Quantum-Information-Theory
Wintersemester 2000/2001

Prof. M. Lewenstein
Institute for Theoretical Physics
University of Hannover

March 27, 2006
Release 1.04

Dear Reader !
This script is based on the lecture by MACIEJ LEWENSTEIN during the winter semester 2000/2001. We, the authors of this script, tried to unify notation and added some additional notes.
You may print out this document for your personal use and you may also give copies (either eletronically or on paper) to third persons as long as those copies are not modified by you in any way and you don't receive any fee except for a reimbursement of copying cost. You may *not* distribute modified versions of this document without prior written consent of the authors.
This version contains the entire lecture and an additional talk given by ANTONIO ACÍN – found in appendix A – which we gratefully acknowledge. We would also like to thank FLORIAN HULPKE and DAGMAR BRUSS for suggestions and corrections.
You can find the latest version of this document (fully hyperlinked) always at
`http://lodda.iqo.uni-hannover.de/download.php?language=en`
If you find any errors, omissions or ambiguous statements (as well as typing mistakes of course) we would be glad to hear about them.
Please mail your comments to
`Helge.Kreutzmann@itp.uni-hannover.de` or
`Kai.Eckert@itp.uni-hannover.de`.
We hope you enjoy reading this script.

KAI ECKERT HELGE KREUTZMANN

Since March 2004, this script is revised and corrected by RODION NEIGOVZEN whom you can reach at `Rodion@Neigovzen.de`.

Contents

Motivation

This lecture intends to describe the theory from the foundations [1] up to the current research front (see e.g. reviews from [2–4], and recent publications from [5] and [6]).
Its main emphasis is on mathematical describtion of the theory, rather than on possible applications, see [7] for those. Although the intention is to prove all theorems, some previuos mathematical knowledge (as found in e.g. [8], [9], [10] and [11]) is expected.
Quantum information theory is strongly related to **entanglement** theory:

1. Quantum "paradoxes" (EPR, SCHRÖDINGER cat, BELL inequalities).

2. Applications in Quantum Information Processing (QIP) (teleportation, cryptography (i.e. for military communications), data compression and quantum computing).

3. Basic and fundamental aspects of quantum mechanics (quantum correlations).

4. Connections to important challenges of modern mathematics (i.e. theory of positive maps on C^* algebras).

1 Introduction

We consider two or sometimes three (quantum) systems which we label A, B and C. They will also be given names of persons: Alice, Bob and Charlie. Each system has a finite HILBERT space and we arrange the systems such, that the entire HILBERT space can be written as

$$\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B \qquad \dim \mathscr{H}_A = M \leq N = \dim \mathscr{H}_B. \tag{1.1}$$

We adopt the notation

$$\{|e_i\rangle\} \in \mathscr{H}_A \qquad |\psi_A\rangle = \sum_{i=1}^{M} a_i |e_i\rangle \tag{1.2}$$

$$\{|f_i\rangle\} \in \mathscr{H}_B \qquad |\psi_B\rangle = \sum_{j=1}^{N} b_j |f_j\rangle. \tag{1.3}$$

The basis used is arbitrary but fixed. All basis changes will be explicitly noted. Thus any state can be written as

$$|\psi\rangle = \sum_{ij} c_{ij} |e_i\rangle \otimes |f_j\rangle \equiv \sum_{ij} c_{ij} |e_j\, f_j\rangle \qquad \in \mathscr{H}_A \otimes \mathscr{H}_B \tag{1.4}$$

where we will omit the direct product sign $\otimes$ in future equations. The dimension of the combined space is

$$\dim \mathscr{H} = \dim \mathscr{H}_A \cdot \dim \mathscr{H}_B = M \cdot N. \tag{1.5}$$

If Alice and Bob have a system with only two possible eigenstates $|0\rangle$ and $|1\rangle$ (each one is said to have a qubit) we can explicitly write down states in the combined HILBERT space in the following way:

$$|\psi^-\rangle = \frac{1}{\sqrt{2}} [|0\rangle|1\rangle - |1\rangle|0\rangle] = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \tag{1.6}$$

Alternatively we can also write the state vectors as vectors of a four dimensional space:

$$|0\rangle_A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_A \qquad |1\rangle_A = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_A \tag{1.7}$$

$$|0\rangle_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_B \qquad |1\rangle_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_B \tag{1.8}$$

$$|\psi_1\rangle = |0\rangle_A |0\rangle_B = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad |\psi^-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \tag{1.9}$$

The translation is done as follows. Write down Alice vector but reserve N components for each of the M components of Alice. Then write Bobs vector into each component of this created vector multiplying Bobs components with the with the corresponding component from Alice. So in the above case for $|\psi_1\rangle$ first write down $|0\rangle_B$ multiplied by 1, then below $|0\rangle_B$ multiplied by 0.
Instead of describing each state by its wave function it is usually more convenient to use the density matrix (usually labeled ρ) instead, since this concept is more general and allows to describe mixed state also.
Each operator O can be written as

$$O = \sum_{ijkl} O_{kl}^{ij} |e_i\rangle \otimes |f_j\rangle\langle e_k| \otimes \langle f_l| \tag{1.10}$$

where both Alice $\{|e_i\rangle\}$ and Bob $\{|f_j\rangle\}$ have an orthonormal basis:

$$\langle e_i|e_j\rangle = \delta_{ij} \qquad\qquad \langle f_k|f_l\rangle = \delta_{kl} \tag{1.11}$$

This way the combined basis in $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$ is also orthogonal and normal. If we denote each pair of indices as one index $k,k' \in \{1,\ldots,NM\}$ then we can write the operator O as

$$O = \sum_{k,k'} O_{k'}^{k} |\psi_k\rangle\langle\psi_{k'}|. \tag{1.12}$$

Def. 1.1 *1. ρ is an operator on $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$*

2. *ρ is hermitian, i.e. $\rho = \rho^\dagger$. $\Leftrightarrow \langle\varphi|(\rho|\psi\rangle) = (\langle\varphi|\rho)|\psi\rangle$. Written component wise this means $\rho_{kk'} = \rho^*_{k'k}$.*
   Using the spectral theorem we can write

$$\rho = \sum_{k,k'} \rho_{kk'} |\psi_k\rangle\langle\psi_{k'}| = \sum_l \lambda_l |\varphi_l\rangle\langle\varphi_l| \tag{1.13}$$

   where λ_l are eigenvalues of ρ and $|\varphi_l\rangle$ its eigenvectors.

3. *$\rho > 0 \Leftrightarrow \forall|\psi\rangle : \langle\psi|\rho|\psi\rangle \geq 0$. This is equivalent to the statement that all eigenvalues $\lambda_l \geq 0$.*

4. *$\mathrm{Tr}(\rho) = \sum_l \lambda_l = 1$.*

The last two definitions allow to interpret the density matrix in terms of probabilities.

Def. 1.2 *The Kernel of ρ is defined as* $\mathrm{K}\{\rho\} = \{|\psi\rangle \in \mathscr{H} : \rho|\psi\rangle = 0\}$.[1]

[1] Note that this is a linear equation.

We have $\rho = \rho^\dagger \Rightarrow \mathrm{K}\{\rho\}$ is a subspace which is spanned by the eigenvectors with zero eigenvalue.

Def. 1.3 *Range of* ρ : $\mathrm{R}\{\rho\} = \{|\psi\rangle \in \mathscr{H} : \exists |\varphi\rangle : \rho|\varphi\rangle = |\psi\rangle\}$.

Since $\rho = \rho^\dagger$ $\mathrm{R}\{\rho\}$ is a linear subspace of $\mathscr{H}$ spanned by the eigenvectors of ρ with $\lambda > 0$, i. e. if $\rho|\phi_1\rangle = |\psi_1\rangle$ and $\rho|\phi_2\rangle = |\psi_2\rangle$ then

$$\rho\left(\alpha|\phi_1\rangle + \beta|\phi_2\rangle\right) = \alpha|\psi_1\rangle + \beta|\psi_2\rangle. \tag{1.14}$$

Further since $\rho = \rho^\dagger$ we have

$$\mathrm{R}\{\rho\} \quad \perp \quad \mathrm{K}\{\rho\}. \tag{1.15}$$

Proof:
Take $|\psi_1\rangle \in \mathrm{R}\{\rho\}$ and $|\psi_2\rangle \in \mathrm{K}\{\rho\}$. Then $\exists_{|\phi\rangle}$ $\rho|\phi\rangle = |\psi_1\rangle$ and

$$\langle\psi_2|\psi_1\rangle = \langle\psi_2|\rho|\phi\rangle = \langle\rho^\dagger\psi_2|\phi\rangle = \langle\rho\,\psi_2|\phi\rangle = 0\cdot|\phi\rangle = 0. \tag{1.16}$$

We call $\mathrm{r}\{\rho\} := \dim\mathrm{R}\{\rho\}$ the rank of ρ. Similar we have $\mathrm{k}\{\rho\} := \dim\mathrm{K}\{\rho\}$. Since eqn. (1.15) we have $\mathrm{r}\{\rho\}+\mathrm{k}\{\rho\} = \dim\mathscr{H}$. We can also proof this by explicit construction:
If we call the eigenvectors of ρ $|\phi_l\rangle$ with $l = 1,\ldots,\mathrm{r}\{\rho\}$ with eigenvalues $\lambda_l > 0$ we can write ρ as

$$\rho = \sum_{l=1}^{\mathrm{r}\{\rho\}} \lambda_l|\phi_l\rangle\langle\phi_l| \qquad \text{and} \qquad |\psi\rangle = \sum_{l=1}^{\mathrm{r}\{\rho\}} a_l|\phi_l\rangle. \tag{1.17}$$

where $|\psi\rangle$ is an arbitrary state in $\mathrm{R}\{\rho\}$ where at least one $a_l \neq 0$. Now we can explicitly construct the state $|\chi\rangle$ which will be projected on $|\psi\rangle$:

$$|\chi\rangle = \sum_{l=1}^{\mathrm{r}\{\rho\}} \frac{a_l}{\lambda_l}|\phi_l\rangle \qquad \Rightarrow \qquad \rho|\chi\rangle = |\psi\rangle \tag{1.18}$$

Without proof we make the general remark that if $A \neq A^\dagger$ we have $\mathrm{R}\{A\} \perp \mathrm{K}\{A^\dagger\}$ and $\mathrm{R}\{A^\dagger\} \perp \mathrm{K}\{A\}$. We will not need this property.
For this lecture we adapt the notation

$$O = \sum_{k,k'}^{\dim\mathscr{H}} O_{kk'}|k\rangle\langle k'| \tag{1.19}$$

$$O^{\mathrm{T}} = \sum_{k,k'}^{\dim\mathscr{H}} O_{kk'}|k'\rangle\langle k| \qquad \text{and thus} \tag{1.20}$$

$$\left(O^{\mathrm{T}}\right)_{kk'} = O_{k'k}. \tag{1.21}$$

Def. 1.4 *Partial transpose:*
We again have a ρ which acts on $\mathscr{H}_A \otimes \mathscr{H}_B$. If we write

$$\rho = \sum_{ijkl} \rho^{ij}{}_{kl} |i\rangle_A \otimes |j\rangle_B \langle k|_A \otimes \langle l|_B \qquad \textit{then} \tag{1.22}$$

$$\rho^{\mathrm{T}_A} = \sum_{ijkl} \rho^{ij}{}_{kl} |k\rangle_A \otimes |j\rangle_B \langle i|_A \otimes \langle l|_B \qquad \textit{and} \tag{1.23}$$

$$\left(\rho^{\mathrm{T}_A}\right)^{ij}{}_{kl} = \rho^{kj}{}_{il}. \tag{1.24}$$

is the partial transpose[2] *with respect to Alice.*

As an example we choose $\mathscr{H}_A = \mathbb{C}^2$ and $\mathscr{H}_B = \mathbb{C}^N$ then ρ is a $2N \times 2N$ matrix which can be written as

$$\rho = (|0\rangle_A\langle 0|)A + (|0\rangle_A\langle 1|)B + (|1\rangle_A\langle 0|)B^\dagger + (|1\rangle_A\langle 1|)C \tag{1.25}$$

where the $N \times N$ matrices $A = A^\dagger$, B and $C = C^\dagger$ act in Bobs space. Now if we explicitly write down ρ we have

$$\rho = \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix} = \rho^\dagger \qquad \rho^{\mathrm{T}} = \begin{pmatrix} A^{\mathrm{T}} & B^* \\ B^{\mathrm{T}} & C^{\mathrm{T}} \end{pmatrix} \tag{1.26}$$

$$\rho^{\mathrm{T}_A} = \begin{pmatrix} A & B^\dagger \\ B & C \end{pmatrix} \qquad \rho^{\mathrm{T}_B} = \begin{pmatrix} A^{\mathrm{T}} & B^{\mathrm{T}} \\ B^* & C^{\mathrm{T}} \end{pmatrix} \tag{1.27}$$

$$\left(\rho^{\mathrm{T}_A}\right)^{\mathrm{T}_B} = \rho^{\mathrm{T}}. \tag{1.28}$$

Partial transposition is a physically strange operation. Transposition can be understood as time inversion, so partial transposition means that e.g. Alice inverses time while Bob does not.
If both Alice and Bob make a local unitary transformation then this transformation remains unitary even if one of them makes a partial transpose:

$$U = U_A \otimes U_B \tag{1.29}$$

$$\rho^{\mathrm{new}} = U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger \tag{1.30}$$

$$(\rho^{\mathrm{new}})^{\mathrm{T}_A} = U_A^* \otimes U_B \rho^{\mathrm{T}_A} U_A^{\mathrm{T}} \otimes U_B^\dagger \tag{1.31}$$

Furthermore $\rho^{\mathrm{T}_A} \geq 0$ iff $(\rho^{\mathrm{new}})^{\mathrm{T}_A} \geq 0$. Also the trace of a partial transposed state remains invariant under local changes of basis.

[2] Notice again that the basis remains fixed albeit arbitrary.

As an example consider a $\mathbb{C}^3 \otimes \mathbb{C}^N$ system:

$$\rho = \begin{pmatrix} A_{00} & A_{01} & A_{02} \\ A_{01}^\dagger & A_{11} & A_{12} \\ A_{02}^\dagger & A_{12}^\dagger & A_{22} \end{pmatrix} \qquad \rho^{\mathrm{T}_B} = \begin{pmatrix} A_{00}^{\mathrm{T}} & A_{01}^{\mathrm{T}} & A_{02}^{\mathrm{T}} \\ A_{01}^* & A_{11}^{\mathrm{T}} & A_{12}^{\mathrm{T}} \\ A_{02}^* & A_{12}^* & A_{22}^{\mathrm{T}} \end{pmatrix} \tag{1.32}$$

where each A_{ij} is an $N \times N$ matrix and $A_{ii} = A_{ii}^\dagger$.

2 Entanglement and Separability

2.1 Entanglement of Pure States

Def. 2.1 *Entanglement of Pure States*
A pure state, i.e. a projector $|\psi\rangle\langle\psi|$ on a vector $\psi \in \mathscr{H}_a \otimes \mathscr{H}_b$, is entangled iff it is not separable, i.e. $|\psi\rangle$ cannot be written as a product vector $|\psi\rangle = |e, f\rangle$.

Theorem 2.1 SCHMIDT *decomposition*
Every $|\psi\rangle \in \mathscr{H}_A \otimes \mathscr{H}_B$ can be represented in an appropriately chosen basis as

$$|\psi\rangle = \sum_{i=1}^{M} a_i |e_i, f_i\rangle \tag{2.1}$$

where the $|e_i\rangle$ ($|f_i\rangle$) form a part of an orthonormal basis in $\mathscr{H}_A$ ($\mathscr{H}_B$) and $a_i \geq 0$, $\sum_{i=1}^{M} a_i^2 = 1$.

In order to proof this theorem we need the following

Theorem 2.2 *Polar (or Singular Value) Decomposition*
Every $M \times N$ matrix A can be represented as

$$A = U A_d V^\dagger. \tag{2.2}$$

where U and V are unitary matrices and A_d is a diagonal, real positive matrix.

Proof:
$B = AA^\dagger$ is a positive, hermitian $M \times M$ matrix. If B is not singular we can invert it to construct

$$U = \frac{1}{\sqrt{B}} A. \tag{2.3}$$

U is an unitary matrix because

$$UU^\dagger = \frac{1}{\sqrt{B}} A A^\dagger \frac{1}{\sqrt{B}} = \frac{1}{\sqrt{B}} B \frac{1}{\sqrt{B}} = \mathbb{1}. \tag{2.4}$$

We can do the same if B is singular but in this case we only operate on the range. Since B is normal ($BB^\dagger = AA^\dagger AA^\dagger = B^\dagger B$) there exists (by the spectral theorem) a basis where B has only entries in the diagonal: $B = V B_d V^\dagger$ with unitary V and therefore also

$$\sqrt{B} = V \sqrt{B_d} V^\dagger. \tag{2.5}$$

Using $A = \sqrt{B}U$ we have $V\sqrt{B_d}V^\dagger U = A$ which leads to the desired result when we rename $V \to U$, $V^\dagger U \to V^\dagger$ and $\sqrt{B_d} \to A_d$.
Using this we are now able to give a proof for the SCHMIDT decomposition:
Every $|\psi\rangle \in \mathscr{H}_A \otimes \mathscr{H}_B$ can be written as

$$\begin{aligned} |\psi\rangle &= \sum_{i,j=1}^{M,M} A_{ij}|i,j\rangle \\ &= \sum_{k,l=1}^{M,M} \sum_{i,j=1}^{M,M} U_{ik} a_k \delta_{kl} V^*_{jl} |i\rangle|j\rangle. \end{aligned} \tag{2.6}$$

where we used the polar decomposition of A in the the second line. Since $\sum_i U_{ik}|i\rangle = |e_k\rangle$ and $\sum_j V^*_{jk}|j\rangle = |e_k\rangle$ we get

$$|\psi\rangle = \sum_{k=1}^{M} a_k |e_k, f_k\rangle. \tag{2.7}$$

$|e_k\rangle$ and $|f_k\rangle$ form an orthonormal basis in $\mathscr{H}_A$, $\mathscr{H}_B$, respectively, because U and V are unitary.
To give an example we take $\mathscr{H}_A \otimes \mathscr{H}_B \subset \mathbb{C}^2 \otimes \mathbb{C}^2$. In this case the SCHMIDT decomposition can contain up to two terms, i.e. up to two SCHMIDT coefficients a_1, a_2. It is obvious that $|\psi\rangle$ is a product vector iff $a_1 = 0$ and $a_2 = 1$ or vice versa. A state with SCHMIDT coefficients $a_1 = a_2 = \frac{1}{\sqrt{2}}$ is a maximally entangled state. Denoting $\{|e_k\rangle\} = \{|f_k\rangle\} = \{|0\rangle, |1\rangle\}$ the possible maximally entangled states can be written as the so called BELL states

$$|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \tag{2.8}$$

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle). \tag{2.9}$$

Observe that the signs in $|\psi^-\rangle$ and $|\phi^-\rangle$ can be absorbed in the definition of $|10\rangle$ and $|11\rangle$ to achieve $a_i \geq 0$ as in the definition.

Def. 2.2 SCHMIDT *rank*
The SCHMIDT *rank is the number of non-vanishing a_i in the* SCHMIDT *decomposition.*

A state is a pure product state iff its SCHMIDT rank is one. Notice that the SCHMIDT rank is unique since there cannot be two SCHMIDT decompositions with different numbers of non-vanishing coefficients[3].

[3] Suppose there are two decompositions for $|\psi\rangle$,

$$|\psi\rangle = \sum_{i=1}^{s} a_i |e_i f_i\rangle \quad \text{and} \quad |\psi\rangle = \sum_{i=1}^{\tilde{s}} \tilde{a}_i |\tilde{e}_i \tilde{f}_i\rangle, \tag{2.10}$$

Def. 2.3 *Entanglement for pure states*

$$E(|\psi\rangle\langle\psi|) = -\operatorname{Tr}(\rho_B \ln \rho_B) \tag{2.11}$$

is a suitable measure for the entanglement of pure states.

Remember that $\rho_B = \operatorname{Tr}_A(\rho)$ acts in $\mathscr{H}_B$ only. We can expand ρ_B (or ρ_A) in the SCHMIDT basis

$$\begin{aligned}\rho_B &= \operatorname{Tr}_A(|\psi\rangle\langle\psi|) \\ &= \operatorname{Tr}_A\left(\sum_k a_k|e_k\rangle|f_k\rangle \sum_l a_l\langle e_l|\langle f_l|\right) \\ &= \sum_{k=1}^{M} a_k^2\,|f_k\rangle\langle f_k| \qquad (2.12)\end{aligned}$$

$$\rho_A = \sum_{k=1}^{M} a_k^2\,|e_k\rangle\langle e_k| \tag{2.13}$$

to express $E(|\psi\rangle\langle\psi|)$ in terms of the a_k:

$$E(|\psi\rangle\langle\psi|) = -\sum_{k=1}^{M} a_k^2 \ln a_k^2 \;\geq\; 0 \tag{2.14}$$

Especially

$$E(|\psi\rangle\langle\psi|) = 0 \quad \text{iff} \;\; a_k = 0\;\forall\, k \;\; \text{except one } a_{k_0} = 1 \tag{2.15}$$

and we observe that $E(|\psi\rangle\langle\psi|)$ is maximal iff all $|e_k\rangle\langle e_k|$ (or $|f_k\rangle\langle f_k|$) come with the same weight:

$$E(|\psi\rangle\langle\psi|) = \max = \ln M \quad \text{iff} \;\; a_k = \frac{1}{\sqrt{M}}\;\forall\, k. \tag{2.16}$$

So $E(|\psi\rangle\langle\psi|)$ is zero for product states and maximal for maximally entangled states.

---

with $\tilde{s} > s$. Because $\{|e_i\rangle_A\}$, $\{|f_i\rangle_B\}$, $\{|\tilde{e}_i\rangle_A\}$ and $\{|\tilde{f}_i\rangle_B\}$ each form sets of orthonormal vectors there is $|x\rangle_A$ such that $|x\rangle_A \in \text{Span}\{|\tilde{e}_i\rangle_A\}$ but $|x\rangle \notin \text{Span}\{|e_i\rangle_A\}$ and thus we get a contradiction because ${}_A\langle x|\psi\rangle = 0$ from the first decomposition and ${}_A\langle x|\psi\rangle \neq 0$ from the second.

2.2 Entanglement and Separability of Mixed States

Def. 2.4 *Entanglement of Mixed States*
A mixed state ρ is entangled iff it is not separable. It is called separable iff it can be represented as

$$\rho = \sum_{i=1}^{K} p_i \, |e_i, f_i\rangle\langle e_i, f_i| \tag{2.17}$$

where $|e_i\rangle \in \mathscr{H}_A$, $|f_i\rangle \in \mathscr{H}_B$ are arbitrary but normalized, $p_i \geq 0$ with $\sum_{i=1}^{K} p_i = 1$, $\mathbb{N}^+ \ni K \leq (\dim \mathscr{H})^2$ with $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$ [12].

We call the state ρ given above separable because it can be created by Alice producing the state $|e_i\rangle$ with probability p_i and Bob correspondingly creating $|f_i\rangle$ with probability p_i. So entangled states are those states that cannot be created using only local operations and classical communication.

2.3 Entanglement Criteria

Theorem 2.3 PERES
If ρ is separable then $\rho^{T_A} \geq 0$ and $\rho^{T_B} = \left(\rho^{T_A}\right)^T \geq 0$.

Proof:
As ρ is separable it can be written as

$$\rho = \sum_{i=1}^{K} p_i |e_i\rangle|f_i\rangle\langle e_i|\langle f_i| = \sum_{i=1}^{K} p_i |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i| \geq 0 \tag{2.18}$$

and we have

$$\begin{aligned}
\rho^{T_A} &= \sum_{i=1}^{K} p_i \left(|e_i\rangle\langle e_i|\right)^{T_A} \otimes |f_i\rangle\langle f_i| \\
&= \sum_{i=1}^{K} p_i |e_i^*\rangle\langle e_i^*| \otimes |f_i\rangle\langle f_i| \\
&= \sum_{i=1}^{K} p_i |e_i^*, \ f_i\rangle\langle e_i^*, \ f_i| \geq 0.
\end{aligned} \tag{2.19}$$

Note that the second line is valid because $A^\dagger = (A^*)^{\mathrm{T}}$.
For arbitrary dimensions this theorem is only valid in the given direction. The only if direction is only valid in special cases:

Theorem 2.4 HORODEKCI
In $\mathbb{C}^2 \otimes \mathbb{C}^2$ *or* $\mathbb{C}^2 \otimes \mathbb{C}^3$ ρ *is separable iff* $\rho^{T_A} \geq 0$.

The method used here to proof that theorem is the method of *subtracting vectors* [13]. We will give the proof in 7 steps.

Lemma 2.1 *A state* ρ *can always be represented as*

$$\rho = \rho' + \Lambda|\psi\rangle\langle\psi| \quad \textit{where} \;\; \rho' \geq 0 \;\; , \;\; |\psi\rangle \in R\{\rho\} \;\; , \;\; \Lambda \leq \frac{1}{\langle\psi|\frac{1}{\rho}|\psi\rangle}. \tag{2.20}$$

Proof:
Taking arbitrary $|\phi\rangle$ we have

$$\begin{aligned} |\langle\phi|\psi\rangle|^2 &= \left|\langle\phi|\sqrt{\rho}\frac{1}{\sqrt{\rho}}|\psi\rangle\right|^2 \\ &\leq \langle\phi|\rho|\phi\rangle\langle\psi|\frac{1}{\rho}|\psi\rangle \end{aligned} \tag{2.21}$$

where ρ^{-1} is defined over $R\{\rho\}$ only and where we used the SCHWARTZ inequality in the second step. Then we get

$$0 \leq \langle\phi|\rho|\phi\rangle\langle\psi|\frac{1}{\rho}|\psi\rangle - |\langle\phi|\psi\rangle|^2 \tag{2.22}$$

$$0 \leq \langle\phi|\rho|\phi\rangle - \frac{|\langle\phi|\psi\rangle|^2}{\langle\psi|\frac{1}{\rho}|\psi\rangle} \tag{2.23}$$

$$0 \leq \langle\phi|\underbrace{\rho - \frac{|\psi\rangle\langle\psi|}{\langle\psi|\frac{1}{\rho}|\psi\rangle}}_{\rho'}|\phi\rangle \tag{2.24}$$

(the last step is due to $|\langle\phi|\psi\rangle|^2 = \langle\phi|\psi\rangle\langle\psi|\phi\rangle$). So we have $\rho = \rho' + \Lambda|\psi\rangle\langle\psi|$ with $\rho' \geq 0$ for all $\Lambda \leq \frac{1}{\langle\psi|\frac{1}{\rho}|\psi\rangle}$.
If we choose the maximal Λ, ρ' no longer contains ψ in its range and the rank of ρ' is diminished by 1:

$$\mathrm{r}\{\rho'\} = \mathrm{r}\{\rho\} - 1 \quad \text{iff} \quad \Lambda = \frac{1}{\langle\psi|\frac{1}{\rho}|\psi\rangle} \tag{2.25}$$

Proof:
...

Lemma 2.2 *If ρ has positive partial transposition (ρ is a PPT state) and if there exists a product vector in the range of ρ, $|e,f\rangle \in \mathrm{R}\{\rho\}$, such that $|e^*,f\rangle \in \mathrm{R}\{\rho^{\mathrm{T}_A}\}$ then ρ can be written as*

$$\rho = \rho' + \Lambda |e,f\rangle\langle e,f| \;\; \textit{with} \;\; \rho' \geq 0\,, \;\; \left(\rho'\right)^{\mathrm{T}_A} \geq 0 \tag{2.26}$$

where

$$\Lambda \leq \min\left\{ \frac{1}{\langle e,f|\frac{1}{\rho}|e,f\rangle}, \; \frac{1}{\langle e^*,f|\frac{1}{\rho^{\mathrm{T}_A}}|e^*,f\rangle} \right\}. \tag{2.27}$$

The proof is clear using lemma 2.1.

Lemma 2.3 *If ρ is a PPT state in $\mathbb{C}^2 \otimes \mathbb{C}^N$ and $\rho|e,f\rangle = 0$ then ρ can be written as*

$$\rho = \rho' + \Lambda |\hat{e},f\rangle\langle \hat{e},f| \;\; \textit{with} \;\; \Lambda = \frac{1}{\langle \hat{e},h|\frac{1}{\rho}|\hat{e},h\rangle} \tag{2.28}$$

where

$$\rho' \geq 0, \; \left(\rho'\right)^{\mathrm{T}_A} \geq 0, \; \langle e|\hat{e}\rangle = 0 \tag{2.29}$$

and

$$\mathrm{r}\{\rho'\} = \mathrm{r}\{\rho\} - 1, \; \mathrm{r}\{\left(\rho'\right)^{\mathrm{T}_A}\} = \mathrm{r}\{\rho^{\mathrm{T}_A}\} - 1. \tag{2.30}$$

This means knowing a product vector in the kernel of ρ makes it possible to diminish the rank of ρ and ρ^{T_A} simultaneously.
Proof:
We partially transpose $\langle e,f|\rho|e,f\rangle = 0$ to get $\langle e^*,f|\rho^{\mathrm{T}_A}|e^*,f\rangle = 0$. Since $\rho^{\mathrm{T}_A} \geq 0$ this implies $\rho^{\mathrm{T}_A}|e^*,f\rangle = 0$.[4]
Because $|e\rangle$ lives in $\mathbb{C}^2$ we always have a unique orthogonal $|\hat{e}\rangle$: $\langle e|\hat{e}\rangle = 0$.
Partially transposing $\langle \hat{e}^*|\rho^{\mathrm{T}_A}|e^*,f\rangle = 0$ and $\langle \hat{e}|\rho|e,f\rangle = 0$ we get

$$\langle e|\rho|\hat{e},f\rangle = 0 \;\; \text{and} \;\; \langle e^*|\rho^{\mathrm{T}_A}|\hat{e}^*,f\rangle = 0 \tag{2.31}$$

and since in $\mathbb{C}^2$ $|\hat{e}\rangle$ is unique there exist some $|h\rangle, \; |\tilde{h}\rangle$ such that

$$\rho|\hat{e},f\rangle = |\hat{e},h\rangle \;\; \text{and} \;\; \rho^{\mathrm{T}_A}|\hat{e}^*,f\rangle = |\hat{e}^*,\tilde{h}\rangle. \tag{2.32}$$

[4]Take $\rho = |\psi^-\rangle\langle\psi^-|$ and $|e,f\rangle = |\psi^+\rangle$ to see that this is not always true for $\rho^{\mathrm{T}_A} \ngeq 0$.

Furthermore

$$|h\rangle = \langle\hat{e}|\rho|\hat{e},f\rangle = \langle\hat{e}^*|\rho^{\mathrm{T}_A}|\hat{e}^*,f\rangle = |\tilde{h}\rangle. \tag{2.33}$$

(In the second step we made the partially transposition with respect to Alice which of course does not change $|h\rangle \in \mathscr{H}_B$).
Now we found $|\hat{e},h\rangle \in \mathrm{R}\{\rho\}$ and $|\hat{e}^*,h\rangle \in \mathrm{R}\{\rho^{\mathrm{T}_A}\}$ and we can use these vectors to rewrite ρ according to lemma 2.2. But since

$$\Lambda_\rho = \frac{1}{\langle\hat{e},h|\underbrace{\frac{1}{\rho}|\hat{e},h\rangle}_{|\hat{e},f\rangle}} = \frac{1}{\langle\hat{e},h|\hat{e},f\rangle} = \frac{1}{\langle h|f\rangle} = \frac{1}{\langle\hat{e}^*,h|\frac{1}{\rho^{\mathrm{T}_A}}|\hat{e}^*,h\rangle} = \Lambda_{\rho^{\mathrm{T}_A}} \tag{2.34}$$

(using eqs. (2.32)) one can choose Λ in lemma 2.2 maximal for both, ρ and ρ^{T_A}, and diminish the rank of ρ, ρ^{T_A} simultaneously.
Under which circumstances can we expect to find a vector in the range of ρ? The following lemma shows that this is always possible if $\mathrm{R}\{\rho\}$ is (at least) a two-dimensional subspace of $\mathbb{C}^2 \otimes \mathbb{C}^2$.

Lemma 2.4 *Every 2-dimensional subspace of $\mathbb{C}^2 \otimes \mathbb{C}^2$ contains a product vector.*

Given $|\chi_1\rangle,\ |\chi_2\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ the question is whether the space spanned by these two vectors does contain a product vector or not. This is not obvious if $|\chi_{1,2}\rangle$ are not product vectors.
Proof:
We are searching for a product vector $|e,f\rangle$ ($|e\rangle \in \mathbb{C}^2$, $|f\rangle \in \mathbb{C}^2$) in the given two dimensional subspace.
Of course we can always find $|\psi_1\rangle$, $|\psi_2\rangle$ spanning the two dimensional subspace orthogonal to the subspace that should contain $|e,f\rangle$: $\langle\psi_1|e,f\rangle = 0 = \langle\psi_2|e,f\rangle$
Using a basis $\{|0\rangle,\ |1\rangle\}$ for Alice we can write

$$|e,f\rangle = (|0\rangle + \alpha|1\rangle)\,|f\rangle \tag{2.35}$$

Note that the proof below does not depend on the normalization. Using Schmidt decomposition we can write ($i \in \{1,2\}$):

$$|\psi_i\rangle = |0\rangle|\phi_i^0\rangle + |1\rangle|\phi_i^1\rangle \tag{2.36}$$

where $|\phi_i^{0,1}\rangle \in \mathbb{C}^2$ are fixed by the chosen basis and $|\psi_i\rangle$.
$\langle\psi_i|e,f\rangle = \left(\langle\phi_i^0| + \alpha\langle\phi_i^1|\right)|f\rangle = 0$ leads to the following matrix equation for α and $|f\rangle = (f_1,\ f_2)^{\mathrm{T}}$:

$$\overbrace{\left[\begin{pmatrix}\langle\phi_1^0|\\ \langle\phi_2^0|\end{pmatrix} + \alpha\begin{pmatrix}\langle\phi_1^1|\\ \langle\phi_2^1|\end{pmatrix}\right]}^{M(\alpha)\,\in\mathbb{C}^2\otimes\mathbb{C}^2}\begin{pmatrix}f_1\\ f_2\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}. \tag{2.37}$$

This equation has a nontrivial solution $\{\alpha,\ |f\rangle\}$ (i.e. we have found a product vector) iff we can find α fulfilling $\det M(\alpha)=0$ which of course is always possible since this is a quadratic equation in $\alpha\in\mathbb{C}$.
Note that this proof can easily be extended to $\mathbb{C}^2\otimes\mathbb{C}^N$.

Lemma 2.5 *If ρ is a PPT state, i.e. $\rho^{\mathrm{T}_A}\geq 0$, acting in $\mathbb{C}^2\otimes\mathbb{C}^2$ and $\mathrm{r}\{\rho\}=2$ then ρ is separable.*

Proof:
$\mathrm{r}\{\rho\}=2$ and by lemma 2.4 there exists a product state $|e,f\rangle$ in the kernel of ρ: $\rho|e,f\rangle=0$.
We use this product state with lemma 2.3 to write ρ as $\rho=\rho'+\Lambda|\hat{e},h\rangle\langle\hat{e},h|$. Since $\mathrm{r}\{\rho'\}=\mathrm{r}\{\rho\}-1=1$, ρ' has to be proportional to a projector. Since $(\rho')^{\mathrm{T}_A}\geq 0$ this projector has to be a projector on a product state[5] which means that we can write ρ as $\rho=|m,n\rangle\langle m,n|+\Lambda|\hat{e},h\rangle\langle\hat{e},h|$ and ρ is separable.

Lemma 2.6 *If in $\mathbb{C}^2\otimes\mathbb{C}^2$ $\mathrm{r}\{\rho\}=\mathrm{r}\{\rho^{\mathrm{T}_A}\}=3$ and ρ is a PPT state and*

$$\exists\ |e,f\rangle\in\mathrm{R}\{\rho\}\quad \textit{such that}\ \ |e^*,f\rangle\in\mathrm{R}\{\rho^{\mathrm{T}_A}\} \tag{2.40}$$

then ρ is separable.

Proof:
We can use lemma 2.2 to reduce the rank of ρ *or* ρ^{T_A} by 1 (taking the maximal Λ), thereby keeping the positivity of both of them:

$$\rho=\rho'+\Lambda|e,f\rangle\langle e,f| \tag{2.41}$$

$$\mathrm{r}\{\rho'\}=\mathrm{r}\{\rho\}-1\ \ \textit{or}\ \ \mathrm{r}\{(\rho')^{\mathrm{T}_A}\}=\mathrm{r}\{\rho^{\mathrm{T}_A}\}-1 \tag{2.42}$$

Now by lemma 2.5 we can show that ρ' or $(\rho')^{\mathrm{T}_A}$ are product states. But ρ' is a product state iff $(\rho')^{\mathrm{T}_A}$ is a product state.

Lemma 2.7 *If $\rho\geq 0$ acting in $\mathbb{C}^2\otimes\mathbb{C}^2$ has $\rho^{\mathrm{T}_A}\geq 0$ and $\mathrm{r}\{\rho\}=3$, $\mathrm{r}\{\rho^{\mathrm{T}_A}\}=3$ then $\exists|e,f\rangle\in\mathrm{R}\{\rho\}$ such that $|e^*,f\rangle\in\mathrm{R}\{\rho^{\mathrm{T}_A}\}$.*

[5] To see why this holds write $\rho'=|\psi\rangle\langle\psi|$ in the basis of the SCHMIDT decomposition of $|\psi\rangle$ as $|\psi\rangle=\alpha|11\rangle+\beta|22\rangle$. Then

$$(\rho')^{\mathrm{T}_A}=|\alpha|^2|11\rangle\langle 11|+\alpha\beta^*|21\rangle\langle 12|+\beta\alpha^*|12\rangle\langle 21|+|\beta|^2|22\rangle\langle 22| \tag{2.38}$$

and we have that $(\rho')^{\mathrm{T}_A}\geq 0$ only if $\alpha=0$ or $\beta=0$ since otherwise

$$[-\alpha\langle 12|+\beta\langle 21|]\,(\rho')^{\mathrm{T}_A}\,[-\alpha^*|12\rangle+\beta^*|21\rangle]=-2|\alpha\beta|^2<0. \tag{2.39}$$

On the other hand if (e.g.) $\beta=0$ then $(\rho')^{\mathrm{T}_A}=|\beta|^2|11\rangle\langle 11|\geq 0$.

This means that there exists a small ε such that

$$\rho - \varepsilon |e, f\rangle\langle e, f| \geq 0 \tag{2.43}$$

$$\rho^{\mathrm{T}_A} - \varepsilon |e^*, f\rangle\langle e^*, f| \geq 0. \tag{2.44}$$

Therefore we can choose an appropriate ε so we can reduce the rank of the density matrices or its partial transpose by one. Having this we are finished (see lemma 2.6).
The proof presented here is not the most simple one but it has the advantage of being extensible to the 2×3 case. See [14] for a simpler proof and [12] for the published version.
Proof:
We use the following notation for this proof:

$$\rho = \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix} \tag{2.45}$$

with $A = A^\dagger$ and $C = C^\dagger$.
A and C are invertible. If one of them is not invertible, e.g. C is not invertible and thus has rank 1 then there exists a vector $|f\rangle$ such that $C|f\rangle = 0$. Thus

$$(0, \langle f|) \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix} \underbrace{\begin{pmatrix} 0 \\ |f\rangle \end{pmatrix}}_{|\psi_f\rangle} = (0, \langle f|) \begin{pmatrix} B|f\rangle \\ 0 \end{pmatrix} = 0 \tag{2.46}$$

thus $\langle \psi_f | \rho | \psi_f \rangle = 0$ and since $\rho \geq 0$ also $\rho |\psi_f\rangle = 0$. This means, $|\psi_f\rangle$ is in the kernel and therefore $B|f\rangle$ must be zero also. This means

$$|\psi_f\rangle = |1\rangle \otimes |f\rangle = |1, f\rangle \in \mathrm{K}\{\rho\} \tag{2.47}$$

which is a product vector in the kernel and we can apply lemma 2.3, reduce $\mathrm{r}\{\rho\}$ and $\mathrm{r}\{\rho^{\mathrm{T}_A}\}$ by one and the proof is completed by lemma 2.5.
We separate the proof into several steps:

1. We can choose the basis in Alice at will:

$$|0\rangle_A = \frac{1}{\sqrt{1+|\alpha|^2}} \begin{pmatrix} 1 \\ \alpha \end{pmatrix} \qquad |1\rangle_A = \frac{1}{\sqrt{1+|\alpha|^2}} \begin{pmatrix} -\alpha^* \\ 1 \end{pmatrix} \tag{2.48}$$

   Using this choice of basis we have

$$B_{\mathrm{new}} = {}_A\langle 0|\rho|1\rangle_A = \frac{1}{1+|\alpha|^2}(1, \alpha^*)\rho \begin{pmatrix} -\alpha^* \\ 1 \end{pmatrix} = \frac{1}{1+|\alpha|^2}\tilde{B}(\alpha^*). \tag{2.49}$$

Using this transformation we have a quadratic equation in α^* in each component of $\tilde{B}$. We choose α^* such that $\det\tilde{B} = 0$. This is possible since $\tilde{B}$ is quadratic in α^* and therefore $\det\tilde{B}$ contains a fourth oder polynomial in α^* which has roots in $\mathbb{C}$.

Using this choice $\tilde{B}$ has rank 1.

2. Next we change the basis in Bobs space:

$$\rho \to \mathbb{1}_A \otimes \frac{1}{\sqrt{C}} \rho \, \mathbb{1}_A \otimes \frac{1}{\sqrt{C}} \tag{2.50}$$

This is not a unitary operation but since it is local and keeps hermicity the separability properties are not changed.

The resulting density matrix is now

$$\rho = \begin{pmatrix} A & B \\ B^\dagger & \mathbb{1} \end{pmatrix} \tag{2.51}$$

Here we introduced *new* matrixes A and B in Bobs space which resulted from the previous basis transformations.

3. $r(\rho) = 3$ means that there exists a vector which fulfills

$$\rho \begin{pmatrix} |f\rangle \\ |\tilde{f}\rangle \end{pmatrix} = 0 \tag{2.52}$$

where $|f\rangle$, $|\tilde{f}\rangle$ in $\mathbb{C}^2$. Using the explicit form of ρ given in eqn. (2.51) we get the constraint

$$|\tilde{f}\rangle = -B^\dagger |f\rangle \qquad \text{or} \tag{2.53}$$

$$\begin{pmatrix} |f\rangle \\ -B^\dagger |f\rangle \end{pmatrix} \in \mathrm{K}\{\rho\} \qquad \text{i.e.} \tag{2.54}$$

$$\rho \begin{pmatrix} |f\rangle \\ -B^\dagger |f\rangle \end{pmatrix} = \begin{pmatrix} (A - BB^\dagger)|f\rangle \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.55}$$

This means that $|f\rangle$ is a vector in the kernel of $A - BB^\dagger$. Since $A - BB^\dagger$ acts in a two dimensional space the rank of $A - BB^\dagger$ is at most one and thus

$$A - BB^\dagger = \lambda P \qquad \text{with} \tag{2.56}$$

$$P = |\psi\rangle\langle\psi| \qquad |f\rangle = |\psi^\perp\rangle \qquad \text{and} \tag{2.57}$$

$$\langle\psi^\perp|\psi\rangle = 0. \tag{2.58}$$

The projector P is unique up to a phase.

We can apply the same arguments to ρ^{T_A}. The only difference is that B is exchanged with $B^\dagger$. The result is

$$A = BB^\dagger + \lambda P \qquad P = |\psi\rangle\langle\psi| \tag{2.59}$$

$$A = B^\dagger B + \tilde{\lambda}\tilde{P} \qquad \tilde{P} = |\tilde{\psi}\rangle\langle\tilde{\psi}|. \tag{2.60}$$

If we compute the difference between those equations we have

$$BB^\dagger - B^\dagger B = \tilde{\lambda}\tilde{P} - \lambda P = \lambda(\tilde{P} - P). \tag{2.61}$$

The last equality can be seen if the trace is taken on both sides. The trace of a commutator is zero, the trace of a projector is one. So $0 = \mathrm{Tr}\left(\lambda - \tilde{\lambda}\right)$ or $\lambda = \tilde{\lambda}$.

4. We choose the basis in Bobs space where

$$BB^\dagger - B^\dagger B = \begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix} = \lambda(\tilde{P} - P). \tag{2.62}$$

This choice is possible since $BB^\dagger - B^\dagger B$ is hermitian and $\mathrm{Tr}\left(BB^\dagger - B^\dagger B\right) = 0$. The *new* operators P and $\tilde{P}$ remain projectors since hermicity and rank are not changed by unitary base transformations.

We now consider the most general states (but disregarding an overall phase as it is irrelevant since we are only interested in projectors):

$$|\psi\rangle = \begin{pmatrix} \sqrt{p} \\ \sqrt{1-p}e^{i\varphi} \end{pmatrix} \qquad |\tilde{\psi}\rangle = \begin{pmatrix} \sqrt{1-\tilde{p}} \\ \sqrt{\tilde{p}}e^{i\tilde{\varphi}} \end{pmatrix} \tag{2.63}$$

Using these vectors, we can evaluate eqn. (2.62) component wise:

$$\Lambda = \lambda\left((1-\tilde{p}) - p\right) \tag{2.64}$$

$$-\Lambda = \lambda\left(\tilde{p} - (1-p)\right) \tag{2.65}$$

$$0 = \lambda\left(\sqrt{1-\tilde{p}}\sqrt{\tilde{p}}e^{-i\tilde{\varphi}} - \sqrt{p}\sqrt{1-p}e^{-i\varphi}\right) \tag{2.66}$$

$$0 = \lambda\left(\sqrt{1-\tilde{p}}\sqrt{\tilde{p}}e^{i\tilde{\varphi}} - \sqrt{p}\sqrt{1-p}e^{i\varphi}\right) \tag{2.67}$$

This system is solvable if $\varphi = \tilde{\varphi}$ and $p = \tilde{p}$ or $p = 1 - \tilde{p}$. In the latter case $|\psi\rangle = |\tilde{\psi}\rangle$ causing $B = B^\dagger$ and thus $\rho = \rho^{T_A}$ which is not the most general case. Therefore we choose $p = \tilde{p}$ and $\Lambda > 0$ we have[6]

$$\Lambda = \lambda(1-2p) \qquad \Rightarrow \qquad p < \frac{1}{2}. \tag{2.68}$$

[6] $\Lambda = 0$ again means $B = B^\dagger$ which is not the most general case.

Now we have

$$\rho = \begin{pmatrix} BB^\dagger + \lambda P & B \\ B^\dagger & \mathbb{1} \end{pmatrix} \qquad \rho^{T_A} = \begin{pmatrix} B^\dagger B + \lambda \tilde{P} & B^\dagger \\ B & \mathbb{1} \end{pmatrix}. \tag{2.69}$$

5. $\forall B$ with $r(B) = 1$ $\exists$ always a unitary K such

$$KBK^\dagger = B^T. \tag{2.70}$$

Proof:
Since B has rank one it can be written as

$$B = \eta |f\rangle\langle g| \qquad \text{and} \tag{2.71}$$
$$B^T = \eta |g^*\rangle\langle f^*|. \tag{2.72}$$

This means that

$$K|f\rangle = |g^*\rangle \qquad \Rightarrow \qquad \langle g|K^\dagger = \langle f^*|. \tag{2.73}$$

Such a transformation exists because if K is unitary, i.e. $K^\dagger K = \mathbb{1}$ then

$$\langle g|f\rangle = \langle f^*|g^*\rangle = \langle f^*|K|f\rangle = \langle g|\tilde{K}K|f\rangle \tag{2.74}$$

where $\tilde{K}$ is an yet unkown linear operator. Comparing both sides we see that $\tilde{K} = K^{-1} = K^\dagger$.

Notice that in the following part of the proof it is not sufficient to only claim that for any B $\exists$ always a unitary K such that

$$KBK^* = B^T. \tag{2.75}$$

For later use we note that (starting with eqn. (2.70))

$$B = K^* B^T K^T = K^* K B K^\dagger K^T \tag{2.76}$$
$$\Leftrightarrow K^\dagger K^T B = B K^\dagger K^T. \tag{2.77}$$

If we define $U = K^\dagger K^T$ we can write this as $UB = BU$.

6. K can be explicitly written as

$$K = e^{i\varphi_0} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.78}$$

Proof:
If we define the real matrix (see eqn. (2.62)) $M = BB^\dagger - B^\dagger B = \lambda(P - \tilde{P})$ then we can write using $K^\dagger K = 1$

$$\begin{aligned} KMK^\dagger &= KBB^\dagger K^\dagger - KB^\dagger BK^\dagger = KBK^\dagger KB^\dagger K^\dagger - KB^\dagger K^\dagger KBK^\dagger \\ &= B^{\mathrm{T}}B^* - B^*B^{\mathrm{T}} = -(B^*(B^\dagger)^* - (B^\dagger)^*B^*) = -M^* \\ &= -M = \lambda(\tilde{P} - P) = \begin{pmatrix} -\Lambda & 0 \\ 0 & \Lambda \end{pmatrix}. \end{aligned} \tag{2.79}$$

Writing down both sides explicitly we have

$$\begin{aligned} -\lambda\left(K|\tilde{\psi}\rangle\langle\tilde{\psi}|K^\dagger - K|\psi\rangle\langle\psi|K^\dagger\right) &= \begin{pmatrix} -\Lambda & 0 \\ 0 & \Lambda \end{pmatrix} \\ &= \lambda\left(|\tilde{\psi}\rangle\langle\tilde{\psi}| - |\psi\rangle\langle\psi|\right). \end{aligned} \tag{2.80}$$

Now we assume again the most general parameterization possible:

$$K|\psi\rangle = e^{i\varphi_1}\begin{pmatrix} \sqrt{1-q} \\ \sqrt{q}e^{i\Theta} \end{pmatrix} \qquad K|\tilde{\psi}\rangle = e^{i\varphi_2}\begin{pmatrix} \sqrt{q} \\ \sqrt{1-q}e^{i\Theta} \end{pmatrix} \tag{2.81}$$

with $q \in \mathbb{R}$ and $\varphi_1, \varphi_2, \Theta \in [0 \ldots 2\pi[$. Using eqn. (2.80) we can now explicitly compare the parameters and see that $q \equiv p$ has to be fulfilled (same Λ). To discover K we make the most general ansatz:

$$\begin{aligned} K|\psi\rangle &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \sqrt{p} \\ \sqrt{1-p}e^{i\varphi} \end{pmatrix} = \begin{pmatrix} a\sqrt{p} + b\sqrt{1-p}e^{i\varphi} \\ c\sqrt{p} + d\sqrt{1-p}e^{i\varphi} \end{pmatrix} \\ &\stackrel{!}{=} e^{i\varphi_1}\begin{pmatrix} \sqrt{1-p} \\ \sqrt{p}e^{i\Theta} \end{pmatrix} \end{aligned} \tag{2.82}$$

Here $a, b, c, d \in \mathbb{C}$. Immediately we see that $a = d = 0$ and

$$b = e^{i(\varphi_1 - \varphi)} \qquad c = e^{i(\varphi_1 + \Theta)}. \tag{2.83}$$

Calculating the conditions using $K|\tilde{\psi}\rangle$ results in the same requirements but with φ_1 replaced by φ_2, thus we see that $\varphi_1 = \varphi_2 \equiv \hat{\varphi}$ has to hold.

The matrix U defined above is now diagonal. We know that

$$\begin{aligned} UB &= \begin{pmatrix} c^*b & 0 \\ 0 & b^*c \end{pmatrix}\begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \begin{pmatrix} c^*bb_1 & c^*bb_2 \\ b^*cb_3 & b^*cb_4 \end{pmatrix} \\ &= \begin{pmatrix} c^*bb_1 & b^*cb_2 \\ c^*bb_3 & b^*cb_4 \end{pmatrix} = BU. \end{aligned} \tag{2.84}$$

We see[7] that either $b = c$ or $b_2 = b_3 \equiv 0$. The latter case means that B is diagonal and thus $BB^\dagger - B^\dagger B = 0$. As has been shown already this conditions is not fulfilled for an arbitrary case.

Therefore we now require $b = c$ and get $\Theta = -\hat{\varphi}$. Thus we can write

$$K = \begin{pmatrix} 0 & \exp(i(\hat{\varphi} - \varphi)) \\ \exp(i(\hat{\varphi} - \varphi)) & 0 \end{pmatrix} = e^{i\varphi_0}\sigma_x \tag{2.85}$$

and furthermore we see that

$$K|\psi\rangle = e^{i\hat{\varphi}} \begin{pmatrix} \sqrt{1-p} \\ \sqrt{p}e^{-i\varphi} \end{pmatrix} = e^{i\hat{\varphi}}|\tilde{\psi}^*\rangle \tag{2.86}$$

$$K|\tilde{\psi}\rangle = e^{i\hat{\varphi}}|\psi^*\rangle. \tag{2.87}$$

Since a phase for K is irrelevant we choose for simplicity $\varphi_0 = \hat{\varphi} - \varphi \equiv 0$.

7. Remembering[8] (2.54) and (2.57) we know that

$$\begin{pmatrix} |\psi^\perp\rangle \\ -B^\dagger|\psi^\perp\rangle \end{pmatrix} \in \mathrm{K}\{\rho\} \qquad \begin{pmatrix} |\tilde{\psi}^\perp\rangle \\ -B|\tilde{\psi}^\perp\rangle \end{pmatrix} \in \mathrm{K}\{\rho^{\mathrm{T}_A}\}. \tag{2.88}$$

If we denote $|e, f\rangle$ as the desired product vector and try $|e\rangle$ as $|e\rangle = \begin{pmatrix} 1 \\ z \end{pmatrix}$ with $z \in \mathbb{C}$ we have

$$|e\rangle \otimes |f\rangle = \begin{pmatrix} |f\rangle \\ z|f\rangle \end{pmatrix} \in \mathrm{R}\{\rho\} \qquad |e^*\rangle \otimes |f\rangle = \begin{pmatrix} |f\rangle \\ z^*|f\rangle \end{pmatrix} \in \mathrm{R}\{\rho^{\mathrm{T}_A}\}. \tag{2.89}$$

The scalar product between a vector from the range and a vector from the kernel has to vanish:

$$\langle\psi^\perp|(1 - zB^\dagger)|f\rangle = 0 \tag{2.90}$$

$$\langle\tilde{\psi}^\perp|(1 - z^*B)|f\rangle = 0 \tag{2.91}$$

Since the subspace is two dimensional we know the states orthogonal to $|\psi^\perp\rangle$. Since we have z still available we can require

$$(1 - zB^\dagger)|f\rangle = |\psi\rangle \qquad (1 - z^*B)|f\rangle \sim |\tilde{\psi}\rangle \tag{2.92}$$

[7] Strictly speaking we see only $\arg(b) = \arg(c)$ but since K is unitary b and c have to be phases.
[8] All vectors are in the base corresponding to Bobs (and Alice) last choice of bases.

which means

$$|f\rangle = \frac{1}{1-zB^\dagger}|\psi\rangle = \eta\frac{1}{1-z^*B}|\tilde{\psi}\rangle \tag{2.93}$$

with an still unknown $\eta \in \mathbb{C}$.
From eqn. (2.86) we know

$$\sigma_x|\tilde{\psi}\rangle = e^{i\varphi}|\psi^*\rangle \tag{2.94}$$

so using $\sigma_i^2 = \mathbb{1}$ and eqn. (2.75) we can rewrite eqn. (2.93):

$$\begin{aligned}\frac{1}{1-zB^\dagger}|\psi\rangle &= \eta\sigma_x\sigma_x\frac{1}{1-z^*B}\sigma_x e^{i\varphi}|\psi^*\rangle \\ &= \eta e^{i\varphi}\sigma_x\frac{1}{1-z^*B^{\mathrm{T}}}|\psi^*\rangle \\ &= \eta e^{i\varphi}\sigma_x\left(\frac{1}{1-zB^\dagger}|\psi\rangle\right)^*\end{aligned} \tag{2.95}$$

So if we write

$$\frac{1}{1-zB^\dagger}|\psi\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{2.96}$$

then we have the following requirement:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \eta e^{i\varphi}\sigma_x\begin{pmatrix} v_1^* \\ v_2^* \end{pmatrix} = \eta e^{i\varphi}\begin{pmatrix} v_2^* \\ v_1^* \end{pmatrix} \tag{2.97}$$

Since

$$\frac{v_1}{v_2} = \frac{v_2^*}{v_1^*} \qquad \Rightarrow \qquad \begin{aligned} v_1 &= ve^{i\vartheta} \\ v_2 &= ve^{i(\vartheta+\delta)}\end{aligned} \tag{2.98}$$

we can now solve both equations for η and check for consistency:

$$e^{i\vartheta} = \eta e^{i(\varphi-\vartheta-\delta)} \qquad \Rightarrow \qquad \eta = e^{i(\delta-\varphi+2\vartheta)} \tag{2.99}$$

$$e^{i(\vartheta+\delta)} = \eta e^{i(\varphi-\vartheta)} \qquad \Rightarrow \qquad \eta = e^{i(\delta-\varphi+2\vartheta)} \tag{2.100}$$

Obviously δ can be chosen arbitrarily. Choosing an appropriate value we can write

$$\frac{1}{1-zB^\dagger}|\psi\rangle \sim \begin{pmatrix} 1 \\ e^{i\tilde{\delta}} \end{pmatrix}. \tag{2.101}$$

We have thus far only computed the relative angle η but have not actually fixed z itself. So we can now choose z such

$$\left(e^{i\tilde{\delta}}, -1\right) \frac{1}{1 - zB^\dagger} |\psi\rangle = 0. \tag{2.102}$$

Since $B^\dagger$ has rank 1 (cf. step 1) $(B^\dagger)^2 = \alpha B^\dagger$ with $\alpha \in \mathbb{C}$. This means

$$\frac{1}{1 - zB^\dagger} = \mathbb{1} + f(z) B^\dagger \tag{2.103}$$

$$\Leftrightarrow\ f(z) = \frac{z}{1 - \alpha z} \tag{2.104}$$

and thus we have to solve for every δ

$$\left(e^{i\tilde{\delta}}, -1\right) \left(\mathbb{1} + \frac{z}{1 - \alpha z} B^\dagger\right) |\psi\rangle = 0 \tag{2.105}$$

$$\Leftrightarrow (1 - \alpha z) \underbrace{\left(e^{i\tilde{\delta}}, -1\right) |\psi\rangle}_{c_1} + (z - \alpha z^2) \underbrace{\left(e^{i\delta}, -1\right) B^\dagger |\psi\rangle}_{c_2} = 0 \tag{2.106}$$

$$\Leftrightarrow c_1 + z(c_2 - \alpha c_1) - z^2(\alpha c_2) = 0. \tag{2.107}$$

This equation has a solution for *any* δ and thus there exists always a product vector in the range of ρ.

Now we are done. The following table lists all possible cases and the lemmas used to reduce the rank or to show separability respectively:

r$\{\rho\}$	r$\{\rho^{\mathrm{T}_A}\}$	lemma(s)
4	4	Use lemma 2.1 to reduce either r$\{\rho\}$ or r$\{\rho^{\mathrm{T}_A}\}$
4	3	Use lemma 2.1 to reduce either r$\{\rho\}$ or r$\{\rho^{\mathrm{T}_A}\}$
3	4	Use lemma 2.1 to reduce either r$\{\rho\}$ or r$\{\rho^{\mathrm{T}_A}\}$
3	3	Because of lemma 2.7 and lemma 2.6 ρ is separable
2	x	Because of lemma 2.5 ρ is separable
x	2	Because of lemma 2.5 ρ is separable

This proof can be extended to the 2×3 case as well.

3 PPT Entangled States

3.1 Definition

Def. 3.1 *PPT entangled state*
A state ρ is called a PPT (partial positive transposed) entangled state (sometimes abbreviated as PPTES) iff

1. *it is entangled and*
2. $\rho^{\mathrm{T}_A} \geq 0 \quad (\rightsquigarrow \rho^{\mathrm{T}_B} \geq 0)$.

Remarks:

- In $\mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathbb{C}^2 \otimes \mathbb{C}^3$ a state is a PPT state iff it is separable.
- In systems with more than two particles also more complicated situations are possible, e.g. $\rho^{\mathrm{T}_A} \geq 0$ but $\rho^{\mathrm{T}_B} < 0$, $\rho^{\mathrm{T}_C} > 0$.
- PPT entangled states are also called bound or hidden entangled states because this type of entanglement is not distillable. See [15] for details.

3.2 A Criterion of Separability

Theorem 3.1 P. HORODECKI
If ρ is separable then

$$\exists \ |e, f\rangle \in \mathrm{R}\{\rho\} \ \textit{ such that } \ |e^*, f\rangle \in \mathrm{R}\{\rho^{\mathrm{T}_a}\}. \tag{3.1}$$

Proof:
Writing ρ as

$$\rho = \sum_{k=1}^{K} \lambda_k |e_k, f_k\rangle\langle e_k, f_k|. \tag{3.2}$$

we see that $|e_k, f_k\rangle$ has to be in the range of ρ[9] and because

$$\rho^{\mathrm{T}_A} = \sum_{k=1}^{K} \lambda_k |e_k^*, f_k\rangle\langle e_k^*, f_k|. \tag{3.3}$$

$|e_k^*, f_k\rangle$ is in the range of ρ^{T_A}.
Remark: We can also extract a stronger formulation for the theorem out of this

[9] This holds because $\forall|\psi\rangle \in \mathrm{K}\{\rho\}$: $0 = \langle\psi|\rho|\psi\rangle = \sum \lambda_k |\langle\psi|e_k, f_k\rangle|^2 \ \Leftrightarrow \ |\psi\rangle \perp \forall|e_k, f_k\rangle$.

proof:
ρ is separable iff

$$\exists\{|e_k, f_k\rangle\}_{k=1,\ldots,K} \quad : \quad \mathrm{R}\{\rho\} = \left[\{|e_k, f_k\rangle\}_{k=1,\ldots,K}\right] \tag{3.4}$$

$$\text{and} \quad : \quad \mathrm{R}\{\rho^{\mathrm{T}_A}\} = \left[\{|e_k^*, f_k\rangle\}_{k=1,\ldots,K}\right]. \tag{3.5}$$

This means that the set of $|e_k, f_k\rangle$, $|e_k^*, f_k\rangle$ span $\mathrm{R}\{\rho\}$ and $\mathrm{R}\{\rho^{\mathrm{T}_A}\}$ respectively.

3.2.1 Example in $\mathbb{C}^2 \otimes \mathbb{C}^4$

$$\left(\begin{array}{cccc|cccc} b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & b \\ 0 & 0 & 0 & b & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{1+b}{2} & 0 & 0 & \frac{\sqrt{1-b^2}}{2} \\ b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & b & 0 & \frac{\sqrt{1-b^2}}{2} & 0 & 0 & \frac{1+b}{2} \end{array}\right) \qquad 1 \geq b > 0 \tag{3.6}$$

For this state to be a PPT state ρ has to be positively defined. We can verify this by showing that the various submatrices are positively defined. We find three types of submatrices:

$$\begin{pmatrix} b & b \\ b & b \end{pmatrix} > 0 \tag{3.7}$$

$$\begin{pmatrix} b \end{pmatrix} > 0 \tag{3.8}$$

$$\begin{pmatrix} b & 0 & b \\ 0 & \frac{1+b}{2} & \frac{\sqrt{1-b^2}}{2} \\ b & \frac{\sqrt{1-b^2}}{2} & \frac{1+b}{2} \end{pmatrix} = \begin{pmatrix} b & 0 & b \\ 0 & 0 & 0 \\ b & 0 & b \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1+b}{2} & \frac{\sqrt{1-b^2}}{2} \\ 0 & \frac{\sqrt{1-b^2}}{2} & \frac{1-b}{2} \end{pmatrix} > 0 \tag{3.9}$$

By the same method one shows $\rho^{\mathrm{T}_A} > 0$.
One can show that all the vectors in the kernel of ρ have to have the form

$$(A, B, C, 0, \kappa C, -A, -B, -C) \quad \text{where} \quad \kappa = \sqrt{\frac{1-b}{1+b}}. \tag{3.10}$$

A, B and C are free parameters, i.e. there are three orthogonal vectors in the kernel. Observe that this construction is not valid in the case of $b = 0$ where $\dim \mathrm{K}\{\rho\} = 6$.

By looking for vectors orthogonal to all these vectors in the kernel we can also identify the vectors in the range of ρ. These are product states:

$$\begin{aligned} |e,f\rangle &= |1,\alpha\rangle \otimes |1,\frac{1}{\alpha},\frac{1}{\alpha^2},\frac{1}{\alpha^3}+\kappa\rangle \\ &= \left(1,\frac{1}{\alpha},\frac{1}{\alpha^2},\frac{1}{\alpha^3}+\kappa,\alpha,1,\frac{1}{\alpha},\frac{1}{\alpha^2}+\kappa\alpha\right) \in \mathrm{R}\{\rho\}. \end{aligned} \tag{3.11}$$

In the same way we find $|e,f\rangle \in \mathrm{R}\{\rho^{\mathrm{T}_A}\}$:

$$|e,f\rangle = |1,\beta\rangle \otimes |\frac{1}{\beta^3}+\kappa,\frac{1}{\beta^2},\frac{1}{\beta},1\rangle \tag{3.12}$$

In order to check if ρ is separable we have to find out whether there is a $|e,f\rangle \in \mathrm{R}\{\rho\}$ such that $|e^*,f\rangle \in \mathrm{R}\{\rho^{\mathrm{T}_A}\}$, i.e. if there are α, β such that

$$|1,\alpha\rangle \otimes |1,\frac{1}{\alpha},\frac{1}{\alpha^2},\frac{1}{\alpha^3}+\kappa\rangle = |1,\beta^*\rangle \otimes |\frac{1}{\beta^3}+\kappa,\frac{1}{\beta^2},\frac{1}{\beta},1\rangle \tag{3.13}$$

From Alice's part we observe that $\beta^* = \alpha$ which means that we get the following conditions:

$$\frac{1}{(\alpha^*)^3}+\kappa = 1, \qquad \frac{1}{(\alpha^*)^2} = \frac{1}{\alpha}, \qquad \frac{1}{\alpha^*} = \frac{1}{\alpha^2}, \qquad 1 = \frac{1}{\alpha^3}+\kappa. \tag{3.14}$$

Using the second (or the third) equation we have $\alpha^2 = \alpha^*$ and we see that α has to be a pure phase, $\alpha = e^{i\phi}$, and furthermore $\alpha^3 = 1$. By the first equation this means $\kappa = 0$ which gives $b = 0$ where our construction is not valid. Thus there exists no $|e,f\rangle \in \mathrm{R}\{\rho\}$ such that $|e^*,f\rangle \in \mathrm{R}\{\rho^{\mathrm{T}_A}\}$ and (by theorem 3.1) ρ is not separable for $0 < b < 1$ (which means that it is PPT entangled because $\rho^{\mathrm{T}_A} > 0$).
Other examples use the so called unextendible product bases (UPB) [16]. These are incomplete orthogonal product bases whose complementary subspace does not contain any product vector.[10]
Let $|\psi_i\rangle$ be such an UPB with n members then it one can observe that

$$\rho = \mathscr{N}\left(\mathbb{1} - \sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|\right) \tag{3.15}$$

[10] In $\mathbb{C}^3 \otimes \mathbb{C}^3$ it is easy to see that such a basis is indeed possible. Take 5 orthogonal product vectors $|e_i,f_i\rangle$, $i = 1\ldots5$. The question is if one can find more product vectors orthogonal to these such that all the vectors span the whole space, especially if one can find a product vector in the orthogonal space?

This $|e,f\rangle$ has to fulfill $\langle e,f|e_i,f_i\rangle = \langle e|e_i\rangle\langle f|f_i\rangle = 0\ \forall i$ but if $\langle e|e_1\rangle = 0 = \langle e|e_2\rangle$ and (in the best case) $\langle f|f_3\rangle = 0 = \langle f|f_4\rangle$ then neither $\langle e|e_5\rangle = 0$ nor $\langle f|f_5\rangle = 0$ is possible since Alice's and Bob's space are 3 dimensional only. For explicit examples see [16].

Also notice that in $\mathbb{C}^2 \otimes \mathbb{C}^N$ there exists no unextendible product bases (with less than $2N$ members).

is a PPT entangled state[11] ($\mathcal{N}$ is a normalization factor).

3.3 Edge States

Def. 3.2 *A PPT entangled state δ is called an* edge state *if for any $\varepsilon > 0$ and any $|e, f\rangle$*

$$\delta' = \delta - \varepsilon |e, f\rangle\langle e, f| \tag{3.16}$$

is not a PPT state (i.e. either $\delta' \ngeq 0$ or $(\delta')^{T_A} \ngeq 0$).

This means that it is not possible to subtract a projection on a product state from an edge state without loosing the property of δ being positive definite and PPT. By lemma 2.2 (which was valid in arbitrary dimensions) this can be put in the following form:

Lemma 3.1 *A PPT entangled state ρ is an edge state iff there exists* no *$|e, f\rangle \in \mathrm{R}\{\rho\}$ such that $|e^*, f\rangle \in \mathrm{R}\{\rho^{T_A}\}$.*

Proof: lemma 2.2 states that if ρ is PPT and there exists $|e, f\rangle \in \mathrm{R}\{\rho\}$ such that $|e^*, f\rangle \in \mathrm{R}\{\rho^{T_A}\}$ then ρ can be decomposed as $\rho = \rho' + \Lambda |e, f\rangle\langle e, f|$ keeping ρ' positive definite and PPT. Since by definition the latter is not true for edge states no such $|e, f\rangle$ can exist.
The importance of the edge states in the discussion of entangled states comes from the possibility to decompose PPT entangled states into a separable state and an edge state as stated in the following lemma which we will not proof here. A proof can be found in [13], [17].

Theorem 3.2 LEWENSTEIN, SANPERA
Every PPT entangled state can be written as

$$\rho = \lambda \rho_s + (1 - \lambda)\delta \tag{3.17}$$

where ρ_s is separable and δ is an edge state and $\lambda \leq 1$.
There exists an optimal decomposition of this form for which λ is maximal.

Notice that λ being maximal means that we put all the information about the entanglement in the edge state. The advantage of the edge state δ as opposed to ρ is that is has generically lower rank.
Figure 1 illustrates the space of all states, separable states, PPT states and PPT entangled states. All these sets except the set of PPT entangled states are convex and compact (i.e. bound and closed). Because of their definition the edge states

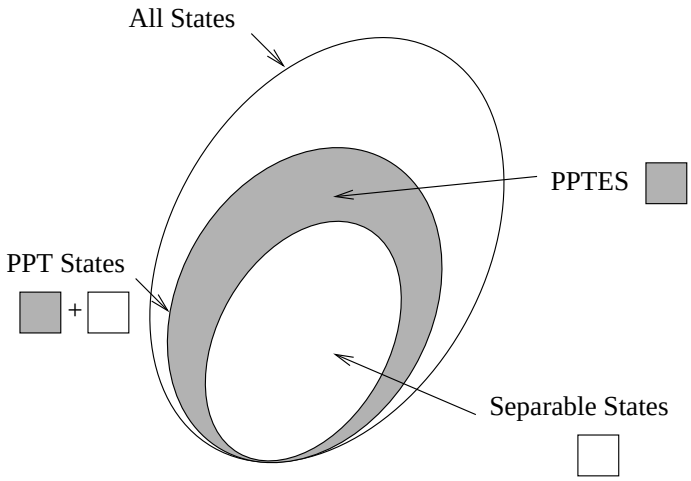

Figure 1: Schematic representation of the space of separable states, entangled states and the PPT entangled states.

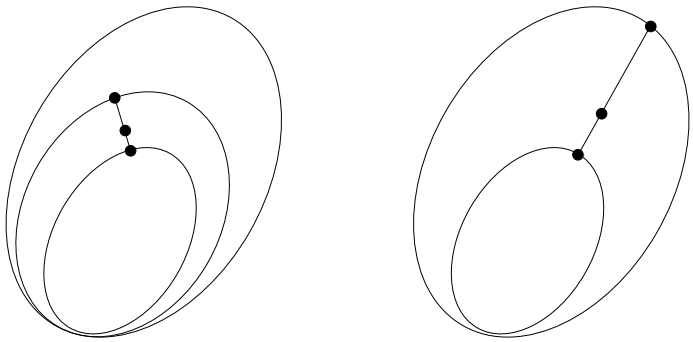

Figure 2: Illustration of lemma 3.2 (left) and lemma 3.3 (right).

can be found on the boundary between the PPT entangled states and the PPT states.
The sum $\rho = a\rho_s + b\delta$ is found by connecting ρ_s and δ by a straight line and dividing the line in the ratio a/b such that ρ is closer to ρ_s if $a > b$ and closer to δ if $b > a$. The left part of figure 2 illustrates the decomposition given in lemma 3.2. That such a decomposition always exists is already obvious from the fact that all the sets are convex.
If we don't care about the PPT entangled states and just look at separable and entangled states it is clear that a similar decomposition has to exist (c.f. also the right part of figure 2). The resulting edge state then lies on the boundary of the entangled states such that subtracting a product projector would result in a not positive definite state. Thus we have the following theorem.

Theorem 3.3 LEWENSTEIN, SANPERA
Every state ρ can be written in a unique way as

$$\rho = \lambda\rho_s + (1-\lambda)\delta \tag{3.18}$$

where ρ_s is separable, $\delta \geq 0$ is entangled and has no product vectors in its range, $\lambda \leq 1$ maximal [12]. Again there exists an optimal decomposition.

[11] Because T_A maps $\mathbb{1}$ to $\mathbb{1}$ and the UPB to another UPB we have $\rho^{T_A} \geq 0$ and furthermore ρ is entangled because by the definition of the UPB there is no product vector in the range.

4 Entanglement Witnesses and Positive Maps

4.1 Entanglement Witnesses

4.1.1 Technical Preface

For several proofs we will need the following

Lemma 4.1 $\mathrm{Tr}\left(\rho^{\mathrm{T}_A}\sigma\right) = \mathrm{Tr}\left(\rho\sigma^{\mathrm{T}_A}\right)$

Proof:
Using the usual notation

$$\sigma = \sum \sigma^{ij}{}_{kl} |ij\rangle\langle kl| \tag{4.1}$$

$$\rho = \sum \rho^{ij}{}_{kl} |ij\rangle\langle kl| \tag{4.2}$$

$$\sigma^{\mathrm{T}_A} = \sum \sigma^{ij}{}_{kl} |kj\rangle\langle il| \tag{4.3}$$

we have

$$\begin{aligned}
\mathrm{Tr}\left(\rho^{\mathrm{T}_A}\sigma\right) &= \mathrm{Tr}\left(\sum_{ijkli'j'k'l'} \rho^{ij}{}_{kl} |kj\rangle\langle il| \sigma^{i'j'}{}_{k'l'} |i'j'\rangle\langle k'l'|\right) \\
&= \sum_{ijkl} \rho^{ij}{}_{kl} \sigma^{il}{}_{kj} \\
&= \mathrm{Tr}\left(\sum_{ijkli'j'k'l'} \rho^{ij}{}_{kl} |ij\rangle\langle kl| \sigma^{i'j'}{}_{k'l'} |k'j'\rangle\langle i'l'|\right) \\
&= \mathrm{Tr}\left(\rho\sigma^{\mathrm{T}_A}\right).
\end{aligned} \tag{4.4}$$

Observation:
The space of linear operators acting on $\mathscr{H}$ (denoted by $\mathscr{B}(\mathscr{H})$) is a HILBERT space itself with the (EUCLEDIAN) scalar product:

$$\langle A|B\rangle = \mathrm{Tr}\left(A^\dagger B\right) \qquad\qquad A,B \in \mathscr{B}(\mathscr{H}) \tag{4.5}$$

This scalar product is equivalent to writing A and B row wise as vectors and scalar multiplying them:

$$\mathrm{Tr}\left(A^\dagger B\right) = \sum_{ij} A^*_{ij} B_{ij} = \sum_{k=1}^{\dim \mathscr{H}^2} a^*_k b_k \tag{4.6}$$

4.1.2 Entanglement Witness

Central to this and the following sections is the HAHN-BANACH theorem which we will present here limited to our situation and without proof (see e.g. [18] for a proof of the more general theorem):

Theorem 4.1 *Let S be a convex compact set in a finite dimensional* BANACH *space. Let ρ be a point in the space with $\rho \notin S$. Then there exists a hyperplane*[12] *that separates ρ from S.*

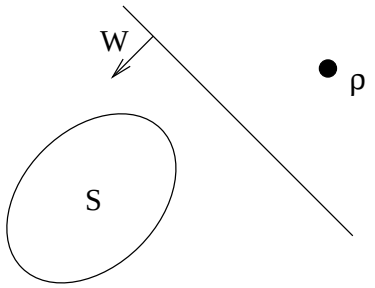

Figure 3: Schematic picture of the HAHN-BANACH theorem. The (unique) unit vector orthonormal to the hyperplane can be used to define *right* and *left* in respect to the hyperplane by using the signum of the scalar product.

Figure 3 motivates the introduction of a new coordinate system located within the hyperplane (supplemented by an orthogonal vector W which is chosen such that it points away from S). Using this coordinate system every state ρ can be characterized by its distance from the plane by projecting ρ onto the chosen orthonormal vector and using the trace as scalar product, i.e. $\mathrm{Tr}(W\rho)$. This measure is either positive, zero or negative. According to our choice of basis in figure 3 every separable state has a positive distance while there are some entangled states with a negative distance. More formally this can be phrased as:

Def. 4.1 *A hermitian operator (an observable) W is called an entanglement witness (EW) iff*

$$\exists\rho \qquad \mathrm{Tr}(W\rho) < 0 \tag{4.7}$$

$$\forall\sigma \in S \qquad \mathrm{Tr}(W\sigma) \geq 0. \tag{4.8}$$

Later on we will choose W such that the set of ρ detected by W is maximized by choosing W tangent to S.

[12] A linear subspace with dimension one less than the dimension of the space itself.

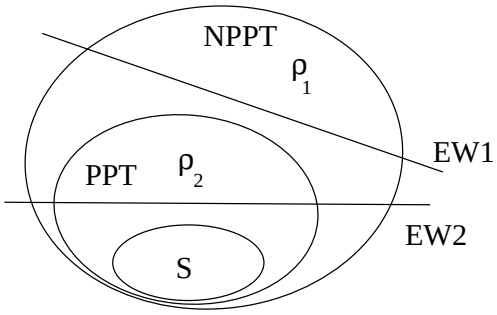

Figure 4: Schematic view of the HILBERT-space with two states ρ_1 and ρ_2 and two witnesses $W1$ and $W2$. $W1$ is a decomposable EW and it does only detect NPPT states like ρ_1. $W2$ is a nd witness and it detects also some PPT states like ρ_2. Note that neither witness detects *all* entangled states.

Def. 4.2 *An EW is decomposable iff there exists operators P, Q with*

$$W = P + Q^{T_A} \qquad P, Q \geq 0. \tag{4.9}$$

Lemma 4.2 *Decomposable EW cannot detect PPT entangled states.*

Proof:
Let δ be a PPT entangled state and EW W be decomposable then

$$\mathrm{Tr}(W\delta) = \mathrm{Tr}(P\delta) + \mathrm{Tr}\left(Q^{T_A}\delta\right) = \mathrm{Tr}(P\delta) + \mathrm{Tr}\left(Q\delta^{T_A}\right) \geq 0. \tag{4.10}$$

Here we used lemma 4.1.

Def. 4.3 *A EW is called non-decomposable entanglement witness (nd-EW) iff there exists at least one PPT entangled state which the witness detects.*

Using these definitions we can restate the consequences of the HAHN-BANACH theorem in several ways:

Theorem 4.2

1. *ρ is entangled iff $\exists$ a witness W such that* $\mathrm{Tr}(\rho W) < 0$.
2. *ρ is a PPT entangled state iff $\exists$ a nd-witness W such that* $\mathrm{Tr}(\rho W) < 0$.
3. *σ is separable iff $\forall$ EW* $\mathrm{Tr}(W\sigma) \geq 0$.

From a theoretical point of view this theorem is quite powerful. However, it is not useful for constructing witnesses that detect a given state ρ.

4.1.3 Examples

1. A decomposable witness

$$W' = P + Q^{T_A} \tag{4.11}$$

detects all separable states σ, i.e.

$$\forall \sigma \in S \qquad \mathrm{Tr}\left(W'\sigma\right) \geq 0. \tag{4.12}$$

Proof:
If σ is separable then it can be written as a convex sum of product vectors (see eqn. (2.17)). So if any product vector $|e,f\rangle$ is detected any separable state will be detected also.

$$\mathrm{Tr}\left(W'|e,f\rangle\langle e,f|\right) = \langle e,f|W'|e,f\rangle \tag{4.13}$$

$$= \underbrace{\langle e,f|P|e,f\rangle}_{\geq 0} + \underbrace{\langle e,f|Q^{T_A}|e,f\rangle}_{\geq 0} \qquad \text{because} \tag{4.14}$$

$$\langle e,f|Q^{T_A}|e,f\rangle = \mathrm{Tr}\left(Q^{T_A}|e,f\rangle\langle e,f|\right) = \mathrm{Tr}\left(Q|e^*,f\rangle\langle e^*,f|\right) \geq 0 \tag{4.15}$$

Here we used lemma 4.1 and $P,Q \geq 0$.

This argumentation shows that $W = Q^{T_A}$ is a suitable witness also.

If we take the simplest case (2×2) we can use

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) \tag{4.16}$$

to create the density matrix

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \qquad Q^{T_A} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}. \tag{4.17}$$

One can quickly verify that indeed $W = Q^{T_A}$ fulfills the witness requirements. Using

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right) \tag{4.18}$$

we can rewrite the witness:

$$W = Q^{T_A} = \frac{1}{2}\left(\mathbb{1} - 2|\psi^-\rangle\langle\psi^-|\right) \tag{4.19}$$

This witness now detects $|\psi^-\rangle$:

$$\mathrm{Tr}\left(W|\psi^-\rangle\langle\psi^-|\right) = -\frac{1}{2} \tag{4.20}$$

Def. 4.4 *The (decomposable) EW W is tangent to S (to P) iff $\exists\sigma \in S$ (exists a $\rho \in P$) with* $\mathrm{Tr}(W\sigma) = 0$ *(*$\mathrm{Tr}(W\rho) = 0$*).*

The witness chosen in eqn. (4.19) is tangent on S because for any state $|e, e^{\perp}\rangle$ (i.e. $|01\rangle$) we have a local unitary transformation

$$U \otimes U|01\rangle = |e, e^{\perp}\rangle \qquad \text{and} \tag{4.21}$$

$$U \otimes U|\psi^-\rangle = e^{i\varphi}|\psi^-\rangle \tag{4.22}$$

because $|\psi^-\rangle$ is a singlet state which must be transformed into a singlet state (with a possible phase) under any unitary transformation.

Now we can calculate

$$\begin{aligned}\langle e, e^{\perp}|\psi^-\rangle\langle\psi^-|e, e^{\perp}\rangle &= \langle 01|U^{\dagger} \otimes U^{\dagger}|\psi^-\rangle\langle\psi^-|U \otimes U|01\rangle \\ &= e^{i\varphi}e^{-i\varphi}\langle 01|\psi^-\rangle\langle\psi^-|01\rangle = \frac{1}{2} \qquad (4.23)\end{aligned}$$

$$\mathrm{Tr}\left(W|e, e^{\perp}\rangle\langle e, e^{\perp}|\right) = \frac{1}{2}\left(1 - 2\frac{1}{2}\right) = 0. \tag{4.24}$$

2. Let ρ be a PPT entangled state with dimension $M \times N$ (and $MN > 6$) then we can write ρ according to theorem 3.2 as

$$\rho = \Lambda\rho_{\mathrm{S}} + (1 - \Lambda)\delta \tag{4.25}$$

where ρ_{S} is a separable state and δ is an edge state and $\Lambda \leq 1$.

Lemma 4.3 *If a nd-EW W detects ρ then it also detects δ, i.e.* $\mathrm{Tr}(W\delta) < 0$.

Proof:

$$0 > \mathrm{Tr}(\rho W) = \underbrace{\mathrm{Tr}(\Lambda\rho_{\mathrm{S}}W)}_{\geq 0} + (1 - \Lambda)\,\mathrm{Tr}(\delta W) \geq (1 - \Lambda)\,\mathrm{Tr}(\delta W) \tag{4.26}$$

Therefore we can concentrate on edge states.

3. We are now looking for nd-EW for edge states.

Def. 4.5

$$\tilde{W} = P_{\mathrm{K}\{\delta\}} + (P_{\mathrm{K}\{\delta^{\mathrm{T}_A}\}})^{\mathrm{T}_A} \tag{4.27}$$

is called a pre-witness. Here $P_{\mathrm{K}\{\delta\}}$ is a projector on the kernel of the edge state δ.

Lemma 4.4 $\forall|e,f\rangle \qquad \langle e,f|\tilde{W}|e,f\rangle \geq \varepsilon > 0.$

Proof:
Let's suppose there exists a state which fulfills

$$0 = \langle e,f|\tilde{W}|e,f\rangle \qquad \text{then} \tag{4.28}$$
$$0 = \langle e,f|P_{\mathrm{K}\{\delta\}}|e,f\rangle + \langle e^*,f|P_{\mathrm{K}\{\delta^{\mathrm{T}_A}\}}|e^*,f\rangle. \tag{4.29}$$

Since any projector fulfills $P \geq 0$ we must have

$$P_{\mathrm{K}\{\delta\}}|e,f\rangle = 0 \qquad \Rightarrow \qquad |e,f\rangle \in \mathrm{R}\{\delta\} \tag{4.30}$$
$$P_{\mathrm{K}\{\delta^{\mathrm{T}_A}\}}|e^*,f\rangle = 0 \qquad \Rightarrow \qquad |e^*,f\rangle \in \mathrm{R}\{\delta^{\mathrm{T}_A}\}. \tag{4.31}$$

This contradicts the properties of edge states as shown in lemma 3.1.

So if we denote

$$0 < \varepsilon_0 = \min_{|e,f\rangle} \langle e,f|\tilde{W}|e,f\rangle \tag{4.32}$$

we can construct a whole family of entanglement witnesses:

$$W = \tilde{W} - \varepsilon\mathbb{1} \qquad 0 < \varepsilon \leq \varepsilon_0 \tag{4.33}$$

W is non-negative on separable states

$$\langle e,f|W|e,f\rangle = \langle e,f|\tilde{W} - \varepsilon\mathbb{1}|e,f\rangle \geq \varepsilon_0 - \varepsilon \geq 0 \tag{4.34}$$

and negative on the edge state δ

$$\mathrm{Tr}\,(W\delta) = \mathrm{Tr}\left(\tilde{W}\delta\right) - \varepsilon = -\varepsilon \tag{4.35}$$

because if we denote a basis of $\mathrm{K}\{\delta\}$ ($\mathrm{K}\{\delta^{\mathrm{T}_A}\}$) with $|k\rangle \in \mathbb{C}^N \otimes \mathbb{C}^M$ ($|\tilde{k}\rangle \in \mathbb{C}^N \otimes \mathbb{C}^M$), $k = 1,\ldots,\dim \mathrm{K}\{\delta\}$ ($\tilde{k} = 1,\ldots,\dim \mathrm{K}\{\delta^{\mathrm{T}_A}\}$) then

$$\mathrm{Tr}\left(P_{\mathrm{K}\{\delta\}}\delta\right) = \mathrm{Tr}\left(\sum_{kk'}|k\rangle\langle k'|\delta\right) = \sum_{kk'}\langle k'|\delta|k\rangle = 0 \tag{4.36}$$

$$\begin{aligned}\mathrm{Tr}\left(P^{\mathrm{T}_A}_{\mathrm{K}\{\delta^{\mathrm{T}_A}\}}\delta\right) &= \mathrm{Tr}\left(P_{\mathrm{K}\{\delta^{\mathrm{T}_A}\}}\delta^{\mathrm{T}_A}\right) = \mathrm{Tr}\left(\sum_{\tilde{k}\tilde{k}'}|\tilde{k}\rangle\langle\tilde{k}'|\delta^{\mathrm{T}_A}\right) \\ &= \sum_{\tilde{k}\tilde{k}'}\langle\tilde{k}'|\delta^{\mathrm{T}_A}|\tilde{k}\rangle = 0.\end{aligned} \tag{4.37}$$

4.2 Positive Maps

4.2.1 Introduction

So far we only considered states in HILBERT spaces and operators acting on these states. Now we go one step further and look at the so-called maps which can be seen as *superoperators* manipulating the operators in HILBERT space. Throughout this section we will denote the various HILBERT spaces by $\mathscr{H}_B$, $\mathscr{H}_C$ and so on and the set of linear operators acting on $\mathscr{H}_B$ as $\mathscr{B}(\mathscr{H}_B)$. We start by defining linear maps:

Def. 4.6 *A linear, self-adjoint map ε is a transformation*

$$\varepsilon: \quad \mathscr{B}(\mathscr{H}_B) \to \mathscr{B}(\mathscr{H}_C) \tag{4.38}$$

which

- *is linear*

$$\varepsilon(\alpha O_1 + \beta O_2) = \alpha\varepsilon(O_1) + \beta\varepsilon(O_2) \quad \forall O_1, O_2 \in \mathscr{B}(\mathscr{H}_B) \quad \alpha, \beta \in \mathbb{C} \tag{4.39}$$

- *and maps hermitian operators to hermitian operators:*

$$\varepsilon(O^\dagger) = (\varepsilon(O))^\dagger \qquad \forall O \in \mathscr{B}(\mathscr{H}_B). \tag{4.40}$$

For brevity we will only write linear map instead of linear self adjoint map.
The following definitions help to further characterize linear maps.

Def. 4.7 *A linear map ε is called trace preserving if*

$$\mathrm{Tr}(\varepsilon(O)) = \mathrm{Tr}(O) \quad \forall O \in \mathscr{B}(\mathscr{H}_B). \tag{4.41}$$

Def. 4.8 *Positive map*
A linear, self adjoint map ε is called positive if

$$\forall\, \rho \in \mathscr{B}(\mathscr{H}_B) \text{ with } \rho \geq 0 \quad \Rightarrow \quad \varepsilon(\rho) \geq 0. \tag{4.42}$$

This means that positive maps have the property of mapping positive operators onto positive operators. It will turn out to be important to consider maps on the tensor product of a positive operator acting on one subsystem A and the identity acting an another subsystem B. In this case we define

Def. 4.9 *Completely positive map*
A positive linear map ε is completely positive if for any tensor extension of the form

$$\begin{aligned} \varepsilon': \quad \mathscr{B}(\mathscr{H}_A \otimes \mathscr{H}_B) &\to \mathscr{B}(\mathscr{H}_A \otimes \mathscr{H}_C) \\ \varepsilon' &= \mathbb{I}_A \otimes \varepsilon \end{aligned} \tag{4.43}$$

ε' is positive.

4.2.2 Examples

Hamiltonian evolution of a quantum system Let $O \in \mathscr{B}(\mathscr{H}_B)$ and U an unitary matrix and define ε by

$$\begin{array}{rccc} \varepsilon: & \mathscr{B}(\mathscr{H}_A) & \to & \mathscr{B}(\mathscr{H}_A) \\ & \varepsilon(O) & = & UOU^\dagger \end{array} \tag{4.44}$$

As an example for this map consider the time-evolution of a density matrix. It can be written as $\rho(t) = U(t)\rho(0)U^\dagger(t)$, i.e. in the form given above.
Clearly this map is linear, self-adjoint, positive and trace-preserving. It is also completely positive because for $0 \leq w \in \mathscr{B}(\mathscr{H}_A \otimes \mathscr{H}_B)$

$$(\mathbb{I}_A \otimes \varepsilon)w = (\mathbb{1}_A \otimes U)w(\mathbb{1}_A \otimes U^\dagger) = \tilde{U}w\tilde{U}^\dagger \tag{4.45}$$

where $\tilde{U}$ is unitary. But then $\langle\psi|\tilde{U}w\tilde{U}^\dagger|\psi\rangle \geq 0$ iff $\langle\psi|w|\psi\rangle \geq 0$ (since positivity is not changed by unitary evolution).

Hamiltonian evolution of a system and its environment Let $\rho \in \mathscr{B}(\mathscr{H}_B)$ (the *system*) and $\sigma \in \mathscr{B}(\mathscr{H}_A)$ (the *environment*) be positive operators and define

$$\begin{array}{rccc} \varepsilon: & \mathscr{B}(\mathscr{H}_B) & \to & \mathscr{B}(\mathscr{H}_B) \\ & \varepsilon(\rho) & = & \mathrm{Tr}_A\left(U\sigma \otimes \rho U^\dagger\right) \end{array} \tag{4.46}$$

where $U \in \mathscr{B}(\mathscr{H}_A \otimes \mathscr{H}_B)$ is unitary. This map describes the time-evolution of a system together with the environment. It is obviously linear, self-adjoint and it is also trace preserving because

$$\begin{aligned} \mathrm{Tr}(\varepsilon(\rho)) &= \mathrm{Tr}_B\left(\mathrm{Tr}_A\left(U\sigma \otimes \rho U^\dagger\right)\right) \\ &= \mathrm{Tr}\left(U\sigma \otimes \rho U^\dagger\right) = \mathrm{Tr}\left(\sigma \otimes \rho U U^\dagger\right) = \mathrm{Tr}(\sigma \otimes \rho)\,. \end{aligned} \tag{4.47}$$

KRAUS' representation of completely positive maps Consider a set of matrices $\{A_i : \mathscr{H}_B \to \mathscr{H}_C\}$ and the map

$$\begin{array}{rccc} \varepsilon: & \mathscr{B}(\mathscr{H}_B) & \to & \mathscr{B}(\mathscr{H}_B) \\ & \varepsilon(\rho) & = & \sum_{i=1}^K A_i \rho A_i^\dagger \end{array} \tag{4.48}$$

This map is obviously linear and self-adjoint. It is trace preserving if and only if

$$\sum_{i=1}^K A_i^\dagger A_i = \mathbb{1}_C. \tag{4.49}$$

It is positive

$$\langle\psi|\varepsilon(\rho)|\psi\rangle = \sum_i \langle\psi|A_i\rho A_i^\dagger|\psi\rangle = \sum_i \langle A_i^\dagger\psi|\rho|A_i^\dagger\psi\rangle \geq 0, \tag{4.50}$$

completely positive because

$$(\mathbb{I}_A \otimes \varepsilon)w = \sum_i (\mathbb{1}_A \otimes A_i)w(\mathbb{1}_A \otimes A_i^\dagger) \tag{4.51}$$

and

$$\langle\psi|(\mathbb{I}_A \otimes \varepsilon)w|\psi\rangle = \sum_i \langle(\mathbb{1}_A \otimes A_i^\dagger)\psi|w|(\mathbb{1}_A \otimes A_i^\dagger)\psi\rangle \geq 0. \tag{4.52}$$

Transposition An example for a positive but not completely positive map is the transposition T defined as:

$$\begin{array}{rccc} \mathrm{T}: & \mathscr{B}(\mathscr{H}_B) & \rightarrow & \mathscr{B}(\mathscr{H}_B) \\ & \mathrm{T}(\rho) & = & \rho^{\mathrm{T}} \end{array} \tag{4.53}$$

Of course this map is positive but it is not completely positive because

$$(\mathbb{I}_A \otimes \mathrm{T})w = w^{\mathrm{T}_B} \tag{4.54}$$

and we know that there are states with $\rho \geq 0$ but $\rho^{\mathrm{T}_B} \leq 0$.

4.2.3 Decomposable Maps

Def. 4.10 *A positive map is called decomposable if and only if it can be written as*

$$\varepsilon = \varepsilon_1 + \varepsilon_2\,\mathrm{T} \tag{4.55}$$

where ε_1, ε_2 are completely positive maps and T *is the operation of transposition introduced in section 4.2.2.*

Theorem 4.3 HORODECKI
A state $\rho \in \mathscr{B}(\mathscr{H}_A \otimes \mathscr{H}_B)$ is separable iff for all positive maps

$$\varepsilon:\ \mathscr{B}(\mathscr{H}_B) \rightarrow \mathscr{B}(\mathscr{H}_C) \tag{4.56}$$

we have

$$(\mathbb{I}_A \otimes \varepsilon)\rho \geq 0. \tag{4.57}$$

Proof:
[$\Rightarrow$] ρ is separable so we can write it as

$$\rho = \sum_{k=1}^{P} p_k |e_k f_k\rangle\langle e_k f_k| = \sum_{k=1}^{P} p_k |e_k\rangle\langle e_k| \otimes |f_k\rangle\langle f_k| \tag{4.58}$$

for some $P > 0$. On this state $(\mathbb{I}_A \otimes \varepsilon)$ acts as

$$(\mathbb{I}_A \otimes \varepsilon)\rho = \sum_{k=1}^{P} p_k |e_k\rangle\langle e_k| \otimes \varepsilon(|f_k\rangle\langle f_k|) \geq 0 \tag{4.59}$$

where the last $\geq$ follows because $|f_i\rangle\langle f_i| \geq 0$ and ε is positive.
[$\Leftarrow$] This direction is not as easy as the only if direction. We will prove it in section 4.2.4.
Note that theorem 4.3 can also be cast into the following form:

Theorem 4.4 HORODECKI
A state $\rho \in \mathscr{B}(\mathscr{H}_A \otimes \mathscr{H}_B)$ is entangled if and only if there exists a positive map $\varepsilon : \mathscr{B}(\mathscr{H}_B) \to \mathscr{B}(\mathscr{H}_C)$ such that

$$(\mathbb{I}_A \otimes \varepsilon)\rho \not\geq 0. \tag{4.60}$$

4.2.4 Jamiołkowski Isomorphism

In order to complete the proof of theorem 4.3 we introduce first the JAMIOŁKOWSKI isomorphism [19] between operators and maps.
Given an operator $E \in \mathscr{B}(\mathscr{H}_B \otimes \mathscr{H}_C)$ and an orthonormal product basis $|k,l\rangle$ we define a map by

$$\begin{aligned} \varepsilon : \mathscr{B}(\mathscr{H}_B) &\to \mathscr{B}(\mathscr{H}_C) \\ \varepsilon(\rho) &= \sum_{k_1,l_1,k_2,l_2} {}_{BC}\langle k_1 l_1|E|k_2 l_2\rangle_{BC}\, |l_1\rangle_{CB}\langle k_1|\rho|k_2\rangle_{BC}\langle l_2| \end{aligned} \tag{4.61}$$

or in short form

$$\varepsilon(\rho) = \mathrm{Tr}_B\left(E\rho^{\mathrm{T}_B}\right). \tag{4.62}$$

This shows how to construct the map ε from a given operator E. To construct an operator from a given map we use the state

$$|\psi^+\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} |i\rangle_{B'}|i\rangle_B \tag{4.63}$$

(where $M := \dim \mathscr{H}_B$) to get

$$M\,(\mathbb{I}_{B'} \otimes \varepsilon)\left(|\psi^+\rangle\langle\psi^+|\right) = E. \tag{4.64}$$

One can see this in the following way:

$$\begin{aligned}
&(\mathbb{I}_{B'} \otimes \varepsilon)\left(|\psi^+\rangle\langle\psi^+|\right) \\
&= (\mathbb{I}_{B'} \otimes \varepsilon)\left(\frac{1}{M}\sum_{i,i'=1}^{M} {}_{B'}|i\rangle\langle i'|_{B'} \otimes {}_{B}|i\rangle\langle i'|_{B}\right) \\
&= \frac{1}{M}\sum_{i,i'=1}^{M} |i\rangle_{B'B'}\langle i'| \otimes \\
&\qquad \otimes \Big(\sum_{k_1,l_1,k_2,l_2} {}_{BC}\langle k_1 l_1|E|k_2 l_2\rangle_{BC}|k_1 l_1\rangle_{BC}\,{}_{BC}\langle k_2 l_2|\,|i\rangle_{BB}\langle i'|\Big) \\
&= \frac{1}{M}\sum_{i,i'=1}^{M} |i\rangle_{B'B'}\langle i'| \otimes \sum_{l_1,l_2} {}_{BC}\langle i l_1|E|i' l_2\rangle_{BC}\,|l_1\rangle_{C}\,{}_{C}\langle l_2| \\
&= \frac{1}{M}\Big(\sum_{i,l_1}|i l_1\rangle\langle i l_1|\Big)E\Big(\sum_{i',l_2}|i' l_2\rangle\langle i' l_2|\Big) = \frac{1}{M}E
\end{aligned} \tag{4.65}$$

Now we can construct the map from the operator and vice versa. This relationship has the following properties:

Lemma 4.5 *1. $E \geq 0$ iff ε is a completely positive map.*

2. E is an entanglement witness iff ε is a positive map.

3. E is a decomposable entanglement witness iff ε is decomposable.

4. E is a non-decomposable entanglement witness iff ε is non-decomposable and positive.

As an example we will give a proof of the "only if" direction of the second statement. Let $E \in \mathscr{B}(\mathscr{H}_B \otimes \mathscr{H}_C)$ be an entanglement witness. Then $\langle e, f|E|e, f\rangle \geq 0$. By the JAMIOŁKOWSKI isomorphism the corresponding map is defined as $\varepsilon(\rho) = \mathrm{Tr}_B\left(E\rho^{\mathrm{T}_B}\right)$ where $\rho \in \mathscr{B}(\mathscr{H}_B)$.
We have to show that

$${}_{C}\langle\phi|\varepsilon(\rho)|\phi\rangle_C = {}_{C}\langle\phi|\,\mathrm{Tr}_B\left(E\rho^{\mathrm{T}_B}\right)|\phi\rangle_C \geq 0 \qquad \forall|\phi\rangle_C \in \mathscr{H}_C. \tag{4.66}$$

Since ρ acts in Bobs space we get (using the spectral decomposition of ρ)

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| \quad \rightsquigarrow \quad \rho^{\mathrm{T}_B} = \sum_i \lambda_i |\psi_i^*\rangle\langle\psi_i^*| \tag{4.67}$$

where all $\lambda_i \geq 0$. Then

$$\begin{aligned} {}_C\langle\phi|\varepsilon(\rho)|\phi\rangle_C &= {}_C\langle\phi|\sum_i \mathrm{Tr}_B\left(E\lambda_i|\psi_i^*\rangle_{B\,B}\langle\psi_i^*|\right)|\phi\rangle_C \\ &= \sum_i \lambda_{i\,BC}\langle\psi_i^*,\phi|E|\psi_i^*,\phi\rangle_{BC} \geq 0. \end{aligned} \tag{4.68}$$

We are now able to proof the $\Leftarrow$ direction of theorem 4.3 or, equivalently, the $\Rightarrow$ direction of theorem 4.4. We thus have to show that taking ρ_{AB} to be entangled there exists a positive map $\varepsilon : \mathscr{B}(\mathscr{H}_A) \to \mathscr{B}(\mathscr{H}_C)$ such that $(\varepsilon \otimes \mathbb{I}_B)\rho$ is not positive definite.
If ρ is entangled then there exists an entanglement witness W_{AB} such that

$$\mathrm{Tr}\,(W_{AB}\rho_{AB}) < 0 \tag{4.69}$$

$$\mathrm{Tr}\,(W_{AB}\sigma_{AB}) \geq 0 \tag{4.70}$$

for all separable σ_{AB}. W_{AB} is an entanglement witness (which detects ρ_{AB}) iff W_{AB}^{T} (note the complete transposition!) is also an entanglement witness (which detects ρ_{AB}^{T})[13]. We define a map by

$$\begin{array}{rccl} \varepsilon: & \mathscr{B}(\mathscr{H}_A) & \to & \mathscr{B}(\mathscr{H}_C) \\ & \varepsilon(\rho) & = & \mathrm{Tr}_A\left(W_{AC}^{\mathrm{T}}\rho_{AB}^{\mathrm{T}_A}\right) \end{array} \tag{4.73}$$

where $\dim\mathscr{H}_C = \dim\mathscr{H}_B \equiv M$. Then

$$(\varepsilon\otimes\mathbb{I}_B)(\rho_{AB}) = \mathrm{Tr}_A\left(W_{AC}^{\mathrm{T}}\rho_{AB}^{\mathrm{T}_A}\right) = \mathrm{Tr}_A\left(W_{AC}^{\mathrm{T}_C}\rho_{AB}\right) = \tilde{\rho}_{CB} \tag{4.74}$$

where we used that Lemma 4.1 and $\mathrm{T} = \mathrm{T}_A \circ \mathrm{T}_C$.
To complete the proof we will show that $\tilde{\rho}_{CB} \not\geq 0$. With the maximally entangled

[13]This holds because

$$\langle ef|W_{AB}^{\mathrm{T}}|ef\rangle = \langle e^* f^*|W_{AB}|e^* f^*\rangle \geq 0 \tag{4.71}$$

(so W_{AB}^{T} is positive on product states when W_{AB} is) and

$$\mathrm{Tr}\left(W_{AB}^{\mathrm{T}}\rho_{AB}^{\mathrm{T}}\right) = \mathrm{Tr}\,(W_{AB}\rho_{AB}) < 0 \tag{4.72}$$

(it detects ρ_{AB}^{T}).

state $|\psi^+\rangle_{CB} = \frac{1}{\sqrt{M}}\sum_i |ii\rangle_{CB}$ where $\{|i\rangle\}$ denotes a real, orthongonal basis we find

$$\begin{aligned}
{}_{CB}\langle\psi^+|\,\mathrm{Tr}_A\left(W_{AC}^{\mathrm{T}_C}\rho_{AB}\right)|\psi^+\rangle_{CB} &= \frac{1}{M}\sum_i {}_{CB}\langle ii|\,\mathrm{Tr}_A\left(W_{AC}^{\mathrm{T}_C}\rho_{AB}\right)\sum_j |jj\rangle_{CB} \\
&= \frac{1}{M}\sum_{ij}\mathrm{Tr}_A\left({}_C\langle i|W_{AC}^{\mathrm{T}_C}|j\rangle_C\,{}_B\langle i|\rho_{AB}|j\rangle_B\right) \\
&= \frac{1}{M}\sum_{ij}\mathrm{Tr}_A\left({}_C\langle j|W_{AC}|i\rangle_C\,{}_B\langle i|\rho_{AB}|j\rangle_B\right) \\
&= \frac{1}{M}\sum_{ij}\mathrm{Tr}_A\left(\mathrm{Tr}_C\left(W_{AC}|i\rangle_{CC}\langle j|\right)\mathrm{Tr}_B\left(\rho_{AB}|j\rangle_{BB}\langle i|\right)\right) \\
&= \frac{1}{M}\mathrm{Tr}_{ABC}\Big(W_{AC}\rho_{AB}\underbrace{\sum_i |i\rangle_{CB}\langle i|}_{\mathbb{1}_{CB}}\underbrace{\sum_j |j\rangle_{BC}\langle j|}_{\mathbb{1}_{BC}}\Big) \\
&= \frac{1}{M}\mathrm{Tr}_{AB}\left(W_{AB}\rho_{AB}\right) < 0. \qquad (4.75)
\end{aligned}$$

This concludes the proof that there exists a map ε with $\varepsilon(\rho) \not\geq 0$.

4.2.5 Comparison of Witnesses and Maps

In this section we developed a strong relation between entanglement witnesses and maps. Notice that an entanglement witness only gives one condition (namely $\mathrm{Tr}(W\rho) < 0$) while for a map $(\varepsilon \otimes \mathbb{I}_B)\rho$ has to be positively definite, i.e. there are many conditions that have to be fulfilled. Thus a map is much stronger.
This can also be seen from the fact that if the map detects ρ_{AB}, i.e. if

$$\mathrm{Tr}_A\left(W_{AC}\rho_{AB}^{\mathrm{T}_A}\right) = \tilde{\rho}_{CB} < 0 \qquad (4.76)$$

then it detects also

$$M_B\rho_{AB}M_B^\dagger = \rho'_{AB} \qquad (4.77)$$

where M_B is invertible ($\det(M_B) \neq 0$). This operation in general changes the trace so it corresponds to a partial measurement. Notice that M_B only acts in Bobs space and thus

$$\mathrm{Tr}_A\left(W_{AC}\left(\rho'_{AB}\right)^{\mathrm{T}_A}\right) = \tilde{\rho}'_{BC} = M_B\tilde{\rho}_{BC}M_B^\dagger. \qquad (4.78)$$

Then if there is a $|\psi\rangle \in \mathscr{H}_{CB}$ such that

$$\langle\psi|\tilde{\rho}_{BC}|\psi\rangle < 0 \qquad (4.79)$$

it follows that

$$\langle \psi' | \tilde{\rho}'_{BC} | \psi' \rangle < 0 \qquad \text{with} \quad |\psi'\rangle = \left(M_B^\dagger\right)^{-1} |\psi\rangle \tag{4.80}$$

because

$$\langle \psi | \underbrace{M_B^{-1} M_B}_{\mathbb{1}} \tilde{\rho}_{BC} \underbrace{M_B^\dagger \left(M_B^\dagger\right)^{-1}}_{\mathbb{1}} |\psi\rangle < 0, \tag{4.81}$$

i.e. the map also detects $M_B \rho_{AB} M_B^\dagger$. A map that detects one entangled state thus detects a complete family of states. This means that given a witness that detects ρ_{AB} we are able to construct a corresponding map that detects not only ρ_{AB} (and all the other states detected by the witness) but also $M_B \rho_{AB} M_B^\dagger$ which does not have to be detected by the witness since it is in general not possible to say whether

$$\mathrm{Tr}_A \left(W_{AB} M_B \rho_{AB} M_B^\dagger \right) < 0 \quad \text{or} \quad \geq 0. \tag{4.82}$$

While the witnesses are much weaker in detecting entanglement we will show in chapter 6 that this concept is able to provide a more detailed classification of entangled states.

5 Classification of Separable States, Entanglement Witnesses and Positive Maps

To classify separable states, entanglement witnesses (EW) and positive maps (PM) we want to remind the reader especially of theorem 3.3 and definition 4.4. We denote the space of separable states with S and the space of PPT states with P where $S \subseteq P$. The following classification is based on [17].

Lemma 5.1 *Let δ be an edge state and $W_\delta = P + Q^{T_B}$ with $\mathrm{R}\{P\} = \mathrm{K}\{\delta\}$ and $\mathrm{R}\{Q\} = \mathrm{K}\{\delta^{T_B}\}$ then*

$$W = W_\delta - \varepsilon \mathbb{1} \tag{5.1}$$

is an non-decomposable EW for

$$0 < \varepsilon \leq \varepsilon_0 = \inf_{|e,f\rangle} \langle e, f|W_\delta|e, f\rangle. \tag{5.2}$$

As shown in eqn. (4.33) W is a witness which detects the PPT entangled edge state δ and is thus non-decomposable (by definition 4.3).

Lemma 5.2 *The state σ is separable iff for all EW's tangent to S $\mathrm{Tr}(W\sigma) \geq 0$.*

The direction $\Rightarrow$ is fulfilled simply by definition of the witness. So we only have to show the other direction.
Proof:
Suppose $\sigma \notin S$. Then $\exists W$ with $\mathrm{Tr}(W\sigma) < 0$. Now we can calculate

$$\varepsilon_0 = \inf_{|e,f\rangle} \langle e, f|W|e, f\rangle \geq 0. \tag{5.3}$$

If $\varepsilon_0 = 0$ then W is tangent to S. But we required $\mathrm{Tr}(W\sigma) \geq 0$ for *any* tangent W which contradicts the assumption $\mathrm{Tr}(W\sigma) < 0$.
If $\varepsilon_0 \neq 0$ then $\tilde{W} = W - \varepsilon_0 \mathbb{1}$ is tangent to S. But we required $\mathrm{Tr}\left(\tilde{W}\sigma\right) \geq 0$ for *any* tangent $\tilde{W}$ which contradicts the assumption $\mathrm{Tr}\left(\tilde{W}\sigma\right) < \mathrm{Tr}(W\sigma) < 0$.
This leads to the following

Lemma 5.3 *If a decomposable witness W is tangent to P at ρ then for any decomposition as in lemma 3.2 W must also be tangent to P at δ and simultaneously to S at ρ_S.*

Proof:

$$\begin{aligned} \mathrm{Tr}(W\rho) = 0 &= \mathrm{Tr}\left(W\left(\Lambda\rho_S + (1-\Lambda)\,\delta\right)\right) \\ &= \Lambda\,\mathrm{Tr}(W\rho_S) + (1-\Lambda)\,\mathrm{Tr}(W\delta) \geq 0 \end{aligned} \tag{5.4}$$

The first addend is not negative because ρ_S is separable and the second addend is not negative because W is a decomposable witness and δ is a PPT state (c.f. eqn. (4.10)). Thus $\mathrm{Tr}(W\rho_S) = \mathrm{Tr}(W\delta) = 0$. Note that the figures 1, 2 and 4 are therefore misleading.

Prop. 5.1 *If an EW W which does not detect any PPTES is tangent to P at some edge state δ then it has the form:*

$$W = P + Q^{T_B} \tag{5.5}$$

with $\mathrm{R}\{P\} \subseteq \mathrm{K}\{\delta\}$ *and* $\mathrm{R}\{Q\} \subseteq \mathrm{K}\{\delta^{T_B}\}$.

Proof:
If W does not detect PPTES then it has to be decomposable, i.e.

$$W = P + Q^{T_B}. \tag{5.6}$$

Since $\mathrm{Tr}(W\delta) = 0$ and $P, Q \geq 0$ we must have $\mathrm{Tr}(P\delta) = 0$ and

$$\mathrm{Tr}\left(Q^{T_B}\delta\right) = \mathrm{Tr}\left(Q\delta^{T_B}\right) = 0 \tag{5.7}$$

which means $R\{P\}$ is orthogonal to the range of δ (i.e. it is in the kernel) and $R\{Q\}$ is orthogonal to the range of δ^{T_B}.

Prop. 5.2 *Any nd-EW W has the form*

$$W = P + Q^{T_B} - \varepsilon\mathbb{1} \qquad \textit{with} \tag{5.8}$$

$$0 < \varepsilon \leq \inf_{|e,f\rangle} \langle e, f|P + Q^{T_B}|e, f\rangle \tag{5.9}$$

and there exists an edge state δ for which P, Q fulfill

$$\mathrm{R}\{P\} \subseteq \mathrm{K}\{\delta\} \qquad\qquad \mathrm{R}\{Q\} \subseteq \mathrm{K}\{\delta^{T_B}\}. \tag{5.10}$$

Proof:
Consider an EW

$$W(\lambda) = W + \lambda\,\mathbb{1} \tag{5.11}$$

which is by lemma 5.1 decomposable for $\lambda > \lambda_0$ (called ε_0 there) and non-decomposable for all $\lambda < \lambda_0$. So for any $\lambda < \lambda_0$ it detects at least one PPTES ρ_λ. Since the set of PPTES is compact the series of ρ_λ converges to the PPT entangled state ρ_{λ_0}. By

construction $W(\lambda_0)$ is decomposable and thus does not detect any PPT entangled states (lemma 4.2) which means that

$$\mathrm{Tr}\left(W(\lambda_0)\rho_{\lambda_0}\right) = 0 \tag{5.12}$$

so $W(\lambda_0)$ is tangent to P at ρ_{λ_0}. Thus by lemma 5.3 there exists an edge state δ with

$$\mathrm{Tr}\left(W(\lambda_0)\delta\right) = 0. \tag{5.13}$$

By proposition 5.1 we know

$$W(\lambda_0) = P + Q^{\mathrm{T}_B} \tag{5.14}$$

and thus

$$W = P + Q^{\mathrm{T}_B} - \varepsilon\mathbb{1} \tag{5.15}$$

with $\varepsilon = \lambda_0$. Hence W is non decomposable for all $0 < \varepsilon \leq \lambda_0$ with $\mathrm{R}\{P\} \subseteq \mathrm{K}\{\delta\}$ and $\mathrm{R}\{Q\} \subseteq \mathrm{K}\{\delta^{\mathrm{T}_B}\}$. Using lemma 5.1 we know

$$\lambda_0 = \inf_{|e,f\rangle} \langle e,f|W_\delta|e,f\rangle. \tag{5.16}$$

Prop. 5.3 *As an extension to proposition 5.2 we consider a nd-EW W of the form*

$$W = P + Q^{\mathrm{T}_B} - \varepsilon\mathbb{1} \qquad \textit{with} \tag{5.17}$$

$$0 < \varepsilon \leq \inf_{|e,f\rangle} \langle e,f|P + Q^{\mathrm{T}_B}|e,f\rangle \tag{5.18}$$

and some HILBERT *spaces $\mathscr{H}_a$ and $\mathscr{H}_b$ which fulfill*

$$\mathrm{R}\{P\} \perp \mathscr{H}_a \qquad\qquad \mathrm{R}\{Q\} \perp \mathscr{H}_b. \tag{5.19}$$

1. *There exists no vector $|e,f\rangle \in \mathscr{H}_a$ such that $|e,f^*\rangle \in \mathscr{H}_b$.*
2. *If $P_{\mathscr{H}_a}$ ($P_{\mathscr{H}_b}$) is a projector onto $\mathscr{H}_a$ ($\mathscr{H}_b$) then*

$$\mathrm{R}\{\mathrm{Tr}_B\left(P_{\mathscr{H}_a}\right)\} = \mathrm{R}\{\mathrm{Tr}_B\left(P_{\mathscr{H}_b}\right)\} \tag{5.20}$$

$$\mathrm{R}\{\mathrm{Tr}_A\left(P_{\mathscr{H}_a}\right)\} = \mathrm{R}\{\mathrm{Tr}_A\left(P_{\mathscr{H}_b}\right)^*\}. \tag{5.21}$$

3. *For $x \in \{a,b\}$ we have*

$$\dim \mathscr{H}_x > \max\left[\mathrm{r}\{\mathrm{Tr}_A\left(P_{\mathscr{H}_x}\right)\}, \mathrm{r}\{\mathrm{Tr}_B\left(P_{\mathscr{H}_x}\right)\}\right]. \tag{5.22}$$

4. *Conjecture: There exists no product vector $|e,f\rangle$ with*

$$\langle e,f|P_{\mathscr{H}_x}|e,f\rangle = 0 \tag{5.23}$$

where $x \in \{a,b\}$.

5.1 Separability in $2 \times N$ Composite Quantum Systems

We will now focus on quantum systems in $\mathbb{C}^2 \otimes \mathbb{C}^N$ dimensions. An example of such a system is a two level atom coupled to an harmonic oscillator. To learn about separability of these states we will again make use of the method of subtracting vectors (see Section 2.3). The results presented here can be found in [20]. In what follows we will always denote an orthogonal basis in $\mathbb{C}^2$ as $\{|0\rangle, |1\rangle\}$ and an orthogonal basis in $\mathbb{C}^N$ as $\{|1\rangle, \ldots, |N\rangle\}$.
Since we want to subtract product vectors from ρ it is important to know in which cases such product vectors can be found in the kernel or the range of ρ. Therefore we start with

Lemma 5.4 *Any subspace $\mathscr{H} \subseteq \mathbb{C}^2 \otimes \mathbb{C}^N$ with $\dim(\mathscr{H}) = M > N$ contains an infinite number of product vectors. If $M = N$ it contains at least one product vector.*

Proof: Let

$$\{|\psi_i\rangle, i = 1, \ldots, 2N - M\} \tag{5.24}$$

be a basis in the orthogonal complement of $\mathscr{H}$. We can write it, using the orthogonal basis specified above, as

$$|\psi_i\rangle = \sum_{k=1}^{N} [A_{ik}|0,k\rangle + B_{ik}|1,k\rangle] \tag{5.25}$$

with A and B being $(2N - M) \times N$ matrices. We can always write a product vector $|e, f\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^N$ as

$$|e, f\rangle = (\alpha|0\rangle_A + |1\rangle_A) \otimes \sum_{k=1}^{N} f_k |k\rangle_B, \quad \alpha \in \mathbb{C} \cup \{\infty\},\ f_k \in \mathbb{C}. \tag{5.26}$$

There exists a product vector in $\mathscr{H}$ iff there exists a solution of $\langle\psi_i|e, f\rangle = 0$, i.e. if all $|\psi_i\rangle$ are orthogonal to $|e, f\rangle$. This conditions yields

$$(\alpha A^* + B^*)\,\vec{f} = 0. \tag{5.27}$$

In the case $M > N$ the number of variables is bigger than the number of equations and thus there exists a solution for every given α, i.e. we can find an infinite number of solutions. For $M = N$ we can find nontrivial solutions only if $\det(\alpha A^* + B^*) = 0$ but since this is a polynomial in α a solution with $\alpha \in \mathbb{C}$ can always be found.
Taking $\alpha \in \mathbb{R}$, i.e. $\alpha = \alpha^*$ in the case $M > N$ we immediately get

Lemma 5.5 *Any subspace $\mathscr{H} \subseteq \mathbb{C}^2 \otimes \mathbb{C}^N$ with $\dim(\mathscr{H}) = M > N$ contains an infinite number of product vectors of the form*

$$|e_r, f\rangle \qquad \textit{where} \qquad |e_r\rangle = |e_r^*\rangle. \tag{5.28}$$

In the following we will work with two subspaces $\mathscr{H}_1,\ \mathscr{H}_2 \in \mathbb{C}^2 \otimes \mathbb{C}^N$. Especially we will choose $\mathscr{H}_1 = \mathrm{R}\{\rho\}$ and $\mathscr{H}_2 = \mathrm{R}\{\rho^{\mathrm{T}_A}\}$. Furthermore let $M_1 = \dim\ \mathscr{H}_1$, $M_2 = \dim\ \mathscr{H}_2$. We define the orthogonal subspaces

$$K_{1,2} = \left\{ |\psi_{i_{1,2}}^{1,2}\rangle,\ i_{1,2} = 1, \ldots, 2N - M_{1,2} \right\} \tag{5.29}$$

where

$$|\psi_i^{1,2}\rangle = \sum_{k=1}^{N} \left[A_{ik}^{1,2} |0,k\rangle + B_{ik}^{1,2} |1,k\rangle \right] \tag{5.30}$$

with $(2N - M_{1,2}) \times N$-matrices A and B.

Lemma 5.6 *1. If $M_1 + M_2 > 3N$ then there exists an infinite number of product states $|e, f\rangle \in \mathscr{H}_1$ such that $|e^*, f\rangle \in \mathscr{H}_2$.*

2. If $M_1 + M_2 \leq 3N$ then there exists a product state $|e, f\rangle \in \mathscr{H}_1$ such that $|e^*, f\rangle \in \mathscr{H}_2$ if we can find an α such that there are at most $N - 1$ linearly independent vectors among the following vectors:

$$\left\{ \alpha \langle \psi_i^1 |0\rangle + \langle \psi_i^1 |1\rangle, \alpha^* \langle \psi_i^2 |0\rangle + \langle \psi_i^2 |1\rangle \right\} \tag{5.31}$$

Proof: Because the subspaces orthogonal to $\mathscr{H}_1$ and $\mathscr{H}_2$ are spanned by $|\psi_i^1\rangle$ and $|\psi_i^2\rangle$, respectively, $|e, f\rangle$ has to fulfill

$$\langle \psi_i^1 | e, f\rangle = 0 \qquad \text{and} \qquad \langle \psi_i^2 | e^*, f\rangle = 0. \tag{5.32}$$

Writing $|e, f\rangle$ as in equation (5.26) we have

$$\left[\alpha (A^1)^* + (B^1)^* \right] \vec{f} = 0 \tag{5.33}$$

$$\left[\alpha^* (A^2)^* + (B^2)^* \right] \vec{f} = 0, \tag{5.34}$$

which can be read as $4N - M_1 - M_2$ equations for $\vec{f}$. In the case $M_1 + M_2 > 3N$ there are more parameters then equations and there exists a solutions for each α, i.e. for each $|e\rangle$.
For $M_1 + M_2 \leq 3N$ consider the $(4N - M_1 - M_2) \times N$ dimensional matrix $M(\alpha, \alpha^*)$ composed of $\alpha (A^1)^* + (B^1)^*$ and $\alpha^* (A^2)^* + (B^2)^*$. There exists a solution of (5.33, 5.34) only if the rank of this matrix is smaller than N. This is the condition

imposed in the lemma. It is interesting to further investigate the conditions that have to be fulfilled to obtain a solution. In the case of $M_1+M_2=3N$ this condition is $\det[M(\alpha,\alpha^*)]=0$ for some α. The determinant is a polynomial of degree $2N-M_1$ in in α and of degree $2N-M_2$ in in α^*.
There is no way to know in advance how many roots such a polynomial has, nor if it has roots at all. E.g. $\alpha\alpha^*+1=0$ has no solutions while $\alpha-(\alpha^*)^2=0$ has infinitely many (all real numbers). If $P^*\neq P$ it is possible to reduce the equation $P(\alpha,\alpha^*)=0$ to an equation $Q(\alpha)=0$ containing only α by solving $P^*(\alpha,\alpha^*)=0$ for α^* and substituting into $P(\alpha,\alpha^*)$. In the end however it has to be checked whether the solutions of $Q(\alpha)=0$ fulfill the original equation. As an example consider $P(\alpha,\alpha^*)=(\alpha^*)^2-\alpha=0$. Then $P^*(\alpha,\alpha^*)=\alpha^2-\alpha^*=0$ and thus $\alpha^*=\alpha^2$. Substitution leads to $\alpha^4-\alpha=0$ which has the four solutions $(0,1,e^{-i2\pi/3},e^{i2\pi/3})$. These are indeed also solutions to $(\alpha^*)^2-\alpha=0$.
If $M_1+M_2<3N$ then all the $N\times N$-subdeterminants of $M(\alpha,\alpha^*)$ have to vanish (i.e. the determinant of the matrix build from the first N rows, the determinant of the matrix build from the first $N-1$ rows together with the $(N+1)^{\text{th}}$ row and so on). This implies that several polynomials in α and α^* have to have common roots.
The main theorem of this chapter makes a statement on the separability of PPT states supported[14] on $\mathbb{C}^2\otimes\mathbb{C}^N$. For this we first note

Lemma 5.7 *If ρ is PPT and supported on $\mathbb{C}^2\otimes\mathbb{C}^N$ then $\mathrm{r}\{\rho\}\geq N$.*

Proof: Let us assume $\mathrm{r}\{\rho\}<N$. Then $\dim\mathrm{K}\{\rho\}\geq N$ and from lemma 5.4 we know that there exist a product vector $|e,f\rangle\in\mathrm{K}\{\rho\}$. Now we can use lemma 2.3 to see that for some $|\hat{e}\rangle$ we can write

$$\rho=\rho_2'+\Lambda|\hat{e},f\rangle\langle\hat{e},f| \tag{5.35}$$

such that $\mathrm{r}\{\rho_2'\}=\mathrm{r}\{\rho\}-1$ and ρ_2' is still PPT. ρ_2' is supported on $\mathbb{C}^2\otimes\mathbb{C}^{N-1}$. Repeating this we can subtract more projectors on product vectors until finally ρ is written as a sum of $\mathrm{r}\{\rho\}$ such projectors. But since we assumed $\mathrm{r}\{\rho\}<N$, there surely is a vector in Bob's space orthogonal to ρ which is a contradiction to the assumption that ρ is supported on $\mathbb{C}^2\otimes\mathbb{C}^N$.
In the case $\mathrm{r}\{\rho\}=N$ it is furthermore possible to make a statement about the separability of ρ as given in the following theorem:

Theorem 5.1 *Let ρ be PPT and supported on $\mathbb{C}^2\otimes\mathbb{C}^N$. If $\mathrm{r}\{\rho\}=N$ then ρ is separable.*

[14] A state ρ acting in $\mathbb{C}^2\otimes\mathbb{C}^N$ is supported on $\mathbb{C}^2\otimes\mathbb{C}^M$ if the minimal subspace $\mathscr{H}\subseteq\mathbb{C}^N$ such that $\mathrm{R}\{\rho\}\subseteq\mathbb{C}^2\otimes\mathscr{H}$ has dimension M.

Proof: The proof is given by induction: The case $N = 1$ is clear. Now assume that the theorem holds for $N-1$. Then if $\mathrm{r}\{\rho\} = N$ then $\dim \mathrm{K}\{\rho\} = N$ and from lemma 5.4 there exists a product vector $|e, f\rangle$ in the kernel of ρ. Then, using again lemma 2.3, we can write

$$\rho = \rho_2' + \Lambda |\hat{e}, f\rangle\langle \hat{e}, f|. \tag{5.36}$$

ρ_2' has rank $N-1$ and is supported on $\mathbb{C}^2 \otimes \mathbb{C}^{N-1}$ and thus we know it is separable. There are two easy consequences of this theorem and the last lemma:

Lemma 5.8 *If ρ is separable on $\mathbb{C}^2 \otimes \mathbb{C}^N$ then it can be written as a convex sum of projectors on N product vectors.*

Lemma 5.9 *If ρ is PPT, supported on $\mathbb{C}^2 \otimes \mathbb{C}^N$ and $\mathrm{r}\{\rho\} = N$ then $\mathrm{r}\{\rho^{\mathrm{T}_A}\} = N$.*

Finally we can make a statement about separability in the special case that ρ is not only PPT but also $\rho = \rho^{\mathrm{T}_A}$:

Theorem 5.2 *If ρ is supported on $\mathbb{C}^2 \otimes \mathbb{C}^N$ and $\rho = \rho^{\mathrm{T}_A}$ then ρ is separable*[15].

Proof: The case of $N = 1$ is clear. Now supposing that the case $N-1$ is true we will proof it for N. If $\mathrm{r}\{\rho\} = N$ then ρ is separable by theorem 5.1. Otherwise as long as $\mathrm{r}\{\rho\} > N$ then by lemma 5.5 there exists $|e,g\rangle = |e^*,g\rangle \in \mathrm{R}\{\rho\}$. Thus there is $\Lambda > 0$ such that

$$\rho = \rho' + \Lambda |e,g\rangle\langle e,g|, \qquad \rho^{\mathrm{T}_A} = \left(\rho'\right)^{\mathrm{T}_A} + \Lambda |e^*,g\rangle\langle e^*,g| \tag{5.37}$$

and $\rho' = (\rho')^{\mathrm{T}_A}$ and $\mathrm{r}\{\rho'\} = \mathrm{r}\{\rho\} - 1$. This subtraction of product projectors can be repeated until $\mathrm{r}\{\rho'\} = N$.

[15] Notice that ρ^{T_B} does not work!

6 Schmidt Number Witnesses

6.1 Introduction

Let's consider the following problem:
Given a mixed state described by ρ how can the entanglement be described (especially: is the state entangled at all) ?
So far we have used witnesses W for this detection where

$$\mathrm{Tr}(W\sigma) \geq 0 \qquad\qquad \mathrm{Tr}(W\rho) < 0 \tag{6.1}$$

for all $\sigma \in S$ and for some entangled ρ. We further found that decomposable witnesses

$$W = aP + (1-a)Q^{\mathrm{T}_B} \tag{6.2}$$

cannot detect PPT entangled states.
For bipartite pure states we have

Def. 6.1 $|\psi\rangle \in \mathscr{H}_a \otimes \mathscr{H}_b$ *with* $\dim \mathscr{H}_a = M \leq \dim \mathscr{H}_b = N$ *has* SCHMIDT *rank* r *if its* SCHMIDT *decomposition reads*

$$|\psi\rangle = \sum_{i=1}^{r\leq M} \alpha_i |e_i\rangle \otimes |f_i\rangle \tag{6.3}$$

with $\alpha_i > 0$ *and* $\sum_i^r \alpha_i^2 = 1$.

The unique SCHMIDT rank[16] describes the number of entangled degrees of freedom.
The problem arises when mixed states are considered because there does not exist a unique SCHMIDT decomposition for them. Instead we define:

Def. 6.2 SCHMIDT *number* k *of the state* ρ *is defined as*

$$k = \min\{r_{max}\} \tag{6.4}$$

where r_{max} *is the maximum* SCHMIDT *rank within a decomposition and the minimum is taken over all decompositions*

For every mixed state ρ there exists an infinite number of developments, i.e.

$$\rho = \sum_i P_i |\psi_i^{r_i}\rangle\langle\psi_i^{r_i}|, \tag{6.5}$$

[16] C.f. definition 2.2 for a discussion.

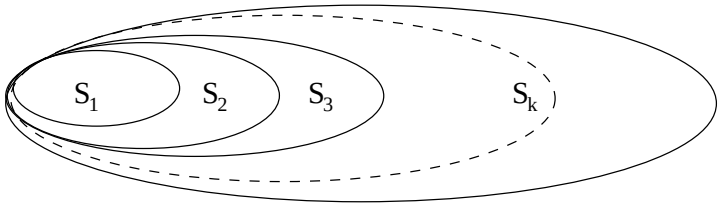

Figure 5: Schematic plot of the set of states with different SCHMIDT number embedded in the space of all states. The subscript denotes the number of entangled degrees of freedom with $S_1 \subset S_2 \subset S_3 \cdots \subset S_k \subset \cdots \subset S_M$.

where $|\psi^{r_i}\rangle$ is a pure state of SCHMIDT rank r_i, is not unique. In every possible decomposition the maximum SCHMIDT rank $r_{\max}$ of the pure states $|\psi^{r_i}\rangle$ has to be determined. The SCHMIDT number is the minimum over all $r_{\max}$ (i.e. over all possible decompositions).
This definition was introduced by TERHAL and HORODECKI.
It is thus possible to catagorize every state ρ by its SCHMIDT number. We denote the whole space of ρ by S_M (remember: $\dim \mathscr{H} = MN$) and the subspace of states with SCHMIDT number $\leq k$ as S_k.
S_k is a compact convex subset of S_M.
How is it possible to determine the SCHMIDT number of an arbitary state ρ acting on $\mathscr{H}_a \otimes \mathscr{H}_b$? The solution is based on the previous discussions regarding entanglement, i.e. we have to find some kind of SCHMIDT number witness (*SNW*). In a first step we generalize the concept of the edge states:

Def. 6.3 *δ is an k-edge state iff $\nexists |\psi^r\rangle \in \mathrm{R}\{\delta\}$ with $r < k$, i.e. there exists no state with* SCHMIDT *number smaller than k in the range of δ.*

Lemma 6.1 *Any $\rho_k \in S_k$ can be written as*

$$\rho_k = (1-p)\rho_{k-1} + p\delta \qquad 1 \geq p > 0 \tag{6.6}$$

where δ is an k-edge state[17] *and $\rho_{k-1} \in S_{k-1}$.*

Lemma 6.2 *The k-edge state δ of eqn. (6.6) has generically lower rank than ρ_k.*

Lemma 6.3 *The k-edge state δ of eqn. (6.6) contains* all *information concerning the* SCHMIDT *number k of ρ_k.*

[17] A proof of this and the following lemmas can be found in [13, 21, 22].

Def. 6.4 *A hermitian operator W is called a* SCHMIDT *number witness (SNW) of class k iff*

$$\forall \sigma \in S_{k-1} : \mathrm{Tr}(W\sigma) \geq 0 \tag{6.7}$$
$$\exists \rho \in S_k : \mathrm{Tr}(W\rho) < 0 \tag{6.8}$$

Therefore every witness which detects entanglement is also a SCHMIDT number witness of class 2.

Lemma 6.4 *Every SNW that detects ρ detects also δ.*

Proof:

$$0 > \mathrm{Tr}(W\rho_k) = (1-p)\,\mathrm{Tr}(W\rho_{k-1}) + p\,\mathrm{Tr}(W\delta) \tag{6.9}$$
$$\Leftrightarrow \qquad \mathrm{Tr}(W\delta) < \frac{p-1}{p}\underbrace{\mathrm{Tr}(W\rho_{k-1})}_{>0} < 0 \tag{6.10}$$

with $0 < p \leq 1$ and definition 6.4.
Thus the knowledge of all SNW of all k edge states fully characterizes all $\rho \in S_k$.

Lemma 6.5 *Given a k-edge state δ, a projector P on the kernel of δ and $\varepsilon = \inf_{\psi^{<k}} \langle \psi^{<k}|P|\psi^{<k}\rangle > 0$, then the operator*

$$W = P - \varepsilon \mathbb{1} \tag{6.11}$$

is a SCHMIDT *number witness for δ, i.e.*

$$\mathrm{Tr}(W\delta) = 0 - \varepsilon < 0 \tag{6.12}$$
$$\mathrm{Tr}(W\rho_{<k}) \geq 0 \tag{6.13}$$

where $\rho_{<k} = |\psi^{<k}\rangle\langle\psi^{<k}|$ is an arbitary state with SCHMIDT *number smaller than k.*

Proof: Since $\mathrm{R}\{\delta\}$ does not contain any $|\psi^{<k}\rangle$ by definition they must be all in the kernel. Furthermore $\mathrm{K}\{\delta\} = \mathrm{R}\{P\}$. So no $|\psi^{<k}\rangle$ can be in the kernel of P and thus $\varepsilon > 0$. Also we have

$$\mathrm{Tr}(W\rho_{<k}) = \mathrm{Tr}(P\rho_{<k}) - \mathrm{Tr}(\varepsilon\mathbb{1}\rho_{<k}) \geq \varepsilon - \varepsilon = 0. \tag{6.14}$$

Lemma 6.6 *Every k-*SCHMIDT *witness can be written in the canonical form*

$$W = \tilde{W} - \varepsilon \mathbb{1} \tag{6.15}$$

with $\mathrm{R}\{\tilde{W}\} = \mathrm{K}\{\delta\}$ with some k-edge state δ and $0 < \varepsilon \leq \inf_{|\psi\rangle \in S_{k-1}} \langle\psi|\tilde{W}|\psi\rangle$.

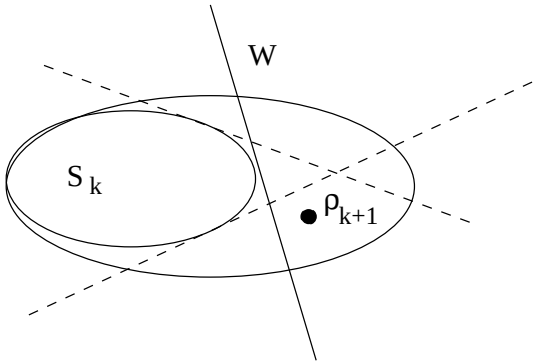

Figure 6: Schematic description of tangent SNW.

Proof:
Since W is an arbitary witness it has to have at least one negative eigenvalue. For simplicity consider W to be in its eigenbasis. Construct $\tilde{W} = W + \varepsilon \mathbb{1}$ where ε is equal to the absolut value of the largest negative eigenvalue of W. By construction the rank of W is reduced by (at least) one and thus $\mathrm{K}\{\tilde{W}\} \neq \emptyset$. Since W is a SNW we know that $\langle \psi^{<k} | \tilde{W} | \psi^{<k} \rangle \geq \varepsilon > 0$ and thus no $|\psi^{<k}\rangle$ is in the kernel of $\tilde{W}$.

Def. 6.5 *A k-*SCHMIDT *witness W is tangent to S_{k-1} at ρ if $\exists$ a state $\rho \in S_{k-1}$ such that* $\mathrm{Tr}(W\rho) = 0$.

Def. 6.6 *An SNW W is optimal if there exists no other SNW W' which detects more states than W.*

Looking at figure 6 motivates again that optimal SCHMIDT witnesses are tangent to S_k.

6.2 Example for a Schmidt Number Witness

Lemma 6.7 *The operator $W : \mathscr{H}_m \to \mathscr{H}_m$,*

$$W = \mathbb{1} - \frac{m}{k-1} P \qquad \textit{with} \tag{6.16}$$

$$P = |\psi_m^+\rangle\langle\psi_m^+| \qquad \textit{and} \qquad |\psi_m^+\rangle = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} |ii\rangle \tag{6.17}$$

is a SCHMIDT *number witness of class k.*

Proof:
The maximum SCHMIDT number is of course m. First we show that W detects a state with $m \geq \mathrm{SN} \geq k$:

$$\begin{aligned}\mathrm{Tr}\left(W|\psi_r^+\rangle\langle\psi_r^+|\right) &= 1 - \frac{1}{r}\sum_{i,l}^{r}\frac{1}{k-1}\sum_{j,k}^{m}\langle ii|jj\rangle\langle kk|ll\rangle \\ &= 1 - \frac{1}{r(k-1)}\sum_{ik}^{r}1 = 1 - \frac{r}{k-1}\end{aligned} \tag{6.18}$$

This is negative for all $r > k-1$ and positive otherwise. So W detects e.g. $|\psi_k^+\rangle$. Furthermore any state $|\psi^{<k}\rangle$ can be written as

$$\rho_{k-1} = \sum_{i=1}^{k-1} p_i\rho_i = \sum_{i=1}^{k-1} p_i \sum_{j=1} q_j|\psi_i^j\rangle\langle\psi_i^j| \tag{6.19}$$

with $\sum_i p_i = 1$, $\sum_j q_j = 1$ and $0 \leq p_i \leq 1$, $0 \leq q_j \leq 1$, i.e. as as sum of density matrices of rank smaller then k which in turn can be written as a convex sum of pure states.
We intend to find a lower bound for $\mathrm{Tr}\,(W\rho_{k-1})$, i.e. an upper bound for $\mathrm{Tr}\,(P\rho_{k-1})$. In eqn. (6.19) we replace ρ_i with the maximal entangled state $|\psi_{k-1}^+\rangle$ as an upper estimate and perform the sum. But for this state we have already shown that $\mathrm{Tr}\left(W|\psi_{k-1}^+\rangle\langle\psi_{k-1}^+|\right) \geq 0$.
This witness is furthermore optimal (not shown here).
Note also that this witness is decomposable:

$$W = P + Q^{\mathrm{T}_A} = \left(1 - \frac{1}{k-1}\right)\mathbb{1} + \frac{2P_a^{\mathrm{T}_A}}{k-1} \tag{6.20}$$

Here $P_a^{\mathrm{T}_A}$ is the partial transposed projector onto the antisymmetric subspace of $\mathbb{C}^m \otimes \mathbb{C}^m$.
As an example consider 2×2 where we can only have $k = 2$ and we have

$$\begin{aligned}W &\overset{!}{=} \left(1 - \frac{1}{1}\right)\mathbb{1} + \frac{2}{1}\left(|\phi^-\rangle\langle\phi^-|\right)^{\mathrm{T}_A} = 2|\psi^-\rangle\langle\psi^-| \\ &= 2\left(\mathbb{1}\frac{1}{2} - |\psi^+\rangle\langle\psi^+|\right) = \mathbb{1} - \frac{2}{2-1}|\psi^+\rangle\langle\psi^+| = W\end{aligned} \tag{6.21}$$

where we used the BELL states (c.f. eqn (2.8)) and their relation as discussed in section 4.1.3.

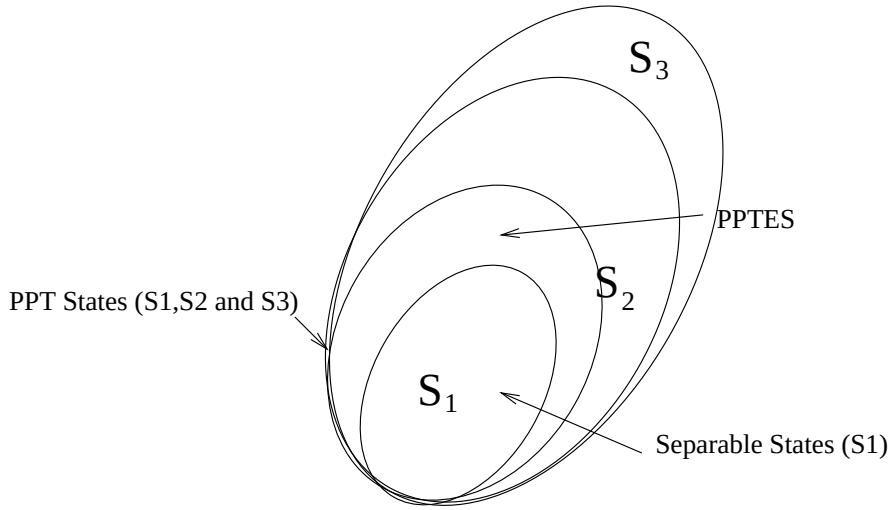

Figure 7: In 3×3 all PPTES have SCHMIDT rank 2.

6.3 The 3×3 Case

By lemma 6.7 we know already a SNW of class 2 and 3:

$$W_2 = \mathbb{1} - 3P \qquad \text{class 2} \tag{6.22}$$

$$W_3 = \mathbb{1} - \frac{3}{2}P \qquad \text{class 3} \tag{6.23}$$

This motivates the following

Conjecture 6.1 *In $\mathscr{H}_3 \otimes \mathscr{H}_3$* all SCHMIDT *number witnesses of class 3 are decomposable which is equivalent to all PPTES have* SCHMIDT *number 2.*

Now we can describe the witnesses more in detail:

Lemma 6.8 *Any SNW of class 2 has the form*

$$W = Q - \varepsilon \mathbb{1} \tag{6.24}$$

where $\mathrm{K}\{Q\}$ *does not contain any product vector, i.e.* $\mathrm{r}\{Q\} \geq 5$.

Proof:
According to lemma 6.6 W can be written this way where Q – according to lemma 6.5 – is a projector on the kernel of an 2-edge state δ.
$\mathrm{K}\{Q\} = \mathrm{R}\{\delta\}$ so by definition of the k-edge-state 6.3 $\mathrm{K}\{Q\}$ cannot contain any state with SCHMIDT rank 1, i.e. any product vector. As shown in footnote 10 on page 27 the maximum subspace created by product vectors has the dimension 5 and thus the dimension of $\mathrm{K}\{Q\}$ must not be larger then 4, i.e. $\mathrm{r}\{Q\} \geq 5$.

Lemma 6.9 *Any SNW of class 3 has the form*

$$W = Q - \varepsilon \mathbb{1} \tag{6.25}$$

where $\mathrm{r}\{Q\} \geq 8$*, i.e. W has at least 8 positive and at most one negative eigenvalue.*

Proof:
Again by lemma 6.6 W can always be written in this form. Simmilarly $\mathrm{K}\{Q\} = \mathrm{R}\{\delta\}$ which means by definition 6.3 that $\mathrm{K}\{Q\}$ cannot contain any state with SCHMIDT rank 2.
Suppose Q had a two dimensional kernel. In this case choosing $|\psi_1\rangle$ and $|\psi_2\rangle$ linearly independent and from the kernel we have $\mathrm{Tr}\left(W|\psi^2\rangle\langle\psi^2|\right) < 0$ with $|\psi^2\rangle \sim |\psi_1\rangle + |\psi_2\rangle$ – which is a contradiction because W should only detect states of SCHMIDT number 3. Thus $\mathrm{K}\{Q\} \leq 1$ or $\mathrm{R}\{Q\} \geq 8$.

Theorem 6.1 *In* $\mathscr{H}_3 \otimes \mathscr{H}_3$ *all PPTES with rank 4 have SN=2.*

Proof:
δ is a PPTES with $\mathrm{r}\{\delta\} = 4$ and thus $\dim \mathrm{K}\{\delta\} = 5$. Therefore by footnote 10 (see also proof of lemma 6.8) there is at least one product vector $|e_1, f\rangle \in \mathrm{K}\{\delta\}$. Since δ is a PPT state $\delta^{\mathrm{T}_A} > 0$ and thus $|e_1^*, f\rangle \in \mathrm{K}\{\delta^{\mathrm{T}_A}\}$.
If we denote an orthogonal basis $|e_i\rangle$ with $i = 1,2,3$ in $\mathscr{H}_A$ we have

$$\langle e_1|\delta|e_i, f\rangle = 0 \qquad i = 2,3 \qquad \text{because} \tag{6.26}$$

$$\langle e_i^*|\delta^{\mathrm{T}_A}|e_1^*, f\rangle = 0 \tag{6.27}$$

and therefore $\delta|e_2, f\rangle$ must be orthogonal to $|e_1\rangle$, i.e.

$$\delta|e_2, f\rangle = |e_2, g\rangle + |e_3, h\rangle =: |\psi^2\rangle \tag{6.28}$$

which has obviously SCHMIDT rank 2.
Applying lemma 2.1 (c.f. eqn. (2.25)) we can write

$$\delta = \delta' + \Lambda|\psi^2\rangle\langle\psi^2| \qquad \text{with} \qquad \Lambda = \frac{1}{\langle\psi^2|\frac{1}{\delta}|\psi^2\rangle} \tag{6.29}$$

and $\mathrm{r}\{\delta'\} = 3$.
Now

$$\delta'|e_1, f\rangle = \delta|e_1, f\rangle - \Lambda|\psi^2\rangle\langle\psi^2|e_1, f\rangle = 0 \tag{6.30}$$

because $|e_1, f\rangle$ is in the kernel of δ and orthogonal to $|\psi^2\rangle$ and

$$\begin{aligned}\delta'|e_2, f\rangle &= \delta|e_2, f\rangle - \Lambda|\psi^2\rangle\langle\psi^2|e_2, f\rangle \\ &= |\psi^2\rangle - \frac{1}{\langle\psi^2|\underbrace{\frac{1}{\delta}|\psi^2\rangle}_{|e_2,f\rangle}}|\psi^2\rangle\langle\psi^2|e_2, f\rangle = 0\end{aligned} \tag{6.31}$$

but

$$\begin{aligned}\delta'|e_3, f\rangle &= \left(\delta - \Lambda|\psi^2\rangle\langle\psi^2|\right)|e_3, f\rangle \\ &= |\Phi^2\rangle = |e_2, \tilde{g}\rangle + |e_3, \tilde{h}\rangle\end{aligned} \tag{6.32}$$

Again using lemma 2.1 we have

$$\delta' = \delta'' + \tilde{\Lambda}|\Phi^2\rangle\langle\Phi^2| \qquad \delta'' > 0 \tag{6.33}$$

and $\mathrm{r}\{\delta''\} = 2$, $\tilde{\Lambda} = (\langle\Phi^2|\frac{1}{\delta}|\Phi^2\rangle)^{-1}$.
It is shown the same way as before that $\delta''|e_i, f\rangle = 0$ for $i = 1, 2, 3$. Since δ'' acts in 3×2 and it is orthogonal to $|f\rangle \in \mathscr{H}_B$, δ'' has at most SN 2. Now in the sum

$$\delta = \delta'' + \tilde{\Lambda}|\Phi^2\rangle\langle\Phi^2| + \Lambda|\psi^2\rangle\langle\psi^2| \tag{6.34}$$

every term has at most SN 2 so the sum can have at most SN 2. But since we started with an entangled state in the first place δ must have SN 2.

A Generalization of the Schmidt Decomposition for the Three Qubit System

A.1 Motivation

So far three approaches regarding this problem have been made:

1. The Barcelona approach [23] and also [24].
2. The approach from SUDBERY et al. [25].
3. The Innsbruck approach [26].

While the first two approaches are very similar, the Innsbruck approach is different.
First we note that entanglement is directly linked to quantum non-locality. If two states $|\psi_1\rangle$ and $|\psi_2\rangle$ can be transformed into each other with probability one by use of only local operations and classical communication then both states have the same entanglement which is equivalent to the possibility to transform one state into the other by unitary transformations:

$$|\psi_1\rangle \sim |\psi_2\rangle \qquad \Leftrightarrow \tag{A.1}$$
$$|\psi_1\rangle = U_1 \otimes U_2 \otimes \cdots \otimes U_n |\psi_2\rangle \tag{A.2}$$

if $|\psi_i\rangle \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}$. This motivates to look at bipartite systems with

$$|\psi_1\rangle, |\psi_2\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \qquad \text{and} \quad d_1 \leq d_2. \tag{A.3}$$

If we expand both states into an orthonormal system

$$|\psi_1\rangle = \sum_{i=1}^{d_1} \alpha_i |ii\rangle \tag{A.4}$$

$$|\psi_2\rangle = \sum_{j=1}^{d_1} \alpha_j |jj\rangle \tag{A.5}$$

then $|\psi_1\rangle \sim |\psi_2\rangle \Leftrightarrow \alpha_i = \beta_i \; \forall i$. If we have e.g. 3 SCHMIDT coefficients and we remember that states have to have the norm one, we can write

$$1 = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \qquad \text{with } \alpha_i > 0 \tag{A.6}$$

and interpret this as a point in entanglement space.

Before continuing, we remember how local transformations act on a two qubit state. If

$$|\psi\rangle = \sum_{ij} t_{ij}|i\rangle|j\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$$
$$= t_{00}|00\rangle + t_{01}|01\rangle + t_{10}|10\rangle + t_{11}|11\rangle \qquad \text{with} \tag{A.7}$$
$$T = \begin{pmatrix} t_{00} & t_{01} \\ t_{10} & t_{11} \end{pmatrix} \qquad \text{and} \tag{A.8}$$
$$|\psi\rangle = (0,1)_{\mathrm{A}}\, T \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathrm{B}} \tag{A.9}$$

then transformations regarding the first index (i) are multiplications of unitary operators (U_1) from left while transformations regarding the second index (j) are multiplications from right (with U_2). Thus we can write

$$T' = U_1 T U_2 \qquad\qquad T = U_1^\dagger \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U_2 \tag{A.10}$$
$$|\psi\rangle = \lambda_1|00\rangle + \lambda_2|11\rangle \qquad \text{in the new basis.} \tag{A.11}$$

Now we want to generalize the decomposition to states

$$|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2. \tag{A.12}$$

Using the same notation as before we can write an arbitrary state as

$$|\psi\rangle = \sum_{ijk} t_{ijk}|ijk\rangle. \tag{A.13}$$

To obtain the maximal physical content of that state (in contrast to mathematical degrees of freedom) we want to obtain a basis in which the maximal number of $\tilde{t}_{ijk} \equiv 0$, i.e. we want to remove all the superfluous information due to a bad choice of the local bases. This is equivalent to diagonalizing a tensor with three indices. The key question is how many coefficients can be always transformed to zero.

A.2 The Barcelona Approach

Since it is difficult to explicitly write down matrices with three indices we split the matrix T into two matrices:

$$T_0 = \begin{pmatrix} t_{000} & t_{001} \\ t_{010} & t_{011} \end{pmatrix} \qquad\qquad T_1 = \begin{pmatrix} t_{100} & t_{101} \\ t_{110} & t_{111} \end{pmatrix} \tag{A.14}$$

Using this notation local transformations on the second (third) subsystem are again simply multiplications of the respective matrices from left (right). Transformations on the first subsystem with

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \tag{A.15}$$

$UU^\dagger = \mathbb{1}$, $\det U = 1$ and thus $|\alpha|^2 + |\beta|^2 = 1$ mix the two matrices:

$$T_0' = \alpha T_0 + \beta T_1 \tag{A.16}$$
$$T_1' = -\beta^* T_0 + \alpha^* T_1 \tag{A.17}$$

Since we still have a free parameter in the transformation we require

$$\det(T_0') = 0 = \det(\alpha T_0 + \beta T_1) \qquad \Leftrightarrow \det(T_0 + x T_1) = 0 \tag{A.18}$$

where $x = \frac{\beta}{\alpha}$ an unbound variable. The determinant is a quadratic equation for complex values and is thus always solvable. We denote the solutions with x_0 and $\bar{x_0}$.
Now we choose transformations in system two and three such that

$$U_2 T_0' U_3 = T_0'' = \begin{pmatrix} \lambda_0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{A.19}$$

This is possible since $\det(U_2 T_0' U_3) = \det(U_2 U_3) \cdot \det(T_0') = 0$ and thus at least one eigenvalue vanishes. With this choice of transformation the second matrix now reads

$$U_2 T_1 U_3 = \begin{pmatrix} \lambda_1 e^{i\varphi} & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix} \tag{A.20}$$

with $\lambda_i \in \mathbb{R}^+$, $0 \leq \lambda_i \leq 1$ and $\sum_i \lambda_i^2 = 1$. All phases except φ are absorbed by redefining the local bases by a phase factor, which is always possible.
Thru this smart choice of local transformations $|\psi\rangle$ now reads

$$|\psi\rangle = \lambda_0|000\rangle + \lambda_1 e^{i\varphi}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle, \tag{A.21}$$

i.e. we have now 6 real parameters. In general we cannot have less than six. This can be shown as follows:
The entire space (regarding its product nature) is $\mathbb{C}^2 \times \mathbb{C}^2 \times \mathbb{C}^2$. It has the complex dimension $2 \cdot 2 \cdot 2 = 8$ or accordingly 16 real parameters. There are several possible counting mechanisms for the minimum number of parameters:

1. In every $\mathbb{C}^2$ subspace we can describe the most general basis for states by

$$|0\rangle = e^{i\Theta}\begin{pmatrix} \sqrt{p} \\ \sqrt{1-p}e^{i\varphi} \end{pmatrix} \qquad |1\rangle = e^{i\Theta}\begin{pmatrix} \sqrt{1-p} \\ -\sqrt{p}e^{i\varphi} \end{pmatrix}. \tag{A.22}$$

   with $p \in \{0\ldots1\}$ and $\varphi, \Theta \in \{0, 2\pi\}$. Since the overall phase Θ corresponds to a rotation of the coordinate system in this subspace (a local rotation) it bears no physical relevance and we can consider it in our choice of t_{ijk}. Therefore we need two real parameters for every subspace and thus six parameters for a general state in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$. Thus we can choose an new basis by an appropriate rotation which transforms the remaining five complex parameters to zero.

2. We must be capable to parameterize the most general transformation on the states. Such transformation belong to

$$U(1) \times SU(2) \times SU(2) \times SU(2). \tag{A.23}$$

   Each local transformation is described by a special unitary transformation (3 parameters instead of 4 because $\det U = 1$) and we can globally add a phase (or collect all local phase to one global phase). So we have $1 + 3 \times 3 = 10$ parameters for the transformation and thus 6 parameters remain in the state independently of the basis chosen.

In eqn. (A.18) we could have have chosen the solution $\bar{x_0}$ instead of x_0; in this case we would have gotten

$$|\psi\rangle = \tilde{\lambda}_0|\tilde{0}\tilde{0}\tilde{0}\rangle + \lambda_1 e^{i\tilde{\varphi}}|\tilde{1}\tilde{0}\tilde{0}\rangle + \tilde{\lambda}_2|\tilde{1}\tilde{0}\tilde{1}\rangle + \tilde{\lambda}_3|\tilde{1}\tilde{1}\tilde{0}\rangle + \tilde{\lambda}_4|\tilde{1}\tilde{1}\tilde{1}\rangle. \tag{A.24}$$

It can be shown that if we require

$$0 < \varphi < \pi \tag{A.25}$$

(or alternatively $\pi < \tilde{\varphi} < 2\pi$) the parameters are uniquely defined. Therefore we can compare the entanglement of two states by decomposing both and comparing the 6 parameters.
It should be noted here that separable states of course have only one $\lambda_i \neq 0$. Besides this criteria, there is is no measure of entanglement, i.e. it is impossible to tell if one state is "more entangled" then another one.

A.3 The Sudbery Approach

Again we describe an arbitrary state by

$$|\psi\rangle = \sum_{ijk} t_{ijk}|ijk\rangle \tag{A.26}$$

and we want to obtain as many zero parameters as possible. To achieve this we choose a new basis which obeys

$$\max_{\alpha,\beta,\gamma} |\langle \alpha,\beta,\gamma|\psi\rangle|^2 = t_{111}^2. \tag{A.27}$$

This fixes the basis in each subsystem:

$$|1\rangle_A := |\alpha\rangle \qquad \text{fixes } |0\rangle_A \tag{A.28}$$
$$|1\rangle_B := |\beta\rangle \qquad \text{fixes } |0\rangle_B \tag{A.29}$$
$$|1\rangle_C := |\gamma\rangle \qquad \text{fixes } |0\rangle_C \tag{A.30}$$

We again obtain the same form for the wave function:

$$|\psi\rangle = \lambda_0|000\rangle + \lambda_1 e^{i\varphi}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle \tag{A.31}$$

i.e. $t_{110} = t_{011} = t_{101} = 0$. If e.g. $t_{011} \neq 0$ then $|\psi\rangle$ would contain the two terms

$$t_{011}|011\rangle + t_{111}|111\rangle = (a|0\rangle + b|1\rangle)\,|11\rangle = |\alpha'\beta\gamma\rangle. \tag{A.32}$$

It can be shown that if e.g. $t_{011} \in \mathbb{R}^+$ then

$$b > \frac{t_{111}^2 - t_{022}^2}{t_{111}^2 + t_{011}^2} < 1 \tag{A.33}$$

and therefore

$$\exists b \text{ such that } |\langle \alpha'\beta\gamma|\psi\rangle|^2 > |\langle \alpha\beta\gamma|\psi\rangle|^2 \tag{A.34}$$

which violates the maximum requirement in eqn. (A.27).
The SUDBERY-criteria cannot determine whether the decomposition is unique or not. Its main advantage lies in the fact that it can be easily extended to system with more qubits.

A.4 The Innsbruck approach

If we look at pure states $\psi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ we know that we can always write them as

$$\psi = \alpha_0|00\rangle + \alpha_1|11\rangle \tag{A.35}$$

where the local basis does not need to be orthogonal.
For generic pure states $\psi \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ we still need only two product states:

$$|\psi\rangle = \alpha|000\rangle + \beta|abc\rangle \tag{A.36}$$

Here also $\langle 0|a\rangle \neq 0$, $\langle 0|b\rangle \neq 0$ and $\langle 0|c\rangle \neq 0$ are possible.
Proof:
From eqns. (A.20) and (A.24) we know the two decompositions possible. In these cases we see that

$$ {}_A\langle 0|\psi\rangle_{ABC} = |0\rangle_B|0\rangle_C \tag{A.37}$$
$$ {}_A\langle \bar{0}|\psi\rangle_{ABC} = |\bar{0}\rangle_B|\bar{0}\rangle_C \tag{A.38}$$

are product states. This means we have to show

$$ {}_A\langle 1|\psi\rangle \sim |bc\rangle_{BC} \qquad\qquad {}_A\langle a_\perp|\psi\rangle \sim |00\rangle_{BC} \tag{A.39}$$

with $\langle a|a_\perp\rangle = 0$.
When we created the states we used the requirement

$$ \det(T_0') = 0 \qquad \Leftrightarrow \qquad px^2 + qx + r = 0 \tag{A.40}$$

in eqn. (A.18) which yields two solutions x_0 and $\bar{x_0}$. But if

$$ q^2 - 4pr = 0 \qquad \text{we have} \qquad x_0 = \bar{x_0} \tag{A.41}$$

which causes $|0\rangle = |\bar{0}\rangle$.
It can be shown that in this case

$$ |\psi\rangle = \lambda_0|000\rangle + \lambda_1 e^{i\varphi}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle \tag{A.42}$$

i.e. $\lambda_4 \equiv 0$. In this case the Innsbruck decomposition is not possible.
If $\lambda_4 \neq 0$ we rewrite eqn. (A.20) as

$$\begin{aligned} |\psi\rangle &= \lambda_0|000\rangle + \lambda_1^{(1)}|100\rangle + \lambda_1^{(2)}|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle \\ &= |d\rangle|00\rangle + |1\rangle\left(\lambda_1^{(2)}|00\rangle + \lambda_2|01\rangle + \lambda_3|10\rangle + \lambda_4|11\rangle\right) \end{aligned} \tag{A.43}$$

where $\lambda_1^{(1)} + \lambda_1^{(2)} = \lambda_1 e^{i\varphi}$ and $|d\rangle = \lambda_0|0\rangle + \lambda_1^{(1)}|1\rangle$.
Now the second term is only a two qubit system where we can use the polar decomposition (2.2) to diagonalize it:

$$ \lambda_1^{(2)}|00\rangle + \lambda_2|01\rangle + \lambda_3|10\rangle + \lambda_4|11\rangle = \sum_{ij} A_{ij}|ij\rangle = \sum_{i=1}^{2} B_i|\tilde{i},\tilde{i}\rangle \tag{A.44}$$

with the diagonal matrix $B = UAV^\dagger$.
To realize the second term in the Innsbruck decomposition of eqn. (A.35) we have to require that either $B_1 = 0$ or $B_2 = 0$, i.e.

$$ \det(B) = 0 = \det(UAV^\dagger) = \det(U)\det(A)\det(V^\dagger) = \det(A) \tag{A.45}$$

since U and V are unitary matrices. This can be rephrased as

$$0 = \lambda_1^{(2)}\lambda_4 - \lambda_2\lambda_3 \qquad \Leftrightarrow \qquad \lambda_1^{(2)} = \frac{\lambda_2\lambda_3}{\lambda_4} \tag{A.46}$$

which we can fulfill for any λ_i (as long as $\lambda_4 \neq 0$ as mentioned above) as we can always adjust with $\lambda_1^{(1)}$.

References

[1] A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Acad. Publ. 1993.

[2] D. Bouwmeester, A.K. Ekert und A. Zeilinger (Eds.). *The Physics of Quantum Information*. Springer Verlag 2000.

[3] M.A. Nielsen und I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press 2000.

[4] G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. Werner und A. Zeilinger. *Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments*. Springer Verlag 2001.

[5] For the approach used in this review, search for papers from the groups of Bruß, Cirac, Lewenstein, and Sanpera on the Los Alamos preprint server.

[6] We have extensively used results by other groups: the Horodecki-family, M.B. Plenio, C.H. Bennett, D.P. DiVicenzo, B.M. Terhal, P.W. Shor, R. Werner, G. Vidal, R. Tarrach, A. Uhlmann, R. Jozsa, B. Schumacher, and many others.

[7] H. Kreutzmann und C. Trump. *Quanteninformationstheorie* 1999. available online from http://fs-maphy.uni-hannover.de.

[8] W. Rudin. *Functional Analysis*. Mc-Graw Hill 1991.

[9] F.R. Gantmacher. *Matrix Theory, Vol. I*.

[10] R.A. Horn und C.R. Johnson. *Matrix Analysis*. Cambridge University Press 1990.

[11] M.L. Mehta. *Elements of Matrix Theory*. Hindustan Publ. Corp. 1977.

[12] S. Karnas und M. Lewenstein. *Separable approximations of density matrices of composite quantum systems*. quant-ph (2000). Preprint: quant-phy/0011066.

[13] M. Lewenstein und A. Sanpera. *Separability and Entanglement of Composite Quantum Systems*. Phys. Rev. Lett. **80** (11), 2261 (1998).

[14] M. Horodecki, P. Horodecki und R. Horodecki. *Separability of mixed states: Necessary and sufficient conditions*. Phys. Lett. A **232**, 1 (1996).

[15] M. Lewenstein, D. Bruss, J.I. Cirac, B. Kraus, M. Kus, J.Samsonowicz, A. Sanpera und R. Tarrach. *Separability and distillability in composite quantum systems - a primer*. J. Mod. Opt. **47**, 2481 (2000).

[16] C.H. Bennet, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin und B.M. Terhal. *Unextendible Product Bases and Bound Entanglement*. Phys. Rev. Lett. **82**, 5385 (1999).

[17] M. Lewenstein, B. Kraus, P. Horodecki und J.I. Cirac. *Characterization of separable states and entanglement witnesses*. Phys. Rev. A **4**, 6304 (2001). Preprint: quant-phy/0005112.

[18] H.W. Alt. *Lineare Funktionalanalysis*. Springer Verlag 1985.

[19] A. Jamiołkowski. Rep. Mat. Phys. **3**, 275 (1972).

[20] B. Kraus, J.I. Cirac, S. Karnas und M. Lewenstein. *Separability in* $2 \times N$ *composite quantum systems*. Phys. Rev. Lett. **61**, 062302 (2000). Preprint: quant-phy/9912010.

[21] A. Sanpera, D. Bruß und M. Lewenstein. *Schmidt number witnesses and bound entanglement*. Phys. Rev. A **63**, 050301 (2001). Preprint: quant-phy/9912010.

[22] A. Sanpera, R. Tarrach und G. Vidal. *Local Description of Quantum Inseparability*. Phys. Rev. A **58** (2), 826 (2000).

[23] A. Acín, A. Andrianov, L. Costa, E. Jane, J.I. Latorre und R. Tarrach. *Generalized Schmidt Decomposition and Classification of Three-Quantum-Bit States*. Phys. Rev. Lett. **85**, 1560 (2000). Preprint: quant-phy/003050.

[24] A. Acín, A. Andrianov, E. Jane und R. Tarrach. *Three-qubit pure-state canonical forms*. J. Phys. A **35** (35), 6725 (2001). Preprint: quant-phy/0009107.

[25] H.A. Carteret, A. Higuchi und A. Sudbery. *Multipartite generalisation of the Schmidt decomposition*. J. Math. Phys. **41**, 7932 (2000). Preprint: quant-phy/0006125.

[26] W. Dür, G. Vidal und J.I. Cirac. *Three qubits can be entangled in two inequivalent ways*. Phys. Rev. A **62**, 062314 (2000). Preprint: quant-phy/0005115.

Index