33-758 QCQI
A Term Paper on Entanglement Distillation

Arun Kannawadi Jayaraman

May 11, 2012

Abstract

'Entanglement distillation' is the procedure of converting a number of non-maximally entangled states into a smaller number of maximally entangled pairs , which can then be used in Quantum Information processes such as Teleportation and Dense coding. Two separated observers, by communicating classically and applying local operations to a supply of not-too-impure entangled states (e.g., singlets shared through a noisy channel), can prepare a smaller number of entangled pairs of arbitrarily high purity (e.g., near-perfect singlets). The term paper will illustrate under what conditions this can be done and how it can be done.

Organization of the paper

The paper begins with a brief introduction to Quantum Mechanics and to Entanglement in Quantum Mechanics. To be able to define 'arbitrarily high purity' states and 'not-too-impure' entangled states, we first need to know how entanglement is measured. We throw some light on different measures that are used such as Von Neumann entropy of the subsystems, Entanglement of Formation, Entanglement of Distillation etc. as measures of entanglement.

Having set the stage, the paper then moves on to give the mathematical details that will be required to distill entanglement from non-maximally entangled pure states. We shall then explain the procedure of 'concentrating' the entanglement present in n pairs of particles in identical, partly entangled pure and mixed states into a smaller number of maximally entangled pairs of particles by means of local operations. We shall also see that this process asymptotically conserves the *entropy of entanglement*. In this paper, we also make a brief mention about entanglement witnesses that can be used to decide whether a particular mixed state is suitable for entanglement distillation or not.

1 Introduction

We very well know now that in Quantum Mechanics, the state of an isolated system (say A) is described by a vector in a Hilbert Space, denoted by $\mathcal{H}_A$. The state of a collection of systems ($A_1, A_2, \ldots, A_N$), isolated from the rest of the universe, is described by a vector in a Hilbert space which is a tensor product of the Hilbert spaces of the individual systems. It is denoted as $\mathcal{H} = \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \ldots \otimes \mathcal{H}_{A_N}$.

The rule for the tensor product is, the basis vectors of the product Hilbert space $\mathcal{H}$ is given by the tensor products of the basis of the individual Hilbert spaces. For example, consider a qubit whose state vectors lie in $\mathcal{H}_1$ such that $dim(\mathcal{H}_1) = 2$ and let the basis vectors be represented as $|0\rangle$and $|1\rangle$. A general state of the qubit is then $|\psi_1\rangle = a|0\rangle + b|1\rangle$. Now consider another qubit which is isolated from the first qubit and the rest of the universe. Its Hilbert space $\mathcal{H}_2$ is isomorphic to $\mathcal{H}_1$. Let the state of the second qubit be represented as $|\psi_2\rangle = c|0\rangle + d|1\rangle$. The vector in $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is then given by the *product state* $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = ac|00\rangle + ad|01\rangle + bc|10\rangle + bd|11\rangle$. Note that the product of the first and the last coefficients is same as that of the middle ones. But it is not necessary that a general vector $\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$ in $\mathcal{H}$ must satisfy this relation. In general $\alpha\delta - \beta\gamma \neq 0$. The two qubits whose states are represented by such a vector are then not isolated from each other and are said to be entangled.

But a state vector is not the most general quantum description of a system. Only *pure* systems are described by vectors in a Hilbert space. It is possible that we may be interested in a *statistical ensemble* of pure states or the system under consideration may be a part of a bigger, isolated system with the other part being unaccessible or uninteresting. Such systems are said to be in a *mixed state.* A system in a pure state could easily evolve interacting with the enviroment and *decohere* giving rise to a mixed state. The description of such a state is given by an operator on the Hilbert space which is called *density operator* or a *density matrix.*

Consider an ensemble of systems in pure states. The system might be in a state $|\psi_i\rangle$ with a probability p_i in which case, the corresponding density operator is given as $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. It follows that the density operator for a pure state $|\psi\rangle$would then be just $|\psi\rangle\langle\psi|$. Thus, we see that the density operator is a more general description of a quantum state.

Any operator ρ that satisfy the following conditions is a valid density operator:

- ρ must be a Hermitian operator i.e. $\rho^\dagger = \rho$.
- ρ must be a positive operator i.e. the eigenvalues of ρ must be non-negative.
- $Tr\,(\rho) = 1$.

We can differentiate the pure states from mixed states by taking the trace of the square of the density operator. For a mixed state, $Tr\left(\rho^2\right) < 1$while for pure states, $Tr\left(\rho^2\right) = 1$.

It is necessary to be able to characterize entanglement in terms of the density operators. We first define a large family of states, known as *separable states*, first defined by Werner in 1989. A separable state, by definition, is of the form $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ with $\sum_i p_i = 1$. It is easy to generalize this form for a system consisting of N subsystems.

All states that are not separable are then called as *entangled states*. One of the reasons for calling them so is a separable state can always be prepared by local operations and classical communication (LOCC). The subsystems cannot have any non-classical correlations in this case. Another reason for defining these to be entangled will come later. Note that this is more general than the product states of the form $\rho_A \otimes \rho_B$.

2 Entanglement

We defined in Section 1 what an entanglement is. Without going much into the details and rigorous math, we shall quickly see what entanglement between two qubits can buy for us. In this section, we shall be consider the entanglement only between two qubits.

In 1935, Einstein, Podolsky and Rosen [8] pointed out a strange property (entanglement) in Quantum Mechanics and argued that the quantum theory must be incomplete. This is popularly known as the EPR paradox. John S. Bell [9] in 1964 formulated their argument mathematically and came up with an inequality. A modified form of Bell's inequality goes by the name CHSH inequality [10]. These inequalities are verifiable experimentally and the 4 *Bell states*, listed below, are found to violate this inequality maximally:

$$|\beta_0\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)$$
$$|\beta_1\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right)$$
$$|\beta_2\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right)$$
$$|\beta_3\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right)$$

2.1 Entanglement as a resource

We shall begin explaining the process of distillation by motivating the reader first with the importance of distillation. Quantum Teleportation and Super-dense coding, discussed in [1], require Alice and Bob to share a Bell state. Teleportation is unfaithful when the shared state is not a Bell state (maximally entangled one). Having a non-maximally entangled states is equivalent to a noisy channel and that introduces errors while transmitting the code as qubits. This is one of the reason why we call a non-maximally entangled state as an *impure state* or an *impure channel* and a maximally entangled state as a *pure channel*. It is more efficient to employ distillation to purify the quantum channel rather than the complicated quantum error-correction schemes.

This is the first statement about the quantity of entanglement. We know intuitively that the product states (or more generally separable states) must have zero entanglement and the Bell states must have the maximum entanglement, lets say 1 in some appropriate units. How do we quantify entanglement in other states?

2.2 Measures of Entanglement

This subsection is taken from [6]. Consider a bipartite system $A-B$. Any measure of entanglement, E, between A and B must satisfy the following properties:

1. For any separable state σ, $E(\sigma) = 0$.
2. E must be invariant under local unitary operations
   i.e. $E\left((U_A \otimes U_B)\,\sigma\left(U_A^\dagger \otimes U_B^\dagger\right)\right) = E(\sigma)$.
3. E must not increase under local unitary operations and classical communication

2.2.1 Von Neumann entropy as a measure of entanglement

Without mathematical proof, we state that any bipartite pure state can be represented in a *Schmidt decomposed* form as $|\psi_{AB}\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$ with $\sum_i \lambda_i^2 = 1$ and without loss of generality, $\lambda_i \geq 0\ \forall i$. The kets of the individual qubits satisfy the orthonormality condition i.e. $\langle i_{A,B} | i'_{A,B} \rangle = \delta_{i,i'}$. This is the most compact form in which the two qubit pure state $|\psi_{AB}\rangle$ could be written in terms of the vectors in the Hilbert space of the individual qubits. The number of non-zero *Schmidt coefficients* λ_is are known as the *Schmidt number* of the state. When the two qubits are unentangled and in a pure state, $|\psi_{AB}\rangle$ is a product state with the Schmidt number being 1 and the modulus of the Schmidt coefficient is unity. For the Bell states, the Schmidt number is 2.

The Schmidt coefficients are invariant under local unitary transformations.

$$(U_A \otimes U_B) \left(\sum_i \lambda_i |i_A\rangle \otimes |i_B\rangle \right) = \sum_i \lambda_i \left(U_A |i_A\rangle \right) \otimes \left(U_B |i_B\rangle \right)$$

It is natural therefore, to think that the measure of entanglement for a bipartite pure state could only be a function of the Schmidt coefficients. We would expect E to be a continuous function of the λ_i s and in addition, we require it to possess the following additive property:

$$E\left(|\Psi_{AB}\rangle \otimes |\Phi_{CD}\rangle \right) = E\left(|\Psi_{AB}\rangle \right) + E\left(|\Phi_{CD}\rangle \right)$$

where $|\Phi_{CD}\rangle$ represents the (pure) state of the bipartite system $C - D$ that is unentangled with $A - B$. These conditions are reminiscent of the conditions imposed on Shannon's entropy. Therefore, the unique measure of entanglement for a pure state $|\psi_{AB}\rangle$ is given by

$$E\left(|\psi_{AB}\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \right) = -\sum_i |\lambda_i|^2 \log\left(|\lambda_i|^2 \right)$$

For pure states, all measures of entanglement are same as Von Neumann entropy. For mixed states, there are a number of measures of entanglement and the values assigned for the amount of entanglement by different measures differ from one another.

In order to be able to come up with a similar measure of entanglement for mixed states, we need to express this in terms of the density operators. We note that $|\lambda_i|^2$s appear as the diagonal elements in the reduced density operator of A defined as $\rho_A = Tr_B\left(\rho_{AB} \right) = \sum_j \langle j_B | \psi_{AB} \rangle \langle \psi_{AB} | j_B \rangle$ where Tr_B refers to the partial trace over the system B. Since the ket vectors of B in the Schmidt decomposition form a basis in $\mathcal{H}_B$ (assuming $dim\left(\mathcal{H}_A \right) \geq dim\left(\mathcal{H}_B \right)$; if not label the subsystem in the higher dimensional Hilbert space as A), we can use these vectors to take the partial trace of ρ_{AB} . In the $\{|i_A\rangle\}$ basis, ρ_A is diagonal with diagonal elements being $|\lambda_i|^2$. Hence we can immediately write

$$E\left(|\psi_{AB}\rangle \right) = -\sum_i |\lambda_i|^2 \log\left(|\lambda_i|^2 \right) = -Tr\left(\rho_A \log\left(\rho_A \right) \right) \equiv S\left(\rho_A \right) = E\left(\rho_{AB} \right)$$

where $S\left(\rho_A \right)$, defined as above, is called the Von Neumann entropy of the subsystem A . Incidentally, it is also equal to the Von Neumann entropy of the subsystem B. Thus, for pure states, $S\left(\rho_A \right) = S\left(\rho_B \right) = E(\rho_{AB})$ can serve as a measure of entanglement. When A and B are unentangled, $S\left(\rho_A \right) = S\left(\rho_B \right) = 0$ since there exists only one Schmidt coefficient whose modulus squared is unity. For Bell states, $S\left(\rho_A \right) = S\left(\rho_B \right) = 1$ which is the maximum that the function $S\left(\right)$can take.

Convexivity of density operators

We shall first discuss about an important property of density operators before we head out to discuss about the entanglement in mixed states. If ρ_1 and ρ_2 are two density operators acting on some Hilbert space $\mathcal{H}$, then $\rho = \alpha\rho_1 + (1-\alpha)\rho_2$ is also a density operator acting on $\mathcal{H}$ for $0 \leq \alpha \leq 1$. More generally, if $\rho_1, \rho_2, \ldots \rho_n$ are density operators acting on $\mathcal{H}$ then $\rho = \alpha_1\rho_1 + \alpha_2\rho_2 + \ldots + \alpha_n\rho_n$ is also a valid density operator provided $\alpha_i \geq 0 \; \forall i$ and $\sum_i \alpha_i = 1$. Conversely any density operator ρ can be expressed as a convex combination of any arbitrary number of density operators.

2.2.2 Entanglement of Formation (for mixed states)

Now consider a mixed state ρ_{AB} of two subsystems A and B. This state can be represented in one or more ways as a convex sum of pure states of A and B as $\rho_{AB} = \sum_i \alpha_i \rho_{AB}^i$ where ρ_{AB}^i s are the pure states of $A-B$. Then we may want to define $E_F(\rho_{AB}) = \sum_i \alpha_i E_F(\rho_{AB}^i) = \sum_i \alpha_i S(\rho_A^i)$. But different realizations of the sum can give rise to different values for $E_F(\rho_{AB})$. It is even possible to get a non-zero entanglement for product states if we defined entanglement for mixed states this way. In order to avoid this ambiguity and to satisfy the condition that entanglement must not increase under LOCC, we minimize the sum over all possible realizations. Therefore,

$$E_F(\rho_{AB}) = \min \sum_i \alpha_i S\left(\rho_A^i\right)$$

The reason for calling it the entanglement of 'formation' is, suppose two persons, Alice and Bob want to create an ensemble of n copies of non-maximally entangled state ρ_{AB} using only local operations and classical communication (LOCC) and m maximally entangled pairs, then (without proof)

$$\lim_{m,n\to\infty} \frac{m}{n} = E_F(\rho_{AB})$$

The above is a statement of conservation of entanglement in the asymptotic limit. An elegant explanation as to why the asymptotic limit is required is given in [1] (pg. 578). The entanglement of formation of a bipartite mixed state can be defined in terms of the number of pure singlets (Bell state $|\beta_3\rangle$) needed to create the state with no further transfer of quantum information.

2.2.3 Concurrence

Finding the entanglement of formation for a generic bipartite system is extremely hard because it involves a minimization over continously infinite number of realizations. But for a pair of qubits in a mixed state, we have a closed form expression given by W. K. Wooters in [3]. We define the function

$$\begin{aligned}\mathcal{E}(x) &= H\left(\frac{1-\sqrt{1-x^2}}{2}\right) \\ &= -\frac{1-\sqrt{1-x^2}}{2}\log_2\left(\frac{1-\sqrt{1-x^2}}{2}\right) - \frac{1+\sqrt{1-x^2}}{2}\log_2\left(\frac{1+\sqrt{1-x^2}}{2}\right)\end{aligned}$$

where $H()$is the binary entropy function ($H(p) = -p\log_2 p - (1-p)\log_2(1-p)$).

Concurrence of a bipartite mixed state ρ is defined as $C(\rho) = \max\{0, +\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}$ where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the eigenvalues of the matrix R defined as follows in the decreasing order. By definition, $R := \sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$ with $\tilde{\rho} = (\sigma_y \otimes \sigma_y)\,\rho *\, (\sigma_y \otimes \sigma_y)$ where the complex conjugation is taken in the basis $\{|\beta_0\rangle, i|\beta_1\rangle, i|\beta_2\rangle, |\beta_3\rangle\}$.

Having defined all the required quantitites, the entanglement of formation for two qubits in a mixed state ρ is given as

$$E_F(\rho) = \mathcal{E}(C(\rho))$$

Since $0 \leq \frac{1-\sqrt{1-C(\rho)^2}}{2} \leq \frac{1}{2}$, $\mathcal{E}(C(\rho))$ is a monotonic function of $C(\rho)$. Interestingly, when $C(\rho) = 0$, $\mathcal{E}(C(\rho)) = 0$ and when $C(\rho) = 1$, $\mathcal{E}(C(\rho)) = 1$. Thus, concurrence itself can be used as a measure of entanglement.

The aim of this subsection is basically to illustrate that finding the entanglement of formation for an arbitrary bipartite system is a hopeless task, given that finding the entanglement of formation for the simplest bipartite system in a closed form is a Herculean task by itself.

2.2.4 Entanglement of distillation

Distillation is the opposite process of formation. Suppose Alice and Bob are provided with a large number (n) copies of a state ρ, and they want to convert them into as many copies as possible (m) of some Bell state, say $|\beta_0\rangle$ with arbitrarily high *fidelity* using local operations and classical communication (LOCC). We have not formally defined fidelity yet, but intuitively, we can understand as the closeness of two quantum states.

Fidelity: For two pure states $|\psi\rangle$and $|\phi\rangle$, fidelity is defined as $F(\psi, \phi) := |\langle\psi|\phi\rangle|$ or sometimes as $|\langle\psi|\phi\rangle|^2$.

For mixed states ρ and σ, $F(\rho, \sigma) := Tr\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)$ or sometimes as its square.

Entanglement of distillation, also known as distillable entanglement, is defined as

$$E_D(\rho) = \lim_{m,n\to\infty} \frac{m}{n}$$

For pure states, $E_D(\rho) = E_F(\rho) = S(\rho_A) = S(\rho_B)$ where ρ_A and ρ_B represent the reduced density operators.

For mixed states, $E_D(\rho) < E_F(\rho)$. This implies that formation of states, in some sense, is an irreversible process. The reason for this is, there exist states ρ that are not separable but have $E(\rho) = 0$. Such states are said to have *bound entanglement* that cannot be distilled or concentrated. This is very analogous to the concept of *exergy* in thermodynamics, the useful work that can be obtained from a system. In the following section, we shall understand the theory and mathematics that is necessary for distillation to be possible and quickly analyze the protocols used in distillation process.

3 Distillation of pure states

We shall begin from the easiest of the starting points. In this section, we consider the case where the qubit pairs shared between Alice and Bob are identical and non-maximally entangled pure states.

3.1 Definitions

All the definitions that follow in this subsection are from [1].

Typical & Atypical sequences

Suppose we have n random variables $X_1, X_2, \ldots, X_n$which are independent and have identical probability distributions (i.i.d random variables). Let X_i be 0 with a probability p and 1 with a probability $1-p$. We form sequences $x_1x_2\ldots x_n$ from the values taken by the random variables $X_1, X_2, \ldots, X_n$ and classify them into two types of sequences namely typical and atypical sequences. From the law of large numbers, we expect that as n gets large, with very high probability, a fraction p of the sequence will be equal to zero and a fraction $1-p$ of the sequence will be equal to 1. The reason for this being, the largest term in the expansion of $(p+(1-p))^n$ is $\binom{n}{np}p^{np}(1-p)^{(1-p)n}$ and is much larger than the other terms.

The sequences $x_1x_2\ldots x_n$ that satisfy this assumption are known as *typical* sequences and those that don't are known as *atypical* sequences.

For typical sequences, $p(x_1x_2\ldots x_n) = p(x_1)p(x_2)\ldots p(x_n) \approx p^{np}(1-p)^{(1-p)n}$. Taking logarithms on both sides, we get

$$-\log_2 p(x_1x_2\ldots x_n) \approx -np\log_2 p - n(1-p)\log_2(1-p) = nH(X)$$

where X is a random variable having the same probability distribution as that of X_i s and $H(X)$ is the (binary) entropy of the random variable X . $H(X) = -p\log_2 p - (1-p)\log_2(1-p)$. There are $2^{nH(X)}$ typical sequences and the probability of getting one of the typical sequences is $2^{-nH(X)}$. This means, if we restrict ourselves to a set of typical sequences rather than a complete of all possible binary sequences of length n, $nH(X)$ bits are enough to represent the typical sequences uniquely. This is one of the techniques used in data compression.

ϵ - typical sequences

We now relax the criterion a little and expand the set of sequences under consideration.

Given $\epsilon > 0$, we say that a binary string $x_1x_2\ldots x_n$ is ϵ - typical if

$$2^{-n(H(X)+\epsilon)} \leq p(x_1x_2\ldots x_n) \leq 2^{-n(H(X)-\epsilon)}$$

and denote the set of all such ϵ - typical sequences of length n by $T(n,\epsilon)$. A useful reformulation of this definiton is

$$|\frac{1}{n}\log_2(p(x_1x_2\ldots x_n)) + H(X)| \leq \epsilon.$$

Theorem of typical sequences (without proof)

1. Fix $\epsilon > 0$. Then for any $\delta > 0$, for sufficiently large n , the probability that a sequence is ϵ - typical is at least $1-\delta$.

2. For any fixed $\epsilon > 0$ and $\delta > 0$, for sufficiently large n , the number of ϵ - typical sequences, $|T(n,\epsilon)|$ satisfies
$$(1-\delta)2^{n(H(X)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(H(X)+\epsilon)}.$$

3. Let $S(n)$ be a collection of size at most 2^{nR}, of length n binary sequences where $R < H(X)$ is fixed. Then for any $\delta > 0$ and for sufficently large n,
$$\sum_{x\in S(n)} p(x) \leq \delta.$$

Majorization

Majorization is an *ordering* on d - dimensional real vectors that is intended to capture the notion that one vector is more or less disordered than another. Theory of majorization is an area of

mathematics that predates quantum mechanics. We shall first understand majorization and then establish its connection with quantum entanglement.

Consider two d - dimensional vectors $\underline{x} = \begin{pmatrix} x_1 & x_2 & \dots & x_d \end{pmatrix}$ and $\underline{y} = \begin{pmatrix} y_1 & y_2 & \dots & y_d \end{pmatrix}$. We shall follow the notation used in [6]. $\underline{x}^{\downarrow}$ is used to represent the vector $\underline{x}$ that is re-ordered so that the components are in non-increasing order. $x_1^{\downarrow}$ denotes the largest component of $\underline{x}$ and $x_d^{\downarrow}$ denotes the smallest component of $\underline{x}$. We say $\underline{x}$ is *majorized* by $\underline{y}$, if $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$ for $k = 1, \dots, d-1$ and strictly equal for $k = d$. We write this mathematically as $\underline{x} \prec \underline{y}$.

The relation between majorization and entanglement is easy to state and surprising, but involves deeper mathematics to prove. Interested reader can refer to Sec 12.5.1 in [1]. If the channel shared by Alice and Bob is in a pure state $|\psi\rangle$ and let us assume that they would like to transform the state to some $|\varphi\rangle$. We define $\rho_\psi \equiv Tr_B\left(|\psi\rangle\langle\psi|\right)$ and $\rho_\varphi \equiv Tr_B\left(|\varphi\rangle\langle\varphi|\right)$ to be the corresponding reduced density operators on the Alice's end. Let λ_ψ and λ_ϕ be the vectors whose entries are the eigenvalues of ρ_ψ and ρ_φ respectively. Then $|\psi\rangle$ may be transformed to $|\varphi\rangle$ by local operations and classical communication *if and only if* $\lambda_\psi \prec \lambda_\varphi$.

3.2 Distillation of pure states - Theory

Consider Alice and Bob sharing n pairs of qubits, each pair being in some arbitrary pure state $|\psi\rangle$. Let the state $|\psi\rangle$ have a Schmidt decomposition

$$|\psi\rangle = \sum_x \sqrt{p(x)}|x_A\rangle|x_B\rangle.$$

For a pair of qubits, the Schmidt number can be only 1 or 2. This follows from the fact that the kets $|x_{A,B}\rangle$ have to be orthogonal in the Hilbert space(s) $\mathcal{H}_{A,B}$ and $\dim\left(\mathcal{H}_{A,B}\right) = 2$. Therefore, it is not possible to have more than two orthogonal kets and hence more than two Schmidt coefficients. Without loss of generality, we can label the two orthonormal kets as $|0_A\rangle$ and $|1_A\rangle$ and similarly $|0_B\rangle$ and $|1_B\rangle$. Note that it is not necessary that the set $\{|0_A\rangle, |1_A\rangle\}$ is same as $\{|0_B\rangle, |1_B\rangle\}$. Additionally, $p\left(0_A\right) = 1 - p\left(1_A\right) = p\left(0_B\right) = 1 - p\left(1_B\right)$.

The point of the exercise above is to emphasize that we have a binary nature for the kets in $\mathcal{H}_A$ and $\mathcal{H}_B$.

Since Alice and Bob have n pair of qubits, the quantum channel between them is represented by the vector

$$|\psi\rangle^{\otimes n} = \sum_{x_1 x_2 \dots x_n} \sqrt{p(x_1)p(x_2)\dots p(x_n)}|x_{1A}x_{2A}\dots x_{nA}\rangle|x_{1B}x_{2B}\dots x_{nB}\rangle.$$

The state possessed by Alice and Bob are denoted as ρ_A and ρ_B . Then,

$$\rho_A = \rho_B = \sum_{x_1 x_2 \ldots x_n} p(x_1)p(x_2)\ldots p(x_n)|x_1 x_2 \ldots x_n\rangle\langle x_1 x_2 \ldots x_n| = \rho^{\otimes n} \equiv \rho_\psi.$$

where $\rho = p(0)|0\rangle\langle 0| + p(1)|1\rangle\langle 1|$ and ρ_ψ is the result of tracing out the other's qubits.

Now define a states $|\varphi_n'\rangle$ by omitting all those sequences of $x_1 x_2 \ldots x_n$ that are not ϵ- typical for some $\epsilon > 0$. Therefore,

$$|\varphi_n'\rangle \equiv \sum_{\underline{x} \in T(n,\epsilon)} \sqrt{p(x_1)p(x_2)\ldots p(x_n)}|x_{1A}x_{1B}\ldots x_{nB}\rangle|x_{1B}x_{2B}\ldots x_{nB}\rangle,$$

where $\underline{x}$ is a short hand to represent the sequence $x_1 x_2 \ldots x_n$. Note that $|\varphi_n'\rangle$ is not a properly normalized quantum state. Therefore, we normalize $|\varphi_n'\rangle$ by redefining $|\varphi_n\rangle \equiv \frac{1}{\sqrt{\langle\varphi_n'|\varphi_n'\rangle}}|\varphi_n'\rangle$.

By statement 1 of the theorem of typical sequences, $F\left(|\psi\rangle^n, |\varphi_n\rangle\right) \to 1$ as $n \to \infty$. A projection onto the ϵ - typical subspace of $\rho^{\otimes n}$ will succeed with arbitrarily high probability. Alice and Bob can perform local projections onto their 'typical' subspace and transform the state of the quantum channel to $|\varphi_n\rangle$. There is a vanishingly small probability that the channel may get projected into the orthogonal subspace in which case, they have to be supplied with a fresh set of n pairs of qubits, each in the state $|\psi\rangle$ again.

By statement 2 of the theorem of typical sequences, the number of terms in $|\varphi_n\rangle$ is at most $2^{n(H(p(x)+\epsilon)} = 2^{n(S(\rho_\psi)+\epsilon)}$, where $S\left(\rho_\psi\right)$ is the Von Neumann entropy of ρ_ψ . The largest Schmidt coefficient appearing in $|\varphi_n'\rangle$ is at most $2^{-n(S(\rho_\psi)-\epsilon)}$ by the definition of typical sequences. The re-normalized state $|\varphi_n\rangle$ can have Schmidt coefficients larger by at most a factor $1/\sqrt{1-\delta}$, since by the theorem of typical sequences , $1-\delta$ is a lower bound on the probability that a sequence is ϵ - typical and may be made arbitrarily close to 1 for sufficiently large n . Therefore, the largest eigenvalue of the state ρ_{φ_n} is at most $2^{-n(S(\rho_\psi)-\epsilon)}/(1-\delta)$. Suppose we choose an m such that

$$\frac{2^{-n(S(\rho_\psi)-\epsilon)}}{1-\delta} \leq 2^{-m}.$$

Then the vector of eigenvalues of ρ_{φ_n} , denoted as λ_{φ_n}, is majorized by the vector $\begin{pmatrix} 2^{-m} & 2^{-m} & \ldots & 2^{-m} \end{pmatrix}$. The vector $\begin{pmatrix} 2^{-m} & 2^{-m} & \ldots & 2^{-m} \end{pmatrix}$ is indeed λ_{β_0} , the eigenvalues of $\rho_{\beta_0} \equiv Tr_B\left(\left(|\beta_0\rangle\langle\beta_0|\right)^{\otimes m}\right)$. Hence, as stated earlier, it is possible to convert the state $|\varphi_n\rangle$ by local operations and classical communication. Taking logarithm on both sides of the equation above, we get

$$-n\left(S\left(\rho_\psi\right) - \epsilon\right) - \log_2\left(1-\delta\right) \geq -m$$

Note the reversal of the inequality sign since we are taking the logarithm of quantities that are less

than 1. For large values of m and n , we can neglect the $\log_2(1-\delta)$ term. Rearranging, we have

$$nS(\rho_\psi) - n\epsilon \leq m$$

Therefore, $\frac{m}{n} \geq S(\rho_\psi)$ and as $n \to \infty$, $\epsilon \to 0$. So

$$\lim_{m,n\to\infty} \frac{m}{n} \equiv E_D(|\psi\rangle) = S(\rho_\psi).$$

Thus for bipartite pure states, the entanglement of formation, the entanglement of distillation and the Von Neumann entropy of the reduced density operators are all same. Interested reader may obtain more information related to this part of the paper from [1].

3.3 Distillation of pure states - Protocol

This subsection is a summary of [4]. Let the state of the n qubit pairs, each in the state $|\psi_{AB}\rangle$, be represented in its Schmidt decomposed form as

$$|\psi_{AB}\rangle^{\otimes n} = \prod_{i=1}^{n} (cos\theta|a_1(i)b_1(i)\rangle + sin\theta|a_2(i)b_2(i)\rangle)$$

When expanded binomially, we get 2^n terms but only $n+1$ coefficients: $cos^n\theta$, $(cos^{n-1}\theta)\,sin\theta$, ... , $sin^n\theta$. All the kets with the same coefficients could be thought of spanning a subspace of the entire Hilbert space. Let Alice perform an incomplete projective measurement, i.e project onto the above mentioned subspaces and not into any of the possible smaller spaces. Let the state be projected onto the K subspace, the subspace whose coefficients have $sin\theta$ raised to the power of K . Either Alice or Bob, lets say Alice, performs the measurement and finds the value assume by the random variable K . Either Alice could simply tell Bob the value of K or Bob could perform the same measurement and find the same K since the state does not leave the subspace after measurement. The probability distribution for this random variable K is a binomial distribution given as

$$Pr(K=k) = \binom{n}{k} \left(cos^2\theta\right)^{n-k} \left(sin^2\theta\right)^k$$

The state after measurement will have a sum of terms whose coefficients are all equal. This implies that the residual state, say Ψ_k , is in a maximally entangled state in a known $2\binom{n}{k}$ subspace as opposed to a partially entangled state in the entire 2^{2n} Hilbert space. Thus, we see that we have partially succeeded in concentrating the entanglement. This state could be used for faithful transmission of quantum information in $\binom{n}{k}$ - dimensional Hilbert spaces or smaller. This method goes by the name *Schmidt projection method.*

If one is particular about obtaining singlets from the n pairs of qubits, then one has to perform the following standardization process. Fix some small, positive ϵ where $\epsilon = 0$ corresponds to the perfect efficiency in transmission. Let Alice and/or Bob perform the incomplete Von Neumann measurements mentioned earlier and come up with outcomes $k_1, k_2, \ldots, k_n$ for the random variables $K_1, K_2, \ldots, K_n$. Let D_m defined as

$$D_m \equiv \binom{n}{k_1}\binom{n}{k_2}\ldots\binom{n}{k_m}$$

be the product of the binomial coefficients $\binom{n}{k}$ for the first m pairs. This sequence is computed till the partial product D_m lies between 2^l and $2^l\,(1+\epsilon)$ for some l . The probability that a particular value of m doesn't lie in this range goes to zero as m increases. $2D_m$ corresponds to the dimension of the Hilbert space in which the states of qubits on Alice's side (or Bob's side) that have been included in the computation live.

Once a suitable m is found, divide the space of $2D_m$ dimensions according to one's convenience into a larger subspace of dimension $2{\times}2^l$ and a smaller subspace of dimension $2\left(D_m - 2^l\right) < \epsilon{\times}2{\times}2^l$ dimensions and make a projective measurement. If the resultant state happens to be in the smaller subspace, then we declare failiure. The probability of this happening is less than ϵ . Therefore, with probability greater than $1-\epsilon$, the residual state is a maximally entangled state of two subsystems in a 2^l dimensional space, one of which is possessed by Alice and the other by Bob. Expressing this state in a Schmidt decomposed form, Alice and Bob can learn the local unitary processes that they need to perform in order to convert this into l pairs of standard singlets or other Bell states.

The expected amount of entanglement distilled in this case is given by

$$\sum_{k=1}^{n-1}\left(cos^2\theta\right)^{n-k}\left(sin^2\theta\right)^{k}\binom{n}{k}\log_2\binom{n}{k}.$$

Thus for $n > 2$, we have shown a method to obtain pure channels i.e Bell states with vanishing low probability of failiure. [4] go further and discuss another method, albeit less effective, to obtain a singlet from a single non-maximally entangled pair of qubits. This being beyond the scope of our presentation, we leave it to the interested reader to pursue and we begin discussing distillation in mixed states.

4 Distillation of mixed states

4.1 Witnessing Entanglement

Before we even try to distill or concentrate entanglement from arbitrary mixed states, it is worthwhile to spend some time knowing if the mixed state given to us has some entanglement that can be distilled. We state a general theorem to find whether a bipartite mixed state is separable or entangled and then give a simpler criterion to determine whether a state involving two qubits is separable or entangled.

4.1.1 Definitions

The following definitions are from [6].

Positive maps: *Positive maps* are maps between the space of operators. They are also known as *superoperators*, because they act on an operator to give another operator, just like how an operator acts on a vector to give another vector. The simplest example is the time evolution in quantum mechanics that takes in a density operator as its input argument and outputs another density operator, the state of the system at a later instant of time.

- A map, Φ , is said to be *linear* if $\Phi(\alpha M + \beta N) = \alpha\Phi(M) + \beta\Phi(N)$ where M and N are two operators in the same operator space.
- A map, Φ , is said to be *self-adjoint* it it maps Hermitian operators onto Hermitian operators. $\Phi(M) = \Phi^{\dagger}(M)\ \forall M \ni: M = M^{\dagger}$.
- A map is said to be *positive* if it is linear, self-adjoint and maps positive operators onto positive operators.
- A map, Φ , is said to be *completely positive* if it is positive and if its extension to any arbitrary anxillary system A of arbitrary dimensions , denoted by $\tilde{\Phi} = I_A \otimes \Phi$ is also positive.

By taking the anxillary system in last statement above to be the rest of the universe, we understand that it is a statement that if performing a physical operation is allowed locally, it must be allowed universally. But if a map is positive and not completely positive, then by performing some local operations, one is forcing the universe to be in an unphysical state and hence is not allowed. This is the idea of witnessing entanglement: If one performs an operation on a subsystem and forces the density operator that represents the state of this subsystem combined with another subsystem to become negative, then it is unphysical and this would mean that there is some entanglement between the two subsystems. This shall become clear as we proceed.

Jamiolkowski isomorphism: To relate Hermitian operators that appear all the time (almost) in quantum mechanics and positive maps that we just introduced, we state the following isomorphism between the two.

For every Hermitian operator A , there exists a positive map Φ_A , such that

$$A = I \otimes \Phi_A \left(|\phi^+\rangle\langle\phi^+| \right),$$

where $|\phi^+\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} |nn\rangle$ is the maximally entangled state of two N - dimensional systems (N could be any number).

We now state an important theorem which tells that positive maps could be used as entanglement witnesses. Interested reader may refer [6] for more information.

4.1.2 Theorem (without proof)

A bipartite mixed state ρ_{AB} is entangled *if and only if* there exists a positve map Λ (not a completely positive map) such that

$$I \otimes \Lambda \left(\rho_{AB} \right) < 0.$$

Thus, the search for witnesses of entanglement and the search for positive maps are one and the same. For a bipartite separable state, $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$, the action of any positive map on it gives back a positive operator (a possible density operator).

$$\left(I \otimes \Lambda \right) \rho_{AB} = \sum_i p_i \rho_A^i \otimes \Lambda \left(\rho_B^i \right) > 0.$$

This remarkable isomorphism between Hermitian operators and positive maps will be turned into a powerful theorem/criterion for a system containing two qubits. Before that, we will need the following mathematical result:

In two dimensions, any positive operator P can be written as

$$P = CP_1 + CP_2 T$$

where CP_1 and CP_2 are two completely positive operators and T is the transposition operation. It is clear that the action of a positive map on a two-dimensional system could possible yield unphysical results only through the action of transposition. Applying this result to a joint state ρ_{AB} of two

qubits, for any positive map P on the subsystem B ,

$$I_A \otimes P_B(\rho_{AB}) = I_A \otimes (CP_1 + CP_2T)_B \rho_{AB}$$

is positive (or negative) only if the transposition of B in the state ρ_{AB} is positive (or negative). Thus, we can conclude that the joint state of the system ρ_{AB} is separable *if and only if* $T_B(\rho_{AB}) > 0$. That is to say that the *partial transpose* is positive. This is a simple test in practice to verify if a bipartite state is separable or not. This test is known as the *Peres' Partial Transpose (PPT) criterion* or sometimes as *Peres-Horodecki condition* and is explained in detail below.

4.1.3 Peres-Horodecki condition

Peres-Horodecki condition, which also goes by the name Peres' Partial Transpose (PPT) condition is a condition to determine whether a bipartite mixed state is separable or entangled.

Consider a generic bipartite state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$. Without loss of generality, we can see that ρ_{AB} can be expressed in the following form:

$$\rho_{AB} = \sum_{i,j,k,l} p_{kl}^{ij} \left(|i\rangle\langle j|\right)_A \otimes \left(|k\rangle\langle l|\right)_B .$$

Its partial transpose with respect to the subsystem B , denoted by $\rho_{AB}^{T_B}$ is then

$$\rho_{AB}^{T_B} \equiv I \otimes T(\rho) = \sum_{i,j,k,l} p_{kl}^{ij} \left(|i\rangle\langle j|\right)_A \otimes \left(|k\rangle\langle l|\right)_B^T = \sum_{i,j,k,l} p_{kl}^{ij} \left(|i\rangle\langle j|\right)_A \otimes \left(|l\rangle\langle k|\right)$$

or equivalently

$$\rho_{AB}^{T_B} = \sum_{i,j,k,l} p_{lk}^{ij} \left(|i\rangle\langle j|\right)_A \otimes \left(|k\rangle\langle l|\right) .$$

If the *positivity* is violated, then we can be certain that the state is not separable and hence entangled. But the converse is not generally true. Entangled states might still satisfy the positivity criterion. This is precisely the reason why $E_D(\rho) \leq E_F(\rho)$. But in the case of two qubits, Peres-Horodecki condition serves as a necessary and sufficient condition to determine separability as given by the generic mathematical results.

Bell's inequalities serve as another witness for entanglement. Peres' Partial Transpose condition detects entanglement in some cases where Bell's inequalities fail to. Thus, Peres-Horodecki condition is a stronger criterion than Bell's inequalities. It turns out that entangled states that are not identified by Peres-Horodecki condition cannot be distilled to form singlets. The entanglement between such systems, that was put in when the state was formed, cannot be recovered to do useful work. Such entanglement are known as *bound entanglement* while those that can be identified by

Peres' Partial Transpose condition are known as *free entanglement* and can be concentrated to do some useful informational work.

4.2 Distilling mixed states - Protocol

This subsection is a summary of [2]. For notational convenience, let us rename the four Bell states in the following manner:

$|\beta_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\Phi^+\rangle$
$|\beta_2\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = |\Phi^-\rangle$
$|\beta_1\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = |\Psi^+\rangle$
$|\beta_3\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = |\Psi^-\rangle$

Let the initial state of any single qubit pair be M , which could potentially be distilled based on the criteria described in the previous subsection. Let us say we are interested in obtaining m number of singlet states $|\Psi^-\rangle$. The purity of the initial state, relative to that of the singlet states, can be expressed by its fidelity $F = \langle\Psi^-|M|\Psi^-\rangle$.

The first step in the protocol is to have Alice and Bob apply *random bilateral rotation* on each shared pair, choosing a random $SU(2)$ rotation for each pair and applying it locally to both the members of the pair to obtain a rotationally symmetric mixed state, which is well known in the quantum information community as the *Werner state*. The authors of [2] claim that the same result could be obtained by choosing elements from a finite set of rotations $\{B_x, B_y, B_z, I\}$ defined as follows:

B_x , B_y , B_z represent *bilateral* rotations - rotations about the x, y and the z - axis respectively by $\pi/2$. More on the properties of these bilateral rotations later.

The resultant states is then represented as

$$W_F = F|\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{3}|\Psi^+\rangle\langle\Psi^+| + \frac{1-F}{3}|\Phi^+\rangle\langle\Phi^+| + \frac{1-F}{3}|\Phi^-\rangle\langle\Phi^-|.$$

The Bell states are orthogonal to each other and since the singlet state is invariant under bilateral rotations, the purity of the Werner state, relative to the singlet state is same as that of the initial mixture relative to the singlet.

$$F = \langle\Psi^-|M|\Psi^-\rangle = \langle\Psi^-|W_F|\Psi^-\rangle.$$

Different initial states M with the same purity generate the same Werner state. Thus, our analysis from now on is general and can be considered only as a function of F , the purity of the

initial state. We now introduce some operators and their action on the Bell states.

1. *Unilateral Pauli rotations* - Rotation about x, y or z - axis by an angle π of *one* qubit in an entangled pair. These operations map the Bell states onto the Bell states (ignoring an overall phase factor) in an one-to-one fashion and leaving no states unchanged.
   σ_x maps $\Psi^{\pm} \rightleftharpoons \Phi^{\pm}$.
   σ_y maps $\Psi^{\pm} \rightleftharpoons \Phi^{\mp}$.
   σ_z maps $\Psi^{\pm} \rightleftharpoons \Psi^{\mp}$ and $\Phi^{\pm} \rightleftharpoons \Phi^{\mp}$.

2. *Bilateral rotations* - $\{B_x, B_y, B_z\}$: Rotations about x, y, z - axis of *both* the qubits in a pair by $\pi/2$. These rotations leave the singlet and one of the triplets invariant and interchange the other two (ignoring the phase factors) .
   B_x maps $\Phi^{+} \rightleftharpoons \Psi^{+}$.
   B_y maps $\Phi^{-} \rightleftharpoons \Psi^{+}$.
   B_z maps $\Phi^{+} \rightleftharpoons \Phi^{-}$.

3. *Bilateral XOR (BXOR)* - Conditionally flip the spin with control qubit on Alice's end and target qubit on Bob's end or vice versa. The table below explicitly shows the source and target states. The first two columns on the left represent the states before the BXOR operation and the last two columns on the right represent the states after the BXOR operation.

Source	Target	Source	Target
$\Phi^{\pm}$	Φ^{+}	NC	NC
$\Psi^{\pm}$	Φ^{+}	NC	Ψ^{+}
$\Psi^{\pm}$	Ψ^{+}	NC	Φ^{+}
$\Phi^{\pm}$	Ψ^{+}	NC	NC
$\Phi^{\pm}$	Φ^{-}	$\Phi^{\mp}$	NC
$\Psi^{\pm}$	Φ^{-}	$\Psi^{\mp}$	Ψ^{-}
$\Psi^{\pm}$	Ψ^{-}	$\Psi^{\mp}$	Φ^{-}
$\Phi^{\pm}$	Ψ^{-}	$\Phi^{\mp}$	NC

*NC = No change

Along with these operations, Alice and Bob can also measure the z - component of the spin of both the qubits in a pair. We shall now show that given two Werner pairs of fidelity $F > \frac{1}{2}$, Alice and Bob, by means of local operations and two-way classical communication, with a probability of $\frac{1}{4}$, can obtain a Werner pair of fidelity $F' > F$. The fidelity F' satisfies the recurrence relation (not proved in this paper)

$$F' = \frac{F^2 + \frac{1}{9}(1-F)^2}{F^2 + \frac{2}{3}F(1-F) + \frac{5}{9}(1-F)^2}.$$

The following is the protocol to achieve this:

- (A1) A unilateral σ_y is performed on each of the two pairs converting them from mostly Ψ^-Werner states to analogously mostly Φ^+ states i.e. Φ^+ has the large component $F > \frac{1}{2}$ and the other three Bell states in equal proportion.

- (A2) A BXOR is performed on the two impure Φ^+states, after which the target pair is locally measured along the z - axis. If the target pair's z - spins come out parallel , as they should for a pure Ψ^+ pair, the unmeasured source pair is retained; otherwise it is discarded.

- (A3) If the source pair was retained, it is converted back to a mostly Ψ^- Werner state by another unilateral σ_y rotation and then made rotationally symmetric by a random bilateral rotation.

Because $F'(F)$ is continuous and is greater than F over the range $\frac{1}{2} < F < 1$, iteration of the protocol described above can distill Werner states of arbitrarily high purity i.e $F_{out} \to 1$ which corresponds to Ψ^- . But unfortunately, the yield (average number of purified output pairs per impure pair) goes to zero as the output fidelity goes to 1. By combining this with another protocol mentioned below, that consumes a fraction of the previously purified Φ^+pairs in order to produce more than the number consumed, like in a breeder reactor, the yield could be increased.

The basic step used in this protocol is the BXOR test, which consists of bilaterally XORing a subset of the impure pairs, used as the sources, into one of the pure Φ^+states, used as a target, followed by the measurement of the target. The BXOR test is very similar to the *parity check* that is done on classical data.

- (B1) Alice and Bob start with n impure pairs, each in the same Werner state W_F with $S(W_F) < 1$ and $n\,[S(W_F) + \delta]$ prepurified Φ^+ states from the method described above. Here, δ is a positive constant that approaches zero for large n .

- (B2) Using the prepurified Φ^+ pairs as targets, Alice and Bob perform BXOR tests on sufficiently many random subsets of the impure pairs to locate all the Ψ states, without differentiating Ψ^+ from Ψ^- . Once found, they are converted back to $\Phi^\pm$states by the action of unilateral σ_x .

- (B3) The impure pairs now contain only $\Phi^\pm$states. Alice and Bob do a bilateral B_y operation to convert the Φ^- to Ψ^+ while leaving the Φ^+ states invariant. Once this is done, they perform BXOR tests on sufficiently many more random subsets to find all the new Ψ^+ states with high probability. Once found, these are converted to the desired Φ^+ form by unilateral σ_x rotations.

The number of BXOR tests per impure pair required to find all the errors, with arbitrarily small chance of failiure, approaches the entropy of the impure pairs, $S(W_F) = -Tr(W \log_2 W)$, in the

limit of large n . This *breeding method* has a yield of $S(W_F)$, producing more pure pairs than consumed if $S(W_F) < 1$. We find that

$$1 - S(W_F) = 1 + F \log_2 F + (1 - F) \log_2 \frac{1 - F}{3}$$

is positive for $F > 0.8107$. Thus, using prepurified Φ^+ states can considerably increase the yield.

It is worthwhile to note that Werner states violate CHSH inequality [10], another witness of entanglement, for $F > \frac{2+3\sqrt{2}}{8} \approx 0.78$ while are not separable for $F > \frac{1}{2}$. As we had mentioned earlier, the Peres' Partial Transpose criterion is stronger than CHSH inequality. And we have successfully shown that entanglement could be distilled from Werner states if $F > \frac{1}{2}$.

We conclude this term paper on entanglement distillation with this. Enormous amount of literature exist on this topic, [2, 4, 5] to mention a few.

References

[1] "Quantum Information and Quantum Computation", *M. A. Nielsen and I. L. Chuang*, Cambridge University Press.

[2] "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels", *C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin and W.K. Wooters*, Phys. Rev. Lett. 76, 722–725 (1996) .

[3] "Entanglement of Formation of an Arbitrary State of Two Qubits", *W. K. Wooters*, Phys. Rev. Lett. 80 : 2245-2248 (1998)

[4] "Concentrating partial entanglement by local operations", *C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher*, Phys. Rev. A 53, 2046–2052 (1996) .

[5] "Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature?", *M. Horodecki, P. Horodecki, R. Horodecki*, Phys. Rev. Lett. 80, 5239–5242 (1998) .

[6] "Introduction to Quantum Information Science", *Vladko Vedral,* Oxford University Press.

[7] "Quantum Computing Explained", *David McMahon,* Wiley-Interscience.

[8] "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", *A. Einstein, B. Podolsky, N. Rosen*, Phys. Rev. 47, 777–780 (1935) .

[9] "On the Einstein Podolsky Rosen paradox", *John S. Bell.*

[10] "Proposed Experiment to Test Local Hidden-Variable Theories", *J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt*, Phys. Rev. Lett. 23, 880–884 (1969) .