Entanglement of Multipartite Quantum States and the Generalized Quantum Search

Thesis by
Robert M. Gingrich

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2002
(Submitted August 23, 2001)

Acknowledgements

I would like to give special thanks to my advisor John Preskill for his support and guidance during my research and to my collaborators, Colin Williams, Christoph Adami and Nicolas Cerf. I would also like to thank David Beckman, Sumit Daftuar, Michael Nielsen, Federico Spedalieri, Anthony Sudbery, David Whitehouse and the rest of the Institute for Quantum Information at Caltech for useful discussions and help during my research. I would like to thank my parents without whom none of this would have been possible for the obvious reason as well as many, many not so obvious reasons. And I would like to thank my friends and coworkers who have made these last five years not only productive but a lot of fun.

Abstract

In chapter 2 various parameterizations for the orbits under local unitary transformations of three-qubit pure states are analyzed. It is shown that the entanglement monotones of any multipartite pure state uniquely determine the orbit of that state. It follows that there must be an entanglement monotone for three-qubit pure states which depends on the Kempe invariant defined in [1]. A form for such an entanglement monotone is proposed. A theorem is proved that significantly reduces the number of entanglement monotones that must be looked at to find the maximal probability of transforming one multipartite state to another.

In chapter 3 Grover's unstructured quantum search algorithm is generalized to use an arbitrary starting superposition and an arbitrary unitary matrix. A formula for the probability of the generalized Grover's algorithm succeeding after n iterations is derived. This formula is used to determine the optimal strategy for using the unstructured quantum search algorithm. The speedup obtained illustrates that a hybrid use of quantum computing and classical computing techniques can yield a performance that is better than either alone. The analysis is extended to the case of a society of k quantum searches acting in parallel.

In chapter 4 the positive map $\Gamma : \rho \rightarrow (\mathrm{Tr}\rho) - \rho$ is introduced as a separability criterion. Any separable state is mapped by the tensor product of Γ and the identity in to a non-negative operator, which provides a necessary condition for separability. If Γ acts on a two-dimensional subsystem, then it is equivalent to partial transposition and therefore also sufficient for 2×2 and 2×3 systems. Finally, a connection between this map for two qubits and complex conjugation in the "magic" basis [2] is displayed.

Contents

List of Figures

Chapter 1

Introduction

At this point in history quantum mechanics is the best description of nature that exists. No repeatable experiment has ever contradicted it, yet it is still not very well understood. Compared to classical mechanics, the principles of quantum mechanics are less intuitive and the mathematics is often more difficult. Nevertheless, if one wants to know what is possible in nature, one must look at the true quantum mechanical description, not the approximation of classical mechanics.

In the last 30 years computers and digital information have become important in our society. This has been made possible by, among other advances, our understanding of computation, algorithms, information compression and error correction. These areas of study have, until recently, been based solely on classical principles and intuition. In the last decade is has been shown that by looking at the true quantum mechanical description new phenomena are possible (e. g., Shor's algorithm, teleportation, Grover's algorithm, quantum error correction). This has led to the studies of quantum computation and quantum information theory.

The concept of entanglement is central to the study of quantum information. All of classical information theory is based on physical systems with zero entanglement. These systems are said to be represented by a separable state. The correlations between the subsystems of an entangled (i. e., non-separable) state cannot be fully explained by classical physics. The first step in understanding what new phenomena are possible with quantum information is to find out what states are non-classical. In chapter 4 a criterion for detecting separability called the "reduction criterion" is investigated. This criterion is shown to be equivalent to the already known Peres criterion in $2 \times N$ systems and to be helpful in the calculation of the entanglement of formation (a particular measure of entanglement) for 2×2 systems.

Entanglement between more than two subsystems is more complicated and hence less well understood than the entanglement between two subsystems. This problem is addressed in chapter 2. A framework for characterizing the set of all measures of entanglement, called entanglement monotones, is proposed. This framework is used to show that there are some important yet undiscovered entanglement monotones for systems with 3 two-dimensional subsystems. Some properties of these entanglement monotones are derived and an explicit form is proposed for one of them.

The study of quantum computation is concerned with algorithms on a quantum computer. Certain algorithms (Shor's algorithm, Grover's algorithm, quantum Fourier transform) for a quantum computer will execute in less time steps than any known classical algorithm. For instance, Grover's algorithm can search an unstructured database of size N in $O(\sqrt{N})$ queries whereas the best classical algorithm takes $O(N)$ queries. If we assume a query time of 1 μsecond, then a search that takes 9.5 hours using Grover's algorithm would take about 20 million years with the best classical algorithm. Unfortunately, no practical quantum computer exists yet that can operate on more than a few qubits (i. e., quantum bits).

In chapter 3 an equation for the computation time of grover's algorithm with an arbitrary starting state, unitary transformation and number of target states is derived. This equation is then used to show that there is a moderate speedup of the algorithm on average if one measures before the peak probability. Also, the idea of parallelizing Grover's search is introduced and the computation time is analyzed.

Chapter 2

Properties of Entanglement Monotones

2.1 Introduction

Entanglement is at the heart of the studies of quantum computation and quantum information theory. It is what separates these studies from their classical counterparts. If we are to understand what new phenomena occur when we look at the true quantum mechanical description of nature as opposed to the approximations of classical mechanics, then we must understand how the quantum mechanical description differs from the classical description. Entanglement is a measure of this difference. While entanglement between two parties is quite well understood [3] [4] [5] [6], the entanglement within a quantum algorithm or in a state shared between many parties involves multipartite entanglement which is just beginning to be understood [7] [8] [9].

An integral part of the study of entanglement is determining the probability of transforming one pure state into another by Local Operations and Classical

Communication (LOCC). For two part systems this problem is solved, or at least reduced to the problem of finding the eigenvalues of a hermitian matrix, by [5] [6]. For a $N \times M$ pure state the Schmidt decomposition tells us we can write

$$|\psi\rangle = \sum_{i=1}^{n} \sqrt{\lambda_i^{\uparrow}} |i\rangle |i'\rangle \tag{2.1}$$

where the $\lambda_i^{\uparrow}$ are in increasing order, $\sum_i \lambda_i^{\uparrow} = 1$, the $|i\rangle$ and $|i'\rangle$ are an orthonormal set of vectors in space A and B respectively, and $n = \min(N, M)$. If we define

$$E_k\left(|\psi\rangle\right) = \sum_{i=1}^{k} \lambda_i^{\uparrow} \quad k = 1, \ldots, n-1 \tag{2.2}$$

then the highest attainable probability of transforming $|\psi\rangle$ to $|\phi\rangle$, $P(|\psi\rangle \to |\phi\rangle)$, is given by [6]

$$P(|\psi\rangle \to |\phi\rangle) = \min_k \frac{E_k\left(|\psi\rangle\right)}{E_k\left(|\phi\rangle\right)} \tag{2.3}$$

The proof of this theorem is constructive so we can actually write down the transformation that gives us $|\phi\rangle$ from $|\psi\rangle$. For pure states of more than two parts no such nice theorem is known. The question of whether two three-qubit pure states can be transformed into each other with non-zero probability by LOCC has been solved by Dür et al. [10] but just getting a reasonable upper bound on that probability when it is a non-zero is unsolved. In this paper I attempt to make some progress towards solving this problem for three-qubit pure states and hopefully shed some light on how we might solve it for larger dimensional spaces and more parts.

One way to find $P(|\psi\rangle \to |\phi\rangle)$ is to look at the entanglement monotones $E(|\psi\rangle)$ for the two states. For the duration of the paper "state" will refer to a pure state unless explicitly called a mixed state. An entanglement monotone, EM, is defined as a function that goes from states to positive real numbers and does not increase under LOCC. As a convention the value of any EM for a separable state is 0. For mixed and pure states of any dimension and number of parts, the following theorem holds [11]:

$$P(\rho \rightarrow \rho') = \min_{E} \frac{E(\rho)}{E(\rho')} \tag{2.4}$$

where the minimization is taken over the set of all EMs [11]. This can be seen by considering $P(\rho \rightarrow \rho')$ as an EM for ρ. The problem is that this minimization is difficult to take since there is no known way to characterize all the entanglement monotones for multipartite states. We would like a "minimal set" of EMs similar to the E_k for the bipartite case in order to take the minimization.

The situation for three or more parts is somewhat different than for bipartite pure states. Firstly, generic $M \times M$ bipartite states have a stabilizer (i. e., the set of unitaries that takes a state to itself) of dimension $M-1$ isomorphic to $U(1)^{\otimes M-1}$ while pure states with more parts generically have a discrete stabilizer. States whose parts are not of the same dimension may have larger stabilizers but bipartite states are the only ones that always have a continuous stabilizer. Secondly, the generalized Schmidt decomposition, however you choose to generalize it [12] [13], has complex coefficients for pure states with three or more parts. This implies that generically these states are not local unitarily equivalent to their complex conjugate states (i. e., the state with each of its coefficients complex conjugated). Also, for bipartite pure states all the local unitary (LU) invariants can be calculated from the eigenvalues of the reduced density matrices but this does not hold for more parts. I will go into more detail about LU invariants in the next section.

The structure of the paper is as follows: in section 2.2 the interconvertibility, behavior under measurement, symmetry properties, parameter ranges and calculability of two generalizations of the Schmidt decomposition of equation (2.1) and the polynomial invariants (defined below) are looked at. In section 2.3 it is shown that the entanglement monotones uniquely determine the orbit of multipartite pure states, and this is used to show that there must be an EM algebraically independent of the known EMs. A form for this EM is proposed and studied. Section 2.4 discusses other monotones that must exist and their properties. Lastly, in section 2.5 a theorem is proved that significantly reduces the number of EMs that must be minimized over to get $P(\rho \rightarrow \rho')$ of equation (2.4).

2.2 Decompositions and Invariants of Three-Qubit Pure States

Let $|\psi\rangle$ be a multipartite state in $\mathcal{H}_1 \otimes \mathcal{H}_2 \ldots \otimes \mathcal{H}_n$ and let $A_k^{(i)} : \mathcal{H}_i \rightarrow \mathcal{H}_i'$ be Krauss operators for an operation on the hilbert space $\mathcal{H}_i$ with $\sum_k A_k^{(i)\dagger} A_k^{(i)} = \mathcal{I}_i$ and $\mathcal{I}_i$ is the identity acting on $\mathcal{H}_i$. A (non-increasing) EM is a real valued function $E\left(|\psi\rangle\right)$ such that

$$E\left(|\psi\rangle\right) \geq \sum_k p_k E\left(\frac{I_1 \otimes \ldots \otimes A_k^{(i)} \otimes \ldots \otimes I_n |\psi\rangle}{\sqrt{p_k}}\right) \tag{2.5}$$

for any state $|\psi\rangle$, operation $A_k^{(i)}$, and space i where

$$p_k = ||I_1 \otimes \ldots \otimes A_k^{(i)} \otimes \ldots \otimes I_n |\psi\rangle||^2. \tag{2.6}$$

This definition for pure states is taken from the definition for a general state in [11]. One can always transform a state into product states and a product state cannot be transformed into anything but another product state so the value of an EM for a product state is chosen to be zero and all other states must have a non-negative value for the EM. Since $A_k^{(i)}$ can be a unitary operator or the inverse of that operator, equation (2.5) implies that all EMs must be invariant under LU. Hence, a first step to understanding the EMs is to look at the LU invariants that parameterize the set of orbits.

There are many ways to find LU invariants for three-qubit states [14] [13] [15] [12] [7] [16] [17], some of which can be generalized to more parts and larger spaces, but for now I will concentrate on the three-qubit case. The three sets of invariants I will look at in this section are the polynomial invariants [14], what I will call the diagonalization decomposition [13] and what I will call the maximization decomposition [12].

2.2.1 The Polynomial Invariants

A general polynomial invariant $P_{\sigma,\tau}\left(|\psi\rangle\right)$ for a state of the form

$$|\psi\rangle = \sum_{i,j,k=0}^{1} t_{ijk}|ijk\rangle \tag{2.7}$$

is written as

$$P_{\sigma,\tau}\left(|\psi\rangle\right) = \sum t_{i_1j_1k_1}\ldots t_{i_nj_nk_n}\bar{t}_{i_1j_{\sigma(1)}k_{\tau(1)}}\ldots\bar{t}_{i_nj_{\sigma(n)}k_{\tau(n)}} \tag{2.8}$$

where σ and τ are permutations on n elements, repeated indices are summed and $\bar{t}$ stands for the complex conjugate of t [14]. If one applies a unitary to any of the qubits in $|\psi\rangle$ and explicitly writes out $P_{\sigma,\tau}\left(|\psi\rangle\right)$ again, it becomes apparent that $P_{\sigma,\tau}\left(|\psi\rangle\right)$ is invariant. Of course, any polynomial in terms of the polynomial invariants $P_{\sigma,\tau}\left(|\psi\rangle\right)$ is another polynomial invariant. In fact, it can be shown that all the polynomial invariants are of this form.

We know from [12] that generic three-qubit states have a discrete stabilizer so the number of independent polynomial invariants is given by

$$\dim\left[\mathcal{C}^2\otimes\mathcal{C}^2\otimes\mathcal{C}^2\right] - 3\dim[SU(2)] - \dim[U(1)] - 1 = 5 \tag{2.9}$$

where the last -1 is due to the fact that we are using normalized states. The five independent continuous invariants are

$$\begin{aligned} I_1 &= P_{e,(12)} \\ I_2 &= P_{(12),e} \\ I_3 &= P_{(12),(12)} \end{aligned}$$

$$\begin{aligned} I_4 &= P_{(123),(132)} \\ I_5 &= |\sum t_{i_1 j_1 k_1} t_{i_2 j_2 k_2} t_{i_3 j_3 k_3} t_{i_4 j_4 k_4} \\ &\quad \times \epsilon_{i_1 i_2} \epsilon_{i_3 i_4} \epsilon_{j_1 j_2} \epsilon_{j_3 j_4} \epsilon_{k_1 i_3} \epsilon_{k_2 i_4}|^2 \end{aligned} \tag{2.10}$$

where $\epsilon_{00} = \epsilon_{11} = 0$, and $\epsilon_{01} = -\epsilon_{10} = 1$ and again repeated indices are summed. I_4 is the Kempe invariant referred to in the abstract. If one writes out I_5 and uses the identity $\epsilon_{ij}\epsilon_{rs} = \delta_{ir}\delta_{js} - \delta_{is}\delta_{jr}$, it can be shown that I_5 is just the sum and difference of 64 polynomials of the form in equation (2.8). With one more discrete invariant,

$$I_6 = \text{sign}\big[\text{Im}\big[P_{(34)(56),(13524)}\big]\big] , \tag{2.11}$$

the LU orbit of a three-qubit state is determined uniquely [13] [18]. I will define sign$[x]$ as 1 for non-negative numbers and -1 otherwise. The polynomial invariants have the advantage of being easy to compute for any state and the four previously known independent EMs [7] are the following simple functions of I_1, I_2, I_3 and I_5

$$\begin{aligned} \tau_{(AB)C} &= 2(1 - I_1) \\ \tau_{(AC)B} &= 2(1 - I_2) \\ \tau_{(BC)A} &= 2(1 - I_3) \\ \tau_{ABC} &= 2\sqrt{I_5}. \end{aligned} \tag{2.12}$$

2.2.2 The Diagonalization Decomposition

The diagonalization decomposition, DD, introduced by Acin et al. [13] is accomplished by first defining matrices $(T_0)_{j,k} = t_{0,j,k}$ and $(T_1)_{j,k} = t_{1,j,k}$, then finding a unitary operation on space A that makes T_0 singular, finding unitaries on space B and C that make T_0 diagonal and using the remaining phase freedom to get rid of as many phases as possible. What is left is a state of the form

$$\begin{aligned} |\psi_{\mathrm{DD}}\rangle &= \sqrt{\mu_0}\,|000\rangle + \sqrt{\mu_1}\,e^{i\phi}|100\rangle \\ &\quad +\sqrt{\mu_2}\,|101\rangle + \sqrt{\mu_3}\,|110\rangle + \sqrt{\mu_4}\,|111\rangle \end{aligned} \tag{2.13}$$

where $\mu_i \geq 0$, $\mu_0+\mu_1+\mu_2+\mu_3+\mu_4 = 1$ and $0 \leq \phi \leq \pi$. Note that generically there are two unitaries that will make T_0 singular, but it can be shown that only one will lead to ϕ between 0 and π. If there is another solution, with ϕ between π and 2π exclusive, it is referred to as the dual state of $|\psi_{\mathrm{DD}}\rangle$. Some nice properties of DD are that there is a 1 to 1 correspondence with the orbits and there are a set of invertible functions between the parameters of the decomposition and the set of polynomial invariants given above. Namely,

$$\begin{aligned} I_1 &= 1-2\mu_0(\mu_2+\mu_4)-2\Delta \\ I_2 &= 1-2\mu_0(\mu_3+\mu_4)-2\Delta \\ I_3 &= 1-2\mu_0(\mu_2+\mu_3+\mu_4) \\ I_4 &= 1-3[(\mu_2+\mu_3)(\mu_0-\mu_4)+\mu_4(1-\mu_4) \\ &\quad -\mu_2\mu_3\mu_0+(1-\mu_0)(\Delta-\mu_1\mu_4)] \\ I_5 &= 4\mu_0^2\mu_4^2 \\ I_6 &= \mathrm{sign}[\sin(\phi)\mu_0^2\sqrt{\mu_1\mu_2\mu_3\mu_4} \\ &\quad \times(\Delta-\mu_4(1-2\mu_0+\mu_1)-\mu_2\mu_3)] \end{aligned} \tag{2.14}$$

where $\Delta = \mu_1\mu_4+\mu_2\mu_3-2\sqrt{\mu_1\mu_2\mu_3\mu_4}\cos(\phi)$ and if we define

$$\begin{aligned} J_1 &= \frac{1}{4}\left(1-I_1-I_2+I_3-2\sqrt{I_5}\right) \\ J_2 &= \frac{1}{4}\left(1-I_1+I_2-I_3-2\sqrt{I_5}\right) \\ J_3 &= \frac{1}{4}\left(1+I_1-I_2-I_3-2\sqrt{I_5}\right) \\ J_4 &= \sqrt{I_5} \\ J_5 &= \frac{1}{4}\left(\frac{5}{3}-I_1-I_2-I_3+\frac{4}{3}I_4-2\sqrt{I_5}\right) \end{aligned} \tag{2.15}$$

then the coefficients are given by

$$\begin{aligned}
\mu_0^{\pm} &= \frac{J_4 + J_5 \pm \sqrt{\Upsilon}}{2(J_1 + J_4)} \\
\mu_i^{\pm} &= \frac{J_i}{\mu_0^{\pm}}, \quad i = 2, 3, 4 \\
\mu_1^{\pm} &= 1 - \mu_0^{\pm} - \frac{J_2 + J_3 + J_4}{\mu_0^{\pm}} \\
\cos(\phi^{\pm}) &= \frac{\mu_1^{\pm}\mu_4^{\pm} + \mu_2^{\pm}\mu_3^{\pm} - J_1}{2\sqrt{\mu_1^{\pm}\mu_2^{\pm}\mu_3^{\pm}\mu_4^{\pm}}} \\
\text{sign}[(\sin(\phi^{\pm})] &= I_6 \,\text{sign}[\sqrt{\mu_1^{\pm}\mu_2^{\pm}\mu_3^{\pm}\mu_4^{\pm}}[J_1 - J_2 J_3 \\
&\quad - J_4(J_2 + J_3 + J_4 - (\mu_0^{\pm})^2)]]
\end{aligned} \tag{2.16}$$

where $\Upsilon = (J_4 + J_5)^2 - 4(J_1 + J_4)(J_2 + J_4)(J_3 + J_4) \geq 0$. The + and − solutions for the coefficients correspond to $|\psi_{\text{DD}}\rangle$ and its dual state. The inversion of the equations for I_i was done independently in [18]. Note that their definition of I_4 is different from the one in this thesis.

Another nice property of the DD is that we can perform an arbitrary measurement on it in space A and stay in the DD form. Since any measurement can be broken into a series of two outcome measurements [19], we can look at the two outcome measurement A_1 and A_2 where $A_1^{\dagger}A_1 + A_2^{\dagger}A_2 = I$. Using the singular value decomposition, we can write $A_i = U_i D_i V$ where V does not depend on i because the two positive hermitian operators $A_1^{\dagger}A_1$ and $A_2^{\dagger}A_2$ sum to the identity and therefore must be simultaneously diagonalizable. The diagonal matrices, D_i, can be written as

$$D_1 = \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}, \quad D_2 = \begin{bmatrix} \sqrt{1 - x^2} & 0 \\ 0 & \sqrt{1 - y^2} \end{bmatrix} \tag{2.17}$$

where $0 \leq x, y \leq 1$ [10]. Since we are only concerned with what orbit the outcomes are in, we may choose the U_i transformation. Also, matrices of the form

$$\begin{bmatrix} e^{i\psi_1} & 0 \\ 0 & e^{i\psi_2} \end{bmatrix} \tag{2.18}$$

where ψ_1 and ψ_2 are real numbers, commute with the D_i matrices so the most general V can be written as

$$\begin{bmatrix} \alpha & \sqrt{1-\alpha^2}e^{i\theta} \\ -\sqrt{1-\alpha^2}e^{-i\theta} & \alpha \end{bmatrix} \tag{2.19}$$

where $0 \leq \alpha \leq 1$ and θ is real. If we choose

$$\begin{aligned} U_1 &= \frac{1}{\sqrt{\gamma}} \begin{bmatrix} y\alpha & -x\sqrt{1-\alpha^2}e^{i\theta} \\ x\sqrt{1-\alpha^2}e^{-i\theta} & y\alpha \end{bmatrix} \\ \gamma &= y^2\alpha^2 + x^2(1-\alpha^2) \end{aligned} \tag{2.20}$$

and similarly for U_2 with (x, y) replaced with $(\sqrt{1-x^2}, \sqrt{1-y^2})$, then in going from $|\psi_{\mathrm{DD}}\rangle$ to $A_1|\psi_{\mathrm{DD}}\rangle$ the DD coefficients undergo the following transformations:

$$\begin{aligned} \mu_0 &\rightarrow \frac{x^2y^2\mu_0}{\gamma} \\ \mu_1 &\rightarrow \frac{1}{\gamma}\left|e^{-i\theta}(x^2-y^2)\alpha\sqrt{\mu_0(1-\alpha^2)} + e^{i\phi}\gamma\sqrt{\mu_1}\right|^2 \\ \mu_i &\rightarrow \mu_i\gamma \qquad i = 2, 3, 4 \\ \phi &\rightarrow \arg\left[e^{-i\theta}(x^2-y^2)\alpha\sqrt{\mu_0(1-\alpha^2)} + e^{i\phi}\gamma\sqrt{\mu_1}\right] \end{aligned} \tag{2.21}$$

and again similarly for $A_2|\psi_{\mathrm{DD}}\rangle$. Things become more complicated when ϕ becomes larger than π and we have a dual solution. In this case we need to transform to the dual state which can be quite tedious. It should also be noted that if we want to plug the new form for the DD coefficients into equations (2.14), then the normalization must be taken into account. The normalization will just be the sum of the new forms for μ_0 through μ_4.

2.2.3 The Maximization Decomposition

The Maximization Decomposition [12], MD, has a somewhat different way of decomposing the three qubit states. First we find the states, $|\phi_A\rangle$, $|\phi_B\rangle$ and $|\phi_C\rangle$ each defined up to an overall phase, that maximize

$$g(|\phi_A\rangle, |\phi_B\rangle, |\phi_C\rangle) = ||\langle\psi|\phi_A\rangle|\phi_B\rangle|\phi_C\rangle||^2 \tag{2.22}$$

and apply a unitary such that $|\phi_A\rangle|\phi_B\rangle|\phi_C\rangle$ becomes $|000\rangle$. Defining $|1\rangle$, up to an overall phase, as the vector perpendicular to $|0\rangle$, then the derivative of g along $|1\rangle$ at the point $|000\rangle$,

$$\begin{aligned} &\lim_{\epsilon\to 0} \frac{g(|0\rangle + \epsilon|1\rangle, |0\rangle, |0\rangle) - g(|0\rangle, |0\rangle, |0\rangle)}{\epsilon} \\ &= 2\mathrm{Re}\left[\langle\psi|100\rangle\langle 000|\psi\rangle\right] \end{aligned} \tag{2.23}$$

must be zero because $g(|0\rangle, |0\rangle, |0\rangle)$ is a maximum. Since we still have phase freedom in $|0\rangle$ and $|1\rangle$ this implies that $\langle\psi|100\rangle = 0$ and similarly for $\langle\psi|010\rangle$ and $\langle\psi|001\rangle$. Using the remaining phase freedom in the choice of $|0\rangle$ and $|1\rangle$, we can eliminate all but one phase leaving us with

$$|\psi_{\mathrm{MD}}\rangle = ae^{i\phi}|000\rangle + b|011\rangle + c|101\rangle + d|110\rangle + f|111\rangle \tag{2.24}$$

where $a^2 + b^2 + c^2 + d^2 + f^2 = 1$, $0 \leq \phi \leq 2\pi$, $0 \leq a, b, c, d, f$ and $b, c, d, f \leq a$. Note that $g(|0_A\rangle, |0_B\rangle, |0_C\rangle) = a^2$. Unfortunately, the parameters as they are given above are not in 1 to 1 correspondence with the orbits. While the decomposition is generically unique, there are choices of the parameters within the given ranges that are not the result of the decomposition. For example, states with $a^2 = \frac{1}{5} + \epsilon$, $b^2 = c^2 = d^2 = f^2 = \frac{1}{5} - \frac{\epsilon}{4}$ and any choice of ϕ have

$$g\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \geq a^2 \tag{2.25}$$

for $\epsilon \leq 0.014$. Hence, these choices of the parameters are not a result of the decomposition. The true ranges of the parameters that would give a 1 to 1 correspondence with the orbits are as yet unknown.

A nice property of the MD is that is it symmetric in particle exchange. Exchanging the particles is equivalent to exchanging b, c and d. This makes the permutation properties of the polynomial invariants easier to see when written in terms of the MD coefficients. They take the following form:

$$\begin{aligned}
I_1 &= 1-2\left((a^2+d^2)(b^2+c^2)+a^2f^2\right)\\
I_2 &= 1-2\left((a^2+c^2)(b^2+d^2)+a^2f^2\right)\\
I_3 &= 1-2\left((a^2+b^2)(c^2+d^2)+a^2f^2\right)\\
I_4 &= 1-3(a^2(1-a^2)-(b^2c^2+b^2d^2+c^2d^2)(1-2a^2)\\
&\quad -2b^2c^2d^2-2abcdf^2\cos(\phi))\\
I_5 &= a^2|\,af^2+4bcde^{i\phi}|^2\\
I_6 &= \text{sign}[abcdf^2\sin(\phi)(a^2(1-2a^2)(1-2a^2-f^2)\\
&\quad -4b^2c^2d^2-2abcdf^2\cos(\phi))].
\end{aligned} \tag{2.26}$$

It is apparent from these equations that I_1, I_2 and I_3 are symmetric in permutations of particles AB, AC and BC respectively and I_4, I_5 and I_6 are symmetric in any permutation of the particles. Unfortunately, the equations in (2.26) are not as easy to invert as those in (2.14). In fact, just calculating the MD coefficients for an arbitrary state is not an easy task, as it is in the case of the polynomial invariants and the DD coefficients, since determining the unitaries for the MD involves maximizing over a six-dimensional space with typically many local maxima.

One more interesting fact about the MD is that $1-a^2$ is a non-increasing EM. We know this because in [17] it is shown that a function of the form

$$E_{k_A,k_B,k_C}(|\psi\rangle)=\max_{\Gamma_A,\Gamma_B,\Gamma_C}||\Gamma_A\otimes\Gamma_B\otimes\Gamma_C|\psi\rangle||^2, \tag{2.27}$$

where Γ_X is a k_X-dimensional projector on system $X=A,B,C$, is a non-decreasing EM and $E_{1,1,1}(|\psi\rangle)=a^2$. The EM $1-a^2$ can be shown to be independent of the τ from equation (2.12) by looking at the gradient vectors of the τ, $1-a^2$ and $N = a^2+b^2+c^2+d^2+f^2$ at, for instance, the point $a=3$, $b,c,d,f=1$ and $\phi=\frac{\pi}{2}$. Since the gradient vectors span a six-dimensional space, $1-a^2$ cannot be written in terms of the τ and N. The problem with using $1-a^2$ as an EM is that one needs to find the global maximum of a six-dimensional space with many local maxima to calculate it. This is a difficult task for most states.

2.3 Fifth Independent EM

In section 2.2 it was shown that all EMs must be invariant under LU and hence are determined by the orbit of the state. For three qubit states this means that EMs are a function of only the polynomial invariants, DD coefficients or MD coefficients. In fact, this determination is unique.

Theorem 1 *The set of all EMs for any multipartite pure state, $|\psi\rangle$, uniquely determine the orbit of the state.*

Proof. Suppose two states $|\psi\rangle$ and $|\phi\rangle$ in $\mathcal{H}_1 \otimes \mathcal{H}_2 \ldots \otimes \mathcal{H}_n$ have the same values for the EMs but lie in different orbits. We know by using equation (2.4) that

$$P(|\psi\rangle \to |\phi\rangle) = P(|\phi\rangle \to |\psi\rangle) = 1 \tag{2.28}$$

so $|\psi\rangle$ can be transformed to $|\phi\rangle$ (and vice versa) by n-party LOCC, n-LOCC, with probability 1. Since EMs are non-increasing with any n-LOCC, they must remain constant during the entire transformation from $|\psi\rangle$ to $|\phi\rangle$ (and vice versa). Also, we know that any EM between a system $X = A, B, \ldots$ and the rest of the systems thought of as one (e. g., between B and $(ACD\ldots)$), I will call these EMs 2-EMs, is also an EM for multipartite states. This is because any n-LOCC on the multipartite state is also a 2-LOCC between X and the rest of the systems, since the 2-EM is non-increasing over 2-LOCC it must also be non-increasing over n-LOCC. In particular, the sum of the lowest k eigenvectors of the reduced density matrices,

$$E_k^X(|\psi\rangle) = \sum_{i=1}^{k} \lambda_i^{\uparrow}(\rho_X(|\psi\rangle)), \tag{2.29}$$

(i. e., the 2-EMs in equation (2.2)) must be EMs. So the $E_k^X(|\psi\rangle)$ must remain unchanged and hence the spectrum of ρ_X is unchanged during the transformation from $|\psi\rangle$ to $|\phi\rangle$. In particular, a measurement on space X, given by A_1 and A_2, must be such that

$$\rho_X\left(\frac{A_i|\psi\rangle}{\sqrt{N}}\right) = U\rho_X\left(|\psi\rangle\right)U^{\dagger} \tag{2.30}$$

where N is the normalization. The only way this can be satisfied is if $\frac{A_i}{\sqrt{N}}$ is a unitary matrix. This means that $|\psi\rangle$ and $|\phi\rangle$ are unitarily equivalent which contradicts our original supposition. □

Since we know there are 5 parameters that determine the orbit of a three qubit state, then by theorem 1 there must be 5 independent, continuous EMs. To the best of the author's knowledge, the only 4 known independent continuous EMs that don't require a difficult maximization over a multidimensional space are the four τ EMs defined in equation (2.12). Any candidate for the fifth independent EM must depend on I_4 since the τ are invertible functions of I_1, I_2, I_3 and I_5 respectively. The following function fulfills that criterion:

$$\sigma_{ABC} = 3 - (I_1 + I_2 + I_3)I_4 \tag{2.31}$$

and numerical results suggest that it is an EM. After generating over 300,000 random states and applying a random operation to each of them, the inequality in equation (2.5) was never violated by σ_{ABC}. Also, note that σ_{ABC} is symmetric in particle permutations as is τ_{ABC}. For the duration of this paper I will assume that σ_{ABC} is an EM. Indeed, it may be that there is a set of measure zero or perhaps just a very small measure for which σ_{ABC} is not a monotone and my numerical test didn't explore this space but there must exist some function of the polynomial invariants which is independent of the τs and is an EM. For it to be useful in improving our upper bound for $P\left(|\psi\rangle \rightarrow |\phi\rangle\right)$, there should be pairs of states $|\psi\rangle$ and $|\phi\rangle$ such that

$$\frac{\sigma_{ABC}(|\psi\rangle)}{\sigma_{ABC}(|\phi\rangle)} < \min_{\tau} \frac{\tau\left(|\psi\rangle\right)}{\tau\left(|\phi\rangle\right)} \tag{2.32}$$

and I have found such states numerically. The largest value of

$$\frac{\sigma_{ABC}(|\psi\rangle)}{\sigma_{ABC}(|\phi\rangle)} - \min_{\tau} \frac{\tau\left(|\psi\rangle\right)}{\tau\left(|\phi\rangle\right)} \tag{2.33}$$

that I found in my limited number of examples was 0.01 and I was able to find examples of states for which $\tau\left(|\psi\rangle\right)/\tau\left(|\phi\rangle\right)$ is greater than one for all τ and $\sigma_{ABC}(|\psi\rangle)/\sigma_{ABC}(|\phi\rangle)$ is less than one.

2.4 Other EMs and the Discrete Invariant

The five independent continuous EMs, $\tau_{(AB)C}$, $\tau_{(AC)B}$, $\tau_{(BC)A}$, τ_{ABC} and σ_{ABC}, can easily be inverted to find I_1 - I_5 but to completely determine the orbit of a state we must also have an EM that will give us the value of the discrete invariant I_6. This is equivalent to finding an EM that is not the same for a state and it complex conjugate state. Note that $I_1, \ldots I_5$ and hence the τ and σ_{ABC} do not change when a state is conjugated, but by looking at any of the sets of LU invariants we can see that generically a state is not LU equivalent to its conjugate. By looking at equation (2.4) we can see that this implies that there must be EMs that are not the same for the generic state and its conjugate. It is also easy to see that for any operation that takes a state $|\psi\rangle$ to its conjugate $|\bar{\psi}\rangle$ with probability p, there is an operation that takes $|\bar{\psi}\rangle$ to $|\psi\rangle$ with the same probability. So, for a generic state $|\psi\rangle$ there must be an EM that goes down for the operation $|\psi\rangle \to |\bar{\psi}\rangle$ and a similar one that goes down the same amount for $|\bar{\psi}\rangle \to |\psi\rangle$. So, EMs of the following form must exist:

$$v^{\pm}\left(|\psi\rangle\right) = \begin{cases} v + v' & \pm I_6 = 1 \\ v & o.w. \end{cases} \tag{2.34}$$

where v and v' are functions of $\tau_{(AB)C}$, $\tau_{(AC)B}$, $\tau_{(BC)A}$, τ_{ABC} and σ_{ABC}.

Also, from [10] we know that there are two classes of three-part entangled states (i. e., states with $\tau_{(AB)C}, \tau_{(AC)B}, \tau_{(BC)A} > 0$) that can be converted into each other with some non-zero probability within the class and zero probability between the classes. Namely, the GHZ-class which contains

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right) \tag{2.35}$$

and has non-zero τ_{ABC} and the W-class which contains

$$|\mathrm{W}\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right) \tag{2.36}$$

and has $\tau_{ABC} = 0$. Looking again at equation (2.4), we see that τ_{ABC} tells us that $P(|\psi_{\mathrm{W}}\rangle \to |\psi_{\mathrm{GHZ}}\rangle) = 0$ but none of the previously defined EMs tell us that $P(|\psi_{\mathrm{GHZ}}\rangle \to |\psi_{\mathrm{W}}\rangle) = 0$. Since the only way to get $P(|\psi_{\mathrm{GHZ}}\rangle \to |\psi_{\mathrm{W}}\rangle) = 0$ is to have an EM that is finite for GHZ-class states and infinite for W-class states or zero for GHZ-class states and non-zero for W-class states, such an EM must exist.

2.5 Finding a Minimal Set

Since $\tau_{(AB)C}$, $\tau_{(AC)B}$, $\tau_{(BC)A}$, τ_{ABC}, σ_{ABC} and $v^{\pm}$ determine the orbit of the state, all other EMs must depend on them. A fairly general way to create new EMs from known EMs is to use what I will call f-type functions.

Definition 1 *A function $f : \mathcal{S} \subset \Re^n \to \Re$ is an f-type function if it satisfies the following:*

1. $f(\vec{0}) = 0$

2. *if* $x_i \geq y_i$ *for all* $i = 1, 2, \dots n$ *then* $f(\vec{x}) \geq f(\vec{y})$ *for* $\vec{x}, \vec{y} \in \mathcal{S}$

3. $f(p\vec{x} + (1-p)\vec{y}) \geq pf(\vec{x}) + (1-p)f(\vec{y})$ *for any* $\vec{x}, \vec{y} \in \mathcal{S}$ *and* $0 \leq p \leq 1$.

For a set of EM, $\{E_i\}$, we have

$$E_i(|\psi\rangle) \geq pE_i\left(\frac{A_1|\psi\rangle}{\sqrt{p}}\right) + (1-p)E_i\left(\frac{A_2|\psi\rangle}{\sqrt{1-p}}\right) \tag{2.37}$$

for any measurement A_1, A_2 and any state $|\psi\rangle$. So, we have

$$\begin{aligned} f[\vec{E}(|\psi\rangle)] &\geq f\left[p\vec{E}\left(\frac{A_1|\psi\rangle}{\sqrt{p}}\right) + (1-p)\vec{E}\left(\frac{A_2|\psi\rangle}{\sqrt{1-p}}\right)\right] \\ &\geq pf\left[\vec{E}\left(\frac{A_1|\psi\rangle}{\sqrt{p}}\right)\right] + (1-p)f\left[\vec{E}\left(\frac{A_2|\psi\rangle}{\sqrt{1-p}}\right)\right] \end{aligned} \tag{2.38}$$

where the first inequality comes from property 2 and the second comes from property 3. Hence, $f(E_1, \dots, E_m)$ is also an EM. We can show that any EM $f(E_1, \dots, E_m)$ that is an f-type function of monotones $E_1, \dots E_m$ does not modify the upper bound on $P(|\psi\rangle \to |\phi\rangle)$ given by

$$P(|\psi\rangle \to |\phi\rangle) \leq \min_i \frac{E_i(|\psi\rangle)}{E_i(|\phi\rangle)}. \tag{2.39}$$

First for the one-dimensional case.

Lemma 1 *If* $f(x)$ *is an* f*-type function with* $n = 1$, *then*

$$\frac{f(x)}{f(y)} \geq \min\left\{\frac{x}{y}, 1\right\} \tag{2.40}$$

for any $x, y \in \mathcal{S}$.

Proof.

Case 1 For $x \geq y$ from property 2 we know $f(x) \geq f(y)$ and hence

$$\frac{f(x)}{f(y)} \geq 1. \tag{2.41}$$

Case 2 For $x < y$ if we choose $p = \frac{x}{y} \epsilon [0,1)$, then we know from properties 1 and 3 that $f(py) \geq pf(y)$ and so

$$\frac{f(x)}{f(y)} \geq \frac{x}{y}. \square \tag{2.42}$$

For n dimensions we have the following theorem (proved with S. Daftuar and D. Whitehouse).

Theorem 2 *If $f(x)$ is an f-type function, then*

$$\frac{f(\vec{x})}{f(\vec{y})} \geq \min\left\{\frac{x_i}{y_i}, 1\right\} \qquad i = 1, 2, \ldots n \tag{2.43}$$

for $\vec{x}, \vec{y} \,\epsilon\, \mathcal{S}$.

Proof. Let

$$c = \min\left\{\frac{x_i}{y_i}\right\} \tag{2.44}$$

then we have

Case 1 If $c \geq 1$ then from property 2 $f(\vec{x}) \geq f(\vec{y})$ and so

$$\frac{f(\vec{x})}{f(\vec{y})} \geq 1. \tag{2.45}$$

Case 2 If $c < 1$ then define

$$z_i = \frac{x_i}{c} \qquad i = 1, 2, \ldots n \tag{2.46}$$

and $g(r) = f(r\vec{z})$. Notice that $g(r)$ is an f-type function with $n = 1$ and hence

$$\frac{g(c)}{g(1)} \geq c \tag{2.47}$$

or substituting in f we have

$$\frac{f(\vec{x})}{f(\vec{z})} \geq c. \tag{2.48}$$

Using $z_i \geq y_i$ and property 2 we have

$$\frac{f(\vec{x})}{f(\vec{y})} \geq c. \square \tag{2.49}$$

For three-qubit states if we take the minimum of $E(|\psi\rangle)/E(|\phi\rangle)$ over $\mathcal{E} = \{\tau_{(AB)C}, \tau_{(AC)B}, \tau_{(BC)A}, \tau_{ABC}, \sigma_{ABC}, v^{\pm}\}$ we are actually taking the minimum over the infinite set of all f-type functions of $\mathcal{E}$. Although from theorem 1 we know that all EMs must be a function of $\mathcal{E}$, it is possible that there exist EMs that are not f-type functions of $\mathcal{E}$. These EMs could cause $P(|\psi\rangle \rightarrow |\phi\rangle)$ to be lower than the minimum of $E(|\psi\rangle)/E(|\phi\rangle)$ over $\mathcal{E}$. The EM mentioned at the end of section 2.4 is an example of such an EM.

2.6 Conclusions and Further Research

Theorem 1 along with theorem 2 implies that there should be a (not necessarily finite) minimal set of EMs, M, for which all EMs for three-qubit states or similarly for any type of multipartite states are f-type functions of M. I conjecture that such a minimal set should be simple since the f-type functions seem to be a rather general way of creating EMs that are functions of other EMs. The difficult part seems to be finding the EMs that are minimal and showing that they are minimal. Using numerical results it seems that the τ may be minimal. I looked at functions of the τ that are almost but not quite f-type such as $\tau^{1.01}$ and numerically tested whether they are EMs or not. None of them were EMs. I cannot say the same for σ_{ABC} and definitely not for $v^{\pm}$ since I do not have an explicit form for the v.

There is further research that may help these problems. If one could invert the equations in (2.26) to write a, b, c, d, f and ϕ in terms of $I_1, \ldots, I_6$ that would allow us to calculate the EM $1 - a^2$ not to mention find the ranges for and calculate the values of a, b, c, d, f and ϕ. The EM $1 - a^2$ could be used to replace σ_{ABC}, or perhaps as an addition to $\mathcal{E}$, and may prove more useful than σ_{ABC}. As far as finding the minimal EMs and showing that they are minimal, the arbitrary measurement on the DD at the end of section (2.2.2) may be useful since it allows us to look at the value of $I_1, \ldots, I_6$ before and after an arbitrary measurement on an arbitrary state with far less parameters than if we didn't take out the LU freedom. Also, it may be able to tell us the maximal probability of transforming the general complex state $|\psi\rangle$ to its conjugate state $|\bar{\psi}\rangle$ and this is a crucial piece of information that is needed to calculate v' in equation (2.34). Unfortunately, most of these tasks involve trying to solve nontrivial equations or systems of equations with many variables which can be difficult or even impossible.

Chapter 3

Generalized Quantum Search

3.1 Introduction

The field of quantum computing has undergone a rapid growth over the past few years. Simple quantum computations have already been performed using nuclear magnetic resonance [20, 21, 22, 23, 24, 25] and nonlinear optics technologies [26, 27]. Recently, proposals for specialized devices that rely on quantum computing have also been made [28]. Such devices are far from being general-purpose computers, nevertheless, they constitute significant milestones along the road to practical quantum computing.

In tandem with these hardware developments, there has been a parallel development of new quantum algorithms. Several important quantum algorithms are now known [29, 30, 31, 32, 33, 34]. Of particular importance is the quantum algorithm for performing unstructured quantum search discovered by Lov Grover in 1996 [31]. Further analysis of this algorithm is given by Jozsa [35] and an optical implimentation is given by Kwiat [36]. Grover's algorithm is able to find a marked item in a virtual "database" containing N items in $O(\sqrt{N})$ computational steps. In contrast, the best classical algorithm requires $O(N/2)$ steps on average, and $O(N)$ steps in the worst case. Thus Grover's algorithm exhibits a quadratic speedup over the best classical counterpart.

Although Grover's algorithm exhibits only a polynomial speedup, it appears to be much more versatile than the other quantum algorithms. Indeed, Grover has shown how his algorithm can be used to speed up almost any other quantum algorithm [37]. More surprisingly, even search problems that contain "structure" in the form of correlations between the items searched over often reduce to an exhaustive search amongst a reduced set of possibilities. Recently, it was shown how Grover's algorithm can be nested to exploit such problem structure [32]. This is significant because NP-hard problems, which are amongst the most challenging computational problems that arise in practice, possess exactly this kind of problem structure.

In order to appreciate the full versatility of Grover's algorithm, it is important to examine all the ways in which it might be generalized. For example, whereas the original Grover algorithm was started from an equally weighted superposition of eigenstates representing all the indices of the items in the database, a natural generalization would be to consider how it performs when started from an arbitrary initial superposition instead. This refinement is important, because if Grover's algorithm is used within some larger quantum computation, it is likely to have to work on a arbitrary starting superposition rather than a specific starting eigenstate. Similarly, the original Grover algorithm uses a particular unitary operator, the Walsh-Hadamard operator, as the basis for a sequence of unitary operations that systematically amplifies the amplitude in the target state at the expense of the amplitude in the non-target states. However, it is now known that this is not the best choice if there is partial information as to the likely location of the target item in the database. In such a situation a different unitary operator is desirable [38]. Hence, it is equally important to understand how Grover's algorithm performs when using an arbitrary unitary operator instead of the Walsh-Hadamard operator.

Each of these refinements have been analyzed in detail *separately* : Biham et al. have considered the case of an arbitrary starting superposition [39], while Grover considered the case of an arbitrary unitary operator [38]. In this paper, we present the analysis of the fully generalized Grover algorithm in which we incorporate both of these effects simultaneously. Our goal is to determine the exact analytic formula for the probability of the fully generalized Grover algorithm succeeding after n iterations when there are r targets amongst N candidates. Having obtained this formula, we will recover the Biham et al. and Grover results as special cases. We will then show that the optimal strategy, on average, for using the fully generalized Grover algorithm consists of measuring the memory register after about 12% fewer iterations than are needed to obtain the maximum probability of success. This result confirms a more restricted case reported in [40]. Finally, we show how to boost the success probability and reduce the required coherence time by running a society of k quantum searches independently in parallel. In particular, we derive an explicit formula connecting the degree of parallelism, i.e., k, to the optimal number of iterations (for each agent in the society) that minimizes the expected search cost overall. We then derive the expected cost of optimal k-parallel quantum search.

3.2 Grover's Algorithm

The problem we have to solve is the following. Given a function $f(x_i)$ on a set $\mathcal{X}$ of input states such that

$$f(x_i) = \begin{cases} 1 & \text{if } x_i \text{ is a target element} \\ 0 & \text{otherwise} \end{cases} . \tag{3.1}$$

How do we find a target element by using the least number of calls to the function $f(x_i)$? In general, there might be r target elements, in which case any one will suffice as the answer.

To solve the problem using Grover's algorithm we first form a Hilbert space with an orthonormal basis element for each input $x_i \in \mathcal{X}$. In this paper, we refer to the basis of input eigenstates as the measurement basis. Let $N = |\mathcal{X}|$ be the cardinality of $\mathcal{X}$. Without loss of generality, we will write the target states as $|t_i\rangle$ (with $i = 1, \cdots r$), and the non-target states as $|l_i\rangle$ (with $i = 1, \cdots N - r$). The function call is to be implemented by a unitary operator that acts as follows:

$$|x_i\rangle|y\rangle \rightarrow |x_i\rangle|y \oplus f(x_i)\rangle \tag{3.2}$$

where $|y\rangle$ is either $|0\rangle$ or $|1\rangle$. By acting on

$$\left(\sum_{i=1}^{N-r} l_i|l_i\rangle + \sum_{j=1}^{r} k_j|t_j\rangle\right) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{3.3}$$

with this operator we construct the state

$$\left(\sum_{i=1}^{N-r} l_i|l_i\rangle - \sum_{j=1}^{r} k_j|t_j\rangle\right) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{3.4}$$

where the r measurement basis states $|t_i\rangle$ are the target states and the $N - r$ measurement basis states $|l_i\rangle$ are the non-target states. If we now disregard the state $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ then all we have done is to invert the phase of the target states. Hence, the operator we have achieved is equivalent to the operator

$$1 - 2\sum_{i=1}^{r} |t_i\rangle\langle t_i| \tag{3.5}$$

although we emphasize that this operation can be performed without knowing the target states *explicitly* but only through the knowledge of $f(x)$.

Next we construct the operator Q defined as

$$Q = -\left(1 - 2|a\rangle\langle a|\right)\left(1 - 2\sum_{i=1}^{r} |t_i\rangle\langle t_i|\right) \tag{3.6}$$

where $|a\rangle$ can be thought of as the state with respect to which an "inversion" is performed. Different choices of $|a\rangle$ give rise to different unitary operators for performing amplitude amplification. In the original Grover algorithm, the state $|a\rangle$ was chosen to be

$$|a\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \mathcal{X}} |x\rangle \tag{3.7}$$

and was obtained by applying the Walsh-Hadamard operator, U, to a starting state $|s\rangle$, i.e., $|a\rangle = U|s\rangle$. Hence, the operation $2|a\rangle\langle a| - 1$, which Grover referred to as "inversion about the average," is equivalent to $-UI_sU^\dagger$ with U being the Walsh-Hadamard operator and I_s being $1 - 2|s\rangle\langle s|$. By knowing more about the structure of the problem, we can choose other vectors $|a\rangle$ that will allow us to find a target state faster. Techniques for doing this are given in [37].

If we write out Q, we get

$$Q = \sum_{i=1}^{r} |t_i\rangle\langle t_i| - \sum_{j=1}^{N-r} |l_j\rangle\langle l_j| + 2|a\rangle\langle a| - 4v|a\rangle\langle t| \tag{3.8}$$

where $|t\rangle$, the normalized projection of $|a\rangle$ onto the space of target states, is given by

$$|t\rangle = \frac{1}{v}\sum_{i=1}^{r} \langle t_i|a\rangle |t_i\rangle \qquad v^2 = \sum_{i=1}^{r} |\langle t_i|a\rangle|^2. \tag{3.9}$$

We can see from this that Q only acts non-trivially on the space spanned by $|a\rangle$ and $|t\rangle$. We can make these vectors an orthonormal basis for this space by using

$$|l\rangle = \frac{1}{\sqrt{1-v^2}}\left(|a\rangle - v|t\rangle\right) \tag{3.10}$$

instead of $|a\rangle$. The vector $|l\rangle$ is just the normalized projection of $|a\rangle$ onto the space of non-target states. The rest of the space (i. e., the space orthogonal to $|t\rangle$ and $|l\rangle$) can be broken up into the space of target states ($\mathcal{S}_T$) and non-target states ($\mathcal{S}_L$). We can now write Q as

$$Q = \cos\phi\left(|t\rangle\langle t| + |l\rangle\langle l|\right) + \sin\phi\left(|t\rangle\langle l| - |l\rangle\langle t|\right) + I_T - I_L, \tag{3.11}$$

where I_T and I_L are the identity operators on ($\mathcal{S}_T$) and ($\mathcal{S}_L$), respectively, and $\phi \equiv \arccos\left[1 - 2v^2\right]$. From this we can see that Q is just a simple rotation matrix on the two-dimensional space spanned by $|l\rangle$ and $|t\rangle$, and acts trivially on the rest of the space. The operator Q has been independently shown by Jozsa [35] to be an exact rotation in the special case of one solution and with $|a\rangle$ given by equation (3.7).

An arbitrary starting superposition $|s\rangle$ for the algorithm can be written as

$$|s\rangle = \alpha|t\rangle + \beta e^{ib}|l\rangle + |s_t\rangle + |s_l\rangle \tag{3.12}$$

where the states $|s_t\rangle$ and $|s_l\rangle$ (which must have a norm less than one if the state $|s\rangle$ is to be properly normalized overall) are the components of $|s\rangle$ in ($\mathcal{S}_T$) and ($\mathcal{S}_L$) respectively. Also, α, β and b are positive real numbers. After n applications of Q on an arbitrary starting superposition $|s\rangle$, we have

$$Q^n|s\rangle = \left(\alpha\cos(n\phi) + \beta e^{ib}\sin(n\phi)\right)|t\rangle + \left(\beta e^{ib}\cos(n\phi) - \alpha\sin(n\phi)\right)|l\rangle + |s_t\rangle + (-1)^n|s_l\rangle. \tag{3.13}$$

If we measure this state our probability of success (i.e., measuring a target state) will be given by two terms. The first term is the squared magnitude of $Q^n|s\rangle$ projected into the space $\mathcal{S}_T$. It is equal to $\langle s_t|s_t\rangle$ and is unchanged by Q. The second term is the squared magnitude of the component of $|t\rangle$ which is given by

$$\begin{aligned} g(n) &\equiv |\langle t|Q^n|s\rangle|^2 \\ &= \left|\alpha\cos(n\phi) + \beta e^{ib}\sin(n\phi)\right|^2 \\ &= \tfrac{\alpha^2+\beta^2}{2} + \tfrac{\alpha^2-\beta^2}{2}\cos(2n\phi) + \alpha\beta\cos(b)\sin(2n\phi) \\ &= \tfrac{\alpha^2+\beta^2}{2} - \tfrac{1}{2}\left|\alpha^2+\beta^2 e^{2ib}\right|\cos(2n\phi+\psi) \end{aligned} \tag{3.14}$$

where $\psi \equiv \arccos\left[\frac{\beta^2-\alpha^2}{|\alpha^2+\beta^2 e^{2ib}|}\right]$. This is the term that is affected by Q, and is the term we wish to maximize. The probability of success after n iterations of Q acting on $|s\rangle$ is thus

$$p(n,r,N) = \langle s_t|s_t\rangle + g(n). \tag{3.15}$$

Assuming that n is continuous (an assumption that we will justify shortly), the maxima of $g(n)$, and hence the maxima of the probability of success of Grover's algorithm, are given by the following:

$$n_j = \frac{-\psi + (1+2j)\pi}{2\phi} \qquad j = 0,1,2,\cdots \tag{3.16}$$

The value of $g(n)$ at these maxima is given by

$$g(n_j) = \frac{\alpha^2+\beta^2}{2} + \frac{1}{2}\left|\alpha^2+\beta^2 e^{2ib}\right|. \tag{3.17}$$

In practice, the optimal n must be an integer and typically the n_j's are not integers. However, since $g(n)$ can be written as

$$g(n_j \pm \delta) = g(n_j) - \phi^2\left|\alpha^2+\beta^2 e^{2ib}\right|\delta^2 + O(\delta^4) \tag{3.18}$$

around n_j and most interesting problems will have $v \ll 1$ and hence $\phi \simeq 2v \ll 1$, simply rounding n_j to the nearest integer will not significantly change the final probability of success. So, we have

$$p(n_{max}, r, N) = \frac{\alpha^2 + \beta^2}{2} + \frac{1}{2}\left|\alpha^2 + \beta^2 e^{2ib}\right| + \langle s_t | s_t \rangle - O(v^2) \tag{3.19}$$

as the probability of measuring a target state after n_{max} applications of Q.

3.3 Recovering the Special Cases

As a check on our fully generalized formula for the probability of success after n iterations, we attempt to recover the corresponding formulae obtained in the analyses of Biham et al. (for a fixed unitary operator and an arbitrary starting superposition) [39] and Grover (for an arbitrary unitary operator and a fixed starting superposition) [38].

In the case of Biham et al., the starting state is arbitrary, but the averaging state $|a\rangle$ is given by

$$|a\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \mathcal{X}} |x\rangle. \tag{3.20}$$

In this case

$$\begin{aligned} v &= \sqrt{\tfrac{r}{N}} \\ |t\rangle &= \tfrac{1}{\sqrt{r}} \textstyle\sum_{i=1}^{r} |t_i\rangle \\ |l\rangle &= \tfrac{1}{\sqrt{N-r}} \textstyle\sum_{i=1}^{N-r} |l_i\rangle. \end{aligned} \tag{3.21}$$

In the analysis of [39] they use $\overline{k}(0)$ and $\overline{l}(0)$ to represent the average amplitudes, in $|s\rangle$, of the target and non-target states respectively, and σ_k and σ_l to represent the standard deviations of those amplitudes. With some algebra one can see that the following relationships connect our notation to theirs:

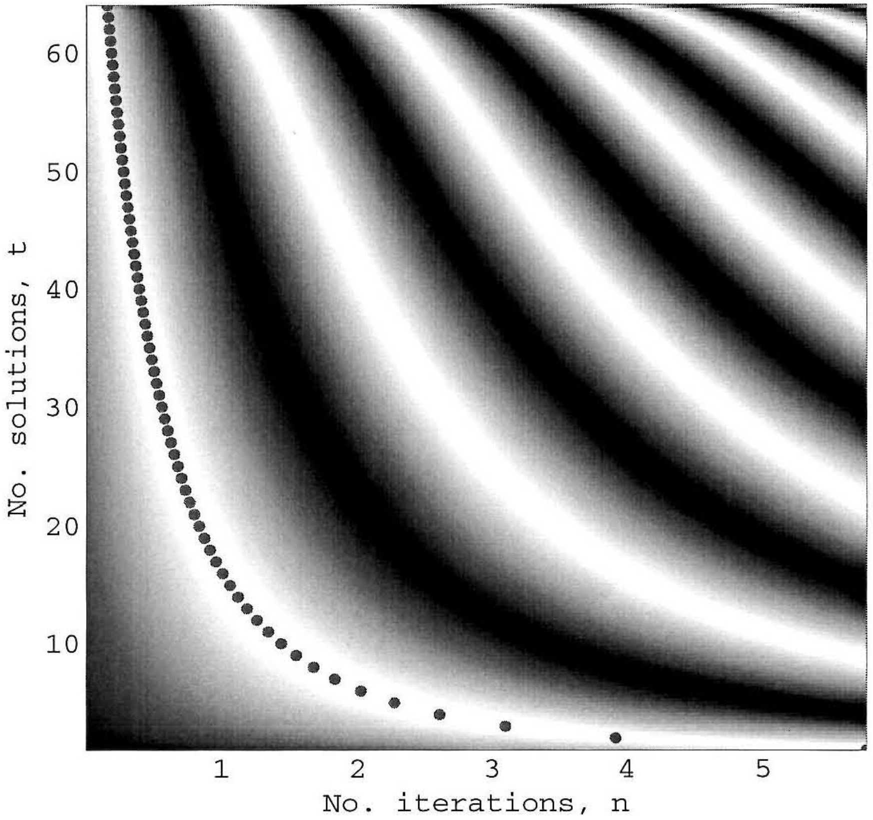

Figure 3.1: Plot of the probability of success of Grover's algorithm after n iterations of amplitude amplification when there are r solutions amongst $N = 64$ possibilities. White regions correspond to probability 1, black regions correspond to probability 0. Note that the success probability is periodic in the number of amplitude amplification iterations for a fixed number of solutions.

$$\begin{aligned} \alpha &\longrightarrow \overline{k}(0)\sqrt{r} \\ \beta e^{ib} &\longrightarrow \overline{l}(0)\sqrt{N-r} \\ \langle s_t|s_t\rangle &\longrightarrow r\sigma_k^2 \\ \langle s_l|s_l\rangle &\longrightarrow (N-r)\sigma_l^2 \\ \phi &\longrightarrow \omega \\ \psi &\longrightarrow 2\mathrm{Re}[\phi] \\ n &\longrightarrow t \\ n_0 &\longrightarrow T. \end{aligned} \tag{3.22}$$

By substituting these relationships into equations (3.14), (3.16), and (3.19), one reproduces the results of [39].

The second special case, in which $|a\rangle$ (with respect to which inversion is done) is an unknown normalized vector, while $|s\rangle$ is given by

$$|s\rangle = |a\rangle = \sqrt{1-v^2}|l\rangle + v|t\rangle \tag{3.23}$$

was considered by Grover. Hence, $\alpha = v$, $\beta = \sqrt{1-v^2}$ and $b = 0$. Also, $|s_t\rangle = |s_l\rangle = 0$. These substitutions lead to $\psi = \phi$. Plugging this into equations (3.16) and (3.19), we get

$$n_{max} = \frac{\pi}{2\phi} - \frac{1}{2} = \frac{\pi}{4v} - \frac{1}{2} - \frac{\pi v}{24} + O(v^2) \tag{3.24}$$

and

$$p(n_{max}) = 1 - O(v^2) \tag{3.25}$$

which agree with the results of [38]. If we examine equation (3.15) in this case, we get

$$p(n) = \frac{1-\cos((1+2n)\phi)}{2} = \sin^2\left((1+2n)\phi/2\right) \tag{3.26}$$

as the probability of measuring a target state after n iterations of Q.

3.4 Application of the Formula for $p(n)$

Next, we show how to apply our analytic formula for the probability of success after n iterations, $P(n)$, to slightly speed up the quantum unstructured search algorithm. Although the speedup we obtain is not dramatic, it is worth making the point that it is possible at all as Zalka has proved, correctly, that Grover's algorithm is exactly optimal [41]. Many people have assumed, therefore, that it is impossible to beat Grover's algorithm. However, by combining techniques of quantum computing with those of classical computing, we show that it is possible to do a little bit better than Grover's algorithm on average. The result we report was apparently discovered previously by Boyer et al. [40] and later by Zalka [42] in the case where $|a\rangle$ is a uniform superposition [as in equation (3.7)]. It is shown here to persist for the more general case when $|s\rangle$ is arbitrary but equal to $|a\rangle$ which is the case treated in [38].

We consider a punctuated quantum search algorithm that works as follows:

Algorithm: Punctuated Quantum Search

1. Run the quantum search algorithm for n iterations.

2. Read the memory register.

3. If the result is a target state halt; else, reset the register to the starting superposition and return to step 1.

The average time, $T_{avg}(n)$, it will take to find a target state if we stop the generalized quantum search algorithm after n iterations of Q is

$$\begin{aligned} T_{avg}(n) &= \sum_{i=1}^{\infty} (1-p(n))^{i-1}\, p(n)\ i\ n \\ &= \frac{n}{p(n)} \\ &= \frac{2n}{1-\cos[(1+2n)\phi]}. \end{aligned} \tag{3.27}$$

We can find the optimal strategy, i.e., the best number of iterations to use before we attempt to measure the register, by minimizing the expected running time T_{avg}. To do this, we set the derivative of T_{avg} to zero and solve for $n = n_{opt}$:

$$\frac{\partial T_{avg}}{\partial n} = \frac{2 - 2\cos[(1+2n)\phi] - 4n\phi\sin[(1+2n)\phi]}{(1-\cos[(1+2n)\phi])^2} = 0. \tag{3.28}$$

Typically, n will be much larger than one, so we can make the approximation $(1+2n)\frac{\phi}{2} \simeq n\phi \equiv x$, so that we obtain

$$\begin{aligned} 1 - \cos 2x &= 2x \sin 2x \\ 2\sin^2 x &= 4x \sin x \cos x \\ 2x &= \tan x. \end{aligned} \tag{3.29}$$

which gives $x_{opt} = 1.1656$ as the lowest positive solution. This solution corresponds to the minimum of the function. Hence the optimal value of n is

$$n_{opt} \simeq \frac{x_{opt}}{\phi} = \frac{1.1656}{\phi}. \tag{3.30}$$

This value of n gives a probability of success of

$$p(n_{opt}) = \sin^2 x_{opt} = 0.8446 \tag{3.31}$$

at each measurement, and corresponds to an average number of iterations of

$$T_{avg}(n_{opt}) \simeq \frac{1}{\phi}\frac{x_{opt}}{\sin^2 x_{opt}} = \frac{1.3801}{\phi}. \tag{3.32}$$

This must be compared to $\frac{\pi}{2\phi} = \frac{1.5708}{\phi}$ iterations if we run Grover's algorithm until the probability is maximal. Thus, we get a 12% reduction of the average computation time by making use of a punctuated algorithm.

It is interesting to note that, if we restrict the analysis some more to the case where $|a\rangle$ is a uniform superposition and where there is only one target state, then we have $\phi = 2/\sqrt{N}$, so that $T_{avg}(n_{opt}) \simeq 0.6900\sqrt{N}$. This is faster than the lower bounds in [43], [44], [40], and [41], but we are using a somewhat different model. They are looking at the minimum time it would take without measuring to find a solution with certainty up to errors from rounding n_{max} to the nearest integer. Instead, the model we use here allows for punctuated measurements and resets of the quantum search algorithm. Nevertheless, the punctuated quantum search algorithm is faster on average. Note that we have assumed that the time it takes to measure, check if a solution was reached, and reset the algorithm is negligible. This is reasonable as checking a solution only requires one function call.

The punctuated quantum search algorithm has another advantage in that it is less sensitive to decoherence. If we wait until we have the maximal probability of measuring a target state, then we must maintain coherence for $\frac{1.5708}{\phi}$ steps as compared to only $\frac{1.1655}{\phi}$ steps for the fastest measure and restart method. This is because we do not need to maintain coherence through the measurment stage of this method. In fact, the punctuated search that takes the same number of steps on average as the standard or maximal probability method (i.e., $\frac{\pi}{2\phi} = \frac{1.5708}{\phi}$ steps) need only maintain coherence for $\frac{\pi}{4\phi} = \frac{0.7854}{\phi}$ steps at a time. This represents only 50% of the coherence time required in the standard Grover method, and corresponds to waiting for a 50% probability of success and then measuring.

3.5 k-Parallel Quantum Search

A way to speed up Grover's algorithm still further is to have a society of k computational agents all running Grover's algorithm independently at the same time. This is promising because the standard deviation

$$\sigma_T = \frac{n}{p(n)}\sqrt{[1-p(n)]} \tag{3.33}$$

in the computation time of punctuated quantum search is fairly large, and hence having multiple searches running may give a considerable speed up.

Suppose that we know that there are exactly r solutions amongst N candidates. Given $p(n,r,N)$, the probability of success for a single agent after n iterations, we can boost the success probability by using k agents acting in parallel. In particular, the probability that at least one agent, in a society of k independent agents, succeeds after each agent has undergone n iterations is given by

$$p_k(n) = 1 - (1-p(n))^k \,. \tag{3.34}$$

Thus, the expected cost, $T^{(k)}_{avg}$, of performing k-parallel quantum search is given by

$$T^{(k)}_{avg}(n) = \sum_{j=1}^{\infty} p_k(n)(1-p_k(n))^{j-1}\; j\; n = \frac{n}{p_k(n)} = \frac{n}{1-\cos^{2k}\left((1+2n)\frac{\phi}{2}\right)} \tag{3.35}$$

As in equation (3.27) we can find the value of n that minimizes the expected cost. To find the mimimum, we find where $\frac{\partial T^{(k)}_{avg}(n)}{\partial n}$ is equal to zero. This derivative is given by

$$\frac{\partial T^{(k)}_{avg}(n)}{\partial n} = \frac{1-\cos^{2k}\left((1+2n)\frac{\phi}{2}\right)\left(1+2k\; n\phi\; \tan\left((1+2n)\frac{\phi}{2}\right)\right)}{\left(1-\cos^{2k}\left((1+2n)\frac{\phi}{2}\right)\right)^2}. \tag{3.36}$$

For $\frac{r}{N} \ll 1$, i.e., when there are very few solutions amongst the items searched over, we have $\phi = \arccos(1-\frac{2r}{N}) \approx 2\sqrt{\frac{r}{N}}$. As before, substituting $x \equiv n\phi \simeq (1+2n)\phi/2$ and realizing that $n \gg 1$, we obtain

$$\frac{\partial T^{(k)}_{avg}(n)}{\partial n} \approx \frac{1-\cos^{2k}(x)\,(1+2kx\tan(x))}{(1-\cos^{2k}(x))^2}. \tag{3.37}$$

In order to find the minimum, we thus have to solve the transcendental equation

$$1 - \cos^{2k}(x) = 2k\ x\ \cos^{2k}(x)\tan(x). \tag{3.38}$$

The variable $x < 1$ provided $n < \frac{1}{2}\left(\sqrt{\frac{N}{r}} - 1\right)$. We know that we can solve the problem with near certainty if we iterate Grover's algorithm to the maximum probability state in $O(\frac{\pi}{4}\sqrt{\frac{N}{r}})$ iterations. Hence, for a large enough number of parallel search agents, k, there is a reasonable chance that the optimum number of iterations, $n_{opt}(r, N, k)$ at which the expected search cost is minimized, satisfies the criterion that $x < 1$. We therefore expand equation (3.37) as a series approximation in x about $x = 0$. Actually, it appears that x scales as $O(1/\sqrt{k})$, so it tends to 0 as k tends to infinity. If we make such an expansion up to order x^2, we get

$$\frac{\partial T_{avg}^{(k)}(n)}{\partial n} \simeq \frac{1}{kx^2}\left(-1 + \frac{3k-1}{6}x^2 + \frac{5k^2-1}{20}x^4 + O(x^6)\right). \tag{3.39}$$

As $\frac{\partial T_{avg}^{(k)}(n)}{\partial n} = 0$ is a second-order equation in x^2, it can be solved analytically. Three of the roots are non-physical, but one corresponds to an approximation to the true minimum of $T_{avg}^{(k)}(n)$. Specifically, we find that $T_{avg}^{(k)}(n)$ is minimized when x is given by

$$x_{opt} \simeq \sqrt{\frac{5 - 15k + \sqrt{5}\sqrt{-31 - 30k + 225k^2}}{-3 + 15k^2}}, \tag{3.40}$$

We note that $x < 1$ for all $k \geq 2$, so that the derivation of the optimum formula is self-consistent. This expression for x_{opt} can be expanded in $1/k^{1/2}$, giving

$$x_{opt} \simeq 1.1118\frac{1}{k^{1/2}} + 0.0829\frac{1}{k^{3/2}} + O\left(\frac{1}{k^{5/2}}\right). \tag{3.41}$$

Using $\phi \simeq 2v = 2\sqrt{r/N}$ and equation (3.41), one gets the corresponding expression for $n_{opt} = x_{opt}/\phi$, i.e., the predicted optimal number of iterations for

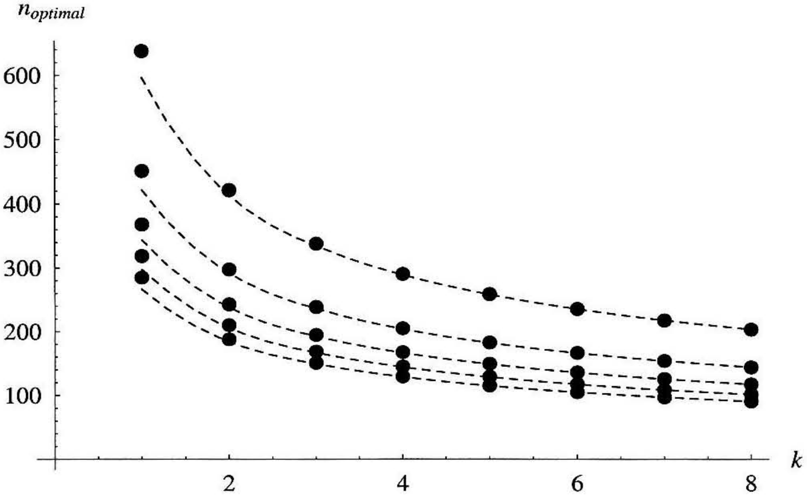

Figure 3.2: Plot of the optimal number of iterations to use in k-parallel quantum search as a function of the degree of parallelism k for $r = 1$ to $r = 5$ solutions (top to bottom in the figure) for the case of a database of size $N = 2^{20}$. The dashed curves correspond to the optima as predicted by our approximate formula for $n_{opt}(r, N, k)$. The points correspond to the exact optima obtained by numerical methods.

each of k quantum searches acting independently in parallel. In Fig. 2, this formula is shown to be in very good agreement with the exact result, obtained by numerical optimization.

Now, if we are only interested in the scaling in N and k of the optimal number of iterations and expected computation time, it is enough to consider the expansion of $\frac{\partial T^{(k)}_{avg}(n)}{\partial n}$ [equation (3.39)] up to order $O(1)$. This simply yields

$$x_{opt} \simeq \sqrt{\frac{6}{3k-1}} \simeq O\left(\frac{1}{\sqrt{k}}\right). \tag{3.42}$$

This formula is only valid at the limit of large k, when x_{opt} tends to zero. The corresponding expression for the optimal number of amplitude amplification iterations is

$$n_{opt} = \frac{x_{opt}}{\phi} \simeq O\left(\frac{1}{\phi\sqrt{k}}\right). \tag{3.43}$$

We can then estimate the expected cost for optimal k-parallel quantum search:

$$T_{avg}^{(k)}(n, r, N) = \frac{1}{\phi}\frac{x_{opt}}{1 - \cos^{2k}(x_{opt})}. \tag{3.44}$$

Again, using the series expansion around $x = 0$, that is, $x/(1 - \cos^{2k}(x)) = 1/(kx) + O(x)$, we get

$$T_{avg}^{(k)}(n, r, N) \simeq \frac{1}{\phi}\frac{1}{kx_{opt}} \simeq O\left(\frac{1}{\phi\sqrt{k}}\right). \tag{3.45}$$

Remembering that $\phi \approx \sqrt{r/N}$, we conclude that T_{avg} scales as $O\left(\sqrt{\frac{N}{rk}}\right)$. Thus, using k agents in parallel simply amounts to having each of them performing a search in a restricted space of size N/k, so that the gain in computation time is of order $O(\sqrt{k})$. Interestingly, this gain is not as good as when parallelizing a classical algorithm.[1] Accordingly, the cumulative time $T_{cumul} = kT_{avg}$, i.e., the sum of the time that all agents spend on quantum search, is *increased* by a factor $O(\sqrt{k})$ with respect to the case of a single agent ($k = 1$).

Our results have implications for the design of prototypical quantum computers. If it is possible to maintain coherence indefinitely, for example, by building fault-tolerance into the computer and by using quantum error correction schemes, our analysis suggests that it is better to use a single agent quantum search. This strategy minimizes the net computational resources expended in solving the problem. However, if coherence time *is* limited, as it most likely

[1] In the latter case, a computation time of order $O(N/r)$ is ideally reduced to $O\left(\frac{N}{rk}\right)$ by using k agents in parallel, so that one has a speedup of order $O(k)$.

will be in prototypical quantum computers, then a parallel punctuated quantum search strategy becomes necessary, with the degree of parallelism set by the desired computation time and desired probability of success. The computational time can be made small by making the degree of parallelism sufficiently large but, of course, at the expense of greater net computational resources being expended on solving the problem.

Let us now consider the situation where the coherence time τ is fixed by some practical considerations, regardless of the value of N and r. The number of agents k must then be of the order of $O\left(\frac{N}{r\tau^2}\right)$ for the parallel time not to exceed the coherence time. This is an interesting result as it implies that the number of agents decreases *quadratically* for an increasing τ. The classical counterpart would be a linear law only.[2] On the other hand, the bad result comes if we reexpress the cumulative computation time for k agents with this value of k:

$$T_{cumul} = kT_{avg} = k\tau = O\left(\frac{N}{r\tau}\right). \tag{3.46}$$

This means that we lose the square root speedup of Grover's algorithm (i.e., T_{serial} does not scale as $\sqrt{N}$) whenever the coherence time is fixed. In order to exploit Grover's quantum speedup, the coherence time τ must necessarily increase as $\sqrt{N}$, i.e., as the square root of the size of the search space.

3.6 Conclusions

In this paper, we have shown how to generalize the analysis of unstructured quantum search to incorporate the effects of an arbitrary starting superposition and an arbitrary unitary operator (or, equivalently, arbitrary state $|a\rangle$) *simulta-*

[2] Classically, if the parallel computation time for each agent is restricted to τ, then the number of agents k should scale as $O\left(\frac{N}{r\tau}\right)$.

neously. We have also shown that, rather than iterating the amplitude amplification operator until the maximum probability of success is attained, i.e., after $O(0.785398\sqrt{N})$ iterations, it is better to measure after only $O(0.6900\sqrt{N})$ iterations (when the probability of success is only 84%). This punctuated strategy is approximately 12% faster than Grover's algorithm on average, and requires a shorter coherence time.

Moreover, an even better quantum search algorithm can be obtained by running k independent quantum searches in parallel, stopping as soon as any of the quantum searches finds a solution. We find that the optimal k-parallel punctuated quantum search strategy is different from that of single agent punctuated quantum search strategy. In general, the higher the degree of classical parallelism, the less (parallel) time is needed to perform the quantum computation. This intuition is captured in equation (3.41), which gives the explicit connection between the optimal number of amplitude amplification iterations $n_{opt} = x_{opt}/\phi$ and the degree of parallelism k. This result is of practical utility to experimental realization of a quantum search algorithm. In particular, in any physical embodiment of a quantum search, there will be some natural coherence time beyond which the computation becomes unreliable. Of course, quantum error correction and fault-tolerant computation allow this time to be extended greatly, arguably indefinitively, if the individual error probability per gate operation can be made sufficiently small. Nevertheless, in practice, this might be extraordinarily difficult to achieve. Instead, if we can predict the degree of parallelism needed so that the quantum search has a good chance of completing within the natural coherence time of the physical system being used as the quantum computer, then the strategy of massive parallelism might provide a realistic alternative to relying solely on quantum error correction. Thus, we see the classical parallelism as an adjunct to quantum error correction rather than a replacement for it. Our results in section refparaQsearch exposes precisely the space/time tradeoff between quantum coherent computing and classical parallelism, at least in the context of unstructured quantum search.

Note: Some of the results obtained in this paper have been derived independently by C. Zalka in a revised (and unpublished [42]) version of Ref. [41].

Chapter 4

Reduction Criterion for Separability

4.1 Introduction

The state of a quantum bipartite system AB is described as *separable* (or classically correlated) if it can be obtained by two parties A and B that prepare their subsystem according to some common instructions (see, e.g., [45, 46]). Mathematically, this means that the density operator ρ characterizing the state of the bipartite system can be written as a convex sum of product states, that is

$$\rho = \sum_i w_i \left(\rho_i^{(A)} \otimes \rho_i^{(B)} \right) \tag{4.1}$$

where the weights w_i satisfy $\sum_i w_i = 1$ and $0 \leq w_i \leq 1$. The w_i's can be viewed as the probability distribution of a classical random variable that is known to both parties A and B and used by them to prepare their subsystem. Namely, if the subsystem A (and B) is prepared in state $\rho_i^{(A)}$ (and $\rho_i^{(B)}$) when the classical variable takes on value i, the state of the joint system AB is given by equation(4.1). A separable state ρ satisfies several interesting properties. The joint statistics of any pair of *local* observables O_A and O_B (measured separately on each subsystem) can be described classically, based on an underlying global "hidden" variable. For example, the quantum expectation value of the product $O_A O_B$ is given by

$$\mathrm{Tr}[\rho(O_A \otimes O_B)] = \sum_i w_i \langle a \rangle_i \langle b \rangle_i \tag{4.2}$$

where $\langle a \rangle_i = \mathrm{Tr}[\rho_i^{(A)} O_A]$ and $\langle b \rangle_i = \mathrm{Tr}[\rho_i^{(B)} O_B]$. In other words, the joint statistics of O_A and O_B can be understood classically, by assuming that the local statistics of the outcomes can be described separately for each $\rho_i^{(A)}$ and $\rho_i^{(B)}$, and that the correlations originate from a hidden variable i distributed according to w_i. Moreover, a separable system always satisfies Bell's inequalities (the converse is not true), so that the latter represent a *necessary* condition for separability (see, e.g., [45]). Note that any joint probability distribution can be represented as a convex combination of product distributions, so that classical probabilities are always separable in the above sense.

The decomposition of a separable state ρ into a convex mixture of product states is not unique in general, but the fact that ρ is separable implies that there must exist at least one such decomposition. If no such decomposition can be found, then ρ is termed *inseparable* or *entangled*, and it can be viewed as *quantum* correlated. Except for the special case where ρ describes a pure state, the distinction between separable and inseparable states appears to be an extraordinarily difficult problem. More precisely, some mixed states can be "weakly" inseparable, in the sense that it is very hard to establish with certainty their inseparability. This is basically due to the difficulty of enumerating explicitly *all* the possible convex combinations of product states in order to detect that a state is actually inseparable. Still, it is possible to find some conditions that *all* separable states must satisfy, therefore allowing the detection of inseparability when a state violates one such condition. The most common example of such a *necessary* condition for separability is the satisfaction of Bell's inequalities. A state that violates Bell's inequalities is inseparable, while a state satisfying them may be separable or weakly inseparable [45].

More recently, a surprisingly simple *necessary* condition for separability has been discovered by Peres [46], which has been shown by Horodecki et al. [47] to be strong enough to *guarantee* separability for bipartite systems of dimension 2×2 and 2×3. If the state ρ is separable, then the operator obtained by applying a *partial* transposition with respect to subsystem A (or B) to ρ must be positive, that is

$$\rho^{T_A} = \left(\rho^{T_B}\right)^* \geq 0. \tag{4.3}$$

Thus, this criterion amounts to checking that all the eigenvalues of the partial transposition of ρ are non-negative, which must be so for all separable states. In Hilbert spaces of dimensions 2×2 and 2×3, this condition is actually *sufficient*, that is, it suffices for ruling out *all* inseparable states [47]. In larger dimensions, however, it is provably *not* sufficient, in the sense that it does not detect some weakly inseparable states [47, 48]. A general necessary *and* sufficient condition for separability in arbitrary dimensions has been found by Horodecki et al. [47], which states that ρ is separable if and only if the tensor product of *any* positive[1] map (acting on A) and the identity (acting on B) maps ρ into a positive operator. Although very important in theory, this criterion is hardly more practical than the definition of separability itself since it involves the characterization of the set of all positive maps. It appears to be useful mainly for 2×2 and 2×3 bipartite systems, where such a general characterization has been found [47].

In this paper, we introduce a positive map, $\Gamma : \rho \to (\mathrm{Tr}\rho) - \rho$, inspired by the structure of the conditional amplitude operator discussed in Ref. [49, 50].

[1] A map is defined as positive if it maps positive operators into positive operators.

This map gives rise to a simple *necessary* condition for separability in arbitrary dimensions. More specifically, it is shown in section 4.2 that any separable state is mapped by the tensor product of Γ (acting on one subsystem, A) and the identity (acting on the other, B) into a non-negative operator. In other words, the eigenvalues of the operator $(\Gamma \otimes I)\rho = (\mathbf{1}_A \otimes \mathrm{Tr}_A \rho) - \rho$ must all be non-negative if ρ is separable, which provides a simple test for separability called *reduction* criterion.[2] In the case where Γ is applied to a two-state system (quantum bit or spin-1/2 particle), as studied in section 4.3, this corresponds to the time-reversal operation applied on one system with respect to the other one. As Peres' criterion has been shown to be unitarily equivalent to such a "local" time-reversal by Sanpera et al. [52], this reduction criterion is simply equivalent to Peres' for $2 \times n$ composite systems. Therefore, it also results in a *sufficient* condition for 2×2 and 2×3 systems, according to Ref. [47]. It also has a very simple geometric representation in the Hilbert-Schmidt representation of the bipartite state. Finally, we demonstrate that the map Γ is connected to the complex conjugation operation in the "magic" basis for two qubits introduced recently by Hill and Wootters [2], which underlies an interesting connection with the entropy of formation [53]. In Appendix A, we illustrate the reduction separability condition by applying it to several separable or inseparable states, and compare it to the separability criterion based on partial transposition.

4.2 Separability of bipartite mixed states of arbitrary dimension

We consider a bipartite quantum system characterized by the density operator ρ_{AB} defined in the joint Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ and $\mathcal{H}_B$

[2] This reduction separability criterion has been independently derived by M. Horodecki and P. Horodecki in Ref. [51].

have arbitrary dimensions d_A and d_B.

Definition 1: Define a linear map Λ which maps Hermitian operators on $\mathcal{H}_{AB}$ into Hermitian operators on $\mathcal{H}_{AB}$:

$$\Lambda : \rho_{AB} \to \lambda_{AB} \equiv \mathbf{1}_A \otimes \rho_B - \rho_{AB} \qquad \text{with } \rho_B = \mathrm{Tr}_A[\rho_{AB}] \tag{4.4}$$

This map commutes with a unitary transformation acting locally on A and B. Indeed, if ρ_{AB} undergoes a unitary transformation of the product form, i. e.,

$$\rho_{AB} \to \rho'_{AB} = (U_A \otimes U_B)\rho_{AB}(U_A^\dagger \otimes U_B^\dagger), \tag{4.5}$$

it is easy to check that $\rho'_B = \mathrm{Tr}_A[\rho'_{AB}] = U_B \rho_B U_B^\dagger$, so that

$$\lambda_{AB} \to \lambda'_{AB} = (U_A \otimes U_B)\lambda_{AB}(U_A^\dagger \otimes U_B^\dagger), \tag{4.6}$$

i.e., λ_{AB} transforms just like ρ_{AB}. As a consequence, the spectrum of λ_{AB} is invariant under a $U_A \otimes U_B$ isomorphism on ρ_{AB}, as expected.

Theorem 1: A *necessary* condition for the separability of the state ρ_{AB} of a bipartite system AB is that it is mapped by Λ into a positive semi-definite operator, i.e., $\Lambda\rho_{AB} \geq 0$.

We need to prove that any separable state is mapped into a positive semi-definite operator λ_{AB}. Consider a *separable* bipartite system AB characterized by a convex combination of product states:

$$\rho_{AB} = \sum_i w_i \left(\rho_A^{(i)} \otimes \rho_B^{(i)}\right) \qquad \text{with } \sum_i w_i = 1 \text{ and } 0 \leq w_i \leq 1 \tag{4.7}$$

where $\rho_A^{(i)}$ and $\rho_B^{(i)}$ are states in $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. It is easy to verify that the operator $\lambda_{AB} = \Lambda\rho_{AB}$ is positive semi-definite,

$$\lambda_{AB} = \sum_i w_i \Big(\underbrace{(\mathbf{1}_A - \rho_A^{(i)})}_{\geq 0} \otimes \underbrace{\rho_B^{(i)}}_{\geq 0} \Big) \geq 0 \tag{4.8}$$

since a sum of positive operators is a positive operator. □

In short, the map Λ reveals non-separability:[3] if $\lambda_{AB} \not\geq 0$, then ρ_{AB} is inseparable. Moreover, it is easy to see that Λ conserves separability since it is linear and maps product states into product operators: if ρ_{AB} is separable, then $\lambda_{AB} \geq 0$ is also separable (or, in general, written as a convex sum of direct products). Let us now calculate the partial traces of λ_{AB}:

$$\lambda_A = \mathrm{Tr}_B[\lambda_{AB}] = \mathbf{1}_A - \rho_A \tag{4.9}$$

$$\lambda_B = \mathrm{Tr}_A[\lambda_{AB}] = (d_A - 1)\rho_B \tag{4.10}$$

where d_A is the dimension of $\mathcal{H}_A$. This shows that Λ does not preserve the trace in general. Indeed, the trace is scaled by an integer factor under Λ, that is, $\mathrm{Tr}[\lambda_{AB}] = (d_A - 1)\mathrm{Tr}[\rho_{AB}]$. Thus, Λ is *trace-preserving* only in the special case where A is a two-state system (i.e., $d_A = 2$). It is also interesting to note that Λ is always *reversible*, the inverse map being given by

$$\Lambda^{-1} : \lambda_{AB} \to (d_A - 1)^{-1}(\mathbf{1}_A \otimes \lambda_B) - \lambda_{AB} = \rho_{AB} \tag{4.11}$$

where λ_B is defined as above. Note that Λ is equal to its inverse Λ^{-1} only if $d_A = 2$. In that case, if λ_{AB} is separable, then $\Lambda^{-1} : \lambda_{AB} \to \rho_{AB} \geq 0$. (The fact that the inverse map reveals inseparability is true in this case only.)

The separability condition based on Λ is illustrated in Appendix A, where we consider several separable and inseparable states. As we will show in section 4.3, $\lambda_{AB} \geq 0$ results in the same condition as Peres' in the case of two quantum bits, in which case it is *sufficient* (see Theorem 4); for larger dimensions, it is only necessary.

Remark 1: Following the approach of Horodecki et al. [47], the map Λ can be written as the tensor product of a positive linear map Γ and the identity, that

[3] This necessary condition for the separability of mixed states is directly related to that based on the conditional amplitude operator (although it is simpler as it does not require the calculation of the latter operator) [50].

is

$$\Lambda = \Gamma \otimes I \qquad \text{with } \Gamma : \rho \to (\mathrm{Tr}\rho) - \rho \tag{4.12}$$

where Γ acts on Hermitian operators in $\mathcal{H}_A$ and the identity acts on operators in $\mathcal{H}_B$. Since Γ is a positive map, $\Lambda = \Gamma \otimes I$ maps *separable* states into *positive* operators [47]. It therefore results in a *necessary* condition for separability, according to Theorem 1. The map Γ commutes with an arbitrary unitary transformation U, that is

$$\Gamma(U\rho U^\dagger) = U(\Gamma\rho)U^\dagger \tag{4.13}$$

which makes the separability condition based on $\Lambda = \Gamma \otimes I$ independent on the basis chosen for A and B. In the same manner, the inverse map Λ^{-1} can be written as

$$\Lambda^{-1} = \Gamma^{-1} \otimes I \qquad \text{with } \Gamma^{-1} : \rho \to \frac{\mathrm{Tr}\rho}{d-1} - \rho \tag{4.14}$$

where d is the dimension of the Hilbert space of ρ. Note that Γ^{-1} is *not* a positive map for $d > 2$, so that Λ^{-1} is in general useless as far as detecting inseparability is concerned. This emphasizes that the reduction separability criterion is quite special in two-dimensions (e.g., for a spin-1/2 particle or a quantum bit), as will be studied in section 4.3. Specifically, we will show that Γ applied to a two-dimensional system can be interpreted as time reversal. Consequently, the map Λ amounts to applying time reversal on subsystem A, while leaving subsystem B unchanged. Such a link between "local" time-reversal and separability has recently been pointed out by Sanpera et al. [52].

Remark 2: It is interesting to consider the classical analog of the maps Γ and $\Lambda = \Gamma \otimes I$ to gain some insight into their physical meaning. First, applying Γ to a classical probability distribution p_i (diagonal ρ) corresponds to the transformation:

$$p_i \to q_i = \sum_k p_k - p_i \tag{4.15}$$

(Obviously, $q_j \geq 0$ is not normalized except for a binary distribution.) The classical analog of $\Lambda = \Gamma \otimes I$ is

$$p_{ij} \to q_{ij} = \left(\sum_k p_{k|j} - p_{i|j} \right) p_j = p_j - p_{ij}. \tag{4.16}$$

Since $p_{i|j}$ is a probability distribution in i, we always have $1 - p_{i|j} \geq 0$ so that $q_{ij} \geq 0$ and the separability criterion is fulfilled. This emphasizes that quantum inseparability ("$q_{ij} < 0$") may be viewed as resulting from a conditional probability that *exceeds* 1 (more precisely, an eigenvalue of $\rho_{A|B}$ which exceeds 1) [50].

Definition 2: Two additional maps from operators on $\mathcal{H}_{AB}$ to operators on $\mathcal{H}_{AB}$ can be defined: the *dual* map

$$\tilde{\Lambda} : \rho_{AB} \to \tilde{\lambda}_{AB} = \rho_A \otimes \mathbf{1}_B - \rho_{AB} \tag{4.17}$$

and the *symmetric* map

$$M : \rho_{AB} \to \mu_{AB} = \mathbf{1}_A \otimes \mathbf{1}_B - \rho_A \otimes \mathbf{1}_B - \mathbf{1}_A \otimes \rho_B + \rho_{AB} \tag{4.18}$$

where $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{Tr}_A[\rho_{AB}]$.

The map Λ which we considered until now is related to the conditional amplitude operator of A conditionally on B, that is $\rho_{A|B}$ [50]. Of course, a similar linear map can be defined using the amplitude operator $\rho_{B|A}$, and exactly the same conclusions follow. This is the *dual* map $\tilde{\Lambda}$ defined in equation (4.17). It is trace-preserving and self-inverse in the case where $d_B = 2$. It can obviously be written as the tensor product $\tilde{\Lambda} = I \otimes \Gamma$, where Γ now acts on operators on $\mathcal{H}_B$, and therefore commutes with a $U_A \otimes U_B$ isomorphism. Since Γ is positive, $\tilde{\Lambda}$ maps separable states into positive (separable) operators, which results in another separability condition, i.e., $\tilde{\lambda}_{AB} \geq 0$. As we will see in section 4.3, the operators λ_{AB} and $\tilde{\lambda}_{AB}$ can be shown to have the same spectrum when $d_A = d_B = 2$, in which case they result in the same separability condition.

However, this property does not hold in larger dimensions, i.e., λ_{AB} and $\tilde{\lambda}_{AB}$ do not have the same spectrum in general (see Appendix A).

We can also construct another linear map by cascading Λ and $\tilde{\Lambda}$ (the order is irrelevant), which results in the *symmetric* map $M = \tilde{\Lambda}\Lambda = \Gamma \otimes \Gamma$ defined in Eq. (4.18). Any separable ρ_{AB} is mapped by M into a separable operator $\mu_{AB} \geq 0$, as expected. The symmetric map also commutes with a $U_A \otimes U_B$ isomorphism,

$$M\left((U_A \otimes U_B)\rho_{AB}(U_A^\dagger \otimes U_B^\dagger)\right) = (U_A \otimes U_B)(M\rho_{AB})(U_A^\dagger \otimes U_B^\dagger), \qquad (4.19)$$

so that the spectrum of $\mu_{AB} = M\rho_{AB}$ is invariant under local transformations on ρ_{AB}. It is also reversible, its inverse map $M^{-1} = \Gamma^{-1} \otimes \Gamma^{-1}$ being given by

$$M^{-1} : \mu_{AB} \to \mathbf{1}_A \otimes \mathbf{1}_B - (d_B - 1)^{-1}(\mu_A \otimes \mathbf{1}_B) - (d_A - 1)^{-1}(\mathbf{1}_A \otimes \mu_B) + \mu_{AB} = \rho_{AB} \qquad (4.20)$$

where $\mu_A = \mathrm{Tr}_B[\mu_{AB}] = (d_B - 1)(\mathbf{1}_A - \rho_A)$ and $\mu_B = \mathrm{Tr}_A[\mu_{AB}] = (d_A - 1)(\mathbf{1}_B - \rho_B)$. As expected, this map is trace-preserving and self-inverse only in the case where $d_A = d_B = 2$. It corresponds then to a time-reversal operation applied to the *joint* system AB. In this case, M by itself is not useful as far as revealing inseparability is concerned since it is positive, i.e., $M\rho_{AB} \geq 0$. Therefore, all inseparable states of two quantum bits are mapped into positive operators just as are separable states. Still, M is important when analyzing the separability of two quantum bits as it is equivalent to the complex conjugation operation in the "magic" basis introduced by Hill and Wootters [2] (see Theorem 6). Whether the positivity of M holds in arbitrary dimensions is not known.

Theorem 2: The reduction separability criterion ($\Lambda\rho_{AB} \geq 0$) is *not* a sufficient condition for the separability of ρ_{AB}.

In order to prove that this criterion is not sufficient, we show that it is possible to find an *inseparable* system with $\lambda_{AB} \geq 0$, i.e., such that its inseparability is *not* revealed by Λ. We will construct such an inseparable system by extending an inseparable component with a separable one, "diluting" the inseparability [50]. Consider an inseparable system $A'B'$ with $\lambda_{A'B'} \not\geq 0$. Let us extend $A'B'$ with a separable system $A''B''$, and apply the reduction criterion to the joint system AB where $A \equiv A'A''$ and $B \equiv B'B''$. Since the joint system is characterized by $\rho_{AB} = \rho_{A'B'} \otimes \rho_{A''B''}$, its associated operator under the map Λ is given by

$$\lambda_{AB} = \Lambda\rho_{AB} = (\mathbf{1}_{A'} \otimes \rho_{B'}) \otimes (\mathbf{1}_{A''} \otimes \rho_{B''}) - \rho_{A'B'} \otimes \rho_{A''B''} \tag{4.21}$$

Using the operators $\lambda_{A'B'} = \Lambda\rho_{A'B'} = \mathbf{1}_{A'} \otimes \rho_{B'} - \rho_{A'B'}$ and $\lambda_{A''B''} = \Lambda\rho_{A''B''} = \mathbf{1}_{A''} \otimes \rho_{B''} - \rho_{A''B''}$ corresponding to Λ applied to each component system, we obtain

$$\lambda_{AB} = \lambda_{A'B'} \otimes \lambda_{A''B''} + \lambda_{A'B'} \otimes \rho_{A''B''} + \rho_{A'B'} \otimes \lambda_{A''B''} \tag{4.22}$$

with $\lambda_{A'B'} \not\geq 0$ and $\lambda_{A''B''} \geq 0$ (since $A''B''$ is separable). The *dilution* of entanglement comes from the fact that the third term on the right-hand side of equation (4.22) is ≥ 0. As a consequence, equation (4.22) *cannot* guarantee that $\lambda_{AB} \not\geq 0$ even though the composite system AB contains an inseparable component as $\lambda_{A'B'} \not\geq 0$ (i.e., even though the sum of the first two terms on the right-hand side of equation (4.22) is $\not\geq 0$). □

Note that, even when both components are inseparable with $\lambda_{A'B'}, \lambda_{A''B''} \not\geq 0$, then $\lambda_{AB} \not\geq 0$ is not necessarily true, so that the inseparability of the joint system AB is not always revealed by Λ.[4] Conversely, equation (4.22) implies that, if both components have $\lambda_{A'B'} \geq 0$ and $\lambda_{A''B''} \geq 0$, then $\lambda_{AB} \geq 0$. It is not difficult to find examples of such inseparable states AB whose inseparability

[4] This property contrasts with the situation prevailing when using the conditional amplitude matrix. If the conditional amplitude operator of each component admits an eigenvalue > 1, then so does the corresponding operator for the whole system.[50]

is masked (i.e., $\lambda_{AB} \geq 0$) by extending an inseparable component $A'B'$ that satisfies $\lambda_{A'B'} \not\geq 0$ with a separable one $A''B''$. For example, let $A'B'$ be one of the Bell states, e.g., $\rho_{A'B'} = |\Phi^+\rangle\langle\Phi^+|$ with $|\Phi^+\rangle = 2^{-1/2}(|00\rangle + |11\rangle)$, and let $A''B''$ be a product of two random quantum bits, i.e., $\rho_{A''B''} = (\mathbf{1}_{A''} \otimes \mathbf{1}_{B''})/4$. Since $\rho_{B'} = \mathbf{1}_{B'}/2$, we have $\lambda_{A'B'} = \mathbf{1}_{A'B'}/2 - \rho_{A'B'} \not\geq 0$, as expected. Using $\lambda_{A''B''} = \rho_{A''B''}$, we see that equation (4.22) yields

$$\lambda_{AB} = (\mathbf{1}_{A'B'} - \rho_{A'B'}) \otimes \rho_{A''B''} \tag{4.23}$$

which is obviously a non-negative operator, so that the inseparability of AB is hidden. The example of weakly inseparable states with a positive partial transpose (see Ref. [48]) is treated in Appendix A, to illustrate that $\lambda_{AB} \geq 0$ is not a sufficient condition in general.

Remark 1: The mechanism of dilution of inseparability can be understood by examining the action of the map Γ on product states. Indeed, when applying $\Lambda = \Gamma \otimes I$ on the state $\rho_{AB} = \rho_{A'B'} \otimes \rho_{A''B''}$, Γ acts on the state $\rho_{A'} \otimes \rho_{A''}$ (B and B' are left unchanged by I). Let us consider a density operator of the product form $\rho = \rho' \otimes \rho''$. Since we have $\mathrm{Tr}(\rho) = \mathrm{Tr}(\rho')\mathrm{Tr}(\rho'')$, we see that it is mapped to

$$\begin{aligned}
\Gamma(\rho' \otimes \rho'') &= \mathrm{Tr}(\rho')\mathrm{Tr}(\rho'') - \rho' \otimes \rho'' \\
&= [\mathrm{Tr}(\rho') - \rho'] \otimes [\mathrm{Tr}(\rho'') - \rho''] + \mathrm{Tr}(\rho') \otimes \rho'' + \rho' \otimes \mathrm{Tr}(\rho'') - 2\rho' \otimes \rho'' \\
&= \Gamma\rho' \otimes \Gamma\rho'' + \Gamma\rho' \otimes \rho'' + \rho' \otimes \Gamma\rho'' \qquad (4.24)
\end{aligned}$$

which implies the relation

$$\Gamma = \Gamma' \otimes \Gamma'' + \Gamma' \otimes I'' + I' \otimes \Gamma'' \tag{4.25}$$

where Γ' (or Γ'') stands for the same map but acting on the subspace of ρ' (or ρ'') while Γ acts on the joint space. Using the same notation for Λ (i.e., Λ' acts on the subspace of $A'B'$ while Λ'' acts on the subspace of $A''B''$), the latter equation gives

$$\Lambda = \Gamma \otimes I = \Lambda' \otimes \Lambda'' + \Lambda' \otimes I'' + I' \otimes \Lambda'' \tag{4.26}$$

which implies equation (4.22). The same reasoning can be applied to the dual map $\tilde{\Lambda} = I \otimes \Gamma$ and to the symmetric map $M = \Gamma \otimes \Gamma$. Thus, even if the maps Λ' and Λ'' reveal inseparability by themselves, the combined map, equation (4.26), is not guaranteed to do so because the non-positivity of $(\Lambda' \otimes \Lambda'')\rho = (\Lambda'\rho') \otimes (\Lambda''\rho'')$ can be masked by one of the last two terms (the one where Λ is applied to the separable component).

Remark 2: It is worth noting that the separability criterion based on the partial transposition [46] does *not* suffer from this dilution of inseparability (even though it is not a sufficient condition in general). Consider, as before, a system AB characterized by $\rho_{AB} = \rho_{A'B'} \otimes \rho_{A''B''}$, where the *inseparable* component $A'B'$ is detected by partial transposition, i.e., $(\rho_{A'B'})^{T_{A'}} \not\geq 0$. Since $(\rho_{AB})^{T_A} = (\rho_{A'B'})^{T_{A'}} \otimes (\rho_{A''B''})^{T_{A''}}$, we have $\mathrm{Tr}_{A''B''}[(\rho_{AB})^{T_A}] = (\rho_{A'B'})^{T_{A'}} \not\geq 0$. Since the partial trace of a non-negative operator is a non-negative operator, this implies that $(\rho_{AB})^{T_A} \not\geq 0$, so that the inseparability of the extended system AB is detected provided that the inseparability of a component of it (here $A'B'$) is detected.

4.3 Separability of two two-dimensional systems

Theorem 3: The map Γ acting on a two-dimensional system corresponds to time-reversal, and is therefore equivalent to applying the complex conjugation operator K followed by a rotation $\mathcal{R}_y$ by an angle π about the y-axis, that is, $\Gamma = \mathcal{R}_y K$.

Let us write the arbitrary state of a two-dimensional quantum system (a quantum bit) in the Bloch-sphere picture:

$$\rho = \frac{1}{2}(1 + \vec{r} \cdot \vec{\sigma}) \tag{4.27}$$

where $\vec{\sigma}$ represent the three Pauli matrices and $\vec{r} = \mathrm{Tr}(\rho\vec{\sigma})$ is a *real* vector in the Bloch sphere (of radius 1). The vector $\vec{r}$ describes the statistics of measurements on the system, as, for example, the quantum expectation value of the spin component along an axis defined by the vector $\vec{v}$ is $\mathrm{Tr}\Big[\rho(\vec{v} \cdot \vec{\sigma})\Big] = (\vec{v}, \vec{r})$. Using equation (4.27), it is straightforward to check that

$$\Gamma\rho = 1 - \rho = \frac{1}{2}(1 - \vec{r} \cdot \vec{\sigma}). \tag{4.28}$$

Thus, Γ performs a *spin-flip*, or, equivalently, performs a *parity* transformation on the Bloch vector $\vec{r} \to -\vec{r}$. This can be viewed as *time-reversal*, and therefore can be decomposed into a complex conjugation K followed by a rotation $\mathcal{R}_y$ of an angle π about the y-axis, that is $\Gamma = \mathcal{T} = \mathcal{R}_y K$ [54]. □

Remark 1: In order to see this explicitly, consider the action of the map $\Lambda = \Gamma \otimes I$ on a product state $|\psi\rangle = |a\rangle \otimes |b\rangle$. Using $\rho_{AB} = P_a \otimes P_b$ with $P_a = |a\rangle\langle a|$ and $P_b = |b\rangle\langle b|$, we have

$$\lambda_{AB} = P_a^{\perp} \otimes P_b \tag{4.29}$$

where $P_a^{\perp} = \Gamma(|a\rangle\langle a|) = \mathbf{1}_A - |a\rangle\langle a|$ is the projector on the subspace orthogonal to $|a\rangle$. In the case where $d_A = 2$, $P_a^{\perp}$ is a rank-one projector as the total trace is preserved. Then, $P_a^{\perp} = |a^{\perp}\rangle\langle a^{\perp}|$, where $|a^{\perp}\rangle$ is a state vector orthogonal to $|a\rangle$.[5] It is easy to check that $|a^{\perp}\rangle$ can be obtained by applying a complex conjugation K on the components of $|a\rangle$ followed by a rotation $\mathcal{R}_y$ of angle π about the y-axis. Indeed, any state $|a\rangle = \alpha|0\rangle + \beta|1\rangle$ (with $|\alpha|^2 + |\beta|^2 = 1$) is

[5] Note that it is impossible to construct a state $|a^{\perp}\rangle$ that is orthogonal to an *arbitrary* state $|a\rangle$ by applying a *unitary* transformation alone.

transformed into $|a^\perp\rangle = -\beta^*|0\rangle + \alpha^*|1\rangle$ by applying the rotation

$$U_y = \exp(-i\pi\sigma_y/2) = -i\sigma_y = \sigma_x\sigma_z = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{4.30}$$

(that is, a bit- and phase-flip) to the state vector $\alpha^*|0\rangle+\beta^*|1\rangle$. The transformed state $|a^\perp\rangle$ is such that $\langle a^\perp|a\rangle = 0$ and $|a^\perp\rangle\langle a^\perp| = \mathbf{1}_A - |a\rangle\langle a|$, as expected. Thus, Γ coincides with time-reversal for a spin-1/2 system ($d_A = 2$) as the latter is equal to complex conjugation K followed by the rotation $\mathcal{R}_y$, i.e., $\mathcal{T} = \mathcal{R}_y K$ [54]. Consequently, Γ is an *antiunitary*[6] operation on state vectors in a two-dimensional Hilbert space (see Appendix B).

Corollary: For the Hilbert-Schmidt decomposition of ρ_{AB}, the map $\Lambda = \Gamma \otimes I$ corresponds to a sign-flip of the Pauli matrices acting on A while leaving those acting on B unchanged.

Let us consider the Hilbert-Schmidt decomposition of an arbitrary state of two quantum bits (or spin-1/2 particles) [55]:

$$\rho_{AB} = \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B + \vec{r}\cdot\vec{\sigma}_A \otimes \mathbf{1}_B + \mathbf{1}_A \otimes \vec{s}\cdot\vec{\sigma}_B + \sum_{m,n=1}^{3} t_{n,m}\, \sigma_A^{(n)} \otimes \sigma_B^{(m)}\right) \tag{4.31}$$

where $\sigma_A^{(n)}$ and $\sigma_B^{(m)}$ stand for the Pauli matrices (with $n = 1, 2, 3$) in the A and B space, respectively. equation (4.31) depends on 15 real parameters, the two three-dimensional vectors $\vec{r}$ and $\vec{s}$, and the 3×3 real matrix $t_{n,m}$. The vectors $\vec{r}$ and $\vec{s}$ correspond to the state of A and B in the Bloch sphere since we have

$$\rho_A = \mathrm{Tr}_B[\rho_{AB}] = \frac{1}{2}(\mathbf{1}_A + \vec{r}\cdot\vec{\sigma}_A) \tag{4.32}$$

$$\rho_B = \mathrm{Tr}_A[\rho_{AB}] = \frac{1}{2}(\mathbf{1}_B + \vec{s}\cdot\vec{\sigma}_B). \tag{4.33}$$

They characterize the reduced systems A and B, that is the local (marginal) statistics of any observable on A or B. The matrix $t_{n,m} = \mathrm{Tr}[\rho_{AB}(\sigma_A^{(n)} \otimes$

[6] For any two state vectors $|a\rangle$ and $|\tilde{a}\rangle$, we have $\langle\tilde{a}^\perp|a^\perp\rangle = \langle\tilde{a}|a\rangle^*$.

$\sigma_B^{(m)})]$ describes the joint statistics of A and B as it characterizes the correlation between the measured spin components along two axes (defined by the vectors $\vec{a}$ and $\vec{b}$): $\mathrm{Tr}\left[\rho(\vec{a}\cdot\vec{\sigma}_A\otimes\vec{b}\cdot\vec{\sigma}_B)\right] = (\vec{a}, t\vec{b})$. Using equations (4.31) and (4.33), it is checked by straightforward calculation that Λ simply flips the sign of the terms in $\vec{\sigma}_A$:

$$\lambda_{AB} = \frac{1}{4}\left(\mathbf{1}_A\otimes\mathbf{1}_B - \vec{r}\cdot\vec{\sigma}_A\otimes\mathbf{1}_B + \mathbf{1}_A\otimes\vec{s}\cdot\vec{\sigma}_B - \sum_{m,n=1}^{3} t_{n,m}\,\sigma_A^{(n)}\otimes\sigma_B^{(m)}\right) \tag{4.34}$$

This implies that $\Lambda = \Gamma\otimes I$ applied to a $2\times n$ system corresponds simply to "local" time-reversal $\mathcal{T}\otimes I$, that is, performing time-reversal on the subsystem A while leaving the subsystem B unchanged [52].

Remark 2: The dual map $\tilde{\Lambda} = I\otimes\Gamma$ flips the sign of the Pauli matrices acting on B while leaving the sign of those acting on A unchanged. The action of the symmetric map $M = \Gamma\otimes\Gamma$ on the Hilbert-Schmidt decomposition of ρ_{AB} is to flip the sign of the Pauli matrices $\vec{\sigma}_A$ and $\vec{\sigma}_B$. This operation corresponds therefore to time-reversal applied to A and B simultaneously, and is equivalent to complex conjugation in the "magic" basis (see Theorem 6). It is worth noting that the set of states that remain invariant under the symmetric map M are those with $\vec{r} = \vec{s} = 0$, that is, mixtures of generalized Bell states (the latter being defined as the states obtained by applying any local transformation to the four Bell states). These states are called "T-states" by Horodecki et al. [55], and are such that the entropy of A and B is maximal, that is $S(\rho_A) = S(\rho_B) = 1$. (The only pure states in this set are the fully entangled states of two qubits, i.e., the generalized Bell states.) Thus, in particular, the (generalized) Bell states are left unchanged by the action of M. In contrast, a (separable) product state $\rho_A\otimes\rho_B$ is mapped into the distinct (non-negative) state $\mu_{AB} = (\mathbf{1}_A - \rho_A)\otimes(\mathbf{1}_B - \rho_B)$. Because of this property, μ_{AB} by itself is uninteresting as far as revealing inseparability is concerned, as mentioned earlier.

Theorem 4: A bipartite system of two-dimensional components A and B characterized by an arbitrary joint density operator ρ_{AB} is separable *if and only if* the operator $\lambda_{AB} = \Lambda\rho_{AB}$ is positive semi-definite.

It is enough to show that Λ is equivalent to a partial transposition up to a completely positive map (in fact, a unitary transformation), since Peres' separability criterion is known to be *necessary* and *sufficient* in this case [47]. Since we are dealing with Hermitian operators, the map $T \otimes I$, where T is the standard transposition of operators on $\mathcal{H}_A$, is equivalent to the "partial conjugation"[7] operation $K \otimes I$, where K is the complex conjugation operator acting on states on $\mathcal{H}_A$. Thus, Theorem 3 reads $\Gamma = \mathcal{R}_y T$. We can now use the fact that any positive map Π acting on density operators in a two-dimensional Hilbert space can be written as [47]

$$\Pi = \Pi_1^{\mathrm{CP}} + \Pi_2^{\mathrm{CP}} T \tag{4.35}$$

where Π_1^{CP} and Π_2^{CP} are completely positive maps (which therefore do not reveal inseparability). With the identification $\Pi_1^{\mathrm{CP}} = 0$ and $\Pi_2^{\mathrm{CP}} = \mathcal{R}_y$, we see that the map Γ can be used rather than the transposition operator T (or K) in order to test the positivity of the operator resulting from applying *any* element of the set of maps $\Pi \otimes I$ on ρ_{AB} (this follows from the reasoning used in Ref. [47]). Thus, using the fact that the complex conjugation operator K is *unitarily* equivalent to Γ, we have shown that $\Lambda\rho_{AB} \geq 0$ results in a necessary *and* sufficient condition for the separability of ρ_{AB}. $\Box$

Remark 1: The map Γ applied to a two-dimensional system is unitarily equivalent to the transposition operator T. Since the spectrum of an operator is conserved by a unitary transformation ($\mathcal{R}_y$), the spectrum of the operator obtained by partial transposition in subspace A, $(T \otimes I)\rho_{AB} = \rho_{AB}^{T_A}$, is the same as

[7] Note that, although K is well-defined, partial conjugation $K \otimes I$ is *only* defined for product state vectors in $\mathcal{H}_{AB}$ [48].

the spectrum of $\lambda_{AB} = \Lambda\rho_{AB}$. Therefore, testing Peres' separability condition or the positivity of λ_{AB} is operationally equivalent, and these conditions can be used interchangeably in the case of two quantum bits, as illustrated in Appendix A. Moreover, λ_{AB} and $\rho_{AB}^{T_A}$ have the same spectrum for $2 \times n$ systems, so that the conditions are also equivalent if Γ is applied on the two-dimensional subsystem. As a consequence, the separability condition based on Λ is necessary *and* sufficient for 2×3 systems, while it is only necessary for $2 \times n$ systems with larger n, just as Peres' condition [47]. Numerical evidence suggests that, for systems with $d_A, d_B > 2$, the reduction condition is weaker than (or equivalent to) the one based on partial transposition.[8]

Remark 2: It is instructive to illustrate Theorem 4 for "T-states" [55], that is, in the case where A and B have a maximal reduced entropy. The T-states ($\vec{r} = \vec{s} = 0$) are such that the reduced density operators are given by $\rho_A = \rho_B = \mathbf{1}/2$, so that the reduced entropies are $S(\rho_A) = S(\rho_B) = 1$. These states are thus completely characterized by the matrix $t_{n,m}$. It has been shown in Ref. [55] that any T-state can be transformed by a unitary transformation of the product form $U_A \otimes U_B$ into a state for which $t_{n,m}$ is *diagonal*. As far as separability is concerned, we can thus restrict ourselves to the class of all states with diagonal t, since these are representative of the entire set of T-states (up to an $U_A \otimes U_B$ isomorphism).

The class of states with diagonal t is a convex subset of the set of T-states, and any state belonging to this subset can be characterized by the real vector $\vec{t} = (t_{11}, t_{22}, t_{33})$ made out of the diagonal elements of t. It was proven in Ref. [55] that an operator ρ_{AB} of the form given by equation (4.31) with $\vec{r} = \vec{s} = 0$ and diagonal t corresponds to a state (i.e., a *positive* unit-trace operator) if and only if the vector $\vec{t}$ belongs to a tetrahedron with vertices $\vec{t_1} = (-1, 1, 1)$,

[8] This has been later proven in Ref. [51].

$\vec{t}_2 = (1,-1,1)$, $\vec{t}_3 = (1,1,-1)$, and $\vec{t}_4 = (-1,-1,-1)$. In other words, any state of this class can be represented by a point inside this tetrahedron. In this representation, the four Bell states $|\Phi^{\pm}\rangle = 2^{-1/2}(|00\rangle \pm |11\rangle)$ and $|\Psi^{\pm}\rangle = 2^{-1/2}(|01\rangle \pm |10\rangle)$ correspond to the vertices of the tetrahedron, that is

$$\begin{aligned}
\vec{t}_1 \quad &: \quad |\Phi^-\rangle\langle\Phi^-| = \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B - \sigma_A^{(x)} \otimes \sigma_B^{(x)} + \sigma_A^{(y)} \otimes \sigma_B^{(y)} + \sigma_A^{(z)} \otimes \sigma_B^{(z)}\right) \\
\vec{t}_2 \quad &: \quad |\Phi^+\rangle\langle\Phi^+| = \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B + \sigma_A^{(x)} \otimes \sigma_B^{(x)} - \sigma_A^{(y)} \otimes \sigma_B^{(y)} + \sigma_A^{(z)} \otimes \sigma_B^{(z)}\right) \\
\vec{t}_3 \quad &: \quad |\Psi^+\rangle\langle\Psi^+| = \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B + \sigma_A^{(x)} \otimes \sigma_B^{(x)} + \sigma_A^{(y)} \otimes \sigma_B^{(y)} - \sigma_A^{(z)} \otimes \sigma_B^{(z)}\right) \\
\vec{t}_4 \quad &: \quad |\Psi^-\rangle\langle\Psi^-| = \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B - \sigma_A^{(x)} \otimes \sigma_B^{(x)} - \sigma_A^{(y)} \otimes \sigma_B^{(y)} - \sigma_A^{(z)} \otimes \sigma_B^{(z)}\right)
\end{aligned}$$

In Ref. [55], it is also shown that a state ρ_{AB} of this T-diagonal class is *separable* if and only if the vector $\vec{t}$ characterizing ρ_{AB} belongs to an octahedron with vertices $\vec{o}_1^{\,\pm} = (\pm 1,0,0)$, $\vec{o}_2^{\,\pm} = (0,\pm 1,0)$, and $\vec{o}_3^{\,\pm} = (0,0,\pm 1)$. Let us consider the action of Λ in this representation. As shown earlier, Λ flips the "spin" $\vec{\sigma}_A$. Within the set of T-states, this amounts to changing the sign of the $t_{n,m}$ matrix, that is, to flipping the sign of the vector $\vec{t}$ for T-diagonal states. Therefore, the criterion for separability $\lambda_{AB} = \Lambda\rho_{AB} \geq 0$ translates, in this representation, to the condition that the "parity" operation on the vector $\vec{t}$ characterizing a separable state results in a positive operator (i.e., a legitimate state). Thus, $-\vec{t}$ must belong to the tetrahedron. It is easy to see that the set of points of the tetrahedron which are such that their image under parity still belongs to the tetrahedron corresponds exactly to the octahedron defined above. Therefore, no inseparable state exists that satisfies $\Lambda\rho_{AB} \geq 0$, so that Λ provides a necessary *and* sufficient condition for separability within the class of T-states, as expected.

Theorem 5: The symmetric map M acting on two two-dimensional systems conserves the spectrum, so that the separability criteria resulting from the map Λ and its dual $\tilde{\Lambda}$ are equivalent.

As a consequence of Theorem 3, $M = \Gamma \otimes \Gamma$ amounts to performing a complex conjugation K (or transposition) of the *joint* density operator in $\mathcal{H}_{AB}$, followed by a tensor product of the rotation $\mathcal{R}_y$ defined by $U_y = \exp(-i\pi\sigma_y/2) = -i\sigma_y$, that is, $U_y \otimes U_y = -\sigma_y \otimes \sigma_y$. Note that, as we are dealing with Hermitian (density) operators, their spectrum is unchanged by K. The same is true for the rotation $U_y \otimes U_y$. Therefore, $\mu_{AB} = M\rho_{AB}$ has the same spectrum as ρ_{AB} when $d_A = d_B = 2$. As Γ is self-inverse ($\Gamma^2 = I$) when $d_A = d_B = 2$, we have the relation $I \otimes \Gamma = (\Gamma \otimes I)(\Gamma \otimes \Gamma)$ or in short $\tilde{\Lambda} = \Lambda M$. This implies that

$$\tilde{\Lambda}\rho_{AB} = \Lambda \left[(U_y \otimes U_y)\rho^*_{AB}(U_y^\dagger \otimes U_y^\dagger)\right] \tag{4.36}$$

which in turn results in

$$\tilde{\lambda}_{AB} = (U_y \otimes U_y)\lambda^*_{AB}(U_y^\dagger \otimes U_y^\dagger) \tag{4.37}$$

as Λ commutes with $U_y \otimes U_y$ and complex conjugation. Since λ_{AB} is Hermitian (just as ρ_{AB}), the latter expression shows that the spectrum of $\tilde{\lambda}_{AB}$ and λ_{AB} are identical, so that the resulting criteria for separability are equivalent. □

Theorem 6: The symmetric map M applied to a bipartite system of two-dimensional components (i.e., global time-reversal) is equivalent to complex conjugation in the "magic" basis introduced in Ref. [53].[9]

Since $\Gamma = \mathcal{R}_y K$, the symmetric map $M = \Gamma \otimes \Gamma$ applied to the state ρ_{AB} of a bipartite system results in

$$M\rho_{AB} = (U_y \otimes U_y)\rho^*_{AB}(U_y^\dagger \otimes U_y^\dagger) \tag{4.38}$$

where $U_y \otimes U_y = -\sigma_y \otimes \sigma_y$. Since M is antiunitary and self-inverse ($M^2 = I$), it is a *conjugation* [57]. It can be written as the complex conjugation operator if expressed in a specific basis. Let us assume that V is the unitary operator (in

[9] This was pointed out independently in Ref. [56], which was brought to our attention after completion of this work.

the joint space) that transforms the product states into the states $\{|e_i\rangle\}$ that form this specific basis, that is

$$|e_1\rangle = V|00\rangle \qquad |e_2\rangle = V|01\rangle \qquad |e_3\rangle = V|10\rangle \qquad |e_4\rangle = V|11\rangle. \tag{4.39}$$

We would like to show that M is equivalent to rotating the states $|e_i\rangle$ into the product states, taking the complex conjugation of the density matrix (in the product basis), and then rotating the product states back to the $|e_i\rangle$'s:

$$M\rho_{AB} = V(V^\dagger \rho_{AB} V)^* V^\dagger = (VV^T)\rho^*_{AB}(VV^T)^\dagger \tag{4.40}$$

where V^T is the transpose of the unitary matrix V. Identifying equations (4.38) and (4.40), we obtain

$$VV^T = U_y \otimes U_y = -\sigma_y \otimes \sigma_y = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{4.41}$$

It is easy to prove that, if V is unitary, then VV^T is unitary and symmetric (but not necessarily Hermitian). In order to find a solution for V that satisfies equation (4.41), we first diagonalize the matrix $\sigma_y \otimes \sigma_y$. Consider the unitary matrix

$$W \equiv \exp\left(-\frac{i\pi}{4}(1-\sigma_x)\otimes(1-\sigma_x)\right) = (1\otimes 1 + 1\otimes\sigma_x + \sigma_x\otimes 1 - \sigma_x\otimes\sigma_x)/2 \tag{4.42}$$

It is in fact a *real* orthogonal matrix, so that $W^{-1} = W^\dagger = W^T$. It can easily be shown that W diagonalizes[10] $\sigma_y \otimes \sigma_y$, that is,

$$W(\sigma_y \otimes \sigma_y)W^T = \sigma_z \otimes \sigma_z \tag{4.43}$$

[10] It is not the only such matrix, as $\sigma_y \otimes \sigma_y$ is obviously also diagonalized by $\exp(-i\frac{\pi}{4}\sigma_x) \otimes \exp(-i\frac{\pi}{4}\sigma_x)$. However, we are looking here for a (real) rotation matrix rather than a general unitary matrix.

Note that the matrix W is self-inverse, i.e., $W^2 = 1$, so that it is also symmetric ($W^T = W$). By multiplying equation (4.41) by W on the left and the right, we obtain

$$WV(WV)^T = -\sigma_z \otimes \sigma_z = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{4.44}$$

which implies that the product WV can be written as a diagonal matrix D:

$$WV = D \equiv \begin{pmatrix} \pm i & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & \pm i \end{pmatrix} \tag{4.45}$$

This yields a (non-unique) solution for the unitary matrix $V = W^T D = WD$ that defines the basis $\{|e_i\rangle\}$. The states $|e_i\rangle$ are thus obtained by applying the rotation matrix W to the product states $\pm i|00\rangle$, $\pm|01\rangle$, $\pm|10\rangle$, and $\pm i|11\rangle$. It is worth noticing at this point that the rotation matrix

$$W = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \tag{4.46}$$

transforms the product states into the four maximally entangled states which are obtained by applying a local transformation $H \otimes 1$ on the four Bell states, i.e.,

$$\begin{aligned} W|00\rangle &= (H\otimes 1)|\Phi^+\rangle = (|00\rangle + |01\rangle + |10\rangle - |11\rangle)/2 \\ W|01\rangle &= (H\otimes 1)|\Psi^+\rangle = (|00\rangle + |01\rangle - |10\rangle + |11\rangle)/2 \\ W|10\rangle &= (H\otimes 1)|\Phi^-\rangle = (|00\rangle - |01\rangle + |10\rangle + |11\rangle)/2 \\ W|11\rangle &= (H\otimes 1)|\Psi^-\rangle = (-|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2 \end{aligned} \tag{4.47}$$

where H is the Hadamard transform. (As a matter of fact, the unitary transformation W corresponds simply to a controlled-NOT gate where the control is in the dual basis $\{|0\rangle+|1\rangle, |0\rangle-|1\rangle\}$ rather than the standard basis.) Therefore, the unitary transformation $V = WD$ is such that the product states are rotated into the four generalized Bell states with the appropriate phases

$$\begin{aligned}
|e_1\rangle = V|00\rangle &= \pm i(H\otimes 1)|\Phi^+\rangle \\
|e_2\rangle = V|01\rangle &= \pm 1(H\otimes 1)|\Psi^+\rangle \\
|e_3\rangle = V|10\rangle &= \pm 1(H\otimes 1)|\Phi^-\rangle \\
|e_4\rangle = V|11\rangle &= \pm i(H\otimes 1)|\Psi^-\rangle .
\end{aligned} \tag{4.48}$$

These states $|e_i\rangle$ are therefore equivalent, up to a local change of basis $H\otimes 1$ and a phase i that are irrelevant here, to the "magic" states introduced in Ref. [53]. (Any four states obtained from the $|e_i\rangle$'s up to an overall phase and a unitary transformation acting locally on each quantum bit are legitimate "magic" states.) This implies that, when expressed in this basis, the symmetric map $M = \Gamma\otimes\Gamma$ reduces the the complex conjugation operation that was used in the context of the calculation of the entropy of formation of a pair of quantum bits (see Refs. [2, 56]). □

Theorem 7: A distinct necessary separability condition for the bipartite state ρ_{AB} is that its support can be spanned by a set of product states which are such that the corresponding product operators obtained by applying Γ to the state vector in $\mathcal{H}_A$ span the support of $\lambda_{AB} = \Lambda\rho_{AB}$.

We only consider this condition in the case where $d_A = 2$. Let us first show that if ρ_{AB} is a separable state, then λ_{AB} is a separable operator obtained by replacing the states $|a\rangle$ in $\mathcal{H}_A$ by projectors $P_a^\perp$ orthogonal to them. Consider the separable state

$$\rho_{AB} = \sum_i w_i \left(|a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i| \right) \tag{4.49}$$

where the $|a_i\rangle \otimes |b_i\rangle$ are pure product states [using the spectral decomposition of $\rho_A^{(i)}$ and $\rho_B^{(i)}$, it is easy to rewrite equation (4.7) into this form]. As a result of Theorem 3, we see that ρ_{AB} is mapped by Λ into the separable operator

$$\lambda_{AB} = \sum_i w_i \left(P_{a_i}^{\perp} \otimes |b_i\rangle\langle b_i| \right) . \tag{4.50}$$

The operator λ_{AB} is a unit-trace operator in the case $d_A = 2$ since each component pure state $|a\rangle \otimes |b\rangle$ is mapped into a *pure* product state, $|a^{\perp}\rangle \otimes |b\rangle$, in which case it simply reads

$$\lambda_{AB} = \sum_i w_i \left(|a_i^{\perp}\rangle\langle a_i^{\perp}| \otimes |b_i\rangle\langle b_i| \right) . \tag{4.51}$$

Let us show that equation (4.51) results in a simple *necessary* condition for separability (distinct from $\lambda_{AB} \geq 0$), inspired from the condition recently proposed by Horodecki [48]. The central point is to note that, if ρ_{AB} is separable, then the ensemble of product states $|a_i\rangle \otimes |b_i\rangle$ span the entire support of ρ_{AB}. (Conversely, any state $|a_i\rangle \otimes |b_i\rangle$ must belong to the support of ρ_{AB} and cannot have a non-vanishing component orthogonal to it.) From equation (4.51), we see that the ensemble of states $|a_i^{\perp}\rangle \otimes |b_i\rangle$ span the entire support of the corresponding separable state λ_{AB} obtained by applying Λ on ρ_{AB} [cf. equation (4.51)]. (Also, any state $|a_i^{\perp}\rangle \otimes |b_i\rangle$ cannot be outside the support of λ_{AB}.) This results in a *necessary* condition for separability which can be stated as follows: if a state ρ_{AB} is separable, then it must be possible to span its support by a set of product states $|a\rangle|b\rangle$ which are such that their image (i.e., the product states obtained by rotating the complex conjugate of state vector $|a\rangle$ in the A space by an angle π about the y-axis while leaving the state vector $|b\rangle$ in the B space unchanged) span the support of the mapped state $\lambda_{AB} = \Lambda\rho_{AB}$. □

4.4 Conclusion

Given a bipartite system characterized by a density operator ρ_{AB}, we construct a simple separability criterion based on the positive linear map $\Gamma : \rho \rightarrow (\mathrm{Tr}\rho) - \rho$. Any separable state ρ_{AB} is mapped by the tensor product of Γ (acting on A) and the identity I (acting on B) into a positive operator. Therefore, a necessary condition for separability is based on checking the non-negativity of the operator $(\Gamma \otimes I)\rho_{AB} = \mathbf{1}_A \otimes \rho_B - \rho_{AB}$. This condition, along with the one based on the dual map $I \otimes \Gamma$, can be shown to be non-sufficient for a system of arbitrary dimension because entanglement dilution can thwart the map's sensitivity. Since Γ commutes with any unitary transformation, the spectrum of the operator $(\Gamma \otimes I)\rho_{AB}$ is invariant under a local unitary transformation $U_A \otimes U_B$, making this reduction criterion independent of the basis in which A and B are expressed.

In the case of a two-dimensional system, Γ is shown to be the time-reversal operator, which flips the sign of the spin matrices (or, equivalently, reverses the Bloch vector characterizing the state of the quantum bit), so that the map $\Gamma \otimes I$ amounts to changing the arrow of time for subsystem A with respect to subsystem B. Such a relation between time-reversal and Peres' partial transposition has been pointed out previously by Sanpera et al. [52], who showed that the partial transposition operator is unitarily equivalent to "local" time-reversal. Thus, the reduction criterion for separability based on $\Gamma \otimes I$ is equivalent to Peres' criterion [46] for $2 \times n$ systems (when applying Γ on the two-dimensional subsystem). As a consequence, it is necessary *and* sufficient for 2×2 and 2×3 systems while it is only necessary for larger systems, just as is Peres' [47]. For systems with $d_A, d_B > 2$, however, the reduction condition is generally weaker than the one based on partial transposition.

Finally, we consider the symmetric map $(\Gamma \otimes \Gamma)\rho_{AB} = \mathbf{1}_A \otimes \mathbf{1}_B - \rho_A \otimes$

$\mathbf{1}_B - \mathbf{1}_A \otimes \rho_B + \rho_{AB}$. The states which are left invariant under this map are mixtures of generalized Bell states, which include the maximally entangled pure states as well as the product of two independent (unentangled) random bits. It can be seen that $\Gamma \otimes \Gamma$ is related to quantum nonlocality even though it does not directly reveal inseparability of two quantum bits. Indeed, it reduces to the complex conjugation in the "magic" basis that has been introduced in the context of the calculation of the entropy of formation of a pair of quantum bits (see Refs. [2, 56]). It might therefore be interesting to look for a simple relation between the map Γ (related to the reduction criterion for inseparability) and the entropy of formation.

.1 Examples

Here we consider several examples illustrating the separability criterion $\lambda_{AB} \geq 0$, and compare it to Peres' criterion [46]. Examples 1-4 deal with states of two quantum bits, and illustrate the fact that the Λ-criterion is necessary and sufficient (the spectrum of λ_{AB} is identical to the spectrum of ρ^{T_A}). Examples 5-6 illustrate that the Λ-condition is not sufficient for systems in larger dimensions (3×3 and 2×4) whose partial transpose is positive (cf. Ref. [47]). In fact, the Λ-condition is equivalent to Peres' condition for $2 \times n$ systems, so that it is also necessary and sufficient for 2×3 systems [47] while it is only necessary for larger n.

Example 1: Consider a Werner state [45] with parameter x ($0 \leq x \leq 1$), that is, a mixture of a fraction x of the singlet state $|\Psi^-\rangle$ and a random fraction $(1-x)$. We shall see that $\lambda_{AB} \geq 0$ is equivalent to Peres' criterion, and is therefore sufficient. Indeed, the joint density matrix

$$\rho_{AB} = x|\Psi^-\rangle\langle\Psi^-| + \frac{(1-x)}{4}(\mathbf{1}\otimes\mathbf{1}) = \begin{pmatrix} \frac{1-x}{4} & 0 & 0 & 0 \\ 0 & \frac{1+x}{4} & -\frac{x}{2} & 0 \\ 0 & -\frac{x}{2} & \frac{1+x}{4} & 0 \\ 0 & 0 & 0 & \frac{1-x}{4} \end{pmatrix} \tag{52}$$

is mapped by Λ into the matrix

$$\lambda_{AB} = \begin{pmatrix} \frac{1+x}{4} & 0 & 0 & 0 \\ 0 & \frac{1-x}{4} & \frac{x}{2} & 0 \\ 0 & \frac{x}{2} & \frac{1-x}{4} & 0 \\ 0 & 0 & 0 & \frac{1+x}{4} \end{pmatrix} \tag{53}$$

which admits three eigenvalues equal to $(1+x)/4$ and a fourth equal to $(1-3x)/4$. The latter becomes negative if $x > 1/3$, so that λ_{AB} is positive semi-definite only if $x \leq 1/3$, which has been proven to be the *exact* threshold for separability (any Werner state with $x \leq 1/3$ is separable as it can be written as a mixture of product states [58]). As expected, the spectrum of λ_{AB} is equal to the spectrum of the partial transpose of ρ_{AB}, so that the Λ-condition is sufficient to ensure separability for Werner states.

Example 2: Consider a mixed state that is made out of a fraction x of the entangled state $|\psi\rangle = a|01\rangle + b|10\rangle$, and fractions $(1-x)/2$ of the separable product states $|00\rangle$ and $|11\rangle$ (see [59]). The joint density matrix is of the form

$$\rho_{AB} = x|\psi\rangle\langle\psi| + \frac{1-x}{2}|00\rangle\langle 00| + \frac{1-x}{2}|11\rangle\langle 11| = \begin{pmatrix} \frac{1-x}{2} & 0 & 0 & 0 \\ 0 & x|a|^2 & xab^* & 0 \\ 0 & xa^*b & x|b|^2 & 0 \\ 0 & 0 & 0 & \frac{1-x}{2} \end{pmatrix} \tag{54}$$

with a and b satisfying $|a|^2 + |b|^2 = 1$. It is mapped by Λ into the matrix

$$\lambda_{AB} = \begin{pmatrix} x|b|^2 & 0 & 0 & 0 \\ 0 & \frac{1-x}{2} & -xab^* & 0 \\ 0 & -xa^*b & \frac{1-x}{2} & 0 \\ 0 & 0 & 0 & x|a|^2 \end{pmatrix}. \tag{55}$$

The eigenvalues of λ_{AB} are $x|a|^2$, $x|b|^2$, and $(1 - x \pm 2x|ab|)/2$. This implies that ρ_{AB} is inseparable if $x > (1 + 2|ab|)^{-1}$, exactly as predicted by Peres using the partial transpose of ρ_{AB}. Since we are dealing with two qubits, this is the exact limit between separability and inseparability [46, 47].

Example 3: In the simpler case where ρ_{AB} is a mixture of a fraction x of the singlet state $|\Psi^-\rangle$ and a fraction $(1-x)$ of the separable product state $|00\rangle$,

$$\rho_{AB} = x|\psi\rangle\langle\psi| + (1-x)|00\rangle\langle 00| = \begin{pmatrix} 1-x & 0 & 0 & 0 \\ 0 & x/2 & -x/2 & 0 \\ 0 & -x/2 & x/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{56}$$

we obtain

$$\lambda_{AB} = \begin{pmatrix} x/2 & 0 & 0 & 0 \\ 0 & 0 & x/2 & 0 \\ 0 & x/2 & 1-x & 0 \\ 0 & 0 & 0 & x/2 \end{pmatrix}. \tag{57}$$

The latter matrix admits two eigenvalues equal to $x/2$ and two eigenvalues equal to $\left(1 - x \pm \sqrt{(1-x)^2 + x^2}\right)/2$, so that its determinant is equal to $-(x/2)^4$. Thus, this state is inseparable whenever $x > 0$, as expected. (It is separable only if it is the pure product state $|00\rangle$.)

Example 4: Consider the class of 2-qubit inseparable states described by Horodecki et al. [47], a mixture of two entangled states:

$$\rho_{AB} = p|\psi_1\rangle\langle\psi_1| + (1-p)|\psi_2\rangle\langle\psi_2| \tag{58}$$

where $|\psi_1\rangle = a|00\rangle + b|11\rangle$ and $|\psi_2\rangle = a|01\rangle + b|10\rangle$, with $a, b > 0$ and satisfying $|a|^2 + |b|^2 = 1$. The joint density matrix

$$\rho_{AB} = \begin{pmatrix} pa^2 & 0 & 0 & pab \\ 0 & (1-p)a^2 & (1-p)ab & 0 \\ 0 & (1-p)ab & (1-p)b^2 & 0 \\ pab & 0 & 0 & pb^2 \end{pmatrix} \tag{59}$$

is mapped by Λ to

$$\lambda_{AB} = \begin{pmatrix} (1-p)b^2 & 0 & 0 & -pab \\ 0 & pb^2 & (p-1)ab & 0 \\ 0 & (p-1)ab & pa^2 & 0 \\ -pab & 0 & 0 & (1-p)a^2 \end{pmatrix} . \tag{60}$$

The latter matrix admits two eigenvalues equal to $\left(p \pm \sqrt{p^2 + 4a^2b^2(1-2p)}\right) \Big/ 2$ and two eigenvalues equal to $\left(1 - p \pm \sqrt{(1-p)^2 + 4a^2b^2(2p-1)}\right) \Big/ 2$, so that its determinant is equal to $-a^4b^4(1-2p)^2$. This state is therefore inseparable whenever $ab \neq 0$ and $p \neq 1/2$, in perfect agreement with Ref. [47].

Example 5: Consider the 3×3 system in a weakly inseparable state introduced by Horodecki [48],

$$\rho_{AB} = \frac{1}{1+8a} \begin{pmatrix} a & 0 & 0 & 0 & a & 0 & 0 & 0 & a \\ 0 & a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 & a & 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1+a}{2} & 0 & \frac{\sqrt{1-a^2}}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a & 0 \\ a & 0 & 0 & 0 & a & 0 & \frac{\sqrt{1-a^2}}{2} & 0 & \frac{1+a}{2} \end{pmatrix} \tag{61}$$

where a is a parameter ($a \neq 0, 1$). As shown in Ref. [48], the partial transpose of this state is positive, although ρ_{AB} is inseparable, which makes the inseparability of ρ_{AB} undetectable using Peres' criterion. It is simple to check that the Λ-mapped matrix

$$\lambda_{AB} = \frac{1}{1+8a} \begin{pmatrix} \frac{1+3a}{2} & 0 & \frac{\sqrt{1-a^2}}{2} & 0 & -a & 0 & 0 & 0 & -a \\ 0 & 2a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\sqrt{1-a^2}}{2} & 0 & \frac{1+3a}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1+3a}{2} & 0 & \frac{\sqrt{1-a^2}}{2} & 0 & 0 & 0 \\ -a & 0 & 0 & 0 & 2a & 0 & 0 & 0 & -a \\ 0 & 0 & 0 & \frac{\sqrt{1-a^2}}{2} & 0 & \frac{1+3a}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2a & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2a & 0 \\ -a & 0 & 0 & 0 & -a & 0 & 0 & 0 & 2a \end{pmatrix} \tag{62}$$

is positive (with a trace equal to 2), so that Λ cannot reveal the inseparability of ρ_{AB} either. Accordingly, the determinant of λ_{AB} is equal to $6a^7(1-a)(5a+3)/(1+8a)^9$ and thus positive. Note that the dual map also yields a positive operator $\tilde{\lambda}_{AB}$ (of trace 2), although the eigenvalues of $\tilde{\lambda}_{AB}$ are distinct from those of λ_{AB}, as is its determinant $\mathrm{Det}(\tilde{\lambda}_{AB}) = 24a^7(1-a^2)/(1+8a)^9$. This example emphasizes the fact that Λ does not result in a sufficient separability condition for 3 × 3 systems, just as Peres' condition [47].

Example 6: Following Horodecki [48], we consider a 2 × 4 system in an inseparable state

$$\rho_{AB} = \frac{1}{1+7b} \begin{pmatrix} b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & b \\ 0 & 0 & 0 & b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1+b}{2} & 0 & 0 & \frac{\sqrt{1-b^2}}{2} \\ b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & b & 0 & \frac{\sqrt{1-b^2}}{2} & 0 & 0 & \frac{1+b}{2} \end{pmatrix} \tag{63}$$

that has a positive partial transpose, where b is a parameter ($b \neq 0, 1$). Applying Λ, we see that

$$\lambda_{AB} = \frac{1}{1+7b} \begin{pmatrix} \frac{1+b}{2} & 0 & 0 & \frac{\sqrt{1-b^2}}{2} & 0 & -b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & -b & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & -b \\ \frac{\sqrt{1-b^2}}{2} & 0 & 0 & \frac{1+b}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b & 0 & 0 & 0 \\ -b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & -b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & -b & 0 & 0 & 0 & 0 & b \end{pmatrix} \tag{64}$$

has eigenvalues 0, b, $2b$, and $\left(1 + 2b \pm \sqrt{(1+2b)^2 - 2b(3+b)}\right)/2$ so that it is always non-negative. Note that the spectrum of λ_{AB} is the same as the spectrum of the partial transpose $\rho_{AB}^{T_A}$ (cf. [48]), as expected. This confirms that the condition based on $\Lambda = \Gamma \otimes I$ and Peres' separability condition are equivalent for $2 \times n$ systems (when Γ is applied to the two-dimensional system and I to the n-dimensional one). In this example, applying the dual map $\tilde{\Lambda} = I \otimes \Gamma$ yields a positive operator which traces to 3.

.2 The antiunitary map Γ

Consider the action of the map $\Gamma : \rho \to (\mathrm{Tr}\rho) - \rho$ on the density operator ρ characterizing a two-dimensional system (e.g., a quantum bit). Since ρ can be written as a linear combination of the unit matrix and the three Pauli matrices $\vec{\sigma}$ with *real* coefficients, it is sufficient to consider the action of Γ on these (Hermitian) basis matrices. We find that Γ is an *antiunitary* operator that leaves the unit matrix unchanged and flips the sign of the Pauli matrices $\sigma_{x,y,z}$,

$$\mathbf{1} \xrightarrow{\Gamma} \mathbf{1} \qquad \sigma_x \xrightarrow{\Gamma} -\sigma_x \qquad \sigma_y \xrightarrow{\Gamma} -\sigma_y \qquad \sigma_z \xrightarrow{\Gamma} -\sigma_z\,. \tag{65}$$

The complex conjugation operator K (or equivalently the transposition, as we deal with Hermitian operators) corresponds to an *antiunitary* operator which acts on the four basis matrices as

$$\mathbf{1} \xrightarrow{K} \mathbf{1} \qquad \sigma_x \xrightarrow{K} \sigma_x \qquad \sigma_y \xrightarrow{K} -\sigma_y \qquad \sigma_z \xrightarrow{K} \sigma_z\,. \tag{66}$$

(Remember that it is enough to consider the action of K on the basis matrices as the coefficients are real.) Also, $\mathcal{R}_y$ is a unitary operation characterized by the unitary matrix $U_y = \exp(-i\pi\sigma_y/2) = -i\sigma_y = \sigma_x\sigma_z$ which maps ρ into $U_y\rho U_y^\dagger = \sigma_y\rho\sigma_y$, so that the basis matrices are transformed according to

$$\mathbf{1} \xrightarrow{\mathcal{R}_y} \mathbf{1} \qquad \sigma_x \xrightarrow{\mathcal{R}_y} -\sigma_x \qquad \sigma_y \xrightarrow{\mathcal{R}_y} \sigma_y \qquad \sigma_z \xrightarrow{\mathcal{R}_y} -\sigma_z\,. \tag{67}$$

It is straightforward to check, using equations (65), (66) and (67), that Γ is the product of K and $\mathcal{R}_y$. (It is a general property of an antiunitary transformation that it can be written as the product of a unitary transformation and a fixed antiunitary operator such as time-reversal.) This can be verified easily by applying $\mathcal{R}_y K$ to a system is in a state given by equation (4.27). We get

$$U_y\rho^* U_y^\dagger = \sigma_y\rho^*\sigma_y = \frac{1}{2}(1 + \sigma_y(\vec{r}\cdot\vec{\sigma}^*)\sigma_y) = \frac{1}{2}(1 - \vec{r}\cdot\vec{\sigma}) = \Gamma\rho \tag{68}$$

where we have used the fact that $\vec{r}$ is a *real* vector and that $\sigma_y\vec{\sigma}\sigma_y = -\vec{\sigma}^*$. This generalizes what was shown in section 4.3 for pure states, namely that if

$|a\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|a^{\perp}\rangle = U_y(\alpha^*|0\rangle + \beta^*|1\rangle) = -\beta^*|0\rangle + \alpha^*|1\rangle$, then we have

$$|a^{\perp}\rangle\langle a^{\perp}| = \Gamma(|a\rangle\langle a|). \tag{69}$$

Bibliography

[1] J. Kempe, Phys. Rev. A **60**, 910 (1999).

[2] S. Hill and W. K. Wootters, Phys. Rev. Lett **78**, 5022 (1997).

[3] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1997).

[4] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).

[5] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).

[6] G. Vidal, Phys. Rev. Lett. **83**, 1046 (1999).

[7] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A **61**, 52306 (2000).

[8] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. A **63**, 012307 (2001).

[9] A. Wong and N. Christensen, Phys. Rev. A **63**, 044301 (2001).

[10] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 62314 (2000).

[11] G. Vidal, J. Mod. Opt. **47**, 355 (2000).

[12] H. Carteret, A. Higuchi, and A. Sudbery, J. Math. Phys. **41**, 7932 (2000).

[13] A. Acin, A. Andrianov, L. Costa, E. Jane, J. I. Latorre, and R. Tarrach, Phys. Rev. Lett. **85**, 1560 (2000).

[14] N. Linden and S. Popescu, Fortsch. Phys. **46**, 567 (1998).

[15] T. A. Brun and O. Cohen, Phys. Lett. A **281**, 88 (2001).

[16] A. Sudbery, J. Phys. A **34**, 643 (2001).

[17] H. Barnum and N. Linden, quant-ph/0103155 (unpublished).

[18] A. Acin, A. Andrianov, E. Jane, and R. Tarrach, quant-ph/0009107 (unpublished).

[19] A. Acin, E. Jane, W. Dür, and G. Vidal, Phys. Rev. Lett. **85**, 4811 (2000).

[20] N. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).

[21] D. G. Cory, M. D. Price, and T. F. Havel, Physica D **120**, 82 (1998).

[22] D. G. C. et al., Phys. Rev. Lett. **81**, 2152 (1998).

[23] I. L. Chuang, N. Gershenfeld, M. G. Kubinec, and D. W. Leung, Proc. Roy. Soc. London A **454**, 447 (1998).

[24] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd, Nature **393**, 143 (1998).

[25] J. A. Jones, M. Mosca, and R. H. Hansen, Nature **393**, 344 (1998).

[26] I. L. Chuang and Y. Yamamoto, Phys. Rev. A **52**, 3489 (1995).

[27] J. D. Franson and T. B. Pittman, in *Proceedings of the First NASA Conference on Quantum Computing and Quantum Communications* (Springer-Verlag, New York, 1998), Vol. 1509.

[28] J. P. Dowling, Phys. Rev. A **57**, 4736 (1998).

[29] D. Deutsch and R. Jozsa, Proc. Royal Society London: Series A **439**, 553 (1992).

[30] P. W. Shor, in *Proc. of the 35th Annual Symposium on Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, New York, 1994), pp. 124–134.

[31] L. K. Grover, in *Proc. 28th Annual ACM Symposium on the Theory of Computing* (Association for Computing Machinery, New York, 1996), p. 212.

[32] N. Cerf, L. K. Grover, and C. P. Williams, Phys. Rev. A **61**, 032303 (2000).

[33] A. Fijany and C. P. Williams, in *Proceedings of the First NASA Conference on Quantum Computing and Quantum Communications* (Springer-Verlag, New York, 1998), Vol. 1509.

[34] G. Brassard, P. Hoyer, and A. Tapp, quant-ph/9805082 (unpublished).

[35] R. Jozsa, quant-ph/9901021 (unpublished).

[36] P. G. Kwiat, J. R. Mitchell, P. D. D. Schwindt, and A. G. White, J. Mod. Opt. **47**, 257 (2000).

[37] L. K. Grover, in *Proc. 30th Annual ACM Symposium on the Theory of Computing* (Association for Computing Machinery, New York, 1998).

[38] L. K. Grover, Phys. Rev. Lett. **80**, 4329 (1998).

[39] E. Biham, O. Biham, D. Biron, M. Grassl, and D. A. Lidar, Phys. Rev. A **60**, 2742 (1999).

[40] M. Boyer, G. Brassard, P. Hoyer, and A. Trapp, Fortsch. Phys. **46**, 493 (1998).

[41] C. Zalka, Phys. Rev. A **60**, 2746 (1999).

[42] C. Zalka, private communication.

[43] L. K. Grover, quant-ph/9809029 (unpublished).

[44] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, SIAM Journal on Computing **26**, 1510 (1997).

[45] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[46] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[47] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[48] P. Horodecki, Phys.Lett. A **232**, 333 (1997).

[49] N. J. Cerf and C. Adami, Phys.Rev.Lett. **79**, 5194 (1997).

[50] N. J. Cerf and C. Adami, Phys. Rev. A. **60**, 893 (1999).

[51] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).

[52] A. Sanpera, R. Tarrach, and G. Vidal, quant-ph/9707041 (unpublished).

[53] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[54] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).

[55] R. Horodecki and M. Horodecki, Phys. Rev. A **54**, 1838 (1996).

[56] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[57] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic Press, New York, 1979).

[58] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[59] N. Gisin, Phys. Lett. A **210**, 151 (1996).