ENTANGLEMENT AND QUANTUM COMMUNICATION COMPLEXITY

by

YORICK HARDY

THESIS

submitted in the fulfilment
of the requirements for the degree

PHILOSOPHIAE DOCTOR

in

APPLIED MATHEMATICS

in the

FACULTY OF SCIENCE

at the

UNIVERSITY OF JOHANNESBURG

PROMOTER: PROF W.-H. STEEB

MAY 2005

UNIVERSITY
OF
JOHANNESBURG

ABSTRACT

Keywords: entanglement, complexity, entropy, measurement

In chapter 1 the basic principles of communication complexity are introduced. Two-party communication is described explicitly, and multi-party communication complexity is described in terms of the two-party communication complexity model. The relation to entropy is described for the classical communication model. Important concepts from quantum mechanics are introduced. More advanced concepts, for example the generalized measurement, are then presented in detail.

In chapter 2 the different measures of entanglement are described in detail, and concrete examples are provided. Measures for both pure states and mixed states are described in detail. Some results for the Schmidt decomposition are derived for applications in communication complexity. The Schmidt decomposition is fundamental in quantum communication and computation, and thus is presented in considerable detail. Important concepts such as positive maps and entanglement witnesses are discussed with examples.

Finally, in chapter 3, the communication complexity model for quantum communication is described. A number of examples are presented to illustrate the advantages of quantum communication in the communication complexity scenario. This includes communication by teleportation, and dense coding using entanglement. A few problems, such as the Deutsch-Jozsa problem, are worked out in detail to illustrate the advantages of quantum communication. The communication complexity of sampling establishes some relationships between communication complexity, the Schmidt rank and entropy. The last topic is coherent communication complexity, which places communication complexity completely in the domain of quantum computation. An important lower bound for the coherent communication complexity in terms of the Schmidt rank is dervived. This result is the quantum analogue to the log rank lower bound in classical communication complexity.

DECLARATION

I hereby declare that the thesis submitted for the Doctor of Philosophy in Applied Mathematics degree to the University of Johannesburg, apart from the help recognised, is my own work and has not been formerly submitted to another university or institution for higher education for a degree.

Yorick Hardy Date

ACKNOWLEDGEMENT

To my parents for always supporting my studies, to Prof. Steeb for guiding me for the past 8 years, and to Prof. Villet for his support and for inspiring me to study applied mathematics.

Contents

Symbols

Constants:

i	$\sqrt{-1}$
h	Planck's constant
$\hbar$	$h/2\pi$

Sets:

$\|A\|$	Cardinality of the set A
$A \times B$	Cartesian product of the sets A and B
$\mathbf{N}$	Set of natural numbers
$\mathbf{R}$	Set of real numbers
$\mathbf{C}$	Set of complex numbers
$\mathbf{C}^2$	$\mathbf{C} \times \mathbf{C}$
$\mathbf{C}^n$	$\overbrace{\mathbf{C} \times \mathbf{C} \times \ldots \times \mathbf{C}}^{n \text{ times}}$

Basic operations:

$\|z\|$	Absolute value of $z \in \mathbf{C}$
$\overline{z}$	Complex conjugate of $z \in \mathbf{C}$
$\Re(z)$	Real component of $z \in \mathbf{C}$
$\Im(z)$	Imaginary component of $z \in \mathbf{C}$
δ_{jk}	Kronecker delta, $\delta_{jk} = 1 \Leftrightarrow j = k$ and $\delta_{jk} = 0 \Leftrightarrow j \neq k$
$+$	Boolean inclusive OR operation
$\cdot$	Boolean AND operation
$\oplus$	Boolean exclusive OR (XOR) operation
$\overline{b}$	Boolean negation

Hilbert spaces:

$\mathcal{H}$	Hilbert space
$l_2(\mathbf{N})$	The infinite-dimensional Hilbert space over $\mathbf{C}$ of all sequences $(x_0, x_1, \ldots)$ with $\sum_{j=0}^{\infty} \lvert x_j\rvert^2 < \infty$ and $\langle (x_0, x_1, \ldots), (y_0, y_1, \ldots)\rangle := \sum_{j=0}^{\infty} \overline{x_j} y_j$
$\oplus$	Direct sum
$\otimes$	Tensor product
$\lvert\psi\rangle$	Ket vector for an element of a Hilbert space
$\langle\psi\rvert$	Bra vector for an element of a Hilbert space
$\langle\psi, \phi\rangle$	Inner product for elements ψ and ϕ of a Hilbert space
$\langle\psi\rvert\phi\rangle$	Inner product between the ket vector $\lvert\psi\rangle$ and ket vector $\lvert\phi\rangle$ in a Hilbert space
$\dim(\mathcal{H})$	dimension of the Hilbert space $\mathcal{H}$
$\{\lvert 0\rangle, \lvert 1\rangle\}$	Arbitrary orthonormal basis for the Hilbert space $\mathbf{C}^2$
$\mathrm{Sch}(\lvert\psi\rangle, \mathcal{H}_A, \mathcal{H}_B)$	Schmidt rank of $\lvert\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$

Matrices / Linear Operators:

$\oplus$	Direct sum
$\otimes$	Tensor product / Kronecker product
I_n	Identity operator on a n-dimensional Hilbert space
$\hat{H}$	Hamilton operator
A^T	Transpose of the matrix A
A^*	Hermitian conjugate /adjoint of the operator A
$[A, B]$	Commutator $AB - BA$ of the operators A and B
$[A, B]_+$	Anti-commutator $AB + BA$ of the operators A and B
$\mathrm{rank}(A)$	Rank of the matrix A
$\mathrm{tr}(A)$	Trace of the linear operator A
$\mathrm{diag}(a_1, a_2, \ldots a_n)$	$n \times n$ diagonal matrix with diagonal entries a_1, a_2, …, a_n
$c_\downarrow$, $c_\uparrow$	Fermi annihilation operator (spin down, spin up)
$c_\downarrow^\dagger$, $c_\uparrow^\dagger$	Fermi creation operator (spin down, spin up)
b	Bose annihilation operator
$b^\dagger$	Bose creation operator
U_H	Walsh-Hadamard transform $\frac{1}{\sqrt{2}}(\lvert 0\rangle + \lvert 1\rangle)\langle 0\rvert + \frac{1}{\sqrt{2}}(\lvert 0\rangle - \lvert 1\rangle)\langle 1\rvert$
U_{NOT}	NOT operation $\lvert 0\rangle\langle 1\rvert + \lvert 1\rangle\langle 0\rvert$
U_{CNOT}	Controlled NOT operation $\lvert 0\rangle\langle 0\rvert \otimes I_2 + \lvert 1\rangle\langle 1\rvert \otimes U_{NOT}$

Chapter 1

Introduction

1.1 Communication Complexity

Complexity is the measure of the mimimum amount of resources required to perform some task. The types of tasks, types of resources and units of complexity can be used to distinguish between different types of complexity measures. For example *computational complexity* [Papad] is the measure of how much time is required to perform some computational task, i.e. *time* is the resource. Usually this can be described in terms of the number of basic operations used for the computation. It is often difficult to describe exactly the complexity, but some idea of the complexity can be obtained in the form of lower and upper bounds. Of course, computational complexity depends on the memory and operations available. Expanding on the above example, the *expected quantum computational complexity* of factoring [Shor94] (where quantum mechanical data and operations are allowed [Deutsch85]) appears to be much less than the classical computational complexity of factoring. Thus it is important to classify complexity according to the resources involved.

Communication complexity [Kushilevitz] is the measure involving only communication resources. We distinguish between a few measures according to what type of communication is involved. Specifically, we are interested in comparing the complexity when quantum mechanical systems and operations are resources, with the classical measures.

For an introduction to classical communication complexity see [Kushilevitz].

For an introduction to quantum computation see [HardyQC] and [Preskill].

1.1.1 Two-party Classical Communication Complexity Measures

Let A and B be two finite sets and $R \subset A \times B$. Let $f : R \to Z$ be a function we wish to compute, where Z is a finite set denoting the possible results of the computation. In most cases we consider $Z = \{0, 1\}$. Two parties, denoted by *Alice (A)* and *Bob (B)*, are involved in the computation. The finite set A denotes the domain (data) of computation for Alice, and the finite set B denotes the domain of computation for Bob. Alice is provided with $a \in A$ and Bob is provided with $b \in B$ with $(a, b) \in R$. A *protocol* is a precise description of what Alice communicates to Bob and what Bob communicates to Alice (for given a and b) in order to compute $f(a, b)$. The communication complexity of f is defined to be the minimum amount of communication over all protocols to compute the function $f(a, b)$. The amount of communication is often measured in *bits*. A *round* in a protocol is a portion of the protocol where only Alice or Bob is communicating. The moment Alice begins communicating her round begins, when Bob begins communicating Alice's round ends and Bob's round begins.

We have not yet defined what resources are available for communication and what units we use for the measure. This will determine the different measures we will consider. First we describe the classical measures (i.e. no quantum resources).

Definition. The *bounded error communication complexity* $C_\epsilon(f)$ (with $0 \leq \epsilon < 1$) is the minimum over all protocols of the maximum amount of communication over all $(a, b) \in R$ to compute $f(a, b)$ with probability at least $1 - \epsilon$. Alice and Bob only have computational resources and classical communication.
■

Definition. The *exact communication complexity* is given by $C_0(f)$.
■

Definition. The *bounded error expected communication complexity* $C^E_\epsilon(f, \mu)$ (with $0 \leq \epsilon < 1$) is the minimum over all protocols of the average amount of communication over $(a, b) \in R$ to compute $f(a, b)$ with probability at

least $1-\epsilon$. The probability distribution μ of elements in R influences the average. Alice and Bob only have computational resources and classical communication.

■

Definition. The *expected communication complexity* is $C_0^E(f, u_R)$, where u_R is the uniform distribution over R.

■

We will assume that $Z = \{0, 1\}$, unless otherwise stated, and that the complexity measures defined above are measured in bits.

Definition. The *communication matrix* M_f of f is defined as

$$M_f := (\chi(R, a, b) f(a, b))_{ab}$$

where

$$\chi(R, a, b) := \begin{cases} 1 & (a, b) \in R \\ 0 & (a, b) \notin R \end{cases}.$$

■

It is possible to obtain bounds in terms of this matrix, for example (for non-trivial and total f) [Kushilevitz]

$$\log_2 \operatorname{rank}(M_f) \leq C_0(f) \leq \operatorname{rank}(M_f) + 1.$$

In the following we define

$$T(f) := \{(a, b) \in R : \ f(a, b) = 1\}$$

$$T(f, A) := \{a \in A : \ \forall\, b \ (a, b) \in R \quad f(a, b) = 1\}$$

$$p_T(f, a) := \frac{|\{b \in B : \ (a, b) \in T(f)\}|}{|T(f)|}.$$

We denote by $|A|$ the cardinality of the finite set A. We have the following result from information theory [Ash, HardyQC]. Given n input symbols with probabilities $p_1, \ldots, p_n$, there exists a base D (i.e. for $D = 2$ we use bits) instantaneous code (prefix code, i.e. uniquely decodable) for the symbols with an average length bounded above by

$$H_D(\{p_1, \ldots, p_n\}) + 1$$

where

$$H_D(\{p_1, \ldots, p_n\}) := -\sum_{j=1}^{n} p_j \log_D p_j$$

is the *Shannon entropy*. Thus we obtain the following bound for the expected communication complexity

$$C_0^E(f, u_R) \leq \frac{|T(f,A)|}{|A|} + \left(1 - \frac{|T(f,A)|}{|A|}\right)(1 + H_2(\,\{p_T(f,a),\, a \in A\}\,) + 1)$$

$$= 1 + \left(1 - \frac{|T(f,A)|}{|A|}\right)(1 + H_2(\,\{p_T(f,a),\, a \in A\}\,)).$$

In other words, we use 1 bit to communicate whether a is in $T(f, A)$ and the remaining bits to specify a according to the probability that $f(a, b) = 1$ and 1 bit for the answer. When $|T(f, A)| = 0$, we only communicate a with the above scheme and the final answer. For a non-uniform distribution we use the definition

$$p_T(f,a) := \frac{\sum_{b \in B} \mu(a,b)\chi(T(f),a,b)}{\sum_{(a,b)\in R} \mu(a,b)\chi(T(f),a,b)}$$

where μ is the distribution over $A \times B$. We note that $\text{rank}(M_f) \leq |T(f, A)|$. In general these bounds are not symmetric with respect to the choice of A and B and the choice of $|T(f, A)|$ or $|T(1 - f, A)|$. Taking this into account we obtain the upper bound

$$C_0^E(f, u_R) \leq \min_{\substack{g \in \{f, 1-f\} \\ \mathcal{A} \in \{A,B\}}} \left\{1 + \left(1 - \frac{|T(g,\mathcal{A})|}{|\mathcal{A}|}\right)(1 + H_2(\,\{p_T(g,a),\, a \in \mathcal{A}\}\,))\right\}.$$

Example. Consider the constant function $f_1(a, b) = 1$. Thus $T(f_1) = R$, $T(f_1, A) = A$ and $p_T(f_1, a) = 1/|A|$. This would be the best case. The upper bound is 1. We define $f_0(a, b) := 1 - f_1(a, b)$. Thus $f_0(a, b) = 0$. In this case $|T(1 - f_1)| = 0$ and $p_T(1 - f_1, a) = 0$ We have $\text{rank}(M_{f_0}) = 0$ and $\text{rank}(M_{f_1}) = 1$. All complexity measures are zero for the functions f_0 and

f_1.
■

Example. Consider the *inner product* function $f_{IP} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ given by

$$f_{IP}(\mathbf{x},\mathbf{y}) := (x_1 \cdot y_1) \oplus (x_2 \cdot y_2) \oplus \cdots \oplus (x_n \cdot y_n)$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T, \mathbf{y} = (y_1, y_2, \ldots, y_n)^T \in \{0,1\}^n$. The symbols $\cdot$ and $\oplus$ denote the Boolean AND and exclusive OR (XOR) operations respectively. We have [Kushilevitz] $\text{rank}(M_{f_{IP}}) = 2^n - 1$. Thus $C_0(f_{IP}) \geq n$. Obviously $C_0(f_{IP}) \leq n$, thus $C_0(f_{IP}) = n$. Let $h(\mathbf{x})$ denote the number of 1's in $\mathbf{x}$. Thus $n - h(\mathbf{x})$ is the number of 0's in $\mathbf{x}$. For given $\mathbf{x}$, we need to determine the number of distinct $\mathbf{y}$ such that $f_{IP}(\mathbf{x},\mathbf{y}) = 1$. This quantity can be determined by noting that $f_{IP}(\mathbf{x},\mathbf{y}) = 1$ implies

$$\sum_{k=1}^{n}(x_k \cdot y_k) = 2j + 1$$

for some $j \in \{0, 1, \ldots, \lfloor \frac{n}{2} \rfloor\}$ where $\lfloor \frac{n}{2} \rfloor$ denotes the floor - i.e. the largest integer which is not greater than $\frac{n}{2}$. In other words an odd number of the $(x_k \cdot y_k)$ are 1, and necessarily then $x_k = 1$. Consequently the number of ways to achieve this is $\binom{h(\mathbf{x})}{2j+1}$. The remaining values for $\mathbf{y}$ satisfying $f_{IP}(\mathbf{x},\mathbf{y}) = 1$ is achieved by selecting the 1's when the corresponding x_k's are 0 in

$$\sum_{k=0}^{n-h(\mathbf{x})} \binom{n-h(\mathbf{x})}{k} = 2^{n-h(\mathbf{x})}$$

ways. We have

$$\begin{aligned} |\{\mathbf{y} \in \{0,1\}^n : \ (\mathbf{x},\mathbf{y}) \in T(f_{IP}\}| &= \sum_{j=0}^{\left\lfloor \frac{h(\mathbf{x})-1}{2} \right\rfloor} \binom{h(\mathbf{x})}{2j+1} \sum_{k=0}^{n-h(\mathbf{x})} \binom{n-h(\mathbf{x})}{k} \\ &= \begin{cases} 2^{n-1} & h(\mathbf{x}) \neq 0 \\ 0 & h(\mathbf{x}) = 0 \end{cases} \end{aligned}$$

where we used for $h(\mathbf{x}) \neq 0$

$$\sum_{k=0}^{n-h(\mathbf{x})} \binom{n-h(\mathbf{x})}{k} = 2^{n-h(\mathbf{x})} \qquad \text{and} \qquad \sum_{j=0}^{\left\lfloor \frac{h(\mathbf{x})-1}{2} \right\rfloor} \binom{h(\mathbf{x})}{2j+1} = \frac{2^{h(\mathbf{x})}}{2}.$$

This result is clearly independent of $\mathbf{x}$ when $h(\mathbf{x}) \neq 0$. Consequently $|T(f)| = (2^n - 1)2^{n-1}$ and

$$p_T(f_{IP}, \mathbf{x}) = \frac{|\{\mathbf{y} \in \{0,1\}^n : \ (\mathbf{x}, \mathbf{y}) \in T(f_{IP})\}|}{(2^n - 1)2^{n-1}} = \begin{cases} \frac{1}{2^n - 1} & h(\mathbf{x}) \neq 0 \\ 0 & h(\mathbf{x}) = 0 \end{cases}.$$

Thus we obtain the upper bound

$$C_0^E(f_{IP}, u_{\{0,1\}^{2n}}) \leq 2 + \log_2(2^n - 1) < n + 2$$

on the expected communication complexity. This is the worst case, when every $\mathbf{y}$ is equally likely to have $f_{IP}(\mathbf{x}, \mathbf{y}) = 1$. In this case it is irrelevant whether Alice or Bob communicate first. We also find that the case $f_{IP}(\mathbf{x}, \mathbf{y}) = 0$ yields the same result.
■

Example. We can completely describe the expected communication complexity of all 16 Boolean functions of two variables as follows

$f(x, y)$	$C_0^E(f, u_{\{0,1\}^2})$
0,1	0
x, y, $\overline{x}$, $\overline{y}$	1
$x + y$, $x + \overline{y}$, $\overline{x} + y$, $\overline{x} + \overline{y}$, $x \cdot y$, $x \cdot \overline{y}$, $\overline{x} \cdot y$, $\overline{x} \cdot \overline{y}$	1.5
$x \oplus y$, $\overline{x \oplus y}$	2

Here $+$ denotes Boolean OR, $\overline{x}$ denotes Boolean negation of x, $\cdot$ denotes Boolean AND and $\oplus$ denotes Boolean exclusive OR.
■

1.1.2 Multi-party Communication Complexity

We can extend the 2-party communication complexity model to an n-party communication complexity model as follows. Let $R \subset A_1 \times A_2 \times \cdots \times A_n$

and let $f : R \rightarrow Z$ be a function we wish to compute. The n parties involved in the computation are denoted by A_1, A_2, ..., A_n. Furthermore A_1 denotes the domain of comuptation for A_1, A_2 the domain of computation for A_2, etc. A_1 is provided with $a_1 \in A_1$, A_2 is provided with $a_2 \in A_2$ and so on. A *protocol* is a precise description of what each party communicates (for given a_1, a_2, ..., a_n) in order to compute $f(a_1, a_2, \ldots, a_n)$. The communication complexity of f is defined to be the minimum amount of communication over all protocols to compute $f(a_1, a_2, \ldots, a_n)$. The amount of communication is often measured in *bits*. A round in a protocol is a portion of the protocol where one party is communicating.

A multi-party communication complexity model can be reduced to a two-party communication complexity model if *broadcasting* is allowed. Defining $A := A_1$ and $B := A_2 \times A_3 \times \ldots \times A_n$. Thus we view A_2, A_3, ..., A_n as a single party working together with its own internal multi-party communication complexity model. If A communicates with the second party B, then all participating A_2, A_3, ..., A_n must receive this communication. If the communication cost for this communication is equivalent to the communication cost for the corresponding two-party interaction then A_1 can be said to *broadcast* to A_2, A_3, ..., A_n. If broadcasting is possible for all parties then the communication complexity can be analysed in terms of two-party communication complexity.

1.2 Quantum Systems

Quantum systems are described by the state of a quantum system and by the operations (Hamilton operators $\hat{H}$ that lead to unitary operators via $U = e^{it\hat{H}/\hbar}$) that can be performed on them.

The state of a quantum system can be descibed by a pure state, or more generally by a mixed state. Furthermore, a state may be entangled (as described in Chapter 2) which can also be considered as a resource.

1.2.1 Pure States

The *pure states* [SteebQM] of a quantum system, are described by normalized vectors $|\psi\rangle$ which are elements of a Hilbert space $\mathcal{H}$ that describes the system.

Specifying a pure state in quantum mechanics is the most that can be said about a physical system. The tensor product of two pure states in a product Hilbert space again represents a pure state for the combined system.

Example. Consider the Hilbert space $\mathbf{C}^2$. Then the following elements may represent the state of a physical system ($\phi \in \mathbf{R}$)

$$\begin{pmatrix}1\\0\end{pmatrix}, \quad \begin{pmatrix}0\\1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}e^{i\phi}\\1\end{pmatrix}.$$

■

The concept of a representing a state as a unit vector leads to the probability interpretation in quantum mechanics. Given a physical system described by the state $|\psi\rangle$, the probability that the system is in the state $|\phi\rangle$ is given by $|\langle\psi|\phi\rangle|^2$ where $\langle\psi|\phi\rangle$ is the inner product between $|\psi\rangle$ and $|\phi\rangle$ in the Hilbert space $\mathcal{H}$ and

$$0 \leq |\langle\psi|\phi\rangle|^2 \leq 1.$$

Example. Consider again the Hilbert space $\mathbf{C}^2$, and

$$|\psi\rangle := \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad |\phi\rangle := \begin{pmatrix}1\\0\end{pmatrix}.$$

Then the probability that a physical system in the state $|\psi\rangle$ is found in the state $|\phi\rangle$ is given by

$$|\langle\psi|\phi\rangle|^2 = \frac{1}{2}$$

where

$$\langle\psi| = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}.$$

■

Example. The *Greenberger-Horne-Zeilinger state* (*GHZ state*) is the state

$$|GHZ\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}$$

in the Hilbert space $\mathbf{C}^8$ and the *W state* is

$$|W\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}+\frac{1}{\sqrt{3}}\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}$$
$$+\frac{1}{\sqrt{3}}\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}.$$

Then

$$|\langle GHZ|W\rangle|^2 = 0.$$

■

The *phase* of a vector $|\psi\rangle$ has no physical significance, however the relative phase does. In other words, for $\alpha \in \mathbf{C}$ with $|\alpha| = 1$,

$$|\langle\alpha\psi|\phi\rangle|^2 = |\langle\psi|\phi\rangle|^2.$$

We use the *Dirac braket notation* where $|\alpha\psi\rangle = \alpha|\psi\rangle$ is a ket vector and $\langle\alpha\psi| = \overline{\alpha}\langle\psi|$ is the corresponding bra vector. Using $|\psi\rangle = |\psi_1\rangle + \alpha|\psi_2\rangle$ yields

$$|\langle\psi|\phi\rangle|^2 = |\langle\psi_1|\phi\rangle|^2 + |\langle\psi_2|\phi\rangle|^2 + \Re(\alpha\langle\psi_1|\phi\rangle\langle\phi|\psi_2\rangle).$$

1.2.2 Mixed States

The statistical mixtures in quantum mechanics lead to statistical quantum mechanics [SteebQM]. The statistical mixture is described by a positive trace class operator ρ on a Hilbert space $\mathcal{H}$, i.e. $\langle\psi|\rho|\psi\rangle \geq 0$ for all $|\psi\rangle$ in the Hilbert space and the trace of ρ is defined (ρ has a square matrix representation).

Definition. Let ρ denote a mixed state on the Hilbert space $\mathcal{H}$. If $\mathrm{tr}\rho = 1$, then ρ is a *density operator*. A density operator is described by a convex linear combination of projection operators $|\psi_j\rangle\langle\psi_j|$ (density operators describing (normalized) pure states)

$$\rho = \sum_{j=1}^{\dim(\mathcal{H})} p_j|\psi_j\rangle\langle\psi_j|$$

where

$$\sum_{j=1}^{\dim(\mathcal{H})} p_j = 1, \qquad \forall j \;\; p_j \geq 0$$

and $|\psi_j\rangle$ are mutually orthogonal normalized pure states, i.e. $\langle\psi_j|\psi_k\rangle = \delta_{jk}$. This representation of a density operator in terms of pure states can be obtained from the spectral decomposition of ρ, and is not necessarily unique.
■

The tensor product of two density operators on the product Hilbert space is again a density operator on the combined system.

If ρ has rank 1, then ρ describes a pure state, i.e. we can write $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle$ is a pure state.

Example. Consider the rank 1 density operator

$$\rho := \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Consequently ρ represents the pure state $\frac{1}{\sqrt{2}}(1 \quad -1)$, i.e.,

$$\rho = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\frac{1}{\sqrt{2}}(1 \quad -1)^T.$$

■

Example. $\rho = \mathrm{diag}(\frac{1}{2}, \frac{1}{2})$ is given by

$$\begin{aligned}\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1 \quad 0) + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}(0 \quad 1) \\ &= \frac{1}{2}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\left(\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}}\right) + \frac{1}{2}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\left(\frac{1}{\sqrt{2}} \quad -\frac{1}{\sqrt{2}}\right).\end{aligned}$$

Thus we have a mixed state.
■

1.2.3 Quantum Operations

Wave functions (vectors) in quantum systems evolve according to the *Schrödinger equation* [SteebQM]

$$i\hbar\frac{d\psi}{dt} = \hat{H}\psi$$

where $\hat{H}$ is a linear self-adjoint operator on the Hilbert space describing the quantum system (called the Hamilton operator). The equation has the formal solution

$$\psi(t) = U(t)\psi(0), \qquad U(t) := e^{-i\hat{H}t/\hbar}$$

where $U(t)$ is a unitary operator and $\psi(0)$ is the initial state.

A density operator $\rho(t)$ is as a convex linear combination of orthogonal projection operators

$$\begin{aligned}\rho(t) &= \sum_{j=1}^{\dim(\mathcal{H})} p_j|\psi_j(t)\rangle\langle\psi_j(t)| \\ &= \sum_{j=1}^{\dim(\mathcal{H})} p_j U(t)|\psi_j(0)\rangle\langle\psi_j(0)|U^*(t) \\ &= U(t)\left(\sum_{j=1}^{\dim(\mathcal{H})} p_j|\psi_j(0)\rangle\langle\psi_j(0)|\right)U^*(t).\end{aligned}$$

Consequently density operators ρ evolve according to

$$\rho(t) = U(t)\rho(0)U(t)^*$$

where * denotes the adjoint or hermitian conjugate. This is the *von Neumann equation* for the evolution of density operators.

Definition. An *observable* is a self-adjoint operator B on the Hilbert space, where the outcomes of the observation is given by the eigenvalues of B. Thus the observation outcomes are always described by real numbers.
■

Observables evolve according to the *Heisenberg equation* [SteebQM]

$$-i\hbar\frac{dB}{dt} = [\hat{H}, B](t)$$

with the formal solution

$$\begin{aligned} B(t) &= \sum_{n=0}^{\infty} \frac{it/\hbar}{n!} \underbrace{[\hat{H}, \ldots, [\hat{H}, [\hat{H}}_{n \text{ times}}, B]]] \\ &= U(t)^* B(0) U(t). \end{aligned}$$

1.2.4 Measurement

Orthogonal Measurement

Measurement is described with repect to an observable B. Suppose the spectral decomposition of B is given by

$$\sum_{j=1}^{\dim(\mathcal{H})} b_j |\phi_j\rangle\langle\phi_j|$$

where the b_j are the real eigenvalues of B. Let $|\psi\rangle$ denote a (normalized) pure state in the Hilbert space $\mathcal{H}$. The state $|\psi\rangle$ is found to be in the state $|\phi_j\rangle$ (with corresponding mesurement outcome b_j) with probability $|\langle\psi|\phi_j\rangle|^2$. In general the b_j may be degenerate, so that we must sum over all probabilities where $|\phi_j\rangle$ corresponding to the b_j. Summing over all measurement outcomes yields the *expectation value* (average measurement outcome)

$$\begin{aligned} b(\psi) &= \sum_{j=1}^{\dim(\mathcal{H})} b_j |\langle\psi|\phi_j\rangle|^2 \\ &= \sum_{j=1}^{\dim(\mathcal{H})} b_j \langle\psi|\phi_j\rangle\langle\phi_j\psi\rangle \\ &= \langle\psi|B|\psi\rangle \\ &= \mathrm{tr}(|\psi\rangle\langle\psi|B). \end{aligned}$$

In the case of a mixed state described by the density operator ρ we have a

statistical mixture of pure states

$$\rho = \sum_{j=1}^{\dim(\mathcal{H})} p_j |\psi_j\rangle\langle\psi_j|$$

The *expectation value* of measurement with respect to an operator B is the average expectation value for measurement of the constituent pure states with respect to B:

$$\begin{aligned} b(\rho) &= \sum_{j=1}^{\dim(\mathcal{H})} p_j b(\psi_j) \\ &= \sum_{j=1}^{\dim(\mathcal{H})} p_j \mathrm{tr}(|\psi_j\rangle\langle\psi_j|B) \\ &= \mathrm{tr}\left(\sum_{j=1}^{\dim(\mathcal{H})} p_j |\psi_j\rangle\langle\psi_j|B\right) \\ &= \mathrm{tr}(\rho B). \end{aligned}$$

To summarise, let B be an observable with k possible measurement outcomes (b_j)

$$B = \sum_{j=1}^{k} b_j \Pi_j$$

where Π_j denotes mutually orthogonal projection operators and

$$\sum_{j=1}^{k} \Pi_j = I.$$

The measurement of a system described by the density operator ρ yields the following *orthogonal measurement*:

1. The outcome b_j is obtained with probability

   $$p(b_j) = \mathrm{tr}(\Pi_j \rho) = \mathrm{tr}(\Pi_j \rho \Pi_j)$$

   where we used the cyclic invariance of the trace and $\Pi_j^2 = \Pi_j$.

2. The expectation value (average measurement value) is given by $\mathrm{tr}(\rho B)$.

3. The state of the system after measurement is in the measured state of the system (i.e. the system is projected onto the state corresponding to the measurement outcome)

$$\rho_{b_j} = \frac{\Pi_j \rho \Pi_j}{p(b_j)}.$$

Example. The density operator for the W state is

$$|W\rangle\langle W| = \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Consider the observable

$$B = \text{diag}(0,1,0,0,0,0,0,0) + 2\text{diag}(0,0,1,0,0,0,0,0)$$
$$+3\text{diag}(0,0,0,0,1,0,0,0).$$

with the measurement outcomes $b_1 = 0$, $b_2 = 1$, $b_3 = 2$, $b_4 = 3$ and the corresponding projection operators

$$\Pi_1 = \text{diag}(1,0,0,1,0,1,1,1), \qquad \Pi_2 = \text{diag}(0,1,0,0,0,0,0,0),$$
$$\Pi_3 = \text{diag}(0,0,1,0,0,0,0,0), \qquad \Pi_4 = \text{diag}(0,0,0,0,1,0,0,0).$$

The probability of the outcome $b_2 = 1$ is given by

$$p_2 := \text{tr}(\Pi_1 B \Pi_1) = \frac{1}{3}.$$

Similarly $p_1 = 0$ and $p_3 = p_4 = \frac{1}{3}$. However, the density operators

$$\rho_1 := \frac{1}{3}\text{diag}(0,1,1,0,1,0,0,0)$$

$$\rho_2 := \frac{1}{6}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

yield the same probabilities for the measurement outcomes with respect to the measurement B.

■

Generalized Measurement

Measurement can be generalized in the sense that an ancilla system (in a well defined state), identified by the Hilbert space $\mathcal{H}_A$, is introduced and allowed to interact with the quantum system identified by the Hilbert space $\mathcal{H}$. The ancilla system is subsequently measured, which may disturb the original system.

Let ρ be a density operator on $\mathcal{H}$ and B an observable on $\mathcal{H}_A$ with k possible measurement outcomes

$$B = \sum_{j=1}^{k} b_j \Pi_j$$

where Π_j denotes mutually orthogonal projection operators and

$$\sum_{j=1}^{k} \Pi_j = I.$$

Thus $I \otimes B$ is an observable on the product Hilbert space $\mathcal{H} \otimes \mathcal{H}_A$. Let $|b\rangle$ be a normalized eigenvector of B, and consequently and eigenvector of only one of the Π_j. The generalized measurement of a system described by the density operator ρ yields the following

1. The system ρ is extended with the ancilla in the normalized pure state $|b\rangle$ which gives the density operator $\rho \otimes |b\rangle\langle b|$.

2. The two system interact via a unitary operator U, i.e. the system is transformed according to $U(\rho \otimes |b\rangle\langle b|)U^*$.

3. The outcome b_j is obtained with probability

$$p(b_j) = \mathrm{tr}((I \otimes \Pi_j)U(\rho \otimes |b\rangle\langle b|)U^*(I \otimes \Pi_j)).$$

4. The expectation value (average measurement value) is given by

$$\mathrm{tr}(U(\rho \otimes |b\rangle\langle b|)U^*(I \otimes B)).$$

5. The state of the system after measurement is in the measured state of the system (i.e. the system is projected onto the state corresponding to the measurement outcome)

$$\sigma_{b_j} = \frac{(I \otimes \Pi_j)U(\rho \otimes |b\rangle\langle b|)U^*(I \otimes \Pi_j)}{p(b_j)}.$$

6. The state of the original system after discarding the ancilla system is given by

$$\rho_{b_j} = \mathrm{tr}_{\mathcal{H}_A}(\sigma_{b_j}).$$

Let ϕ_B denote an orthonormal basis for $\mathcal{H}_A$ of normalized eigenvectors of B, and so $|b\rangle \in \phi_B$. Writing the unitary operator U as

$$U = \sum_{|j\rangle,|k\rangle \in \phi_B} U_{jk} \otimes |j\rangle\langle k|$$

we find that the constraint $UU^* = U^*U = I$ yields

$$\begin{aligned} U^*U &= \sum_{|j\rangle,|k\rangle,|l\rangle,|m\rangle \in \phi_B} U^*_{jk}U_{lm} \otimes |k\rangle\langle j|l\rangle\langle m| \\ &= \sum_{|j\rangle,|k\rangle,|m\rangle \in \phi_B} U^*_{jk}U_{jm} \otimes |k\rangle\langle m| \\ &= I \end{aligned}$$

and

$$UU^* = \sum_{|j\rangle,|k\rangle,|l\rangle,|m\rangle \in \phi_B} U_{jk}U^*_{lm} \otimes |j\rangle\langle k|m\rangle\langle l|$$

$$= \sum_{|j\rangle,|k\rangle,|l\rangle \in \phi_B} U_{jk} U_{lk}^* \otimes |j\rangle\langle l|$$
$$= I.$$

Equating to I we find

$$\sum_{|j\rangle \in \phi_B} U_{jk}^* U_{jm} = \delta_{km} I$$

and

$$\sum_{|k\rangle \in \phi_B} U_{jk} U_{lk}^* = \delta_{jl} I.$$

Consequently we find

$$(I \otimes \Pi_j) U(\rho \otimes |b\rangle\langle b|) U^* (I \otimes \Pi_j)$$
$$= (I \otimes \Pi_j) \left(\sum_{|m\rangle,|n\rangle \in \phi_B} U_{mb} \rho U_{nb}^* \otimes |m\rangle\langle n| \right) (I \otimes \Pi_j)$$
$$= \sum_{|m\rangle,|n\rangle \in \phi_B} (U_{mb} \rho U_{nb}^*) \otimes (\Pi_j |m\rangle\langle n| \Pi_j).$$

Applying the definition of ρ_{b_j} yields

$$\rho_{b_j} = \sum_{|m\rangle \in \phi_B, \Pi_j |m\rangle = |m\rangle} \frac{1}{p(b_j)} U_{mb} \rho U_{mb}^*$$

where

$$p(b_j) = \mathrm{tr} \left(\sum_{|m\rangle \in \phi_B, \Pi_j |m\rangle = |m\rangle} U_{mb} \rho U_{mb}^* \right).$$

Assuming that the b_j are non-degenerate (i.e. that measurement yields maximal information) and that $|j\rangle$ is the eigenvector of Π_j (for the eigenvalue 1) we have k operators U_{mb} where $m = 1, 2, \ldots, k$ and the *generalized measurement*:

1. $\sum_{m=1}^{k} U_{mb}^* U_{mb} = I.$

2. The outcome b_j is obtained with probability

$$p(b_j) = \mathrm{tr}\left(U_{jb}\rho U_{jb}^*\right).$$

3. The state of the system ρ after the measurement outcome b_j is given by

$$\rho_{b_j} = \frac{1}{p(b_j)} U_{jb}\rho U_{jb}^*.$$

Thus for any k operators U_{mb} satisfying

$$\sum_{m=1}^{k} U_{mb}^* U_{mb} = I$$

we can construct a unitary operator U on the product Hilbert space $\mathcal{H}\otimes\mathcal{H}_A$ for $\mathcal{H}_A$ of appropriate dimension, and B that yields the generalized measurement described above.

Example. Consider the Hilbert space $\mathcal{H} = \mathbf{C}^2$ and the ancilla Hilbert space $\mathcal{H}_A = \mathbf{C}^2$. We construct a generalized measurement described by $\alpha, \beta \in \mathbf{C}$

$$A_0 = \alpha|0\rangle\langle 0| + \beta|1\rangle\langle 1|, \qquad A_1 = \alpha|1\rangle\langle 1| + \beta|0\rangle\langle 0|$$

where $\{|0\rangle, |1\rangle\}$ forms an orthonormal basis in $\mathbf{C}^2$ and $|\alpha|^2 + |\beta|^2 = 1$. Consequently

$$\begin{aligned} A_0^*A_0 + A_1^*A_1 &= (\overline{\alpha}|0\rangle\langle 0| + \overline{\beta}|1\rangle\langle 1|)(\alpha|0\rangle\langle 0| + \beta|1\rangle\langle 1|) \\ &\quad +(\overline{\alpha}|1\rangle\langle 1| + \overline{\beta}|0\rangle\langle 0|)(\alpha|1\rangle\langle 1| + \beta|0\rangle\langle 0|) \\ &= |0\rangle\langle 0| + |1\rangle\langle 1| = I_2. \end{aligned}$$

For $\alpha = 0$ or $\beta = 0$, A_0 and A_1 describe an orthogonal measurement. To construct the unitary operator

$$U := U_{00}\otimes|0\rangle\langle 0| + U_{01}\otimes|0\rangle\langle 1| + U_{10}\otimes|1\rangle\langle 0| + U_{11}\otimes|1\rangle\langle 1|$$

we use the the commutators given by

$$[A_0, A_0^*] = [A_1, A_1^*] = [A_0, A_1] = [A_0, A_1^*] = 0$$

and set

$$U_{00} := A_0, \qquad U_{01} := A_1^*, \qquad U_{10} := A_1, \qquad U_{11} := -A_0^*$$

to satisfy that U is unitary. One orthogonal measurement that implements the generalized measurement is described by the projection operators

$$\Pi_1 = |0\rangle\langle 0|, \qquad \Pi_2 = |1\rangle\langle 1|$$

and the corresponding observable

$$B := -1\Pi_1 + 1\Pi_2.$$

Thus the generalized measurement of ρ, a density operator on $\mathcal{H}$, may proceed as follows

1. Perform U on $\rho \otimes |0\rangle\langle 0|$.
2. Measure the ancillary system with respect to the observable B.

■

Chapter 2

Entanglement

2.1 von Neumann Entropy

The von Neumann entropy plays a similar role in quantum computing to the role Shannon entropy plays in classical computing [HardyQC].

Let ρ denote a *density operator*, i.e., a bounded positive semi-definite operator with unit trace. Suppose further that ρ is a density operator on the finite-dimensional product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $n = \dim(\mathcal{H}_B)$. We define

$$\rho_A := \mathrm{tr}_B \rho, \qquad \rho_B := \mathrm{tr}_A \rho.$$

Let $|\phi_1\rangle, \ldots, |\phi_n\rangle$ be an orthonormal basis in $\mathcal{H}_B$. Thus

$$\rho_A = \sum_{j=1}^{n} (I \otimes \langle\phi_j|)\rho(I \otimes |\phi_j\rangle).$$

For all $|x\rangle \in \mathcal{H}_A$ we have

$$\langle x|\rho_A|x\rangle = \sum_{j=1}^{n} (\langle x| \otimes \langle\phi_j|)\rho(|x\rangle \otimes |\phi_j\rangle) \geq 0$$

since ρ is positive semidefinite. Thus ρ_A is positive semidefinite. Furthermore $\mathrm{tr}\rho_A = \mathrm{tr}_A(\mathrm{tr}_B\rho) = \mathrm{tr}\rho = 1$. Thus ρ_A is also a density operator on $\mathcal{H}_A$, similarly ρ_B is a density operator on $\mathcal{H}_B$.

Definition. The *von Neumann entropy* $S_b(\rho)$ of a density operator ρ is defined as

$$S_b(\rho) := -\mathrm{tr}\rho \log_b \rho$$

where the base b of the log determines the units. For $b = 2$ the units are qubits. We use the convention $S := S_2$.

■

Let ρ be a density operator on a finite-dimensional Hilbert space. The spectral decomposition of ρ gives

$$\rho = \sum_{j=1}^{n} \lambda_j |\phi_j\rangle\langle\phi_j|$$

where the dimension of the Hilbert space is n, and $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$ are the eigenvalues of ρ with corresponding (orthonormal) eigenstates $|\phi_1\rangle, \ldots, |\phi_n\rangle$ and

$$\sum_{j=1}^{n} \lambda_j = 1.$$

Since the eigenvalues of a density operator ρ are non-negative and sum to one, it is obvious that $S_b(\rho) \geq 0$:

$$\begin{aligned}
S_b(\rho) &= -\mathrm{tr}\left[\left(\sum_{j=1}^{n} \lambda_j |\phi_j\rangle\langle\phi_j|\right) \log_b \left(\sum_{j=1}^{n} \lambda_j |\phi_j\rangle\langle\phi_j|\right)\right] \\
&= -\mathrm{tr}\left[\left(\sum_{j=1}^{n} \lambda_j |\phi_j\rangle\langle\phi_j|\right) \left(\sum_{j=1}^{n} (\log_b \lambda_j) |\phi_j\rangle\langle\phi_j|\right)\right] \\
&= -\mathrm{tr}\left[\sum_{j=1}^{n} \lambda_j (\log_b \lambda_j) |\phi_j\rangle\langle\phi_j|\right] \\
&= -\sum_{j=1}^{n} \lambda_j \log_b \lambda_j \geq 0
\end{aligned}$$

since $0 \leq \lambda_j \leq 1$ and $\log_b \lambda_j \leq 0$.

Example. Let $|\psi\rangle$ denote a pure state in a Hilbert space. Thus the density operator ρ_ψ for $|\psi\rangle$ is given by

$$\rho_\psi = |\psi\rangle\langle\psi|$$

and we find that $S_b(\rho_\psi) = 0$ due to the fact that one eigenvalue of ρ_ψ is 1 and the thus remaining eigenvalues of ρ are 0.
■

Example. Consider the Hilbert space $\mathbf{C}^2$ with an orthonormal basis $\{|0\rangle, |1\rangle\}$. Let

$$\rho_{01} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|.$$

Thus the two eigenvalues of ρ_{01} are both $\frac{1}{2}$. Then

$$S_2(\rho_{01}) = 1.$$

■

The von Neumann entropy is a measure of the uncertainty of the pure state represented in the density operator.

Example. Since $S_b(\rho_\psi) = 0$, from the previous examples, there is no uncertainty that ρ_ψ represents $|\psi\rangle$ up to global phase. Thus $|\phi\rangle = e^{i\theta}|\psi\rangle$ leads to

$$\rho = |\phi\rangle\langle\phi| = |\psi\rangle\langle\psi|.$$

The density operator ρ_{01} is an equal mixture of the $|0\rangle$ and $|1\rangle$ states. Since $S_2(\rho_{01}) = 1$ the uncertainty in the pure state represented in ρ_{01} is maximal.
■

Definition. The *quantum relative entropy* [Henderson99] between two density operators ρ and σ is defined by

$$S_b(\rho||\sigma) := \mathrm{tr}(\rho \log_b \rho - \rho \log_b \sigma) = -S_b(\rho) - \mathrm{tr}\rho \log_b \sigma.$$

■

Consider the term $\rho \log_b \sigma$ where ρ and σ are density operators on a finite-dimensional Hilbert space of dimension n. Let λ_j and $|\phi_j\rangle$ be the corresponding eigenvalues and (orthonormal) eigenstates of ρ. Similarly, let μ_k and $|\psi_k\rangle$

be the corresponding eigenvalues and (orthonormal) eigenstates of σ. Thus we have

$$\begin{aligned}\rho \log_2 \sigma &= \left(\sum_{j=1}^{n} \lambda_j |\phi_j\rangle\langle\phi_j|\right)\left(\sum_{k=1}^{n} \log_b \mu_k |\psi_k\rangle\langle\psi_k|\right) \\ &= \sum_{j,k=1}^{n} \lambda_j \log_b \mu_k \langle\phi_j|\psi_k\rangle |\phi_j\rangle\langle\psi_k|.\end{aligned}$$

Taking the trace using the basis $\{|\psi_k\rangle,\ k = 1, 2, \ldots, n\}$ yields

$$\begin{aligned}\mathrm{tr}(\rho \log_b \sigma) &= \sum_{j,k,l=1}^{n} \lambda_j (\log_b \mu_k) \langle\phi_j|\psi_k\rangle\langle\psi_l|\phi_j\rangle\langle\psi_k|\psi_l\rangle \\ &= \sum_{j,k=1}^{n} \lambda_j (\log_b \mu_k) \langle\phi_j|\psi_k\rangle\langle\psi_k|\phi_j\rangle \\ &= \sum_{j,k=1}^{n} \lambda_j (\log_b \mu_k) |\langle\phi_j|\psi_k\rangle|^2.\end{aligned}$$

Thus we obtain

$$\begin{aligned}S_b(\rho||\sigma) &= \sum_{j=1}^{n} \lambda_j \log_b \lambda_j - \sum_{j,k=1}^{n} \lambda_j (\log_b \mu_k) |\langle\phi_j|\psi_k\rangle|^2 \\ &= \sum_{j=1}^{n} \lambda_j \log_b \lambda_j - \sum_{j=1}^{n} \lambda_j \log_b \left(\prod_{k=1}^{n} \mu_k^{|\langle\phi_j|\psi_k\rangle|^2}\right) \\ &= \sum_{j=1}^{n} \lambda_j \log_b \lambda_j - \sum_{j=1}^{n} \lambda_j \log_b \nu_j\end{aligned}$$

where, for convenience, we defined

$$\nu_j := \left(\prod_{k=1}^{n} \mu_k^{|\langle\phi_j|\psi_k\rangle|^2}\right).$$

If $\nu_j = 0$ for some j where $\lambda_j \neq 0$ then $S_B(\rho||\sigma) \geq 0$ trivially. In the following we assume $\nu_j = 0$ if and only if $\lambda_j = 0$. Since

$$\sum_{k=1}^{n} \mu_k = \sum_{k=1}^{n} |\langle \phi_j | \psi_k \rangle|^2 = \sum_{j=1}^{n} |\langle \phi_j | \psi_k \rangle|^2 = 1$$

we find that

$$0 \leq \nu_j \leq 1, \qquad \alpha := \sum_{j=1}^{n} \nu_j \leq 1.$$

We determine the maximum value of

$$\sum_{j=1}^{n} \lambda_j \log_b x_j, \qquad \sum_{j=1}^{n} x_j = \alpha$$

which can be formulated as a *Lagrange multiplier problem*, i.e. we find the critical points of

$$f(x_1, \ldots, x_n) := \sum_{j=1}^{n} \lambda_j \log_b x_j - \theta \left(\sum_{j=1}^{n} x_j - \alpha \right)$$

where θ is the Lagrange multiplier. Thus we find

$$\frac{\partial f}{\partial x_j} = \begin{cases} \frac{\lambda_j}{x_j \ln b} - \theta & \lambda_j \neq 0 \\ 0 & \lambda_j = 0 \end{cases}.$$

Since $\frac{\partial f}{\partial x_j} = \frac{\partial f}{\partial x_k} = 0$ we have, when $\lambda_j \neq 0$ and $\lambda_k \neq 0$,

$$x_k = \frac{\lambda_k}{\lambda_j} x_j.$$

We seek a maximum, thus we require that $\lambda_j = 0$ implies $x_j = 0$. Inserting $x_k \neq 0$ into the constraint yields

$$\sum_{j=1}^{n} x_j = x_k + \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{\lambda_j}{\lambda_k} x_k = \alpha.$$

Thus $x_k = \alpha \lambda_k$, since $\sum_{j=1}^{n} \lambda_j = 1$. Thus we have a maximum and consequently

$$\begin{aligned} S_b(\rho||\sigma) &= \sum_{j=1}^{n} \lambda_j \log_b \lambda_j - \sum_{j=1}^{n} \lambda_j \log_b \nu_j \\ &\geq \sum_{j=1}^{n} \lambda_j \log_b \lambda_j - \sum_{j=1}^{n} \lambda_j \log_b(\alpha\lambda_j) \\ &= -\log_b \alpha \geq 0. \end{aligned}$$

The above inequality is known as *Klein's inequality* [Nielsen].

Consider the special case $\sigma = \rho_A \otimes \rho_B$. The eigenvalues of σ are the products of eigenvalues $\lambda_{A,j}$ of ρ_A and $\lambda_{B,k}$ of ρ_B. The corresponding (orthonormal) eigenstates are $|\phi_{A,j}\rangle \otimes |\phi_{B,k}\rangle$ composed of the orthonormal eigenstates of ρ_A and ρ_B corresponding to $\lambda_{A,j}$ and $\lambda_{B,k}$. Consequently

$$\begin{aligned} S_b(\rho||\rho_A \otimes \rho_B) &= -S_b(\rho) - \mathrm{tr}\rho \log_b(\rho_A \otimes \rho_B) \\ &= -S_b(\rho) - \mathrm{tr}_A\left(\mathrm{tr}_B\left(\rho \log_b(\rho_A \otimes \rho_B)\right)\right) \\ &= -S_b(\rho) - \mathrm{tr}_A\Bigg(\mathrm{tr}_B\Bigg(\\ &\qquad \rho \sum_{j=1}^{m}\sum_{k=1}^{n} \log_b(\lambda_{A,j}\lambda_{B,k}) |\phi_{A,j}\rangle\langle\phi_{A,j}| \otimes |\phi_{B,k}\rangle\langle\phi_{B,k}| \Bigg)\Bigg) \\ &= -S_b(\rho) - \mathrm{tr}_A\left(\mathrm{tr}_B\left(\rho \sum_{j=1}^{m} \log_b \lambda_{A,j} |\phi_{A,j}\rangle\langle\phi_{A,j}| \otimes I_n\right)\right) \\ &\qquad -\mathrm{tr}_B\left(\mathrm{tr}_A\left(\rho \sum_{k=1}^{n} \log_b \lambda_{B,k} I_m \otimes |\phi_{B,k}\rangle\langle\phi_{B,k}|\right)\right) \\ &= -S_b(\rho) - \mathrm{tr}_A\left(\rho_A \sum_{j=1}^{m} \log_b \lambda_{A,j} |\phi_{A,j}\rangle\langle\phi_{A,j}|\right) \\ &\qquad -\mathrm{tr}_B\left(\rho_B \sum_{k=1}^{n} \log_b \lambda_{B,k} |\phi_{B,k}\rangle\langle\phi_{B,k}|\right) \\ &= S_b(\rho_A) + S_b(\rho_B) - S_b(\rho). \end{aligned}$$

From Klein's inequality, $S(\rho||\rho_A \otimes \rho_B) \geq 0$, we obtain the property of sub-additivity for the von Neumann entropy:

$$S_b(\rho) \leq S_b(\rho_A) + S_b(\rho_B).$$

2.2 Schmidt Number and Schmidt Rank

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two finite-dimensional (with dimension greater than 1) Hilbert spaces with the underlying field $\mathbf{C}$. Let $|\psi\rangle$ denote a pure state in the product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Definition. The *Schmidt number* (also called the *Schmidt rank*) of $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ over $\mathcal{H}_A \otimes \mathcal{H}_B$ is the smallest non-negative integer $\mathrm{Sch}(|\psi\rangle, \mathcal{H}_A, \mathcal{H}_B)$ such that $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{j=1}^{\mathrm{Sch}(|\psi\rangle,\mathcal{H}_A,\mathcal{H}_B)} |j\rangle_A \otimes |j\rangle_B$$

where $|j\rangle_A \in \mathcal{H}_A$ and $|j\rangle_B \in \mathcal{H}_B$.
■

Theorem. Suppose $\{|1_A\rangle, |2_A\rangle, \ldots, |\dim(\mathcal{H}_A)_A\rangle\}$ is an orthonormal basis in $\mathcal{H}_A$ and $\{|1_B\rangle, |2_B\rangle, \ldots, |\dim(\mathcal{H}_B)_B\rangle\}$ is an orthonormal basis in $\mathcal{H}_B$. Writing $|\psi\rangle$ in these bases as

$$\sum_{j=1}^{\dim(\mathcal{H}_A)} \sum_{k=1}^{\dim(\mathcal{H}_B)} \psi_{jk} |j_A\rangle \otimes |k_B\rangle$$

then

$$\mathrm{Sch}(|\psi\rangle, \mathcal{H}_A, \mathcal{H}_B) = \mathrm{rank}(M) = \mathrm{rank}(\mathrm{tr_A}(|\psi\rangle\langle\psi|))$$

where M is a $\dim(\mathcal{H}_A) \times \dim(\mathcal{H}_B)$ matrix with entries $(M)_{jk} = \psi_{jk}$.
■

Theorem. Let $\mathcal{H}_A$, $\mathcal{H}_B$, $\mathcal{H}_C$ and $\mathcal{H}_D$ denote arbitrary finite-dimensional Hilbert spaces. Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and $|\phi\rangle \in \mathcal{H}_C \otimes \mathcal{H}_D$. Then $|\psi\rangle \otimes |\phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$. Now let $|\theta\rangle \in \mathcal{H}_A \otimes \mathcal{H}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_D$ denote the state corresponding to $|\psi\rangle \otimes |\phi\rangle$ with the Hilbert spaces reordered. It follows that [NielsenBenasque]

$$\mathrm{Sch}(|\theta\rangle, \mathcal{H}_A \otimes \mathcal{H}_C, \mathcal{H}_B \otimes \mathcal{H}_D) = \mathrm{Sch}(|\psi\rangle, \mathcal{H}_A, \mathcal{H}_B)\mathrm{Sch}(|\phi\rangle, \mathcal{H}_C, \mathcal{H}_D).$$

■

Theorem. Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ where $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_C$ denote arbitrary finite-dimensional Hilbert spaces then

$$\mathrm{Sch}(|\psi\rangle, \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{H}_C) \leq \dim(\mathcal{H}_B)\mathrm{Sch}(|\psi\rangle, \mathcal{H}_A, \mathcal{H}_B \otimes \mathcal{H}_C).$$

■

Let

$$\rho := |\psi\rangle\langle\psi|, \qquad \rho_A = \mathrm{tr}_{\mathcal{H}_B}(\rho), \qquad \rho_B = \mathrm{tr}_{\mathcal{H}_A}(\rho).$$

Theorem. ρ_A and ρ_B have the same nonzero eigenvalues $\lambda_1, \ldots, \lambda_k$ (with the same multiplicities) and any extra dimensions are made up with zero eigenvalues, where $k \leq \min(\dim(\mathcal{H}_A), \dim(\mathcal{H}_B))$. Here k is the Schmidt rank of $|\psi\rangle$. There is no need for $\mathcal{H}_A$ and $\mathcal{H}_B$ to have the same dimension, so the number of zero eigenvalues of ρ_A and ρ_B can differ. Furthermore, the state $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{j=1}^{k} \sqrt{\lambda_j}|j\rangle_A \otimes |j\rangle_B$$

where $|j\rangle_A$ (respectively $|j\rangle_B$) are orthonormal eigenvectors of ρ_A in $\mathcal{H}_A$ (respectively ρ_B in $\mathcal{H}_B$) belonging to the eigenvalue λ_j. This expression is called the *Schmidt polar form* or *Schmidt decomposition* [HardyQC]. The Schmidt number is the number of non-zero λ_j. If the eigenvalues λ_j are non-degenerate, then the Schmidt decomposition can be obtained from ρ_A and ρ_B exclusively. Clearly a separable state has Schmidt number 1 and an entangled state has Schmidt number greater than 1. Furthermore

$$S_b(\rho_A) = S_b(\rho_B) = -\sum_{j=1}^{k} \lambda_j \log_b \lambda_j.$$

■

Let $\{|\alpha_1\rangle, \ldots, |\alpha_{\dim(\mathcal{H}_A)}\rangle\}$ and $\{|\beta_1\rangle, \ldots, |\beta_{\dim(\mathcal{H}_B)}\rangle\}$ be orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Thus we can write

$$|\psi\rangle = \sum_{i=1}^{\dim(\mathcal{H}_A)} \sum_{j=1}^{\dim(\mathcal{H}_B)} c_{ij}|\alpha_i\rangle \otimes |\beta_j\rangle.$$

The Schmidt decomposition then follows from the singular value decomposition of the matrix C with entries $(C)_{ij} = c_{ij}$, where λ_j are the singular values of C. From the *singular value decomposition* of C we find that the matrix can be written in the form

$$C = USV^*$$

where U is a $\dim(\mathcal{H}_\mathcal{A}) \times \dim(\mathcal{H}_\mathcal{B})$ matrix with mutually orthogonal column vectors, S is a $\dim(\mathcal{H}_\mathcal{B}) \times \dim(\mathcal{H}_\mathcal{B})$ diagonal matrix and V is a $\dim(\mathcal{H}_\mathcal{B}) \times \dim(\mathcal{H}_\mathcal{B})$ matrix with mutually orthogonal column vectors. Thus we write

$$U = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n), \qquad \mathbf{u}_k \in \mathbf{C}^m, \quad k = 1, 2, \ldots, n$$

$$S = \mathrm{diag}(s_1, s_2, \ldots, s_n)$$

$$V = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)^*, \qquad \mathbf{v}_k \in \mathbf{C}^n, \quad k = 1, 2, \ldots, n.$$

The numbers s_k $(k = 1, 2, \ldots, \dim(\mathcal{H}_B))$ are the singular values. Consequently

$$USV = \sum_{k=1}^{n} s_k \mathbf{u}_k \mathbf{v}_k^*$$

$$c_{ij} = \sum_{k=1}^{n} s_k u_{ki} v_{kj}^*$$

where u_{ki} denotes the i-th component of $\mathbf{u}_k$. Finally

$$|\psi\rangle = \sum_{i=1}^{\dim(\mathcal{H}_A)} \sum_{j=1}^{\dim(\mathcal{H}_B)} \sum_{k=1}^{\dim(\mathcal{H}_B)} s_k u_{ki} v_{kj}^* |\alpha_i\rangle \otimes |\beta_j\rangle$$

$$= \sum_{k=1}^{\dim(\mathcal{H}_B)} s_k \left[\left(\sum_{i=1}^{\dim(\mathcal{H}_A)} u_{ki}|\alpha_i\rangle \right) \otimes \left(\sum_{j=1}^{\dim(\mathcal{H}_B)} v_{kj}^*|\beta_j\rangle \right) \right].$$

The states

$$\sum_{i=1}^{\dim(\mathcal{H}_A)} u_{ki}|\alpha_i\rangle$$

are mutually orthogonal due to the mutual orthogonality of the $\mathbf{u}_k$ ($k = 1, 2, \ldots, \dim(\mathcal{H}_A)$). Similarly the states

$$\sum_{j=1}^{\dim(\mathcal{H}_B)} v_{kj}^*|\beta_j\rangle$$

are mutually orthogonal. Thus we have deduced the *Schmidt polar form* described above where $s_k = \sqrt{\lambda_k}$ and

$$|j_A\rangle = \sum_{i=1}^{\dim(\mathcal{H}_A)} u_{ki}|\alpha_i\rangle$$

$$|j_B\rangle = \sum_{j=1}^{\dim(\mathcal{H}_B)} v_{kj}^*|\beta_j\rangle.$$

Example. Consider $\dim(\mathcal{H}_A) = 3$, $\dim(\mathcal{H}_B) = 2$ and

$$|\psi\rangle := \frac{1}{\sqrt{2}}(1, 0, 0, 0, 0, 1)^T.$$

We have

$$\rho = |\psi\rangle\langle\psi|, \quad \rho_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \rho_2 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The eigenvalues of ρ_1 are $\frac{1}{2}, \frac{1}{2}$ and 0. The eigenvalues of ρ_2 are $\frac{1}{2}$ and $\frac{1}{2}$. Thus

$$\rho_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}(1, 0, 0) + \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}(0, 0, 1)$$

$$\rho_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1, 0) + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}(0, 1)$$

$$|\psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In this case the eigenvalues are degenerate.
■

The definition of the Schmidt number for pure states can also be applied to matrices (linear operators).

Definition. The *Schmidt rank* of a linear operator $L : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_B$ over $\mathcal{H}_A \otimes \mathcal{H}_B$ is the smallest non-negative integer $\mathrm{Sch}(L, \mathcal{H}_A, \mathcal{H}_B)$ such that L can be written as

$$L = \sum_{j=1}^{\mathrm{Sch}(L,\mathcal{H}_A,\mathcal{H}_B)} L_{j,A} \otimes L_{j,B}$$

where $L_{j,A} : \mathcal{H}_A \to \mathcal{H}_A$ and $L_{j,B} : \mathcal{H}_B \to \mathcal{H}_B$ are linear operators.
■

The *Hilbert-Schmidt inner product* of two linear operators A and B acting on the same Hilbert space $\mathcal{H}$ is given by

$$\langle A, B \rangle = \mathrm{tr}(B^* A).$$

The linear operators A and B are said to be orthogonal operators if $\langle A, B \rangle = 0$. Thus the linear operators form a Hilbert space with the Hilbert-Schmidt inner product. This leads to the *operator-Schmidt decomposition* [Nielsen2003]

$$L = \sum_{j=1}^{\mathrm{Sch}(L,\mathcal{H}_A,\mathcal{H}_B)} s_j L_{j,A} \otimes L_{j,B}$$

$$\langle L_{j,A}, L_{k,A} \rangle = \langle L_{j,B}, L_{k,B} \rangle = \delta_{jk}, \quad j, k = 1, 2, \ldots, \mathrm{Sch}(L, \mathcal{H}_A, \mathcal{H}_B)$$

where the values s_j are found from the singular value decomposition.

Example. Consider the unitary operator (*controlled NOT* or *CNOT gate*)

$$U_{CNOT} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Defining

$$|0\rangle := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad |1\rangle := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$$|I_2\rangle := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad |U_{NOT}\rangle := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and using the Hilbert-Schmidt inner product we find

$$\begin{aligned} \langle 0|0\rangle &= \langle 1|1\rangle = 1 \\ \langle 0|1\rangle &= \langle 1|0\rangle = 0 \\ \langle I_2|I_2\rangle &= \langle U_{NOT}|U_{NOT}\rangle = 2 \\ \langle I_2|U_{NOT}\rangle &= \langle U_{NOT}|I_2\rangle = 0. \end{aligned}$$

Consequently we can write U_{CNOT} as a state in a Hilbert space

$$|U_{CNOT}\rangle = |0\rangle \otimes |I_2\rangle + |1\rangle \otimes |U_{NOT}\rangle.$$

We simply apply the technique for the Schmidt decomposition to obtain the operator-Schmidt decomposition for U_{CNOT}. Taking the partial trace of $|U_{CNOT}\rangle\langle U_{CNOT}|$ using the basis $\{|0\rangle \otimes I_{2\times 2}, |1\rangle \otimes I_{2\times 2}\}$ yields

$$\rho_A = |I_2\rangle\langle I_2| + |U_{NOT}\rangle\langle U_{NOT}|$$

and taking the partial trace using the basis $\{\frac{1}{\sqrt{2}} I_{2\times 2} \otimes |I_2\rangle, \frac{1}{\sqrt{2}} I_{2\times 2} \otimes |U_{NOT}\rangle\}$ yields

$$\rho_B = 2|0\rangle\langle 0| + 2|1\rangle\langle 1|.$$

Here $I_{2\times 2}$ is the identity operator on 2×2 matrices. We find that ρ_A and ρ_B have the same eigenvalues 2 with multiplicity 2, and orthonormal eigenvectors $\{\frac{1}{\sqrt{2}}|I_2\rangle, \frac{1}{\sqrt{2}}|U_{NOT}\rangle\}$ and $\{|0\rangle, |1\rangle\}$ respectively. Thus the Schmidt decomposition is given by

$$|U_{CNOT}\rangle = \sqrt{2}|0\rangle \otimes \frac{1}{\sqrt{2}}|I_2\rangle + \sqrt{2}|1\rangle \otimes \frac{1}{\sqrt{2}}|U_{NOT}\rangle,$$

so that the operator-Schmidt decomposition is given by

$$U_{CNOT} = \sqrt{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sqrt{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

■

Example. Let U_q be a unitary operation on the Hilbert space $\mathbf{C}^n \otimes \mathbf{C}^2$. We can write U_q in the form

$$U_q = Q_0 \otimes |0\rangle\langle 0| + Q_1 \otimes |0\rangle\langle 1| + Q_2 \otimes |1\rangle\langle 0| + Q_3 \otimes |1\rangle\langle 1| \tag{2.1}$$

where Q_0, Q_1, Q_2 and Q_3 are $n \times n$ matrices over $\mathbf{C}$ (some of which may be the zero matrix), and $\{|0\rangle, |1\rangle\}$ is an orthonormal basis in $\mathbf{C}^2$. Next we determine the constraints on Q_0, Q_1, Q_2 and Q_3 from the condition $U_q^* U_q = I_n \otimes I_2$. Since U_q is a unitary operator we have

$$\begin{aligned} U_q U_q^* &= (Q_0 Q_0^* + Q_1 Q_1^*) \otimes |0\rangle\langle 0| + (Q_0 Q_2^* + Q_1 Q_3^*) \otimes |0\rangle\langle 1| \\ &\quad + (Q_2 Q_2^* + Q_3 Q_3^*) \otimes |1\rangle\langle 1| + (Q_2 Q_0^* + Q_3 Q_1^*) \otimes |1\rangle\langle 0| \\ &= I_n \otimes I_2 = U_q^* U_q. \end{aligned}$$

I_n denotes the $n \times n$ identity matrix and 0_n denotes the $n \times n$ zero matrix. We find the conditions

$$\begin{aligned} Q_0 Q_0^* + Q_1 Q_1^* &= I_n \qquad Q_0^* Q_0 + Q_2^* Q_2 = I_n \\ Q_2 Q_2^* + Q_3 Q_3^* &= I_n \qquad Q_1^* Q_1 + Q_3^* Q_3 = I_n \\ Q_2 Q_0^* + Q_3 Q_1^* &= 0_n \qquad Q_1^* Q_0 + Q_3^* Q_2 = 0_n. \end{aligned} \tag{2.2}$$

First we determine the implications for the Schmidt rank from (2.2). Let A be an $n \times n$ matrix over $\mathbf{C}$. Then $AA^* = 0_n$ implies $A = 0_n$ [SteebMC]. We note that if $Q_1 = 0_n$ the diagonal elements of $Q_2^* Q_2$ and $Q_2 Q_2^*$ must be zero. Furthermore $Q_2^* Q_2$ and $Q_2 Q_2^*$ are positive semi-definite so that $Q_2^* Q_2 = Q_2 Q_2^* = 0_n$ (the eigenvalues of $Q_2^* Q_2$ and $Q_2 Q_2^*$ are zero [SteebMC]). So we conclude that $Q_2 = 0_n$. Similarly $Q_2 = 0_n$ implies $Q_1 = 0_n$ and $Q_0 = 0_n \iff Q_3 = 0_n$. In other words when one of Q_0, Q_1, Q_2 and Q_3 is the zero matrix, the non-zero matrices of Q_0, Q_1, Q_2 and Q_3 are unitary. This result is independent of the chosen basis $\{|0\rangle, |1\rangle\}$. Thus we find the Schmidt rank of U_q over $\mathbf{C}^n \otimes \mathbf{C}^2$ is either 1, 2 or 4.

The matrices $Q_0 Q_0^*$, $Q_1 Q_1^*$, $Q_2 Q_2^*$, $Q_3 Q_3^*$, $Q_0^* Q_0$, $Q_1^* Q_1$, $Q_2^* Q_2$ and $Q_3^* Q_3$ are positive semi-definite. From equations (3.2) we deduce the following. Let U_0 be a unitary matrix which diagonalizes $Q_0 Q_0^*$. From

$$Q_0 Q_0^* + Q_1 Q_1^* = I_n$$

it follows that

$$U_0 Q_0 Q_0^* U_0^* + U_0 Q_1 Q_1^* U_0^* = I_n.$$

Since I_n is a diagonal matrix we find that U_0 also diagonalizes the positive semi-definite matrix $Q_1 Q_1^*$. Furthermore, $U_0^*|e_j\rangle$ is simultaneously an eigenvector for $Q_0 Q_0^*$ and $Q_1 Q_1^*$ where $|e_j\rangle$ is an element of the standard basis in $\mathbf{C}^n$ and $j = 1, 2, \ldots, n$. The eigenvalues corresponding to the eigenvector $U_0^*|e_j\rangle$ for $Q_0 Q_0^*$ and $Q_1 Q_1^*$ sum to 1 and are both non-negative.

Now suppose none of the operators are unitary (Schmidt rank 4 over $\mathbf{C}^n \otimes \mathbf{C}^2$), in other words the matrices Q_0, Q_1, Q_2 and Q_3 are linearly independent. The matrices Q_0, Q_1, Q_2 and Q_3 span a 4-dimensional vector space which is isomorphic to the space spanned by $\{|0\rangle\langle 0|, |0\rangle\langle 1|, |1\rangle\langle 0|, |1\rangle\langle 1|\}$. The matrices $\{|0\rangle\langle 0|, |0\rangle\langle 1|, |1\rangle\langle 0|, |1\rangle\langle 1|\}$ are linearly independent, with the matrix representations in the orthonormal basis $\{|0\rangle, |1\rangle\}$ of $\mathbf{C}^2$

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

We define

$$\begin{aligned} U_{NOT} &:= |0\rangle\langle 1| + |1\rangle\langle 0| \\ U_{NOT}(1) &:= U_{NOT} \otimes I \\ U_{NOT}(2) &:= I \otimes U_{NOT} \\ U_{CNOT}(1,2) &:= |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U_{NOT} \\ U_{CNOT}(2,1) &:= I \otimes |0\rangle\langle 0| + U_{NOT} \otimes |1\rangle\langle 1| \\ U_{SWAP} &:= |00\rangle\langle 00| + |10\rangle\langle 01| + |01\rangle\langle 10| + |11\rangle\langle 11|. \end{aligned}$$

The vector space isomorphisms which do not violate the conditions (2.2) are

1) $Q_0 \to |0\rangle\langle 0|, \quad Q_1 \to |1\rangle\langle 0|, \quad Q_2 \to |0\rangle\langle 1|, \quad Q_3 \to |1\rangle\langle 1|$

$$|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| + |0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|$$

$$= U_{SWAP}$$

2) $Q_0 \to |1\rangle\langle 0|, \quad Q_1 \to |0\rangle\langle 0|, \quad Q_2 \to |1\rangle\langle 1|, \quad Q_3 \to |0\rangle\langle 1|$

$$|1\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |0\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 0| + |0\rangle\langle 1| \otimes |1\rangle\langle 1|$$

$= U_{NOT}(1)U_{SWAP}$

3) $Q_0 \rightarrow |0\rangle\langle 1|, \quad Q_1 \rightarrow |1\rangle\langle 1|, \quad Q_2 \rightarrow |0\rangle\langle 0|, \quad Q_3 \rightarrow |1\rangle\langle 0|$

$|0\rangle\langle 1| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |0\rangle\langle 1| + |0\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 0| \otimes |1\rangle\langle 1|$

$= U_{NOT}(2)U_{SWAP}$

4) $Q_0 \rightarrow |1\rangle\langle 1|, \quad Q_1 \rightarrow |0\rangle\langle 1|, \quad Q_2 \rightarrow |1\rangle\langle 0|, \quad Q_3 \rightarrow |0\rangle\langle 0|$

$|1\rangle\langle 1| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |0\rangle\langle 0| \otimes |1\rangle\langle 1|$

$= U_{NOT}(1)U_{NOT}(2)U_{SWAP}$

5) $Q_0 \rightarrow |0\rangle\langle 0|, \quad Q_1 \rightarrow |1\rangle\langle 1|, \quad Q_2 \rightarrow |0\rangle\langle 1|, \quad Q_3 \rightarrow |1\rangle\langle 0|$

$|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |0\rangle\langle 1| + |0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 0| \otimes |1\rangle\langle 1|$

$= U_{CNOT}(1,2)U_{SWAP}$

6) $Q_0 \rightarrow |1\rangle\langle 0|, \quad Q_1 \rightarrow |0\rangle\langle 1|, \quad Q_2 \rightarrow |1\rangle\langle 1|, \quad Q_3 \rightarrow |0\rangle\langle 0|$

$|1\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 0| + |0\rangle\langle 0| \otimes |1\rangle\langle 1|$

$= U_{NOT}(1)U_{CNOT}(1,2)U_{SWAP}$

7) $Q_0 \rightarrow |0\rangle\langle 1|, \quad Q_1 \rightarrow |1\rangle\langle 0|, \quad Q_2 \rightarrow |0\rangle\langle 0|, \quad Q_3 \rightarrow |1\rangle\langle 1|$

$|0\rangle\langle 1| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| + |0\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|$

$= U_{NOT}(2)U_{CNOT}(1,2)U_{SWAP}$

8) $Q_0 \rightarrow |1\rangle\langle 1|, \quad Q_1 \rightarrow |0\rangle\langle 0|, \quad Q_2 \rightarrow |1\rangle\langle 0|, \quad Q_3 \rightarrow |0\rangle\langle 1|$

$|1\rangle\langle 1| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |0\rangle\langle 1| \otimes |1\rangle\langle 1|$

$= U_{NOT}(1)U_{NOT}(2)U_{CNOT}(1,2)U_{SWAP}$

9) $Q_0 \rightarrow |0\rangle\langle 0|, \quad Q_1 \rightarrow |1\rangle\langle 0|, \quad Q_2 \rightarrow |1\rangle\langle 1|, \quad Q_3 \rightarrow |0\rangle\langle 1|$

$|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 0| + |0\rangle\langle 1| \otimes |1\rangle\langle 1|$

$= U_{CNOT}(2,1)U_{SWAP}$

10) $Q_0 \rightarrow |1\rangle\langle 0|, \quad Q_1 \rightarrow |0\rangle\langle 0|, \quad Q_2 \rightarrow |0\rangle\langle 1|, \quad Q_3 \rightarrow |1\rangle\langle 1|$

$|1\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |0\rangle\langle 1| + |0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|$

$$= U_{NOT}(1)U_{CNOT}(2,1)U_{SWAP}$$

11) $Q_0 \to |1\rangle\langle 1|, \quad Q_1 \to |0\rangle\langle 1|, \quad Q_2 \to |0\rangle\langle 0|, \quad Q_3 \to |1\rangle\langle 0|$

$$|1\rangle\langle 1| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |0\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 0| \otimes |1\rangle\langle 1|$$

$$= U_{NOT}(2)U_{CNOT}(2,1)U_{SWAP}$$

12) $Q_0 \to |0\rangle\langle 1|, \quad Q_1 \to |1\rangle\langle 1|, \quad Q_2 \to |1\rangle\langle 0|, \quad Q_3 \to |0\rangle\langle 0|$

$$|0\rangle\langle 1| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |0\rangle\langle 0| \otimes |1\rangle\langle 1|$$

$$= U_{NOT}(1)U_{NOT}(2)U_{CNOT}(2,1)U_{SWAP}$$

The operators Q_0, Q_1, Q_2 and Q_3 can be written in polar form $Q_j = U_j H_j$, where U_j is unitary any H_j is positive semi-definite for $j = 0, 1, 2, 3$. Inserting the polar form into equations (2.2) yields the set of equations

$$H_0^2 + H_2^2 = I_n$$

$$H_0 H_2 = -H_2 H_0 (U_3^* U_2)^* (U_1^* U_0)$$

$$H_1 = (U_3^* U_2) H_2 (U_3^* U_2)^* = (U_1^* U_0) H_2 (U_1^* U_0)^*$$

$$H_3 = (U_3^* U_2) H_0 (U_3^* U_2)^* = (U_1^* U_0) H_0 (U_1^* U_0)^*.$$

Now suppose $Q_j = U_j \Pi_j$ where U_j is unitary and Π_j is a projection operator for $j = 0, 1, 2, 3$. This is a special case of the polar decomposition of Q_j. Inserting these assumptions into equations (2.2) yields

$$\Pi_0 + \Pi_2 = I_n$$

$$\Pi_1 = (U_3^* U_2) \Pi_2 (U_3^* U_2)^* = (U_1^* U_0) \Pi_2 (U_1^* U_0)^*$$

$$\Pi_3 = (U_3^* U_2) \Pi_0 (U_3^* U_2)^* = (U_1^* U_0) \Pi_0 (U_1^* U_0)^*.$$

Consequently we also find

$$\Pi_0 \Pi_2 = \Pi_0 (I_n - \Pi_0) = 0_n.$$

Thus Π_0 and Π_2 decompose the Hilbert space $\mathbf{C}^n$ into two orthogonal subspaces. Additionally, Π_1 and Π_3 also decompose $\mathbf{C}^n$ into two orthogonal subspaces and is equivalent to the decomposition given by $\{\Pi_0, \Pi_2\}$ under

the unitary transformation $U_1^* U_0$ (or $U_3^* U_2$). This allows us to decompose U_q as a product of three matrices each of Schmidt rank no greater than 2

$$\begin{aligned} U &= U_0\Pi_0 \otimes |0\rangle\langle 0| + U_1\Pi_1 \otimes |0\rangle\langle 1| + U_2\Pi_2 \otimes |1\rangle\langle 0| + U_3\Pi_3 \otimes |1\rangle\langle 1| \\ &= U_0\Pi_0 \otimes |0\rangle\langle 0| + U_0\Pi_2 U_0^* U_1 \otimes |0\rangle\langle 1| + U_2\Pi_2 \otimes |1\rangle\langle 0| + U_2\Pi_0 U_2^* U_3 \otimes |1\rangle\langle 1| \\ &= (U_0 \otimes |0\rangle\langle 0| + U_2 \otimes |1\rangle\langle 1|) \\ &\quad \times (\Pi_0 \otimes I_2 + \Pi_2 \otimes U_{NOT}) \\ &\quad \times (I \otimes |0\rangle\langle 0| + [\Pi_0 U_2^* U_3 + \Pi_2 U_0^* U_1] \otimes |1\rangle\langle 1|). \end{aligned}$$

The matrix $\Pi_0 U_2^* U_3 + \Pi_2 U_0^* U_1$ is obviously unitary.
■

Example. Thus for example U_{SWAP} has $U_0 = U_3 = I_2$, $U_1 = U_2 = U_{NOT}$, $\Pi_0 = \Pi_1 = |0\rangle\langle 0|$ and $\Pi_2 = \Pi_3 = |1\rangle\langle 1|$. Consequently

$$\begin{aligned} U_{SWAP} &= (I_2 \otimes |0\rangle\langle 0| + U_{NOT} \otimes |1\rangle\langle 1|) \\ &\quad \times (|0\rangle\langle 0| \otimes I_2 + |1\rangle\langle 1| \otimes U_{NOT}) \\ &\quad \times (I \otimes |0\rangle\langle 0| + [|0\rangle\langle 0| U_{NOT} + |1\rangle\langle 1| U_{NOT}] \otimes |1\rangle\langle 1|) \\ &= U_{CNOT}(2,1) U_{CNOT}(1,2) U_{CNOT}(2,1). \end{aligned}$$

■

2.3 Entanglement for Pure States

Entanglement [HardyQC, Bruß2002, Preskill, Rieffel98, Steeb2000a, Steeb2000b] is the characteristic trait of quantum mechanics which enforces its entire departure from classical lines of thought. We consider entanglement of pure states. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two finite-dimensional Hilbert spaces. Thus a basic question in quantum computing [HardyQC] is as follows: given a normalized state $|z\rangle$ in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, can two normalized states $|x\rangle$ and $|y\rangle$ in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively be found such that

$$|x\rangle \otimes |y\rangle = |z\rangle.$$

In other words, what is the condition on $|z\rangle$ such that $|x\rangle$ and $|y\rangle$ exist? If

no such $|x\rangle$ and $|y\rangle$ exist then $|u\rangle$ is said to be *entangled*. A state which is not entangled is said to be *separable*.

Example. Let $\mathcal{H}_A = \mathcal{H}_B = \mathbf{C}^2$ with the standard basis

$$\{\, |0\rangle := (1,0)^T,\ |1\rangle := (0,1)^T \,\}.$$

Then the elements of the standard basis

$$|0\rangle := (1,0,0,0)^T = (1,0)^T \otimes (1,0)^T = |0\rangle \otimes |0\rangle$$

$$|1\rangle := (0,1,0,0)^T = (1,0)^T \otimes (0,1)^T = |0\rangle \otimes |1\rangle$$

$$|2\rangle := (0,0,1,0)^T = (0,1)^T \otimes (1,0)^T = |1\rangle \otimes |0\rangle$$

$$|3\rangle := (0,0,0,1)^T = (0,1)^T \otimes (0,1)^T = |1\rangle \otimes |1\rangle$$

in $\mathbf{C}^4$ are clearly not entangled, whereas the elements of the *Bell basis*

$$|\Phi^+\rangle := \frac{1}{\sqrt{2}}(1,0,0,1)^T, \quad |\Phi^-\rangle := \frac{1}{\sqrt{2}}(1,0,0,-1)^T,$$

$$|\Psi^+\rangle := \frac{1}{\sqrt{2}}(0,1,1,0)^T, \quad |\Psi^-\rangle := \frac{1}{\sqrt{2}}(0,1,-1,0)^T$$

cannot be expressed as product states and are therefore entangled pure states. The state $|\Phi^+\rangle$ is known as the *EPR state* named after Einstein, Podolsky and Rosen. Note that Einstein, Podolsky and Rosen did not consider the finite-dimensional case [Bohm] in their famous paper [EPR] but the infinite-dimensional case.

■

2.3.1 Separability Criteria

Clearly the Schmidt number of a pure quantum state is 1 if and only if the quantum state is separable. Thus the Schmidt number (equal to 1) gives a necessary and sufficient condition for separability. Below we consider a measure of entanglement, the so-called *entanglement of formation*, and then describe its relationship to separability criteria. We consider zero entanglement to mean separable and non-zero (positive) entanglement to be non-separable.

The measure for entanglement (of formation) for pure states $E(|u\rangle\langle u|)$ is defined as follows [Preskill, Bennett96]

$$E(|u\rangle\langle u|) := S_b(\rho_{\mathcal{H}_A}) = S_b(\rho_{\mathcal{H}_B})$$

where the *density operators* for the state $|u\rangle$ are defined as

$$\rho_{\mathcal{H}_A} := \mathrm{tr}_{\mathcal{H}_B}|u\rangle\langle u|, \qquad \rho_{\mathcal{H}_B} := \mathrm{tr}_{\mathcal{H}_A}|u\rangle\langle u|$$

and

$$S_b(\rho) := -\mathrm{tr}\rho \log_b \rho.$$

tr denotes the trace and $\mathrm{tr}_{\mathcal{H}_A}$ denotes the partial trace over $\mathcal{H}_A$. We use the base b for the logarithm $\log_b$. We have $0 \log_b 0 = 0$ and $1 \log_b 1 = 0$. In most cases we consider $b = 2$.

Thus $0 \leq E \leq \log_b \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\}$. If $E = 1$ we call the pure state maximally entangled. If $E = 0$, the pure state is not entangled. We note that

$$S_b(\rho) = -\sum_{j=1}^{k} \lambda_j \log_b \lambda_j$$

where $\{ \lambda_j, j = 1, \ldots, k \}$ are the eigenvalues of ρ and ρ is an operator on a k-dimensional Hilbert space.

Next we describe the conditions for separability of pure states and the relationship to the definition given above. The cases $m = n = 2$ and $m = n = 3$ have been discussed in [Steeb2000a, Steeb2000b]. In the following we consider the general case for finite-dimensional Hilbert spaces. We prove that if $|z\rangle$ can be written as the Kronecker product $|x\rangle \otimes |y\rangle$ then $E(|z\rangle\langle z|) = 0$ and conversely if $E(|z\rangle\langle z|) = 0$ it follows that $|z\rangle$ can be written as $|x\rangle \otimes |y\rangle$.

Conditions for Separability

Let $m := \dim(\mathcal{H}_A)$, $n := \dim(\mathcal{H}_B)$,

$$\{|j\rangle_{\mathcal{H}_A}, j = 0, 1, \ldots, m-1 \}$$

be an orthonormal basis for $\mathcal{H}_A$ and

$$\{|j\rangle_{\mathcal{H}_B}, j = 0, 1, \ldots, n-1\}$$

be an orthonormal basis for $\mathcal{H}_B$. Thus

$$\{|j\rangle_{\mathcal{H}_A} \otimes |k\rangle_{\mathcal{H}_B}, j = 0, 1, \ldots, m-1, k = 0, 1, \ldots, n-1\}$$

is an orthonormal basis for $\mathcal{H}_A \otimes \mathcal{H}_B$. We consider arbitrary normalized states $|x\rangle$, $|y\rangle$ and $|z\rangle$ in the Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_A \otimes \mathcal{H}_B$ respectively. We can identify these states with the vectors $(x_0, x_1, \ldots, x_{m-1})^T \in \mathbf{C}^m$, $(y_0, y_1, \ldots, y_{m-1})^T \in \mathbf{C}^n$ and $(z_0, z_1, \ldots, z_{nm-1})^T \in \mathbf{C}^{mn}$ as follows

$$|x\rangle := \sum_{j=0}^{m-1} x_j |j\rangle_{\mathcal{H}_A} \quad |y\rangle := \sum_{j=0}^{n-1} y_j |j\rangle_{\mathcal{H}_B}$$

$$|z\rangle := \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} z_{in+j} |i\rangle_{\mathcal{H}_A} \otimes |j\rangle_{\mathcal{H}_B}$$

where

$$\sum_{j=0}^{m-1} |x_j|^2 = 1, \quad \sum_{j=0}^{n-1} |y_j|^2 = 1, \quad \sum_{j=0}^{nm-1} |z_j|^2 = 1.$$

Thus we see that the tensor product is equivalent to considering the Kronecker product [SteebMC]. To ensure that $|z\rangle$ is the product state of $|x\rangle$ and $|y\rangle$ we must have

$$\forall i \in \{0, 1, \ldots, m-1\},\ j \in \{0, 1, \ldots, n-1\} \qquad x_i y_j = z_{in+j}.$$

Thus we have the following conditions for separability

$$\begin{aligned} &x_i y_j x_k y_l = z_{in+j} z_{kn+l} = z_{in+l} z_{kn+j} \\ &i = 0, 1, \ldots, m-1, \quad j = 0, 1, \ldots, n-1, \\ &k = i+1, i+2, \ldots, m-1, \quad l = j+1, j+2, \ldots, n-1. \end{aligned} \tag{2.3}$$

There are

$$\left(\sum_{i=0}^{m-1} (m-1-i)\right)\left(\sum_{j=0}^{n-1} (n-1-j)\right) = \frac{mn(m-1)(n-1)}{4}$$

such equations. Supposing equations (2.3) hold then a "factorization" is given by

$$x_j = \left(\sum_{k=0}^{n-1} |z_{jn+k}|^2\right)^{\frac{1}{2}} e^{i\alpha_j} \qquad \alpha_j{=}(1-\delta_{j,0})(\arg(z_{jn}) - \beta_0)$$

$$y_j = \left(\sum_{k=0}^{m-1} |z_{kn+j}|^2\right)^{\frac{1}{2}} e^{i\beta_j} \qquad \beta_j{=}\arg(z_j).$$

Theorem. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two finite-dimensional Hilbert spaces, and let $|z\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. Then $E(|z\rangle\langle z|) = 0$ iff the conditions (2.3) hold.

Proof. We will omit the subscripts $\mathcal{H}_A$ and $\mathcal{H}_B$ since the Hilbert space will be clear from the context. Since

$$|z\rangle\langle z| = \sum_{i,k=0}^{m-1} \sum_{j,l=0}^{n-1} z_{in+j}\overline{z_{kn+l}}(|i\rangle\langle k|) \otimes (|j\rangle\langle l|),$$

taking the partial trace over $\mathcal{H}_A$ yields the density operator

$$\rho_{\mathcal{H}_B} = \mathrm{tr}_{\mathcal{H}_A}|z\rangle\langle z| = \sum_{i=0}^{m-1} (\langle i| \otimes I_n)|z\rangle\langle z|(|i\rangle \otimes I_n)$$

$$= \sum_{i=0}^{m-1} \sum_{j,l=0}^{n-1} z_{in+j}\overline{z_{in+l}}|j\rangle\langle l|$$

and taking the partial trace over $\mathcal{H}_B$ yields the density operator

$$\rho_{\mathcal{H}_A} = \mathrm{tr}_{\mathcal{H}_B}|z\rangle\langle z| = \sum_{i=0}^{n-1} (I_m \otimes \langle i|)|z\rangle\langle z|(I_m \otimes |i\rangle)$$

$$= \sum_{j=0}^{n-1} \sum_{i,k=0}^{m-1} z_{in+j}\overline{z_{kn+j}}|i\rangle\langle k|$$

where I_m is the $m \times m$ unit matrix. From the condition (2.3) we find that

$$\overline{z_k} \cdot \overline{z_{in+l}} = \overline{z_l} \cdot \overline{z_{in+k}}, \qquad \overline{z_{ln}} \cdot \overline{z_{kn+j}}{=}\overline{z_{kn}} \cdot \overline{z_{ln+j}}$$

so that

$$\overline{z_k} \sum_{i=0}^{m-1} z_{in+j}\overline{z_{in+l}} = \overline{z_l} \sum_{i=0}^{m-1} z_{in+j}\overline{z_{in+k}}$$

and

$$\overline{z_{ln}} \sum_{j=0}^{n-1} z_{in+j} \overline{z_{kn+j}} = \overline{z_{in}} \sum_{j=0}^{n-1} z_{in+j} \overline{z_{ln+j}}.$$

Consequently we find $\text{rank}(\rho_{\mathcal{H}_A}) = \text{rank}(\rho_{\mathcal{H}_B}) = 1$. Thus exactly one of the eigenvalues of $\rho_{\mathcal{H}_A}$ and exactly one of the eigenvalues of $\rho_{\mathcal{H}_B}$ are non-zero. Since $\rho_{\mathcal{H}_A}$ and $\rho_{\mathcal{H}_B}$ are positive operators with unit trace, they have eigenvalues $\{ 0, 1 \}$ with appropriate multiplicities. Thus $S_b(\rho_{\mathcal{H}_A}) = S_b(\rho_{\mathcal{H}_B}) = 0$.

For the converse, suppose $S_b(\rho_{\mathcal{H}_A}) = S_b(\rho_{\mathcal{H}_B}) = 0$. We will consider $S_b(\rho_{\mathcal{H}_A}) = 0$. Since $\rho_{\mathcal{H}_A}$ is a positive operator with unit trace the eigenvalues

$$\{ \lambda_j , j = 1, \ldots, m \}$$

must satisfy $0 \leq \lambda_j \leq 1$ and $\sum_{j=1}^{m} \lambda_j = 1$. Since $-\lambda_j \log_b \lambda_j \geq 0$ we find that all the eigenvalues except one are zero. The eigenspace corresponding to the eigenvalue has dimension $m - 1$, i.e. there exist $m - 1$ linearly independent vectors $|a_j\rangle$ $(j = 1, 2, \ldots, m)$ such that $\rho_{\mathcal{H}_A}|\mathbf{a}\rangle = 0$. Thus each of the vectors

$$(\langle i|\rho_{\mathcal{H}_A})^* = \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \overline{z_{in+j}} z_{kn+j} |k\rangle$$

are orthogonal to all $m - 1$ of the $|a_j\rangle$, and as a consequence are all scalar multiples of each other. Thus we find

$$\forall i, k, l \in \{0, 1, \ldots, m-1\} \qquad \sum_{j=0}^{n-1} z_{in+j} \overline{z_{kn+j}} = \alpha_{il} \sum_{j=0}^{n-1} z_{ln+j} \overline{z_{kn+j}}.$$

$$\forall i, k, l \in \{0, 1, \ldots, m-1\} \qquad \sum_{j=0}^{n-1} (z_{in+j} - \alpha_{il} z_{ln+j}) \overline{z_{kn+j}} = 0.$$

This is the inner product [SteebQM] between $\sum_{j=0}^{n-1} (z_{in+j} - \alpha_{il} z_{ln+j}) |j\rangle$ and $\sum_{j=0}^{n-1} z_{kn+j} |j\rangle$. Since $\dim(\mathcal{H}_A) = m$, the vectors (for all i, k, l) cannot all be orthogonal and non-zero. Suppose

$$|c_{il}\rangle := \sum_{j=0}^{n-1} (z_{in+j} - \alpha_{il} z_{ln+j}) |j\rangle$$

is non-zero. The vectors $\sum_{j=0}^{n-1} z_{in+j} |j\rangle$ and $\sum_{j=0}^{n-1} z_{ln+j} |j\rangle$ define at most a two-

dimensional vector space. If $z_{in+j} = z_{ln+j} = 0$ for all j then $z_{in+j} = \alpha_{il} z_{ln+j}$ trivially, where we choose $\alpha_{il} = 1$. If this is not the case the vector $|c_{il}\rangle$ is a linear combination of these two vectors and yet is orthogonal to both – a contradiction. Thus we conclude

$$\forall j \in \{0, 1, \ldots, n-1\} \qquad z_{in+j} = \alpha_{ik} z_{kn+j}.$$

If $\alpha_{ik} = 0$, then $z_{in+j} = z_{kn+j} = 0$ and $z_{in+j} z_{kn+l} = 0 = z_{kn+j} z_{in+l}$. Thus we consider the case $\alpha_{ik} \neq 0$. In this case we have $\alpha_{ik} = \alpha_{ki}^{-1}$. Now we consider

$$z_{in+j} z_{kn+l} = \alpha_{ik} z_{kn+j} \alpha_{ki} z_{in+l} = z_{in+l} z_{kn+j}.$$

Thus the two conditions for separability are equivalent.
■

2.3.2 Example: Hubbard Model

As an example we consider the two-point Hubbard model [Steeb2001a]. We wish to know when an entangled state results for given parameters t and U as well as the time τ required for the system to evolve to these states.

The two-point Hubbard model with cyclic boundary conditions is given by

$$\hat{H} = t(c_{1\uparrow}^\dagger c_{2\uparrow} + c_{1\downarrow}^\dagger c_{2\downarrow} + c_{2\uparrow}^\dagger c_{1\uparrow} + c_{2\downarrow}^\dagger c_{1\downarrow}) + U(n_{1\uparrow} n_{1\downarrow} + n_{2\uparrow} n_{2\downarrow})$$

where

$$n_{j\uparrow} := c_{j\uparrow}^\dagger c_{j\uparrow}, \quad n_{j\downarrow} := c_{j\downarrow}^\dagger c_{j\downarrow}, \qquad j = 1, 2.$$

The Fermi operators $c_{j\uparrow}^\dagger$,$c_{j\downarrow}^\dagger$,$c_{j\uparrow}$,$c_{j\downarrow}$ obey the anti-commutation relations

$$[c_{j,\sigma}^\dagger, c_{k,\sigma'}]_+ = \delta_{\sigma\sigma'} \delta_{jk} I, \quad [c_{j,\sigma}^\dagger, c_{k,\sigma'}^\dagger]_+ = [c_{j,\sigma}, c_{k,\sigma'}]_+ = 0.$$

$\hat{H}$ commutes with the total number operator $\hat{N}$, and the total spin operator $\hat{S}_z$ in the z direction

$$\hat{N} := \sum_{j=1}^{2} (c_{j\uparrow}^\dagger c_{j\uparrow} + c_{j\downarrow}^\dagger c_{j\downarrow}) \qquad \hat{S}_z := \frac{1}{2} \sum_{j=1}^{2} (c_{j\uparrow}^\dagger c_{j\uparrow} - c_{j\downarrow}^\dagger c_{j\downarrow}).$$

We consider the subspace with two electrons, $N = 2$ and $S_z = 0$. A basis for

2 particles with total spin 0 is

$$|s_1\rangle := c^\dagger_{1\uparrow}c^\dagger_{1\downarrow}|0\rangle, \quad |s_2\rangle := c^\dagger_{1\uparrow}c^\dagger_{2\downarrow}|0\rangle, \quad |s_3\rangle := c^\dagger_{2\uparrow}c^\dagger_{1\downarrow}|0\rangle, \quad |s_4\rangle := c^\dagger_{2\uparrow}c^\dagger_{2\downarrow}|0\rangle$$

where $\langle 0|0\rangle = 1$. Applying $\hat{H}$ to the basis gives

$$\hat{H}|s_1\rangle = t|s_2\rangle + t|s_3\rangle + U|s_1\rangle$$

$$\hat{H}|s_2\rangle = t|s_1\rangle + t|s_4\rangle$$

$$\hat{H}|s_3\rangle = t|s_1\rangle + t|s_4\rangle$$

$$\hat{H}|s_4\rangle = t|s_2\rangle + t|s_3\rangle + U|s_4\rangle.$$

Identifying $|s_i\rangle$ with elements $\mathbf{e}_i$ of the standard basis in $\mathbf{C}^4$ yields the matrix representation of $\hat{H}$

$$\hat{H} = \begin{pmatrix} U & t & t & 0 \\ t & 0 & 0 & t \\ t & 0 & 0 & t \\ 0 & t & t & U \end{pmatrix}.$$

Suppose a Hamilton operator $\hat{K}$ can be written as $\hat{K} = A_1 \otimes I_2 + I_2 \otimes A_2$ where $A_1, A_2 \in M^2$ and I_2 is the 2×2 identity matrix. Then we have [SteebMC]

$$\begin{aligned} \exp(-i\hat{K}\tau/\hbar) &= \exp\left(-i\frac{\tau}{\hbar}A_1 \otimes I_2 - i\frac{\tau}{\hbar}I_2 \otimes A_2\right) \\ &= \exp\left(-i\frac{\tau}{\hbar}A_1\right) \otimes \exp\left(-i\frac{\tau}{\hbar}A_2\right). \end{aligned}$$

In this case separable states remain separable under time evolution in the model, and entangled states remain entangled under time evolution in the model. For the matrix representation of $\hat{H}$, however we have

$$\hat{H} = tV_{NOT} \otimes I_2 + tI_2 \otimes V_{NOT} + \mathrm{diag}(U, 0, 0, U), \qquad V_{NOT} := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The diagonal matrix $\mathrm{diag}(U, 0, 0, U)$ cannot be written in the form $A_1 \otimes I_2 + I_2 \otimes A_2$. Thus we conclude that almost all initial separable states evolve into entangled states under the time evolution of the model.

The eigenvalues of $\hat{H}$ are $E_1 = 0$, $E_2 = U$ and $E_{3,4} = \frac{1}{2}(U \pm \sqrt{U^2 + 16t^2})$.

For all U and t the Hamilton operator $\hat{H}$ can be written as (spectral representation)

$$\hat{H} = \sum_{j=1}^{4} E_j |x_j\rangle\langle x_j|$$

where E_j are the eigenvalues and $|x_j\rangle$ are the corresponding orthonormal eigenstates of $\hat{H}$.

To find the time evolution for the two-point Hubbard model we solve the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial\tau} = \hat{H}\psi \quad \Rightarrow \quad |\psi(\tau)\rangle = U_{\hat{H}}(\tau)|\psi(0)\rangle$$

where $|\psi(0)\rangle$ is the initial state and the unitary operator $U_{\hat{H}}(\tau)$ is given by

$$U_{\hat{H}}(\tau) := e^{-i\hat{H}\tau/\hbar}\,.$$

For $t = 0$, the unitary transformation $U_{\hat{H}}(\tau)$ implements a phase change.

For the initial state $|s_1\rangle$ we find the condition for separability for $U_{\hat{H}}(\tau)|s_1\rangle$ is

$$\begin{aligned} &UE_3m^4\left(E_4\exp\left(\frac{i\tau}{2\hbar}\left(E_4-E_3\right)\right)-E_3\exp\left(\frac{i\tau}{2\hbar}\left(E_3-E_4\right)\right)\right)^2 \\ &\quad = \frac{1}{4}\exp\left(-i\frac{\tau}{\hbar}U\right) \end{aligned} \tag{2.4}$$

where

$$m := \frac{1}{\sqrt{2}}\left((E_3-U)^2+4t^2\right)^{-\frac{1}{2}}\,.$$

To satisfy the imaginary part of equation (2.4) we find the condition

$$E_4\sin\left(\frac{\tau}{\hbar}E_4\right) - E_3\sin\left(\frac{\tau}{\hbar}E_3\right) = 0.$$

Thus we find $U_{\hat{H}}(\tau)|s_1\rangle$ is entangled when the condition is not satisfied.

For the initial state $|s_2(\tau)\rangle$ we find the condition for separability for $U_{\hat{H}}(\tau)|s_2\rangle$

is

$$\begin{aligned} & UE_4^{-1}m^4\left(E_4^2\exp\left(\frac{i\tau}{2\hbar}\left(E_4-E_3\right)\right)-E_3^2\exp\left(\frac{i\tau}{2\hbar}\left(E_3-E_4\right)\right)\right)^2 \\ & \quad =\frac{1}{4}\exp\left(i\frac{\tau}{\hbar}U\right). \end{aligned} \tag{2.5}$$

To satisfy the imaginary part of equation (2.5) we find the condition

$$E_4^2\sin\left(\frac{\tau}{\hbar}E_3\right)-E_3^2\sin\left(\frac{\tau}{\hbar}E_4\right)=0.$$

Thus we find $U_{\hat{H}}(\tau)|s_2\rangle$ is entangled when the condition is not satisfied.

For $U_{\hat{H}}(\tau)|s_3\rangle$ (respectively $U_{\hat{H}}(\tau)|s_4\rangle$) we find the condition for separability is identical to that for $U_{\hat{H}}(\tau)|s_2\rangle$ (respectively $U_{\hat{H}}(\tau)|s_1\rangle$).

Next we determine the conditions that the states

$$|\Phi^+\rangle:=\frac{1}{\sqrt{2}}(|s_1\rangle+|s_4\rangle),\qquad |\Phi^-\rangle:=\frac{1}{\sqrt{2}}(|s_1\rangle-|s_4\rangle)$$

$$|\Psi^+\rangle:=\frac{1}{\sqrt{2}}(|s_2\rangle+|s_3\rangle),\qquad |\Psi^-\rangle:=\frac{1}{\sqrt{2}}(|s_2\rangle-|s_3\rangle)$$

are entangled under time evolution of the model. They are maximally entangled states. These are the *Bell states* and form a basis in $\mathbf{C}^4$.

For $U_{\hat{H}}(\tau)|\Phi^+\rangle$ we find the condition for separability

$$2UE_3m^4\left(E_4\exp\left(\frac{i\tau}{2\hbar}\left(E_4-E_3\right)\right)-E_3\exp\left(\frac{i\tau}{2\hbar}\left(E_3-E_4\right)\right)\right)^2=0.$$

Thus when $U_{\hat{H}}(\tau)|\Phi^+\rangle$ is not entangled $U_{\hat{H}}(\tau)|s_1\rangle$ and $U_{\hat{H}}(\tau)|s_4\rangle$ are entangled and vice versa. For $U\neq 0$ the condition becomes

$$E_4=E_3\exp\left(i\frac{\tau}{\hbar}\left(E_3-E_4\right)\right)$$

which has no real solutions for τ. Thus $U_{\hat{H}}(\tau)|\Phi^+\rangle$ is entangled for all τ.

For $U_{\hat{H}}(\tau)|\Phi^-\rangle$ we find the condition for separability

$$\exp\left(-2i\frac{\tau}{\hbar}U\right) = 0.$$

Of course, this equation cannot be satisfied. Thus $U_{\hat{H}}(\tau)|\Phi^-\rangle$ is entangled for all τ.

For $U_{\hat{H}}(\tau)|\Psi^+\rangle$ we find the condition for separability

$$-2UE_4^{-1}m^4\left(E_4^2\exp\left(\frac{i\tau}{2\hbar}\left(E_4-E_3\right)\right)-E_3^2\exp\left(\frac{i\tau}{2\hbar}\left(E_3-E_4\right)\right)\right)^2=0.$$

Thus when $U_{\hat{H}}(\tau)|\Psi^+\rangle$ is not entangled $U_{\hat{H}}(\tau)|s_2\rangle$ and $U_{\hat{H}}(\tau)|s_3\rangle$ are entangled and vice versa. We find that $U_{\hat{H}}(\tau)|\Psi^+\rangle$ is entangled for all τ.

For $U_{\hat{H}}(\tau)|\Psi^-\rangle$ we find the condition for separability, $\frac{1}{2} = 0$, cannot be satisfied. Thus $U_{\hat{H}}(\tau)|\Psi^-\rangle$ is entangled for all τ.

The behaviour described above can be understood if we realize that the Hubbard model admits a discrete symmetry under the change $1 \to 2$, $2 \to 1$. Thus we have a finite group with two elements. We obtain two irreducible representations [SteebPB2]. The group-theoretical reduction leads to the invariant subspaces

$$S_1 = \left\{|\psi_1\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle + c_{2\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle),\quad |\psi_2\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle + c_{2\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle)\right\}$$

$$S_2 = \left\{|\psi_3\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle - c_{2\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle),\quad |\psi_4\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle - c_{2\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle)\right\}.$$

These four states can be considered as the Bell states. In the Bell basis the matrix representation of the Hubbard model is given by (where we use the ordering $|\Phi^+\rangle$, $|\Psi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^-\rangle$)

$$\begin{pmatrix} U & 2t & 0 & 0 \\ 2t & 0 & 0 & 0 \\ 0 & 0 & U & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} U & 2t \\ 2t & 0 \end{pmatrix} \oplus \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix}$$

where $\oplus$ denotes the direct sum. Thus the ground state is not a Bell state, but it is also not a separable state. We used a program to perform the symbolic calculations and most of the simplification of the separability conditions given

above. The program makes use of SymbolicC++ [SymbolicC++] to do the calculations and simplification. The Hubbard model can also be considered in Bloch representation. The entanglement can then be considered in the momentum space.

2.3.3 Examples: Fermi-Bose Systems

Entanglement has been studied in detail for finite-dimensional quantum systems and to a lesser extent for infinite-dimensional quantum systems (see [SteebPQC, HardyQC] and references therein). Chi and Lee [Chi2003] have described a class of 2 parameter density operators in a product space of 2-dimensional by n-dimensional Hilbert space for which they could obtain a lower bound and tight upper bound on the entanglement of formation. Wang and Sanders [Wang2001] discussed the generation of multipartite entangled coherent states and provide further references on entangled coherent states. Keyl et al. [Keyl2003] discussed the entanglement of infinite-dimensional systems and quantum states exhibiting infinite entanglement. In [Hardy2004] and section 2.4.1 we described entanglement for the Hubbard model, and discussed entanglement and phonon coupling. Hines et al. [Hines2003] considered the entanglement of two-mode Bose-Einstein condensates in the Bose-Hubbard model and macroscopic quantum superposition (Schrödinger cat like) states.

In this section we consider entanglement of a Fermi system with a Bose system in terms of coherent states and macroscopic quantum superposition states [Hardy2004]. For example we can consider entanglement of the two-dimensional Hubbard model with phonon coupling [Steeb86]

$$\hat{H} = t \sum_{\sigma\in\{\uparrow,\downarrow\}} (c_{1\sigma}^{\dagger} c_{2\sigma} + c_{2\sigma}^{\dagger} c_{1\sigma}) \otimes I_b + U(n_{1\uparrow} n_{1\downarrow} + n_{2\uparrow} n_{2\downarrow}) \otimes I_b$$

$$+ I_c \otimes \omega b^{\dagger} b + k \sum_{\sigma\in\{\uparrow,\downarrow\}} (c_{1\sigma}^{\dagger} c_{2\sigma} + c_{2\sigma}^{\dagger} c_{1\sigma}) \otimes (b^{\dagger} + b)$$

where b, $b^{\dagger}$ are the Bose annihilation and creation operators for the vibrational mode, respectively. This means we have $[b, b^{\dagger}] = I_b$. In this case we have a product space of the states given above and the number states $|n\rangle = (n!)^{-1/2}(b^{\dagger})^n|0\rangle$ with $b|0\rangle = 0$, $\langle 0|0\rangle = 1$ and $n = 0, 1, 2, \ldots$. For $S_z = 0$ the basis in the product space is now given as

$$S_1' = \{\, |\psi_1\rangle \otimes |n\rangle,\ |\psi_2\rangle \otimes |n\rangle, \quad n = 0, 1, 2, \dots \}$$
$$S_2' = \{\, |\psi_3\rangle \otimes |n\rangle, \quad n = 0, 1, 2, \dots \}$$
$$S_3' = \{\, |\psi_4\rangle \otimes |n\rangle, \quad n = 0, 1, 2, \dots \}$$

where $\langle n|m\rangle = \delta_{nm}$ and

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle + c_{2\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle), \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle + c_{2\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle)$$
$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle - c_{2\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle), \quad |\psi_4\rangle = \frac{1}{\sqrt{2}}(c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger|0\rangle - c_{2\downarrow}^\dagger c_{1\uparrow}^\dagger|0\rangle)\,.$$

For the subspace S_1' we obtain

$$\begin{aligned}\hat{H}|\psi_1\rangle \otimes |n\rangle &= 2t|\psi_2\rangle \otimes |n\rangle + (U + n\omega)|\psi_1\rangle \otimes |n\rangle \\ &\quad + 2k(n+1)^{1/2}|\psi_2\rangle \otimes |n+1\rangle + 2kn^{1/2}|\psi_2\rangle \otimes |n-1\rangle \\ \hat{H}|\psi_2\rangle \otimes |n\rangle &= 2t|\psi_1\rangle \otimes |n\rangle + n\omega|\psi_2\rangle \otimes |n\rangle \\ &\quad + 2k(n+1)^{1/2}|\psi_1\rangle \otimes |n+1\rangle + 2kn^{1/2}|\psi_1\rangle \otimes |n-1\rangle\,.\end{aligned}$$

In the subspace S_2' we have the energy levels $U + n\omega$ $(n = 0, 1, 2, \dots)$ and in the subspace S_3' we find $n\omega$ $(n = 0, 1, 2, \dots)$. In both cases the eigenvalues do not depend on k. This allows to study entanglement between Bose and Fermi states. For example, we could consider entangled states such as

$$\frac{1}{\sqrt{2}}(|\psi_1\rangle \otimes |0\rangle + |\psi_2\rangle \otimes |1\rangle).$$

First we give an explicit expression for the entanglement of states in a one-Fermi system coupled to an arbitrary quantum system. Then we proceed to elaborate on this expression for specific cases in a one-Bose system, providing explicit formulas for the entanglement for these special cases. The Hilbert space for the one-Fermi system is $\mathbf{C}^2$. For the one-Bose system we have the Hilbert space $l_2(\mathbf{N})$. Thus we work in the product Hilbert space $\mathbf{C}^2 \otimes l_2(\mathbf{N})$.

Entanglement

Consider the product Hilbert space $\mathbf{C}^2 \otimes \mathcal{H}$ where $\mathcal{H}$ denotes an arbitrary Hilbert space. For the one-Bose system we would set $\mathcal{H} = l_2(\mathbf{N})$. An arbitrary pure state in this product Hilbert space can be written as

$$|\psi\rangle := |0\rangle \otimes |\phi_0\rangle + |1\rangle \otimes |\phi_1\rangle$$

where $|\phi_0\rangle, |\phi_1\rangle \in \mathcal{H}$ and $\{\, |0\rangle, |1\rangle \,\}$ forms an orthonormal basis in $\mathbf{C}^2$. The condition for the state $|\psi\rangle$ to be normalized, i.e. $\langle\psi|\psi\rangle = 1$, leads to the constraint

$$\langle\phi_0|\phi_0\rangle + \langle\phi_1|\phi_1\rangle = 1. \tag{2.6}$$

If we assume that $|\phi_0\rangle$ and $|\phi_1\rangle$ have identical norms, then $|\psi\rangle$ takes the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |\varphi_0\rangle + |1\rangle \otimes |\varphi_1\rangle)$$

where $|\phi_0\rangle = \frac{1}{\sqrt{2}}|\varphi_0\rangle$, $|\phi_1\rangle = \frac{1}{\sqrt{2}}|\varphi_1\rangle$ and $|\varphi_0\rangle, |\varphi_1\rangle$ are normalized. Defining the reduced density matrices (using the partial trace)

$$\rho_1 := \mathrm{tr}_{\mathbf{C}^2}(|\psi\rangle\langle\psi|), \qquad \rho_2 := \mathrm{tr}_{\mathcal{H}}(|\psi\rangle\langle\psi|)$$

the entanglement of $|\psi\rangle$ is given by

$$E(|\psi\rangle) := -\mathrm{tr}(\rho_1 \log_2 \rho_1) = -\mathrm{tr}(\rho_2 \log_2 \rho_2).$$

Straightforward calculation yields

$$\rho_1 = |\phi_0\rangle\langle\phi_0| + |\phi_1\rangle\langle\phi_1|$$

and

$$\rho_2 = \langle\phi_0|\phi_0\rangle|0\rangle\langle 0| + \langle\phi_1|\phi_0\rangle|0\rangle\langle 1| + \langle\phi_0|\phi_1\rangle|1\rangle\langle 0| + \langle\phi_1|\phi_1\rangle|1\rangle\langle 1|.$$

Applying the constraint (2.6) we find that the non-zero eigenvalues of ρ_1 and ρ_2 are given by

$$\lambda(\langle\phi_0|\phi_0\rangle, |\langle\phi_0|\phi_1\rangle|^2) := \frac{1}{2}\left(1 + \sqrt{(1 - 2\langle\phi_0|\phi_0\rangle)^2 + 4|\langle\phi_0|\phi_1\rangle|^2}\right)$$

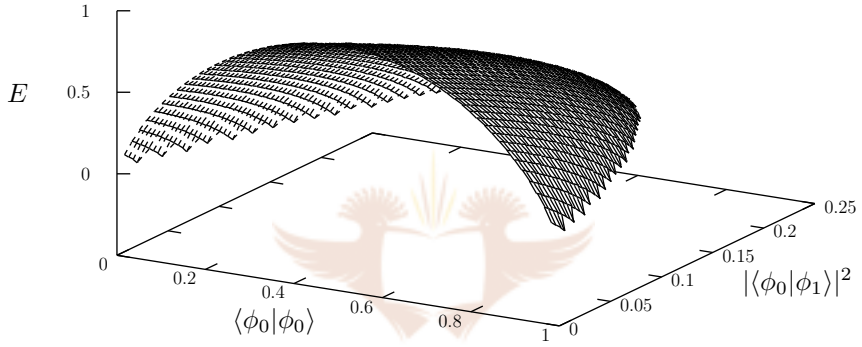
and $1 - \lambda$. Thus

$$E(|\psi\rangle) = -\lambda \log_2 \lambda - (1-\lambda)\log_2(1-\lambda).$$

The entanglement is described exclusively by $\langle\phi_0|\phi_0\rangle$ and $|\langle\phi_0|\phi_1\rangle|^2$. Furthermore we have the inequality

$$|\langle\phi_0|\phi_1\rangle|^2 \leq \langle\phi_0|\phi_0\rangle - \langle\phi_0|\phi_0\rangle^2 \leq \frac{1}{4}. \tag{2.7}$$

The figure below shows the entanglement E as a function of $\langle\phi_0|\phi_0\rangle$ and $|\langle\phi_0|\phi_1\rangle|^2$. The maximum value of E is achieved for $\langle\phi_0|\phi_1\rangle = 0$ and $\langle\phi_0|\phi_0\rangle = \frac{1}{2}$.



Fock States

Let $\{\,|n\rangle : n = 0, 1, 2, \ldots\}$ be the *number states* (*Fock states*) [SteebPQC]. The scalar product between the number states $|n\rangle$ and $|m\rangle$ is given by $\langle n|m\rangle = \delta_{nm}$. For the number states we use $|\phi_0\rangle = c_0|m\rangle$ and $|\phi_1\rangle = c_1|n\rangle$, where $c_0, c_1 \in \mathbf{C}$. In other words

$$|\psi\rangle = c_0|0\rangle \otimes |m\rangle + c_1|1\rangle \otimes |n\rangle.$$

The condition (2.6) leads to $|c_0|^2 + |c_1|^2 = 1$. Since the scalar products are given by $\langle\phi_0|\phi_0\rangle = |c_0|^2$ and

$$|\langle\phi_0|\phi_1\rangle|^2 = |c_0|^2(1-|c_0|^2)\delta_{mn},$$

we have

$$E(|\psi\rangle) = -|c_0|^2 \log_2 |c_0|^2 - (1-|c_0|^2)\log_2(1-|c_0|^2).$$

For $|c_0|^2 = |c_1|^2 = \frac{1}{2}$ we obtain maximum entanglement. This is analogous to the entanglement in $\mathbf{C}^2 \otimes \mathbf{C}^2$ since the Hilbert space spanned by $\{ |m\rangle, |n\rangle \}$, for fixed m and n, is isomorphic to $\mathbf{C}^2$.

Coherent States

The *coherent states* [SteebPQC, Hines2003] are defined by

$$b|\beta\rangle = \beta|\beta\rangle$$

where b is the Bose annihilation operator and $\beta \in \mathbf{C}$. The scalar product between the coherent states $|\alpha\rangle$ and $|\beta\rangle$ is given by

$$\langle\alpha|\beta\rangle = e^{-\frac{1}{2}(|\alpha|^2+|\beta|^2)+\overline{\alpha}\beta}.$$

It follows that

$$|\langle\alpha|\beta\rangle|^2 = e^{-|\alpha-\beta|^2}.$$

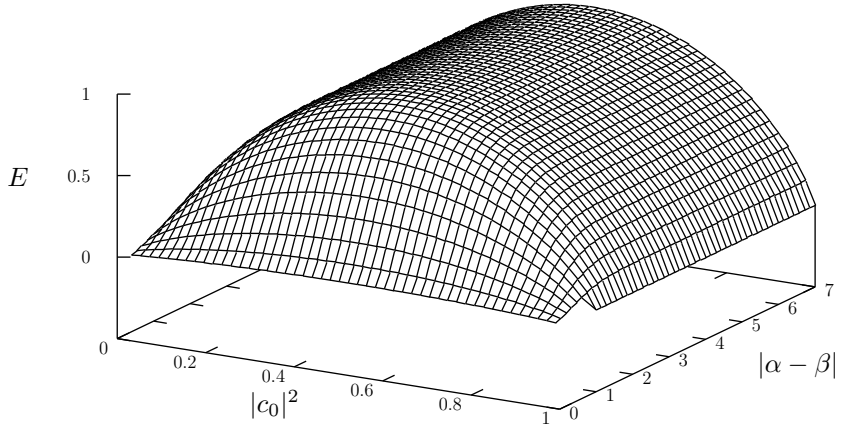
For the coherent states we use $|\phi_0\rangle = c_0|\alpha\rangle$ and $|\phi_1\rangle = c_1|\beta\rangle$ where $c_0, c_1 \in \mathbf{C}$. Thus

$$|\psi\rangle = c_0|0\rangle \otimes |\alpha\rangle + c_1|1\rangle \otimes |\beta\rangle.$$

The condition (2.6) leads to $|c_0|^2 + |c_1|^2 = 1$. We find that $\langle\phi_0|\phi_0\rangle = |c_0|^2$ and

$$|\langle\phi_0|\phi_1\rangle|^2 = |c_0|^2(1 - |c_0|)^2 e^{-|\alpha-\beta|^2}.$$

The behaviour with respect to the coherent states tends to the behavior of the number states as $|\alpha - \beta| \to \infty$. The maximal entanglement is reached asymptotically as $|\alpha - \beta| \to \infty$ when $|c_0|^2 = \frac{1}{2}$. The following figure shows the entanglement E as a function of $|c_0|^2$ and $|\alpha - \beta|$.



Macroscopic Quantum Superposition States

For *macroscopic quantum superposition states* (defined in [Itano97] as *Schrödinger cat states*) we use $|\phi_0\rangle = c_0|\beta\rangle$ and $|\phi_1\rangle = c_1|-\beta\rangle$ where $c_0, c_1 \in \mathbf{C}$ and $|\beta\rangle$ and $|-\beta\rangle$ are coherent states. This means that we consider a special case of the coherent states discussed above with $\alpha \leftarrow \beta$ and $\beta \leftarrow -\beta$.

Consider next the superposition for *macroscopic quantum superposition states*, $|\phi_0\rangle = c_0(|\alpha\rangle + |-\alpha\rangle)$ and $|\phi_1\rangle = c_1(|\beta\rangle + |-\beta\rangle)$ where $c_0, c_1 \in \mathbf{C}$, i.e.,

$$|\psi\rangle = c_0|0\rangle \otimes (|\alpha\rangle + |-\alpha\rangle) + c_1|1\rangle \otimes (|\beta\rangle + |-\beta\rangle).$$

The condition (2.6) for this case gives

$$2|c_0|^2(1 + e^{-2|\alpha|^2}) + 2|c_1|^2(1 + e^{-2|\beta|^2}) = 1. \tag{2.8}$$

Consequently $\langle\phi_0|\phi_0\rangle = 2|c_0|^2(1 + e^{-2|\alpha|^2})$, and using (2.8)

$$|\langle\phi_0|\phi_1\rangle|^2 = \langle\phi_0|\phi_0\rangle(1 - \langle\phi_0|\phi_0\rangle)\frac{(e^{-\frac{1}{2}|\alpha-\beta|^2} + e^{-\frac{1}{2}|\alpha+\beta|^2})^2}{(1 + e^{-2|\alpha|^2})(1 + e^{-2|\beta|^2})}.$$

It is convenient to define the real valued quantities

$$p_{00} := \langle\phi_0|\phi_0\rangle = 2|c_0|^2(1 + e^{-2|\alpha|^2})$$

$$p_{01} := \frac{(e^{-\frac{1}{2}|\alpha-\beta|^2} + e^{-\frac{1}{2}|\alpha+\beta|^2})^2}{(1 + e^{-2|\alpha|^2})(1 + e^{-2|\beta|^2})}.$$

Thus we obtain $|\langle\phi_0|\phi_1\rangle|^2 = p_{00}(1-p_{00})p_{01}$. The maximum entanglement occurs when $\langle\phi_0|\phi_1\rangle = 0$ and $\langle\phi_0|\phi_0\rangle = \frac{1}{2}$. Since the above equation implies $\langle\phi_0|\phi_1\rangle \neq 0$ for $\langle\phi_0|\phi_0\rangle = \frac{1}{2}$, the maximum entanglement is approached asymptotically for $\alpha = 0$, $|\beta| \to \infty$ or $\beta = 0$, $|\alpha| \to \infty$. This is due to the fact that $|\alpha-\beta|^2 = |\alpha|^2+|\beta|^2-2\Re(\alpha\overline{\beta})$ and $|\alpha+\beta|^2 = |\alpha|^2+|\beta|^2+2\Re(\alpha\overline{\beta})$. In other words, for $\alpha, \beta \neq 0$ one term shrinking in the numerator of p_{01} implies the other is growing.

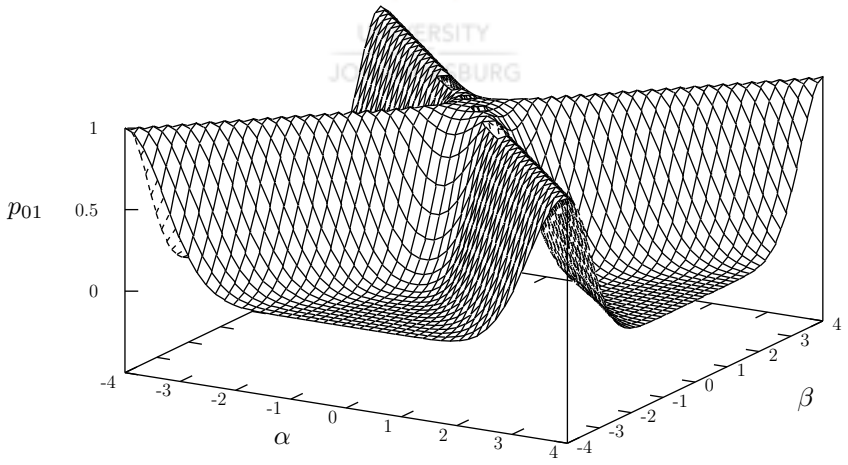
To find the entanglement we first determine the eigenvalues of ρ_1 and ρ_2 which are now given by

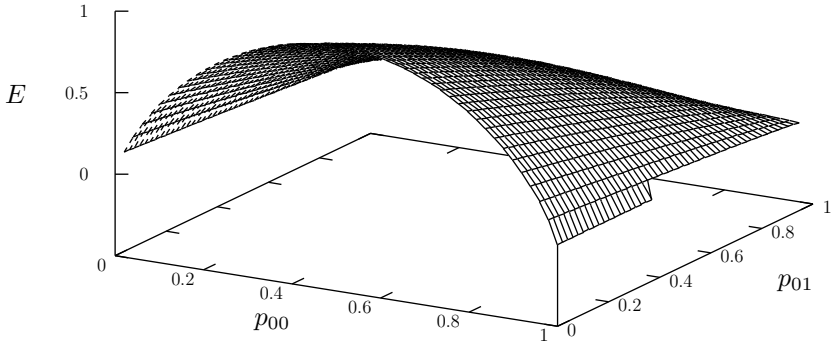
$$\lambda = \frac{1}{2}\left(1 + \sqrt{1 - 4(1-p_{01})p_{00}(1-p_{00})}\right).$$

The first figure below shows the values for p_{01} on the domain

$$\{\,(\alpha,\beta) \in \mathbf{C}\times\mathbf{C} \,:\, \Im(\alpha) = 0,\, \Im(\beta) = 0\,\}$$

and the next figure shows the entanglement. It is interesting to note that for these macroscopic quantum superposition states the inequality (2.5) results exclusively from p_{01}.





Coherent/Fock States

As a final example we consider the case when $|\phi_0\rangle$ is described by a number state and $|\phi_1\rangle$ is described by a coherent state, i.e.

$$|\psi\rangle = c_0|0\rangle \otimes |n\rangle + c_1|1\rangle \otimes |\alpha\rangle.$$

The scalar product between a number state $|n\rangle$ and a coherent state $|\alpha\rangle$ is given by

$$\langle n|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \frac{\alpha^n}{\sqrt{n!}}.$$

The condition (2.4) for this case gives

$$|c_0|^2 + |c_1|^2 = 1.$$

Consequently $\langle\phi_0|\phi_0\rangle = |c_0|^2$, and

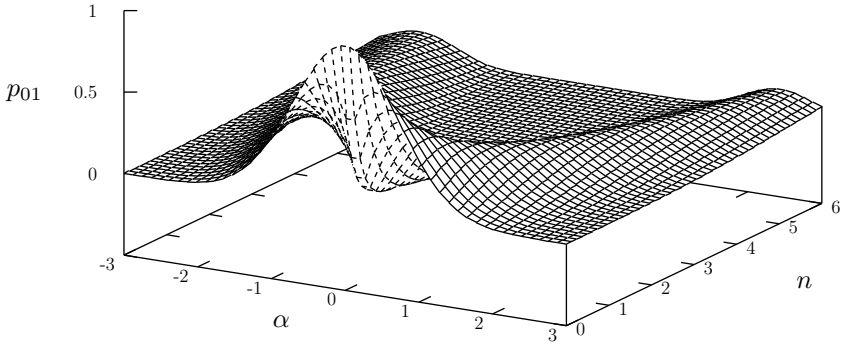
$$|\langle\phi_0|\phi_1\rangle|^2 = |c_0|^2(1 - |c_0|^2)e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}.$$

It is again convenient to define the quantities p_{00} and p_{01} as

$$p_{00} := \langle\phi_0|\phi_0\rangle = |c_0|^2, \qquad p_{01} := e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}.$$

Thus we obtain $|\langle\phi_0|\phi_1\rangle|^2 = p_{00}(1 - p_{00})p_{01}$. The entanglement can again be determined from p_{00} and p_{01} and proceeds as described above for macroscopic quantum superposition states. The figure below describes p_{01} on the domain

$\{ (\alpha, n) \in \mathbf{C} \times \mathbf{R} : \Im(\alpha) = 0 \}$, although we only consider integer n.



2.4 Entanglement for Mixed States

Since a density operator ρ is described by a linear combination of density operators describing normalized pure states, the entanglement of a mixed state can be defined in terms of the entanglement of the constituent pure states in the linear combination

$$\rho = \sum_j p_j \rho_j, \quad \sum_j p_j = 1, \quad \forall j \; p_j \geq 0, \quad \forall j \; \text{rank}(\rho_j) = \text{tr}(\rho_j) = 1$$

where the ρ_j represent mutually orthogonal pure states. Let ρ describe a mixed state in the product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ and $\mathcal{H}_\mathcal{B}$ are the Hilbert spaces describing two physical systems under consideration.

Example. The density operator ρ is said to be *separable* if it can be written in the form

$$\rho = \sum_j p_j \rho_j^A \otimes \rho_j^B, \quad \sum_j p_j = 1, \quad \forall j \; p_j \geq 0,$$

$$\forall j \; \text{rank}(\rho_j^A) = \text{tr}(\rho_j^A) = \text{rank}(\rho_j^B) = \text{tr}(\rho_j^B) = 1$$

where the $\rho_j^A \otimes \rho_j^B$ represent mutually orthogonal pure states, ρ_j^A is a density operator on the Hilbert space $\mathcal{H}_A$ and ρ_j^B is a density operator on the Hilbert space $\mathcal{H}_B$. A density operator which is not separable is said to be *entangled*.

■

Example. The Bell states $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$ and $|\Psi^-\rangle$ are entangled pure states. We consider the density operator

$$\rho := \frac{1}{4}|\Phi^+\rangle\langle\Phi^+| + \frac{1}{4}|\Phi^-\rangle\langle\Phi^-| + \frac{1}{4}|\Psi^+\rangle\langle\Psi^+| + \frac{1}{4}|\Psi^-\rangle\langle\Psi^-|.$$

Since

$$\begin{aligned}\rho &= \frac{1}{4}\mathrm{diag}(1,1,1,1) \\ &= \frac{1}{4}|0\rangle\langle 0| \otimes |0\rangle\langle 0| + \frac{1}{4}|0\rangle\langle 0| \otimes |1\rangle\langle 1| + \frac{1}{4}|1\rangle\langle 1| \otimes |0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| \otimes |1\rangle\langle 1|,\end{aligned}$$

we find ρ does not represent an entangled mixed state, even though it is described by the (entangled) Bell states.
■

Example. The *Werner state* [Werner89, Bennett96]

$$\rho_w := \frac{5}{8}|\Phi^+\rangle\langle\Phi^+| + \frac{1}{8}|\Phi^-\rangle\langle\Phi^-| + \frac{1}{8}|\Psi^+\rangle\langle\Psi^+| + \frac{1}{8}|\Psi^-\rangle\langle\Psi^-|$$

is an entangled (non-separable) mixed state.
■

2.4.1 Entanglement Witnesses

Consider the finite-dimensional Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$.

Definition. An *entanglement witness* [Cirac2002, Horodecki96] W on the product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_\mathcal{B}$ is a linear operator on $\mathcal{H}_A \otimes \mathcal{H}_\mathcal{B}$ such that W is not positive semi-definite and

$$\langle\psi|W|\psi\rangle \geq 0$$

for all separable $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, where $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$.
■

Let $s := \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\}$ and

$$\rho = \sum_{j=1}^{s} p_j |j_A\rangle\langle j_A| \otimes |j_B\rangle\langle j_B|$$

$$= \sum_{j=1}^{s} p_j (|j_A\rangle \otimes |j_B\rangle)(\langle j_A| \otimes \langle j_B|)$$

be a separable mixed state on $\mathcal{H}_A \otimes \mathcal{H}_\mathcal{B}$, where

$$\{|1_A\rangle, |2_A\rangle, \ldots, |\dim(\mathcal{H}_A)_A\rangle\}$$

and

$$\{|1_B\rangle, |2_B\rangle, \ldots, |\dim(\mathcal{H}_B)_B\rangle\}$$

denote orthonormal basis for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. Then

$$\mathrm{tr}(W\rho) = \sum_{k=1}^{\dim(\mathcal{H}_\mathrm{A})} \sum_{l=1}^{\dim(\mathcal{H}_\mathrm{B})}$$

$$\langle k_A| \otimes \langle l_B| \left(W \sum_{j=1}^{s} p_j (|j_A\rangle \otimes |j_B\rangle)(\langle j_A| \otimes \langle j_B|) \right) |k_A\rangle \otimes |l_B\rangle$$

$$= \sum_{j=1}^{s} p_j (\langle j_A| \otimes \langle j_B|) W (|j_A\rangle \otimes |j_B\rangle) \geq 0$$

Consequently, a necessary criteria for separability of a mixed state ρ is that $\mathrm{tr}(W\rho) \geq 0$ for all entanglement witnesses W.

Example. Consider an arbitrary pure state ρ which is not separable, thus we write

$$\rho := |k_{AB}\rangle\langle k_{AB}|$$

where $|k_{AB}\rangle$ is a normalized entangled pure state. Defining the entanglement witness W as

$$W := I - \mu_{k_{AB}} |k_{AB}\rangle\langle k_{AB}|$$

where I is the identity operator on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ and

$$\mu_{k_{AB}} := \left(\max_{\substack{|a\rangle\otimes|b\rangle\in\mathcal{H}_A\otimes\mathcal{H}_B \\ \langle a|a\rangle=\langle b|b\rangle=1}} \left|(\langle a|\otimes\langle b|)(|k_{AB}\rangle)\right|^2 \right)^{-1}$$

it is clear that W is not positive definite. Furthermore, for $|\psi\rangle$ separable we have

$$\langle\psi|W|\psi\rangle = \langle\psi_A|\psi_A\rangle\langle\psi_B|\psi_B\rangle - \mu_{k_{AB}}|\langle\psi|k_{AB}\rangle|^2 \geq 0$$

and

$$\mathrm{tr}\left(\rho - \mu_{k_{AB}}|k_{AB}\rangle\langle k_{AB}|\right) = 1 - \mu_{k_{AB}} < 0$$

since $\mu_{k_{AB}} > 1$. Thus W is an entanglement witness for ρ and ρ is entangled.
■

Theorem. [Cirac2002, Horodecki96] A mixed state ρ is separable if and only if $\mathrm{tr}(W\rho) \geq 0$ for all entanglement witnesses W.
■

This results follows from the *Hahn-Banach separation theorem* [Garret98, Garret05] which is as follows.

Theorem. Given a locally convex, topological vector space V, a non-empty compact convex subset X, and non-empty closed convex set Y with $X \cap Y = \emptyset$. Then there exists a continuous linear functional λ on V and constants $c_1, c_2 \in \mathbf{R}$ with $c_1 < c_2$ such that for all $x \in X$ and $y \in Y$

$$\Re(\lambda(x)) \leq c_1 < c_2 \leq \Re(\lambda(y))$$

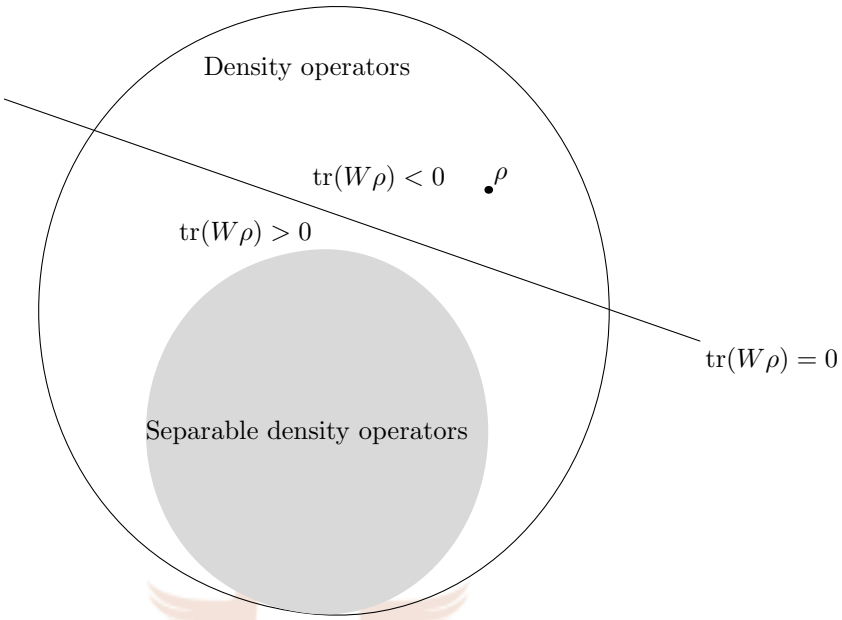
where $\Re(\lambda(x))$ is the real part of the complex number $\lambda(x)$.
■

The set of separable density operators is clearly convex, and can be shown to be compact. Thus, according to the Hahn-Banach theorem, there exists a hyperplane separating the separable density operators from an entangled mixed state ρ (which on its own forms a closed convex set). This hyperplane is described by W according to

$$\{\sigma \,:\, \mathrm{tr}(W\sigma) = 0\}$$

where $\mathrm{tr}(W\sigma) = \langle W, \sigma\rangle$ is the scalar product between the entanglement

witness W and the density operator σ.



Example. Consider

$$W = I_4 - 2|\Phi^+\rangle\langle\Phi^+| = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

The eigenvalues of W are -1 and 1 with multiplicity 3, with corresponding eigenvectors given by the Bell basis. Clearly W is not positive semi-definite. Using $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, with $|\psi_A\rangle, |\psi_B\rangle \in \mathbf{C}^2$, we find

$$\begin{aligned}
\langle\psi|W|\psi\rangle &= \langle\psi_A|\psi_A\rangle\langle\psi_B|\psi_B\rangle - |\langle\psi_A|0\rangle\langle\psi_B|0\rangle + \langle\psi_A|1\rangle\langle\psi_B|1\rangle|^2 \\
&= \Big(|\langle\psi_A|0\rangle|^2 + |\langle\psi_A|1\rangle|^2\Big)\Big(|\langle\psi_B|0\rangle|^2 + |\langle\psi_B|1\rangle|^2\Big) \\
&\quad - |\langle\psi_A|0\rangle\langle\psi_B|0\rangle + \langle\psi_A|1\rangle\langle\psi_B|1\rangle|^2 \\
&\geq \Big(|\langle\psi_A|0\rangle|^2 + |\langle\psi_A|1\rangle|^2\Big)\Big(|\langle\psi_B|0\rangle|^2 + |\langle\psi_B|1\rangle|^2\Big) \\
&\quad - \Big(|\langle\psi_A|0\rangle\langle\psi_B|0\rangle| + |\langle\psi_A|1\rangle\langle\psi_B|1\rangle|\Big)^2
\end{aligned}$$

$$= \Big(|\langle\psi_A|0\rangle\langle\psi_B|1\rangle| - |\langle\psi_A|1\rangle\langle\psi_B|0\rangle|\Big)^2 \geq 0.$$

Thus W is an entanglement witness. Furthermore,

$$\mathrm{tr}(W|\Phi^+\rangle\langle\Phi^+|) = \langle\Phi^+|W|\Phi^+\rangle = -1,$$

and consequently W is a witness for the entangled Bell state $|\Phi^+\rangle$.
■

2.4.2 Positive Maps

Definition. Let L denote the space of linear operators on a Hilbert space $\mathcal{H}$. A linear map $\epsilon : L \to L$ is called a *positive map* if for all positive semidefinite $\rho \in L$ the operator $\epsilon(\rho)$ is also positive semidefinite.
■

Example. Consider the unitary evolution $U(t)$ of a density operator ρ

$$\rho(t) = U(t)\rho(0)U^*(t).$$

The linear transform (with respect to ρ)

$$\epsilon_t(\rho(0)) = U(t)\rho(0)U^*(t)$$

implements the time evolution and is a clearly a positive map.
■

Definition. Let $B = \{|1\rangle, |2\rangle, \ldots, |\dim(\mathcal{H})\rangle\}$ denote an orthonormal basis for the Hilbert space $\mathcal{H}$. The *transpose* of a linear operator

$$A := \sum_{j=1}^{\dim(\mathcal{H})} \sum_{k=1}^{\dim(\mathcal{H})} a_{jk}|j\rangle\langle k|$$

is given by

$$\epsilon_{T(B)}(A) := \sum_{j=1}^{\dim(\mathcal{H})} \sum_{k=1}^{\dim(\mathcal{H})} a_{jk}|k\rangle\langle j|.$$

The transpose operation $\epsilon_{T(B)}$ is a positive map.

■

Definition. Let L_A denote the space of linear operators on a Hilbert space $\mathcal{H}_A$, L_B the space of linear operators on the Hilbert space $\mathcal{H}_B$ and L_{AB} the space of linear operators on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. The extension $\epsilon \otimes I$ of the linear map $\epsilon : L_A \to L_A$ is defined by

$$(\epsilon \otimes I)\left(\sum_{k=1}^{\dim(L_A)} A_k \otimes B_k\right) = \sum_{k=1}^{\dim(L_A)} \epsilon(A_k) \otimes B_k$$

where $A_k \in L_A$ and $B_k \in L_B$ for $k = 1, 2, \ldots, \dim(L_A)$. Similarly the extension $I \otimes \mu$ of $\mu : L_B \to L_B$ can be defined.

■

Definition. A positive map $\epsilon : L \to L$ is a *completely positive map* if all possible extensions ($\epsilon \otimes I$ or $I \otimes \epsilon$) of the map to arbitrary Hilbert spaces ($L \otimes L_{ext}$ or $L_{ext} \otimes L$) where L_{ext} is the space of linear operators on the arbitrary Hilbert space $\mathcal{H}_{ext}$.

■

Every completely positive map can be written in the form

$$\epsilon(\rho) = \sum_k A_k \rho A_k^*$$

where A_k is a linear operator on the Hilbert space. Furthermore, if

$$\sum_k A_k A_k^* = I$$

then ϵ is trace preserving, i.e. $\mathrm{tr}(\epsilon(\rho)) = \mathrm{tr}(\rho)$.

Example. *Partial transposition* $\epsilon_{T(B)} \otimes I$ in the basis B of the Hilbert space $\mathcal{H}_A$ of a state ρ in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is not completely positive. Consider the Bell state in the Hilbert state $\mathbf{C}^2 \otimes \mathbf{C}^2$

$$\begin{aligned}|\Phi^+\rangle\langle\Phi^+| = \frac{1}{2}(&|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| \\ &+|1\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)\end{aligned}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Thus we have

$$\begin{aligned}(\epsilon_{T(\{|0\rangle,|1\rangle\})} \otimes I)(|\Phi^+\rangle\langle\Phi^+|) = \frac{1}{2}(&|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| \\ &+|0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.\end{aligned}$$

This last matrix has as eigenvalues 1 and -1 and consequently is not positive semidefinite.

■

Now consider the linear map $\epsilon : L_A \to L_B$

$$\epsilon_{W,B}(\rho) = \mathrm{tr}_A(W\epsilon_{T(B)}(\rho) \otimes I)$$

where tr_A denotes the partial trace over the first system and $\epsilon_{T(B)}$ denotes the partial transpose over the first system in the orthonormal basis $B := \{|1\rangle, |2\rangle, \ldots, |\dim(\mathcal{H}_A)\rangle\}$. Given $\epsilon_{W,B}$ and B it is possible to determine $W \in L_A \otimes L_B$ as

$$W = (I \otimes \epsilon_{W,B}) \sum_{j,k=1}^{\dim(\mathcal{H}_A)} \Big(|j\rangle\langle k| \otimes |j\rangle\langle k|\Big)$$

where we write $|j\rangle$ for $j = 1, 2, \ldots, \dim(\mathcal{H}_A)$ as the orthonormal basis B for $\mathcal{H}_A$. This can be verified as follows

$$\begin{aligned}&\mathrm{tr}_A\left((I \otimes \epsilon_{W,B})\left(\sum_{j,k=1}^{\dim(\mathcal{H}_A)} |j\rangle\langle k| \otimes |j\rangle\langle k|\right)\sum_{l,m=1}^{\dim(\mathcal{H}_A)} p_{ml}|l\rangle\langle m| \otimes I\right) \\ &= \mathrm{tr}_A\left(\sum_{j,k=1}^{\dim(\mathcal{H}_A)} |j\rangle\langle k| \otimes \epsilon_{W,B}(|j\rangle\langle k|)\sum_{l,m=1}^{\dim(\mathcal{H}_A)} p_{ml}|l\rangle\langle m| \otimes I\right)\end{aligned}$$

$$= \mathrm{tr}_A \left(\sum_{j,k,m=1}^{\dim(\mathcal{H}_A)} p_{mk} |j\rangle\langle m| \otimes \epsilon_{W,B}(|j\rangle\langle k|) \right)$$

$$= \sum_{j,k=1}^{\dim(\mathcal{H}_A)} p_{jk} \epsilon_{W,B}(|j\rangle\langle k|)$$

$$= \epsilon_{W,B}(\rho)$$

where we used

$$\rho := \sum_{l,m=1}^{\dim(\mathcal{H}_A)} p_{lm} |l\rangle\langle m|.$$

Consequently

$$\epsilon_{T(B)}(\rho) = \sum_{l,m=1}^{\dim(\mathcal{H}_A)} p_{ml} |l\rangle\langle m|.$$

Thus we have an isomorphism between $\epsilon_{W,B} : L_A \to L_B$ and $W : \mathcal{H}_A \otimes \mathcal{H}_\mathcal{B} \to \mathcal{H}_A \otimes \mathcal{H}_\mathcal{B}$.

Theorem. A mixed state ρ on $\mathcal{H}_A \otimes \mathcal{H}_\mathcal{B}$ is separable if and only if $(I \otimes \epsilon)(\rho)$ is positive semidefinite for every positive map $\epsilon : L_B \to L_A$.
■

This a consequence of the isomorphism described above. Consider the separable and positive semidefinite density operator

$$\rho := \sum_{k=1}^{s} p_k |k\rangle\langle k| \otimes |k\rangle\langle k|$$

where $s := \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\}$ and $\{|1\rangle, |2\rangle, \ldots, |s\rangle\}$ forms an orthonormal basis in $\mathcal{H}_A$ and similarly in $\mathcal{H}_B$ (i.e. the bases are isomorphic). Thus

$$(I \otimes \epsilon)(\rho) = \sum_{k=1}^{s} p_k |k\rangle\langle k| \otimes \epsilon(|k\rangle\langle k|)$$

which remains positive semidefinite since ϵ is a positive map and the convex linear combination of positive semidefinite operators is positive semidefinite.

Suppose $(I\otimes\epsilon)(\rho)$ is positive semidefinite for every positive map $\epsilon : L_B \to L_A$. We note that for any Hilbert space $\langle\rho_1, \rho_2\rangle := \mathrm{tr}(\rho_1\rho_2^*)$ defines an inner product for the Hilbert space of corresponding linear operators. Obviously for the density operator ρ we have $\rho^* = \rho$. Let W be an arbitrary entanglement witness, then

$$\begin{aligned}
\mathrm{tr}(W\rho) = \langle W, \rho\rangle &= \left\langle (I \otimes \epsilon_{W,B}) \left(\sum_{j,k=1}^{\dim(\mathcal{H}_A)} |j\rangle\langle k| \otimes |j\rangle\langle k| \right), \rho \right\rangle \\
&= \left\langle \sum_{j,k=1}^{\dim(\mathcal{H}_A)} |j\rangle\langle k| \otimes |j\rangle\langle k|, (I \otimes \epsilon_{W,B})^*(\rho) \right\rangle \\
&= \mathrm{tr}\left(\left(\sum_{j,k=1}^{\dim(\mathcal{H}_A)} |j\rangle\langle k| \otimes |j\rangle\langle k| \right) (I \otimes \epsilon_{W,B})^*(\rho) \right) \\
&= \left(\sum_{j=1}^{\dim(\mathcal{H}_A)} \langle j| \otimes \langle j| \right) (I \otimes \epsilon_{W,B})^*(\rho) \left(\sum_{k=1}^{\dim(\mathcal{H}_A)} |k\rangle \otimes |k\rangle \right) \\
&= \left(\sum_{j=1}^{\dim(\mathcal{H}_A)} \langle j| \otimes \langle j| \right) (I \otimes \epsilon_{W,B})(\rho) \left(\sum_{k=1}^{\dim(\mathcal{H}_A)} |k\rangle \otimes |k\rangle \right) \\
&\geq 0.
\end{aligned}$$

We used the fact that the adjoint linear map $(I \otimes \epsilon_{W,B})^*$ is also an extended positive map $I\otimes\epsilon$ where $\epsilon : L_B \to L_A$ is a positive map, and that $(I\otimes\epsilon_{W,B})(\rho)$ is self-adjoint due to the fact that $(I \otimes \epsilon_{W,B})(\rho)$ can be written in the form

$$(I \otimes \epsilon_{W,B})(\rho) = \sum_k (I \otimes A_k)\rho(I \otimes A_k)^*$$

where ρ is self-adjoint.

2.4.3 Entanglement of Formation

Here we introduce a measure of entanglement for mixed states. It is called the *entanglement of formation* [Bennett96]. Let A and B be two finite-dimensional quantum systems. We define for the coupled system $A \otimes B$

$$\rho_A := \mathrm{tr}_B(\rho_{AB})$$

where tr_B denotes the partial trace over B, i.e., we use $I \otimes |\beta_j\rangle$ as the basis for the trace where $|\beta_j\rangle$ is an orthonormal basis in B and I is the identity operator on A. The measure of entanglement $E(AB)$ is then defined as

$$E(AB) := S(\rho_A).$$

This describes the entanglement for pure states.

Example. For the state $|\psi\rangle := \frac{1}{\sqrt{3}}(|00\rangle + |01\rangle + |10\rangle)$ we find

$$\rho_A = \mathrm{tr}_B|\psi\rangle\langle\psi| = \frac{1}{3}(2|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|).$$

Thus

$$\begin{aligned} E(\psi) &= -\left(\frac{1}{2} + \frac{\sqrt{5}}{6}\right)\log_2\left(\frac{1}{2} + \frac{\sqrt{5}}{6}\right) - \left(\frac{1}{2} - \frac{\sqrt{5}}{6}\right)\log_2\left(\frac{1}{2} - \frac{\sqrt{5}}{6}\right) \\ &\approx 0.55 \end{aligned}$$

where we used the eigenvalues $\frac{1}{2} \pm \frac{\sqrt{5}}{6}$ of ρ_A.
■

For mixed states we take the minimum entanglement as defined above over all possible ensembles of pure states realizing ρ_{AB}. First we define the set $En(\rho_{AB})$ all ensembles realizing ρ_{AB} by

$$\begin{aligned} &\{\,(p_1, |\psi_1\rangle), \ldots, (p_n, |\psi_n\rangle)\,\} \in En(\rho_{AB}) \\ \Leftrightarrow\ & p_1, p_2, \ldots, p_n \geq 0, \quad \sum_{j=1}^{n} p_j = 1, \quad \sum_{j=1}^{n} p_j |\psi_j\rangle\langle\psi_j| = \rho_{AB}. \end{aligned}$$

The entanglement of formation is then the minimum over all possible ensembles

$$E(AB) := \min_{P \in En(\rho_{AB})} \sum_{(p,|\psi\rangle)\in P} pS(\mathrm{tr}_B|\psi\rangle\langle\psi|).$$

If we have that the Hilbert space for both A and B is $\mathbf{C}^2$, then we can calculate the entanglement of formation using a quantity called the *concurrence*

$$C(\rho_{AB}) := \max\{\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4, 0\}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ are the eigenvalues of ρ_{AB}. The entanglement of

formation can be calculated from

$$E(AB) = H_2\left(\frac{1+\sqrt{1-(C(\rho_{AB}))^2}}{2}\right)$$

where

$$H_2(p) := -p\log_2 p - (1-p)\log_2(1-p)$$

is the *Shannon entropy*.

Example. Consider the density matrix (*Werner state*) [Werner89, Bennett96]

$$\rho_w := r|\phi^+\rangle\langle\phi^+| + \frac{1-r}{4}I_4$$

where $|\phi^+\rangle = \frac{1}{\sqrt{2}}(1,0,0,1)^T$ is the Bell state, and $0 \leq r \leq 1$.

We have

$$\rho_w = \begin{pmatrix} (1+r)/4 & 0 & 0 & r/2 \\ 0 & (1-r)/4) & 0 & 0 \\ 0 & 0 & (1-r)/4 & 0 \\ r/2 & 0 & 0 & (1+r)/4 \end{pmatrix}.$$

Thus $\mathrm{tr}(\rho_w) = 1$. The eigenvalues are $(1+r)/4 + r/2 = (1+3r)/4$ and $(1-r)/4$ with multiplicity 3, so that

$$\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 = (1+3r)/4 - 3(1-r)/4 = (3r-1)/2.$$

The concurrence is

$$C(\rho_w) = \max\{(3r-1)/2, 0\}.$$

Hence for $r = \frac{1}{2}$ the concurrence is

$$C(\rho_w) = \max\left\{\frac{1}{4}, 0\right\} = \frac{1}{4}.$$

Thus $E(AB) \approx 0.1176$ [Bennett96].

■

2.4.4 Relative Entropy of Entanglement

Definition. The *relative entropy of entanglement* of a density operator ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is given by

$$\min_{\sigma \in D(\mathcal{H}_A \otimes \mathcal{H}_B)} S_b(\rho||\sigma)$$

where $D(\mathcal{H}_A \otimes \mathcal{H}_B)$ is the set of all separable density operators in $\mathcal{H}_A \otimes \mathcal{H}_B$. ■

If we further assume that ρ represents a pure state, i.e. $\lambda_1 = 1$ and $\lambda_j = 0$ for $j = 2, 3, \ldots, n$ we find

$$\begin{aligned} S_b(\rho||\sigma) &= \mathrm{tr}(\rho \log_b \rho - \rho \log_2 \sigma) \\ &= -\mathrm{tr}(\rho \log_b \sigma) \\ &= -\sum_{k=1}^{n} (\log_b \mu_k) |\langle \phi_1 | \psi_k \rangle|^2. \end{aligned}$$

Example. Let ρ represent a pure state $|\phi\rangle$ on $\mathcal{H}_A \otimes \mathcal{H}_B$, $m = \dim \mathcal{H}_A$ and $n = \dim \mathcal{H}_B$. Let $\sigma \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$. Thus we find that the eigenvalues μ_{kl} have corresponding eigenstates $|\psi_{A,k}\rangle \otimes |\psi_{B,l}\rangle$ where $\{|\psi_{A,k}\rangle,\ k = 1, 2, \ldots, m\}$ and $\{|\psi_{B,l}\rangle,\ l = 1, 2, \ldots, n\}$ are orthonormal basis in $\mathcal{H}_A$ and $\mathcal{H}_B$), respectively. Some of the the μ_{kl} may be zero. Thus we find

$$S_b(\rho||\sigma) = -\sum_{k=1}^{m} \sum_{l=1}^{n} (\log_2 \mu_{kl}) |\langle \phi | (|\psi_{A,k}\rangle \otimes |\psi_{B,l}\rangle)|^2.$$

Since we seek a minimum we make the assumption

$$\log_2 \mu_{kl} = 0 \quad \Rightarrow \quad \langle \phi | (|\psi_{A,k}\rangle \otimes |\psi_{B,l}\rangle) = 0.$$

Now consider the entanglement of formation. First we determine

$$\begin{aligned} \rho_A &:= \mathrm{tr}_{\mathcal{H}_B} \rho \\ &= \sum_{l=1}^{n} (I \otimes \langle \psi_{B,l}|) |\phi\rangle\langle\phi| (I \otimes |\psi_{B,l}\rangle). \end{aligned}$$

It is convenient to find the eigenvalues λ_j with corresponding orthonormal eigenstates $|\phi_{A,k}\rangle$ for $k = 1, 2, \ldots, m$ of ρ_A as follows

$$\begin{aligned}
\lambda_k &= \langle\phi_{A,k}|\rho_A|\phi_{A,k}\rangle \\
&= \lambda_k\langle\phi_{A,k}|\phi_{A,k}\rangle \\
&= \langle\phi_{A,k}|\left(\sum_{l=1}^{n}(I\otimes\langle\psi_{B,l}|)|\phi\rangle\langle\phi|(I\otimes|\psi_{B,l}\rangle)\right)|\phi_{A,k}\rangle \\
&= \sum_{l=1}^{n}(\langle\phi_{A,k}|\otimes\langle\psi_{B,l}|)|\phi\rangle\langle\phi|(|\phi_{A,k}\rangle\otimes|\psi_{B,l}\rangle) \\
&= \sum_{l=1}^{n}|\langle\phi|(|\phi_{A,k}\rangle\otimes|\psi_{B,l}\rangle)|^2.
\end{aligned}$$

Thus we find for the entanglement of formation

$$\begin{aligned}
S_b(\rho_A) &= -\mathrm{tr}(\rho_A\log_b\rho_A) \\
&= -\sum_{k=1}^{m}\lambda_k\log_b\lambda_k \\
&= -\sum_{k=1}^{m}\sum_{l=1}^{n}(\log_b\lambda_k)|\langle\phi|(|\phi_{A,k}\rangle\otimes|\psi_{B,l}\rangle)|^2.
\end{aligned}$$

Consequently

$$\begin{aligned}
S_b(\rho||\sigma)-S_b(\rho_A) &= -\sum_{k=1}^{m}\sum_{l=1}^{n}(\log_b\mu_{kl})|\langle\phi|(|\psi_{A,k}\rangle\otimes|\psi_{B,l}\rangle)|^2 \\
&\quad+\sum_{k=1}^{m}\sum_{l=1}^{n}(\log_b\lambda_k)|\langle\phi|(|\phi_{A,k}\rangle\otimes|\psi_{B,l}\rangle)|^2 \\
&= \sum_{k=1}^{m}\sum_{l=1}^{n}\Big[(\log_b\lambda_k)|\langle\phi|(|\phi_{A,k}\rangle\otimes|\psi_{B,l}\rangle)|^2 \\
&\quad-(\log_b\mu_{kl})|\langle\phi|(|\psi_{A,k}\rangle\otimes|\psi_{B,l}\rangle)|^2\Big].
\end{aligned}$$

If σ is such that $\mu_{kl}=\lambda_k$ and $|\psi_{A,k}\rangle=|\phi_{A,k}\rangle$ then $S(\rho||\sigma)=S(\rho_A)$.

Since ρ represents a pure state we have that $S_b(\rho)=0$ and from the *Akari-Lieb inequality* [Nielsen]

$$S_b(\rho)\geq|S_b(\rho_A)-S_b(\rho_b)|$$

where $S_b(\rho) = 0$, so that $S_b(\rho_A) = S_b(\rho_B)$. Since we seek $\sigma \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ that minimizes $S_b(\rho||\sigma)$ and

$$S_b(\rho||\rho_A \otimes \rho_B) = S_b(\rho_A) + S_b(\rho_B) - S_b(\rho) = 2S_b(\rho_A)$$

the following inequality holds

$$S_b(\rho||\sigma) \leq S_b(\rho||\rho_A \otimes \rho_B) = 2S_b(\rho_A).$$

Consequently

$$S_b(\rho||\sigma) - S(\rho_A) \geq S_b(\rho||\sigma) - \frac{1}{2}S_b(\rho||\sigma) = \frac{1}{2}S_b(\rho||\sigma) \geq 0.$$

The last inequality is *Klein's inequality*. Thus the minimum is given by $S_b(\rho_A)$, i.e. when ρ represents a pure state, the entanglement of formation and entropy of entanglement are identical.
■

2.4.5 Distillable Entanglement

Distillable entanglement [Bennett96, Bruß2002, Rains99] measures the maximum dimension of a maximally entanglement state that can be obtained from an infinite number of copies of a state under given constraints on what operations may be performed on the states. Thus different classes of operations yield different measures of distillable entanglement. We allow subsets of generalized measurements.

Let $A = \{A_1, \ldots, A_k\}$ and $B = \{B_1, \ldots, B_l\}$ denote generalized measurements. For $k = 1$ A_1 describes a unitary operation.

Definition. *Composition* $A \circ B$ is given by application of B, and then application of A. In other words, B is applied to ρ and yields the outcome corresponding to B_j and consequently we obtain the state

$$\rho_{B,j} := \frac{B_j \rho B_j^*}{\mathrm{tr}(B_j \rho B_j^*)}$$

and then A is applied to ρ_B yielding the outcome corresponding to A_n giving the state

$$\rho_{A\circ B,j,n} := \frac{A_n B_j \rho B_j^* A_n^*}{\mathrm{tr}(A_n B_j \rho B_j^* A_n^*)}.$$

Furthermore

$$\sum_{j=l}^{k}\sum_{n=1}^{k} B_j^* A_n^* A_n B_j = \sum_{j=l}^{k} B_j^* \left(\sum_{n=1}^{k} A_n^* A_n\right) B_j = I.$$

In other words we have generalized measurement according to

$$A \circ B := \{A_n B_j \,|\, A_n \in A \text{ and } B_j \in B\}.$$

■

Definition. The *tensor product* $A \otimes B$ is a generalized measurement given by

$$A \otimes B := \{A_n \otimes B_j \,|\, A_n \in A \text{ and } B_j \in B\}.$$

■

A further operation that reorders two measurement operations which have the same outcome may also be defined [Rains99] . However if we insist, as in chapter 1, that all measurement outcomes are distinct then this operation may be omitted.

Definition. A *class of operations* is a set of generalized measurements including the identity which is closed under composition.
■

Definition. A generalized measurement of cardinality 1 is said to be *non-measuring*, otherwise it is said to be *measuring*.
■

In the literature 5 classes of operations are often considered [Rains99]

1. Local operations:
   The set of all local unitary operation $U_A \otimes U_B$ on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ where U_A and U_B are unitary. In other words a density operator ρ is transformed according to $\rho \to (U_A \otimes U_B)\rho(U_A \otimes U_B)^*$.

2. 1-Local operations:
   Local operations and generalized measurement on $\mathcal{H}_A$ (i.e. generalized measurement of the form $A \otimes I$ where A is a generalized measurement on $\mathcal{H}_A$) and classical communication.

3. 2-Local operations:
   Local operations and independent generalized measurement on $\mathcal{H}_A$ ($A \otimes I$, for generalized measurement A) and on $\mathcal{H}_B$ ($I \otimes B$ for generalized measurement B) and classical communication.

4. Separable operations:
   The set of all generalized measurements of the form

   $$M := \{A_1 \otimes B_1, \ldots, A_k \otimes B_k\}$$

   i.e. generalized measurements which do not disturb separability.

5. Positive partial transpose operations:
   The class of generalized measurements which have when composed with the positive partial transpose are completely positive.

[Rains99] provides 4 definitions for the distillable entanglement which are equivalent provided the relevant class of operations include the 1-local operations. Here we give two of the convenient and intuitive definitions.

Definition. Let C be a class of operations and ρ be a density operator on $\mathcal{H}_A \otimes \mathcal{H}_B$. The *distillable entanglement* $D_C(\rho)$ is given by

$$D_C(\rho) := \lim_{n\to\infty} \frac{1}{n} \max_{T_n \in C} \sum_{j=1}^{k} p_j \log_2 m_j$$

where the maximum (implicitly over m_1, m_2, ..., m_k which determines the maximum Schmidt ranks of the the maximally entangled state $|\Phi_{m_j}\rangle$ below) is taken over all $T_n = \{T_{n,1}, \ldots, T_{n,k}\} \in C$ (i.e. sequences of operations from C acting on density operators on $\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n}$ since C is closed under composition) such that the fidelity

$$F_{m_j}(T_{n,j}(\rho)) := \langle \Phi_{m_j} | T_{n,j}(\rho) | \Phi_{m_j} \rangle$$

is high (and increasing for n increasing) where

$$|\Phi_{m_j}\rangle := \frac{1}{\sqrt{m_j}} \sum_{l=1}^{m_j} |l_A\rangle \otimes |l_B\rangle$$

is a maximally entangled state with $|l_A\rangle \in \mathcal{H}_A^{\otimes n}$ and $|l_B\rangle \in \mathcal{H}_B^{\otimes n}$ for $l = 1, \ldots, m_j$ orthonormal states in their respective Hilbert spaces. The quantity $0 \leq p_j \leq 1$ denotes the probability that $T_{n,j}$ is performed in a measuring operation.

■

An equivalent definition [Rains99] is as follows.

Definition. Let C be a class of operations and ρ be a density operator on $\mathcal{H}_A \otimes \mathcal{H}_B$. The distillable entanglement $D_C(\rho)$ is given by

$$D_C(\rho) := \lim_{n\to\infty} \frac{1}{n} \max_{T_{n,m}\in C} \log_2 m$$

where the maximum (implicitly over m, which determines the maximum Schmidt rank of the the maximally entangled state $|\Phi_m\rangle$ below) is taken over all non-measuring $T_{n,m} \in C$ (i.e. sequences of operations from C acting on density operators on $\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n}$ since C is closed under composition) such that the fidelity

$$F_m(T_{n,m}(\rho)) := \langle \Phi_m | T_{n,m}(\rho) | \Phi_m \rangle$$

satisfies the limit

$$\lim_{n\to\infty} F_m(T_{n,m}(\rho)) = 1$$

when the maximum is attained and

$$|\Phi_m\rangle := \frac{1}{\sqrt{m}} \sum_{j=1}^{m} |j_A\rangle \otimes |j_B\rangle$$

is a maximally entangled state with $|j_A\rangle \in \mathcal{H}_A^{\otimes n}$ and $|j_B\rangle \in \mathcal{H}_B^{\otimes n}$ for $j = 1, \ldots, m$ orthonormal states in their respective Hilbert spaces.

■

Example. As a simple example we consider the distillation under 1-local operations of a pure state in the Hilbert space $\mathbf{C}^2$

$$|\psi\rangle := \alpha|00\rangle + \beta|11\rangle$$

with $|\alpha|^2 + |\beta|^2 = 1$. If this state can be distilled, it will be transformed into the Bell state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle.$$

We could attempt to construct a separable unitary to perform this transform, however one component of a generalized measurement on the first qubit will achieve this for $k \in \mathbf{C}$, $k \neq 0$

$$A_0 := \frac{k}{\alpha}|0\rangle\langle 0| + \frac{k}{\beta}|1\rangle\langle 1|$$

which transforms the density operator $\rho := |\psi\rangle\langle\psi|$ to

$$\frac{(A_0 \otimes I)\rho(A_0 \otimes I)^*}{\mathrm{tr}(A_0 \otimes I)\rho(A_0 \otimes I)^*} = |\Phi^+\rangle\langle\Phi^+|$$

with probability

$$p_0 := \mathrm{tr}(A_0 \otimes I)\rho(A_0 \otimes I)^* = 2|k|^2.$$

Of course this is only possible for $|\alpha^2|, |\beta|^2 \neq 0$, i.e. for $\alpha = 0$ or $\beta = 0$ there is no distillable entanglement. To complete the generalized measurement we must find A_1 such that

$$A_0^* A_0 + A_1^* A_1 = I.$$

Since $A_0^* A_0$ is clearly positive semidefinite we can use the polar decomposition $A_1 = U_1 H_1$, where U_1 is unitary and H_1 is positive semidefinite. We obtain

$$A_1 = U_1\sqrt{I - A_0^* A_0} = U_1\left(\sqrt{1 - \left|\frac{k}{\alpha}\right|^2}|0\rangle\langle 0| + \sqrt{1 - \left|\frac{k}{\beta}\right|^2}|1\rangle\langle 1|\right).$$

Consequently it is necessary that $k \leq \min\{|\alpha|^2, |\beta|^2\}$.

2.5 Entanglement Capability

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ denote two finite-dimensional Hilbert spaces. Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. The *von Neumann entropy* is given by

$$E(|\psi\rangle) := -\mathrm{tr}_A(\rho_A \log_2 \rho_A)$$

where $\rho_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|)$. The *entanglement capability* of an operator H on $\mathcal{H}_A \otimes \mathcal{H}_B$ is defined as [Wang2003]

$$E(H) := \max_{|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B} \Gamma(t)|_{t\to 0}$$

where

$$\Gamma(t) := \frac{dE(\exp(-iHt)|\psi(0)\rangle)}{dt}$$

is the *state entanglement rate*.

Example. Consider the Hamilton operator

$$H = X_A \otimes X_B$$

where the linear operator X_A acts on $\mathcal{H}_A$ and the linear operator X_B acts on $\mathcal{H}_B$. Assume that $X_A = X_A^{-1}$ and $X_B = X_B^{-1}$, and consequently $H = H^{-1}$. Let $\rho_{AB}(t) := |\psi(t)\rangle\langle\psi(t)|$ and $\rho_A(t) := \mathrm{tr}_B(\rho_{AB}(t))$. We have

$$\rho_{AB}(t) = \exp(-iHt)\rho_{AB}(0)\exp(iHt)$$

and

$$i\frac{d\rho_{AB}}{dt}(t) = [H, \rho_{AB}(t)].$$

Thus

$$\frac{d\rho_A}{dt}(t) = \mathrm{tr}_B[H, \rho_{AB}(t)].$$

$$\Gamma(t) = -\frac{d}{dt}\mathrm{tr}_\mathrm{A}(\rho_A \log_2 \rho_A) = -\mathrm{tr}_\mathrm{A}\left(\frac{d}{dt}\rho_A \log_2 \rho_A\right)$$

$$= -\mathrm{tr}_\mathrm{A}\left(\frac{d\rho_A}{dt}\log_2 \rho_A + \rho_A \frac{d}{dt}\log_2 \rho_A\right) = -\mathrm{tr}_\mathrm{A}\left(\frac{d\rho_A}{dt}\log_2 \rho_A\right)$$

$$= i\mathrm{tr}_\mathrm{A}\left(\mathrm{tr}_B[H, \rho_{AB}]\log_2 \rho_A\right)$$

since $\mathrm{tr}_A(\rho_A \frac{d}{dt}\log_2 \rho_A) = 0$. Let

$$|\psi(0)\rangle = \sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_j}|\phi_j\rangle \otimes |\eta_j\rangle$$

be the Schmidt decomposition of $|\psi(0)\rangle$, where $\lambda_j > 0$ with

$$\sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \lambda_j = 1,$$

and

$$\{\, |\phi_1\rangle, \ldots, |\phi_{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)}\rangle \,\}, \qquad \{\, |\eta_1\rangle, \ldots, |\eta_{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)}\rangle \,\}$$

are orthonormal sets of states. $\mathrm{Sch}(|\psi(0)\rangle, \mathcal{H}_A, \mathcal{H}_B)$ denotes the Schmidt rank of $|\psi(0)\rangle$. Thus

$$\begin{aligned}
\mathrm{tr}_B[H, \rho_{AB}(0)] &= \sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} (I \otimes \langle\eta_j|)\,[H, \rho_{AB}(0)]\,(I \otimes |\eta_j\rangle) \\
&= \sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} (I \otimes \langle\eta_j|)\,[X_A \otimes X_B, \rho_{AB}(0)]\,(I \otimes |\eta_j\rangle) \\
&= \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}\langle\eta_n|X_B|\eta_m\rangle X_A|\phi_m\rangle\langle\phi_n| \\
&\quad - \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}|\phi_m\rangle\langle\phi_n|X_A\langle\eta_n|X_B|\eta_m\rangle \\
&= \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}\langle\eta_n|X_B|\eta_m\rangle \Big[X_A, |\phi_m\rangle\langle\phi_n|\Big]
\end{aligned}$$

where

$$\rho_{AB}(0) = \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}(|\phi_m\rangle \otimes |\eta_m\rangle)(\langle\phi_n| \otimes \langle\eta_n|).$$

Since

$$\rho_A(0) = \sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \lambda_j |\phi_j\rangle\langle\phi_j|,$$

$$\log_2 \rho_A(0) = \sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} (\log_2 \lambda_j) |\phi_j\rangle\langle\phi_j|$$

we find

$$\begin{aligned}
\Gamma(t)|_{t\to 0} &= i\mathrm{tr}_{\mathrm{A}}(\mathrm{tr}_B([H, \rho_{AB}(0)] \log_2 \rho_A(0))) \\
&= i \sum_{j=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \langle\phi_j|\mathrm{tr}_B[H, \rho_{AB}(0)] \log_2 \rho_A(0)|\phi_j\rangle \\
&= i \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}(\log_2 \lambda_n)\langle\eta_n|X_B|\eta_m\rangle\langle\phi_n|X_A|\phi_m\rangle \\
&\quad -i \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}(\log_2 \lambda_m)\langle\eta_n|X_B|\eta_m\rangle\langle\phi_n|X_A|\phi_m\rangle \\
&= i \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}\left(\log_2 \frac{\lambda_n}{\lambda_m}\right)\langle\eta_n|X_B|\eta_m\rangle\langle\phi_n|X_A|\phi_m\rangle.
\end{aligned}$$

Taking the absolute value and applying the triangle inequality yields

$$\begin{aligned}
\Gamma(t)|_{t\to 0} &\leq \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \sqrt{\lambda_m\lambda_n}\left|\log_2 \frac{\lambda_n}{\lambda_m}\right| |\langle\eta_n|X_B|\eta_m\rangle||\langle\phi_n|X_A|\phi_m\rangle| \\
&= \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} (\lambda_m + \lambda_n)|\langle\eta_n|X_B|\eta_m\rangle||\langle\phi_n|X_A|\phi_m\rangle| \\
&\quad \times\sqrt{\frac{\lambda_n}{\lambda_m + \lambda_n}\frac{\lambda_m}{\lambda_m + \lambda_n}}\left|\log_2 \frac{\lambda_n}{\lambda_m}\right| \\
&\leq \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} (\lambda_m + \lambda_n)|\langle\eta_n|X_B|\eta_m\rangle||\langle\phi_n|X_A|\phi_m\rangle| \\
&\quad \times \max_{x\in(0,1)} \sqrt{x(1-x)}\log_2\left(\frac{x}{1-x}\right)
\end{aligned}$$

$$\leq 2 \max_{x\in(0,1)} \sqrt{x(1-x)} \log_2 \left(\frac{x}{1-x}\right)$$
$$\approx 1.9123$$

where we used

$$\sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} \lambda_m |\langle \eta_n | X_B | \eta_m \rangle| |\langle \phi_n | X_A | \phi_m \rangle|$$
$$\leq \sum_{m,n=1}^{\mathrm{Sch}(|\psi(0)\rangle,\mathcal{H}_A,\mathcal{H}_B)} |\langle \eta_n | X_B | \eta_m \rangle| |\langle \phi_n | X_A | \phi_m \rangle|$$
$$\leq 1,$$

since $X_A^2 = I$ and $X_B^2 = I$.
■

Example. Consider the Hamilton operator [Childs2003]

$$\hat{H} = \mu_x \sigma_x \otimes \sigma_x + \mu_y \sigma_y \otimes \sigma_y, \quad \mu_x, \mu_y \in \mathbf{R}$$

where σ_x and σ_y are Pauli spin matrices. Thus we have

$$\hat{H} = \begin{pmatrix} 0 & 0 & 0 & \mu_x - \mu_y \\ 0 & 0 & \mu_x + \mu_y & 0 \\ 0 & \mu_x + \mu_y & 0 & 0 \\ \mu_x - \mu_y & 0 & 0 & 0 \end{pmatrix}.$$

The eigenvalues are $\mu_x - \mu_y$ with corresponding eigenvector

$$|\phi^+\rangle = 1/\sqrt{2}(1,0,0,1)^T,$$

$\mu_y - \mu_x$ with corresponding eigenvector

$$|\phi^-\rangle = 1/\sqrt{2}(1,0,0,-1)^T,$$

$\mu_x + \mu_y$ with corresponding eigenvector

$$|\psi^+\rangle = 1/\sqrt{2}(0,1,1,0)^T$$

and $-\mu_x - \mu_y$ with corresponding eigenvector

$$|\psi^-\rangle = 1/\sqrt{2}(0,1,-1,0)^T.$$

The eigenvectors are the Bell states. Consider

$$|\psi_{max}\rangle := \begin{pmatrix} 0 \\ \sqrt{x_0} \\ -i\sqrt{1-x_0} \\ 0 \end{pmatrix}$$

$$= (\sqrt{x_0} - i\sqrt{1-x_0})\frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + (\sqrt{x_0} + i\sqrt{1-x_0})\frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

where $x_0 \in (0,1)$ yields the maximum α from

$$\alpha = 2\sqrt{x_0(1-x_0)}\log_2\left(\frac{x_0}{1-x_0}\right) = \max_{x\in(0,1)} 2\sqrt{x(1-x)}\log_2\left(\frac{x}{1-x}\right).$$

Thus

$$\exp(-i\hat{H}t)|\psi_{max}\rangle = \sum_{j=0}^{\infty}\frac{(-it)^j}{j!}\begin{pmatrix} 0 & 0 & 0 & \mu_x-\mu_y \\ 0 & 0 & \mu_x+\mu_y & 0 \\ 0 & \mu_x+\mu_y & 0 & 0 \\ \mu_x-\mu_y & 0 & 0 & 0 \end{pmatrix}^j |\psi_{max}\rangle$$

$$= \sum_{j=0}^{\infty}\frac{(-it)^{2j}(\mu_x+\mu_y)^{2j}}{(2j)!}\begin{pmatrix} 0 \\ \sqrt{x_0} \\ -i\sqrt{1-x_0} \\ 0 \end{pmatrix}$$

$$+ \sum_{j=0}^{\infty}\frac{(-it)^{2j+1}(\mu_x+\mu_y)^{2j+1}}{(2j+1)!}\begin{pmatrix} 0 \\ -i\sqrt{1-x_0} \\ \sqrt{x_0} \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ \sqrt{x_0}\cos(t(\mu_x+\mu_y)) - \sqrt{1-x_0}\sin(t(\mu_x+\mu_y)) \\ -i\left(\sqrt{x_0}\sin(t(\mu_x+\mu_y)) + \sqrt{1-x_0}\cos(t(\mu_x+\mu_y))\right) \\ 0 \end{pmatrix}.$$

Defining

$$a_1 := \sqrt{x_0}\cos(t(\mu_x+\mu_y)) - \sqrt{1-x_0}\sin(t(\mu_x+\mu_y))$$

$$a_2 := \sqrt{x_0}\sin(t(\mu_x+\mu_y)) + \sqrt{1-x_0}\cos(t(\mu_x+\mu_y))$$

$$\rho_{max}(t) := \exp(-i\hat{H}t)|\psi_{max}\rangle\langle\psi_{max}|\exp(i\hat{H}t)$$

we find

$$\rho_{max}(t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a_1^2 & ia_1\overline{a_2} & 0 \\ 0 & -ia_2\overline{a_1} & a_2^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathrm{tr}_{\mathbf{C}^2}(\rho_{max}(t)) = \begin{pmatrix} a_1^2 & 0 \\ 0 & a_2^2 \end{pmatrix}.$$

Thus

$$\begin{aligned}
\Gamma(t) &= -\frac{d}{dt}\left(a_1^2\log_2 a_1^2 + a_2^2\log_2 a_2^2\right) \\
&= -\left(4a_1\frac{da_1}{dt}\log_2 a_1 + \frac{2}{\ln 2}a_1\frac{da_1}{dt} + 4a_2\frac{da_2}{dt}\log_2 a_2 + \frac{2}{\ln 2}a_2\frac{da_2}{dt}\right) \\
&= 4(\mu_x+\mu_y)a_1a_2\left(\log_2\frac{a_1}{a_2}\right) \\
&= 4(\mu_x+\mu_y)\left(\sqrt{x_0(1-x_0)}\cos(2t(\mu_x+\mu_y)) + \frac{1}{2}(2x_0-1)\sin(2t(\mu_x+\mu_y))\right) \\
&\quad\times\log_2\left(\frac{\sqrt{x_0}\cos(2t(\mu_x+\mu_y)) - \sqrt{1-x_0}\sin(2t(\mu_x+\mu_y))}{\sqrt{x_0}\sin(2t(\mu_x+\mu_y)) + \sqrt{1-x_0}\cos(2t(\mu_x+\mu_y))}\right)
\end{aligned}$$

where we used that a_1 and a_2 satisfy the system of linear differential equations

$$\frac{da_1}{dt} = -a_2, \qquad \frac{da_2}{dt} = a_1.$$

Since $\hat{H}$ is asymptotically equivalent to $(\mu_x+\mu_y)\sigma_x\otimes\sigma_x$ and

$$E((\mu_x+\mu_y)\sigma_x\otimes\sigma_x) \leq (\mu_x+\mu_y)\alpha$$

and using that

$$\Gamma(0) = (\mu_x+\mu_y)2\sqrt{x_0(1-x_0)}\log_2\frac{x_0}{1-x_0} = \alpha(\mu_x+\mu_y),$$

we find $E(\hat{H}) = \alpha(\mu_x + \mu_y)$.
■

Chapter 3

Quantum Communication and Complexity

3.1 Quantum Communication Complexity Without Entanglement

Since classical computation and communication is a subset of the model of quantum computation and communication, results from classical communication complexity transfer also to quantum communication complexity under restricted domains of computation and computational processes.

Consider the Hilbert spaces $\mathcal{H}_A := \mathbf{C}^{2^n}$, $\mathcal{H}_B := \mathbf{C}^{2^m}$, and $\mathcal{H}_C := \mathbf{C}^{2^k}$. The dimensions are powers of 2 so that $\mathcal{H}_A$ can represent n classical bits, $\mathcal{H}_B$ can represent m classical bits and $\mathcal{H}_C$ can represent k classical bits. The Hilbert space $\mathcal{H}_C$ describes the quantum system used for the communication channel.

Definition. A *quantum protocol* is a sequence of unitary operators

$$U_{BC,1}, U_{AC,2}, U_{BC,3}, \ldots, U_{AC,l}$$

where $U_{AC,j}$ is a unitary operator on $\mathcal{H}_A \otimes \mathcal{H}_C$ and $U_{BC,j}$ is a unitary operator on $\mathcal{H}_C \otimes \mathcal{H}_B$. The length or cost of the protocol is l. The sequence does not necessarily end on $U_{AC,l}$, sequences with odd length will end on $U_{BC,l}$.
■

Definition. The *quantum communication complexity* [Kremer95] $Q_\epsilon(f)$ of $f : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}$ is the minimum length l over all protocols such that

$$(U_{AC,l} \otimes I)(I \otimes U_{BC,l-1}) \cdots (U_{AC,2} \otimes I)(I \otimes U_{BC,1})(|x\rangle \otimes |0\rangle \otimes |y\rangle)$$

yields the measurement outcome 1 with probability greater than $1 - \epsilon$ with respect to the orthogonal measurement given by the observable

$$O_{f(x,y)} := (-1)^{f(x,y)}(I_{2^n} \otimes |0\rangle\langle 0| \otimes I_{2^{m+k-1}}) \\ -(-1)^{f(x,y)}(I_{2^n} \otimes |1\rangle\langle 1| \otimes I_{2^{m+k-1}})$$

for all $x \in \{0,1\}^n$ and $y \in \{0,1\}^m$.
■

We have the following theorems [Kremer95].

Theorem. $Q_{\frac{1}{3}}(f) \leq kC_{\frac{1}{3}}(f)$ for some constant $k \in \mathbf{R}$ with $k > 0$.
■

Theorem. Let

$$P := (U_{AC,l} \otimes I)(I \otimes U_{BC,l-1}) \cdots (U_{AC,2} \otimes I)(I \otimes U_{BC,1})$$

be a protocol of length l. Then $P(|x\rangle \otimes |0\rangle \otimes |y\rangle)$ for $x \in \{0,1\}^n$ and $y \in \{0,1\}^m$ can be written in the form

$$P(|x\rangle \otimes |0\rangle \otimes |y\rangle) = \sum_{j=(j_0,j_1,\ldots,j_{n-1}) \in \{0,1\}^l} |v_{A,j}\rangle \otimes |j_0\rangle \otimes |v_{B,j}\rangle$$

where $|v_{A,j}\rangle \in \mathcal{H}_A$ and $|v_{B,j}\rangle \in \mathcal{H}_B$ and j_0 is the value of the lowest order bit of j. Furthermore, $|v_{A,j}\rangle$ is determined exclusively by the protocol and x and $|v_{B,j}\rangle$ is determined exclusively by the protocol and y.
■

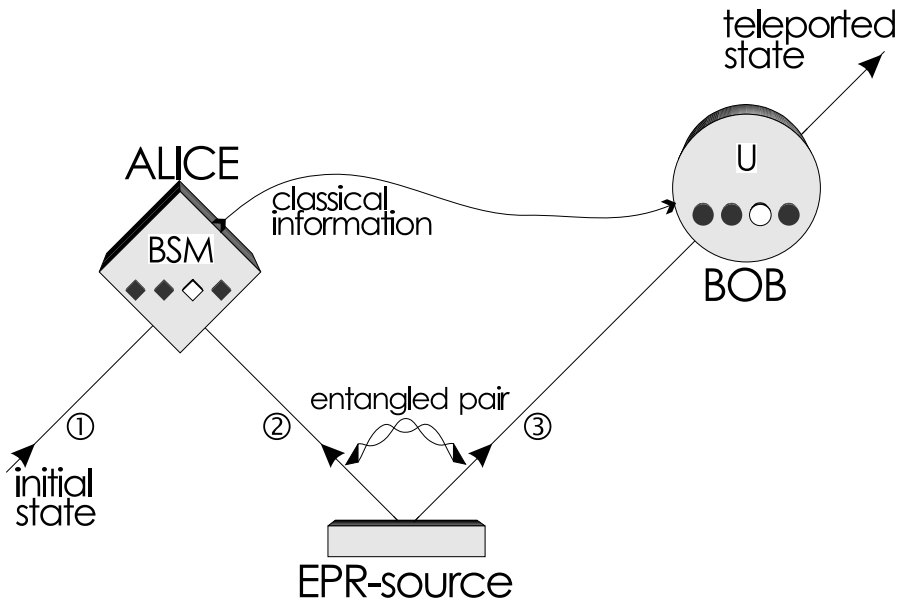
Thus the quantum communication complexity with no error is bounded above by the Schmidt rank over $\mathcal{H}_A \otimes (\mathcal{H}_C \otimes \mathcal{H}_B)$ of $P(|x\rangle \otimes |0\rangle \otimes |y\rangle)$ for all $x \in \{0,1\}^n$ and $y \in \{0,1\}^m$.

It is useful to introduce entanglement as a resource for quantum communication, this changes the nature of the communication and allows the use of quantum correlation to determine some properties without communication.

The use of entanglement in discussed in the following sections.

3.2 Teleportation

Quantum teleportation is the disembodied transport of an unknown quantum state $|\psi\rangle$ from one place to another. All protocols for accomplishing such transport require nonlocal correlations, or entanglement, between systems shared by sender and receiver. The sender is normally called *Alice* and the receiver is called *Bob*. Most attention has focused on teleporting the states of finite-dimensional systems, such as the two-dimensional polarization of a photon or the discrete level structure of an atom. First proposed in 1993 by Charles Bennett and his colleagues [Bennett93, Bouwmeester97, Rieffel98] quantum teleporation thus allows physicists to take a photon or any other quantum scale particle such as an atom and transfer its properties (such as the polarization) to another photon even if the two photons are on opposite sides of the galaxy. This scheme transports the particle's properties to the remote location and not the particle itself. The state of the original particle must be destroyed to create an exact reconstruction at the other end. This is a consequence of the no cloning theorem. A role in the teleportation scheme is played by an entangled ancillary pair of particles which will be initially shared by Alice and Bob. Thus it is possible to communicate quantum information using classical communication when prior entanglement is available as a resource.



3.2.1 Teleportation Algorithm

Suppose particle 1 (in the following we assume that it is a spin-$\frac{1}{2}$ particle) which Alice wants to teleport is in the initial state

$$|\phi\rangle_1 = a|0\rangle_1 + b|1\rangle_1$$

and the entangled pair of particles 2 and 3 shared by Alice and Bob is in the state

$$|\phi\rangle_{23} = \frac{1}{\sqrt{2}}(|0\rangle_2 \otimes |1\rangle_3 - |1\rangle_2 \otimes |0\rangle_3)\,.$$

Alice gets particle 2 and Bob particle 3. This entangled state contains no information on the individual particles 2 and 3. It only indicates that the two particles will be in opposite states. Alice then performs a joint Bell-state measurement on the initial particle 1 and particle 2 projecting them also onto an entangled state. After Alice has sent the result of her measurement as classical information to Bob, he can perform a unitary transformation on the other ancillary particle resulting in it being in the state of the original particle.

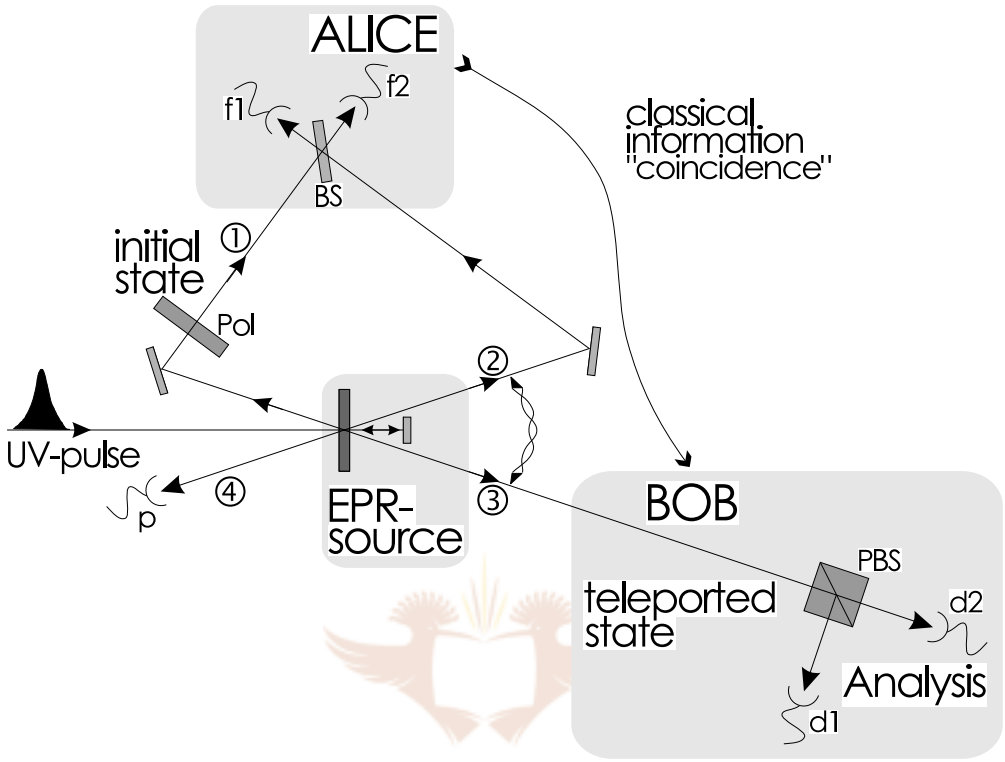
Most experiments are done with photons which are spin-1 particles. The information to be teleported is the polarization state of the photon. The *Innsbruck experiment* is a simplified version of the teleportation described above. In this experiment photons are used. The photon is a particle with spin 1 and rest mass 0. If the photon moves in positive z direction, i.e. the wave vector $\mathbf{k}$ is given by $(0, 0, k)^T$ we have the wave functions

$$\phi_1(z,t) = (2V)^{-1/2}(-\mathbf{e}_1 - i\mathbf{e}_2)\exp(i(kz - \omega t))$$

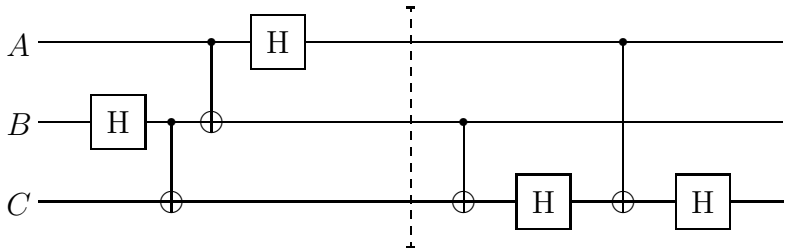
$$\phi_{-1}(z,t) = (2V)^{-1/2}(\mathbf{e}_1 - i\mathbf{e}_2)\exp(i(kz - \omega t))$$

where $\mathbf{e}_1 := (1, 0, 0)^T$ and $\mathbf{e}_2 := (0, 1, 0)^T$. Thus we have two transverse waves. Although the photon is a spin-1 particle the vectors $\mathbf{s}$ and $\mathbf{k}$ can only be parallel (or antiparallel). The wave ϕ_1 is in a state of positive helizity and the wave ϕ_{-1} is in a state of negative helizity. In the *Innsbruck experiment* at the sending station of the quantum teleporter, Alice encodes photon M with a specific state: 45 degree polarization. This photon travels towards a *beamsplitter*. A beamsplitter partially transmits and partially reflects the photons with a resultant change in phase. A 50/50 beamsplitter can be in-

terpreted as a Hadamard transform. Meanwhile two additional entangled photons A and B are created. Thus they have complementary polarizations. For example, if photon A is later measured to have horizontal (0 degrees) polarization, then the other photon B must collapse into the complementary state of vertical (90 degrees) polarization. Now entangled photon A arrives at the beamsplitter at the same time as the message photon M. The beamsplitter causes each photon either to continue towards detector 1 or change course and travel to detector 2. in 1/4 of all cases, in which the two photons go off into different detectors, Alice does not know which photon went to which detector. Owing to the fact that the two photons are now indistinguishable, the message photon M loses its original identity and becomes entangled with A. The polarization value for each photon is now indeterminate, but since the two photons travel towards different detectors Alice knows that the two photons must have complementary polarizations. Since message particle M must have complementary polarization to particle A, then the other entangled particle B must now attain the same polarization value as M. Therefore teleportation is successful. The receiver Bob sees that the polarization value of the particle B is 45 degrees, which is the initial value of the message photon. In the experimental version of this setup executed at the University of Innsbruck, the 45-degree polarization would always fire when detector 1 and detector 2 fired. Except in rare instances attributable to background noise, it was never the case that the 135-degree polarization detector fired in coincidence with detectors 1 and 2.



Teleportation can also be understood using the quantum circuit shown in the following figure.



In the figure A is the input $|\psi\rangle$, B the input $|0\rangle$ and C the input $|0\rangle$. Now we study what happens when we feed the product state (input)

$$|\psi\rangle \otimes |0\rangle \otimes |0\rangle$$

into the quantum circuit. From the circuit we have the following eight 8×8

unitary matrices

$$U_1 = I_2 \otimes U_H \otimes I_2, \qquad U_2{=}I_2 \otimes U_{XOR}, \qquad U_3{=}U_{XOR} \otimes I_2,$$
$$U_4 = U_H \otimes I_2 \otimes I_2 \qquad U_5{=}I_2 \otimes U_{XOR}, \qquad U_6{=}I_2 \otimes I_2 \otimes U_H,$$
$$U_7 = I_4 \oplus U_{NOT} \oplus U_{NOT}, \quad U_8{=}I_2 \otimes I_2 \otimes U_H$$

where $\oplus$ denotes the direct sum of matrices [SteebMC] and

$$U_{NOT} := |0\rangle\langle 1| + |1\rangle\langle 0|, \qquad U_{XOR} := |0\rangle\langle 0| \otimes I_2 + |1\rangle\langle 1| \otimes U_{NOT}\,.$$

Applying the first four unitary matrices $U_4U_3U_2U_1$ to the input state we obtain

$$\frac{a}{2}(|000\rangle + |100\rangle + |011\rangle + |111\rangle) + \frac{b}{2}(|010\rangle - |110\rangle + |001\rangle - |101\rangle)\,.$$

This state can be rewritten as

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes (\frac{a}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle))$$
$$+\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \otimes (\frac{b}{\sqrt{2}}(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle))$$

Applying all eight unitary matrices $U_8U_7U_6U_5U_4U_3U_2U_1$ to the input state we obtain

$$\frac{a}{2}(|000\rangle + |100\rangle + |010\rangle + |110\rangle) + \frac{b}{2}(|011\rangle + |111\rangle + |001\rangle + |101\rangle)\,.$$

This state can be rewritten as

$$\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes |\psi\rangle.$$

Thus the state $|\psi\rangle$ will be transferred to the lower output, where both other outputs will come out in the state $(|0\rangle + |1\rangle)/\sqrt{2}$. If the two upper outputs are measured in the standard basis ($|0\rangle$ versus $|1\rangle$), two random classical bits will be obtained in addition to the quantum state $|\psi\rangle$ on the lower output.

Consider the case when the the qubit to be teleported is one qubit of an entangled pair. The first two qubits are entangled. Applying the teleportation

algorithm to the second, third and fourth qubits yields

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \otimes |00\rangle \rightarrow \frac{1}{\sqrt{2}}(|0000\rangle + |1001\rangle).$$

The first and last qubits are now entangled, whereas the first and second are no longer entangled. Thus we have achieved *entanglement swapping*.

3.2.2 Teleportation by Measurement

Consider the following states (a, $b \in \mathbf{C}$)

$$|\psi\rangle := a|0\rangle + b|1\rangle, \quad |a|^2 + |b|^2 = 1$$

$$|\phi\rangle := |\psi\rangle \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)\,.$$

On the other hand we have

$$\begin{aligned}
|\phi\rangle = & \frac{1}{2\sqrt{2}}(a|000\rangle + a|110\rangle + b|001\rangle + b|111\rangle) \\
& + \frac{1}{2\sqrt{2}}(a|000\rangle - a|110\rangle - b|001\rangle + b|111\rangle) \\
& + \frac{1}{2\sqrt{2}}(a|011\rangle + a|101\rangle + b|010\rangle + b|100\rangle) \\
& + \frac{1}{2\sqrt{2}}(a|011\rangle - a|101\rangle - b|010\rangle + b|100\rangle) \\
= & \frac{1}{2\sqrt{2}}(|00\rangle + |11\rangle) \otimes (a|0\rangle + b|1\rangle) + \frac{1}{2\sqrt{2}}(|00\rangle - |11\rangle) \otimes (a|0\rangle - b|1\rangle) \\
& + \frac{1}{2\sqrt{2}}(|01\rangle + |10\rangle) \otimes (a|1\rangle + b|0\rangle) + \frac{1}{2\sqrt{2}}(|01\rangle - |10\rangle) \otimes (a|1\rangle - b|0\rangle)\,.
\end{aligned}$$

We measure in the *Bell basis*

$$\left\{\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),\ \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),\ \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle),\ \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\right\}.$$

From the state $|\phi\rangle$ we can see that the first two qubits are in each of the Bell states with equal probability. Thus if we measure we obtain a result corresponding to each of the Bell states and can perform a transform to obtain $|\psi\rangle$ in the last qubit as follows

Bell State	Transform
$\frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	I_2
$\frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$	$\|0\rangle\langle 0\| - \|1\rangle\langle 1\|$
$\frac{1}{\sqrt{2}}(\|01\rangle + \|10\rangle)$	U_{NOT}
$\frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	$\|0\rangle\langle 1\| - \|1\rangle\langle 0\|$

After measurement and applying the corresponding transform we obtain $|\psi\rangle$ as the last qubit. So if Alice and Bob initially share the entangled pair

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Alice can perform a measurement in the Bell basis on her qubit and her part of the entangled pair and sends the result (two bits) to Bob who applies the corresponding transform to his part of the entangled pair. The state $|\psi\rangle$ is thus *teleported* from Alice's qubit to Bob's qubit. The Bell basis is obtained by applying the unitary operator $U_{CNOT}(U_H \otimes I_2)$ to the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. The transforms are unitary and therefore invertible. Thus we can also measure the first two qubits in the computational basis after applying

$$(U_H \otimes I_2)U_{CNOT}.$$

3.3 Dense Coding

Dense coding [Bennett92, Rieffel98] uses an EPR pair and a single qubit to transmit two classical bits of information. The EPR pair is shared ahead of time between the transmitter and receiver. Only a single qubit is needed to transfer the two classical bits of information. This is possible due to entanglement. Entanglement makes it possible to transform a two qubit state to one of four orthogonal states by interacting with only one qubit. This choice can be identified with two classical bits of information. Dense

coding illustrates the use of entanglement as an information resource. The scheme is also secure in the sense that communication is only possible between the the two parties which share the EPR pair.

The transmitter (Alice) and the receiver (Bob) each have one quantum subsystem which together form the quantum system of an already prepared EPR state

$$|\psi\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle).$$

We let the first system denote Alice's quantum subsystem (qubit) of the EPR state and the second system denote Bob's quantum subsystem (qubit) of the EPR state. Alice can transform $|\psi\rangle$ to any one of the Bell basis states according to

$$\begin{aligned}
\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes I_2\right)|\psi\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv \Phi^+ \\
\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes I_2\right)|\psi\rangle &= \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) \equiv \Psi^+ \\
\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes I_2\right)|\psi\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \equiv \Phi^- \\
\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes I_2\right)|\psi\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \equiv \Psi^-
\end{aligned}$$

where I_2 is the 2×2 unit matrix.

Alice has two bits representing the values 0,1,2 or 3. She transforms $|\psi\rangle$ according to the following table.

Value	Initial state	Transformed state
0	$\|\psi\rangle$	$\frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$
1	$\|\psi\rangle$	$\frac{1}{\sqrt{2}}(\|10\rangle + \|01\rangle)$
2	$\|\psi\rangle$	$\frac{1}{\sqrt{2}}(-\|10\rangle + \|01\rangle)$
3	$\|\psi\rangle$	$\frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$

The transformations are obviously unitary. Alice then sends her quantum qubit to Bob. Now Bob applies a controlled NOT using the first (Alice's)

qubit as the control and then applies the Walsh-Hadamard transform to the first qubit. Finally a controlled NOT is applied to yield the data. The following table describes the quantum state after the transformation.

Value	Initial state	Transformed state	Final state
0	$\frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	$\|00\rangle$	$\|00\rangle$
1	$\frac{1}{\sqrt{2}}(\|10\rangle + \|01\rangle)$	$\|01\rangle$	$\|01\rangle$
2	$\frac{1}{\sqrt{2}}(-\|10\rangle + \|01\rangle)$	$\|11\rangle$	$\|10\rangle$
3	$\frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$	$\|10\rangle$	$\|11\rangle$

Thus the transformed state uniquely determines the value 0, 1, 2 or 3. Consequently it is possible to communicate two bits of classical information when prior entanglement is available as a resource.

3.4 Quantum Two-party Communication Complexity

We can describe some of the measures using notation from quantum mechanics and quantum computing [HardyQC, SteebQM], provided $Z = \{0, 1\}$. Suppose that $\{|a\rangle : \ a \in A\}$ and $\{|b\rangle : \ b \in B\}$ are orthonormal basis for the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Thus we can write the communication matrix as

$$M_f = \sum_{a \in A, b \in B} f(a, b)|a\rangle\langle b|.$$

We also find that the upper bound on the expected communication complexity obtained previously can be written as

$$C_0^E(f, u_R) = 1 - \mathrm{tr}(\rho_C \log_2 \rho_C)$$

where the density matrix ρ_C is given by (the definition of p_T is given in chapter 1)

$$\rho_C := \sum_{a \in A} p_T(f, a)|a\rangle\langle a|$$

is essentially a classical mixture.

A consequence of Holevo's bound [Holevo98] is that the communication of quantum systems, in our case *qubits* (two-level quantum states), cannot be used to communicate any more information than is possible classically. However if entanglement is already shared between the two parties, it is possible to decrease the communication required. For example, using superdense coding [Nielsen, HardyQC, Barenco95] we can reduce the communication by a factor of 2. In other words we can communicate half of the information before we know what we want to communicate. Since the first half of the communication is independent of the function to be computed, we choose to ignore this communication in our measure. Thus, for example [Nielsen]

$$C_0(f_{IP}) = \left\lceil \frac{n}{2} \right\rceil$$

when Alice and Bob share n entangled pairs prior to obtaining a and b, where $\lceil x \rceil$ denotes the ceiling, i.e. the greatest integer less than x.

3.5 Substituting Entanglement for Classical Communication

In this section we describe a few examples from the literature [Cleve, Buhrman97b] which illustrate that entanglement may be substituted for classical communication and thus reduce the communication complexity.

Example. Consider $x_A, x_B, x_C \in \{0,1\}$ such that $x_A + x_B + x_C = 1 \ mod\ 2$. We wish to calculate $x_A \cdot x_B \cdot x_C$. Let $SU(2)$ be the special unitary group defined by

$$SU(2) = \{U \in M(2) \,:\, \det(U) = 1,\, \langle Ux, Uy\rangle = \langle x, y\rangle \ \ \forall x, y \in \mathbf{C}^2\}.$$

Using the mapping $f : \{0,1\} \to SU(2)$

$$f_1(0) := U_H \qquad f_1(1) := I_2$$

where U_H is the Walsh-Hadamard transform, we define the map from the triple (x_A, x_B, x_C) to linear operators acting on three qubits

$$f_3(x_A, x_B, x_C) := f_1(x_A) \otimes f_1(x_B) \otimes f_1(x_C).$$

Let

$$|\psi\rangle := \frac{1}{2}(|001\rangle + |010\rangle + |100\rangle - |111\rangle).$$

For each triple (x_A, x_B, x_C) we calculate

$$|\phi\rangle := f_3(x_A, x_B, x_C)|\psi\rangle.$$

Let s_A, s_B, s_C denote the result (0 or 1) of measuring the first, second and third qubit of $|\phi\rangle$ respectively in the computational basis. We have

$$(x_A, x_B, x_C) \in \{\, (0,0,1), (0,1,0), (1,0,0), (1,1,1) \,\}.$$

We note the symmetry in the state $|\psi\rangle$. A qubit can be exchanged with any other without our being able to determine any difference. Thus we need only to calculate the transform for $(0,0,1)$ and $(1,1,1)$. For $(1,1,1)$ we have $f_3(1,1,1)|\psi\rangle = I_2 \otimes I_2 \otimes I_2|\psi\rangle = |\psi\rangle$. Measuring the qubits yields

$$(s_A, s_B, s_C) \in \{\, (0,0,1), (0,1,0), (1,0,0), (1,1,1) \,\}$$

with equal probability. In each case we find $s_A + s_B + s_C = 1 \ mod\ 2$. For $(0,0,1)$ we have $f_3(0,0,1) = U_H \otimes U_H \otimes I_2$. We have

$$|\psi\rangle = \frac{1}{2}(|01\rangle + |10\rangle) \otimes |0\rangle + \frac{1}{2}(|00\rangle - |11\rangle) \otimes |1\rangle.$$

Thus

$$f_3(0,0,1)|\psi\rangle = \frac{1}{2}(|00\rangle - |11\rangle) \otimes |0\rangle + \frac{1}{2}(|01\rangle + |10\rangle) \otimes |1\rangle.$$

We find that measuring the qubits yields

$$(s_A, s_B, s_C) \in \{\, (0,0,0), (0,1,1), (1,0,1), (1,1,0) \,\}$$

with equal probability. In each case we find $s_A + s_B + s_C = 0 \ mod\ 2$. Thus we find that $s_A + s_B + s_C = x_A \cdot x_B \cdot x_C \ mod\ 2$. Suppose Alice, Bob and Carol each have a bit string $(x_{A,1}, \ldots, x_{A,n})$, $(x_{B,1}, \ldots, x_{B,n})$ and $(x_{C,1}, \ldots, x_{C,n})$, respectively. They want to calculate

$$f(\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C) = \sum_{j=1}^{n} x_{A,j} \cdot x_{B,j} \cdot x_{C,j} \ mod\ 2$$

sharing (communicating) as little information as possible. If Alice, Bob and Carol share n triplets of qubits in the state $|\psi\rangle$ they can calculate $s_{A,1}, \ldots, s_{A,n}$, $s_{B,1}, \ldots, s_{B,n}$ and $s_{C,1}, \ldots, s_{C,n}$ respectively as above. Thus

$$f(\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C) = \sum_{j=1}^{n} (s_{A,j} + s_{B,j} + s_{C,j}) \ mod\ 2.$$

If Alice, Bob and Carol calculate

$$S_{A|B|C} = \sum_{j=1}^{n} S_{A|B|C,j} \ mod\ 2,$$

Bob and Carol need only to send one bit each (S_B and S_C) to Alice for Alice to compute $f(\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C) = S_A + S_B + S_C$, for any n. In other words the communication complexity is 2. Classically, for $n \geq 3$, 3 bits of communication is required [Cleve].

■

Example. We consider $x, y, z \in \{0, 1, 2, 3\}$ such that $x + y + z = 0 \ mod\ 2$. We wish to determine

$$f(x, y, z) := \frac{(x + y + z) \ mod\ 4}{2}.$$

When $x + y + z$ is even, $(x + y + z) \ mod\ 4 \ \in \{0, 2\}$. Now when $x + y + z = 0 \ mod\ 2$ then $f(x, y, z) \in \{0, 1\}$. Using the binary representation for $x = x_1 x_0$, $y = y_1 y_0$ and $z = z_1 z_0$ where $x_0, x_1, y_0, y_1, z_0, z_1 \in \{0, 1\}$ Since $x + y + z = 0 \ mod\ 2$ the least significant bit of the sum must be zero. The least significant bit is given by $x_0 \oplus y_0 \oplus z_0 = 0$. Thus we have

$$(x_0, y_0, z_0) \in \{\, (0,0,0), (0,1,1), (1,0,1), (1,1,0) \,\}.$$

We find that

$$f(x, y, z) = x_1 \oplus y_1 \oplus z_1 \oplus (x_0 + y_0 + z_0).$$

XOR is denoted by "$\oplus$" and OR is denoted by "$+$". We use the mapping

$$f_1(0) = I_2 \qquad f_1(1) = U_H.$$

Thus we can map from the triple (x_0, y_0, z_0) to linear operators acting on three qubits

$$f_3(x_0, y_0, z_0) = f_1(x_0) \otimes f_1(y_0) \otimes f_1(z_0).$$

Let

$$|\psi\rangle := \frac{1}{2}(|000\rangle - |011\rangle - |101\rangle - |110\rangle).$$

For each triple (x_0, y_0, z_0) we determine

$$|\phi\rangle := f_3(x_0, y_0, z_0)|\psi\rangle.$$

Let s_x, s_y, s_z denote the result (0 or 1) of measuring the first, second and third qubit of $|\phi\rangle$ respectively in the computational basis. We note the symmetry in the state $|\psi\rangle$. A qubit can be exchanged with any other without our being able to determine any difference. Thus we need only calculate the transform for $(0, 0, 0)$ and $(0, 1, 1)$. For $(0, 0, 0)$ we have $f_3(0, 0, 0)|\psi\rangle = I_2 \otimes I_2 \otimes I_2|\psi\rangle = |\psi\rangle$. Measuring the qubits yields

$$(s_x, s_y, s_z) \in \{\, (0, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0) \,\}$$

with equal probability. In each case we find $s_x + s_y + s_z = 0 \ mod\ 2$. For $(0, 1, 1)$ we have $f_3(0, 1, 1) = I_2 \otimes U_H \otimes U_H$. Note that

$$|\psi\rangle = \frac{1}{2}|0\rangle \otimes (|00\rangle - |11\rangle) - \frac{1}{2}|1\rangle \otimes (|01\rangle + |10\rangle).$$

Thus

$$f_3(0, 1, 1)|\psi\rangle = \frac{1}{2}|0\rangle \otimes (|01\rangle + |10\rangle) - \frac{1}{2}|1\rangle \otimes (|00\rangle - |11\rangle).$$

We find that measuring the qubits yields

$$(s_x, s_y, s_z) \in \{\, (0, 1, 0), (1, 0, 0), (0, 0, 1), (1, 1, 1) \,\}$$

with equal probability. In each case we find $s_x + s_y + s_z = 1 \ mod\ 2$. We find that $(s_x + s_y + s_z \ mod\ 2) = x_0 + y_0 + z_0$. Thus for three parties to calculate $f(x, y, z)$, where each party has one of the x, y and z, it is sufficient for each party to send one bit ($x_1 \oplus s_x$ or $y_1 \oplus s_y$ or $z_1 \oplus s_z$) to the other parties to calculate $f(x, y, z)$. In other words each party can calculate

$$x_1 \oplus s_x \oplus y_1 \oplus s_y \oplus y_1 \oplus s_y \oplus z_1 \oplus s_z = x_1 \oplus y_1 \oplus z_1 \oplus (x_0 + y_0 + z_0)$$

$$= f(x, y, z)$$

after communication. In other words three bits broadcast to all parties are sufficient to calculate $f(x, y, z)$, the communication complexity is 3 bits. Classically 4 bits are necessary to be broadcast [Buhrman97b].
■

3.6 Deutsch-Jozsa Relation

3.6.1 Deutsch's Problem

Deutsch's problem[Deutsch92, HardyQC] is given as follows. Suppose we have a function

$$f : \{0, 1\} \rightarrow \{0, 1\}.$$

There are four such functions, the constant functions which map all inputs to 0 or all inputs to 1, and the varying functions which have $f(0) \neq f(1)$. In other words

$f_1(0) = 0,$	$f_1(1) = 0$	constant
$f_2(0) = 1,$	$f_2(1) = 1$	constant
$f_3(0) = 0,$	$f_3(1) = 1$	varying
$f_4(0) = 1,$	$f_4(1) = 0$	varying

The task is to determine for such a function if it is constant or varying using only one calculation of the function. In the classical case it is necessary to compute f twice before it is known whether it is constant or varying. For example if $f(0) = 0$, the function could be f_1 or f_3, similarly for any other single evaluation two of the functions have the same value.

In quantum computing the following solution was found [Vedral98]. The function is implemented on quantum hardware with the unitary transformation U_f such that

$$U_f|x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle$$

where $\oplus$ denotes the XOR operation. We apply the transformation to the state

$$|\psi\rangle := \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle).$$

This gives

$$U_f|\psi\rangle = \frac{1}{2}(|0\rangle \otimes |0 \oplus f(0)\rangle - |0\rangle \otimes |1 \oplus f(0)\rangle) + \frac{1}{2}(|1\rangle \otimes |0 \oplus f(1)\rangle - |1\rangle \otimes |1 \oplus f(1)\rangle).$$

Since U_f is linear and quantum mechanics allows the superposition of states we have calculated the function f with each input twice. This feature of quantum parallelism makes it possible to solve the problem. Applying the Walsh-Hadamard transform U_H to the first qubit yields

$$\begin{aligned}(U_H \otimes I_2)U_f|\psi\rangle = \frac{1}{2\sqrt{2}}(|0\rangle \otimes (\quad & |0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle \\ & + |0 \oplus f(1)\rangle - |1 \oplus f(1)\rangle)) \\ + \frac{1}{2\sqrt{2}}(|1\rangle \otimes (& |0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle \\ & - |0 \oplus f(1)\rangle + |1 \oplus f(1)\rangle)).\end{aligned}$$

If f is constant we have $f(0) = f(1)$ and we find

$$(U_H \otimes I_2)U_f|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes (|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle).$$

If f is varying we have

$$0 \oplus f(0) = 1 \oplus f(1),$$

$$1 \oplus f(0) = 0 \oplus f(1)$$

and we find

$$(U_H \otimes I_2)U_f|\psi\rangle = \frac{1}{\sqrt{2}}|1\rangle \otimes (|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle).$$

Thus measuring the first qubit as zero indicates the function f is constant and measuring the first qubit as one indicates the function f is varying. The function f was only calculated once using U_f and so the problem is solved.

3.6.2 Deutsch-Jozsa Problem

The generalization of Deutsch's problem is the Deutsch-Jozsa problem. Given $f : \{0,1\}^n \to \{0,1\}$ we need to determine whether f is constant or maps to an equal number (2^{n-1}) of 0's and 1's. Furthermore it is given that

$$\Big|\{x \in \{0,1\}^n \,:\, f(x) = 1\}\Big| \in \{0, 2^{n-1}, 2^n\}.$$

Now we use

$$U_f|x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle$$

where $x \in \{0,1\}^n$ and $y \in \{0,1\}$. Starting with

$$\begin{aligned}|\phi\rangle &:= \left(\overbrace{U_H \otimes U_H \otimes \cdots \otimes U_H}^{n+1 \text{ times}}\right) \overbrace{|0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle}^{n \text{ times}} \otimes |1\rangle \\ &= \left(\frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle\right) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\end{aligned}$$

and then applying U_f yields

$$\begin{aligned}U_f|\phi\rangle &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} |x\rangle \otimes \Big(|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle\Big) \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)}|x\rangle \otimes (|0\rangle - |1\rangle).\end{aligned}$$

The probability P of finding $U_f|\phi\rangle$ in the state

$$\left(\overbrace{U_H \otimes U_H \otimes \cdots \otimes U_H}^{n+1 \text{ times}}\right) \overbrace{|0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle}^{n \text{ times}} \otimes |1\rangle$$

is given by

$$P = \left| \overbrace{\langle 0| \otimes \langle 0| \otimes \cdots \otimes \langle 0|}^{n \text{ times}} \otimes \langle 1| \left(\overbrace{U_H \otimes U_H \otimes \cdots \otimes U_H}^{n+1 \text{ times}} \right) \times \right.$$

$$\left. \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)} |x\rangle \otimes (|0\rangle - |1\rangle) \right|^2$$

$$= \left| \frac{1}{2^n} \sum_{x,y \in \{0,1\}^n} (-1)^{f(x)} \langle x|y\rangle \right|^2$$

$$= \left| \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \right|^2 .$$

Thus when f is constant we have $P = 1$ and when f is alternating $P = 0$. Consequently measuring the first n qubits as 0 (n bits) corresponds to f constant and nonzero corresponds to f mapping to an equal number of zeros and ones.

3.6.3 Communication Problem

The communication problem corresponding to the Deutsch-Jozsa problem is described by the Deutsch-Jozsa relation [Brassard, Brassard99]. Alice must compute $f_A(x, y)$ and Bob must compute $f_B(x, y)$ where $x, y \in \{0, 1\}^n$ and $f_A : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}^k$, $f_B : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}^k$. The functions f_A and f_B are constrained by

$$f_A(x, y) = f_B(x, y) \qquad \Leftrightarrow \qquad x = y.$$

If we further insist that x and y are equal or differ in exactly half of the bits we have the Deutsch-Jozsa relation. The task is to achieve k exponentially less than $n = 2^k$, i.e. k should be of order $\log_2 n$.

It is known [Brassard99] that there exists a positive constant c such that (classically) the relation cannot be established with fewer than $k = cn$ bits of communication for sufficiently large n. Thus the relation $k = O(\log_2 n)$ cannot be satisfied in the classical case.

However, suppose Alice prepares $2k = 2\log_2 n$ qubits in the state

$$|\psi\rangle := \frac{1}{\sqrt{2^{2k}}} \sum_{a=(a_0,a_1,\ldots,a_{k-1})\in\{0,1\}^k} |a_0\rangle \otimes \ldots \otimes |a_{k-1}\rangle \otimes |a_0\rangle \otimes \ldots \otimes |a_{k-1}\rangle$$

$$= \frac{1}{\sqrt{2^{2k}}} \sum_{a\in\{0,1\}^k} |a\rangle \otimes |a\rangle.$$

We denote by $N : \{0,1\}^n \to \mathbf{N}$

$$N(a = (a_0, a_1, \ldots, a_{k-1})) = \sum_{j=0}^{k-1} a_j 2^j$$

the natural number corresponding to the binary representation a. Now Alice and Bob both obtain

$$x = (x_0, x_1, \ldots, x_{n-1}) \in \{0,1\}^n \quad \text{and} \quad y = (y_0, y_1, \ldots, y_{n-1}) \in \{0,1\}^n.$$

Let

$$F_j(x) := \sum_{a\in\{0,1\}^k} (-1)^{x\delta_{jN(a)}} |a\rangle\langle a|.$$

Alice then applies $F_A := F_0(x_0)F_1(x_1)\ldots F_{n-1}(x_{n-1})$ and similarly Bob applies $F_B := F_0(y_0)F_1(y_1)\ldots F_{n-1}(y_{n-1})$ to obtain

$$(F_A \otimes F_B)|\psi\rangle = \frac{1}{\sqrt{2^{2k}}} \sum_{a\in\{0,1\}^k} F_A|a\rangle \otimes F_B|a\rangle$$

$$= \frac{1}{\sqrt{2^{2k}}} \sum_{a\in\{0,1\}^k} (-1)^{x_{N(a)}}|a\rangle \otimes (-1)^{y_{N(a)}}|a\rangle$$

$$= \frac{1}{\sqrt{2^{2k}}} \sum_{a\in\{0,1\}^k} (-1)^{x_{N(a)}+y_{N(a)}}|a\rangle \otimes |a\rangle.$$

Finally Alice and Bob apply the Walsh-Hadamard transform to each of their qubits:

$$|\phi\rangle := \left(\bigotimes_{l=1}^{2k} U_H\right)(F_A \otimes F_B)|\psi\rangle$$

$$= \frac{1}{\sqrt{2^{2k}}} \sum_{a\in\{0,1\}^k} (-1)^{x_{N(a)}+y_{N(a)}} \left(\bigotimes_{l=1}^{k} U_H\right)|a\rangle \otimes \left(\bigotimes_{l=1}^{k} U_H\right)|a\rangle$$

$$= \frac{1}{\sqrt{2^{4k}}} \sum_{a,b,c\in\{0,1\}^k} (-1)^{x_{N(a)}+y_{N(a)}+\langle a,b\rangle+\langle a,c\rangle}|b\rangle \otimes |c\rangle$$

where

$$b = (b_0, b_1, \ldots, b_{k-1}) \in \{0,1\}^k$$

$$c = (c_0, c_1, \ldots, c_{k-1}) \in \{0,1\}^k$$

$$\langle a,b\rangle := \sum_{j=0}^{k} a_j b_j, \qquad \langle a,c\rangle := \sum_{j=0}^{k} a_j c_j.$$

The probability of finding $|\phi\rangle$ in the state $|d\rangle \otimes |d\rangle$ where $d \in \{0,1\}^k$ is given by

$$|(\langle d| \otimes \langle d|)|\phi\rangle|^2 = \left| \frac{1}{\sqrt{2^{4k}}} \sum_{a,b,c\in\{0,1\}^k} (-1)^{x_{N(a)}+y_{N(a)}+\langle a,b\rangle+\langle a,c\rangle} \langle d|b\rangle\langle d|c\rangle \right|^2$$

$$= \frac{1}{2^{4k}} \left| \sum_{a,b,c\in\{0,1\}^k} (-1)^{x_{N(a)}+y_{N(a)}+\langle a,d\rangle+\langle a,d\rangle} \right|^2$$

$$= \frac{1}{2^{2k}} \left| \sum_{a\in\{0,1\}^k} (-1)^{\delta_{x_{N(a)},y_{N(a)}}} \right|^2$$

where $\delta_{x_{N(a)},y_{N(a)}}$ is the Kronecker delta, and is 1 when $x_{N(a)} = y_{N(a)}$ and 0 otherwise. Thus $|(\langle d| \otimes \langle d|)|\phi\rangle|^2 = 2^{-k}$ when $x = y$ and $|(\langle d| \otimes \langle d|)|\phi\rangle|^2 = 0$ when x and y differ in exactly half of the bits. Thus measurement in the computational basis always yields d for both alice and bob where $d \in \{0,1\}^k$ when $x = y$ but never when x and y differ in exactly half of the bits. Now Alice and Bob can output the result of their measurements of each of their qubits. The outputs are the same if and only if $x = y$. The only communication is the initial k qubits that Alice sends to Bob.

Thus the Deutsch-Jozsa relation can be achieved by communicating only $k = \log_2 n$ qubits.

3.7 Communication Complexity of Sampling

In the following we use concepts from chapter 1, referenced and described below. The *communication matrix* M_f of $f : A \times B \to \{0,1\}$, for A and B two finite sets, is defined as

$$M_f := (f(a,b))_{ab}.$$

In the following we define

$$\begin{aligned} T(f) &:= \{(a,b) \in A \times B : \ f(a,b) = 1\} \\ T(f,A) &:= \{a \in A : \ \forall b \ (a,b) \in A \times B \quad f(a,b) = 1\} \\ p_T(f,a) &:= \frac{|\{b \in B : \ (a,b) \in T(f)\}|}{|T(f)|}. \end{aligned}$$

Consider the pure state ($|T(f)| \neq 0$)

$$|\psi_f\rangle := \frac{1}{\sqrt{|T(f)|}} \sum_{a \in A, b \in B} f(a,b)|a\rangle \otimes |b\rangle.$$

Suppose Alice has the physical system which is described by a state in the Hilbert space $\mathcal{H}_A$ and Bob has the physical system described by a state in $\mathcal{H}_B$. Of course Alice's part of the system may in fact not be describable as a pure state, her part may be described by the density operator

$$\begin{aligned} \rho_A &:= \mathrm{tr}_{\mathcal{H}_B}(|\psi_f\rangle\langle\psi_f|) \\ &= \frac{1}{|T(f)|} \sum_{a,c \in A, b \in B} f(a,b) f(c,b) |a\rangle\langle c| \\ &= \frac{1}{|T(f)|} M_f M_f^T. \end{aligned}$$

We define the *von Neumann entropy* as

$$S(\rho) := -\mathrm{tr}(\rho \log_2 \rho)$$

where ρ is a density operator. We also note that

$$\mathrm{rank}(\rho_A) = \mathrm{rank}(M_f)$$

since

$$\text{rank}(M_f M_f^T) = \min\{\text{rank}(M_f), \text{rank}(M_f^T)\} = \text{rank}(M_f).$$

Now suppose Alice and Bob wish to share the state $|\psi_f\rangle$. Furthermore they wish to prepare the state with as little communication as possible. This is an example of the problem of *sampling* [Ambainis]. Given b, Bob can calculate $f(a, b)$ for any given a. Thus communication complexity is the measure of how much Bob should know about a given f. Similarly, the communication complexity of sampling is the amount of information Bob should have on a given f in order to prepare $|\psi_f\rangle$. In this case the measure of information is entanglement [Bennett96, HardyQC]. In other words Alice and Bob must share the appropriate amount of entanglement in order to prepare $|\psi_f\rangle$.

We notice that ρ_A can be written as

$$\rho_A = \sum_{a\in A} p_T(f, a)|a\rangle\langle a| + \frac{1}{|T(f)|} \sum_{\substack{b\in B, c\in A \\ c\neq a}} f(a, b)f(c, b)|a\rangle\langle c|.$$

The off-diagonal terms represent correlation between the different a's and c's which Alice may use as input for the computation. In other words, suppose $a, c \in A$ and $a \neq c$ with $f(a, b) = f(c, b)$ for all $b \in B$. In this case Bob does not need to distinguish between a and c, and it is possible to reduce the expected communication complexity accordingly. $S(\rho_A)$ is a measure of the entanglement [Bennett96, HardyQC] of $|\psi_f\rangle$. It describes how many qubits Alice should send to Bob in order for them to prepare $|\psi_f\rangle$. If this value is zero, $f(a, b)$ can be written as a product $g(a)h(b)$ and the communication complexity is $C_0(f) = 1$ or

$$C_0(f) = 1 + S(\rho_A)$$

and

$$C_0^E(f) \leq 1 + S(\rho_A).$$

Example. Let

$$|\psi_f\rangle := \frac{1}{\sqrt{2}} \left(|00\rangle_A \otimes |00\rangle_B + |11\rangle_A \otimes |11\rangle_B\right).$$

Then

$$\mathrm{tr}_B(|\psi_f\rangle\langle\psi_f|) = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|.$$

Thus the entanglement measure yields 1 *ebit* of entanglement. From the non-zero probability amplitudes of $|\psi_f\rangle$ (i.e. $|00\rangle_A \otimes |00\rangle_B$ and $|11\rangle_A \otimes |11\rangle_B$) we note that

$$\begin{aligned} f(a_1, a_2, b_1, b_2) &= a_1 \cdot a_2 \cdot b_1 \cdot b_2 + \overline{a_1} \cdot \overline{a_2} \cdot \overline{b_1} \cdot \overline{b_2} \\ &= a_1 \cdot a_2 \cdot b_1 \cdot b_2 + \overline{(a_1 + a_2 + b_1 + b_2)}\,. \end{aligned}$$

Thus two bits of communication are necessary to compute $f(a_1, a_2, b_1, b_2)$.
■

Definition. Consider the state

$$|\psi\rangle := \sum_{j=1}^{\dim(\mathcal{H}_A)} \sum_{k=1}^{\dim(\mathcal{H}_B)} \psi_{jk}|j\rangle \otimes |k\rangle$$

where $|j\rangle$ for $j = 1, 2, \ldots, \dim(\mathcal{H}_A)$ forms an orthonormal basis in $\mathcal{H}_A$ and $|k\rangle$ for $k = 1, 2, \ldots, \dim(\mathcal{H}_B)$ forms an orthonormal basis in $\mathcal{H}_B$ and

$$\sum_{j=1}^{\dim(\mathcal{H}_A)} \sum_{k=1}^{\dim(\mathcal{H}_B)} |\psi_{jk}|^2 = 1.$$

A quantum protocol is said to *q-sample* the state $|\psi\rangle$ to within ϵ error if the protocol transforms the separable initial state $|\phi_i\rangle := |\phi_{i,A}\rangle \otimes |\phi_{i,B}\rangle$ with $|\phi_{i,A}\rangle \in \mathcal{H}_A$ and $|\phi_{i,B}\rangle \in \mathcal{H}_B$ to the state $|\phi_o\rangle$ with $|\langle\phi_o|\psi\rangle|^2 \geq 1 - \epsilon$.

Equivalently, a quantum protocol is said to *q-sample* the state $|\psi\rangle$ to within ϵ error if the protocol transforms the separable initial state $|\phi_i\rangle := |\phi_{i,A}\rangle \otimes |\phi_{i,B}\rangle$ with $|\phi_{i,A}\rangle \in \mathcal{H}_A$ and $|\phi_{i,B}\rangle \in \mathcal{H}_B$ to the state $|\phi_o\rangle$ with

$$||\,\mathrm{tr}_B|\phi_d\rangle\langle\phi_d|\,||^2 \leq 2\epsilon$$

where $||\cdot||$ denotes the Hilbert-Schmidt norm and $|\phi_d\rangle := |\phi_o\rangle - |\psi\rangle$.
■

Theorem. Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, $0 \leq \epsilon \leq \frac{1}{2}$ and K_ϵ be defined by

$$\sum_{j=1}^{K_\epsilon} \lambda_j \geq 1 - \epsilon$$

where λ_1, λ_2, ...are the decreasing eigenvalues of the matrix $\mathrm{tr}_B|\psi\rangle\langle\psi|$. The minimum communication cost $Q_\epsilon(|\psi\rangle)$, in qubits, of q-sampling $|\psi\rangle$ is bounded by

$$\lceil \log_2 K_{2\epsilon} \rceil \leq Q_\epsilon(|\psi\rangle) \leq \lceil \log_2 K_\epsilon \rceil.$$

■

3.8 Other Examples

[Brassard, Ta-Shma] provide reviews of quantum communication complexity, including the exponential separation between quantum and classical communication complexity in [Raz]. Further examples include quantum whispers [HardyL], lower bounds described in [Klauck, Buhrman00] and rounds in quantum communication [Klauck01].

In the following section we consider a natural model for communication complexity in the quantum setting.

3.9 Coherent Quantum Communication Complexity

In the context of quantum computation all operations on quantum data are represented by unitary operators. Thus [Nielsen] suggested the *coherent quantum communication complexity*. This is the number of qubits which must be communicated to perform the quantum operation. In other words, we use local unitary operations and communication of qubits. Let $Q_{CC,0}$ denote the coherent quantum communication complexity without error for two parties in qubits. Let U be a unitary operator acting on the finite-dimensional Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ where $\mathcal{H}_A$ is the Hilbert space describing the system belonging to Alice and $\mathcal{H}_B$ is the Hilbert space describing the system belonging to Bob.

Theorem. Nielsen [Nielsen, NielsenBenasque] proved that

$$\log_2 \mathrm{Sch}(U, \mathcal{H}_A, \mathcal{H}_B) \leq Q_{CC,0}(U)$$

where $\mathrm{Sch}(U, \mathcal{H}_A, \mathcal{H}_B)$ denotes the Schmidt rank over $\mathcal{H}_A \otimes \mathcal{H}_B$.

Proof. We follow the proof in [Nielsen] for the lower bound. Let Alice be the first party in the protocol achieving $Q_{CC,0}(U)$, and Bob the second. They wish to compute U over their quantum data (qubits). Let U_s denote the unitary operation performed after s steps in the protocol, and $\mathcal{H}_A(s)$ and $\mathcal{H}_B(s)$ the respective Hilbert spaces describing the systems of Alice and Bob. For $s = 0$ we have

$$U_0 = I \quad \Rightarrow \quad \mathrm{Sch}(U_0, \mathcal{H}_A(0), \mathcal{H}_B(0)) = 1.$$

For each step of the protocol we have one of the following

1) Alice performs a unitary operation on her quantum data and sends a qubit to Bob.

2) Bob performs a unitary operation on his quantum data and sends a qubit to Alice.

We describe the first case, and the second case follows by symmetry. Suppose

$$U_s = \sum_{i=1}^{\mathrm{Sch}(U_s, \mathcal{H}_A(s), \mathcal{H}_B(s))} A_{s,i} \otimes B_{s,i}$$

is a minimal decomposition of the unitary matrix U_s, and that Alice sends a qubit Q to Bob after applying her transform $U_{A,s}$. $\mathcal{H}_A(s)$ denotes Alice's system for this step (s), and $\mathcal{H}_B(s)$ denotes Bob's. We rewrite U_s in terms of A' and Q where $\mathcal{H}_A(s) = A' \otimes Q = \mathbf{C}^{\dim(\mathcal{H}_A(s))/2} \otimes \mathbf{C}^2$

$$A_{s,i} = A'_{s,i} \otimes Q_{s,i},$$

$$U_{s+1} = U_{A,s} \sum_{i=1}^{\mathrm{Sch}(U_s, \mathcal{H}_A(s), \mathcal{H}_B(s))} (A'_{s,i} \otimes Q_{s,i}) \otimes B_{s,i}.$$

Since $U_{A,s}$ operates on some quantum data and the single qubit Q, the matrix $U_{A,s}$ is of Schmidt rank over $A' \otimes Q$ is 1, 2 or 4. For the case 1, Alice applies

a separable unitary transform, i.e. achieves no meaningful communication. Consequently for the lower bound we only consider the cases Schmidt rank 2 or 4. Thus we can write

$$U_{A,s} = \sum_{j=1}^{4} U_{A',s,j} \otimes U_{Q,s,j},$$

$$U_{s+1} = \sum_{i=1}^{\mathrm{Sch}(U_s,\mathcal{H}_A(s),\mathcal{H}_B(s))} \sum_{j=1}^{4} (U_{A',s,j} A'_{s,i} \otimes U_{Q,s,j} Q_{s,i}) \otimes B_{s,i}.$$

We note that the Schmidt rank of

$$\sum_{j=1}^{4} (U_{A',s,j} A'_{s,i} \otimes U_{Q,s,j} Q_{s,i})$$

is at most twice that of

$$A'_{s,i} \otimes Q_{s,i}$$

since $U_{A,s}$ is of Schmidt rank 2 or 4. For the Schmidt rank

$$\mathrm{Sch}(U_{s+1}, \mathcal{H}_A(s+1), \mathcal{H}_B(s+1))$$

we use

$$\mathcal{H}_A(s+1) := \mathbf{C}^{\dim(\mathcal{H}_A(s))/2},$$

$$\mathcal{H}_B(s+1) := \mathbf{C}^2 \otimes \mathcal{H}_B(s).$$

Thus we find that

$$\mathrm{Sch}(U_{s+1}, \mathcal{H}_A(s+1), \mathcal{H}_B(s+1)) \leq 2\,\mathrm{Sch}(U_s, \mathcal{H}_A(s), \mathcal{H}_B(s)).$$

We find the same if Bob sends a qubit to Alice, from which follows

$$\mathrm{Sch}(U, \mathcal{H}_A, \mathcal{H}_B) \leq 2^{Q_{CC,0}(U)} \quad \Rightarrow \quad \log_2 \mathrm{Sch}(U, \mathcal{H}_A(s), \mathcal{H}_B(s)) \leq Q_{CC,0}(U).$$

■

Bibliography

[Ambainis]
Ambainis A., L. J. Schulman, A. Ta-Shma, U. Vazirani and A. Wigderson, "The Quantum Communication Complexity of Sampling", *SIAM Journal on Computing*, **32**, 1570–1585, 2003.

[Ash]
Ash R. B., *Information Theory*, Dover Publications, 1990.

[Barenco95]
Barenco A. and A. K. Ekert, "Dense coding based on quantum entanglement", *Journal of Modern Optics* **42**, 1253–1259, 1995.

[Bennett92]
Bennett C. H. and S.J. Wiesner, "Communication via one- and two-particle operations on Einstein-Podolsky-Rosen states", *Physical Review Letters* **69**, 2881-2884, 1992.

[Bennett93]
Bennett C. H., G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, "Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels", *Physical Review Letters* **70** 1895–1899, 1993.

[Bennett96]
Bennett C. H., D.P. DiVincenzo, J.A. Smolin and W.K. Wootters, "Mixed State Entanglement and Quantum Error Correction", *Physical Review A*, **54**, 3824–3851, 1996.

[Bohm]
Bohm D., *Quantum Theory*, Prentice-Hall, New York, 1951.

[Bouwmeester97]
Bouwmeester D., J-W. Pan, K. Mattle, M. Eible, H. Weinfurter and

A. Zeilinger, "Experimental Quantum Teleportation", *Nature*, **390**, 575, 1997.

[Brassard]
Brassard G., "Quantum Communication Complexity", *Foundations of Physics*, **33**, 1593 – 1616, 2003.

[Brassard99]
Brassard G., R. Cleve and A. Tapp, "The cost of exactly simulating quantum entanglement with classical communication", *Physical Review Letters*, **83**, 1874–1877, 1999.

[Brußß2002]
Bruß D., "Characterizing entanglement", *Journal of Mathematical Physics*, **43**, 4237-4251, 2002.

[Buhrman97a]
Buhrman H., W. van Dam, P. Høyer and A. Tapp, "Multiparty Quantum Communication Complexity", *Physical Review A*, **60**, 2737–2741, 1999.

[Buhrman97b]
Buhrman H., R. Cleve and W. van Dam, "Quantum Entanglement and Communication Complexity", *SIAM Journal on Computing*, **30**, 1829–1841, 2001.

[Buhrman98]
Buhrman H., R. Cleve and A. Wigderson, "Quantum vs Classical Communication and Computation", *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, 63–68, 1998.

[Buhrman00]
Buhrman H. and R. de Wolf, "Communication Complexity Lower Bounds by Polynomials", *Proceedings of the 16th Annual Conference on Computational Complexity*, 120, 2001.

[Chi2003]
Dong Pyo Chi and Soojon Lee, "Entanglement for a two-parameter class of states in 2 $\otimes$ n quantum system" *Journal of Physics A: Mathematical and General*, **36**, 11503–11511, 2003.

[Childs2003]
Childs A. M., D. Leung, F. Verstraete and W. van Dam, "Asymptotic entanglement capacity of the Ising and anisotropic Heisenberg interactions", *Quantum Information and Computation*, **3**, 97–105, 2003.

[Cirac2002]
Cirac J. I., "Quantum Information: Entanglement, Purification, Error Correction, and Quantum Optical Implementations", in *Fundamentals of Quantum Information*, Heiss D. (Ed.), Springer, Berlin, 2002.

[Cleve]
Cleve R. and H. Buhrman, "Substituting Quantum Entanglement for Communication", *Physical Review Letters*, **83**, 1874–1877, 1999.

[Deutsch85]
Deutsch D., "Quantum theory, the Church-Turing principle and the universal quantum computer", *Proceedings of the Royal Society of London A*, **400**, 97–117, 1985.

[Deutsch92]
Duetsch D. and R. Jozsa, "Rapid solution of problems by quantum computation", *Proceedings of the Royal Society of London A*, **439**, 553–558, 1992.

[EPR]
Einstein, A., B. Podolsky and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", *Physical Review*, **47**, 777–780, 1935.

[Garret98]
Garret P., "Hahn-Banach Theorems", 1998,
`http://www.math.umn.edu/~garrett/m/fun/hahn_banach.pdf`.

[Garret05]
Garret P., "Normed and Banach Spaces", 2005,
`http://www.math.umn.edu/~garrett/m/fun/Notes/03_banach.pdf`.

[HardyL]
Hardy L. and W. van Dam, "Quantum Whispers", *Physical Review A*, **59**, 2635–2640, 1999.

[HardyQC]
Hardy Y. and W.-H. Steeb, *Classical and Quantum Computing*, Birkhäuser, 2001.

[Hardy2004]
Hardy Y. and W.-H. Steeb, "Fermi-Bose Systems, Macroscopic Quantum Superposition States and Entanglement", *International Journal of Theoretical Physics*, 2004.

[Henderson99]
Henderson L. and V. Vedral, "Information, Relative Entropy of Entanglement and Irreversibility", *Physical Review Letters*, **84**, 2263–2266, 2000.

[Hines2003]
Hines A. P., R. H. McKenzie and G. J. Milburn, "Entanglement of two-mode Bose-Einstein condensates", *Physical Review A*, **67**, 013609, 2003.

[Holevo98]
Holevo A. S., "Coding Theorems for Quantum Channels", *Tamagawa University Research Review*, **4**, 1998.

[Horodecki96]
Horodecki M., P. Horodecki and R. Horodecki, "Separability of Mixed States: necessary and sufficient conditions", *Physics Letters A*, **223**, 1–8, 1996.

[Itano97]
Itano W. M., C. Monroe, D. M. Meekhof, D. Leibfried, B. E. King and D. J. Wineland, "Quantum harmonic oscillator state synthesis and analysis", in *Atom Optics*, eds M.G. Prentiss and W. D. Phillips, Proceedings of SPIE 2995, **43**, 1997.

[Keyl2003]
Keyl M., D. Schlingemann and R. F. Werner, "Infinitely entangled states", *Quantum Information and Computation*, **3**, 281–306, 2003.

[Klauck]
Klauck H., "Lower bounds for quantum communication complexity", *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, 288 – 297, 2001.

[Klauck01]
Klauck H., A. Nayak, A. Ta-Shma and D. Zuckerman, "Interaction in Quantum Communication and the Complexity of Set Disjointness", *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, 124 – 133, 2001.

[Kremer95]
Kremer I., *Quantum Communication*, Masters thesis, Hebrew University of Jerusalem, 1995.

[Nielsen]
Nielsen M. A., *Quantum information theory*, Ph.D. thesis, University of New Mexico, 1998.

[Nielsen2003]
Nielsen M. A., C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow and A. Hines, "Quantum dynamics as a physical resource", *Physical Review A*, **67**, 052301, 2003.

[NielsenBenasque]
Nielsen M. A., "Entanglement and distributed quantum computation", Talk at the Benasque Center for Physics,
`http://theory.caltech.edu/~mnielsen/talks/benasque00/talk.ps`.

[Kushilevitz]
Kushilevitz E. and N. Nisan, *Communication Complexity*, Cambridge University Press, 1997.

[Papad]
Papadimitriou C. H., *Computational Complexity*, Addison-Wesley, 1994.

[Preskill]
Preskill J., *Quantum Information and Computation*,
`http://www.theory.caltach.edu/~preskill/ph229`.

[Rains99]
Rains E. M., "Rigorous treatment of distillable entanglement", *Physical Review A*, **60**, 173–178, 1999.

[Raz]
Raz R., "Exponential Separation of Quantum and Classical Communication Complexity", *Proceedings of the 31st Symposium on Theory of Computing*, 358-367, 1999.

[Rieffel98]
Rieffel E. and W. Polak, "An Introduction to Quantum Computing for Non-Physicists", *ACM Computing Surveys*, **32**, 300 - 335, 2000.

[Shor94]
Shor P. W., "Proceedings of the 35th Annual Symposium on the Foundations of Computer Science", edited by S. Goldwasser (Los Alamitos, CA: IEEE Computer Society Press), p. 124, 1994.

[Spector]
Spector L., H. Barnum, H. J. Bernstein and N. Swamy, "Finding a Better-than-Classical Quantum AND/OR Algorithm using Genetic Programming", *Proceedings of the 1999 Congress on Evolutionary Computation*, IEEE Press, 1999.

[Steeb2000a]
Steeb W.-H. and Y. Hardy, "Entangled quantum states and a C++ implementation", *International Journal of Modern Physics C*, **11**, 69–77, 2000.

[Steeb2000b]
Steeb W.-H. and Y. Hardy, "Entangled quantum states", *International Journal of Theoretical Physics*, **39**, 2765, 2000.

[Steeb2001a]
Steeb W.-H. and Y. Hardy, "Fermi systems, Hubbard model and a SymbolicC++ implementation", *International Journal of Modern Physics C*, **12**, 235–245, 2001.

[Steeb2002a]
Steeb W.-H. and Y. Hardy, "Entangled Quantum States and the Kronecker Product", *Zeitschrift für Naturforschung*, **57a**, 400, 2002.

[SteebPB2]
Steeb W.-H., *Problems and Solutions in Theoretical and Mathematical Physics*, World Scientific Publishing, Volume II, Advanced Level, Singapore, 2003.

[SteebPQC]
Steeb W.-H. and Y. Hardy, *Problems and Solutions in Quantum Computing and Quantum Information*, World Scientific, Singapore, 2004.

[SteebQM]
Steeb W.-H., *Hilbert Spaces, Wavelets, Generalised Functions and Modern Quantum Mechanics*, Kluwer Academic Publishers, 1998.

[SteebMC]
Steeb W.-H., *Matrix Calculus and Kronecker Product with Applications and C++ Programs*, World Scientific, 1997.

[Steeb86]
Steeb W.-H., J. A. Louw, C. M. Villet and A. Kunick, "Finite Hubbard-Model with Phonon Coupling", *Physica Scripta*, **34**, 245, 1986.

[SymbolicC++]
Tan Kiat Shi, W.-H. Steeb and Y. Hardy, *SymbolicC++: An Introduction to Computer Algebra using Object-Oriented Programming*, Springer-Verlag, London 2000.

[Ta-Shma]
Ta-Shma A., "Classical versus Quantum Communication Complexity", *SIGACT News: Complexity Theory Column 23*.

[Vedral98]
Vedral V. and M. B. Plenio, "Basics of Quantum Computation", *Progress in Quantum Electronics*, **22**, 1–3, 1998.

[Wang2001]
Wang X. and B. C. Sanders, "Multipartite entangled coherent states", *Physical Review A*, **65**, 012303, 2002.

[Wang2003]
Wang X. and B. C. Sanders, "Entanglement capability of self-inverse Hamiltonian evolution", *Physical Review A*, **68**, 014301, 2003.

[Werner89]
Werner R. F., "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model", *Physical Review A*, **40**, 4277, 1989.

[Wootters98]
Wootters W. K., "Entanglement of Formation of an Arbitrary State of Two Qubits", *Physical Review Letters*, **80**, 2245, 1998.

Index