, in general, also difficult to calculate for mixed states, as it involves a minimization over all possible decompositions. So far, there has been more analytic progress for $E_{\rm F}$ than forn $E_{\rm C}$ and $E_{\rm D}$. Notable cases include (i) Wootters' formula for arbitary two qubits [13]; (ii) Terhal and Vollbrecht's formula [14] for *isotropic* states for two qu-*dit*s (d-level parties); and (iii) Vollbrecht and Werner's formula [15] for generalized Werner states of two qu-dits. But these are more advanced results for this course.

Additivity? One of the central issues in entanglement theory is the so-called *additivity* of entanglement, i.e., whether the entanglement of formation, defined as an average quantity, equals the entanglement cost, which is defined asymptotically. Shor [16] has established that the additivity problem of entanglement of formation is equivalent to three other additivity problems: the strong superadditivity of the entanglement of formation, the additivity of the minimum output entropy of a quantum channel, and the additivity of the Holevo classical capacity of a quantum channel. However, further discussion of these additivity problems is beyond the scope of this dissertation.

VIII. ENTANGLEMENT VIA A DISTANCE MEASURE

As any mixture of separable density matrices is still, by definition, separable, any separable state can be expressed as a sum of two separable states

$$\rho_s = p\,\rho_s^1 + (1-p)\rho_s^2, \tag{22}$$

unless it is the extremal point, viz. a pure product state. Thus we see that the set of separable states is a *convex* set. This leads to another type of entanglement measure: the shortest "distance" $E(\rho)$ from an entangled state to the convex set $D_{\rm s}$ of separable mixed states [17], i.e., $E(\rho)\equiv \min_{\sigma\in D_{\rm s}} d(\rho||\sigma)$. One example of such an entanglement measure is the relative entropy of entanglement,

$$E_{\rm R}(\rho)\equiv \min_{\sigma\in D_{\rm s}} {\rm Tr}\left(\rho\log\rho-\rho\log\sigma\right), \tag{23}$$

where the distance measure d is defined to be the relative entropy of two states:

$$d(\rho||\sigma)\equiv {\rm Tr}\big[\rho\log\rho-\rho\log\sigma\big]. \tag{24}$$

We remark that the relative entropy is non-negative, but it is also not symmetric, i.e., $d(\rho||\sigma)\ne d(\sigma||\rho)$. For pure states this definition of entanglement reduces to the entropy of entanglement.

Another example is the Bures metric of entanglement $E_{\rm B}(\rho)$, defined via

$$E_{\rm B}(\rho)\equiv \min_{\sigma\in D_{\rm s}} \big[2-2F(\rho,\sigma)\big], \tag{25}$$

where $F(\rho,\sigma)\equiv\big({\rm Tr}\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\big)^2$ is called the *fidelity* and *is* symmetric. For two pure states $\rho=|\psi\rangle\langle\psi|$ and $\sigma=|\phi\rangle\langle\phi|$, the distance $d(\rho||\sigma)\equiv\big(2-2F(\rho,\sigma)\big)$ reduces to $2(1-|\langle\psi|\phi\rangle|^2)$.

We shall give more discussion of the relative entropy of entanglement later on.

IX. WOOTTERS' FORMULA

The entanglement of formation defined in Eq. (21) is, in general, difficult to calculate. However, for two-qubit systems, Wootters [13] has provided and proved the following formula:

$$E_{\rm F}(\rho)=h\left(\frac{1}{2}[1+\sqrt{1-C(\rho)^2}]\right), \tag{26}$$

where $h(x)\equiv -x\log_2 x-(1-x)\log_2(1-x)$, and $C(\rho)$, the *concurrence* of the state ρ, is defined as

$$C(\rho)\equiv\max\{0,\sqrt{\lambda_1}-\sqrt{\lambda_2}-\sqrt{\lambda_3}-\sqrt{\lambda_4}\}, \tag{27}$$

in which $\lambda_1,\dots,\lambda_4$ are the eigenvalues of the matrix $\rho(\sigma_y\otimes\sigma_y)\rho^*(\sigma_y\otimes\sigma_y)$ in nonincreasing order and σ_y is a Pauli spin matrix. $E_{\rm F}(\rho)$, $C(\rho)$ and the *tangle* $\tau(\rho)\equiv C(\rho)^2$ are equivalent measures of entanglement, inasmuch as they are monotonic functions of one another. For pure state $a|00\rangle+b|01\rangle+c|10\rangle+d|11\rangle$ the concurrence C is $2|ad-bc|$, as discussed before.

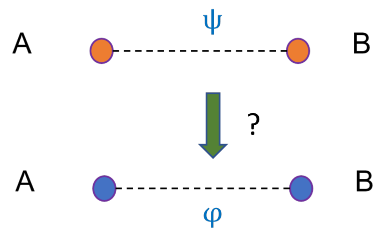

FIG. 6. The illustration of entanglement conversion of a single copy.

X. CRITERIA FOR GOOD ENTANGLEMENT MEASURES

From the discussions above, we know that entanglement is a nonlocal quantum characteristic which cannot be generated by local operations and classical communication between two distant parties. Local unitary transformations are simply changes of local basis, and cannot change the amount of entanglement. These, as well as other considerations lead to postulates for entanglement measures. For a given state ρ a good entanglement measure should, at least, satisfy the following conditions [17–19]:
C1. (a) $E(\rho) \geq 0$; (b) $E(\rho) = 0$ if ρ is not entangled.
C2. For any state ρ and any local unitary transformation (i.e., a unitary transformation of the form $U_A \otimes U_B$), the entanglement should remain unchanged.
C3. Local operations, classical communication and postselection (i.e., keeping certain measurement outcomes and discarding the rest) should not increase the expectation value of the entanglement.
C4. Entanglement is convex under discarding information: $\sum_i p_i E(\rho_i) \geq E(\sum_i p_i \rho_i)$.

There are additional postulates such as continuity and additivity, but C1-C4 are those widely accepted in the literature. Note that in C1b we do not say "$E(\rho) = 0$ *iff* ρ is not entangled." This is because cases exist in which $E_{\rm F}(\rho) > 0$ although $E_{\rm D}(\rho) = 0$, i.e., there can exist *bound entanglement* [10], entanglement that cannot be distilled.

XI. ENTANGLEMENT TRANSFORMATION FOR SINGLE COPY

We have seen using many copies for entanglement distillation and cost. In the limit of infinite copies of bipartite pure states, these two processes are reversible. However, in practice, one has only finite copies. In the extreme limit, what if Alice and Bob shares a single copy of ψ; can this be converted to another state ϕ (using only local operation and classical communication)? This is illustrated in Fig. 6.

Nielsen's majorization criterion. In order to discussion Nielsen's criterion [20], we need to acquaint ourselves with the notion of majorization.

Majorization: for two sets of numbers $\mathbf{x}$ & $\mathbf{y}$ (e.g. square of Schmidt coefficients in decreasing order), $\mathbf{x}$ is majorized by $\mathbf{y}$ ($\mathbf{x} \prec \mathbf{y}$) if

$$\sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} y_j, \text{ for } k = 1, \ldots, d.$$

This is illustrated in Fig. 7. Now that we understand majorization for two sets of numbers, we can state Nielsen's result. Nielsen showed that the transform $\psi \to \phi$ of two bipartite pure states is possible with probability 1 if and only if the square of their Schmidt coefficients (denoted by λ_ψ, λ_ϕ, respectively) satisfy the majorization relation,

$$\psi \to \phi, \quad \text{iff } \lambda_\psi \prec \lambda_\phi. \tag{28}$$

Furthermore, if the majorization condition is not satisfied, the maximum probability of conversion is given by

$$p_{\max}\big(\psi \to \phi\big) = \min_{1 \leq m \leq n} \frac{1 - \sum_{j=1}^{m-1} \lambda_j(\psi)}{1 - \sum_{j=1}^{m-1} \lambda_j(\phi)}.$$

An example. Let us consider $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and $|\phi(p)\rangle = \sqrt{1-p}|00\rangle + \sqrt{p}|11\rangle$. Assume $p \leq 1/2$, As

$$\{1/2, 1/2\} \prec \{1-p, p\},$$

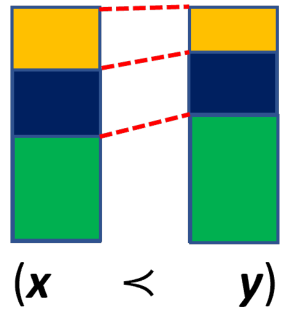

FIG. 7. The illustration of majorization

we see that the conversion $|\Phi^+\rangle \to |\phi(p)\rangle$ is possible.

Note that we write $|\phi(p)\rangle$ in the Schmidt form above; even if it were not in the Schmidt form, the conversion is still possible as long as the Schmidt coefficients are the same, because the transformation to the Schmidt form can be done by local unitary transformations. A question for the readers to think is how can such conversion be done? What is the procedure?

Another example. Let us study another example.

$$|\psi\rangle = \sqrt{\frac{1}{2}}|11\rangle + \sqrt{\frac{2}{5}}|22\rangle + \sqrt{\frac{1}{10}}|33\rangle, \tag{29}$$

$$|\phi\rangle = \sqrt{\frac{3}{5}}|11\rangle + \sqrt{\frac{1}{5}}|22\rangle + \sqrt{\frac{1}{5}}|33\rangle. \tag{30}$$

As one can check that λ_ψ and λ_ϕ have no majorization relation, therefore, we conclude that

$$|\psi\rangle \nrightarrow |\phi\rangle,\ |\phi\rangle \nrightarrow |\psi\rangle. \tag{31}$$

A natural question is that what is the maximum conversion probability?

Entanglement catalysis. This phenomenon was first discussed by Jonathan and Plenio [21]. To understand this, we use a third example with

$$|\psi_1\rangle = \sqrt{0.4}|00\rangle + \sqrt{0.4}|11\rangle + \sqrt{0.1}|22\rangle + \sqrt{0.1}|33\rangle, \tag{32}$$

$$|\psi_2\rangle = \sqrt{0.5}|00\rangle + \sqrt{0.25}|11\rangle + \sqrt{0.25}|22\rangle. \tag{33}$$

This is also an example where

$$|\psi_1\rangle \nrightarrow |\psi_2\rangle, \text{as}\, \lambda_{\psi_1} \nprec \lambda_{\psi_2}.$$

However, if they have access to another shared (entangled) state,

$$|\phi\rangle = \sqrt{0.6}|44\rangle + \sqrt{0.4}|55\rangle,$$

the combined systems have the square of Schmidt coefficients being

$$|\psi_1\rangle \otimes |\phi\rangle : 0.24, 0.24, 0.16, 0.16, 0.06, 0.06, 0.04, 0.04 \tag{34}$$

$$|\psi_2\rangle \otimes |\phi\rangle : 0.30, 0.20, 0.15, 0.15, 0.10, 0.10, 0.00, 0.00. \tag{35}$$

Then

$$|\psi_1\rangle|\phi\rangle \to |\psi_2\rangle|\phi\rangle \text{ is allowed.}$$

A natural question that can be further investigated is the condition that entanglement catalysis exists?

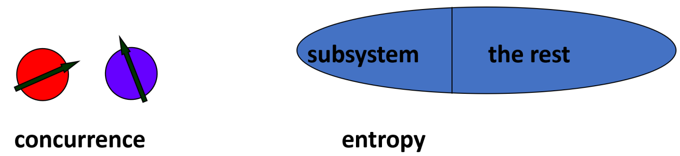

FIG. 8. The illustration of bipartite entanglement.

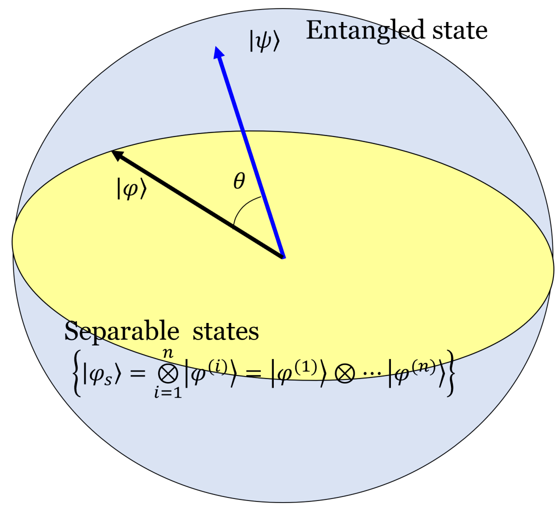

FIG. 9. The illustration of the geometric measure of entanglement for pure states.

XII. MULTIPARTITE ENTANGLEMENT

We have seen the concurrence and entanglement entropy previously, and they are essentially bipartite, i.e., their entanglement is between two parts. In previous units, we have seen many examples of multipartite states, such as the GHZ state,

$$|\text{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2},$$

and three-qubit redudant encoding,

$$|\psi_3\rangle = a|000\rangle + b|111\rangle,$$

as well as all quantum error correction codes. In this section, we will study two multipartite entanglement measures: (i) Vedral '04: relative entropy of entanglement (ER); (ii) Wei & Goldbart '03: geometric measure of entanglement (GME). We have seen the relative entropy of entanglement in Sec. VIII, where we were discussing bipartite mixed-state entanglement. Its definition applies to the multipartite case as the separable states $\sigma \in \mathcal{D}_S$ are simply multipartite separable states. In the case of GHZ state, the closest separable state is $\sigma^* = (|000\rangle\langle 000| + |111\rangle\langle 111|)/2$ and then $E_R(\text{GHZ}) = 1$. However, this measure is hard to compute, even for pure states, e.g., the W state,

$$|\text{W}\rangle = (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}.$$

It turns out that the GME provides a lower bound on the ER for pure states.

A. Geometric measure of entanglement*

[This section can be skipped in the first reading.] The underlying idea of this entanglement measure in the case of pure states is very simple, as illustrated in Fig. 9. The discussion here based Ref. [22].

We begin with an examination of entangled *pure* states, and of how one might quantify their entanglement by making use of simple ideas of Hilbert space geometry. Let us start by developing a quite general formulation, appropriate for multi-partite systems comprising n parts, in which each part can have a distinct Hilbert space. Consider a general n-partite pure state

$$|\psi\rangle = \sum_{p_1\cdots p_n} \chi_{p_1p_2\cdots p_n}|e_{p_1}^{(1)}e_{p_2}^{(2)}\cdots e_{p_n}^{(n)}\rangle, \tag{36}$$

where $\{e_{p_k}^{(k)}\}$ is the local basis of the k-th party, e.g., the spin $\uparrow$ or $\downarrow$. One can envisage a geometric definition of its entanglement content via the distance

$$d = \min_{|\phi\rangle} \| \, |\psi\rangle - |\phi\rangle \| \tag{37}$$

between $|\psi\rangle$ and the nearest of the separable states $|\phi\rangle$ (or equivalently the angle between them). Here

$$|\phi\rangle \equiv \otimes_{i=1}^{n}|\phi^{(i)}\rangle = |\phi^{(1)}\rangle \otimes |\phi^{(2)}\rangle \otimes \cdots \otimes |\phi^{(n)}\rangle \tag{38}$$

is an arbitrary separable (i.e., Hartree) n-partite pure state, the index $i = 1 \ldots n$ labels the parties, and a state vector of part i is written as

$$|\phi^{(i)}\rangle \equiv \sum_{p_i} c_{p_i}^{(i)} \, |e_{p_i}^{(i)}\rangle. \tag{39}$$

It seems natural to assert that the more entangled a state is, the further away it will be from its best unentangled approximant (and, correspondingly, the wider will be the angle between them). We emphasize that we only compare the entangled pure state to the set of *pure* unentangled state. We shall extend to mixed states via the so-called convex-hull construction later. Another approach one might take is to compare any entangled state to the set of unentalged states, including both pure and mixed. The Bures measure, introduced earlier in Sec. VIII, is such an example.

To actually find the nearest separable state, it is convenient to minimize, instead of d, the quantity d^2, i.e.,

$$\||\psi\rangle - |\phi\rangle\|^2, \tag{40}$$

subject to the constraint $\langle\phi|\phi\rangle = 1$. In fact, in solving the resulting stationarity condition one may restrict one's attention to the subset of solutions $|\phi\rangle$ that obey the further condition that each factor $|\phi^{(i)}\rangle$ obeys its own normalization condition $\langle\phi^{(i)}|\phi^{(i)}\rangle = 1$. Thus, by introducing a Lagrange mulitplier Λ to enforce the constraint $\langle\phi|\phi\rangle = 1$, differentiating with respect to the independent amplitudes, and then imposing the further condition $\langle\phi^{(i)}|\phi^{(i)}\rangle = 1$, one arrives at the *nonlinear eigenproblem* for the stationary $|\phi\rangle$:

$$\sum_{p_1\cdots\widehat{p_i}\cdots p_n} \chi^*_{p_1p_2\cdots p_n} c_{p_1}^{(1)} \cdots \widehat{c_{p_i}^{(i)}} \cdots c_{p_n}^{(n)} = \Lambda \, c_{p_i}^{(i)*}, \tag{41a}$$

$$\sum_{p_1\cdots\widehat{p_i}\cdots p_n} \chi_{p_1p_2\cdots p_n} c_{p_1}^{(1)*} \cdots \widehat{c_{p_i}^{(i)*}} \cdots c_{p_n}^{(n)*} = \Lambda \, c_{p_i}^{(i)} \;, \tag{41b}$$

where the eigenvalue Λ is associated with the Lagrange multiplier enforcing the constraint $\langle\phi|\phi\rangle=1$, and the symbol $\widehat{\ }$ denotes the exclusion of the corresponding term or factor. In a form independent of the choice of basis within each party, Eqs. (41) read

$$\langle\psi|\Big(\mathop{\otimes}_{j(\neq i)}^{n} |\phi^{(j)}\rangle\Big) = \Lambda\langle\phi^{(i)}|, \tag{42a}$$

$$\Big(\mathop{\otimes}_{j(\neq i)}^{n} \langle\phi^{(j)}|\Big)|\psi\rangle = \Lambda|\phi^{(i)}\rangle. \tag{42b}$$

From Eqs. (41) or (42), e.g., by taking inner product of both sides of Eq. (42a) with $|\phi^{(i)}\rangle$ one readily sees that

$$\Lambda = \langle\psi|\phi\rangle = \langle\phi|\psi\rangle \tag{43}$$

and thus the eigenvalues Λ are real, in $[-1,1]$, and independent of the choice of the local basis $\big\{|e_{p_i}^{(i)}\rangle\big\}$. Hence, the spectrum Λ is the cosine of the angle between $|\psi\rangle$ and $|\phi\rangle$; the largest, $\Lambda_{\max}$, which we call the *entanglement eigenvalue*, corresponds to the closest separable state and is equal to the maximal overlap

$$\Lambda_{\max} = \max_{\phi} ||\langle\phi|\psi\rangle||, \tag{44}$$

where $|\phi\rangle$ is an arbitrary separable pure state.

Although, in determining the closest separable state, we have used the squared distance between the states, there are alternative (basis-independent) candidates for entanglement measures which are related to it in an elementary way: the distance, the sine, or the sine squared of the angle θ between them (with $\cos\theta \equiv \mathrm{Re}\,\langle\psi|\phi\rangle$). We shall adopt $E_1 \equiv 1 - \Lambda_{\max}^2$ as our entanglement measure because, as we shall see, when generalizing E_1 to mixed states we have been able to show that it satisfies a set of criteria demanded of entanglement measures.

Bipartite systems. In bi-partite applications, the eigenproblem (41) is in fact *linear*, and solving it is actually equivalent to finding the Schmidt decomposition [23]. Moreover, the entanglement eigenvalue is equal to the maximal Schmidt coefficient. To be more precise, in bi-partite systems the stationarity conditions (41) reduce to the linear form

$$\sum_{p_1} \chi^*_{p_1 p_2} c^{(1)}_{p_1} = \Lambda\, c^{(2)*}_{p_2}, \tag{45a}$$

$$\sum_{p_1} \chi_{p_1 p_2} c^{(1)*}_{p_1} = \Lambda\, c^{(2)}_{p_2}, \tag{45b}$$

$$\sum_{p_2} \chi^*_{p_1 p_2} c^{(2)}_{p_2} = \Lambda\, c^{(1)*}_{p_1}, \tag{45c}$$

$$\sum_{p_2} \chi_{p_1 p_2} c^{(2)*}_{p_2} = \Lambda\, c^{(1)}_{p_1}. \tag{45d}$$

Eliminating $c^{(2)}_p$ between Eqs. (45a) and (45d) and, similarly, eliminating $c^{(1)}_p$ between Eqs. (45b) and (45c) gives

$$\sum_{p_1'\, p_2} \chi_{p_1 p_2} \chi^*_{p_1' p_2} c^{(1)}_{p_1'} = \Lambda^2\, c^{(1)}_{p_1}, \tag{46a}$$

$$\sum_{p_1\, p_2'} \chi_{p_1 p_2} \chi^*_{p_1 p_2'} c^{(2)}_{p_2'} = \Lambda^2\, c^{(2)}_{p_2}, \tag{46b}$$

or equivalently

$$\mathrm{Tr}_2\big(|\psi\rangle\langle\psi|\big)|\phi^{(1)}\rangle = \Lambda^2\, |\phi^{(1)}\rangle, \tag{47a}$$

$$\mathrm{Tr}_1\big(|\psi\rangle\langle\psi|\big)|\phi^{(2)}\rangle = \Lambda^2\, |\phi^{(2)}\rangle. \tag{47b}$$

Now, solving the above equations is equivalent to finding the Schmidt decomposition for $|\psi\rangle$. To see this, first recall that Schmidt decomposability guarantees that an arbitrary pure bi-partite state

$$|\psi\rangle = \sum_{p_1 p_2} \chi_{p_1 p_2} |e^{(1)}_{p_1}\rangle \otimes |e^{(2)}_{p_2}\rangle \tag{48}$$

in the (product) Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ (with factor dimensions d_1 and d_2) can always be expressed in the form

$$|\psi\rangle = \sum_{k=1}^{\min\{d_1,d_2\}} \lambda_k\, |\tilde{e}^{(1)}_k\rangle \otimes |\tilde{e}^{(2)}_k\rangle. \tag{49}$$

Here, $\lambda_k \geq 0$, $\sum_k \lambda_k^2 = 1$, the $|\tilde{e}^{(1)}_k\rangle$'s and $|\tilde{e}^{(2)}_k\rangle$'s are orthonormal, respectively, in $\mathcal{H}_1$ and $\mathcal{H}_2$. Moreover, the new (tilde) bases, as well as the λ_k's, follow as the solution of the eigenproblems of the reduced density matrix that one obtains by tracing $|\psi\rangle\langle\psi|$ over party 1 or 2:

$$\mathrm{Tr}_2\big(|\psi\rangle\langle\psi|\big)|\tilde{e}^{(1)}_k\rangle = \lambda_k^2 |\tilde{e}^{(1)}_k\rangle, \tag{50a}$$

$$\mathrm{Tr}_1\big(|\psi\rangle\langle\psi|\big)|\tilde{e}^{(2)}_k\rangle = \lambda_k^2 |\tilde{e}^{(2)}_k\rangle. \tag{50b}$$

These are Eqs. (47); thus we see that determining entanglement for bi-partite pure states is equivalent to finding their Schmidt decomposition, except that one only needs the *largest* Schmidt coefficient $\Lambda_{\max} = \lambda_{\max}$.

Example: for $|\psi\rangle = \sqrt{1-p}|00\rangle + \sqrt{p}|11\rangle$, it $\Lambda_{\max} = \max\{\sqrt{1-p}, \sqrt{p}\}$.

Multipartite systems. By contrast, for the case of three or more parts, the eigenproblem is a *nonlinear* one. For example, in the setting of tri-partite systems, the stationarity conditions (41) associated with the pure state $|\psi\rangle = \sum_{p_1p_2p_3} \chi_{p_1p_2p_3} |e^{(1)}_{p_1} e^{(2)}_{p_2} e^{(3)}_{p_3}\rangle$ become

$$\sum_{p_2p_3} \chi^*_{p_1p_2p_3} c^{(2)}_{p_2} c^{(3)}_{p_3} = \Lambda\, c^{(1)*}_{p_1}, \tag{51a}$$

$$\sum_{p_1p_3} \chi^*_{p_1p_2p_3} c^{(1)}_{p_1} c^{(3)}_{p_3} = \Lambda\, c^{(2)*}_{p_2}, \tag{51b}$$

$$\sum_{p_1p_2} \chi^*_{p_1p_2p_3} c^{(1)}_{p_1} c^{(2)}_{p_2} = \Lambda\, c^{(3)*}_{p_3}. \tag{51c}$$

Note the nonlinear structure of this tri-partite (and, in general, any $n \geq 3$-partite) eigenproblem. As far as we are aware, for nonlinear eigenproblems such as these, one cannot take advantage of the simplicity of linear eigenproblems, for which one can address the task of determining the eigenvalues directly (via the characteristic polynomial), without having to address the eigenvectors. Hence, even for systems comprising qubits, one is forced to proceed numerically. This is consistent with the notion that no Schmidt decomposition exists beyond bi-partite systems.

Two explicitly forms.

- $E_1(\psi) \equiv 1 - \Lambda^2_{\max}(\psi)$.
- $E_g(\psi) \equiv -2\log_2 \Lambda_{\max}(\psi)$.

The latter form can be used for infinite systems and gives an entanglement per particle in the thermodynamic limit,

$$\mathcal{E}_g(\psi) = \lim_{N\to\infty} \frac{E_g(\psi)}{N}.$$

Lower bound on the relative entropy of entanglement. We mention earlier that the GME provides a lower bound on ER. The exact statement is that

$$E_R(\psi) \geq E_g(\psi).$$

It is known that for permutation-invariant states, the above relation becomes an equality. In particular, this gives the value of $E_R(\mathrm{W})$ for the W state.

Examples. For the GHZ state, its $\Lambda_{\max} = 1/\sqrt{2}$. For the W state, its $\Lambda_{\max} = 2/3$, and the closest product state is $|\phi^*\rangle = (\sqrt{2/3}|0\rangle + \sqrt{1/3}|1\rangle)^{\otimes 3}$. From these, we also obtain their relative entropy of entanglement (where the lower bound from GME becomes an equality),

$$E_R(W) = -2\log_2(2/3) \approx 1.16833, \quad E_R(\mathrm{GHZ}) = 1.$$

B. Extension to mixed states

The extension of the geometric measure to mixed states ρ can be made via the use of the *convex roof* (or *hull*) construction [indicated by "co"], as was done for the entanglement of formation by Wootters and collaborators. The essence is a minimization over all decompositions $\rho = \sum_i p_i\, |\psi_i\rangle\langle\psi_i|$ into pure states, i.e.,

$$E(\rho) \equiv (\mathrm{co}\, E_{\mathrm{pure}})(\rho) \equiv \min_{\{p_i,\psi_i\}} \sum\nolimits_i p_i\, E_{\mathrm{pure}}(|\psi_i\rangle). \tag{52}$$

This convex hull construction ensures that the measure gives zero for separable states; however, in general it also complicates the task of determining mixed-state entanglement.

Now, any good entanglement measure E should, at least, satisfy the following criteria (c.f. Refs. [17–19, 24]) listed above, but repeated here,

C1. (a) $E(\rho)\geq 0$; (b) $E(\rho)=0$ if ρ is not entangled.

C2. Local unitary transformations do not change E.

C3. Local operations and classical communication (LOCC) (as well as post-selection) do not increase the expectation value of E [25].

C4. Entanglement is convex under the discarding of information, i.e., $\sum_i p_i\, E(\rho_i) \geq E(\sum_i p_i\, \rho_i)$.

The first requirement simply states that the entanglement is a non-negative quantity, and it vanishes for unentangled states. Local unitary transformations are simply a change of local basis, and entanglement should be invariant under such a change. The third criterion simply requires that the average entanglement cannot be increased during manipulations that are local, which reflects the fact that entanglement is a nonlocal resource. The last criterion states the fact that, for a set of states, entanglement can never be increased if the information about which state is which is lost. The issue of the desirability of additional features, such as continuity and additivity, requires further investigation, but C1-C4 are regarded as the minimal set. If a measure satisfies C1-C4, it is called an *entanglement monotone* [24].

Does the geometric measure of entanglement obey C1-4? The answer is affirmative but we will not give detailed proof here. We remark that from the definition (52) it is evident that C1 and C2 are satisfied, provided that $E_{\rm pure}$ satisfies them. Furthermore, the convex-hull construction automatically fulfills C4. The consideration of C3 seems to be more delicate and whether or not it holds depends on the explicit form of $E_{\rm pure}$. In Appendix **??** we show that the choice of taking $E_{\sin^2}$ as the entanglement measure *does* satisfy C3.

XIII. GEOMETRIC MEASURE OF ENTANGLEMENT AND ONE-WAY QUANTUM COMPUTER

`‘‘Entanglement, like most good things in life, must be consumed in moderation’’ (Bacon, 2009).`
We have seen that measurement-based quantum computation uses entanglement as a resource and proceeds computation by performing single-qubit measurements. The high value of initial entanglement is crucial. However, it turns out that too much entanglement cannot necessarily be useful.

Gross, Flammia, and Eisert (Gross et al., 2009) [26] found that random states generically have a high amount of entanglement and if the entanglement of a quantum state is too high, then using it for MBQC cannot offer any speedup for computation and is no better than random coin tossing. What they mean by being too high is expressed in terms the geometric entanglement of a state,

$$E_g(\psi) = -2\log_2 \Lambda_{\max}(\psi) > n - \delta,$$

where n is the number of qubits and delta is a small constant.

A similar conclusion that random states drawn uniformly from the state space (or in a more technical term, from the Haar measure) are useless for MBQC was reached by Bremner, Mora, and Winter (Bremner et al., 2009) [27].

XIV. QUANTUM DATA COMPRESSION

It is also known as the Noiseless Quantum Shannon Channel Coding Theorem.

To communicate quantum information by directly transmitting qubits may be costly. The idea of quantum data compression (QDC) is to ask the question whether we can compress the message into fewer qubits so as to minimize the cost of transmission. Schmacher has provided an answer to achieve this [12]. One excellent review of quantum data compression is by Preskill `http://www.theory.caltech.edu/people/preskill/ph229`, which we follow here.

Let us go straight to the procedure of QDC. Suppose Alice needs to communicate with Bob through some noiseless quantum channel as efficiently as possible, that is, she hopes to compress her message using as few qubits as possible. The message consists of letters represented by some states $|\phi_x\rangle$. Since on average, the frequency of each letter's appearance may not be equal but is some probability p_x, the message can be said to be drawn from an ensemble of states:

$$\{|\phi_x\rangle, p_x\}, \tag{53}$$

so each letter has a density matrix $\boldsymbol{\rho} = \sum_x p_x|\phi_x\rangle\langle\phi_x|$. As we will see in the following, the lowest number of qubits per letter needed to encode is set by the von Neumann entropy $S(\rho) = -{\rm tr}(\rho\log_2\rho)$. If we try to compress into fewer qubits, the fidelity of compression will be ruined.

If the total length of the message is n, then the message has a density matrix which is a direct product of n letter density matrices

$$\boldsymbol{\rho}^{\otimes n} \equiv \underbrace{\boldsymbol{\rho}\otimes\cdots\otimes\boldsymbol{\rho}}_{n\ \rho\text{'s}}. \tag{54}$$

The procedure for quantum data compression goes as follows,
1) Diagonalize $\boldsymbol{\rho}$. Work in the orthonormal basis in which $\boldsymbol{\rho}$ is diagonal. If $\boldsymbol{\rho}$ has eigenvalues (arranged decreasingly) $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$(the number d depends on whether the state $|\phi\rangle$ is a qubit, tri-bit, or d-bit), then $\boldsymbol{\rho}^{\otimes n}$ has eigenvalues of the form (i.e., the eigenvalues are obtained by choosing n values from $\lambda_1, \lambda_2, \ldots, \lambda_d$)

$$\lambda(\{k_i\}) \equiv \prod_{i=1}^{d} \lambda_i^{k_i} = \lambda_1^{k_1} \lambda_2^{k_2} \cdots \lambda_d^{k_d}, \tag{55}$$

where $\sum_i^d k_i = n$, and each eigenvalue $\lambda(\{k_i\})$ occurs $N(\{k_i\}) = n!/(k_1!k_2! \cdots k_d!)$ times. We will restrict ourselves to qubits, i.e., $d = 2$: $\lambda(k_1, k_2 = n - k_1) = \lambda_1^{k_1}\lambda_2^{k_2}$, and $N(k_1, k_2) = C_{k_1}^n$.
2) Given a set of tolerances δ (tolerance for using slightly more qubits than the asymptotically optimal case) and ϵ (tolerance for not projecting onto the "typical" subspace), find the typical subspace Λ and its dual subspace $\Lambda^{\perp}$:

2a) First find out the smallest number $D(n)$ of necessary largest eigenvalues (suppose they are $\lambda_{n,1} \geq \lambda_{n,2} \geq \cdots \geq \lambda_{n,D}$) and corresponding eigenvectors($|\lambda_{n,1}\rangle, |\lambda_{n,2}\rangle, \cdots, |\lambda_{n,D}\rangle$) of $\boldsymbol{\rho}^{\otimes n}$ such that the sum of these eigenvalues is larger than $1 - \epsilon$, but the value $D(n)$ may be larger than $2^{n(S(\rho)+\delta)}$. Increase n and repeat the step until when $n > n_0$

$$D(n) \leq 2^{n(S(\rho)+\delta)}, \tag{56}$$

where n_0 is the smallest number such that the above inequality is satisfied. Note that it is sufficient to use at most some $n(S(\rho) + \delta)$ qubits to represent any state in Λ. This means there exists a unitary transformation $\boldsymbol{U}$ which takes any state $|\phi_\Lambda\rangle$ in Λ to

$$\boldsymbol{U}|\phi_\Lambda\rangle = |\phi_{\text{compressed}}\rangle|0_{\text{rest}}\rangle, \tag{57}$$

where $|\phi_{\text{compressed}}\rangle$ is a state of $n(S(\rho) + \delta)$ qubits, and $|0_{\text{rest}}\rangle$ is a state $|0\rangle \otimes \cdots |0\rangle$ of $(n - n(S + \delta))$ qubits.

2b) Those eigenvectors corresponding to the first $D(n)$ largest eigenvalues span a typical subspace Λ, the remaining spanning a dual subspace $\Lambda^{\perp}$. Note this division into two subspaces can be represented by a projection operator $\boldsymbol{E}$ which projects onto Λ and the complement of which $\boldsymbol{1} - \boldsymbol{E}$ to $\Lambda^{\perp}$. The condition that the sum of eigenvalues of eigenvectors in Λ is larger than $1 - \epsilon$ can be rewritten as

$$\text{tr}(\boldsymbol{\rho}^n \boldsymbol{E}) > 1 - \epsilon. \tag{58}$$

This means states in Λ have much higher overlap with any state drawn from the ensemble than those in $\Lambda^{\perp}$.
3) Prepare the input state $|\psi\rangle = |\phi_1\rangle \cdots |\phi_n\rangle$, where $|\phi_i\rangle$ belongs to the ensemble in the Eq.(53). Make the unitary transformation $\boldsymbol{U}$ on $|\psi\rangle$, and measure the state of the last $(n - n(S + \delta))$ qubits mentioned above. If the result is $|0_{\text{rest}}\rangle$, Alice successfully compresses $|\psi\rangle$ onto $|\psi_{\text{compressed}}\rangle|0_{\text{rest}}\rangle$, and she simply sends $|\psi_{\text{compressed}}\rangle$ to Bob. On the other hand, if Alice gets the results other than $|0_{\text{rest}}\rangle$, she fails to compress her message and the best she can do is send a state $|0'_{\text{compressed}}\rangle$ which is the compressed state corresponding to the largest eigenvector $|\lambda_{n,1}\rangle$ in Λ,

$$\boldsymbol{U}|\lambda_{n,1}\rangle = |0'_{\text{compressed}}\rangle|0_{\text{rest}}\rangle. \tag{59}$$

We note that the input state $|\psi\rangle$ has much higher overlap with states in Λ than any other states in $\Lambda^{\perp}$, the result for Alice to get $|0_{\text{rest}}\rangle$ is of high probability (larger than $1 - \epsilon$).
4)Bob, after receiving $|\psi_{\text{compressed}}\rangle$, appends $|0_{\text{rest}}\rangle$ to it, and applies the inverse unitary transformation $\boldsymbol{U^{-1}}$. On average, Bob receives a density matrix

$$\boldsymbol{\rho}\text{'}^n = \boldsymbol{E}|\psi\rangle\langle\psi|\boldsymbol{E} + |\lambda_{n,1}\rangle\langle\lambda_{n,1}|\langle\psi|(\boldsymbol{1} - \boldsymbol{E})|\psi\rangle. \tag{60}$$

The averaged fidelity $\overline{F}$ of this procedure over the ensemble of possible messages $\{|\psi_i\rangle, p'_i\}$ can be shown to be larger than $1 - 2\epsilon$. It can also be shown that if we try to compress the message into $n(S(\rho) - \delta)$ qubits, the fidelity will be arbitrarily small for sufficient large n.

Finally, we note quantum data compression cannot compress messages drawn from a completely(maximally) mixed state, since $S(\rho_{\text{ completely mixed}}) = 1$.

A. An Example

The example we will discuss shortly is for a small n. From previous discussion, we know for any given δ and ϵ, we can always find a number n_0, such that for any $n > n_0$, the procedure succeeds with the prescibed tolerances. In fact, for a given n, as ϵ becomes smaller, the necessary δ increases, which means we need more qubits to compress, as can be seen from the average number of qubits necessary to encode is $n(S(\rho)+\delta)$. On the other hand, for a given n, as δ decreases (we require fewer qubits to encode), ϵ increases, which means the average fidelity decreases, as can be seen from $\overline{F} > 1-2\epsilon$. Hence, there is some tradeoff between δ, and ϵ for a fixed finite n.

Suppose the ensemble consists of $\{(|H\rangle, p_H = \frac{1}{2}), (|D\rangle, p_D = \frac{1}{2})\}$, where $|H\rangle$ is the state of horizontal polarization while $|D\rangle$ is 45°,

$$|H\rangle = \begin{pmatrix}1\\0\end{pmatrix},\ |D\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}. \tag{61}$$

The density matrix is

$$\rho = \frac{1}{2}|H\rangle\langle H| + \frac{1}{2}|D\rangle\langle D| = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4}\end{pmatrix}, \tag{62}$$

where the matrix is in $|H\rangle$ and $|V\rangle$ basis, and has two eigenvectors and eigenvalues

$$\begin{aligned} |Q\rangle = |22.5^\circ\rangle &= \begin{pmatrix}\cos\frac{\pi}{8}\\ \sin\frac{\pi}{8}\end{pmatrix},\ \lambda_Q = \cos^2\frac{\pi}{8} \\ |\overline{Q}\rangle = |112.5^\circ\rangle &= \begin{pmatrix}-\sin\frac{\pi}{8}\\ \cos\frac{\pi}{8}\end{pmatrix},\ \lambda_{\overline{Q}} = \sin^2\frac{\pi}{8}. \end{aligned} \tag{63}$$

The von Neumann entropy $S(\rho)$ is

$$S(\rho) = -\lambda_Q \log_2 \lambda_Q - \lambda_{\overline{Q}} \log_2 \lambda_{\overline{Q}} \approx 0.60088, \tag{64}$$

so the minimal number of qubits per letter needed to encode is about 0.6009.

Suppose Alice needs to send 3 letters to Bob, but she can afford only two qubits. Since $3 * S(\rho) \approx 1.8$, it's possible to compress 3 letters using only 2 qubits with high fidelity. Note this means $\delta \approx 0.2$ and $D(n=3) = 2$. The eigenvalues and eigenvectors of ρ^3 are

$$\begin{aligned} &\lambda_1 = \cos^3\tfrac{\pi}{8},\ \lambda_2 = \lambda_3 = \lambda_4 = \cos^2\tfrac{\pi}{8}\sin\tfrac{\pi}{8}, \\ &\lambda_5 = \lambda_6 = \lambda_7 = \cos\tfrac{\pi}{8}\sin^2\tfrac{\pi}{8},\ \lambda_8 = \sin^3\tfrac{\pi}{8}. \end{aligned}$$

$$\begin{aligned} &|1\rangle = |QQQ\rangle,\ |2\rangle = |QQ\overline{Q}\rangle,\ |3\rangle = |Q\overline{Q}Q\rangle,\ |4\rangle = |\overline{Q}QQ\rangle, \\ &|5\rangle = |Q\overline{Q}\overline{Q}\rangle,\ |6\rangle = |\overline{Q}Q\overline{Q}\rangle,\ |7\rangle = |\overline{Q}\overline{Q}Q\rangle,\ |8\rangle = |\overline{Q}\overline{Q}\overline{Q}\rangle. \end{aligned} \tag{65}$$

The subspace Λ is spanned by $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}$, while its dual subspace $\Lambda^\perp$ is by $\{|5\rangle, |6\rangle, |7\rangle, |8\rangle\}$, and

$$\begin{aligned} P_\Lambda &\equiv \mathrm{tr}(\rho^{\otimes 3}\boldsymbol{E}) = \sum_{i=1}^{4}\lambda_i \approx 0.9419 \\ P_\Lambda^\perp &\equiv \mathrm{tr}(\rho^{\otimes 3}(\mathbf{1}-\boldsymbol{E})) = \sum_{i=5}^{8}\lambda_i \approx 0.0581. \end{aligned} \tag{66}$$

Alice and Bob both agree on the form of the unitary transformation that they will use,

$$\mathrm{U}\begin{pmatrix}|1\rangle\\|2\rangle\\|3\rangle\\|4\rangle\end{pmatrix} \rightarrow \begin{pmatrix}|HHH\rangle\\|HVH\rangle\\|VHH\rangle\\|VVH\rangle\end{pmatrix} \quad \mathrm{U}\begin{pmatrix}|5\rangle\\|6\rangle\\|7\rangle\\|8\rangle\end{pmatrix} \rightarrow \begin{pmatrix}|HHV\rangle\\|HVV\rangle\\|VHV\rangle\\|VVV\rangle\end{pmatrix}; \tag{67}$$

this transformation is unitary since it is simply the transformation between two sets of orthonormal bases of 3 qubits.

Alice prepares her message in a state $|\psi\rangle$, which can be expanded in the basis of $|1\rangle \cdots |8\rangle$,

$$|\psi\rangle = \sum_{i=1}^{8} a_i |i\rangle, \tag{68}$$

where from Eq.(66) we have

$$\sum_{i=1}^{4}|a_i|^2 = P_\Lambda \gg \sum_{i=5}^{8}|a_i|^2 = P_\Lambda^\perp. \tag{69}$$

Then Alice applies the unitary transformation $\boldsymbol{U}$ on $|\psi\rangle$ followed by a measurement on the third qubit. If the result is $|H\rangle$, she sucessfully projects $|\psi\rangle$ into the likely subspace Λ. At this stage, the total state is

$$a_1|HHH\rangle + a_2|HVH\rangle + a_3|VHH\rangle + a_4|VVH\rangle = |\psi_{\text{compressed}}\rangle|H\rangle, \tag{70}$$

where $|\psi_{\text{compressed}}\rangle \equiv a_1|HH\rangle + a_2|HV\rangle + a_3|VH\rangle + a_4|VV\rangle$. She simply sends this two-qubit state $|\psi_{\text{compressed}}\rangle$ to Bob. Upon receiving $|\psi_{\text{compressed}}\rangle$, Bob appends a third qubit $|H\rangle$ to it, and does the inverse transformation $\boldsymbol{U^{-1}}$ to get

$$|\psi'\rangle = \boldsymbol{U^{-1}}(|\psi_{\text{compressed}}\rangle|H\rangle) = \sum_{i=1}^{4} a_i|i\rangle, \tag{71}$$

which has high resemblance to the initial $|\psi\rangle$, that is

$$F_1 \equiv |\langle\psi|\psi'\rangle|^2 = P_\Lambda \approx 0.9419. \tag{72}$$

On the other hand, if, when Alice measures the third qubit and gets $|V\rangle$, she fails to project $|\psi\rangle$ into Λ but $\Lambda^\perp$ instead, the best she can do is send a qubit state $|HH\rangle$. After Bob receives it and decompresses it, he gets

$$|\psi''\rangle = \boldsymbol{U^{-1}}(|HH\rangle|H\rangle) = |1\rangle, \tag{73}$$

which has overlap with the initial $|\psi\rangle$:

$$F_2 \equiv |\langle\psi|\psi''\rangle|^2 = |a_1|^2 = \lambda_1 \approx 0.6219. \tag{74}$$

The fidelity of this procedure is $F = P_\Lambda F_1 + P_\Lambda^\perp F_2 = 0.9234$.

How good is this? Let us compare it to the case when Alice sends the first two letters without compressing and asks Bob to guess the third letter. Since both $|H\rangle$ and $|D\rangle$ from the ensemble have higher overlap with $|Q\rangle$ than with $|\overline{Q}\rangle$, the best guess he can make is $|Q\rangle$. The fidelity of this procedure is

$$F = \frac{1}{2}|\langle H|Q\rangle|^2 + \frac{1}{2}|\langle D|Q\rangle|^2 = 0.8535, \tag{75}$$

which is smaller than the case when we do compression.

XV. CONCLUDING REMARKS

In this unit, we have discussed entanglement of quantum states, entanglement of formation and distillation, entanglement entropy, Schmidt decomposition, majorization, quantum Shannon theory, etc.

It is a good time to check whether you have achieved the following Learning Outcomes:
After this Unit, you'll be able to understand the basics of quantum information and entanglement theory.

Suggested reading: N&C 12.2, 12.5. R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," Rev. Mod. Phys. 81, 865 (2009) [28] at `https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.81.865`.

---

[1] R. F. Werner, Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).
[2] A. Peres, Separability criterion for density matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[3] In general, if a state has PPT, no entanglement can be distilled out from it [10]. Horodecki et al. showed that the PPT criterion is also sufficient for 2×2 (two qubits) and 2×3 (a qubit and a qutrit) [4]. But the state can be either unentangled or entangled. When the state has PPT and is also entangled, it is called a bound entangled state.
[4] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A **223**, 1 (1996).
[5] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Volume of the set of separable states, Phys. Rev. A **58**, 883 (1998).
[6] G. Vidal and R. F. Werner, Computable measure of entanglement, Phys. Rev. A **65**, 032314 (2002).
[7] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).
[8] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, Phys. Rev. Lett. **76**, 722 (1996).
[9] M. Horodecki, P. Horodecki, and R. Horodecki, Inseparable two spin- $\frac{1}{2}$ density matrices can be distilled to a singlet form, Phys. Rev. Lett. **78**, 574 (1997).
[10] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature?, Phys. Rev. Lett. **80**, 5239 (1998).
[11] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).
[12] B. Schumacher, Quantum coding, Phys. Rev. A **51**, 2738 (1995).
[13] W. K. Wootters, Entanglement of formation of an arbitrary state of two qubits, Phys. Rev. Lett. **80**, 2245 (1998).
[14] B. M. Terhal and K. G. H. Vollbrecht, Entanglement of formation for isotropic states, Phys. Rev. Lett. **85**, 2625 (2000).
[15] K. G. H. Vollbrecht and R. F. Werner, Entanglement measures under symmetry, Phys. Rev. A **64**, 062307 (2001).
[16] P. W. Shor, Equivalence of additivity questions in quantum information theory, Communications in Mathematical Physics **246**, 453 (2004).
[17] V. Vedral and M. B. Plenio, Entanglement measures and purification procedures, Phys. Rev. A **57**, 1619 (1998).
[18] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying entanglement, Phys. Rev. Lett. **78**, 2275 (1997).
[19] M. Horodecki, P. Horodecki, and R. Horodecki, Limits for entanglement measures, Phys. Rev. Lett. **84**, 2014 (2000).
[20] M. A. Nielsen, Conditions for a class of entanglement transformations, Phys. Rev. Lett. **83**, 436 (1999).
[21] D. Jonathan and M. B. Plenio, Entanglement-assisted local manipulation of pure quantum states, Phys. Rev. Lett. **83**, 3566 (1999).
[22] T.-C. Wei and P. M. Goldbart, Geometric measure of entanglement and applications to bipartite and multipartite quantum states, Phys. Rev. A **68**, 042307 (2003).
[23] A. Shimony, Degree of entanglement a, Annals of the New York Academy of Sciences **755**, 675 (1995).
[24] G. Vidal, Entanglement monotones, Journal of Modern Optics **47**, 355 (2000).
[25] This requirement does not contradict with distillation, as it takes into account the cases when distillation fails.
[26] D. Gross, S. T. Flammia, and J. Eisert, Most quantum states are too entangled to be useful as computational resources, Phys. Rev. Lett. **102**, 190501 (2009).
[27] M. J. Bremner, C. Mora, and A. Winter, Are random pure states useful for quantum computation?, Phys. Rev. Lett. **102**, 190502 (2009).
[28] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).