BIPARTITE ENTANGLEMENT
-
CRITERIA, MEASURES, APPLICATIONS

Contents

1. Introduction 3
    1.1. Notation and Terminology 3
2. Separability criteria 4
    2.1. Partial Transpose 4
    2.2. Necessary and sufficient conditions 5
    2.3. Separability criterion and positive operators 6
    2.4. Low dimensional case 11
3. Distillability 11
    3.1. Relation to Peres' criterion 12
    3.2. The reduction criterion 14
    3.3. A sufficient condition for distillability 16
4. Irreversibility of distillation 18
    4.1. Definition of terms 18
    4.2. Lower bound on the entanglement cost 18
    4.3. Irreversibility with catalyst 20
5. Negativity 20
    5.1. Definition 20
    5.2. Monotonicity 21
6. Teleportation 23
    6.1. Observations on channels and fidelity 24
    6.2. Maximal singlet fraction and teleportation fidelity 26
7. Teleportation and dense coding 27
    7.1. Dense Coding ⇒ Teleportation Scheme 28
    7.2. Teleportation Scheme ⇒ Dense Coding 32
8. Conclusion 33

References 34

1. Introduction

One of the most striking differences between classical and quantum information theory is the concept of entanglement where, intuitively speaking, two systems "feel each other" such that a measurement "over here" influences a measurement "over there", regardless of the spatial separation. This somewhat strange idea is the very essence of phenomena such as teleportation and dense coding. In this essay the key aspects of entanglement will be discussed from a mathematical point of view. Physical experiments are not included.

Only bipartite entanglement for finite dimensional systems will be treated but these can be general mixed states. Different criteria for separability based on positive maps will be presented. The relation to distillability will be made clear, although distillation protocols are not discussed in detail. The process of distillation will be shown to be irreversible. Negativity is presented as an entanglement measure and its monotonicity is proven. As a first application of entanglement as a physical resource teleportation is introduced and its fidelity is related to the maximal singlet fraction. For the "tight" case the equivalence of all teleportation and dense coding schemes is shown.

1.1. **Notation and Terminology.** $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ will be a finite dimensional Hilbert space. $\mathcal{H}_A$ and $\mathcal{H}_B$ will generally not be assumed to have the same dimension. $\mathbb{1}$ denotes the identity in $\mathcal{B}(\mathcal{H})$, the set of bounded linear operators acting on $\mathcal{H}$, while I is the identity linear map acting on $\mathcal{B}(\mathcal{H})$. A distinction between elements in $\mathcal{B}(\mathcal{H})$ and matrices is neglected, though it is always pointed out when a particular basis is chosen. All inner products $\langle .|.\rangle$ are defined to be linear in the second component and anti-linear in the first. I use the term *positive operator* to refer to positive semi-definite operators and thus *not* meaning strictly positive. † denotes the Hermitian conjugate whereas * stands for complex conjugation. Vectors will always be written in the $|\,\rangle$ ket notation. The *support* of an operator is a synonym for its image space.

Finally the most important definition:

Definition 1.1.1. We call a density operator ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ *separable* if it can be written as a convex combination

$$\rho = \sum_i p_i \, \rho_i^A \otimes \rho_i^B, \tag{1.1.1}$$

where $\rho_i^{A/B}$ acts on $\mathcal{H}_{A/B}$, $\sum_i p_i = 1$ and $\forall i \; p_i > 0$. If this is not possible, the state is *inseparable* or *entangled.*

2. Separability criteria

The first part of this essay will discuss criteria which determine whether a general mixed state ρ is separable or not. Asher Peres [1] observed that the partial transpose of ρ must have only non-negative eigenvalues if ρ is separable.

2.1. **Partial Transpose.**

Definition 2.1.1. Let $\{|e_i\rangle\}$ be a basis for $\mathcal{H}_A$ and $\{|f_i\rangle\}$ be a basis for $\mathcal{H}_B$. With this choice of basis the density matrix has the entries

$$\rho_{m\mu,n\nu} = \langle e_m \otimes f_\mu |\, \rho \,| e_n \otimes f_\nu \rangle \tag{2.1.1}$$

Latin indices refer to the first subsystem, Greek indices to the second. The *partial transpose* ρ^{T_2} is then defined by

$$\rho^{T_2}_{m\mu,n\nu} = \rho_{m\nu,n\mu}. \tag{2.1.2}$$

Proposition 2.1.2. *If ρ is separable then ρ^{T_2} is also a density operator.*

Proof. If ρ is separable then by (1.1.1)

$$\rho^{T_2} = \sum_i p_i\, \rho_i^A \otimes (\rho_i^B)^T$$

Observe that it is enough to show that $(\rho_i^B)^T$ is a density matrix, as then ρ will also be non-negative and have $\operatorname{tr}\rho = 1$. But clearly the transpose of a non-negative matrix is again non-negative. Furthermore, transposition preserves the traces. So $(\rho_i^B)^T$ is another density matrix. □

From this observation we deduce the following necessary condition:

> If ρ is separable then ρ^{T_2} has only positive eigenvalues.

This criterion is generally *not* sufficient [2]. It will later be discussed how a violation of this criterion can be quantified and hence used as a measure of entanglement. In the above I have followed the convention in [3], where the transposition acts on the second system. This does not make any difference as the criterion holds if and only if it holds when working with ρ^{T_1}. This is obvious, as $\rho^{T_1} = (\rho^{T_2})^T$, the (full) transpose of ρ^{T_2}. Hence, they have the same characteristic equation and thus the same eigenvalues.

Proposition 2.1.2 poses the question whether the partial transposition could be replaced by another, or maybe even arbitrary positive map to obtain a useful criterion. In other words, for which class of positive operators Λ, is $(\Lambda \otimes I)\rho$ *not* a positive operator given that ρ is inseparable? Clearly, completely positive maps will always be useless for such a test. In some cases it turns out that the partial transposition is characteristic for this class of positive maps and that we also obtain a sufficient condition (see Proposition 2.4.1).

2.2. **Necessary and sufficient conditions.** In this section some of the ideas developed above will be generalized to obtain a sufficient and necessary condition for separability. The following lemma plays an important role as it tells us how we can tell an inseparable state apart from the class of separable states.

Lemma 2.2.1. *For any inseparable state $\rho \in \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B)$ there exists a Hermitian operator A such that*

$$\operatorname{tr}(A\rho) < 0 \quad \textit{and} \quad \operatorname{tr}(A\sigma) \geq 0 \tag{2.2.1}$$

for any separable σ.

Proof. We want to invoke the following Separation Theorem (a consequence of the Hahn-Banach Theorem). See e.g. [4] for a proof.

Separation Theorem: Let E be a real locally convex space equipped with a seminorm and let A and B be convex subsets of E with $A \cap B = \emptyset$. Then if A is open there exists a linear functional $f : E \to \mathbb{R}$ and $\xi \in \mathbb{R}$ s.t.

$$f(x) < \xi \leq f(y) \quad \forall x \in A, y \in B.$$

In our setting $E = Herm(\mathcal{H})$ is the set of Hermitian operators acting on $\mathcal{H}$, which is a real vector space. As a norm on $Herm(\mathcal{H})$ we can take the standard norm $\|A\| := \sup_{\|v\|=1} v^T A v$. This space is also (globally) convex as the convex sum of positive operators is again positive. The set A is the one element set $A = \{\rho\}$ which is trivially both convex and open. The set B is the set of all separable density operators, which is also convex. As ρ is assumed to be inseparable $A \cap B = \emptyset$ is satisfied. Thus we can apply the Separation theorem to deduce the existence of a real functional f on $Herm(\mathcal{H})$ and $\xi \in \mathbb{R}$ such that

$$f(\rho) < \xi \leq f(\sigma)$$

for any inseparable σ. Observe that any linear functional g can be written as $g(\sigma) = \operatorname{tr}(\sigma \widetilde{A})$ where the matrix $\widetilde{A}$ depends on g. To see this in this particular case, observe that $Herm(\mathcal{H})$ is generated by matrices of the form $\alpha E_{ij} + \alpha^* E_{ji}$ where E_{ij} has zeros everywhere except at the position i (row), j (column) where it has a one. Applying f to this generator and remembering that f is real, one notes that

$\widetilde{A}$ has to be Hermitian and one can read off the matrix entry $\widetilde{A}_{ij}$ and define the matrix by these entries.

$$\begin{array}{lrcl} \text{Hence} & \operatorname{tr}(\rho\widetilde{A}) & < \xi \leq & \operatorname{tr}(\sigma\widetilde{A}) \\ \Leftrightarrow & \operatorname{tr}(\rho\widetilde{A}) - \xi & < 0 \leq & \operatorname{tr}(\sigma\widetilde{A}) - \xi \\ \Leftrightarrow & \operatorname{tr}(\rho\widetilde{A}) - \operatorname{tr}(\rho\xi\mathbb{1}) & < 0 \leq & \operatorname{tr}(\sigma\widetilde{A}) - \operatorname{tr}(\sigma\xi\mathbb{1}) \\ \Leftrightarrow & \operatorname{tr}(\rho(\widetilde{A} - \xi\mathbb{1})) & < 0 \leq & \operatorname{tr}(\sigma(\widetilde{A} - \xi\mathbb{1})) \end{array}$$

Here we have used the fact $\operatorname{tr}(\rho) = \operatorname{tr}(\sigma) = 1$. By defining $A := \widetilde{A} - \xi\mathbb{1}$ we obtain equation (2.2.1).

□

Theorem 2.2.2. *A state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ is separable iff*

$$\operatorname{tr}(\rho A) \geq 0 \tag{2.2.2}$$

for any Hermitian operator A satisfying $\operatorname{tr}((P \otimes Q)A) \geq 0$ *for all projections P and Q acting on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively.*

Although this theorem gives us the desired necessary and sufficient conditions we will see in the next section how it can be rewritten in a more useful form in terms of positive operators.

Proof. $\Rightarrow$: If ρ is separable then by definition 1.1.1 it can be written as $\rho = \sum_i p_i\, \rho_i^A \otimes \rho_i^B$. But each ρ_i^A and ρ_i^B is a convex linear combination of pure density matrices. Pure density matrices are exactly rank one projections. So using linearity of the trace and noting that all coefficients involved are non-negative the first implication follows.

$\Leftarrow$: Let ρ be inseparable. Assume the condition (2.2.2) *does* always hold and try to deduce a contradiction. By Lemma 2.2.1 we could then find a Hermitian operator A such that $\operatorname{tr}(\rho A) < 0$ but $\operatorname{tr}(\rho\sigma) \geq 0$ for all separable states σ. In particular we would have $\operatorname{tr}((P \otimes Q)A)$ for all $P \otimes Q$, where P and Q are projections which is a contradiction and thus proves the theorem. □

2.3. **Separability criterion and positive operators.** The aim of this section is to prove the following

Theorem 2.3.1. *A state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ is separable iff for any positive map $\Lambda : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A)$ the operator $(I \otimes \Lambda)\rho$ is positive.*

Peres' criterion (Proposition 2.1.2) is an immediate corollary. The surprising fact that for low dimensions it is also sufficient to *only* consider the transposition operator will be proved later (see Proposition 2.4.1). As the proof is rather involved, it is broken up into several steps.

Proof. To deduce this theorem from Theorem 2.2.2 we need to relate operators A, which are positive on products of projections, to positive maps Λ. It will be shown that this can be achieved by the map $S : \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B)$ which is defined by

$$[A^\dagger \otimes B, S(\Lambda)] := [B, \Lambda(A)]$$

Here $[.\,,.]$ is the inner product on the space of operators defined as $[A_1, A_2] = \mathrm{tr}(A_1^\dagger A_2)$ for any operators A_1 and A_2.

We will now establish some properties of the map S. (Also see [5].)

Claim 1

$$S(B][A) = A^\dagger \otimes B$$

where $(B][A)$ is an element of $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ and it acts as $(B][A)C = [A, C]B$.

Proof of Claim 1:

$$\begin{aligned} & [C \otimes D, S(B][A)] \\ = \; & [D, (B][A)C^\dagger] && \text{(definition of } S\text{)} \\ = \; & [D, [A, C^\dagger]B] && \text{(definition of (B][A))} \\ = \; & [A, C^\dagger][D, B] && \text{(linearity in 2nd component)} \\ = \; & [C, A^\dagger][D, B] && \text{(cyclicity of trace)} \\ = \; & [C \otimes D, A^\dagger \otimes B] \end{aligned}$$

As this holds for all C and D, we deduce $S(B][A) = A^\dagger \otimes B$ which finishes the proof of Claim 1.

Claim 2

$$S(\Lambda) = \sum_i E_i^\dagger \otimes \Lambda(E_i) \tag{2.3.1}$$

where $\{E_i\}$ is any orthonormal basis in $\mathcal{H}_A$.

Proof of Claim 2: From Claim 1 it follows that

$$\begin{aligned} S(B][A) &= A^\dagger \otimes B \\ &= \textstyle\sum_i (E_i^\dagger][E_i^\dagger) A^\dagger \otimes B && \text{(completeness for conjugate orthonormal basis)} \\ &= \textstyle\sum_i [A, E_i] E_i^\dagger \otimes B && \text{(definition of } (E_i^\dagger][E_i^\dagger) \text{ and cyclicity of trace)} \\ &= \textstyle\sum_i E_i^\dagger \otimes [A, E_i]B \\ &= \textstyle\sum_i E_i^\dagger \otimes (B][A)E_i \end{aligned}$$

As linear maps of the form $(B][A)$ generate $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$, we deduce that equation (2.3.1) holds for any Λ which finishes the proof of Claim 2.

Claim 3

A map $\Lambda : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ sends Hermitian operators on $\mathcal{H}_A$ to Hermitian operators on $\mathcal{H}_B$ if and only if $S(\Lambda) \in \mathcal{H}_A \otimes \mathcal{H}_B$ is Hermitian.

Proof of Claim 3:

$$\begin{aligned}
[A \otimes B, S(\Lambda)^\dagger] &= [S(\Lambda), A^\dagger \otimes B^\dagger] && \text{(cyclicity of trace)}\\
&= [A^\dagger \otimes B^\dagger, S(\Lambda)]^* && \text{(anti-symmetry of inner product)}\\
&= [B^\dagger, \Lambda(A)]^* && \text{(definition of } S\text{)}\\
&= [\Lambda(A), B^\dagger] && \text{(anti-symmetry of inner product)}\\
&= [B, \Lambda(A)^\dagger] && \text{(cyclicity of trace)}
\end{aligned}$$

Now $[B, \Lambda(A)^\dagger] = [B, \Lambda(A^\dagger)]$ for all A and B iff $\Lambda(A)^\dagger = \Lambda(A^\dagger)$. But as $[B, \Lambda(A^\dagger)] = [A \otimes B, S(\Lambda)]$ we obtain

$$[A \otimes B, S(\Lambda)^\dagger - S(\Lambda)] = 0$$

for all A and B iff $\Lambda(A^\dagger) = \Lambda(A)^\dagger$. This finishes the proof of Claim 3.

Claim 4

A map $\Lambda : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ sends positive operators to positive operators iff $\langle x \otimes y, S(\Lambda)(x \otimes y)\rangle \geq 0$ for all $x \in \mathcal{H}_A$ and $y \in \mathcal{H}_B$ [6].

Proof of Claim 4: Observe that the set of positive operators on $\mathcal{H}_A$ can be generated by positive linear combinations of projections $P_x = |x\rangle\langle x|$ where $x \in \mathcal{H}_A$ and $\langle x|x\rangle = 1$ and clearly any such projection is a positive operator. (In the following proof the normalization is of no relevance.) This can be seen by diagonalizing the positive operator. Thus Λ preserves positive operators iff

$$\langle y|\Lambda(P_x)y\rangle \geq 0 \tag{2.3.2}$$

for all $x \in \mathcal{H}_A$ and $y \in \mathcal{H}_B$.

$$\begin{aligned}
\Lambda(A) &= \Lambda(\textstyle\sum_i (E_i][E_i)A) && \text{(completeness)}\\
&= \textstyle\sum_i [E_i, A]\Lambda(E_i)
\end{aligned}$$

Hence,
$$\begin{aligned}
\Lambda(P_x) &= \textstyle\sum_i [E_i, P_x]\Lambda(E_i) && \text{(using the formula above)}\\
&= \textstyle\sum_i \operatorname{tr}(E_i^\dagger |x\rangle\langle x|)\Lambda(E_i)\\
&= \textstyle\sum_i \langle x|E_i^\dagger x\rangle\Lambda(E_i)
\end{aligned}$$

Substituting this into (2.3.2) we get

$$\begin{aligned} & \textstyle\sum_i \langle x|E_i^\dagger x\rangle\langle y|\Lambda(E_i)y\rangle \geq 0 \\ \Leftrightarrow\quad & \textstyle\sum_i \langle x\otimes y|(E_i^\dagger \otimes \Lambda(E_i))x\otimes y\rangle \geq 0 \\ \Leftrightarrow\quad & \langle x\otimes y|(\textstyle\sum_i E_i^\dagger \otimes \Lambda(E_i))x\otimes y\rangle \geq 0 \\ \Leftrightarrow\quad & \langle x\otimes y|S(\Lambda)x\otimes y\rangle \geq 0 \qquad \text{(Claim 2)} \end{aligned}$$

This finishes the proof of Claim 4.

Remark 2.3.2. The condition in Claim 4 is equivalent to

$$\operatorname{tr}(S(\Lambda)(|x\rangle\langle x|\otimes|y\rangle\langle y|)) \geq 0$$

for all $x \in \mathcal{H}_A$ and $y \in \mathcal{H}_B$. This is then equivalent to

$$\operatorname{tr}(S(\Lambda)(P\otimes Q)) \geq 0$$

for all projections $P \in \mathcal{B}(\mathcal{H}_A)$ and $Q \in \mathcal{B}(\mathcal{H}_B)$. Note that positive maps preserve Hermiticity as we can write any Hermitian operator as a linear combination of positive operators (see (5.2.1)) and then use linearity and the fact that positive operators are Hermitian. So we have that Λ is positive iff $S(\Lambda)$ is Hermitian and $\operatorname{tr}(S(\Lambda)(P\otimes Q)) \geq 0$ for all projections P and Q.

Claim 5:

The map S also establishes an isomorphism, i.e. a structure-preserving 1-1 correspondence between $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ and $\mathcal{B}(\mathcal{H}_A)\otimes\mathcal{B}(\mathcal{H}_B)$ (in fact an isometry with respect to the inner product $[.\,,.]$).

Proof of claim 5:

- Injective: By (2.3.1)
$$\begin{aligned} & S(\Lambda_1) = S(\Lambda_2) \\ \Leftrightarrow\quad & \textstyle\sum_i E_i^\dagger \otimes (\Lambda_1 - \Lambda_2)(E_i) = 0 \\ \Leftrightarrow\quad & \textstyle\sum_i (\Lambda_1 - \Lambda_2)(E_i) = 0 \\ \Leftrightarrow\quad & \Lambda_1 = \Lambda_2 \end{aligned}$$
- Surjective: We know $S(B][A) = A^\dagger \otimes B$ and $\mathcal{B}(\mathcal{H}_A)\otimes\mathcal{B}(\mathcal{H}_B)$ is spanned by operators of this form.

Claim 5 shows that we can write $A = S(\Lambda)$ in Theorem 2.2.2. Together with the remark after Claim 4 this theorem becomes

Theorem 2.3.3. *A state ρ acting on $\mathcal{H}_A\otimes\mathcal{H}_B$ is separable iff*

$$\operatorname{tr}(\rho\, S(\Lambda)) \geq 0 \tag{2.3.3}$$

for any positive map Λ.

By choosing the basis $\{E_{ij}\}$ for $\mathcal{B}(\mathcal{H}_A)$ where $E_{ij} = |i\rangle\langle j|$ for an orthonormal basis $\{e_i\}$ of $\mathcal{H}_A$, and using the explicit formula (2.3.1) for $S(\Lambda)$ this becomes

$$\begin{aligned}
\mathrm{tr}(\rho \textstyle\sum_{ij} E_{ij}^{\dagger} \otimes \Lambda(E_{ij})) &\geq 0 \\
\mathrm{tr}(\rho\,(I \otimes \Lambda) \textstyle\sum_{ij} E_{ji} \otimes E_{ij}) &\geq 0 \\
\mathrm{tr}(\rho\,(I \otimes \Lambda T) \textstyle\sum_{ij} E_{ji} \otimes E_{ji}) &\geq 0
\end{aligned}$$

Here $T : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_A)$ is the transposition operator with respect to the basis $\{e_i\}$. Note that as Λ runs over all positive maps so does $\Lambda T = \Lambda'$. This follows as T is positive and a bijection ($T^2 = I$). Hence we can drop the $'$.

For simplicity let $E_0 := \sum_{ij} E_{ji} \otimes E_{ji}$. Now E_0 is Hermitian and $(I \otimes \Lambda)$ is positive and hence preserves Hermiticity. So the condition above becomes

$$\begin{aligned}
\mathrm{tr}(\rho\,(I \otimes \Lambda)E_0) &\geq 0 && \\
\mathrm{tr}(\rho\,[(I \otimes \Lambda)E_0]^{\dagger}) &\geq 0 && \text{(Hermiticity)} \\
\mathrm{tr}([(I \otimes \Lambda)E_0]^{\dagger}\rho) &\geq 0 && \text{(cyclicity of trace)} \\
[(I \otimes \Lambda)E_0, \rho] &\geq 0 && \text{(definition of inner product)} \\
[E_0, (I \otimes \Lambda^{\ddagger})\rho] &\geq 0 && \text{(adjoint map)} \\
\mathrm{tr}(E_0(I \otimes \Lambda^{\ddagger}\rho)) &\geq 0 && \text{(definition of inner product, Hermiticity of } E_0\text{)}
\end{aligned}$$

and this has to hold for all adjoint maps $\Lambda^{\ddagger} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A)$ of positive maps $\Lambda : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$. But for positive maps Λ

$$\begin{aligned}
& \langle y|\Lambda(|x\rangle\langle x|)|y\rangle && \geq 0 && \text{(positivity)} \\
\Leftrightarrow\quad & \mathrm{tr}(\Lambda(|x\rangle\langle x|)|y\rangle\langle y|) && \geq 0 && \\
\Leftrightarrow\quad & \mathrm{tr}([\Lambda(|x\rangle\langle x|)]^{\dagger}|y\rangle\langle y|) && \geq 0 && \text{(}\Lambda\text{ preserves Hermiticity)} \\
\Leftrightarrow\quad & [(\Lambda(|x\rangle\langle x|), |y\rangle\langle y|] && \geq 0 && \\
\Leftrightarrow\quad & [(|x\rangle\langle x|, \Lambda^{\ddagger}(|y\rangle\langle y|)] && \geq 0 && \text{(definition of adjoint map)} \\
\Leftrightarrow\quad & \langle x|\Lambda^{\ddagger}(|y\rangle\langle y|)|x\rangle && \geq 0 &&
\end{aligned}$$

So Λ is positive iff $\Lambda^{\ddagger}$ is positive. As there is a 1-1 correspondence between maps and their adjoints, ρ is separable iff $\mathrm{tr}(E_0(I \otimes \Lambda^{\ddagger}\rho))$ holds for all positive maps $\Lambda : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A)$.

Finally, we observe that if ρ is separable, then $I \otimes \Lambda\,\rho$ will be positive for any positive Λ. Conversely, if $I \otimes \Lambda\,\rho$ is positive for any positive Λ then, as discussed before, $\mathrm{tr}(P(I \otimes \Lambda\,\rho)) \geq 0$ for any projector P. But it can easily be seen that up to a positive constant $E_0^2 = E_0$, and so E_0 is a projection. Hence, the condition $\mathrm{tr}(E_0(I \otimes \Lambda\, \rho)) \geq 0$ is satisfied for any positive Λ and Theorem 2.3.1 follows

□

2.4. **Low dimensional case.** This theorem now allows us to analyze Peres' criterion in more detail. In particular we will see that for low dimensions it is also sufficient.

Proposition 2.4.1. *A density operator* ρ *acting on* $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ *where* $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$ *or* $\mathcal{H}_A = \mathbb{C}^2$, $\mathcal{H}_B = \mathbb{C}^3$ *(or equivalently* $\mathcal{H}_A = \mathbb{C}^3$, $\mathcal{H}_B = \mathbb{C}^2$ *by passing to the transposed maps) is separable iff its partial transposition* ρ^{T_2} *is also a density operator.*

Proof. $\Rightarrow$: See Proposition 2.1.2.

$\Leftarrow$: The proof depends crucially on the following fact. If $\Lambda : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is a positive map then it can be decomposed as

$$\Lambda = \Lambda_1^{CP} + \Lambda_2^{CP} T$$

Here the Λ_i^{CP} are completely positive maps and T is the transposition map with respect to a fixed basis. This result concerning the decomposition only holds when Λ acts on operators of dimension 2x2 (see [7]) and 2x3 (or 3x2) (see [8]).

Now as the Λ_i^{CP} are completely positive maps, the maps $\Lambda_i = I \otimes \Lambda_i^{CP}$ are also positive. Hence

$$\begin{aligned} & (I \otimes \Lambda)\rho \\ = & (\Lambda_1 + \Lambda_2(I \otimes T))\rho \\ = & \Lambda_1 \rho + \Lambda_2 \rho^{T_2} \end{aligned}$$

Thus if ρ^{T_2} is a density operator and hence positive, the operator $I \otimes \Lambda\, \rho$ will be positive for all positive maps Λ. So the proposition follows from Theorem 2.3.1.

□

As mentioned before, in general for higher dimensions the partial transpose criterion is *not* sufficient [2].

3. Distillability

Most protocols for applications that make use of quantum entanglement, e.g. teleportation or dense coding, require a pair of maximally entangled particles, or in other words *pure* entanglement. However, in practice it is only ever possible to produce impure entanglement. This gives rise to the concept of distillation (sometimes also referred to as purification) where from a number of N copies of a partly entangled state ρ by means of LOCC (local quantum operations and classical communication) a smaller number of n maximally entangled Bell states is obtained. A natural question to ask in this context is:

Can every entangled state be used for distillation?

In [9] it was shown that we can use any inseparable pair of qubits for distillation, regardless of how small the amount of entanglement. So one might conjecture that this is always true, even for higher dimensions. Surprisingly, it turns out that this is not true generally [10]. In fact, the question of distillability is closely related to Peres' criterion and this relationship will be discussed here.

3.1. **Relation to Peres' criterion.** By Postulate Three any action on the N copies $\rho^{\otimes N}$ can be described by a superoperator Λ. Using Kraus' Representation Theorem this can be written as

$$\Lambda(\rho^{\otimes N}) = \frac{1}{\mathrm{tr}} \sum_i M_i^\dagger \rho^{\otimes N} M_i,$$

where $\mathrm{tr} = \mathrm{tr}(\sum_i M_i^\dagger \rho^{\otimes N} M_i)$. Now if the actions are required to be local, this becomes

$$\Lambda(\rho^{\otimes N}) = \frac{1}{\mathrm{tr}} \sum_{ij} (A_i^\dagger \otimes B_j^\dagger)\, \rho^{\otimes N} (A_i \otimes B_j),$$

where A_i acts on $\mathcal{H}_A^{\otimes N}$ and B_j on $\mathcal{H}_B^{\otimes N}$. Note that the operators are allowed to act on the *whole collection* of particles in either Alice's or Bob's part of the full Hilbert space. Classical communication allows the $\{A_i\}$ and $\{B_j\}$ to be correlated. So the most general operation under LOCC can be written as

$$\Lambda(\rho^{\otimes N}) = \frac{1}{\mathrm{tr}} \sum_i (A_i^\dagger \otimes B_i^\dagger)\, \rho^{\otimes N} (A_i \otimes B_i).$$

Note that the operators $\{A_i\}$ and $\{B_i\}$ need not be square and for our purposes they will generally not be. If ρ is distillable then there has to be a finite N such that the outcome, by an appropriate choice of $\{A_i\}$ and $\{B_i\}$, is an inseparable state $\widetilde{\rho}_2$ of two qubits $\widetilde{\rho}_2 = \frac{1}{\mathrm{tr}} \sum_i (A_i^\dagger \otimes B_i^\dagger)\, \rho^{\otimes N} (A_i \otimes B_i)$. The normalization factor tr will in the following discussion be neglected, as all considerations here are independent of positive factors.

As $\widetilde{\rho}_2$ is assumed to be inseparable at least one of the summands also has to correspond to an inseparable state as inseparability cannot be attained by summing separable states. Without loss of generality the two qubit state

$$\rho_i = (A_i^\dagger \otimes B_i^\dagger)\, \rho^{\otimes N} (A_i \otimes B_i) \tag{3.1.1}$$

is inseparable for a henceforth fixed index i. This relation to an inseparable two qubit state will later allow us to apply the sufficiency of the partial transposition in this case to the higher dimensional $\rho^{\otimes N}$.

Now as the operators A_i and B_i act on two dimensional Hilbert spaces, they can be represented as

$$A_i = |\psi_A\rangle\langle 0_A| + |\phi_A\rangle\langle 1_A|$$
$$B_i = |\psi_B\rangle\langle 0_B| + |\phi_B\rangle\langle 1_B|$$

for orthonormal bases $\{|0_{A/B}\rangle, |1_{A/B}\rangle\}$ and (non-normalized) vectors $\psi_{A/B}, \phi_{A/B} \in \mathcal{H}_{A/B}^{\otimes N}$.

Let P_A be the two-dimensional projection operators projecting onto the support of A_i, i.e. the space spanned by $|\psi_A\rangle$ and $|\phi_A\rangle$ and similar for P_B. So as a projection acts as the identity on its image space we have $P_A A_i = A_i$ and $P_B B_i = B_i$. Substituting this into (3.1.1) and using the fact $P^\dagger_{A/B} = P_{A/B}$ we obtain

$$\rho_i = (A_i^\dagger \otimes B_i^\dagger)(P_A \otimes P_B\, \rho^{\otimes N} P_A \otimes P_B)(A_i \otimes B_i).$$

It is obvious that product operators cannot transform a separable state into an inseparable one. So if ρ_i is inseparable then

$$\rho' = P_A \otimes P_B\, \rho^{\otimes N} P_A \otimes P_B$$

must also be inseparable.

In order to apply Peres' partial transposition criterion to this state ρ' we have to choose a basis to compute the transpose in. Let $\{|e_1\rangle, |e_2\rangle\}$ be an orthonormal basis for the support of P_A, i.e. the space spanned by $|\psi_A\rangle$ and $|\phi_A\rangle$ and correspondingly choose an orthonormal basis $\{|f_1\rangle, |f_2\rangle\}$ for the support of P_B. Extend both of these to obtain $\{|e_i\rangle\}$, $\{|f_i\rangle\}$ as orthonormal bases for $\mathcal{H}_A^{\otimes N}$ and $\mathcal{H}_B^{\otimes N}$ respectively. Thus we obtain $\{|e_i\rangle \otimes |f_j\rangle\}$ as an orthonormal basis for $\mathcal{H}^{\otimes N}$. With respect to this basis the matrix of ρ' has entries

$$\rho'_{m\mu,n\nu} = \langle P_A e_m \otimes P_B f_\mu|\, \rho^{\otimes N} |P_A e_n \otimes P_B f_\nu\rangle,$$

which is only possibly non-zero for $1 \leq m, n, \mu, \nu \leq 2$ as the projections vanish otherwise. This defines a 4×4 matrix M, which can be thought of as describing a two-qubit state. Only this matrix M changes under partial transposition as the rest of ρ' is zero. If M remains positive under partial transposition then by the sufficiency of Peres' criterion (see Proposition 2.4.1) M represents a separable state. When this two-qubit state is embedded into $\mathcal{H}^{\otimes N}$ it remains separable. Hence ρ'

would also be separable, which is a contradiction. So the partial transpose M^{T_2} must have a negative eigenvalue.

Now M consists of *all* non-zero entries of ρ'. So ${\rho'}^{T_B}$ is the embedding of M^{T_2} into the higher dimensional space and must also have a negative eigenvalue corresponding to, say, $|\theta\rangle \in \mathcal{H}^{\otimes N}$. In fact, this vector must be spanned by $\{|e_1 \otimes f_1\rangle, |e_1 \otimes f_2\rangle, |e_2 \otimes f_1\rangle, |e_2 \otimes f_2\rangle\}$ as these vectors span the support of ρ' and hence, by symmetry, of ${\rho'}^{T_B}$. But $P_A \otimes P_B$ acts as the identity on this subspace. Thus it follows that

$$\langle\theta|\,(\rho^{\otimes N})^{T_B}|\theta\rangle = \langle\theta|\,{\rho'}^{T_B}|\theta\rangle < 0,$$

which means that the state $\rho^{\otimes N}$ violates the partial transposition criterion. But the partial transpose of $\rho^{\otimes N}$ is given by $(\rho^{T_B})^{\otimes N}$, the N-fold tensor product of ρ^{T_B}. So $\rho^{\otimes N}$ violates the criterion iff ρ does.

Recapitulating, for ρ to be distillable we must have ρ^{T_B} *not* positive. For 2×2 states this criterion is also sufficient for distillability [9]. For higher dimensions there are density matrices which remain positive under partial transposition (and hence are *not* distillable) but which are inseparable [10]. This kind of entanglement which cannot be distilled or extracted is called *bound entanglement.*

3.2. **The reduction criterion.**

Although we mentioned that not every entangled state is distillable one might at least speculate:

> Can every state violating Peres' criterion be distilled?

The answer is presently unknown. In the following discussion we will establish a separability criterion which is weaker than the partial transposition but which, if violated, guarantees distillability. We will apply some observations established earlier, relating to positive maps Λ. The positive map which will be considered here and which will lead to the desired criterion is

$$\Lambda(A) = \mathbb{1}\,\mathrm{tr}(A) - A. \tag{3.2.1}$$

To see that this is positive consider

$$\begin{aligned} & \langle y|\Lambda(|x\rangle\langle x|)|y\rangle \\ = \quad & \langle y|y\rangle\langle x|x\rangle - \langle y|x\rangle\langle x|y\rangle \\ = \quad & \langle y|y\rangle\langle x|x\rangle - |\langle y|x\rangle|^2 \qquad \geq 0 \quad \text{(Cauchy-Schwartz)} \end{aligned}$$

As this holds for all $|x\rangle$ and $|y\rangle$ we deduce that Λ is indeed positive.

In particular we saw that for any positive map Λ we obtain the necessary criterion that $I \otimes \Lambda\, \rho \geq 0$ for any separable ρ. For the Λ as defined in (3.2.1) this becomes

$$\rho^A \otimes \mathrm{tr}(\rho^B)\mathbb{1} - \rho = \rho^A \otimes \mathbb{1} - \rho \geq 0,$$

and dually we also have for a separable ρ

$$\mathbb{1} \otimes \rho^B - \rho \geq 0.$$

As this criterion involves the reduced density operators it is generally referred to as *reduction* criterion.

Lemma 3.2.1. *Let ρ be the density matrix of a two qubit system. Then*

$$I \otimes \Lambda\, \rho \geq 0 \quad \Leftrightarrow \quad I \otimes T\rho \geq 0$$

where Λ is as above and T is the transposition operator with respect to a fixed basis.

So in the two qubit case both criteria are equivalent so that the reduction criterion is then also sufficient.

Proof.

$$\Lambda(A) = \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} = \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right)^T = \sigma_y A^T \sigma_y$$

where σ_y is the Pauli matrix $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ which is unitary.

$$\begin{array}{rcll} \text{So} & \langle x | I \otimes \Lambda\, \rho\, | x \rangle & \geq 0 & \forall |x\rangle \in \mathcal{H} \\ \Leftrightarrow & \langle (\mathbb{1} \otimes \sigma_y) x | I \otimes \Lambda\, \rho\, | (\mathbb{1} \otimes \sigma_y) x \rangle & \geq 0 & \forall |x\rangle \in \mathcal{H} \\ \Leftrightarrow & \langle x | I \otimes T\rho\, | x \rangle & \geq 0 & \forall |x\rangle \in \mathcal{H} \end{array}$$

□

How is the reduction criterion related to Peres' partial transposition criterion? Unfortunately, it is strictly weaker. That it is certainly not stronger follows from the fact that we can write $\Lambda = \Lambda^{CP} T$ where Λ^{CP} is a completely positive map. (See [11] for a proof.) From this it can be easily seen that $I \otimes \Lambda\, \rho = I \otimes \Lambda^{CP} \rho^{T_B} \geq 0$ if Peres' criterion is satisfied. Now the Werner states for $N \geq 3$ satisfy the reduction criterion but violate Peres' partial transpose criterion. Hence, the reduction criterion is strictly weaker. (The Werner states [12] W_N for an $N \times N$ system are given by

$$W_N := (N^3 - N)^{-1} \left((N - \phi) I + (N\phi - 1) V \right)$$

where $-1 \leq \phi \leq 1$ and V is defined by $V \psi \otimes \theta = \theta \otimes \psi$.)

Before we will go one to see the relationship to distillability let us pause a second to contemplate a simple possible strengthening of this criterion. We saw that Peres' criterion could *not* be strengthened by simply applying it to $\rho^{\otimes N}$ rather than just to ρ. But does this trick work in this case? Let $\widetilde{\Lambda} = I \otimes \Lambda$. Then

$$\widetilde{\Lambda}(\rho_1 \otimes \rho_2) = (\rho_1 \otimes \rho_2)^A \otimes \mathbb{1} - \rho_1 \otimes \rho_2 = (\rho_1^A \otimes \mathbb{1}) \otimes (\rho_2^A \otimes \mathbb{1}) - \rho_1 \otimes \rho_2$$

and so

$$\begin{aligned} & \widetilde{\Lambda}(\rho_1) \otimes \widetilde{\Lambda}(\rho_2) \\ = \quad & (\rho_1^A \otimes \mathbb{1} - \rho_1) \otimes (\rho_2^A \otimes \mathbb{1} - \rho_2) \\ = \quad & \widetilde{\Lambda}(\rho_1 \otimes \rho_2) + 2\rho_1 \otimes \rho_2 - \rho_1 \otimes (\mathbb{1} \otimes \rho_2^A) - (\mathbb{1} \otimes \rho_1^A) \otimes \rho_2 \\ = \quad & \widetilde{\Lambda}(\rho_1 \otimes \rho_2) - \rho_1 \otimes \widetilde{\Lambda}(\rho_1) - \widetilde{\Lambda}(\rho_2) \otimes \rho_2 \end{aligned}$$

So if both $\widetilde{\Lambda}(\rho_1) \geq 0$ and $\widetilde{\Lambda}(\rho_2) \geq 0$ then automatically $\widetilde{\Lambda}(\rho_1 \otimes \rho_2) \geq 0$. So considering the action on products does not yield a more sensitive criterion.

3.3. **A sufficient condition for distillability.**

Theorem 3.3.1. *If a state ρ violates the reduction criterion then it is distillable. Conversely, if a state can be distilled by a protocol consisting of two steps:*

1. *one-side, single-pair filtering*
2. *a protocol that distills a state if and only if $F > 1/N$,*

then the state ρ violates the criterion.

Here $F := \langle \Psi^+ | \, \rho \, | \Psi^+ \rangle$ is the *singlet fraction* and will be discussed in detail later. This theorem justifies the use of the reduction criterion, even if it is weaker than the partial transpose criterion.

Proof. The full proof will *not* be given here but some important aspects of it will be pointed out. See [11] for a complete proof. A crucial part of this proof is the classification of all states ρ such that

$$(U^\dagger \otimes {U^*}^\dagger) \, \rho \, (U \otimes U^*) = \rho \tag{3.3.1}$$

for all unitary operators U. Here * denotes complex conjugation. For this classification fix a product basis $\{|e_i\rangle \otimes |f_j\rangle\}$ for $\mathcal{H}$. It is also convenient to use the shorthand notation $|mn\rangle := |e_m \otimes f_n\rangle$.

So the condition of invariance becomes

$$\langle Ue_m \otimes U^* f_n|\, \rho \,|Ue_p \otimes U^* f_q\rangle = \langle e_m \otimes f_n|\, \rho \,|e_p \otimes f_q\rangle$$

for all unitary operators U. In particular we can consider in turn operators of the form

$$\begin{aligned} U_r|r\rangle &= \; -|r\rangle \quad , \quad U_r|i\rangle = |i\rangle \quad \text{for all } i \neq r \quad \text{and} \\ U_r|r\rangle &= \; i|r\rangle \quad , \quad U_r|i\rangle = |i\rangle \quad \text{for all } i \neq r \quad , \end{aligned}$$

which act on the product basis as

$$\begin{aligned} U_r \otimes U_r^* |ij\rangle &= (-1)^{\delta_{ir}+\delta_{jr}} |ij\rangle \qquad \text{and} \\ U_r \otimes U_r^* |ij\rangle &= (i)^{\delta_{ir}-\delta_{jr}} |ij\rangle \qquad \text{respectively.} \end{aligned}$$

Using Hermiticity and the normalization of ρ and also real 2×2 rotations one obtains a family of *noisy singlet* states ρ_F, which are parameterizable in terms of the singlet fraction $F := \langle \Psi^+ |\, \rho \,| \Psi^+ \rangle$, $0 \leq F \leq 1$ and which are a combination of a maximally mixed state and the generalized singlet state. This family describes *all* states invariant under the action above. A generalized singlet is given by $|\Psi^+\rangle := \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |e_i \otimes e_i\rangle$. (Note that in the 2-dimensional case this gives a slight modification of the "traditional" singlet.) Explicitly, we have

$$\rho_F = \frac{N^2}{N^2-1} \left[(1-F)\frac{I}{N^2} + \left(F - \frac{1}{N^2} \right) P^+ \right] ,$$

where $P^+ := |\Psi^+\rangle\langle\Psi^+|$.

Using the reduction criterion, these states can be seen to be inseparable for $F > 1/N$. But states ρ_F with $F \leq 1/N$ can be produced from a proper product σ state by $U \otimes U^*$ twirling:

$$\rho_F = \int dU \, (U \otimes U^*) \, \sigma (U^\dagger \otimes U^{*\dagger})$$

Here dU denotes uniform probability distribution proportional to the Haar measure on the unitary group of dimension N. By taking this average over all U, the outcome will automatically satisfy the invariance condition (3.3.1). (Observe the similarity to the technique used in the proof of Maschke's Theorem.) Furthermore, twirling preserves F. So by choosing the separable $\sigma = |\phi_A\rangle\langle\phi_A| \otimes |\phi_B\rangle\langle\phi_B|$, where $|\phi_A\rangle = |1\rangle$ and $|\phi_B\rangle = a|1\rangle + b|2\rangle$, we obtain the state ρ_F with $F = |a|^2/N$. As twirling does not transform product states into inseparable states we see that ρ_F is separable for $0 \leq F \leq 1/N$. Hence, ρ_F is separable iff $0 \leq F \leq 1/N$. This observation can then be used to prove Theorem 3.3.1. See [11] for details and a description of the distillation protocol. ☐

4. IRREVERSIBILITY OF DISTILLATION

The existence of bound entangled states, i.e. entangled states from which no entanglement can be distilled, seems to be strong evidence in favour of the hypothesis that entanglement distillation is irreversible. By this we mean, that the amount of pure entanglement, which can be distilled from a large number of copies of a state may be strictly smaller than the amount needed to create this product state. To turn this into a mathematical statement, we will need the definitions of entanglement of distillation, entanglement cost and entanglement of formation.

4.1. Definition of terms.

The *entanglement of distillation* measures the proportional yield, i.e. the asymptotic ratio of the number m of output singlets $|\Psi\rangle = (|00\rangle + \||11\rangle)/\sqrt{2}$ to the number n of input states ρ. Suppose we have a distillation protocol which transforms $\rho^{\otimes n} \to \rho'_n$ such that

$$\lim_{n\to\infty} D(\rho'_n, |\Psi\rangle\langle\Psi|^{\otimes m}) \to 0,$$

where D is the *Bures distance* given by $D(\rho,\rho') := 2\sqrt{1 - F_U(\rho,\rho')}$ and F_U is the *Uhlmann fidelity* defined as $F_U(\rho,\rho') := \operatorname{tr}\left\{(\rho^{1/2}\rho'\rho^{1/2})^{1/2}\right\}$. The number m will depend on n. We define the maximal asymptotic ratio m/n achievable under any possible LOCC distillation protocol as the *entanglement of distillation* $E_D(\rho)$.

By considering the inverse of this process, i.e. a transformation $(|\Psi\rangle\langle\Psi|)^{\otimes m} \to \rho_m$ such that

$$\lim_{m\to\infty} D(\rho_m, \rho^{\otimes n}) \to 0,$$

we obtain the *entanglement cost* $E_C(\rho)$, which is defined as the minimal asymptotic ration m/n, also with respect to all possible LOCC protocols.

The distillation process is irreversible if $E_D(\rho) < E_C(\rho)$. For *pure* states it has been shown that we must have $E_D(\rho) = E_C(\rho)$ [13]. (Also see the excellent discussion in [14].) Note that the mere existence of bound entangled states, for which by definition $E_D = 0$, does not yet constitute a proof as the amount of entanglement needed to create these states could vanish asymptotically. It is therefore perceivable that also $E_C = 0$. By establishing a lower bound on E_C we will show that this is not the case, which proves the irreversibility hypothesis.

4.2. Lower bound on the entanglement cost.

Theorem 4.2.1. *Let P be the operator projecting onto the support of ρ. If for all n we have $\langle e\otimes f|P^{\otimes n}|e\otimes f\rangle \leq \alpha^n$ for all normalized product vectors $|e\otimes f\rangle \in \mathcal{H}^{\otimes n}(= {\mathcal{H}_A}^{\otimes n} \otimes {\mathcal{H}_B}^{\otimes n})$ then $E_C(\rho) \geq -\log\alpha$.*

Proof. The entanglement cost satisfies [15]

$$E_C(\rho) = \lim_{n\to\infty} \frac{E_f(\rho^{\otimes n})}{n}, \tag{4.2.1}$$

where $E_f := \min_{\mathcal{E}} \sum_i p_i E(|\psi_i\rangle\langle\psi_i|)$ is the *entanglement of formation*. The minimization is over the set of all ensembles of pure states $\mathcal{E} = \{p_i, |\psi_i\rangle\langle\psi_i|\}$, such that $\rho^{\otimes n} = \sum_i p_i\, |\psi_i\rangle\langle\psi_i|$. Here E is the *entropy of entanglement* which, for a pure state $|\psi\rangle\langle\psi|$, is defined as $\operatorname{tr}\left(-(|\psi\rangle\langle\psi|)^A \log(|\psi\rangle\langle\psi|)^A\right)$, the van Neumann entropy of the eigenvalues of the reduced density matrix $(|\psi\rangle\langle\psi|)^A$, also referred to as the entropy of entanglement. (For pure states $|\psi\rangle\langle\psi|$ the value does not depend on whether the subsystem $\mathcal{H}_A$ or $\mathcal{H}_B$ is chosen.)

As the $\{\psi_i\}$ are pure states, we can write them in terms of their Schmidt decomposition as $|\psi_i\rangle = \sum_k c_{i,k}|e_{i,k} \otimes f_{i,k}\rangle$ for orthonormal $\{|e_{i,k}\rangle\}$ and $\{|f_{i,k}\rangle\}$. Since every $|\psi_i\rangle$ lies in the support of $\rho^{\otimes n}$, we must have $|\psi_i\rangle\langle\psi_i| \leq P^{\otimes n}$ for all i. In particular this implies

$$|c_{i,k}|^2 = \langle e_{i,k} \otimes f_{i,k}|(|\psi_i\rangle\langle\psi_i|)|e_{i,k} \otimes f_{i,k}\rangle \leq \langle e_{i,k} \otimes f_{i,k}|P^{\otimes n}|e_{i,k} \otimes f_{i,k}\rangle \leq \alpha^n.$$

The last inequality follows from the hypothesis of the theorem.

The reduced density operator $(|\psi_i\rangle\langle\psi_i|)^A$ becomes $\sum_k |c_{i,k}|^2 |e_{i,k}\rangle\langle e_{i,k}|$, which is diagonal. Thus

$$E(\psi_i) = \sum_k -|c_{i,k}|^2 \log|c_{i,k}|^2 \leq \sum_k -|c_{i,k}|^2 \log \alpha^n = -n\log(\alpha). \tag{4.2.2}$$

The inequality follows from the monotonicity of log and the last equation uses $\sum_k |c_{i,k}|^2 = 1$.

As equation (4.2.2) holds for all i, we have that for all ensembles $\{p_i, |\psi_i\rangle\langle\psi_i|\}$

$$\frac{\sum_i p_i E(|\psi_i\rangle\langle\psi_i|)}{n} \geq -\log(\alpha)$$

As this hold for all n, we can use equation (4.2.1) to deduce the result.

□

Hence to prove the irreversibility it is sufficient to exhibit a single bound entangled state ρ for which $E_D(\rho) = 0$ but for which there exists $\alpha < 1$ in the hypothesis of the previous theorem. See [16] for such an example of a state acting on $\mathbb{C}^3 \otimes \mathbb{C}^3$ and for the corresponding calculations.

4.3. **Irreversibility with catalyst.** The result above can be strengthened as it can be shown that borrowing pure entanglement as a catalyst does not help to increase the distillation rate for bound entangled states. In fact we have the following

Theorem 4.3.1. *Given* n *copies of a state* σ_b *with a positive partial transposition and* n' *copies of some other state* σ*, the number of singlet states* $P^+ := |\Psi^+\rangle\langle\Psi^+|$ *that can by asymptotically distilled from them is, at most, the number of singlets required to create* $\sigma^{\otimes n'}$*, i.e.*

$$E_D(\sigma_b^{\otimes n} \otimes \sigma^{\otimes n'}) \leq n' E_C(\sigma).$$

Proof. See [16]. □

5. NEGATIVITY

This section is devoted to the discussion of the entanglement measure *negativity.* Although it does not have a direct physical interpretation it has certain appealing properties as it is easy to compute and gives bounds on the singlet distance, on the teleportation distance and distillation rates at certain error levels. See [17] for a more detailed discussion of negativity and its applications.

5.1. **Definition.** The negativity is based directly on the partial transposition criterion for separability and measures by how much a given state ρ fails to satisfy this criterion. In detail, let ρ^{T_B} as before denote the partial transpose and let $||\,\rho^{T_B}||_1 = \operatorname{tr}\sqrt{(\rho^{T_B})^\dagger \rho^{T_B}}$ be its trace norm. If ρ^{T_B} has eigenvalues $\{\lambda_i\}$ then this norm can be written as $\sum_i |\lambda_i|$. Suppose $\{\mu_i\}$ are the *negative* eigenvalues of ρ^{T_B}. This set is only non-empty if ρ is *not* separable, so for a *separable* ρ we have $\operatorname{tr}\rho = ||\,\rho\,||_1 = 1$. In either case we can always write

$$1 = \operatorname{tr}(\rho^{T_B}) = \sum_i \lambda_i = ||\,\rho^{T_B}||_1 - 2|\sum_j \mu_j|.$$

Hence, the last sum measures the amount by which ρ^{T_B} fails to have only positive eigenvalues. Thus it is intuitively sensible to define the *negativity* as

$$\mathcal{N}(\rho) := \frac{||\,\rho^{T_B}||_1 - 1}{2} = |\sum_j \mu_j|. \tag{5.1.1}$$

The second equality only holds as $\operatorname{tr}(\rho) = 1$. For a general non-normalized Hermitian matrix it is more sensible to define the negativity as the absolute sum of all its negative eigenvalues but we shall only be concerned with density matrices.

5.2. **Monotonicity.** Any measure of entanglement should vanish for separable states and should be non-increasing under LOCC. That the first requirement is satisfied follows from the discussion on Peres' criterion earlier. The second requirement is not so obvious and demands closer examination.

To prove the monotonicity of $\mathcal{N}$ under LOCC we shall need some easy facts concerning the decomposition of Hermitian matrices into sums and the fact that $\mathcal{N}$ is convex. This last fact, i.e. that $\mathcal{N}\left(\sum_i p_i\, \rho_i\right) \leq \sum_i p_i\, \mathcal{N}(\rho_i)$ for all convex linear combinations of Hermitian operators ρ_i, follows immediately from the definition $\mathcal{N}(\rho) = (||\, \rho^{T_B}\,||_1 - 1)/2$ and the observations that the norm $||.||_1$ both satisfies the triangle inequality and is linear for positive factors.

Any Hermitian operator A can be written as the difference of two positive (and hence also Hermitian) operators. As we are only dealing with the finite dimensional case these can be normalized such that

$$A = a_+\rho^+ - a_-\rho^- \tag{5.2.1}$$

for density matrices $\rho^\pm$ and $a_\pm \geq 0$. To see that this is always possible, diagonalize A and split it into two diagonal matrices, one with all positive eigenvalues, one with all negative eigenvalues. Then normalize appropriately.

Lemma 5.2.1. *For any Hermitian matrix A there is a decomposition of the form (5.2.1) which minimizes $a_+ + a_-$. For this decomposition we have $||A||_1 = a_+ + a_-$ and a_- equals the absolute value of the sum of negative eigenvalues of A.*

Proof. For ease of notation I will write $\mathcal{N}$ for the absolute value of the sum of negative eigenvalues of A, although this identification is only correct if $\operatorname{tr} A = 1$. Note that as $\operatorname{tr} A = a_+ - a_-$, minimizing $a_+ + a_-$ is equivalent to minimizing $\operatorname{tr} A + 2a_-$. Thus it is sufficient to minimize a_-.

Let P^- be the projection operator onto the subspace corresponding to negative eigenvalues of A. So $\mathcal{N} = -\operatorname{tr}(AP^-)$. We can rewrite the decomposition (5.2.1) as $a_+\rho^+ = A + a_-\rho^-$ which is a positive operator. So we must have

$$0 \leq \operatorname{tr}\left((A + a_-\rho^-)P^-\right) = -\mathcal{N} + a_- \operatorname{tr}(\rho^- P^-)$$

But $0 \leq \operatorname{tr}(\rho^- P^-) \leq 1$ and so we obtain $\mathcal{N} \leq a_-$. We can obtain equality, and hence minimize a_-, by choosing $a_-\rho^- = -P^- A P^-$ corresponding to the obvious (Jordan) decomposition mentioned earlier. □

We thus obtain the alternative formula for the negativity

$$\mathcal{N} = \inf \left\{ a_- | A^{T_B} = a_+ \rho^+ - a_- \rho^- \right\}, \tag{5.2.2}$$

where the infimum is over all density matrices $\rho^{\pm}$ and $a_{\pm} \geq 0$. As mentioned above this only agrees with (5.1.1) if $\operatorname{tr} A = 1$, which is the case of interest for us.

We are now in a position to prove the monotonicity of $\mathcal{N}$ under LOCC.

Theorem 5.2.2. *If ρ is prepared by LOCC such that it yields state ρ'_i with probability p_i, the* expected *negativity is non-increasing, i.e.*

$$\mathcal{N}(\rho) \geq \sum_i p_i \mathcal{N}(\rho'_i).$$

(Also see [18] for a general discussion and classification of entanglement monotones.)

It is worth pointing out that as only the expected negativity is considered the parties involved could "gamble". They could try to "force" the state into a state with higher negativity by making appropriate measurements. This would fail most of the time and would lead to a loss of entanglement. But although the expected yield will be lower than the initial amount there is a (small) chance of surpassing this limit.

Proof. As before in the discussion of bound entanglement we consider the most general form of LOCC where have a superoperator $\mathcal{F}$ consisting of a family of completely positive maps $\mathcal{M}_i$, such that $\sum_i \mathcal{M}_i(\mathbb{1}) \leq \mathbb{1}$ and $\operatorname{tr} \sum_i \mathcal{M}_i(A) \leq \operatorname{tr} A$ for all A .

With the notation from the statement of the theorem we have $\mathcal{M}_i(\rho) = p_i\, \rho'_i$. If any of the $\mathcal{M}_i$ can be decomposed into the sum of two completely positive maps, i.e. $\mathcal{M}_i = \mathcal{M}_i^{(1)} + \mathcal{M}_i^{(2)}$ then we can invoke the convexity of $\mathcal{N}$ to deduce that the expected negativity after decomposing will be no smaller than in the "compressed" case. Thus it is sufficient to prove the result when all $\mathcal{M}_i$ are "pure". In particular this means that each $\mathcal{M}_i$ has only a single summand in its Kraus decomposition. This last claim holds as operators of the form $\mathcal{Q}(\rho) = R\, \rho\, R^\dagger$ are necessarily completely positive and thus would give a further decomposition. (Simply write out what it means to be completely positive to see that $\mathcal{Q}$ satisfies the requirement.)

So we can assume that each completely positive linear map $\mathcal{M}$ is of the form

$$\mathcal{M}_i(\rho) = (A_i \otimes B_i)\, \rho\, (A_i^\dagger \otimes B_i^\dagger).$$

Here we must have $A_i A_i^\dagger \leq \mathbb{1}_A$ and $B_i B_i^\dagger \leq \mathbb{1}_B$ (as $\sum_i \mathcal{M}_i(\mathbb{1}) \leq \mathbb{1}$).

We will need the identity

$$\begin{aligned} & \mathcal{M}_i(\rho)^{T_B} \\ = \ & \Big((A_i \otimes B_i)\, \rho\, ({A_i}^\dagger \otimes B_i^\dagger)\Big)^{T_B} \\ = \ & (A_i \otimes B_i^*)\, \rho^{T_B} ({A_i}^\dagger \otimes {B_i^*}^\dagger) \\ = \ & \widetilde{\mathcal{M}}_i(\rho^{T_B}), \end{aligned}$$

where

$$\widetilde{\mathcal{M}}_i(\rho) := (A_i \otimes B_i^*)\, \rho\, ({A_i}^\dagger \otimes {B_i^*}^\dagger)$$

It is obvious that $\widetilde{\mathcal{M}}_i$ is also completely positive and that we have $\operatorname{tr} \widetilde{\mathcal{M}}_i(A) \leq \operatorname{tr} A$.

Hence we obtain

$$\begin{aligned} & p_i\, (\rho_i')^{T_B} = \mathcal{M}_i(\rho)^{T_B} = \widetilde{\mathcal{M}}_i(\rho^{T_B}) = \widetilde{\mathcal{M}}_i\left((1+\mathcal{N})\rho^+ - \mathcal{N}\rho^-\right) \\ = \ & (1+\mathcal{N})\widetilde{\mathcal{M}}_i(\rho^+) - \mathcal{N}\, \widetilde{\mathcal{M}}_i(\rho^-). \end{aligned}$$

Now $\operatorname{tr} \widetilde{\mathcal{M}}_i(\rho^\pm) \leq 1$, so we can rewrite the last equation above as

$$(\rho_i')^{T_B} = \frac{1+N}{c_i^+ p_i} \widetilde{\rho}^+ - \frac{N}{c_i^- p_i} \widetilde{\rho}^-, \tag{5.2.3}$$

where $N = \mathcal{N}(\rho)$, $c_i^\pm = 1/\mathrm{tr}(\widetilde{\mathcal{M}}_i(\rho^\pm)) \geq 1$ and $\widetilde{\rho}^\pm = c_i^\pm \widetilde{\mathcal{M}}_i(\rho)^\pm$ are density matrices.

Comparing (5.2.3) with (5.2.2), we see that $\mathcal{N}(\rho_i') \leq N/(c_i^- p_i)$ (as the negativity is the infimum). As $\sum_i \mathcal{M}_i(\mathbb{1}) \leq \mathbb{1}$ we have $\sum_i 1/c_i \leq 1$. Hence

$$\sum_i p_i\, \mathcal{N}(\rho_i') \leq \sum_i N/c_i \leq N = \mathcal{N}(\rho),$$

which establishes the monotonicity. □

6. TELEPORTATION

I use the term *teleportation* to mean sending quantum information using only LOCC and given entanglement. The general setting is as follows. Alice and Bob, which could be spatially separated, share an entangled pair of particles in a given state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^d \otimes \mathbb{C}^d$. Charlie wants to send his (unknown) state ψ acting on $\mathcal{H}_C = \mathbb{C}^d$ to Bob but he only has access to Alice's particle and the parties are allowed trace preserving local operations and classical communication (TPLOCC) in which Alice can act on $\mathcal{H}_A \otimes \mathcal{H}_C$ and Bob on his Hilbert space $\mathcal{H}_B$. Not much generality is lost by assuming that all Hilbert spaces involved have the same dimension as smaller spaces can be embedded in bigger ones. Let T, which

I will refer to as the teleportation protocol, denote the LOCC operations involved. Then Bob's final state will be given by

$$\rho^{\psi}_{\text{Bob}} = \operatorname{tr}_{A,C}[T(\psi \otimes \rho)].$$

This then defines a quantum channel $\Lambda_{T,\rho}$ which sends a general input state ψ to

$$\Lambda_{T,\rho}(\psi) = \rho^{\psi}_{\text{Bob}}.$$

The relationship here between states ρ and channels $\rho_{T,\rho}$ is *not* the isomorphism which will be seen below in (6.1.1). In fact, this relation is not even a 1-1 correspondence and it depends on the protocol T.

6.1. **Observations on channels and fidelity.** The observations introduced in this section will be used in the next section to establish a relation between the maximal singlet fraction and the optimal teleportation fidelity. The proofs are, if at all, only outlined. See [19] for detailed proofs.

Observation 1

$$\rho_{\Lambda} = (I \otimes \Lambda_{\rho}) P_{+} \tag{6.1.1}$$

establishes an isomorphism between density matrices ρ acting on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^d \otimes \mathbb{C}^d$ which satisfy $\rho^A := \operatorname{tr}_B \rho = I/d$ and completely positive trace-preserving maps Λ acting on $\mathcal{B}(\mathbb{C}^d)$. In the following a completely positive trace-preserving map will be called a *channel*.

Generally, the mathematical equivalence above might not be a physical one. Given the channel, one can always obtain the related state by sending one half of the singlet P^+ down the channel. But if Alice and Bob share the state ρ can they create the corresponding channel Λ? The answer is yes in the case of the depolarizing channel, which corresponds to a noisy singlet state and this fact is e.g. used in the proof of Observation 4 which is not given here.

The channel fidelity $f(\Lambda)$ measures how close on average the input state $\sigma \in \mathbb{C}^d$ is to the output state $\Lambda(\sigma)$. It is defined as

$$f(\Lambda) := \int d\phi \, \langle\phi|\Lambda(|\phi\rangle\langle\phi|)|\phi\rangle,$$

where the integral is performed with respect to the uniform distribution $d\phi$ over all pure input states. It gives the average probability over all input states that the output state passes the test of being the input state. For a state ρ we define $f(\rho) := f(\Lambda_{T_0,\rho})$ as the fidelity of the standard teleportation channel [20] using ρ, where the original scheme is adapted such that P^+ guarantees perfect transmission rather than the true singlet state used in the original scheme.

As before when talking about distillability, $F(\rho) := \langle \Psi^+ |\, \rho \,| \Psi^+ \rangle$ denotes the singlet fraction. For a channel Λ we define $F(\Lambda) := F(\rho_\Lambda)$, using (6.1.1). So that explicitly $F(\Lambda) = \langle \Psi | (I \otimes \Lambda) P^+ | \Psi \rangle$.

Observation 2

For a noisy singlet state ρ_p given by

$$\rho_p := p\, P^+ + (1-p) \frac{\mathbb{1} \otimes \mathbb{1}}{d^2}, \quad 0 \le p \le 1,$$

(which is a natural generalization of a 2×2 Werner state), it can be shown that we have

$$f = p + (1-p) \frac{1}{d}$$

and

$$F = p + (1-p) \frac{1}{d^2}.$$

Thus we can deduce

$$f = \frac{Fd + 1}{d+1}. \tag{6.1.2}$$

As any one of the parameters p, f and F uniquely determine the noisy singlet state ρ_p, I shall also write ρ_f and ρ_F depending on the parametrization being used.

Observation 3

Recall that we mentioned earlier when we established criteria for distillability that the states ρ_p are the only states invariant under $U \otimes U^*$ transformations such that any state ρ becomes

$$\rho_F = \int dU \, (U \otimes U^*) \, \rho \, (U^\dagger \otimes {U^*}^\dagger)$$

under the action of twirling, where $F = \operatorname{tr} \rho \, P_+ = F(\rho)$.

Observation 4

There is a similar notion for the twirling of channels. Here Alice subjects the state ψ to be sent to a random unitary transformation $U \psi U^\dagger$. She then sends it to Bob via the channel Λ and informs him of the chosen transformation. Bob on receiving the particle then applies the inverse $U^\dagger\_U$ to it. It can be shown that this twirling turns any channel Λ into a depolarizing channel Λ^{dep} such that $F(\Lambda) = F(\Lambda^{dep})$. Furthermore, it also leaves the channel fidelity f invariant.

Observation 5

For any channel Λ, $f(\Lambda)$ and $F(\Lambda)$ satisfy

$$f(\Lambda) = \frac{F(\Lambda)d + 1}{d+1}$$

Proof. Given any channel apply twirling to it. This leads to a depolarizing channel. But the equation above holds for noisy singlets and hence for any depolarizing channel. Twirling leaves f and F unchanged. So the result must hold for an arbitrary channel. $\square$

6.2. **Maximal singlet fraction and teleportation fidelity.**

Theorem 6.2.1. *Let F_{max} be the maximal possible singlet fraction which can be obtained from a given state ρ by means of TPLOCC. Let f_{max} be the maximal teleportation fidelity via the state ρ attainable by TPLOCC (i.e. $f_{max}(\rho) = \max_T f(\Lambda_{T,\rho})$). Then these two quantities are related by*

$$f_{max} = \frac{F_{max}d + 1}{d+1}. \tag{6.2.1}$$

Proof. $\leq$: Suppose we are given ρ which corresponds to a teleportation channel Λ of fidelity f_{max}. We want to obtain a state ρ' by TPLOCC for which (6.2.1) is satisfied. By Observation 5 the singlet fraction F of the channel satisfies $f_{max} = (Fd+1)/(d+1)$. By definition this equals $F(\rho')$ where $\rho' = I \otimes \Lambda P_+$. This state can be created by teleporting (and hence only using TPLOCC) half of the singlet state using the given channel Λ.

$\geq$: Suppose that by TPLOCC a state ρ' with $F(\rho') = F_{max}$ has been obtained. Apply twirling to this state . The result will be the noisy singlet state $\rho_{F_{max}}$ for which (6.1.2) holds. As f and F are invariant under twirling the original state ρ' also satisfies this equation.

$\square$

This theorem in combination with the discussion of bound entanglement allows us to prove the following

Corollary 6.2.2. *The optimal fidelity of teleportation f_{max} which can be obtained using a bound entangled state as a resource equals the classical teleportation fidelity f_{cl}.*

Proof. We shall first establish a formula for the classical teleportation fidelity f_{cl}. Due to Theorem 6.2.1 it suffices to find the maximal singlet fraction F_{max} which can be attained under TPLOCC. In classical teleportation there is no shared entanglement so that F_{max} has to be obtained from separable states. As discussed in the proof of Theorem 3.3.1, $F > 1/d$ corresponds to an entangled state as twirling produces a noisy singlet state for which this is true and twirling does not transform separable into entangled states. But $F = 1/d$ can be attained by the depolarizing channel which is separable for this value of F. Hence we deduce from (6.2.1) that $f_{cl} = 2/(d+1)$.

If we are now given a bound entangled state then this must satisfy $F \leq 1/d$ as we saw in the discussion earlier. Hence, we immediately see from (6.2.1) that such a state can just as well be discarded in favour of classical teleportation.

$\square$

7. Teleportation and dense coding

Another important application of entanglement is *dense coding*, sometimes also referred to as superdense coding. As for a general d-dimensional Hilbert space at most d (orthonormal) states can be told apart with certainty, one might think that by transmitting such a state it is not possible to send more information than by sending a letter from an alphabet of size d via a classical noiseless channel. Dense coding allows us to effectively transmit a letter from an alphabet of size up to d^2 by only transmitting a single d-dimensional state, provided we are given entanglement [21].

As in the case of teleportation our main resource is a density matrix ω acting on $\mathcal{H}_A \otimes \mathcal{H}_B$. The schemes considered here are all noiseless and dense coding and teleportation are supposed to work with perfect reliability. Werner [22] proved the equivalence of all such teleportation and dense coding schemes. This equivalence only holds in the *tight* case when all resources are being used optimally. Generally, teleportation becomes easier as the dimension of the Hilbert space $\mathcal{H}_C$, which is the source of states to be teleported, decreases. On the other hand, dense coding is facilitated if the dimension of the maximally entangled state is increased while the number of messages to be transmitted is decreased. In the tight case we have $\dim \mathcal{H}_A = \dim \mathcal{H}_B = \dim \mathcal{H}_C = d$ and $|X| = d^2$ and the following discussion will only deal with this setup. For ease of notation and as all Hilbert spaces have the same dimension I will simply write $\mathcal{H}$ to denote a d-dimensional Hilbert space.

Here I will define what is meant exactly by teleportation and dense coding schemes. In both cases we have a density operator ω, which will turn out to be maximally entangled, acting on $\mathcal{H} \otimes \mathcal{H}$. Furthermore, in both cases we have for each $x \in X$

- a channel (i.e. unital superoperator) $\Lambda_x : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_B)$,
- a positive operator F_x on $\mathcal{H} \otimes \mathcal{H}$, $\sum_x F_x = \mathbb{1}$.

We say that we have a teleportation scheme if these objects satisfy

$$\sum_{x\in X} \operatorname{tr}(\rho \otimes \omega)(F_x \otimes \Lambda_x(A)) = \operatorname{tr} \rho A, \tag{7.0.2}$$

for all density operators ρ on $\mathcal{H}$ and all Hermitian $A \in \mathcal{B}(\mathcal{H})$.

We think of Alice as applying a measurement to the combined state of Charlies particle and her half of the entangled state ω. She then sends the result x corresponding to the outcome F_x to Bob who then applies the corresponding transformation Λ_x to his particle. If (7.0.2) holds then the probability of measuring any outcome of A on Charlie's initial particle is the same as for the corresponding teleported state which is now in Bob's possession, regardless of the initial measurement x.

For a dense coding scheme we require

$$\operatorname{tr}(\omega(\Lambda_x \otimes I)(F_y)) = \delta_{xy} \tag{7.0.3}$$

to hold.

Here Alice first makes a chosen transformation Λ_x corresponding to the message x she wishes to send. If (7.0.3) holds then Bob can, on receiving Alice's particle, by measuring the observable F_y deduce with certainty which message Alice intended to send.

In the next sections we will show how these two equations are equivalent and that the objects involved have to satisfy the following set of equations.

$$\omega = |\Omega\rangle\langle\Omega| \tag{7.0.4}$$

$$F_x = |\Phi_x\rangle\langle\Phi_x| \tag{7.0.5}$$

$$\Lambda_x(A) = U_x^{\dagger} A U_x \tag{7.0.6}$$

$$|\Phi_x\rangle = |(U_x \otimes \mathbb{1})\Omega\rangle \tag{7.0.7}$$

Here the U_x are unitary operators and $|\Omega\rangle$ is maximally entangled.

7.1. **Dense Coding $\Rightarrow$ Teleportation Scheme.** Suppose ω is decomposed into pure states such that $\omega = \sum_i \lambda_i \, \omega_i$ where $\sum_i \lambda_i = 1$. Then (7.0.3) becomes

$$\sum_i \lambda_i \operatorname{tr}(\omega_i(\Lambda_x \otimes I)(F_y)) = \delta_{xy}$$

But each $0 \leq \operatorname{tr}(\omega_i(\Lambda_x \otimes I)(F_y)) \leq 1$ as Λ_x preserves trace and $F_y \leq \mathbb{1}$. So we must have $\operatorname{tr}(\omega_i(\Lambda_x \otimes I)(F_y)) = \delta_{xy}$ for each pure component as $\sum_i \lambda_i = 1$. Thus we can for the moment assume that $\omega = |\Omega\rangle\langle\Omega|$ is pure (or alternatively only work with one pure component). We will later deduce that ω is indeed pure by showing that only one pure component is consistent with other elements derived.

Lemma 7.1.1. *Let $\mathcal{H}$ be a D-dimensional Hilbert space, and $\sigma_x, F_x \in \mathcal{B}(\mathcal{H})$, for $x \in X$ where $|X| = D$. Suppose that each σ_x is a density operator, F is an observable, and $\operatorname{tr}(\sigma_x F_y) = \delta_{xy}$ for $x, y \in X$. Then there exists an orthornormal basis $\{|\phi_x\rangle\}$ such that*

$$\sigma_x = F_x = |\phi_x\rangle\langle\phi_x|.$$

Proof. Fix x and consider σ_x. As this is a density operator there exists an orthonormal basis $\{|\phi_x\rangle\}$ such that $\sigma_x = \sum_z \lambda_z |\phi_z\rangle\langle\phi_z|$. We only have to show that this basis does what we want, i.e. that it diagonalizes *all* σ_x and F_x, with eigenvalues either 0 or 1. Without loss of generality we may by relabelling assume that $\lambda_x \neq 0$, as if all $\lambda_z = 0$ then $\sigma_x = 0$ which would violate the trace condition. Thus we obtain $1 = \operatorname{tr}(\sigma_x F_x) = \sum_z \lambda_z \langle\phi_z|F_x\phi_z\rangle$. But we have $0 \leq \langle\phi_z|F_x\phi_z\rangle \leq 1$ for all z as $0 \leq F_x \leq \mathbb{1}$. As $\sum_z \lambda_z = 1$, $0 \leq \lambda_z \leq 1$ and $\lambda_x \neq 0$ we deduce $\langle\phi_x|F_x\phi_x\rangle = 1$ and hence $|F_x\phi_x\rangle = |\phi_x\rangle$ (again using $F_x \leq \mathbb{1}$). This also implies for all x: $|\phi_x\rangle = \mathbb{1}|\phi_x\rangle = \sum_z |F_z\phi_x\rangle = |\phi_x\rangle + \sum_{z\neq x} |F_z\phi_x\rangle$. So $0 = \sum_{z\neq x}\langle\phi_x|F_z\phi_x\rangle$, where each term is between 0 and 1. Hence, $\langle\phi_x|F_z\phi_x\rangle = 0$ for $z \neq x$. As this holds for all x we have $|F_z\phi_x\rangle = 0$ for $z \neq x$. Thus we have diagonalized all F_x as desired. That the σ_y are also diagonal follows by repeating the argument for a fixed F_x, which has now been shown to be also a density operator and noting that we can reuse the same orthornormal basis. □

Note that this lemma also applies to $\alpha_x(F_y) = \delta_{xy}$ for a collection $\{\alpha_x\}$ of functionals $\alpha_x : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ for which $\alpha_x(\mathbb{1}) = 1$ and $\alpha_x(M^\dagger M) \geq 0$, as there is an isomorphism between such functionals and density operators. Hence we can apply the lemma with $D = d^2$ to the functionals $\operatorname{tr}(\omega(\Lambda_x \otimes I)(\_))$ from (7.0.3) and thus we obtain equation (7.0.5). It will be shown later that the $\{|\Phi_x\rangle\}$ are maximally entangled.

The next small step will be to show that $|\Omega\rangle$ has full Schmidt rank. Suppose there existed a projection $P : \mathcal{H}_B \to \mathcal{H}_B$ such that, for all $x \in X$, $0 = \operatorname{tr}(\omega(\Lambda_x \otimes I)(\mathbb{1} \otimes P))$. Let ρ_x be the density matrix associated with the functional such that $\operatorname{tr}(\omega(\Lambda_x \otimes I)(\mathbb{1} \otimes P)) = \operatorname{tr}(\rho_x(\mathbb{1} \otimes P))$. Then, by Lemma (7.1.1), $\rho_x = |\Phi_x\rangle\langle\Phi_x|$. So $0 = \langle\Phi_x|(\mathbb{1} \otimes P)\Phi_x\rangle = \langle(\mathbb{1} \otimes P)\Phi_x|(\mathbb{1} \otimes P)\Phi_x\rangle = ||(\mathbb{1} \otimes P)|\Phi_x\rangle||^2$ for all x. But the $\{|\Phi_x\rangle\}$ form a basis for $\mathcal{H}$, so this cannot happen for $P \neq 0$. To deduce that $|\Omega\rangle$ has full Schmidt rank, we rewrite the equation above as

$$0 = \mathrm{tr}(\rho_x(\mathbb{1} \otimes P)) = \mathrm{tr}(\omega(\Lambda_x \otimes I)(\mathbb{1} \otimes P)) = \mathrm{tr}(\omega(\mathbb{1} \otimes P)) = \langle\Omega|(\mathbb{1} \otimes P)\Omega\rangle$$

But $|\Omega\rangle$ is pure and hence has a Schmidt decomposition $\sum_k \lambda_k |e_k \otimes e_k\rangle$. Thus the above equation becomes $0 = \sum_k |\lambda_k|^2 ||Pe_k||^2$. As this can never be satisfied for $P \neq 0$ we deduce that $|\Omega\rangle$ has full Schmidt rank. In particular, this then implies

$$\begin{aligned} & ((A \otimes \mathbb{1}) - (A' \otimes \mathbb{1}))|\Omega\rangle = 0 \\ \Leftrightarrow \quad & \textstyle\sum_k \lambda_k \, (A - A')|e_k\rangle \otimes |e_k\rangle = 0 \\ \Leftrightarrow \quad & A = A', \end{aligned}$$

as $\lambda_k \neq 0$ for all k. This will be used below to show that all Kraus summands are proportional.

Let $\Lambda_x(A) = \sum_\alpha K_{x,\alpha}^\dagger A \, K_{x,\alpha}$ be the Kraus decomposition of Λ_x. Then the dense coding equation (7.0.3) becomes

$$\sum_\alpha |\langle\Omega|(K_{x,\alpha}^\dagger \otimes \mathbb{1})\Phi_y\rangle|^2 = \delta_{xy}.$$

Hence, $\langle\Omega|(K_{x,\alpha}^\dagger \otimes \mathbb{1})\Phi_y\rangle = \langle(K_{x,\alpha} \otimes \mathbb{1})\Omega|\Phi_y\rangle = 0$ for all $y \neq x$ and all α (for every fixed x). But the $\{|\Phi_y\rangle\}$ form an orthonormal basis. Thus we must have

$$(K_{x,\alpha} \otimes \mathbb{1})\Omega = c_\alpha \Phi_x, \tag{7.1.1}$$

and so

$$((K_{x,\alpha} \otimes \mathbb{1}) - c_{\alpha,\beta}(K_{x,\beta} \otimes \mathbb{1}))\Omega = 0.$$

Invoking the previous observation we conclude that all $K_{x,\alpha}$ are proportional for a fixed x and that Λ_x can be written with a single Kraus summand. As $\Lambda_x(\mathbb{1}) = \mathbb{1}$ we deduce that $U_x := K_x$ is unitary so that (7.0.6) follows. As both sides of (7.1.1) are normalized, we have (7.0.7) after absorbing any for F_x and Λ_x irrelevant phase factor into U_x.

We still have to show that $|\Phi_x\rangle$ (and hence $|\Omega\rangle$) is maximally entangled. Observe that the orthonormality of $\{|\Phi_x\rangle\}$ implies

$$\delta_{xy} = \langle\Phi_x|\Phi_y\rangle = \mathrm{tr}(|\Omega\rangle\langle\Omega|U_x^\dagger U_y \otimes \mathbb{1}) = \mathrm{tr}(\rho \, U_x^\dagger U_y)$$

where ρ is the reduced density operator of $|\Omega\rangle\langle\Omega|$. Thus we can apply the following proposition to deduce that each subsystem is maximally mixed. (See [22] for a proof.)

Proposition 7.1.2. *Let* $\{U_x\} \in \mathcal{B}(\mathcal{H})$ *be unitaries in a* d*-dimensional Hilbert space* $\mathcal{H}$*, and* ρ *a density operator such that* $\mathrm{tr}(U_x^\dagger \, \rho \, U_y) = \delta_{xy}$*. Then* $\rho = (1/d)\mathbb{1}$*.*

The dense coding scheme includes all objects necessary for a teleportation scheme. As we have established the claimed formulae relating these objects, it remains to show that these objects define a teleportation scheme such that equation (7.0.2) is satisfied. As one-dimensional operators of the form $\rho = |\phi_1\rangle\langle\phi_2|$, $A = |\psi_1\rangle\langle\psi_2|$ generate $\mathcal{B}(\mathcal{H})$ and everything is linear, it is sufficient to prove (7.0.2) for these types of operators and this will actually show that the teleportation equation holds for *all* ρ and A. Substituting ρ and A into (7.0.2), one obtains for an individual term of the sum

$$\langle\phi_2 \otimes \Omega|\Phi_x \otimes U_x^\dagger\psi_1\rangle\langle\Phi_x \otimes U_x^\dagger\psi_2|\phi_1 \otimes \Omega\rangle.$$

Making all the relevant substitutions for the terms involved and using Lemma 7.1.3 (see below) the two scalar products turn out to be $(1/d)\langle\phi_2|\psi_1\rangle$, $(1/d)\langle\psi_2|\phi_1\rangle$ respectively. Thus (7.0.2) becomes

$$\sum_x \frac{1}{d^2}\langle\phi_2|\psi_1\rangle\langle\psi_2|\phi_1\rangle = \frac{1}{d^2}\sum_x \mathrm{tr}(\rho A) = \mathrm{tr}(\rho A)$$

as desired.

It now only remains to show that indeed ω is pure. But as we have seen, a single (pure) component determines all other objects completely using the given F_x and Λ_x. In particular we have $|\Omega\rangle = (U_x^\dagger \otimes \mathbb{1})|\Phi_x\rangle$, so that $\omega = |\Omega\rangle\langle\Omega|$ is pure.

In the proof above the following lemma was used.

Lemma 7.1.3. *Let* $|\Omega\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ *be the maximally entangled vector* $\Omega = (1/\sqrt{d})\sum_k |e_k \otimes e_k\rangle$*, where* $\{|e_k\rangle\}$ *is the standard basis of* $\mathbb{C}^d$*. Let* $M \in \mathcal{B}(\mathbb{C}^d)$*, and* $\mu \in \mathbb{C}$*. Then the equation*

$$\langle\phi \otimes \Omega|(\mathbb{1} \otimes M \otimes \mathbb{1})\Omega\psi\rangle = \mu\langle\phi|\psi\rangle$$

holds for all ϕ*,* $\psi \in \mathbb{C}^d$ *iff* $M = d\mu\mathbb{1}$*.*

Proof. The proof is straightforward and follows from inserting the definition of Ω into the LHS. It then uses the fact that $(1/d)\langle\phi|M^T\psi\rangle = \mu\langle\phi|\psi\rangle$ for all ϕ, ψ iff $M^T = d\mu\mathbb{1} = M$. □

7.2. **Teleportation Scheme $\Rightarrow$ Dense Coding.** Conversely, suppose we are given a teleportation scheme. Again we want to deduce that all objects are related via equations (7.0.4) - (7.0.7) and that (7.0.3) is satisfied. The complete proof is far too long and only some points will be highlighted. Again see [22] for details.

Firstly, observe that if the result (7.0.2) holds for all density operators ρ then it holds for *all* (not necessarily Hermitian) $\rho \in \mathcal{B}(\mathcal{H})$. The quickest way to see this is to decompose a general $M \in \mathcal{B}(\mathcal{H})$ into a (complex) linear combination of density operators. As a first step we can split $M = (1/2)(M + M^\dagger) + (1/2)(M - M^\dagger)$ into a sum of a Hermitian $R := (1/2)(M + M^\dagger)$ and an anti-Hermitian term $S := (1/2)(M - M^\dagger)$. This second term can then be used to define $T := iS$, such that $T^\dagger = T$. But each Hermitian operator can be written as a linear combination of density operators as seen before in the section on negativity. Thus, we can apply the equation to each bit and then use linearity to put the bits back together.

Remark 7.2.1. Note that this small observation was not made in the original paper [22]. Hence, the result proven there strictly speaking requires (7.0.2) to hold for *all* (not just density operators) $\rho \in \mathcal{B}(\mathcal{H})$. Furthermore, this observation allows us to prove a slightly stronger result than in [22] as we can assume that A is Hermitian (or even also a density operator) rather than any arbitrary bounded operator, thus limiting ourselves to the case when $\mathrm{tr}(\rho\, A)$ has a nice probabilistic interpretation and not worrying about complex measurements.

Secondly, we have to comment on why it is again justified to initially assume that ω is pure. The following lemma, a corollary of the so-called Radon-Nikodym Theorem for completely positive maps, is both used for this justification and as a crucial ingredient of the proof.

Lemma 7.2.2. *Let $T_\alpha : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be completely positive maps such that $\sum_\alpha T_\alpha = I$. Then there are positive numbers t_α such that $T_\alpha = t_\alpha I$.*

Proof. By decomposing each T_α into its Kraus summands we obtain a finer decomposition of I. So we may as well assume that $T_\alpha(A) = K_\alpha^\dagger A K_\alpha$ with a single Kraus factor. Consider the case when A is of the form $A = |\psi\rangle\langle\psi|$. Then

$$T_\alpha(A) = |K_\alpha^\dagger \psi\rangle\langle K_\alpha^\dagger \psi| \leq \sum_\alpha |K_\alpha^\dagger \psi\rangle\langle K_\alpha^\dagger \psi| = \sigma_\alpha T_\alpha(A) = |\psi\rangle\langle\psi|.$$

But every term in the sum is a positive rank one operator, and as the overall sum is also a rank one operator we conclude that $K_\alpha^\dagger|\psi\rangle = \lambda(\psi)\psi$, $\lambda(\psi) \in \mathbb{C}$. As $|\psi\rangle$ was arbitrary, *every* $|\psi\rangle \in \mathcal{H}$ is an eigenvector of $K_\alpha^\dagger$ which implies that $K_\alpha^\dagger$ is a multiple of the identity. □

This Lemma can then be used as follows. If $\omega = \sum_\alpha \omega_\alpha$ then (7.0.2) becomes

$$\sum_\alpha \operatorname{tr} \sum_{x\in X} (\rho \otimes \omega_\alpha)(F_x \otimes \Lambda_x(A)) = \sum_\alpha \operatorname{tr}(\rho T_\alpha(A)) = \operatorname{tr} \rho A$$

Here the channels T_α are defined by

$$\operatorname{tr}(\rho T_\alpha(A)) := \operatorname{tr} \sum_{x\in X} (\rho \otimes \omega_\alpha)(F_x \otimes \Lambda_x(A)) \qquad \forall \rho$$

So the T_α satisfy the hypothesis of the lemma and thus we conclude that each T_α is proportional to the identity, which implies that (7.0.2) also holds when we replace ω by one component ω_α. So we can for the moment assume that ω is pure and then proceed later as in the previous section.

A starting point for the actual proof is that by fixing an arbitrary maximally entangled vector $|\Xi\rangle \in \mathcal{H} \otimes \mathcal{H}$ we can establish a 1-1 correspondence between operators $A \in \mathcal{B}(\mathcal{H})$ and vectors $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ via $|\Psi\rangle = (A \otimes \mathbb{1})|\Xi\rangle$. In particular, we can write $|\Omega\rangle = (W \otimes \mathbb{1})|\Xi\rangle$.

The proof proceeds by making all objects involved as "concrete" as possible. Thus the Λ_x and F_x are decomposed into their Kraus summands and their spectral decomposition respectively. For the latter the correspondence above is then used.

This turns the LHS of (7.0.2) into a sum over three indices. Lemma 7.2.2 is then applied again to deduce as before that each term in the sum must be a multiple of the identity.

To obtain more explicit expressions for the terms involved, A and ρ are supposed to be of the form $A = |\psi_1\rangle\langle\psi_2|$ and $\rho = |\phi_1\rangle\langle\phi_2|$ for arbitrary $|\phi_i\rangle$, $|\psi_i\rangle$, which is legitimate by the remark at the beginning of this section.

The rather long proof also uses additional lemmas and corollaries which, together with the complete proof can be found in [22].

8. Conclusion

Recapitulating, a detailed overview of ideas related to bipartite entanglement has been given. The partial transposition criterion was introduced and its essence was generalized using positive maps to obtain necessary and sufficient conditions for separability. In the low dimensional case the sufficiency of Peres' criterion was discussed. Making further use of positive maps a sufficient criterion for distillability was presented and the existence of bound entanglement shown. In the development of this criterion the technique of twirling and the invariant states under $U \otimes U^*$ transformations were introduced. Distillation was shown to be irreversible. The monotonicity under LOCC of negativity as a computable measure of entanglement was established. Ideas related to twirling were reused in relating the optimal teleportation fidelity to the max singlet fraction. Finally, dense coding as another application of bipartite entanglement was explained and its equivalence to teleportation schemes in the tight case was outlined.

References

[1] A. Peres, Physical Review Letters **77**, pp. 1413 (1996)
[2] P. Horodecki, Physics Letters A, **232**, pp. 333 (1997)
[3] M. Horodecki, P. Horodecki, R. Horodecki, Physics Letters A, **223**, pp. 1 (1996)
[4] G. Allan, Introduction to Banach Spaces and Algebras, notes for Part III course (2003)
[5] J. de Pillis, Pacific Journal of Mathematics, **23**, pp. 129 (1967)
[6] A. Jamiolkowski, Reports on Mathematical Physics, **3**, pp. 275 (1972)
[7] E. Størmer, Acta Mathematica, **110**, pp. 233 (1963)
[8] S. Woronowicz, Reports on Mathematical Physics, **10**, pp. 165 (1976)
[9] M. Horodecki, P. Horodecki, R. Horodecki, Physical Review Letters, **78**, pp. 574 (1997)
[10] M. Horodecki, P. Horodecki, R. Horodecki, Physical Review Letters, **80**, pp. 5239 (1998)
[11] M. Horodecki, P. Horodecki, Physical Review A, **59**, pp. 4206 (1999)
[12] R. Werner, Physical Review A, **40**, pp. 4277 (1989)
[13] C. Bennett, H. Bernstein, S. Popescu, B. Schumacher, Physical Review A, **53**, pp. 2046 (1996)
[14] M. Nielsen, I. Chuang, Quantum Computation and Quantum Information, CUP (2000)
[15] P. Hayden, M. Horodecki, B. Terhal, Journal of Physics A, **34**, pp. 6891 (2001)
[16] C. Vidal, J. Cirac, Physical Review Letters, **86**, pp. 5803 (2001)
[17] G. Vidal, R. Werner, Physical Review A, **65**, 032315 (2002)
[18] G. Vidal, Journal of Modern Optics, **47**, pp. 355 (2000)
[19] M. Horodecki, P. Horodecki, R. Horodecki, Physical Review A, **60**, pp. 1888 (1999)
[20] C. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W. Wooters, Physical Review Letters, **70**, pp. 1895 (1993)
[21] C. Bennett, S. Wiesner, Physical Review Letters, **69**, pp. 2881 (1992)
[22] R. Werner, Journal of Physics A, **34**, pp. 7081 (2001)