

ENTANGLED QUANTUM DYNAMICS



Matthew Leifer
School of Mathematics

October 2003

A DISSERTATION SUBMITTED TO THE UNIVERSITY OF BRISTOL
IN ACCORDANCE WITH THE REQUIREMENTS OF THE DEGREE
OF DOCTOR OF PHILOSOPHY IN THE FACULTY OF SCIENCE

Abstract

This thesis addresses the connections between quantum entanglement and quantum dynamics. The central theme is that there is a resource contained in interactions between quantum systems, as described by Hamiltonians, Unitary operators and measurements, that can be classified and quantified in a similar way to entanglement in quantum states. In particular, the problem of using an interaction Hamiltonian to simulate the action of another Hamiltonian efficiently via local operations is addressed and, for two qubits, the results can be conveniently summarized by a majorization-like partial order on the space of Hamiltonians.

Secondly, the relationship between interaction resources and entanglement is analyzed by considering the ability of a unitary operator to generate entanglement. Analytic results are presented for the case of a single application of a two-qubit unitary where the system acted upon is not entangled to any ancillary systems. Numerical results are presented for the case where such ancillas are present. The notion of processing multiple copies of a unitary operator collectively is discussed and is shown to be unnecessary for generating the maximum amount of entanglement per application of the operation.

Finally, a different kind of connection between dynamics and entanglement is investigated by finding ways in which the entanglement properties of unknown, multi-party quantum states can be determined efficiently by measurements on multiple copies of the state. Specifically, networks for directly measuring invariants under local unitary transformations and stochastic local operations and classical communication are presented and their efficiency is compared to alternative methods based on estimating the coefficients of the state.

Acknowledgements

I would like to thank my PhD supervisor Noah Linden. Throughout he has guided me towards interesting problems in quantum information theory. Also, his advice on writing, presentation and academic careers has proved invaluable. I would like to thank my other local co-authors, Sandu Popescu, Leah Henderson and Andreas Winter, for many useful discussions. Of my other co-authors, special thanks go to the group at IBM, particularly Debbie Leung and Charlie Bennett, for their open exchange of ideas and results.

Thanks are due to the rest of the Quantum Information group at Bristol for significantly developing my understanding of quantum mechanics, information and computation. Richard Josza and Sandu Popescu have been especially influential in this regard through the many group discussions and talks. Thanks also go to my fellow PhD students, Dan Collins, Stuart Presnell, Simone Severini, David Roberts and Nick Jones, all of whom have helped me at some point over the past three years.

The staff and postgraduates in the Mathematics department have provided me with a lot of support and particular thanks are due to Stephen Plasting, Brian Winn, Kristín Hákonardóttir and Miguel Marques dos Santos, for remaining friends with me throughout my PhD. On a similar note, thanks to those who have helped to organize the university Postgraduate Society with me during my time in Bristol. Finally, I have to thank my parents, Jeffrey and Janet Leifer, for their support throughout my education.

My PhD was funded by the Engineering and Physical Sciences Research Council, studentship award number 00800478.

Author's Declaration

I declare that the work in this thesis was carried out in accordance with the Regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and do not necessarily represent those of the University of Bristol. The thesis has not been presented to any other university for examination either in the United Kingdom or overseas.

Matthew Leifer

Date: October 2003

Contents

Abstract	ii
Acknowledgements	iii
Author's Declaration	iv
1 Introduction	1
1.1 The Quantum Formalism	3
1.1.1 Quantum States	3
1.1.2 Quantum evolutions	4
1.1.3 Measurements	5
1.2 Composite systems and entanglement	6
1.2.1 Using Bell States as a Communication Resource	8
1.3 Quantifying Entanglement as a Resource	13
1.3.1 Bipartite pure state entanglement	13
1.3.2 Bipartite mixed state entanglement	21
1.3.3 Multi-party entanglement	23
1.4 Dynamics as an Information Resource	24
1.5 Examples of using Quantum Dynamics as a Communication Resource	26
1.5.1 Unitary Operators	27
1.5.2 Measurements	32
1.6 Overview	32

2	Hamiltonian Simulation	35
2.1	Introduction	35
2.2	The 2-qubit simulation protocol	36
2.3	Normal form for two qubit Hamiltonians	40
2.4	Simulation of normal form two-qubit Hamiltonians	41
2.4.1	Optimization over \mathcal{P}_H	43
2.5	s-majorization	47
2.6	Conclusions	50
2.7	Related Work	51
2.A	Proof that $\mathcal{C}_H = \mathcal{P}_H$	52
2.B	More general simulation protocols	55
3	The Entangling Capacity of Quantum Gates	61
3.1	Introduction	61
3.2	Canonical form for two-qubit unitary operators	63
3.3	Single Copy Entangling Capacity	69
3.3.1	Purity of States in the Optimal Protocol	69
3.3.2	Single Application with no ancillas	71
3.3.3	Ancillas	76
3.4	Collective Processing	77
3.5	Conclusions	82
3.6	Related Work	83
4	Measuring Polynomial Invariants of Quantum States	85
4.1	Introduction	85
4.2	Polynomial Invariants under LU transformations	87
4.2.1	Pure states	87
4.2.2	Mixed states	89
4.3	Measuring Invariants under LU transformations	90
4.3.1	Network construction	90
4.4	Polynomial invariants under SLOCC	93

4.4.1	Pure states	94
4.4.2	Mixed states	95
4.4.3	Examples of $SL(2)^n$ invariants	96
4.5	Measuring SLOCC invariants	97
4.6	The Structural Physical Approximation	99
4.7	Evaluation	101
4.7.1	Statistical analysis of the network	102
4.7.2	Comparison to methods based on state estimation	104
4.8	Conclusions	107
4.A	Statistical Inference	108
4.B	Integrals over Haar measure for two-qubits	110
5	Conclusions	113
A	Notation and Conventions	117
A.1	Hilbert Spaces	117
A.2	Bases	117
A.3	Operators	118
Bibliography		120

List of Tables

1.1	States involved in the protocol for using a CNOT to communicate 1 cbit in each direction.	28
-----	---	----

List of Figures

1.1	Schematic diagram of the teleportation protocol.	9
1.2	Schematic diagram of the superdense coding protocol.	11
1.3	Schematic diagram of pure state entanglement distillation and dilution.	15
2.1	Network diagram for the two-qubit Hamiltonian simulation protocol	36
2.2	\mathcal{P}_H for the case $(h_1, h_2, h_3) = (1, 0, 0)$ viewed from the direction $(1, 1, 1)$	44
2.3	\mathcal{P}_H for the generic case $h_1 > h_2 > h_3 > 0$ viewed from the direction $(1, 1, 1)$	44
3.1	Single-copy entangling capacity and optimal initial entanglement for a general two-qubit unitary of the form of eq.(3.3) when no ancillas are allowed. Crosses show the entangling capacity and diamonds show the minimum initial entanglement of a state that achieves the capacity.	78
3.2	Single-copy entangling capacity for the CNOT family. Crosses are for no ancillas, diamonds are for one ancilla on each side and the line shows the equivalent result from [60] when the starting state is restricted to be a product between Alice and Bob	78
3.3	Single-copy entangling capacity for the DCNOT family. Crosses are for no ancillas and diamonds are for one ancilla on each side.	79

3.4	Single-copy entangling capacity for the SWAP family. Crosses are for no ancillas and diamonds are for one ancilla on each side.	79
3.5	A general entanglement generation protocol when two applications of the unitary U_{AB} are available.	80
4.1	Diagrammatic representation of the quartic two-qubit LU invariant $J_{(\sigma,\tau)}$, given in eq.(4.3). The first index of each term is represented by a circle and the second by a square. A line joins indices that are contracted with a δ .	88
4.2	General construction of network to measure polynomial LU invariants.	91
4.3	Network for measuring the 2-qubit quartic invariant.	92
4.4	Diagrammatic representation of the 3-tangle. The first index of each term is represented by a circle, the second by a square and the third by a triangle. A line joins indices that are contracted with an ϵ .	97
4.5	No. copies required $\times \epsilon$ in the asymptotic limit for the case where $b_1 = b_2 = 0$. The solid line is for the method based on estimating the state coefficients and the dashed line if for the network.	106

Chapter 1

Introduction

Despite the fact that quantum theory is now almost a century old, it is only recently that its most counter-intuitive implications have been found to have practical applications. Specifically, there is currently a revolution in the fields of information processing and computation because quantum systems appear to have radically different computational abilities from systems that obey classical physics. Of particular significance, is an efficient algorithm found by Shor [84], that uses quantum mechanics to factorize composite numbers. This is a problem of great interest in cryptography, and is generally thought to be impossible to solve efficiently using an ordinary classical computer.

At the most basic level, this revolution has come about by replacing the classical unit of information, the bit or *cbit*, by its quantum equivalent, the *qubit*. A cbit is a physical system that can be prepared in one of two definite states, usually denoted by 0 and 1. In contrast, a quantum two-level system or qubit has a continuous range of definite states in which it can be prepared, called pure states.

One of the central goals of information theory is to find efficient ways of achieving basic communication tasks between two or more separated parties¹. Such communication tasks include the transmission of messages and the dis-

¹The standard convention of naming the first two parties Alice and Bob will be used in this thesis.

tribution of secure keys for cryptography. The parties will typically have some communication resources available to them, which they would like use as efficiently as possible.

For example they may have classical resources, such as telephone lines, to send cbits to one another. These classical channels may introduce noise into the signal, but this can be compensated for by encoding some redundancy into the data, for example by sending each cbit three times and deciding on the true value by majority vote. One may then ask what is the minimum number of bits that have to be sent through the noisy channel in order to communicate 1 cbit perfectly. Another way of stating this is to say that the noisy channel can be used to *simulate* the perfect channel and the number of uses required to do this quantifies the amount of communication resources contained in the channel, relative to the perfect channel (see [30] for details).

Quantum mechanics introduces new resources into the framework, such as quantum channels for sending qubits to one another. An example might be an optical fibre for sending single-photons. One may then ask questions about the resource content of these quantum channels. For example, how many uses of a noisy quantum channel does it take to simulate a perfect one? There are also new types of question that can be introduced, such as how many uses of a quantum channel does it take to simulate a classical one and can a classical channel be used to simulate a quantum one?²

In addition to qubit channels, there are many other features of quantum mechanics that do not exist in classical physics and one of the most important of these is entanglement. Entanglement has a long history in quantum mechanics and it is responsible for many of the most counter intuitive features of the theory, particularly the EPR paradox [39] and Bell's proof of non-locality [7]. For a long time, it was regarded as a problematic concept and the majority of research that tackled the issue was aimed at highlighting its seemingly paradoxical implications. However, more recently, it has been realized that

²For general introductions to quantum information theory see [74, 78].

entanglement can be turned to our advantage and regarded as a new kind of resource in information theory. Also, from a practical point of view, experiments in many quantum systems are starting to reach the stage where entanglement can be manipulated in a controlled way.

The central theme of this thesis is that quantum *dynamics* can be quantified as a resource in a similar way to entanglement. Before introducing this concept, the basic formalism of quantum mechanics is reviewed in §1.1, in part to establish the notation used throughout the rest of this thesis. Then, in §1.2 the formalism for discussing composite systems and entanglement is reviewed and some important applications of entanglement: teleportation and superdense coding, are introduced in §1.2.1. Some of the ways in which entanglement can be quantified as a resource are reviewed in §1.3. The idea of using quantum dynamics as a resource is introduced in §1.4 and illustrated with some examples of quantum gates in §1.5. The introduction concludes in §1.6 with an overview of the main results presented in this thesis.

1.1 The Quantum Formalism

1.1.1 Quantum States

The state of a quantum system is described by a positive semi-definite density operator, ρ , on a Hilbert space, \mathcal{H} . $\mathfrak{B}(\mathcal{H})$ will be used to denote the space of linear operators on the Hilbert space \mathcal{H} , so these conditions can be written as $\rho \in \mathfrak{B}(\mathcal{H}), \rho \geq 0$. It is usually required that ρ is normalized such that $\text{Tr}(\rho) = 1$.

In the special case where ρ is of rank 1, the state is *pure* and it can alternatively be represented by a vector $|\psi\rangle \in \mathcal{H}$, where $\rho = |\psi\rangle\langle\psi|$. In doing so, an arbitrary global phase is introduced, since $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ have the same density matrix. The choice of this phase does not affect the physical predictions that can be made about the state, so it can be chosen to have any convenient

value. In what follows, \mathcal{H} will usually be finite dimensional.

The most basic system in quantum information theory is the qubit, for which $\mathcal{H} = \mathbb{C}^2$, a 2 dimensional complex Hilbert space. The standard orthonormal basis for \mathbb{C}^2 , often called the *computational* basis, is given by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.1)$$

and a pure qubit state can be written in this basis as

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \quad (1.2)$$

1.1.2 Quantum evolutions

The evolution of a closed quantum system in the absence of measurements is described by the Schrödinger equation

$$i\hbar \frac{\partial \rho(t)}{\partial t} = [H, \rho(t)] \quad (1.3)$$

where H is the self-adjoint Hamiltonian operator, which represents the energy of the system and $[H, \rho(t)] = H\rho(t) - \rho(t)H$. For pure states, this can be written as

$$-i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H |\psi(t)\rangle \quad (1.4)$$

The resulting evolution is given by $\rho(t) = U(t)\rho(0)U(t)^\dagger$, where $U(t)$ is the unitary time evolution operator $U(t) = e^{-\frac{iH}{\hbar}t}$. For pure states this can alternatively be written as $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ ³. By suitably engineering the Hamiltonian of the system, any unitary operator may be realized as the time evolution operator. Controlling the evolution of a quantum system in this way is experimentally challenging, but this thesis is mainly concerned with the general limits to information processing imposed by the laws of quantum mechanics rather than the problems associated with generating particular physical evolutions. Thus, in this thesis it will usually be assumed that any unitary operator can be generated by engineering Hamiltonians in this way.

³In the rest of this thesis units are chosen such that $\hbar = 1$ and $U(t) = e^{-iHt}$.

A convenient basis for operators on \mathbb{C}^2 is given by the identity matrix and the 3 Pauli matrices.

$$\sigma_0 = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.5)$$

It is sometimes convenient to use a vector notation for the Pauli matrices $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^T$.

1.1.3 Measurements

A measurement in quantum mechanics is described by a set of operators M_j satisfying

$$\sum_j M_j^\dagger M_j = I \quad (1.6)$$

where I is the identity operator on \mathcal{H} . The labels on the operators represent the different possible outcomes and when the measurement is performed on a state ρ , the outcome j occurs with probability

$$p_j = \text{Tr} \left(M_j^\dagger M_j \rho \right) \quad (1.7)$$

After the measurement, if the outcome was j then the state becomes

$$\rho_j = \frac{M_j \rho M_j^\dagger}{\text{Tr} \left(M_j^\dagger M_j \rho \right)} \quad (1.8)$$

When the initial state of the system is pure, (1.7) may be replaced by

$$p_j = \langle \psi | M_j^\dagger M_j | \psi \rangle, \quad (1.9)$$

and (1.8) may be replaced by

$$|\psi_j\rangle = \frac{M_j |\psi\rangle}{\sqrt{\langle \psi | M_j^\dagger M_j | \psi \rangle}} \quad (1.10)$$

If the state after the measurement is not of interest, it is convenient to describe the measurement by the operators $E_j = M_j^\dagger M_j$, since the measurement

operators only appear in (1.7) and (1.9) in this combination. Generally, each E_j is a positive operator and $\sum_j E_j = I$. The set of operators E_j is known as a Positive Operator Valued Measure (POVM).

A particularly important class of measurements are the orthogonal projective measurements. For these measurements, each $M_j = \Pi_j$ is a projector satisfying

$$\Pi_j \Pi_k = \delta_{jk} \Pi_j \quad (1.11)$$

Such measurements can alternatively be described by a self adjoint operator $A = \sum_j \lambda_j \Pi_j$ called an observable. Here, the λ_j 's are the distinct real eigenvalues of A and λ_j is the value of the observable A when the outcome j is observed. Using, (1.7) and (1.11), the mean value of A is given by

$$\begin{aligned} \langle A \rangle &= \sum_j p_j \lambda_j = \sum_j \lambda_j \text{Tr}(\Pi_j \rho) \\ &= \text{Tr}\left(\sum_j \lambda_j \Pi_j \rho\right) = \text{Tr}(A \rho) \end{aligned} \quad (1.12)$$

which reduces to

$$\langle A \rangle = \langle \psi | A | \psi \rangle \quad (1.13)$$

for pure states.

When each projector is of rank 1, then they can be written in terms of vectors as $\Pi_j = |\phi_j\rangle\langle\phi_j|$. (1.11) and (1.6) imply that these vectors form a complete orthonormal basis, so this type of measurement is often called a measurement in the basis $\{|\phi_j\rangle\}$.

1.2 Composite systems and entanglement

Entanglement occurs in systems of two or more parties, so it is necessary to introduce the quantum formalism for composite systems. The focus is on the bipartite case here, but the generalization to more than 2 parties proceeds in the obvious manner.

Suppose there are two parties, Alice and Bob, who each have their own Hilbert space, \mathcal{H}_A and \mathcal{H}_B respectively. According to quantum mechanics the

state space of their combined system is given by $\mathcal{H}_A \otimes \mathcal{H}_B$, the tensor product of the two subsystems. $\mathcal{H}_A \otimes \mathcal{H}_B$ is spanned by the vectors $|i\rangle_A \otimes |j\rangle_B$, where $\{|i\rangle_A\}, \{|j\rangle_B\}$ are basis vectors for $\mathcal{H}_A, \mathcal{H}_B$ respectively. For example, if Alice and Bob both have a qubit then the state space is $\mathbb{C}^2 \otimes \mathbb{C}^2$, which is spanned by the vectors $|0\rangle_A \otimes |0\rangle_B, |0\rangle_A \otimes |1\rangle_B, |1\rangle_A \otimes |0\rangle_B, |1\rangle_A \otimes |1\rangle_B$. The tensor product symbol is often omitted from these vectors, so the vector $|0\rangle_A \otimes |0\rangle_B$ might be written as $|0\rangle_A |0\rangle_B$ or simply $|00\rangle_{AB}$ for example.

The description of the state space of a composite system as a tensor product is quite general, but when discussing entanglement it is useful to imagine that Alice and Bob are separated by a large distance and that their quantum systems are well localized in the spatial part of the wave-function. This allows their systems to be treated as distinguishable particles. The labels A and B then refer implicitly to these distinct, localized spatial wave-functions.

The overall quantum state of a bipartite system is a density operator $\rho_{AB} \in \mathfrak{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Alice's reduced density operator, $\rho_A = \text{Tr}_B(\rho_{AB})$, is defined by tracing over any orthonormal basis for Bob's system. For example, if Bob's system is a qubit then $\rho_A = \langle 0|_B \rho_{AB} |0\rangle_B + \langle 1|_B \rho_{AB} |1\rangle_B$. Bob's reduced density operator is defined in a similar way: $\rho_B = \text{Tr}_A(\rho_{AB})$.

If the overall state is pure then it can alternatively be defined as a vector $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$. This state is *separable* if it can be written as $|\psi\rangle_{AB} = |\phi\rangle_A \otimes |\eta\rangle_B$, for some $|\phi\rangle_A \in \mathcal{H}_A$ and $|\eta\rangle_B \in \mathcal{H}_B$. If the state is separable then both Alice and Bob's reduced states will be pure states. Non-separable states are called *entangled* states and in this case the reduced states will both be mixed.

For density matrices a separable state is defined as one that can be written $\rho_{AB} = \sum_j p_j (|\psi_j\rangle_A \otimes |\phi_j\rangle_B) (\langle\psi_j|_A \otimes \langle\phi_j|_B)$, for some vectors $|\psi_j\rangle_A \in \mathcal{H}_A, |\phi_j\rangle_B \in \mathcal{H}_B$ and some probabilities $p_j > 0, \sum_j p_j = 1$. Separable states are an important class of states because they can be prepared without the need for any entangling interaction or prior entanglement between Alice and Bob's subsystems. For example, one way of generating an ensemble of states

described by ρ_{AB} , is for Alice to generate a random variable with outcomes j distributed according to the probabilities p_j . Then, if outcome j is obtained, she prepares the state $|\psi_j\rangle_A$. Finally, she lets Bob know which state she prepared via a classical channel and Bob prepares the corresponding $|\phi_j\rangle_B$.

An example of an entangled pure state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is

$$|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B) \quad (1.14)$$

In fact, an orthonormal basis of entangled states is given by (1.14) and the three additional states

$$\begin{aligned} |\phi^-\rangle_{AB} &= \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |0\rangle_B - |1\rangle_A \otimes |1\rangle_B) \\ |\psi^+\rangle_{AB} &= \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |0\rangle_B) \\ |\psi^-\rangle_{AB} &= \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B) \end{aligned} \quad (1.15)$$

These states are called *Bell states* and they play a central role in many protocols in quantum information theory. The unit of entanglement, the *ebit*, is defined to be the amount of entanglement contained in a Bell state. The role of ebits will become clear in §1.2.1, where the use of Bell states in two important protocols: teleportation [9] and superdense coding [14], is reviewed and in §1.3, which reviews the quantification of entanglement in quantum states.

1.2.1 Using Bell States as a Communication Resource

The two communication protocols reviewed here provide an operational interpretation of the ebit and are of importance for discussing the quantification of entanglement. They will also be used in §1.5 and in chapter 3 to provide bounds on the entanglement and classical communication capabilities of quantum gates.

Teleportation

Suppose Alice has a single qubit in an unknown state $|\psi\rangle$ that she would like to transmit to Bob, but there is no quantum channel available between

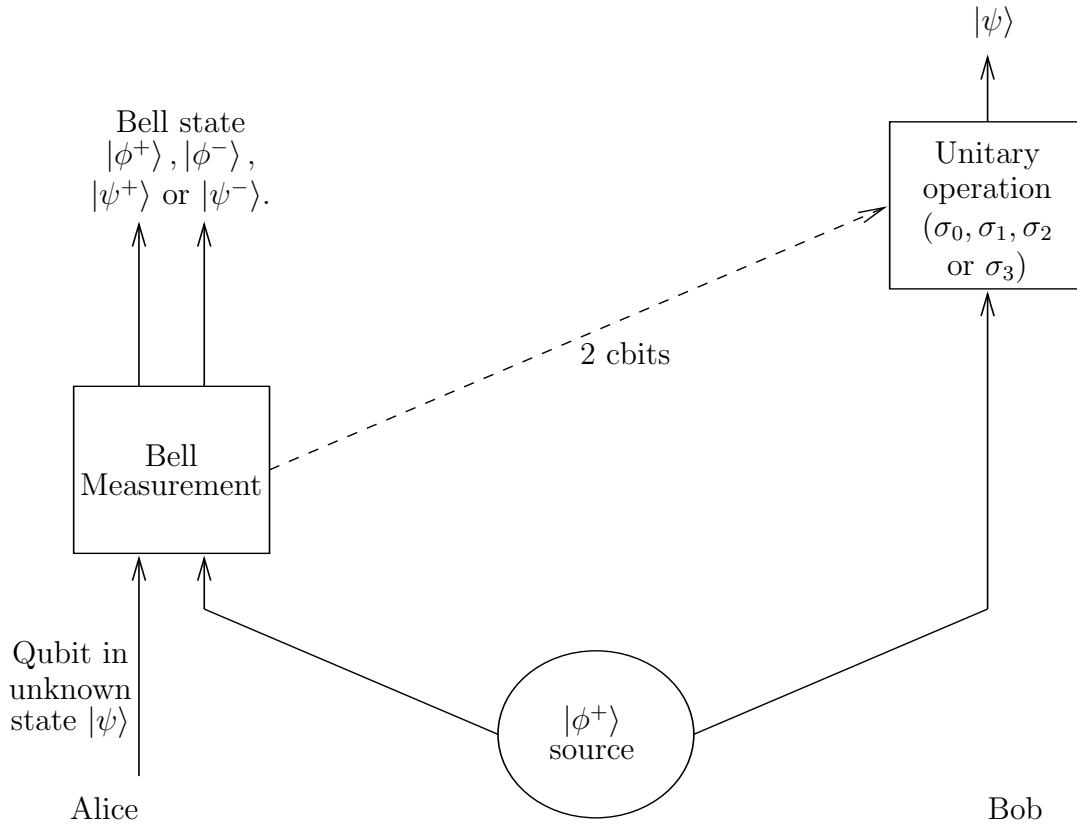


Figure 1.1: Schematic diagram of the teleportation protocol.

them. However, they do share a Bell state $|\phi^+\rangle$ and they have the ability to communicate classical information. They can achieve the task by using the entangled state and transmitting two cbits from Alice to Bob, by a protocol known as teleportation. It is illustrated schematically in fig. 1.1. To see how teleportation works, denote the space of Alice's half of the entangled state by A , the space of Bob's half by B and the space of the unknown state by A' . A general unknown pure state can be written as $|\psi\rangle_{A'} = \alpha|0\rangle_{A'} + \beta|1\rangle_{A'}$, where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. The total initial state of Alice's and Bob's

systems can be written as

$$\begin{aligned}
 |\psi\rangle_{A'} \otimes |\phi^+\rangle_{AB} &= (\alpha|0\rangle + \beta|1\rangle)_{A'} \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{AB} \\
 &= \frac{1}{2} \left[\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{A'A} \otimes (\alpha|0\rangle + \beta|1\rangle)_B \right. \\
 &\quad + \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)_{A'A} \otimes (\alpha|0\rangle - \beta|1\rangle)_B \\
 &\quad + \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)_{A'A} \otimes (\alpha|1\rangle + \beta|0\rangle)_B \\
 &\quad \left. + \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)_{A'A} \otimes (\alpha|1\rangle - \beta|0\rangle)_B \right] \\
 &= \frac{1}{2} [|\phi^+\rangle_{A'A} \otimes |\psi\rangle_B + |\phi^-\rangle_{A'A} \otimes \sigma_3 |\psi\rangle_B \\
 &\quad + |\psi^+\rangle_{A'A} \otimes \sigma_1 |\psi\rangle_B + |\psi^-\rangle_{A'A} \otimes i\sigma_2 |\psi\rangle_B]
 \end{aligned} \tag{1.16}$$

From the last line of this equation, one can see that if Alice performs a measurement in the Bell basis on her two qubits A and A' described by the four projectors $|\phi^\pm\rangle_{A'A} \langle\phi^\pm|_{AA'}$, $|\psi^\pm\rangle_{A'A} \langle\psi^\pm|_{AA'}$, then each measurement outcome will occur with probability $\frac{1}{4}$. The system will be left in one of the four states $|\phi^+\rangle_{A'A} \otimes |\psi\rangle_B$, $|\phi^-\rangle_{A'A} \otimes \sigma_3 |\psi\rangle_B$, $|\psi^+\rangle_{A'A} \otimes \sigma_1 |\psi\rangle_B$, $|\psi^-\rangle_{A'A} \otimes i\sigma_2 |\psi\rangle_B$. The resulting state of Bob's system will be Alice's original unknown state multiplied by one of the four unitary operators $\sigma_0, \sigma_3, \sigma_1, i\sigma_2$ depending on which outcome was obtained by Alice. If Alice tells Bob which outcome she obtained, then Bob can undo the unitary operator by applying the inverse and recover Alice's original unknown state. Since there are four outcomes, this can be done by sending 2 cbits from Alice to Bob.

The resources used in this teleportation protocol can be conveniently summarized by the relation

$$1 \text{ ebit} + 2 \text{ cbits}_{A \rightarrow B} \Rightarrow 1 \text{ qubit}_{A \rightarrow B} \tag{1.17}$$

which should be read as 1 Bell state and the transmission of 2 cbits via a classical channel from Alice to Bob can be used to simulate the transmission of 1 qubit via a quantum channel from Alice to Bob.

Teleportation is an important primitive in quantum information theory, and it gives meaning to the concept of an ebit, which can be regarded as the amount of entanglement required to achieve a perfect teleportation.

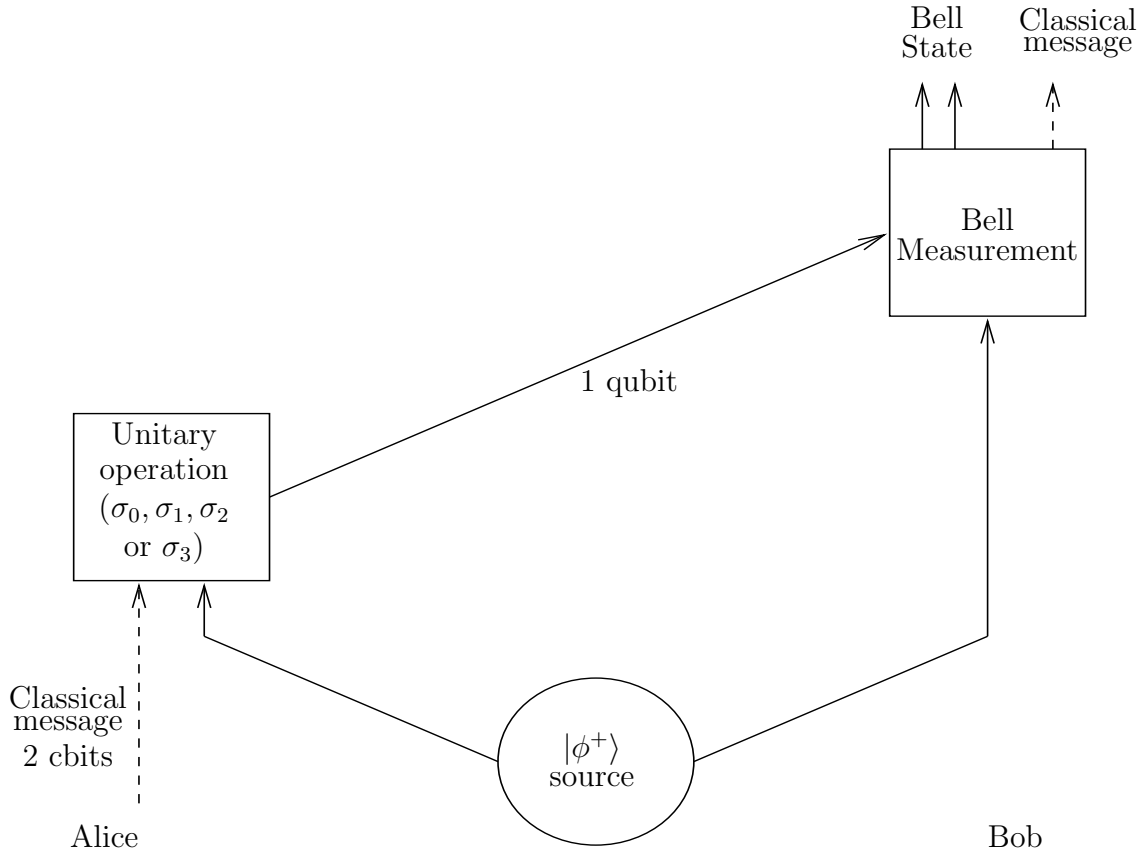


Figure 1.2: Schematic diagram of the superdense coding protocol.

Superdense Coding

In addition to transmitting quantum states, a quantum channel can be used to transmit a cbits from Alice to Bob. One simple way of doing this is for Alice to send the state $|0\rangle$ to Bob if the cbit to be sent is 0 and $|1\rangle$ if the message is 1. Bob can then perform a measurement in the basis $\{|0\rangle, |1\rangle\}$ to recover the initial message. However, a qubit is specified by continuous parameters, so one might suppose that this can be exploited to transmit more than one cbit from Alice to Bob for each qubit that is transmitted. It has been shown [72] that this is not possible if the transmitted qubit is not entangled to any other qubits in Bob's possession. On the other hand, if Alice and Bob share a Bell state to begin with, then Alice can transmit two cbits to Bob by sending a single qubit. The protocol for doing this is called superdense coding and it is illustrated in fig. 1.2.

To achieve superdense coding, suppose Alice and Bob start with the Bell

state $|\phi^+\rangle_{AB}$. Alice then applies one of the following unitary operators to her half of the state, depending on which 2 cbit message is to be sent.

$$\begin{aligned}
 \sigma_0 & \text{ if the message is } 00 \\
 \sigma_1 & \text{ if the message is } 01 \\
 \sigma_2 & \text{ if the message is } 10 \\
 \sigma_3 & \text{ if the message is } 11
 \end{aligned}
 \tag{1.18}$$

The resulting state will be one of the four Bell states, up to a global phase. If Alice then transmits her half of the entangled state to Bob, he can recover the original message by performing a measurement in the Bell basis.

The resources used in this protocol can be conveniently summarized by the following relation.

$$1 \text{ ebit} + 1 \text{ qubit}_{A \rightarrow B} \Rightarrow 2 \text{ cbits}_{A \rightarrow B}
 \tag{1.19}$$

The superdense coding protocol gives another operational meaning to the notion of an ebit. It is the amount of entanglement required to achieve the perfect transmission of 2 cbits with only a single use of a qubit channel.

Note that both the teleportation and superdense coding protocols could alternatively be achieved using any state of the form $U_A \otimes U_B |\phi^+\rangle_{AB}$, where $U_A, U_B \in U(2)$, instead of $|\phi^+\rangle_{AB}$. For example, all of the Bell states (1.15) are of this form. To do this, Alice and Bob would simply have to apply the inverse operations U_A^\dagger, U_B^\dagger at the beginning of the protocol. Thus, all these states contain an ebit of entanglement because they are equivalent with respect to the communication tasks they can be used to achieve. This notion of equivalence, called *local* or *local unitary (LU)* equivalence, can be extended to all states and is of central importance in the classification and quantification of entanglement.

1.3 Quantifying Entanglement as a Resource

The teleportation and superdense coding protocols demonstrate that Bell states may be used as a communication resource. What about other entangled states? As with all resources in information theory, their ability to perform communication tasks can be quantified by asking what other resources they can be used to simulate. In particular, this can be done by asking how many of them are needed to create a Bell state.

Before this can be done, the actions that the parties should be allowed to perform in these protocols must be established. Clearly, any action which can create Bell states starting from no initial entanglement, such as allowing quantum interactions or quantum channels between the parties, must be ruled out. In contrast, the teleportation and superdense coding protocols consume entanglement and hence provide an insight into the type of action that should be allowed. In both of these protocols, Alice and Bob's actions always consist of performing operations locally on their own subsystems and communicating with one another via classical channels. Generally, this type of action is known as *Local Operations and Classical Communication (LOCC)*. LOCC is quite a general class of actions to allow because it typically consumes entanglement and never generates entanglement on average. Some results on the quantification of entanglement under LOCC are reviewed in the following sections.

1.3.1 Bipartite pure state entanglement

Suppose Alice and Bob share a bipartite state $|\phi\rangle_{AB}$ in a finite dimensional Hilbert space $\mathcal{H}_{AB} = \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, where \mathbb{C}^{d_A} is the Hilbert space of Alice's portion of the system and \mathbb{C}^{d_B} is the Hilbert space of Bob's portion of the system. The entanglement properties of $|\phi\rangle_{AB}$ will be unaffected if either Alice or Bob perform a reversible operation, such as a unitary operation, on their portion of the system. This leads to the notion of local equivalence of

states defined as follows.

$$\begin{aligned} |\phi'\rangle_{AB} \sim |\phi\rangle_{AB} \quad \text{iff} \quad & \exists U_A \in U(d_A), U_B \in U(d_B) \\ \text{s.t.} \quad & |\phi'\rangle_{AB} = U_A \otimes U_B |\phi\rangle_{AB} \end{aligned} \quad (1.20)$$

Since the entanglement properties of locally equivalent states are the same, it is convenient to use a unique canonical representative of each equivalence class under the relation (1.20). This is provided by the Schmidt form [82], which is given by

$$|\phi\rangle_{AB} \sim \sum_{j=1}^d \sqrt{p_j} |j\rangle_A \otimes |j\rangle_B \quad (1.21)$$

where $\sum_j p_j = 1$, $p_1 \geq p_2 \geq \dots \geq p_d \geq 0$ and $d = \min(d_A, d_B)$. It follows from this that all states with the same Schmidt form are locally equivalent. The number of non-zero coefficients in this form is called the Schmidt number of the state and is denoted $N_{\text{Sch}}(|\phi\rangle_{AB})$. The Schmidt decomposition is used extensively in chapter 3.

The proof of (1.21) can be illustrated simply for the case where $d_A = d_B = d$. In this case, a general pure state can be written as $|\phi\rangle_{AB} = \sum_{j,k=1}^d \alpha_{jk} |j\rangle_A \otimes |k\rangle_B$, where $\sum_{jk} |\alpha_{jk}|^2 = 1$. A local unitary operation then transforms this state to

$$\begin{aligned} |\phi'\rangle_{AB} &= U_A \otimes U_B |\phi\rangle_{AB} \\ &= \sum_{jk} \alpha_{jk} U_A |j\rangle_A \otimes U_B |k\rangle_B \\ &= \sum_{jkmn} \alpha_{mn} (U_A)_{mj} |j\rangle_A \otimes (U_B)_{nk} |k\rangle_B \end{aligned} \quad (1.22)$$

where $(U_A)_{jk} = \langle j|_A U_A |k\rangle_A$ are the components of the matrix representation of U_A in the computational basis and similarly $(U_B)_{jk} = \langle j|_B U_B |k\rangle_B$. Thus, regarding α_{jk} as the components of a $d \times d$ matrix, the transformed coefficients can be written as

$$\alpha' = U_A^T \alpha U_B \quad (1.23)$$

By the singular value decomposition [17], U_A and U_B can be chosen such that α' is diagonal with components given by the singular values of α , which are the eigenvalues of $\sqrt{\alpha\alpha^\dagger}$. Furthermore, the normalization of α implies that the squares of the singular values sum to one. Finally, the singular values may

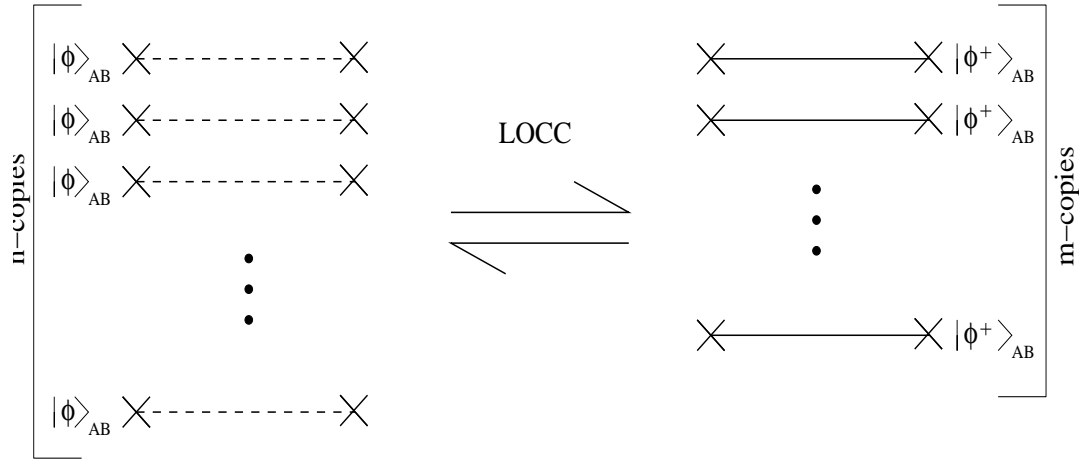


Figure 1.3: Schematic diagram of pure state entanglement distillation and dilution.

be placed in any order along the diagonal of α' , since the permutation matrices required to do this are unitary. Thus, choosing them to be in non-increasing order makes the decomposition unique.

For the case where $d_A \neq d_B$ the matrix α will not be square. This can be dealt with by embedding the smaller subsystem in a larger Hilbert Space and padding the matrix α with zeroes appropriately.

Note that the matrix $\alpha\alpha^\dagger$ is simply Alice's reduced density operator expressed as a matrix in the computational basis. Similarly, the matrix $\alpha^\dagger\alpha$ is Bob's reduced density operator in this basis and it has the same eigenvalues. Thus, the Schmidt coefficients are given by the square roots of the eigenvalues of either of the reduced density operators and the Schmidt number is given by their rank.

Entanglement distillation and dilution

Since an arbitrary state, $|\phi\rangle_{AB}$, is locally equivalent to its Schmidt form, questions about how many Bell states can be created from $|\phi\rangle_{AB}$ via LOCC can be addressed without loss of generality by restricting to states of this form. The problem is also greatly simplified by allowing Alice and Bob to process multiple copies of $|\phi\rangle_{AB}$ collectively. The general framework is illustrated in

fig. 1.3. Initially, Alice and Bob start with n copies of $|\phi\rangle_{AB}$ and by performing collective LOCC on them they obtain m Bell states. Similarly, they could start with Bell states and perform LOCC to obtain the state $|\phi\rangle_{AB}$. Protocols for achieving these tasks are known as entanglement *distillation* and *dilution* respectively.

The two most commonly used entanglement measures are the entanglement of distillation E_D and the entanglement cost E_C , defined by

$$\begin{aligned} E_D(|\phi\rangle_{AB}) &= \lim_{n \rightarrow \infty} \max_{|\phi\rangle_{AB}^{\otimes n} \xrightarrow{\text{LOCC}} |\phi^+\rangle_{AB}^{\otimes m}} \left(\frac{m}{n}\right) \\ E_C(|\phi\rangle_{AB}) &= \lim_{n \rightarrow \infty} \min_{|\phi^+\rangle_{AB}^{\otimes m} \xrightarrow{\text{LOCC}} |\phi\rangle_{AB}^{\otimes n}} \left(\frac{m}{n}\right) \end{aligned} \quad (1.24)$$

E_D is the upper limit on the number of Bell states that can be “distilled” per copy of the state $|\psi\rangle_{AB}$ and E_C is the lower limit on the number of Bell states it “costs” to form each copy of the state by dilution.

For bipartite pure states, the process of converting states to singlets via LOCC is asymptotically reversible, i.e. E_D and E_C are the same. They are given by the following formula [8].

$$\begin{aligned} E_D(|\phi\rangle_{AB}) = E_C(|\phi\rangle_{AB}) &= -\text{Tr} \rho_A \log_2 \rho_A = -\text{Tr} \rho_B \log_2 \rho_B \\ &= -\sum_j p_j \log_2 p_j \end{aligned} \quad (1.25)$$

The quantity $E = -\text{Tr} \rho_A \log_2 \rho_A$ is also known as the *entropy of entanglement*.

To illustrate how this works, consider a protocol for distilling two-qubit states that achieves this asymptotic limit. Without loss of generality, suppose Alice and Bob start with n copies of the state $|\phi\rangle_{AB} = \sqrt{1-p}|00\rangle_{AB} + \sqrt{p}|11\rangle_{AB}$. The total state of their system will be

$$|\psi\rangle^{\otimes n} = \left(\sqrt{1-p}|00\rangle + \sqrt{p}|11\rangle \right)^{\otimes n} \quad (1.26)$$

The next step is for Alice to perform a measurement on her subsystem described by projectors onto subspaces spanned by computational basis states with the same number of 1’s. For example the first two projectors would be

$\Pi_0 = |0\rangle^{\otimes n} \langle 0|^{\otimes n}$, $\Pi_1 = |1000 \dots 0\rangle \langle 1000 \dots 0| + |0100 \dots 0\rangle \langle 0100 \dots 0| + \dots + |0000 \dots 1\rangle \langle 0000 \dots 1|^4$.

Expanding (1.26) gives 2^n terms, of which there are $\binom{n}{k}$ terms where exactly k of Alice's qubits are in the state $|1\rangle$. Each of these terms has a coefficient $\sqrt{p^k(1-p)^{n-k}}$. Thus, the probability of outcome k is

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (1.27)$$

i.e. a binomial distribution with parameter p . For large n this distribution will be strongly peaked around the mean value $k = np$. If Alice obtains the measurement outcome k then the state will be an equal superposition of $\binom{n}{k}$ orthogonal terms, where each term is a product of a state with exactly k ones on Alice's side with the same state on Bob's side. Such a state is entangled and it remains to show how it can be transformed into Bell states.

In the large n limit⁵, the number of terms in the superposition will be

$$\binom{n}{np} = \frac{n!}{(np)!(n-np)!} \approx \frac{\left(\frac{n}{e}\right)^n}{\left(\frac{np}{e}\right)^{np} \left(\frac{n-np}{e}\right)^{n-np}} = 2^{nh(p)} \quad (1.28)$$

where $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$ is the Shannon entropy and Stirling's approximation in the form $x! \approx (x/e)^x$ has been used.

If $nh(p)$ happens to be an integer, then there is a local unitary transformation that takes the state to a product of $nh(p)$ Bell states. This is because a product of m Bell states has Schmidt number 2^m and the Schmidt coefficients are all equal. The existence of a local unitary transformation then follows because all states with the same Schmidt decomposition are locally equivalent. Of course, $nh(p)$ need not be an integer, but if Alice and Bob perform the above protocol for l batches of n copies of the state then they will end up with a state with $N = \binom{n}{k_1} \binom{n}{k_2} \dots \binom{n}{k_l}$ equally weighted terms in its Schmidt decomposition, where each k_j is close to np . For any $\epsilon > 0$ this number will

⁴If the $|0\rangle$ and $|1\rangle$ states are the spin down and spin up states of a spin- $\frac{1}{2}$ particle then this is a measurement of the total spin of Alice's system.

⁵For simplicity, the asymptotics are only sketched here. A more rigorous treatment can be found in [8].

eventually be close to a power of 2 for some l , i.e. $\forall \epsilon > 0, \exists r_l \in \mathbb{Z}^+$ such that

$$2^{r_l} \leq N < 2^{r_l}(1 + \epsilon) \quad (1.29)$$

When this occurs, Alice can project the state into a subspace of dimension 2^{r_l} , succeeding with probability $(1 - \epsilon)$, and then Alice and Bob can perform local unitary operations to obtain r_l Bell states. r_l will typically be close to $nlh(p)$, so they have succeeded in distilling $h(p)$ Bell states for each copy of the state they started with.

In this description, only local operations by Alice and Bob have been mentioned so far. However, Alice would also have to communicate her measurement results to Bob so that he knows the subspace in which the resulting state lies in order to implement his local unitary transformation. Equivalently, he could simply replicate the measurements performed by Alice on his side because the form of the state guarantees that he will obtain the same outcome and not disturb the state further.

To complete the proof of (1.25), one needs to show that the state $|\phi\rangle$ can be diluted using $h(p)$ Bell states for each copy of the state. This can be done using quantum data compression, details of which can be found in [83]. Specifically, Alice prepares multiple copies of the state $|\phi\rangle$ locally, compresses one qubit from each state and then uses Bell states to teleport the compressed states to Bob who can then decompress them.

The reversibility of entanglement distillation and dilution for pure states is used in chapter 3 in the discussion of collective processing of quantum operations. The entropy of entanglement is also used as an entanglement measure in that chapter.

Single-copy entanglement manipulation

Whilst the entropy of entanglement is the unique measure of entanglement for bipartite pure states in the sense discussed above [77, 96], there are still other questions that could be asked to give a more detailed classification of

entanglement. For example, given just a single copy of the state $|\phi\rangle_{AB}$, which other states can be obtained with certainty by LOCC? The answer to this question was provided in [73], where the following result was obtained. First define the quantities

$$M_k(|\phi\rangle_{AB}) = \sum_{j=1}^k p_j \quad (1.30)$$

where p_j are the Schmidt coefficients of $|\phi\rangle_{AB}$ for $j \leq N_{\text{Sch}}(|\phi\rangle_{AB})$ and $p_j = 0$ otherwise. Then $|\phi'\rangle_{AB}$ can be obtained with certainty from a single copy of $|\phi\rangle_{AB}$ by LOCC iff

$$M_k(|\phi\rangle_{AB}) \leq M_k(|\phi'\rangle_{AB}) \quad (1.31)$$

with equality for $k \geq \max(N_{\text{Sch}}(|\phi\rangle_{AB}), N_{\text{Sch}}(|\phi'\rangle_{AB}))$. (1.31) is an example of a majorization relation and these are discussed in more detail in chapter 2, where they are compared to the rather similar classification of two-qubit Hamiltonians that is derived there. The quantities M_k are known as *entanglement monotones* and they can also be regarded as entanglement measures.

The majorization result shows that a Bell state cannot be distilled with certainty from a single copy of any partially entangled two-qubit state. However, it can be done if a probability of failure is tolerated. To illustrate this, suppose Alice and Bob share a state $|\phi\rangle_{AB} = \sqrt{p}|00\rangle_{AB} + \sqrt{1-p}|11\rangle_{AB}$, where $\frac{1}{2} < p < 1$. Alice can perform a measurement described by the operators $M_1 = \sqrt{\frac{1-p}{p}}|0\rangle_A\langle 0|_A + |1\rangle_A\langle 1|_A$, $M_2 = \sqrt{\frac{2p-1}{p}}|0\rangle_A\langle 0|_A$. Outcome 1 occurs with probability $2(1-p)$ and results in the production of the Bell state $|\phi^+\rangle_{AB}$. Outcome 2 represents a failure, since the resulting state is a product $|00\rangle_{AB}$.

In fact, this is the optimal procedure generating a Bell state from a single copy and it is an example of a general result proved in [92], where the optimal probability of generating a single copy of a state $|\phi'\rangle_{AB}$ from a single copy of the state $|\phi\rangle_{AB}$ via LOCC is shown to be

$$P(|\phi\rangle_{AB} \xrightarrow{\text{LOCC}} |\phi'\rangle_{AB}) = \min_k \frac{M_k(|\phi\rangle_{AB})}{M_k(|\phi'\rangle_{AB})} \quad (1.32)$$

However, the relations (1.31) and (1.32) do not represent the end of the story as far as a single copy pure state entanglement manipulation is concerned. In particular, it is possible to convert between some pairs of states that do not satisfy (1.31) via LOCC and boost the probability (1.32) if Alice and Bob have access to another entangled state, even if this second state is left unchanged by the protocol [56]. This effect is known as entanglement *catalysis* because the second state plays a role analogous to a catalyst in a chemical reaction. For example, suppose Alice and Bob want to convert the state

$$|\phi\rangle_{AB} = \sqrt{\frac{2}{5}}|11\rangle_{AB} + \sqrt{\frac{2}{5}}|22\rangle_{AB} + \sqrt{\frac{1}{10}}|33\rangle_{AB} + \sqrt{\frac{1}{10}}|44\rangle_{AB} \quad (1.33)$$

into the state

$$|\phi'\rangle_{AB} = \sqrt{\frac{1}{2}}|11\rangle_{AB} + \sqrt{\frac{1}{4}}|22\rangle_{AB} + \sqrt{\frac{1}{4}}|33\rangle_{AB} \quad (1.34)$$

via LOCC. These states do not satisfy (1.31) and (1.32) indicates that the maximum probability of success is $\frac{4}{5}$. However, if they also share the state $|\psi\rangle_{A'B'} = \sqrt{\frac{3}{5}}|11\rangle_{A'B'} + \sqrt{\frac{2}{5}}|22\rangle_{A'B'}$, then the states $|\phi\rangle_{AB} \otimes |\psi\rangle_{A'B'}$ and $|\phi'\rangle_{AB} \otimes |\psi\rangle_{A'B'}$ do satisfy (1.31) and the conversion may be performed with certainty.

Axiomatic approaches to entanglement

So far, only measures of entanglement that have an operational interpretation in terms of transformations between states that can be achieved via LOCC have been considered. However, one can also take a more abstract approach and construct systems of axioms that entanglement measures should obey [88, 49, 96]. Then, one can construct general functions that satisfy the axioms and use them as entanglement measures. Probably the most important of these axioms is that entanglement measures must be monotonically decreasing under LOCC.

Such an approach is most useful for quantifying mixed state entanglement, for which there are fewer generally applicable results about the operational measures of entanglement than for pure states. However, there are some benefits to taking this approach in the pure state case as well. In particular, all

entanglement measures on pure two-qubit states, satisfying monotonicity under LOCC, must be monotonic functions of one another since, for this case, (1.31) is a total order on the states. This is a useful property because some of the more general entanglement measures are easier to work with analytically and properties of the operational measures can then be derived using monotonicity. This technique is used in chapter 3, making use of the following two entanglement measures. The linearized entropy of entanglement, defined by

$$R(|\phi\rangle_{AB}) = 1 - \text{Tr}(\rho_A^2) = 1 - \sum_j p_j^2 \quad (1.35)$$

and for two-qubit states the concurrence [101], defined by

$$C(|\phi\rangle_{AB}) = \left| \langle \phi |_{AB} \sigma_2^{(A)} \otimes \sigma_2^{(B)} | \phi^* \rangle_{AB} \right| = 4p(1-p) \quad (1.36)$$

where $p_1 = p$ and $p_2 = 1 - p$. The concurrence is related to the entropy of entanglement by

$$E(C) = h\left(\frac{1}{2}\left(1 + \sqrt{1 - C^2}\right)\right) \quad (1.37)$$

where $h(x) = -x \log_2 x - (1 - x) \log_2(1 - x)$ is the Shannon entropy. This is a convex function of C (i.e. upwardly curving).

1.3.2 Bipartite mixed state entanglement

For mixed states, one can define E_D and E_C in a similar way⁶. However, no general formula has been found for them, so it is useful to consider alternative entanglement measures.

Entanglement of formation

The entanglement of formation is defined to be

$$E_F(\rho) = \min_{\{|\psi_j\rangle\}} \sum_j p_j E(|\psi_j\rangle) = \min_{\{|\psi_j\rangle\}} \langle E \rangle \quad (1.38)$$

⁶However, there are a few subtleties concerning protocols which do not distill pure Bell states exactly, but only approach them in the asymptotic limit [11, 79].

where the minimum is taken over all ensembles $\{p_j, |\psi_j\rangle\}$ such that $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$. A decomposition of ρ that achieves this minimum is called an optimal decomposition of ρ . An ensemble of states with density operator ρ could be prepared with $E_F(\rho)$ singlets and LOCC by preparing the optimal state $|\psi_j\rangle$ with probability p_j and then discarding the information about which state was prepared. It is not known if multiple copies of ρ could be prepared with less than E_F singlets by making use of entanglement between the copies, i.e. it is not known if $E_F = E_C$ ⁷. Nevertheless, E_F is still a useful entanglement measure and it is used in chapter 3. An important property of E_F that is used in that chapter is its convexity. That is, for any set of density matrices ρ_j and probabilities p_j , the following holds.

$$E_F \left(\sum_j p_j \rho_j \right) \leq \sum_j p_j E_F (\rho_j) \quad (1.39)$$

This is a simple consequence of the minimization in the definition (1.38)

The general formula for E_F is only known for the case of two-qubits [101] and the procedure to calculate it in this case is as follows. Firstly, the concurrence for a mixed state is defined in a similar way to the entanglement of formation, i.e.

$$C(\rho) = \min_{|\psi_j\rangle} \sum_j p_j C(|\psi_j\rangle) = \min_{|\psi_j\rangle} \langle C \rangle \quad (1.40)$$

From the convexity of (1.37) it follows that for any particular decomposition $\langle E \rangle \geq E(\langle C \rangle)$. Thus, the lower bound $E_F(\rho) \geq E(C(\rho))$ is obtained. Note that this bound would be an equality if there is a decomposition that minimizes $\langle C \rangle$ in which each state $|\psi_j\rangle$ has the same value of $C(|\psi_j\rangle)$. The formula for E_F given in [101] is obtained by constructing such a decomposition.

The result can be conveniently summarized by defining the spin-flipped density operator

$$\tilde{\rho} = \sigma_2 \otimes \sigma_2 \rho^T \sigma_2 \otimes \sigma_2 \quad (1.41)$$

⁷However, E_C can be defined as the regularized version of the entanglement of formation $E_C(\rho_{AB}) = \lim_{n \rightarrow \infty} (\frac{1}{n} E_F(\rho_{AB}^{\otimes n}))$.

and denoting the eigenvalues of $\tilde{\rho}\rho$ by λ_j , ordered such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. The concurrence for two-qubit mixed states is found to be

$$C(\rho) = \max \left\{ 0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4} \right\} \quad (1.42)$$

and the existence of a decomposition in which each state has this concurrence implies that the entanglement of formation is given by

$$E_F(\rho) = h \left(\frac{1}{2} \left(1 + \sqrt{1 - C^2} \right) \right) \quad (1.43)$$

Note that there are further interesting features of mixed state entanglement, which have not been reviewed here, such as the existence of *bound entangled* states, from which no Bell states can be distilled despite the fact that they are entangled [48].

1.3.3 Multi-party entanglement

Currently, there is no unique way of quantifying entanglement for $n > 2$ parties. Even for pure states of 3 qubits there is no single state, like $|\phi^+\rangle_{AB}$ for pure bipartite states, into which all states can be converted reversibly in the asymptotic limit. Indeed, for each n , there exists genuine n -party entanglement which is not reversibly convertible into entangled states shared by $< n$ parties [69]. For 3-qubits, it is known that more than one type of genuine three party entanglement exists [36, 2].

Despite these complications, some of the concepts from bipartite entanglement can be usefully generalized. Particularly important is the idea that locally invariant quantities, like the Schmidt coefficients of bipartite states, can be used to characterize entanglement. Invariants under Local Unitary (LU) and more general transformations, such as Stochastic Local Operations and Classical Communication (SLOCC), have been extensively studied in this context [68, 23, 70, 86, 24, 44, 89, 90, 91, 54, 87, 53]. These will be described in more detail in chapter 4, where some methods of measuring them are presented.

Multi-party entanglement is of particular interest in quantum computing, which involves entangled states of large numbers of qubits. It has been shown [58], at least in pure state models, that multi-party entanglement is necessary for quantum algorithms to exhibit exponential speedup over their classical counterparts. In other words, an efficient simulation of a quantum algorithm is possible on a classical computer if this entanglement is not present at some point during the execution of the algorithm. However, the same paper also shows that this is not a sufficient condition. This provides one motivation for studying the entanglement properties of quantum dynamics, since the operations involved in a quantum algorithm play a role that is perhaps more fundamental than the states.

1.4 Dynamics as an Information Resource

The quantification of entanglement presented so far is appropriate when discussing quantum communication protocols, such as teleportation or cryptography, where the local manipulation of quantum states is the primary focus of interest. However, in quantum computing and quantum control theory, the dynamics of a quantum system, as described by Hamiltonians, unitary operators and measurements, is the central focus of interest. There are essentially three areas in which quantum dynamics can be considered as an information resource.

- *Computation* - Quantum gates, such as the *controlled-not (CNOT)* gate defined in the next section, can be used to perform the steps of a quantum algorithm. The first question that arises in this area is to establish which sets of gates are universal for quantum computation, i.e. which sets of gates can be combined to generate an arbitrary unitary operation. One example of such a universal set is the CNOT and all single qubit unitaries [5]. Once a universal set is established, the difficulty of generating a particular unitary operation may be quantified by asking how many gates

from the universal set are needed to generate it. To address issues of computational complexity it is sufficient to establish whether a uniform family of unitaries representing a quantum algorithm with input size n can be generated with a number of gates that is polynomial in n or not. However, for the practical purposes of building a quantum computer it is also relevant to establish the exact number of universal gates required to generate a particular gate to within a given tolerance of error.

- *Communication* - Quantum dynamics can also be used to perform communication tasks. For example, if Alice and Bob have the ability to generate a particular interaction Hamiltonian or to perform a quantum gate or measurement on their joint system, then this might be used to generate entanglement. This in turn might be used for teleportation or superdense coding. Dynamical resources can also be used to transmit classical communication directly or to simulate the action of some other quantum dynamics via LOCC. Also, generating a particular quantum evolution can itself be regarded as a communication task that can be achieved using other types of resource. For example, it is possible to generate a quantum gate acting on an unknown state of two separated parties using only entanglement and LOCC. Some examples of all these types of protocol are given in §1.5.
- *Statistics* - The final way in which quantum dynamics can be regarded as a resource is in statistics, particularly via the role of measurements in estimation theory. For example, suppose there is a source which produces quantum states ρ which depend on an unknown parameter θ . The first question that arises is to establish which measurements are sufficient to infer the value of θ . When the unknown parameters are all the coefficients of the state, this is known as quantum state tomography [85, 64, 32, 42], which is a well established field of research. One might then try to establish the minimum number of measurements and copies of the state

required to determine θ to a given accuracy. This is known as quantum parameter estimation [6]. A relatively new area of this research is to look at these questions when θ is a parameter related to the entanglement properties of ρ . In particular, chapter 4 is concerned with these questions for the local invariants of multi-party quantum states.

These three areas are all closely linked to one another. In particular, some communication problems can be used to establish lower bounds on computational problems. For example, the minimum number of CNOTs required to generate a unitary by LOCC places a lower bound on the number of CNOTs needed to generate the gate using the universal set of the CNOT and all single qubit unitaries. This is because the single qubit gates are a subset of the operations allowed in LOCC. Also, in parameter estimation, restricting the communication resources available will generally affect the quality of estimates that can be achieved. For example, the measurements used might be restricted to those that can be implemented by LOCC between the multiple copies of the state. Clearly then, any theory that attempts to classify and quantify dynamics as a physical resource must encompass all three of these areas.

1.5 Examples of using Quantum Dynamics as a Communication Resource

Of the three areas discussed in the previous section, the use of dynamics as a communication resource is the most recent development. To illustrate this, some examples of the use of quantum gates as a communication resource are given and the analogous ideas for quantum measurements are briefly discussed. The gate examples were first given in [29, 40].

1.5.1 Unitary Operators

Example: The CNOT gate

Consider the CNOT gate, which acts as follows on the computational basis.

$$\begin{aligned}
 |00\rangle_{AB} &\rightarrow |00\rangle_{AB} \\
 |01\rangle_{AB} &\rightarrow |01\rangle_{AB} \\
 |10\rangle_{AB} &\rightarrow |11\rangle_{AB} \\
 |11\rangle_{AB} &\rightarrow |10\rangle_{AB}
 \end{aligned} \tag{1.44}$$

This gate performs a bit-flip on Bob's qubit if the state of Alice's qubit is $|1\rangle_A$ and does nothing to Bob's qubit if Alice's qubit is in the state $|0\rangle_A$. Alice's qubit is called the control and Bob's is the target and this will sometimes be shown by denoting the gate as $CNOT_{A \rightarrow B}$ when the distinction between the control and target qubits is important. It is possible to generate 1 ebit of entanglement with a CNOT by acting on the product state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_A \otimes |0\rangle_B$ to obtain $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{AB} = |\phi^+\rangle_{AB}$.

It is also possible to use a CNOT to send 1 cbit from Alice to Bob and 1 cbit from Bob to Alice simultaneously. To do this, Alice and Bob start with the Bell state $|\phi^+\rangle_{AB}$. If Alice wishes to encode a 0 then she does nothing and if she wishes to encode a 1 then she applies σ_1 to her qubit. Similarly, if Bob wishes to encode a 0 he does nothing and if he wishes to encode a 1 he applies σ_3 to his qubit. They then apply the CNOT operation. The states corresponding to each message before and after applying the CNOT are shown in table 1.1. In each case, the final state is a product. Alice's state depends only on Bob's message and her state when Bob's message is 0 is orthogonal to her state when Bob's message is 1. Bob's state depends on Alice's message in a similar way. Therefore, they can both determine each other's messages with certainty via a projective measurement. Specifically, Alice measures the observable σ_1 on her qubit and Bob measures the observable σ_3 on his.

Conversely, one may ask how much entanglement and classical communication is required to implement a CNOT. This can be done by using 1 ebit

Messages		State prepared	State after CNOT
Alice	Bob		
0	0	$ \phi^+\rangle_{AB}$	$\frac{1}{\sqrt{2}}(0\rangle + 1\rangle)_A \otimes 0\rangle_B$
0	1	$ \phi^-\rangle_{AB}$	$\frac{1}{\sqrt{2}}(0\rangle - 1\rangle)_A \otimes 0\rangle_B$
1	0	$ \psi^+\rangle_{AB}$	$\frac{1}{\sqrt{2}}(0\rangle + 1\rangle)_A \otimes 1\rangle_B$
1	1	$ \psi^-\rangle_{AB}$	$\frac{1}{\sqrt{2}}(0\rangle - 1\rangle)_A \otimes 1\rangle_B$

Table 1.1: States involved in the protocol for using a CNOT to communicate 1 cbit in each direction.

and 1 cbit of classical communication in each direction. To see this, note that implementing a CNOT requires the transformation of an arbitrary unknown two-qubit state, $|\psi\rangle_{AB} = (\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle)_{AB}$ to the state $\text{CNOT}_{A \rightarrow B}|\psi\rangle_{AB} = (\alpha|00\rangle + \beta|01\rangle + \gamma|11\rangle + \delta|10\rangle)_{AB}$. Suppose Alice and Bob also share the Bell state $|\phi^+\rangle_{A'B'}$. The first step is for Alice to perform a parity measurement on her qubits A and A' , defined by the projectors $\Pi_+ = |00\rangle_{AA'}\langle 00|_{AA'} + |11\rangle_{AA'}\langle 11|_{AA'}$, $\Pi_- = |01\rangle_{AA'}\langle 01|_{AA'} + |10\rangle_{AA'}\langle 10|_{AA'}$. If she obtains the $+$ outcome the resulting state will be

$$(\alpha|0000\rangle + \beta|0010\rangle + \gamma|1101\rangle + \delta|1111\rangle)_{AA'BB'} \quad (1.45)$$

and if she obtains the $-$ outcome the state will be

$$(\alpha|0101\rangle + \beta|0111\rangle + \gamma|1000\rangle + \delta|1010\rangle)_{AA'BB'} \quad (1.46)$$

In either case, Alice can disentangle the qubit A' from the rest of the system by performing a local $\text{CNOT}_{A \rightarrow A'}$ operation and then this qubit can be discarded. After discarding A' , the states (1.46) and (1.45) differ only by a bit-flip of qubit B' . Therefore, Bob can correct the state obtained from the $-$ outcome to that of the $+$ outcome by performing a σ_1 operation on this qubit. In order for Bob to know whether he needs to do this, Alice needs to communicate 1 cbit to him to inform him of her measurement outcome. After this, the state of the remaining three qubits will be

$$(\alpha|000\rangle + \beta|010\rangle + \gamma|101\rangle + \delta|111\rangle)_{ABB'} \quad (1.47)$$

1.5. Examples of using Quantum Dynamics as a Communication Resource

The next step is for Bob to perform a local $CNOT_{B' \rightarrow B}$ operation, which gives

$$(\alpha |000\rangle + \beta |010\rangle + \gamma |111\rangle + \delta |101\rangle)_{ABB'} \quad (1.48)$$

Bob then performs a measurement on his qubit B' in the basis

$\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_{B'}, \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)_{B'}\}$. This results in one of the following two states

$$(\alpha |00\rangle + \beta |01\rangle + \gamma |11\rangle + \delta |10\rangle)_{AB} \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_{B'} \quad (1.49)$$

$$(\alpha |00\rangle + \beta |01\rangle - \gamma |11\rangle - \delta |10\rangle)_{AB} \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)_{B'} \quad (1.50)$$

In both cases, the qubit B' is disentangled from the rest of the system and may be discarded. In the first case, the CNOT has been performed as desired, but in the second case Alice needs to apply σ_3 to her qubit A . In order for her to know whether to do this or not, Bob has to transmit 1 cbit to Alice.

Taken together, these three protocols show that a CNOT can be generated with the same resources as it can be used to generate, namely 1 ebit and 1 cbit in each direction. This implies that a CNOT cannot generate more entanglement and classical communication resources than these and that it cannot be generated with less.

To see this, suppose that a CNOT could generate more than 1 ebit. Then, it would be possible to use an ebit to generate a CNOT via LOCC and use the resulting CNOT to generate more than 1 ebit of entanglement. However, entanglement cannot increase under LOCC, so this is not possible.

Similarly, a CNOT operation cannot be used to transmit more than 1 cbit in each direction because this would allow faster than light communication, violating causality. To see this, suppose the CNOT could be used to transmit more than 1 cbit in each direction. Alice and Bob can implement a protocol to generate a CNOT, but instead of transmitting the required classical information they simply guess each other's message. Since they both have to guess a cbit, this will succeed with probability one quarter. They can then use the imperfect, but instantaneous CNOT to communicate more than 1 cbit in

each direction. More than 2 cbits communicated correctly with probability one quarter represents a non-zero amount of information. Thus Alice and Bob have managed to communicate some information to one another instantaneously.

This sort of reversibility argument can be applied to any quantum operation when it has been shown that it can generate the same amount of entanglement and classical communication resources as it takes to generate the operation. For the remaining examples, the optimality of the protocols discussed will be implied by demonstrating this reversibility.

The interaction of the CNOT with the other communication resources in the above protocols can be conveniently summarized by the following relations.

$$\begin{aligned}
 1 \text{ CNOT} & \Rightarrow 1 \text{ ebit} \\
 1 \text{ CNOT} + 1 \text{ ebit} & \Rightarrow 1 \text{ cbit}_{A \rightarrow B} + 1 \text{ cbit}_{B \rightarrow A} \\
 1 \text{ ebit} + 1 \text{ cbit}_{A \rightarrow B} + 1 \text{ cbit}_{B \rightarrow A} & \Rightarrow 1 \text{ CNOT}
 \end{aligned} \tag{1.51}$$

Note that although a CNOT can be used to generate either an ebit or 1 cbit in each direction, it does not appear to be possible to generate both simultaneously. Also, the second line shows that an ebit is consumed in the classical communication protocol. Thus, although the protocols are reversible if the entanglement or classical communication resources are counted on their own, they are not reversible with respect to the total amount of resources generated and consumed.

Example: The SWAP gate

As a second example, consider the SWAP gate, which acts on product states as follows.

$$\text{SWAP}_{AB} |\psi\rangle_A \otimes |\phi\rangle_B = |\phi\rangle_A \otimes |\psi\rangle_B \tag{1.52}$$

The SWAP gate can be used to generate 2 ebits of entanglement if Alice and Bob are allowed to hold ancillary qubits in addition to the qubits that the SWAP acts on. Let the ancillary qubits be labelled A' and B' respectively. Suppose Alice and Bob prepare the state $|\phi^+\rangle_{AA'} \otimes |\phi^+\rangle_{BB'}$, or equivalently

1.5. Examples of using Quantum Dynamics as a Communication Resource

any product of locally maximally entangled states. This state contains no entanglement between Alice and Bob. After the application of the SWAP to qubits A and B the state becomes $|\phi^+\rangle_{AB'} \otimes |\phi^+\rangle_{BA'}$, which contains 2 ebits.

The SWAP gate can also be used to send 2 cbits in each direction simultaneously. Alice and Bob start with the state $|\phi^+\rangle_{AB'} \otimes |\phi^+\rangle_{BA'}$. Alice then encodes her message in qubit A and Bob encodes his in qubit B using the same encoding as in the superdense coding protocol. Applying a SWAP operation then acts like a bi-directional quantum channel, sending the qubit A to Bob and the qubit B to Alice. They can then discover each other's message using the decoding part of the superdense coding protocol.

Conversely, the SWAP, or indeed any two-qubit operation, can be implemented with 2 ebits and 2 cbits in each direction. To do this, Alice teleports her qubit A to Bob, Bob implements the operation locally and then teleports the qubit back to Alice. This also implies that no two-qubit operation can generate more than 2 ebits or transmit more than 2 cbits in each direction, and hence the SWAP operation is optimal in this sense.

The SWAP example also shows that ancillary qubits may be required for optimal entanglement generation and classical communication. Indeed, when there are no ancillas the entanglement in a state is always invariant under the SWAP operation. In chapter 3, it will be shown that the presence of ancillas typically increases the amount of entanglement that can be generated from an arbitrary two-qubit gate.

The results for the SWAP operation can be summarized by the relations

$$\begin{aligned}
 1 \text{ SWAP} & \Rightarrow 2 \text{ ebits} \\
 1 \text{ SWAP} + 2 \text{ cbits}_{A \rightarrow B} & \Rightarrow 2 \text{ cbits}_{A \rightarrow B} + 2 \text{ cbits}_{B \rightarrow A} \quad (1.53) \\
 2 \text{ ebits} + 2 \text{ cbits}_{A \rightarrow B} + 2 \text{ cbits}_{B \rightarrow A} & \Rightarrow 1 \text{ SWAP}
 \end{aligned}$$

Comparing the results for the CNOT and SWAP gates shows that 2 CNOTs are equivalent to a SWAP under LOCC, since both can be reversibly converted into 2 ebits of entanglement.

1.5.2 Measurements

As well as unitary gates, one can also perform communication protocols using quantum measurements as a resource. Some authors have looked at the amount of entanglement required to implement measurements in different scenarios [34, 57, 81]. The entanglement and classical communication properties of measurements are slightly more subtle than those of unitary gates, since they typically depend not only on the POVM $\{E_j\}$ to be implemented, but also on the measurement operators M_j (where $E_j = M_j^\dagger M_j$) used to realize it. They also depend on how the classical information obtained about the measurement outcomes is distributed amongst the parties. For example, all two-qubit measurements can be implemented with 2 ebits and 2 cbits in each direction by Alice teleporting her half of the state to be measured to Bob, who performs the measurement locally and then teleports Alice's state back to her. However, in this scheme, only Bob learns the outcome of the measurement and further classical communication would be required if Alice is to know the outcome as well.

Nevertheless, examples of converting measurements to entanglement and classical communication and back again can be developed along similar lines to the unitary gate examples discussed above. In currently ongoing work, I have found that a measurement in the Bell basis has very similar properties to the SWAP gate and a measurement of parity has similar properties to the CNOT gate in this regard. This work is currently unpublished, but I believe that similar results can be derived for a large class of other measurements as well.

1.6 Overview

The remainder of this thesis is structured as follows.

In chapter 2, I discuss the problem of *Hamiltonian simulation*. Given an interaction Hamiltonian H , which other Hamiltonians may be simulated effi-

ciently by interspersing evolutions under H by local operations? The results for two-qubit Hamiltonians are conveniently summarized by a simple partial order relation on the space of Hamiltonians, similar to the majorization relation (1.31) for states. This work has been published in [10].

Chapter 3, is devoted to the entanglement generating capabilities of quantum operations. It generalizes some of the notions in §1.5.1 to arbitrary two-qubit operations. I define the *entangling capacity* of a quantum operation as the maximum amount of entanglement per operation that can be generated by repeated application of the operation on an arbitrary input state and LOCC. Analytical results are presented for the case of a single application of a two-qubit unitary with the additional restriction that no ancillas are allowed in the input state. Numerical results are presented for the case where ancillas are allowed. Finally, I show that collective processing of many copies of a quantum operation is not necessary to achieve the entangling capacity. The results of this chapter have been published in [62].

The focus of chapter 4 is to find ways of determining the entanglement properties of quantum states directly by measuring multiple copies of an unknown quantum state. In particular, I consider the multi-party case and determine networks for measuring the invariants of quantum states under LU and SLOCC transformations. I also analyze the statistical efficiency of these networks when only a finite number of copies of the unknown state are available. The results of this chapter have been presented in [63].

Finally, chapter 5 concludes with some open questions and a general discussion of the possible applications of the results presented in this thesis.

Chapter 2

Hamiltonian Simulation

2.1 Introduction

One of the central problems in quantifying the entanglement properties of a quantum interaction is to determine which other interactions it can be used to simulate under LOCC. In many physical implementations of quantum computing, interactions between qubits proceed by some Hamiltonian, which is always present in the system. In addition, it is usually possible to perform local operations on the individual qubits and thus simulate the effect of a different interaction Hamiltonian. This chapter describes the most efficient protocols for doing this with 2-qubit interactions [10].

Suppose Alice and Bob each have a qubit, and the qubits interact with each other via a non-local Hamiltonian H , which is continually acting on the joint system according to the Schrödinger equation. Additionally, they are each allowed to perform local operations on their own qubit. To simplify the problem, it is assumed that the local operations can be done arbitrarily fast compared to the interaction. This assumption is justified by the fact that the coupling coefficients in the interaction terms of typical Hamiltonians used in experiments aimed at demonstrating the principles of quantum computing are small compared to the coefficients of the single-particle Hamiltonians that can be generated. If Alice and Bob were to do nothing, the evolution of the qubits

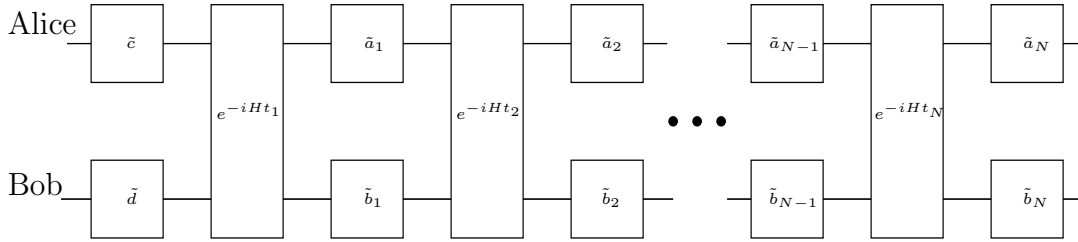


Figure 2.1: Network diagram for the two-qubit Hamiltonian simulation protocol

for time t will be given by the operator $U = e^{-iHt}$. However, they could instead allow the Hamiltonian to act for a series of shorter times t_1, t_2, \dots, t_N such that $\sum_{j=1}^N t_j = t$ interspersed with fast local unitary operations. This leads to an evolution that would have occurred if there were a different interaction Hamiltonian H' present acting for a time t' . We say that the protocol simulates the action of H' running for t' with efficiency t'/t .

If the protocol is required to work with the same efficiency for all possible values of t , then not only must it result in $e^{-iH't'}$, but it must simulate the entire dynamics that would have occurred if H' had been present, since it must work for arbitrarily small t . This is called *dynamics* or *time independent simulation* and this is the case dealt with here. Note that this is different to *gate* or *finite time simulation*, where H is used to generate $e^{-iH't'}$ just for a particular value of t' . Since this work has been published, optimal protocols for two qubit gate simulation have also been found [59, 95].

2.2 The 2-qubit simulation protocol

For now, the analysis is restricted to the case where Alice and Bob are only allowed to evolve their joint system according to H , interspersed with single-qubit unitaries on their systems. In appendix 2.B it is shown that this is the most general type of protocol needed to achieve the optimal efficiency if the object is to simulate H' with certainty. The evolution will be given by the

unitary operator

$$e^{-iH't'} = \left(\prod_{j=1}^N \tilde{a}_j \otimes \tilde{b}_j e^{-iHt_j} \right) \tilde{c} \otimes \tilde{d} \quad (2.1)$$

where $\tilde{a}_j, \tilde{b}_j, \tilde{c}, \tilde{d} \in U(2)$ (see fig. 2.1 for a network diagram). It is convenient to rewrite this in terms of a conjugation action as follows:

$$e^{-iH't'} = \left(\prod_{j=1}^N a_j \otimes b_j e^{-iHt_j} a_j^\dagger \otimes b_j^\dagger \right) c \otimes d = \left(\prod_{j=1}^N e^{-i[a_j \otimes b_j H a_j^\dagger \otimes b_j^\dagger] t_j} \right) c \otimes d \quad (2.2)$$

where $a_j, b_j, c, d \in U(2)$ are suitably defined in terms of $\tilde{a}_j, \tilde{b}_j, \tilde{c}$ and \tilde{d} . Specifically, a_j can be defined recursively by $a_1 = \tilde{a}_1$, $a_{j+1} = a_j \tilde{c}_j \tilde{a}_{j+1}$ and then $c = a_N \tilde{c}_N$. Similarly for b_j and d .

The most general 2-qubit Hamiltonian is given by

$$H = \gamma I_2 \otimes I_2 + \vec{\alpha} \cdot \vec{\sigma} \otimes I_2 + I_2 \otimes \vec{\beta} \cdot \vec{\sigma} + \sum_{j,k=1}^3 R_{jk} \sigma_j \otimes \sigma_k \quad (2.3)$$

where γ, R_{jk} are real numbers and $\vec{\alpha}, \vec{\beta}$ are real 3-dimensional vectors. In dynamics simulation, the protocol must work for arbitrarily small t . Therefore, let δt be a short evolution time. Then, the Baker-Campbell-Hausdorff formula gives

$$e^{-i(H\delta t + O(\delta t^2))} = e^{-i\gamma\delta t} \left(e^{-i\vec{\alpha} \cdot \vec{\sigma} \delta t} \otimes e^{-i\vec{\beta} \cdot \vec{\sigma} \delta t} \right) e^{-i \sum_{j,k=1}^3 R_{jk} \sigma_j \otimes \sigma_k \delta t} \quad (2.4)$$

To first order in δt , the $\vec{\alpha}$ and $\vec{\beta}$ terms generate a purely local evolution. The γ term generates an global phase, which is physically irrelevant. Thus, any interaction is due to the R_{jk} term only. Therefore, by continuously interspersing the evolution due to H with local unitary transformations, the $\vec{\alpha}, \vec{\beta}$ and γ terms may be removed. Note that any local unitary evolution may be simulated by choosing $c \otimes d$ appropriately. Thus, the local terms and the $c \otimes d$ operation may be neglected, and only Hamiltonians of the form $\sum_{j,k} R_{jk} \sigma_j \otimes \sigma_k$ need to be considered. Expanding eq. (2.2) to first order in t yields

$$sH' = \sum_{j=1}^N p_j a_j \otimes b_j H a_j^\dagger \otimes b_j^\dagger \quad (2.5)$$

where $s = t'/t$ is the efficiency and $p_j = t_j/t$, which implies that $\sum_j p_j = 1$.

Definition 2.1. H' can be efficiently simulated by H ($H' \prec_S H$) if there exists a simulation protocol with $s = 1$. If $H' \prec_S H$ and $H \prec_S H'$ then they are equivalent ($H' \approx_S H$).

Note that it is always possible to simulate the zero Hamiltonian by choosing $p_{1,2,3,4} = \frac{1}{4}$, $a_j = \sigma_j$ for $j = 1, 2, 3$, $a_4 = I_2$ and $b_j = I_2$ for $j = 1, 2, 3, 4$. This can be generalized to bipartite Hamiltonians on arbitrary dimensional Hilbert spaces [10]. Thus, if there is a protocol with $s > 1$, then there is also a protocol with $s = 1$, since the zero Hamiltonian may be simulated for the remaining time.

Also, it is possible to simulate an arbitrary 2-qubit Hamiltonian with some simulation factor s , providing at least one of the coefficients R_{jk} is non-zero. To see this, note that for any σ_i, σ_j there exist unitary operations $U_{ij\pm}$ such that

$$U_{ij\pm}\sigma_i U_{ij\pm}^\dagger = \pm\sigma_j \quad (2.6)$$

These operations form the single-qubit *Clifford Group* [43], which is generated by the following two matrices

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad (2.7)$$

It can be assumed without loss of generality that $R_{33} > 0$, since otherwise a conjugation with $U_{33-} \otimes I_2$ may be applied. Then by choosing $p_{1,2,3,4} = \frac{1}{4}$, $a_{1,2} = I_2$, $a_{3,4} = \sigma_3$, $b_{1,3} = I_2$, $b_{2,4} = \sigma_3$, the Hamiltonian $\sigma_3 \otimes \sigma_3$ may be simulated with some simulation factor, although this is not necessarily the most efficient simulation. $\sigma_3 \otimes \sigma_3$ can be used to simulate an arbitrary product Hamiltonian of the form $\sigma_j \otimes \sigma_k$ by conjugating with Clifford Group operations. An arbitrary 2-qubit Hamiltonian may then be simulated by short evolutions under each $\sigma_j \otimes \sigma_k$ with time ratios according to R'_{jk} . The Clifford Group construction, and hence this result generalizes to higher dimensional Hilbert spaces as well.

The problem we are concerned with can now be expressed in two equivalent ways.

Let H be arbitrary. The *optimal simulation* problem, given H' , is to find a solution, $\{p_j, a_j, b_j\}$ such that s is maximal. The *efficient simulation problem* is to find all H' , such that $H' \prec_S H$

Definition 2.2. *The optimal simulation factor $s_{H|H'}$ is the maximal s such that $sH' \prec_S H$.*

The two problems are equivalent because it is always possible to simulate an arbitrary H' using any H with some efficiency. The efficient simulation problem can be solved by finding the optimal solution for all H' . Then all H' with $s_{H|H'} \leq 1$ can be efficiently simulated. The optimal simulation problem can be solved by finding the efficiently simulated sH' with maximal value of s . Thus, throughout the rest of this chapter, the solutions to both problems are discussed interchangeably.

In the following sections, necessary and sufficient conditions for $H' \prec_S H$ are derived for arbitrary H, H' , the optimal simulation factor $s_{H|H'}$ is found and the optimal simulation strategy in terms of $\{p_j, a_j, b_j\}$ is given. These results endow the set of two qubit Hamiltonians with a partial order, \prec_S , which for each H induces a convex set $\{H' : H' \prec_S H\}$. The convexity arises because if $H' \prec_S H$ and $H'' \prec_S H$ then $\lambda H' + (1 - \lambda)H''$ can be simulated with unit efficiency by time sharing the evolution according to H between the protocols for simulating H' and H'' in the ratio $\lambda : (1 - \lambda)$. This set has a simple geometric interpretation and allows the relation \prec_S to be succinctly characterized by a majorization-like relation. The geometric and majorization interpretations offer two different methods to calculate the optimal protocol and simulation factor.

2.3 Normal form for two qubit Hamiltonians

Following the discussion in the previous section, a general 2-qubit interaction Hamiltonian is taken to be of the form $K = \sum_{jk} R_{jk} \sigma_j \otimes \sigma_k$.

Definition 2.3. *The normal form of a two qubit Hamiltonian is $H = \sum_j h_j \sigma_j \otimes \sigma_j$, where $h_1 \geq h_2 \geq |h_3|$ are the singular values of the matrix R with elements R_{jk} and $h_3 = \text{sgn}(\det(R))|h_3|$*

Theorem 2.1. *Let H be the normal form of K . Then $H \approx_S K$. Furthermore, it only takes a one stage protocol to simulate one of them with the other, so the resulting evolutions are equivalent under local unitary transformations, i.e.*

$$\exists a, b \in U(2) \quad \text{s.t.} \quad e^{-iH} = a \otimes b e^{-iK} a^\dagger \otimes b^\dagger \quad (2.8)$$

Proof. Conjugating the evolution resulting from K acting for a time t by a local unitary $a \otimes b$ gives

$$e^{-iK't} = a \otimes b e^{-iKt} a^\dagger \otimes b^\dagger = e^{-i(a \otimes b)K(a^\dagger \otimes b^\dagger)t} \quad (2.9)$$

where

$$\begin{aligned} K' &= (a \otimes b)K(a^\dagger \otimes b^\dagger) \\ &= \sum_{jk} R_{jk} (a \sigma_j a^\dagger) \otimes (b \sigma_k b^\dagger) \\ &= \sum_{jk} R_{jk} (\sum_l P_{jl} \sigma_l) \otimes (\sum_m Q_{km} \sigma_m) \\ &= \sum_{lm} (P^T R Q)_{lm} \sigma_l \otimes \sigma_m \equiv \sum_{lm} R'_{lm} \sigma_l \otimes \sigma_m \end{aligned}$$

and $P, Q \in SO(3)$, since conjugating $\vec{r} \cdot \vec{\sigma}$ by an $SU(2)$ matrix is equivalent to rotating the vector \vec{r} by a matrix in $SO(3)$. Thus it remains to show that it is always possible to choose P and Q , such that $K' = H$.

Let $R = O_1 D O_2$ be a singular value decomposition of R , where $O_1, O_2 \in O(3)$, and $D = \text{diag}(h_1, h_2, |h_3|)$ is the diagonal matrix whose entries are the

singular values of R . Then

$$\begin{aligned}
 R &= O_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(O_1) \end{pmatrix} \times D \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(O_1)\det(O_2) \end{pmatrix} \\
 &\quad \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(O_2) \end{pmatrix} O_2 \\
 &= \tilde{O}_1 \times \begin{pmatrix} h_1 & 0 & 0 \\ 0 & h_2 & 0 \\ 0 & 0 & |h_3|\text{sgn}(\det(R)) \end{pmatrix} \times \tilde{O}_2
 \end{aligned}$$

Note that $\tilde{O}_1, \tilde{O}_2 \in SO(3)$ for all $O_1, O_2 \in O(3)$. Thus, choosing $P = \tilde{O}_1^T, Q = \tilde{O}_2$ gives the required decomposition. \square

Now, some notations suggested by the above argument can be introduced.

Definition 2.4. *The Pauli representation of K is the 3×3 real matrix R . D_K denotes the Pauli representation of the normal form of K .*

2.4 Simulation of normal form two-qubit Hamiltonians

Due to theorem 2.1, only simulations that take the normal form of H to the normal form of H' need to be considered. Thus, eq. (2.5) can be re-expressed as

$$sD_{H'} = \sum_j p_j P_j D_H Q_j \quad (2.10)$$

where $P_j, Q_j \in SO(3)$. Since H and H' are in their normal form, $h_1 \geq h_2 \geq |h_3|$ and $h'_1 \geq h'_2 \geq |h'_3|$. Without loss of generality, two further assumptions can be made. Firstly, assume $h_3 \geq 0$, since if $h_3 < 0$ then eq. (2.10) can be right multiplied by $S = \text{diag}(1, 1, -1)$

$$s(D'_H S) = \sum_j p_j P_j (D_H S) (S Q_j S) \quad (2.11)$$

where $SQ_jS \in SO(3)$ and $D_H S = \text{diag}(h_1, h_2, |h_3|)$ is of the desired form. Secondly, note that $s_{H|H'} = \frac{1}{a} s_{H'|aH} = a s_{aH'|H}$. The protocol is unchanged if eq. (2.10) is divided by $\text{Tr}(D_H)$. Thus, the normalization $\text{Tr}(D_H) = 1$ can be assumed.

The $D_{H'}$ that can be efficiently simulated are precisely the diagonal subset of the convex hull of the set $\{PD_HQ : P, Q \in SO(3)\}$. This convex diagonal subset will be denoted \mathcal{C}_H . $D_{H'} = 0 \in \mathcal{C}_H$ because it was noted earlier that the zero Hamiltonian can always be simulated. Indeed, $D_{H'} = 0$ is an interior point of \mathcal{C}_H because it was noted in §2.2 that it is always possible to simulate any non-zero H' with some efficiency s . It follows that the optimal solution $\forall D_{H'} \neq 0$ is a boundary point of \mathcal{C}_H . Thus, the problem of optimal or efficient simulation can be rephrased as follows:

Given H, H' can be efficiently simulated iff $D_{H'} \in \mathcal{C}_H$. The optimal simulation $s_{H|H'} D_{H'}$ is the unique intersection of the half-line $\lambda D_{H'}, (\lambda \geq 0)$ with the boundary of \mathcal{C}_H . An optimal protocol is then obtained by decomposing $s_{H|H'} D_{H'}$ in terms of the extreme points of \mathcal{C}_H .

In fact, it suffices to consider the convex hull \mathcal{P}_H of a subset $P_{24} \in \mathcal{C}_H$. The 24 elements of P_{24} are obtained from D_H by permuting the diagonal elements and putting an even number of - signs. The proof that $\mathcal{C}_H = \mathcal{P}_H$ is given in appendix 2.A, since it depends of the characterization of \mathcal{P}_H given in the next

section. The elements of P_{24} are explicitly given by $\pi_j D_H \pi_j s_k$, where

$$\begin{aligned}
 \pi_0 = I_3, \pi_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \pi_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \pi_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}, \\
 \pi_4 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \pi_5 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \\
 s_0 = I_3, s_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, s_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, s_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
 \end{aligned} \tag{2.12}$$

In the next section, the geometry of \mathcal{P}_H is investigated and optimal protocols for simulating any H' are found. Then, in §2.5, the solution is restated in terms of a majorization-like relation.

2.4.1 Optimization over \mathcal{P}_H

Since all the matrices in P_{24} and \mathcal{P}_H are diagonal, their elements can be represented by real 3-dimensional vectors. \mathcal{P}_H is a polytope with 24 (not necessarily distinct) vertices that are elements of P_{24} . By plotting the vertices in P_{24} , the equations of the boundary faces can be determined, and these give a useful characterization of \mathcal{P}_H in terms of a set of inequalities. First consider the simple case $(h_1, h_2, h_3) = (1, 0, 0)$, for which there are 6 distinct vertices: $(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)$ and \mathcal{P}_H is an octahedron, as shown in fig. (2.2)¹. There are a few other simple cases, such as $h_3 = 0$ and $h_1 = h_2$, for which the polytope can be constructed in the same way, but the most complicated case is the generic case, $h_1 > h_2 > h_3 > 0$, as shown in fig. (2.3).

As in fig. (2.2), fig. (2.3) is viewed from the direction $(1, 1, 1)$. Three faces are removed to show the structure in the back. There are 3 types of faces. There are 6 identical rectangular purple faces on the planes $x = \pm h_1, y =$

¹Diagrams in this section are due to D. Leung.

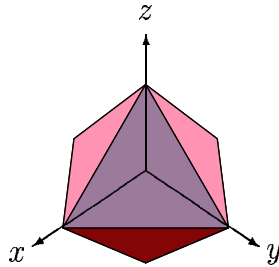


Figure 2.2: \mathcal{P}_H for the case $(h_1, h_2, h_3) = (1, 0, 0)$ viewed from the direction $(1, 1, 1)$.

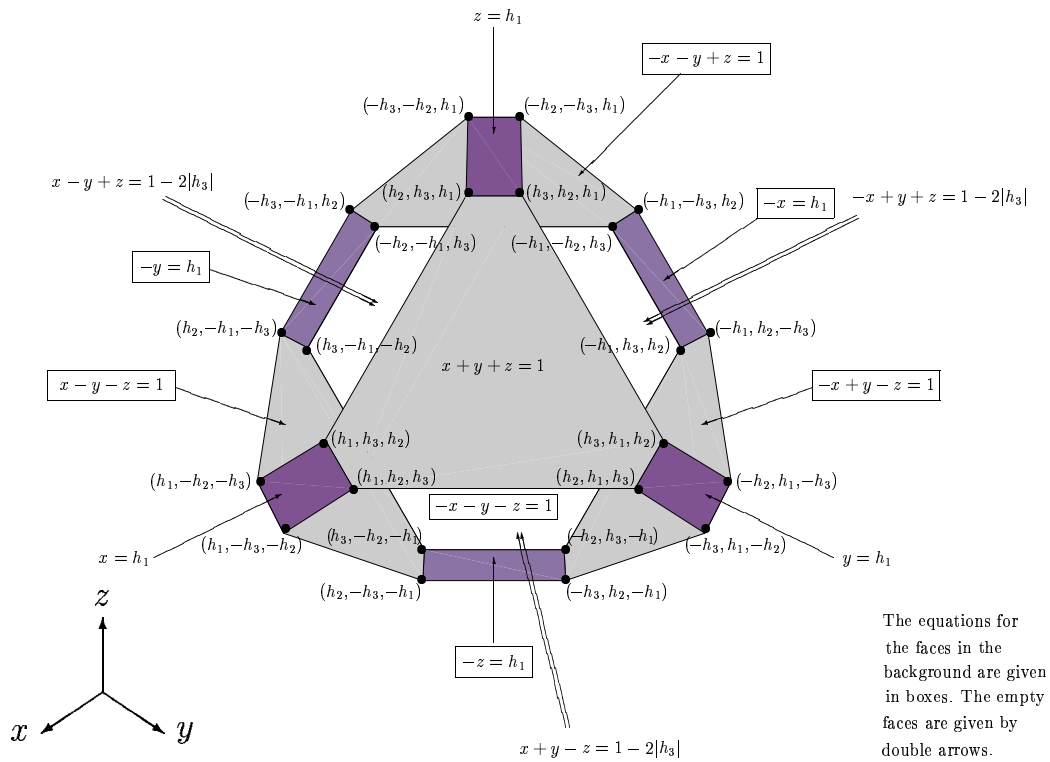


Figure 2.3: \mathcal{P}_H for the generic case $h_1 > h_2 > h_3 > 0$ viewed from the direction $(1, 1, 1)$.

$\pm h_1, z = \pm h_1$. There are two groups of 4 identical hexagonal faces. The first group of 4 consists of the 3 light blue faces in the back, and the light blue face in the front. These are the truncated faces of the original octahedron, lying on the planes $x + y + z = 1, -x + y - z = 1, -x - y + z = 1, x - y - z = 1$. The second group consists of the 3 empty faces in the front, and the white face in the back. They are *inside* the original octahedron and are parallel to the original faces. They lie on the planes $-x - y - z = 1 - 2h_3, -x + y + z = 1 - 2h_3, x - y + z = 1 - 2h_3, x + y - z = 1 - 2h_3$. Note that each hexagon in one group has a parallel counterpart in the other group. All together, there are 7 pairs of parallel faces, so \mathcal{P}_H may be characterized by the following inequalities.

$$(x, y, z) \in \mathcal{P}_H \text{ iff } \begin{cases} |x| \leq h_1, |y| \leq h_1, |z| \leq h_1 \\ -(1 - 2h_3) \leq +x + y + z \leq 1 \\ -(1 - 2h_3) \leq -x - y + z \leq 1 \\ -(1 - 2h_3) \leq +x - y - z \leq 1 \\ -(1 - 2h_3) \leq -x + y - z \leq 1 \end{cases} \quad (2.13)$$

Recall from §2.4 that the optimal simulation problem can be restated as, given D_H and $D_{H'}$, find the unique intersection of the half-line $\lambda D_{H'}$, ($\lambda > 0$) with the boundary of \mathcal{P}_H . This fact can be used to explicitly work out $s_{H'|H}$, i.e. the value of λ in the intersection, as a function of $D_{H'}$.

The intersection is given by a vector $\vec{v} = s_{H'|H}(h'_1, h'_2, h'_3)$, so that

$$s_{H'|H} = \frac{\|\vec{v}\|_1}{\|(h'_1, h'_2, h'_3)\|_1} = \frac{\|\vec{v}\|_1}{\|H'\|_1}, \quad (2.14)$$

where $\|\vec{v}\|_1$ for a vector \vec{v} is the sum of the absolute values of the entries, and $\|H'\|_1 \equiv \|(h'_1, h'_2, h'_3)\|_1$. The set \mathcal{P}_H has only 3 types of boundary faces. Therefore, there are only 3 possibilities where the intersection can occur:

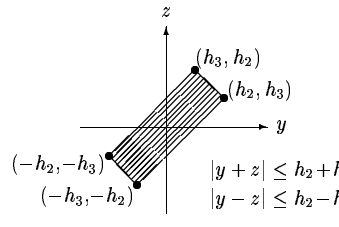
1. On the group of faces given by $x + y + z = 1, -x + y - z = 1, -x - y + z = 1, x - y - z = 1$. In this case, $\|\vec{v}\|_1 = 1$, and $s_{H'|H} = \frac{1}{\|H'\|_1}$.
2. On the group of faces $x + y - z = 1 - 2h_3, x - y + z = 1 - 2h_3, -x + y + z =$

$1 - 2h_3, -x - y - z = 1 - 2h_3$. In this case, $\|\vec{v}\|_1 = 1 - 2h_3$, and $s_{H'|H} = \frac{1-2h_3}{\|H'\|_1}$.

3. On the group of faces $x = \pm h_1, y = \pm h_1, z = \pm h_1$. In this case, $\vec{v} = \frac{h_1}{h'_1}(h'_1, h'_2, h'_3)$, and $\|\vec{v}\|_1 = \frac{h_1}{h'_1}\|H'\|_1$ is not constant on the face, and $s_{H'|H} = \frac{h_1}{h'_1}$.

Note that when H' is in a normal form, \vec{v} can only fall on one face in each case. These faces are respectively $x + y + z = 1$, $x + y - z = 1 - 2h_3$, and $x = h_1$. The (h'_1, h'_2, h'_3) belonging to each case can be characterized as follows.

- Case 1. Note that the face of \mathcal{P}_H on $x + y + z = 1$ is the convex hull of (h_1, h_2, h_3) and all permutations of the entries. The hexagon contains exactly all vectors \vec{v} majorized by (h_1, h_2, h_3) , $\vec{v} \prec (h_1, h_2, h_3)$ (see §2.5 for definition of majorization). Hence, (h'_1, h'_2, h'_3) is in case 1 if and only if it is proportional to some $\vec{v} \prec (h_1, h_2, h_3)$.
- Case 3. In this case, $\vec{v} = (h_1, \frac{h_1 h'_2}{h'_1}, \frac{h_1 h'_3}{h'_1})$ (note $h'_1/h_1 \geq 0$). Thus (h'_1, h'_2, h'_3) is in case 3 iff $(\frac{h_1 h'_2}{h'_1}, \frac{h_1 h'_3}{h'_1})$ is within the rectangle with vertices $(h_2, h_3), (h_3, h_2), (-h_2, -h_3), (-h_3, -h_2)$.



$$(2.15)$$

Hence, (h'_1, h'_2, h'_3) is of case 3 iff

$$\left| \frac{h_1 h'_2}{h'_1} + \frac{h_1 h'_3}{h'_1} \right| \leq h_2 + h_3 \quad \text{and} \quad \left| \frac{h_1 h'_2}{h'_1} - \frac{h_1 h'_3}{h'_1} \right| \leq h_2 - h_3 \quad (2.16)$$

$$\text{iff} \quad \frac{h_1}{h_2 + h_3} \leq \frac{h'_1}{h'_2 + h'_3} \quad \text{and} \quad \frac{h_1}{h_2 - h_3} \leq \frac{h'_1}{h'_2 - h'_3} \quad (2.17)$$

- Case 2. This contains all (h'_1, h'_2, h'_3) not in case 1 or 3.

The fact that the intersection is on one of the boundary faces means that it can be easily decomposed as a convex combination of at most 3 vertices in

P_{24} ². The decomposition directly translates to an optimal protocol with at most 3 types of conjugation.

2.5 *s*-majorization

The problem of Hamiltonian simulation can also be analyzed from the perspective of a majorization-like relation, which provides a natural and compact language to present the results of the previous section.

Recall the standard notions of majorization and submajorization as defined in the space of n -dimensional real vectors. Let u denote an n -dimensional real vector and u_i , $i = 1, \dots, n$, its n components. Define u^\downarrow to be a vector with components $u_1^\downarrow \geq u_2^\downarrow \geq \dots \geq u_n^\downarrow$, corresponding to the components u_i decreasingly ordered. Let v be another n -dimensional vector and v^\downarrow its ordered version. Then u is submajorized or weakly majorized by v , written $u \prec_w v$, if

$$u_1^\downarrow \leq v_1^\downarrow, \tag{2.18}$$

$$u_1^\downarrow + u_2^\downarrow \leq v_1^\downarrow + v_2^\downarrow, \tag{2.19}$$

$$\vdots \tag{2.20}$$

$$u_1^\downarrow + u_2^\downarrow + \dots + u_n^\downarrow \leq v_1^\downarrow + v_2^\downarrow + \dots + v_n^\downarrow. \tag{2.21}$$

In case of equality in the last equation, u is *majorized* by v , which is written as $u \prec v$.

These notions can be extended to real matrices by comparing the corresponding vectors of singular values. Suppose M and N are two $n \times n$ real matrices. Then, M is majorized by N , $M \prec N$, when $\text{sing}(M) \prec \text{sing}(N)$, where $\text{sing}(M)$ denotes the vector of singular values of the matrix M . Thus, majorization endows the set of real matrices with a partial order, and a notion of equivalence,

$$M \sim PMQ \quad \forall P, Q \in O(n) \tag{2.22}$$

²This is a consequence of Carathéodory's theorem which states that an interior point of a convex n -dimensional polytope can always be written as a convex combination of $\leq n + 1$ boundary points [104].

because the transformation $M \rightarrow PMQ$ preserves the singular values. A “convex sum” characterization of weak majorization

$$M \prec_w N \Leftrightarrow M = \sum_i p_i P_i N Q_i \quad (2.23)$$

also holds [17], with the meaning that N always weakly majorizes an (left and right) orthogonal mixing of itself.

Likewise, the notion of special majorization, s-majorization for short, can be introduced. In close analogy with majorization, consider again the $n \times n$ real matrices M , but this time transform them by acting on the right and on left with *special* orthogonal matrices $P, Q \in SO(n)$ with determinant $+1$. Proceeding back to front, introduce the equivalence relation

$$M \sim_s PMQ, \quad (2.24)$$

where $P, Q \in SO(n)$. Following the singular value decomposition of M , $M \sim_s \text{diag}(d_1, d_2, \dots, d_{n-1}, \text{sg}(\det M) d_n)$ where $d_1 \geq d_2 \geq \dots \geq d_n \geq 0$ are the singular values of M . This suggests the following rearrangement of a real vector u with components u_i . Let $|u|$ be the real vector with components $|u_i|$. Then, the s-ordered vector $u^{\downarrow s}$ is constructed as follows:

$$(|u|_1^{\downarrow}, |u|_2^{\downarrow}, \dots, |u|_{n-1}^{\downarrow}, \text{sg}(\prod_i u_i) |u|_n^{\downarrow}). \quad (2.25)$$

In other words, the absolute values of u_i are arranged in decreasing order, the sign of the last element is chosen to be the product of all the original signs. Then define the s-majorization relation, which will be denoted by the symbol \prec_s , directly on real matrices by means of the “convex sum” characterization,

$$M \prec_s N \Leftrightarrow M = \sum_i p_i P_i N Q_i. \quad (2.26)$$

That is, M is s-majorized by N when M is a (left and right) special orthogonal mixing of N . This definition applies for real vectors (when M and N are diagonal matrices). The remainder of this section is restricted to the $n = 3$ case, which is relevant for the two-qubit Hamiltonian simulation problem. The

results in §2.4.1 imply the following alternative definition of s-majorization which is equivalent to eq. (2.13).

Definition 2.5. *Let (u_1, u_2, u_3) and (v_1, v_2, v_3) be s-ordered. Then (u_1, u_2, u_3) is s-majorized by (v_1, v_2, v_3) , denoted $u \prec_s v$, if and only if*

$$\begin{aligned} u_1 &\leq v_1, \\ u_1 + u_2 - u_3 &\leq v_1 + v_2 - v_3, \\ u_1 + u_2 + u_3 &\leq v_1 + v_2 + v_3. \end{aligned} \tag{2.27}$$

If the real vectors u and v are not s-ordered, then $u \prec_s v$ when $u^{\downarrow s} \prec_s v^{\downarrow s}$.

Note that when applied to s-ordered vectors u and v , the s-majorization relation implies the weak majorization relation, that is $u \prec_s v \Rightarrow u \prec_w v$. This can be seen by summing the second and third inequalities in (2.27). Moreover, when $\text{sg}(\Pi_i u_i) = \text{sg}(\Pi_i v_i)$ and $\sum_i u_i = \sum_i v_i$, then \prec_w , \prec_s , and \prec are all equivalent.

Now s-majorization can be related to Hamiltonian simulation. Consider the polytope \mathcal{P}_H with the set of vertices P_{2^4} associated to $h = (h_1, h_2, h_3)$, which describes the normal form of the Hamiltonian H . The Hamiltonians H' , characterized by vectors h' , that can be efficiently simulated with H , $H' \prec_S H$, are given by just comparing h and h' according to the s-majorization relation.

Theorem 2.2. *Let h and h' be real, s-ordered 3-dimensional vectors. Then $h' \in \mathcal{P}_H$ iff $h' \prec_s h$.*

Proof. When $h = (h_1, h_2, h_3)$ is s-ordered, but not necessarily satisfying $\|h\|_1 = 1$ and $h_3 \geq 0$, the polytope \mathcal{P}_H associated can still be characterized by a simple modification of eq. (2.13):

$$D_{H'} \in \mathcal{P}_H \Leftrightarrow \begin{cases} \forall_i |h'_i| \leq h_1 \\ -(h_1 + h_2 - h_3) \leq h'_1 + h'_2 + h'_3 \leq h_1 + h_2 + h_3 \\ -(h_1 + h_2 - h_3) \leq -h'_1 - h'_2 + h'_3 \leq h_1 + h_2 + h_3 \\ -(h_1 + h_2 - h_3) \leq h'_1 - h'_2 - h'_3 \leq h_1 + h_2 + h_3 \\ -(h_1 + h_2 - h_3) \leq -h'_1 + h'_2 - h'_3 \leq h_1 + h_2 + h_3 \end{cases} \tag{2.28}$$

Furthermore, when h' is also s -ordered, the above reduces to

$$\begin{aligned} h'_1 &\leq h_1, \\ h'_1 + h'_2 - h'_3 &\leq h_1 + h_2 - h_3, \\ h'_1 + h'_2 + h'_3 &\leq h_1 + h_2 + h_3. \end{aligned} \tag{2.29}$$

which is exactly the condition for $h' \prec_s h$. \square

Theorem 2.3. *Let $H = \sum_i h_i \sigma_i \otimes \sigma_i$ and $H' = \sum_i h'_i \sigma_i \otimes \sigma_i$ be two-qubit Hamiltonians in their normal forms. Then*

$$H' \prec_S H \iff h' \prec_s h. \tag{2.30}$$

The optimal simulation factor is given by $s_{H'|H} = \max_{h' \prec_s h} s$.

2.6 Conclusions

In this chapter the optimal dynamics simulation protocols for two-qubit Hamiltonians have been derived. The results give rise to the s -majorization relation, which is a partial order on the 2-qubit Hamiltonians, similar to the partial order on entangled states given by (1.31). Here, the majorization arises in a different way, but it provides an intriguing glimpse of the possible connections between entanglement in quantum states and the non-local properties of quantum operations.

In this chapter, only simulation protocols that operate on a single copy of the system on which we are trying to perform the simulation have been considered. More generally, protocols could act in a “blockwise” fashion, by simulating $(e^{-iH't})^{\otimes n}$ using $(e^{-iHt})^{\otimes m}$, or the evolution according to H could be time-shared amongst many copies of the system. By analogy with the manipulation of entanglement in states, it seems likely that these sort of protocols could have greater efficiency and possibly also be reversible. Finally, only protocols that succeed in the simulation with certainty have been considered here, but a simulation might also be stochastic and fail with some finite probability,

in which case the expected cost of the simulation could be considered. These more general protocols would be interesting topics for future research.

2.7 Related Work

Since the publication of these results, there has been much interest in the Hamiltonian simulation problem.

Some of the results can be generalized by allowing additional resources. For example, Vidal and Cirac [94] have shown that allowing classical communication between the parties in addition to local operations does not improve the optimal simulation factor. In the same paper, they also showed that ancillas are needed to achieve the optimal simulation when each party has a Hilbert space of dimension > 2 in contrast to the result of appendix 2.B. Finally, they showed that the s-majorization relation can be alternatively formulated as an ordinary majorization relation on the eigenvalues of the Hamiltonians in their normal form. In another paper [93] they demonstrated a catalysis effect, whereby the presence of additional entanglement in the system allows certain Hamiltonians to be efficiently simulated that could not be efficiently simulated otherwise, despite the fact that the entanglement is not consumed by the protocol. This is similar to the catalysis effect in the conversion of quantum states under LOCC discussed in §1.3.1.

There are a few direct generalizations of bipartite Hamiltonian simulation to higher dimensional Hilbert spaces. Recently it has been shown by Childs et. al. [26] that any product Hamiltonian, of the form $H_A \otimes H_B$ can simulate any other Hamiltonian of this form reversibly. Necessary and sufficient conditions for efficient simulation in arbitrary dimensions have not yet been found, but Chen has found a necessary condition based on algebraic geometry [25].

Simulation has also been investigated in a number of quite different regimes. The problem of gate simulation, as opposed to the dynamics simulation discussed here, was first investigated prior to the publication of the results in

this chapter [59], but has since been solved for the two-qubit case [95] by breaking the problem down into a minimization over a discrete set of dynamics simulation problems, which can each be solved via the results presented here. Hamiltonian simulation has also been considered for multi-party systems [98, 55, 100, 99, 65, 19, 33, 75] and efficient protocols have been found for some particular instances. Simulation for systems described by continuous variables has also recently been discussed [61].

Most of the Hamiltonian simulation protocols developed so far require continuous switching of the local unitary operations. Whilst these are appropriate for finding the fundamental limitations on simulation imposed by quantum mechanics, they are far from being experimentally feasible. However, Haselgrove et. al. [46] have recently shown for the two-qubit case that the cost in simulation time of restricting the protocols to have a fixed number of steps is not too great.

2.A Proof that $\mathcal{C}_H = \mathcal{P}_H$

Recall that \mathcal{C}_H is defined as the diagonal subset of the convex hull of $\{PD_HQ : P, Q \in SO(3)\}$, where D_H is the Pauli representation of a two-qubit Hamiltonian in its normal form (i.e. $D_H = \text{diag}(h_1, h_2, h_3)$). \mathcal{P}_H is the convex hull of $P_{24} \subset \mathcal{C}_H$ obtained by permuting the diagonal elements of D_H and putting an even number of $-$ signs on the diagonal.

By definition $\mathcal{P}_H \subseteq \mathcal{C}_H$, so the result can be proved by showing that $\mathcal{C}_H \subseteq \mathcal{P}_H$. \mathcal{C}_H consists of Hamiltonians that have Pauli representations that can be expressed as $D_{H'} = \sum_j p_j P_j D_H Q_j$, where $P_j, Q_j \in SO(3)$, $p_j > 0$ and $\sum_j p_j = 1$. Since $D_{H'}$ is diagonal, only the diagonal elements of each $P_j D_H Q_j$ contribute to $D_{H'}$. Whilst it is possible for an individual $P_j D_H Q_j$ term to have off-diagonal elements, these elements have to cancel out in the sum. Thus, to show that $\mathcal{C}_H \subseteq \mathcal{P}_H$, it suffices to show that the diagonal part of each $P_j D_H Q_j \in \mathcal{P}_H$ because any $D_{H'} \in \mathcal{C}_H$ will then be in \mathcal{P}_H by convexity.

In what follows, $P_j D_H Q_j$ will be denoted as $P D_H Q$ without the j subscripts because only a single term is considered. The diagonal elements of $P D_H Q$ can be represented as a three-dimensional vector (g_1, g_2, g_3) . The proof proceeds by showing that the components of this vector satisfy the inequalities (2.13) and thus the vector belongs to \mathcal{P}_H . Since $D_H = \text{diag}(h_1, h_2, h_3)$,

$$g_i = (P D_H Q)_{ii} = \sum_j P_{ij} h_j Q_{ji} = \sum_j P_{ij} Q_{ij}^T h_j \quad (2.31)$$

The vectors (h_1, h_2, h_3) and (g_1, g_2, g_3) are linearly related by

$$\begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = P * Q^T \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \quad (2.32)$$

where $*$ denotes the entry-wise multiplication of two matrices, otherwise known as the Hadamard or Schur product. Expanding (2.31) explicitly gives

$$g_i = P_{i1} Q_{i1}^T h_1 + P_{i2} Q_{i2}^T h_2 + P_{i3} Q_{i3}^T h_3 \quad (2.33)$$

Recall that in §2.4 it was assumed that all the h_j 's are positive, so by the triangle inequality

$$|g_i| \leq |P_{i1} Q_{i1}^T| h_1 + |P_{i2} Q_{i2}^T| h_2 + |P_{i3} Q_{i3}^T| h_3 \quad (2.34)$$

Now, since $P, Q \in SO(3)$, their columns and rows are orthonormal vectors. Hence $(|P_{i1}|, |P_{i2}|, |P_{i3}|)$ and $(|Q_{i1}^T|, |Q_{i2}^T|, |Q_{i3}^T|)$ are unit vectors and their inner product $|P_{i1} Q_{i1}^T| + |P_{i2} Q_{i2}^T| + |P_{i3} Q_{i3}^T| \leq 1$. This argument, which is used frequently, will be referred to as the ‘‘inner product argument’’. The first group of inequalities (2.13) follows from (2.34) and the inner product argument, since

$$\begin{aligned} |g_i| &\leq (|P_{i1} Q_{i1}^T| + |P_{i2} Q_{i2}^T| + |P_{i3} Q_{i3}^T|) \max_i h_i \\ &\leq \max_i h_i = h_1 \end{aligned} \quad (2.35)$$

The second group of inequalities can again be proved by the triangle inequality followed by the inner product argument

$$\begin{aligned} \sum_i |g_i| &= \sum_i \left| \sum_j P_{ij} Q_{ij}^T h_j \right| \\ &\leq \sum_j (\sum_i |P_{ij}| |Q_{ij}^T|) |h_j| \leq \sum_j h_j = 1 \end{aligned} \quad (2.36)$$

from which

$$\begin{aligned} g_1 + g_2 + g_3 &\leq 1, & g_1 - g_2 - g_3 &\leq 1, \\ -g_1 + g_2 - g_3 &\leq 1, & -g_1 - g_2 + g_3 &\leq 1. \end{aligned} \quad (2.37)$$

can be obtained. The final set of inequalities can be proved by the following method, which is illustrated for $g_1 + g_2 + g_3$.

$$\begin{aligned} g_1 + g_2 + g_3 &= \begin{pmatrix} P_{11}Q_{11}^T \\ +P_{21}Q_{21}^T \\ +P_{31}Q_{31}^T \end{pmatrix} h_1 + \begin{pmatrix} P_{12}Q_{12}^T \\ +P_{22}Q_{22}^T \\ +P_{32}Q_{32}^T \end{pmatrix} h_2 + \begin{pmatrix} P_{13}Q_{13}^T \\ +P_{23}Q_{23}^T \\ +P_{33}Q_{33}^T \end{pmatrix} h_3 \\ &= \lambda_1 h_1 + \lambda_2 h_2 + \lambda_3 h_3 \end{aligned} \quad (2.38)$$

where λ_i is the coefficient of h_i in the parenthesis. By the inner product argument $|\lambda_i| \leq 1$ and it can also be shown that $\sum_i \lambda_i \geq -1$. To see this, note that

$$\begin{aligned} \sum_i \lambda_i &= P_{11}Q_{11}^T + P_{21}Q_{21}^T + P_{31}Q_{31}^T + P_{12}Q_{12}^T + P_{22}Q_{22}^T + P_{32}Q_{32}^T \\ &\quad + P_{13}Q_{13}^T + P_{23}Q_{23}^T + P_{33}Q_{33}^T \\ &= \text{tr}(PQ) \end{aligned} \quad (2.39)$$

Since $P, Q \in SO(3)$, $PQ \in SO(3)$. Every $SO(3)$ matrix is a spatial rotation, having an eigenvalue $+1$ that corresponds to an eigenvector in the direction of the axis of rotation. Moreover, since the determinant of an $SO(3)$ matrix is 1, the other two eigenvalues are of the form $e^{\pm i\phi}$. Therefore, the trace is $1 + 2 \cos \phi \geq -1$. From this, the final inequality for $g_1 + g_2 + g_3$ can be proved as follows

$$\begin{aligned} g_1 + g_2 + g_3 &\geq \lambda_1 h_1 + \lambda_2 h_2 + (-1 - \lambda_1 - \lambda_2) h_3 \\ &= \lambda_1 (h_1 - h_3) + \lambda_2 (h_2 - h_3) - h_3 \end{aligned} \quad (2.40)$$

$$\geq -h_1 - h_2 + h_3 \quad (2.41)$$

$$= -(1 - 2h_3)$$

where (2.41) is the minimum of (2.40) attained at $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 1$. This completes the proof of the inequalities for $g_1 + g_2 + g_3$. The remaining

three inequalities

$$\begin{aligned}
 +g_1 - g_2 - g_3 &\geq -(1 - 2h_3) \\
 -g_1 + g_2 - g_3 &\geq -(1 - 2h_3) \\
 -g_1 - g_2 + g_3 &\geq -(1 - 2h_3)
 \end{aligned} \tag{2.42}$$

can be proved similarly. For example, the previous argument can be applied to the expression

$$g_1 - g_2 - g_3 = \begin{pmatrix} P_{11}Q_{11}^T \\ -P_{21}Q_{21}^T \\ -P_{31}Q_{31}^T \end{pmatrix} h_1 + \begin{pmatrix} P_{12}Q_{12}^T \\ -P_{22}Q_{22}^T \\ -P_{32}Q_{32}^T \end{pmatrix} h_2 + \begin{pmatrix} P_{13}Q_{13}^T \\ -P_{23}Q_{23}^T \\ -P_{33}Q_{33}^T \end{pmatrix} h_3 \tag{2.43}$$

by defining a new $SO(3)$ matrix \tilde{P} to be

$$\tilde{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} P \tag{2.44}$$

Thus, (g_1, g_2, g_3) satisfies (2.13), the defining inequalities of \mathcal{P}_H . The diagonal part of any PD_HQ is in \mathcal{P}_H and hence $\mathcal{C}_H = \mathcal{P}_H$.

2.B More general simulation protocols

This chapter has focussed on deterministic simulation protocols that involve no ancillary systems and where Alice and Bob's local actions are unitary. Denote this class of operations by LU and the Hilbert space of Alice's (Bob's) qubit in this type of protocol by \mathcal{H}_A (\mathcal{H}_B).

Some immediate generalizations would be to allow Alice and Bob to have ancillary systems $\mathcal{H}_{A'}$ and $\mathcal{H}_{B'}$ of arbitrary dimension and allow Alice to perform local unitary operations on $\mathcal{H}_A \otimes \mathcal{H}_{A'}$ and similarly for Bob on $\mathcal{H}_B \otimes \mathcal{H}_{B'}$. The states of $\mathcal{H}_{A'}$ and $\mathcal{H}_{B'}$ are assumed to be initially uncorrelated. Denote this type of protocol by LU + anc. Similarly, Alice and Bob might be allowed to perform more general trace-preserving quantum operations, either only on \mathcal{H}_A and \mathcal{H}_B (LO) or on a larger system including ancillas that are initially uncorrelated (LO + anc).

Clearly, $\text{LU} \subset \text{LU} + \text{anc} \subset \text{LO} + \text{anc}$ and $\text{LU} \subset \text{LO} \subset \text{LO} + \text{anc}$. However, $\text{LU} + \text{anc}$, LO and $\text{LO} + \text{anc}$ are all equivalent in terms of the efficiency of simulation they can be used to achieve. To see this, consider first $\text{LU} + \text{anc}$ and $\text{LO} + \text{anc}$. It is well known that any trace preserving operation can be implemented by performing a unitary operation on a larger Hilbert space and then discarding the extra degrees of freedom (see [74] for example). Thus, the only differences between $\text{LU} + \text{anc}$ and $\text{LO} + \text{anc}$ is that measurements and tracing out some of the ancillary degrees of freedom are allowed in the latter. However, these are not necessary for the following two reasons.

Firstly, measurements can be delayed until the end of the protocol by replacing the measurements with controlled unitary operations that store the measurement outcomes in an ancilla. Then, any operations that depend on the measurement results can be replaced by unitary operations controlled on the state of these ancillas. Secondly, any simulation protocol must end with a state of the form $e^{-iH't'} |\psi\rangle_{AB} \otimes |\phi\rangle_{A'B'}$ for any input state $|\psi\rangle_{AB}$ regardless of any measurement outcomes because only deterministic protocols are being considered here. Replacing measurements with controlled unitaries will still lead to a final state of this form, so no actual measurements or discarding of ancillas are needed.

This means that LO and $\text{LO} + \text{anc}$ are no more powerful than $\text{LU} + \text{anc}$. Conversely, due to the second reason given above, any $\text{LU} + \text{anc}$ protocol can be viewed as an LO protocol. Thus LO , $\text{LO} + \text{anc}$ and $\text{LU} + \text{anc}$ are all equivalent.

It remains to determine whether $\text{LU} + \text{anc}$ is more powerful than LU . This has been shown to be true in general [94], but in fact they are equivalent for two-qubit simulation protocols. To show this, the following lemma is used, which is valid for bipartite dynamics simulation protocols in arbitrary dimensions.

Lemma 2.4. Any LU + anc simulation protocol can be described by

$$sH' = \sum_j p_j M_j \otimes N_j (H \otimes I_{A'B'}) M_j^\dagger \otimes N_j^\dagger \quad (2.45)$$

where s is the simulation factor, the p_j 's form a probability distribution, each M_j is a linear operator $M_j : \mathcal{H}_{AA'} \rightarrow \mathcal{H}_A$ and each N_j is a linear operator $\mathcal{H}_{BB'} \rightarrow \mathcal{H}_B$.

Proof. The most general LU + anc simulation protocol may be written as

$$\begin{aligned} e^{-iH't'} |\psi\rangle_{AB} \otimes (V_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'}) \\ = c \otimes d \prod_{j=1}^N a_j \otimes b_j e^{-iHt_j} a_j^\dagger \otimes b_j^\dagger |\psi\rangle_{AB} \otimes |0\rangle_{A'} \otimes |0\rangle_{B'} \end{aligned} \quad (2.46)$$

where equality holds for all $|\psi\rangle_{AB} \in H_{AB}$. Here, a_j, c are unitary operators acting on $\mathcal{H}_{AA'}$, b_j, d are unitary operators acting on $\mathcal{H}_{BB'}$ and $|0\rangle_{A'} \otimes |0\rangle_{B'}$ is the initial state of the ancillas, which can be chosen to be an arbitrary product state. The operator, $V_{A'B'}$ is the residual unitary acting on $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$ and it may generate entanglement between the two ancillary systems.

This must be valid for all $t = \sum_j t_j$, t' and in particular, when they are both arbitrarily small. Thus, defining $p_j = t_j/t$ and $s = t'/t$ and expanding (2.46) to first order in t and t' gives

$$\begin{aligned} c \otimes d \left(I_{AB A'B'} - it \sum_j p_j a_j \otimes b_j H \otimes I_{A'B'} a_j^\dagger \otimes b_j^\dagger \right) |0\rangle_{A'} \otimes |0\rangle_{B'} \\ = (I_{AB} - itsH') \otimes (V_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'}) \end{aligned} \quad (2.47)$$

where each term is an operator on $\mathcal{H}_A \otimes \mathcal{H}_B$. It follows that

$$c |0\rangle_{A'} \otimes d |0\rangle_{B'} = I_{AB} \otimes (V_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'}) + O(t) \quad (2.48)$$

This implies that $V_{A'B'}$ must be a product operator to zeroth order in t , i.e.

$$V_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'} = V_{A'} |0\rangle_{A'} \otimes V_{B'} |0\rangle_{B'} + O(t) \quad (2.49)$$

and

$$\begin{aligned} c |0\rangle_{A'} &= I_A \otimes V_{A'} |0\rangle_{A'} + O(t) \\ d |0\rangle_{B'} &= I_B \otimes V_{B'} |0\rangle_{B'} + O(t) \end{aligned} \quad (2.50)$$

Defining $\hat{c} = I_A \otimes V_{A'}^\dagger c$, $\hat{d} = I_B \otimes V_{B'}^\dagger d$ and $\hat{V}_{A'B'} = V_{A'}^\dagger \otimes V_{B'}^\dagger V_{A'B'}$ and writing the most general $O(t)$ terms in (2.49) and (2.50) gives

$$\begin{aligned}
 \hat{c} |0\rangle_{A'} &= (I_{AA'} - itK_{AA'}) |0\rangle_{A'} + O(t^2) \\
 \hat{d} |0\rangle_{B'} &= (I_{BB'} - itK_{BB'}) |0\rangle_{B'} + O(t^2) \\
 \hat{V}_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'} &= (I_{A'} - itK_{A'}) |0\rangle_{A'} \otimes (I_{B'} - itK_{B'}) |0\rangle_{B'} \\
 &\quad - itK_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'} + O(t^2)
 \end{aligned} \tag{2.51}$$

where all the K 's are arbitrary self-adjoint operators acting on the systems indicated. Pre-multiplying (2.47) by $V_{A'}^\dagger \otimes V_{B'}^\dagger$ and substituting (2.51) into the resulting equation gives

$$\begin{aligned}
 sH' \otimes I_{A'B'} |0\rangle_{A'} \otimes |0\rangle_{B'} &= \left(\sum_j p_j a_j \otimes b_j (H \otimes I_{A'B'}) a_j^\dagger \otimes b_j^\dagger + K_{AA'} \right. \\
 &\quad \left. + K_{BB'} - K_{A'} - K_{B'} - K_{A'B'} \right) |0\rangle_{A'} \otimes |0\rangle_{B'}' \\
 &\quad + O(t)
 \end{aligned} \tag{2.52}$$

where the equal zeroth order terms have been subtracted from both sides and the remaining terms have been divided by t .

Multiplying this equation on the left by $\langle 0|_{A'} \otimes \langle 0|_{B'}$, all the K terms become local. However, since H' has no local terms, their contributions must vanish and this gives

$$sH' = \langle 0|_{A'} \otimes \langle 0|_{B'}' \left(\sum_j p_j a_j \otimes b_j (H \otimes I_{A'B'}) a_j^\dagger \otimes b_j^\dagger \right) |0\rangle_{A'} \otimes |0\rangle_{B'} \tag{2.53}$$

Substituting $M_j = \langle 0|_{A'} a_j$ and $N_j = \langle 0|_{B'} b_j$ gives the desired result. \square

Note that, in the case where there are no ancillas and a_j, b_j act on $\mathcal{H}_A, \mathcal{H}_B$ only, (2.5) is recovered.

Theorem 2.5. *LU and LU + anc are equivalent for two-qubit Hamiltonian simulation.*

Proof. The result can be proved by considering a single term in (2.45), $M \otimes N (H_{AB} \otimes I_{A'B'}) M^\dagger \otimes N^\dagger$, where the subscripts j have been dropped. This is done by showing that the contribution to H' given by this term can be

obtained by a convex sum of conjugations by local unitaries acting on \mathcal{H}_A and \mathcal{H}_B only, i.e. by an equally efficient LU protocol. First note this term can be written as a composition of two linear maps that act on the spaces $\mathfrak{B}(\mathcal{H}_A)$ and $\mathfrak{B}(\mathcal{H}_B)$ respectively. To do this, define the maps $\mathcal{E}_A(T_A) = M(T_A \otimes I_{A'})M^\dagger$ and $\mathcal{E}_B(T_B) = N(H \otimes I_{B'})N^\dagger$, which gives

$$M \otimes N (H_{AB} \otimes I_{A'B'}) M^\dagger \otimes N^\dagger = \mathcal{E}_A \circ \mathcal{E}_B (H) \quad (2.54)$$

\mathcal{E}_A and \mathcal{E}_B are unital, i.e. $\mathcal{E}_A(I) = I$ and similarly for \mathcal{E}_B . This can be shown for \mathcal{E}_A by

$$\begin{aligned} \mathcal{E}_A(I_A) &= M(I_{AA'})M^\dagger \\ &= \langle 0|_{A'} a_{AA'} I_{AA'} a_{AA'}^\dagger |0\rangle_{A'} \\ &= I_A \end{aligned} \quad (2.55)$$

and similarly for \mathcal{E}_B . Moreover, they are both completely positive, since they can be given operator sum representations by expanding $I_{A'}$ and $I_{B'}$ in orthonormal bases. For example, \mathcal{E}_A can be written as

$$\mathcal{E}_A(T) = \sum_j F_j T F_j^\dagger \quad (2.56)$$

where $F_j = M|j\rangle_{A'}$ and $\{|j\rangle_{A'}\}$ is an orthonormal basis for $\mathcal{H}_{A'}$. Generally, the maps are neither trace non-increasing or trace non-decreasing, but

$$\begin{aligned} \text{Tr} \left(\sum_j F_j^\dagger F_j \right) &= \text{Tr} \left(\sum_j \langle j|_{A'} a_{AA'}^\dagger |0\rangle_{A'} \langle 0|_{A'} a_{AA'} |j\rangle_{A'} \right) \\ &= \text{Tr} \left(\langle 0|_{A'} a_{AA'}^\dagger \sum_j (|j\rangle_{A'} \langle j|_{A'}) a_{AA'} |0\rangle_{A'} \right) \\ &= \text{Tr} \left(\langle 0|_{A'} a_{AA'}^\dagger a_{AA'} |0\rangle_{A'} \right) \\ &= \text{Tr} (\langle 0|_{A'} I_{AA'} |0\rangle_{A'}) \\ &= \text{Tr} (I_A) = 2 \end{aligned} \quad (2.57)$$

The maps \mathcal{E}_A and \mathcal{E}_B can be replaced by a convex combination of conjugations with local unitary operations and hence they can be simulated with unit efficiency by an LU protocol. The proof of this is shown explicitly for \mathcal{E}_A , but the construction for \mathcal{E}_B is the same.

Firstly, each F_j can be written in its singular value decomposition $F_j = V_j Q_j W_j$, where V_j and W_j are unitary and

$$Q_j = \begin{bmatrix} q_{j1} & 0 \\ 0 & q_{j2} \end{bmatrix} \quad (2.58)$$

where $q_{jk} \geq 0$. Using this, \mathcal{E}_A can be written as

$$\begin{aligned} \mathcal{E}_A(T) &= \sum_j V_j Q_j W_j T W_j^\dagger Q_j V_j^\dagger \\ &= \sum_j \frac{1}{2} (q_{j1}^2 + q_{j2}^2) V_j \tilde{Q}_j W_j T W_j^\dagger \tilde{Q}_j V_j^\dagger \end{aligned} \quad (2.59)$$

where

$$\tilde{Q}_j = \sqrt{2} \begin{bmatrix} \cos \theta_j & 0 \\ 0 & \sin \theta_j \end{bmatrix} \quad \text{and} \quad \cos \theta_j = \frac{q_{j1}}{\sqrt{q_{j1}^2 + q_{j2}^2}} \quad (2.60)$$

Without affecting the simulation, the map \mathcal{E}_A can be replaced by

$$\tilde{\mathcal{E}}_A(T) = \sum_j \frac{1}{2} (q_{j1}^2 + q_{j2}^2) V_j \mathcal{F}_j \left(W_j T W_j^\dagger \right) V_j^\dagger \quad (2.61)$$

where $\mathcal{F}_j(T) = (1 - \cos \theta \sin \theta) I T I + \cos \theta \sin \theta \sigma_3 T \sigma_3$. To see this, compare the map \mathcal{F}_j with conjugation by \tilde{Q}_j .

$$\begin{aligned} \tilde{Q}_j I \tilde{Q}_j &= I + \cos(2\theta_j) \sigma_3 & \tilde{Q}_j \sigma_1 \tilde{Q}_j &= \sin(2\theta_j) \sigma_1 \\ \tilde{Q}_j \sigma_2 \tilde{Q}_j &= \sin(2\theta_j) \sigma_2 & \tilde{Q}_j \sigma_3 \tilde{Q}_j &= \cos(2\theta_j) I + \sigma_3 \end{aligned} \quad (2.62)$$

$$\begin{aligned} \mathcal{F}_j(I) &= I & \mathcal{F}_j(\sigma_1) &= \sin(2\theta_j) \sigma_1 \\ \mathcal{F}_j(\sigma_2) &= \sin(2\theta_j) \sigma_2 & \mathcal{F}_j(\sigma_3) &= \sigma_3 \end{aligned} \quad (2.63)$$

They differ only when the input has an I or σ_3 component. This will not affect the Hamiltonian simulation for two reasons. Firstly, the normal form Hamiltonian has no I terms and none are generated by the conjugations with $a_j \otimes b_j$. Secondly, the action on σ_3 only differs by the addition of an I term. This will generate a purely local evolution and can be eliminated with another local unitary operation. Finally, note that $\tilde{\mathcal{E}}_A$ can be implemented by an LU protocol and $\sum_j \frac{1}{2} (q_{j1}^2 + q_{j2}^2) = \frac{1}{2} \sum_j \text{Tr} \left(F_j^\dagger F_j \right) = 1$, so it is indeed a convex combination of the individual terms and can thus be implemented with unit efficiency. \square

Chapter 3

The Entangling Capacity of Quantum Gates

3.1 Introduction

The fundamental resource used in many quantum information protocols, such as cryptography and teleportation, is the entanglement in a quantum state. A major theme of investigation in quantum information theory is the analysis and characterization of entanglement properties of quantum states under local operations and classical communication (LOCC). One issue is how to extract the entanglement in a quantum state. The simplest protocols involve taking a single copy of the quantum state and using LOCC to extract the entanglement [73]. An important realization is that, in general, collective processing (i.e. processing more than one copy of the state at a time) is more efficient than individual copy processing. Indeed, for mixed states [67], there are examples where no entanglement can be extracted at all if one only has one copy, but collective processing does allow extraction of entanglement. The fact that *asymptotic* collective processing (i.e. processing of infinitely many copies) is necessary for the *reversible* extraction of entanglement is a key building block in the general theory of entanglement [8, 69].

In contrast, the fundamental resource considered in this thesis is a non-local quantum operation, such as an interaction Hamiltonian or a unitary gate. These can be used, along with local actions, to perform the steps of quantum algorithms and generate entangled states. Conversely, an entangled state and LOCC can be used to apply a non-local operation to an arbitrary state, enabling distributed quantum processing. In analogy to the interconversion of states, it is natural to ask whether non-local operations and entangled states may be converted between reversibly and whether optimal conversion requires collective processing.

This chapter focuses on the problem of generating entanglement from the action of two-qubit unitary operations on pure states. Suppose that Alice and Bob share a state $|\psi\rangle$ in their combined Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ and that they are able to implement an operation $U_{AB} \in U(4)$ on any non-local two-qubit subspace. They would like to maximize the amount of entanglement that they generate per application of U_{AB} . This maximum is the entangling capacity, \mathcal{EC}_E , of U_{AB} . For single applications of U_{AB} , the entangling capacity is given by

$$\mathcal{EC}_E(U_{AB}) = \max_{|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B} [E(U_{AB}|\psi\rangle) - E(|\psi\rangle)] \quad (3.1)$$

where E is an entanglement measure and U_{AB} acts on one qubit in \mathcal{H}_A and one in \mathcal{H}_B .

In §3.2, the useful decomposition of two-qubit unitaries that was introduced in [60, 59] is reviewed. §3.3 concerns the single-copy entangling capacity. In §3.3.1, an argument due to [66, 12] that shows that the single-copy entangling capacity can be achieved when U_{AB} is only allowed to act on pure states is reviewed. This argument is then extended to show that pure states can still be used if the entangling capacity is to be achieved using the minimal amount of initial entanglement. In §3.3.2 and §3.3.3, the amount of entanglement that can be generated by a single use of a quantum operation when Alice and Bob share initial entanglement is found; this work extends [60] where the authors considered entangling capacities of unitaries but did not allow initial entangle-

ment; it also extends [38], which allowed initial entanglement but only unitary transformations infinitesimally close to the identity (i.e. Hamiltonians). In the case where ancillas are not allowed (§3.3.2), analytic results about the entangling capacities of unitaries are derived. It generally helps to start with an entangled state, although this is dependent on the entanglement measure used in (3.1). §3.3.3 concerns the case where ancillas are allowed; mostly numerical results are described here, however these numerical results allow us to conclude, quite generally, that allowing initial entanglement can increase the entangling capacity even when ancillas are available.

The final part of this chapter (§3.4) concerns collective processing of quantum operations. As described above, collective processing is a key idea in understanding entanglement properties of quantum states. The main result, essentially that collective processing of quantum operations does not help in generating quantum entanglement, is in stark contrast to the situation for processing of quantum states. To conclude, the implications of these results for the interconvertibility of quantum operations and the classification of their entanglement properties are discussed.

3.2 Canonical form for two-qubit unitary operators

The entanglement properties of a unitary operation, such as its ability to simulate other operations and generate entanglement, are invariant under local unitary operations applied before or after the operation. This gives a notion of local equivalence of operations

$$U_{AB} \sim U'_{AB} \text{ iff } U'_{AB} = V_A^{(1)} \otimes V_B^{(2)} U_{AB} W_A^{(1)} \otimes W_B^{(2)} \quad (3.2)$$

where $V^{(1)}, V^{(2)}, W^{(1)}, W^{(2)}$ are local unitaries acting on the systems indicated.

In order to simplify the calculations in this chapter, the following decomposition of two-qubit unitary operators is used.

Theorem 3.1. *Any two-qubit unitary, $U_{AB} \sim U_d$, where*

$$U_d = e^{i \sum_{j=1}^3 \alpha_j \sigma_j^A \otimes \sigma_j^B} \quad (3.3)$$

$\frac{\pi}{4} \geq \alpha_1 \geq \alpha_2 \geq |\alpha_3| \geq 0$ and $\sigma_{1,2,3}$ are the Pauli matrices.

Since, U_d has the same entangling capacity as U , this form will always be used in what follows¹. Note that the eigenvalues of U_d are given by $e^{i\lambda_j}$ where

$$\begin{aligned} \lambda_1 &= -\alpha_1 + \alpha_2 + \alpha_3 \\ \lambda_2 &= +\alpha_1 - \alpha_2 + \alpha_3 \\ \lambda_3 &= +\alpha_1 + \alpha_2 - \alpha_3 \\ \lambda_4 &= -\alpha_1 - \alpha_2 - \alpha_3 \end{aligned} \quad (3.4)$$

The corresponding eigen-basis is given by $U_d |\Phi_j\rangle = e^{i\lambda_j} |\Phi_j\rangle$ and is the Bell basis. For later convenience, the following phase convention is chosen:

$$\begin{aligned} |\Phi_1\rangle &= \frac{-i}{\sqrt{2}} (|00\rangle - |11\rangle) \\ |\Phi_2\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \\ |\Phi_3\rangle &= \frac{-i}{\sqrt{2}} (|01\rangle + |10\rangle) \\ |\Phi_4\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \end{aligned} \quad (3.5)$$

With this choice of phase convention, this basis is known as the “magic” basis [47]. Using the magic basis often simplifies calculations related to the entanglement of two-qubit states and unitaries.

To prove, theorem 3.1 it must be shown that any U_{AB} is locally equivalent to an operator that is diagonal in the magic basis. This relies on the following three lemmas.

Lemma 3.2. *A local unitary operation $W^{(1)} \otimes W^{(2)} \in SU(2) \times SU(2)$ acts on the coefficients of states expressed in the magic basis as an $SO(4, \mathbb{R})$ matrix.*

That is

$$W_A^{(1)} \otimes W_B^{(2)} |\Phi_j\rangle_{AB} = \sum_{k=1}^4 O_{jk} |\Phi_k\rangle_{AB} \quad (3.6)$$

where O_{jk} are the components of an $SO(4, \mathbb{R})$ matrix O .

¹In fact, when considering the entangling capacity, the condition $\alpha_3 \geq 0$ can be assumed. This is because $e^{i \sum_{j=1}^2 \alpha_j \sigma_j^A \otimes \sigma_j^B - \alpha_3 \sigma_3^A \otimes \sigma_3^B} \sim \left(e^{i \sum_{j=1}^3 \alpha_j \sigma_j^A \otimes \sigma_j^B} \right)^*$ and \mathcal{EC}_E is invariant under conjugation

Proof. This can be shown by demonstrating that linear combinations of the generators of $SU(2) \times SU(2)$ act like the generators of $SO(4, \mathbb{R})$ on the magic basis. $SU(2) \times SU(2)$ is generated by the six operators $\sigma_j \otimes I, I \otimes \sigma_j, j = 1, 2, 3$ and $SO(4, \mathbb{R})$ rotations of vectors expressed in the magic basis are generated by the six operators

$$X_{jk} |\Phi_m\rangle = i\delta_{mj} |\Phi_k\rangle - i\delta_{mk} |\Phi_j\rangle \quad (3.7)$$

where $1 \leq j < k \leq 4$.

To see how this construction works, take X_{12} , the generator of rotations in the $|\Phi_1\rangle, |\Phi_2\rangle$ plane, as an example. The operators $\sigma_3 \otimes I$ and $I \otimes \sigma_3$ act as follows on the magic basis

$$\begin{aligned} \sigma_3 \otimes I |\Phi_1\rangle &= -i |\Phi_2\rangle & , & \quad I \otimes \sigma_3 |\Phi_1\rangle = -i |\Phi_2\rangle \\ \sigma_3 \otimes I |\Phi_2\rangle &= i |\Phi_1\rangle & , & \quad I \otimes \sigma_3 |\Phi_2\rangle = i |\Phi_1\rangle \\ \sigma_3 \otimes I |\Phi_3\rangle &= -i |\Phi_4\rangle & , & \quad I \otimes \sigma_3 |\Phi_3\rangle = i |\Phi_4\rangle \\ \sigma_3 \otimes I |\Phi_4\rangle &= i |\Phi_3\rangle & , & \quad I \otimes \sigma_3 |\Phi_4\rangle = -i |\Phi_3\rangle \end{aligned} \quad (3.8)$$

and setting $X_{12} = -\frac{1}{2}(\sigma_3 \otimes I + I \otimes \sigma_3)$ gives the generator required by (3.7). The remaining five generators are constructed in a similar way. They are all linearly independent and span the Lie algebra of $SU(2) \times SU(2)$ and hence $SU(2) \times SU(2)$ acts as $SO(4, \mathbb{R})$ on the magic basis. \square

The next two lemmas are modified versions of the usual polar and singular value decompositions for matrices.

Lemma 3.3. *Any unitary matrix, U , may be written as $U = OS$, where $S = \sqrt{U^T U}$ is a complex unitary symmetric matrix and O is a real orthogonal matrix.*

Proof. Firstly for two complex vectors \vec{v}, \vec{w} , let $\langle \vec{v} | \vec{w} \rangle = \sum_j v_j^* w_j$ denote the usual inner product and let $\vec{v}^T \vec{w} = \sum_j v_j w_j$.

$S = \sqrt{U^T U}$ is a unitary matrix, so its spectral decomposition is of the form $S = \sum_j e^{i\lambda_j} |\vec{v}_j\rangle \langle \vec{v}_j|$, where \vec{v}_j are a complete set of orthonormal eigenvectors

and $\lambda_j \in \mathbb{R}$. Moreover, since S also is symmetric, the eigenvectors \vec{v}_j can be chosen to have real components. To see this, note that

$$S^*S = (S^T)^\dagger S = S^\dagger S = I \quad (3.9)$$

Then, from the eigenvalue equation

$$\begin{aligned} S\vec{v}_j &= e^{i\lambda_j}\vec{v}_j \\ \Rightarrow S^*S\vec{v}_j &= e^{i\lambda_j}S^*\vec{v}_j \\ \Rightarrow S^*\vec{v}_j &= e^{-i\lambda_j}\vec{v}_j \end{aligned} \quad (3.10)$$

Thus, the vectors \vec{v}_j will also be eigenvalues of the matrix S^* . This means that they will be eigenvectors of the real symmetric matrix $S + S^*$ and the eigenvectors of such a matrix can be chosen to have real components.

Next, define the vectors $\vec{w}_j = e^{-i\lambda_j}U\vec{v}_j$. These vectors also have real components, which can be checked by comparing

$$\begin{aligned} \langle \vec{w}_j | \vec{w}_j \rangle &= \langle \vec{v}_j | U^\dagger e^{i\lambda_j} e^{-i\lambda_j} U | \vec{v}_j \rangle \\ &= \langle \vec{v}_j | U^\dagger U | \vec{v}_j \rangle \\ &= \langle \vec{v}_j | \vec{v}_j \rangle = 1 \end{aligned} \quad (3.11)$$

with

$$\begin{aligned} \vec{w}_j^T \vec{w}_j &= e^{-2i\lambda_j} \vec{v}_j^T U^T U \vec{v}_j \\ &= e^{-2i\lambda_j} \vec{v}_j^T S^2 \vec{v}_j \\ &= \vec{v}_j^T \vec{v}_j = 1 \end{aligned} \quad (3.12)$$

and noting that these two forms should only coincide if the vectors \vec{w}_j have real components.

Finally, define a real, orthogonal matrix $O = \sum_j \vec{w}_j \vec{v}_j^T$. We have that $OS\vec{v}_j = e^{i\lambda_j}\vec{w}_j = U\vec{v}_j$. Since the action of OS and U coincide on a complete orthonormal basis, $U = OS$. \square

Lemma 3.4. *Any unitary operator, U , may be written as $U = PDQ$, where P and Q are real orthogonal matrices and D is a diagonal matrix with diagonal elements given by the eigenvalues of $\sqrt{U^T U}$.*

Proof. From the previous lemma, $U = OS$, where O is a real orthogonal matrix and S is a symmetric unitary matrix. It was also shown that the eigenvectors of S can be chosen to have real components, so it follows that $S = Q^T D Q$ for some real orthogonal matrices R and Q . The proof is completed by setting $P = OQ^T$. \square

Proof of theorem 3.1. To prove the decomposition, it must be shown that U_{AB} is locally equivalent to a unitary that is diagonal in the magic basis. This can be done by making use of an isomorphism between operators acting on $\mathbb{C}^2 \otimes \mathbb{C}^2$ and states in $\mathbb{C}^4 \otimes \mathbb{C}^4$ [103]. Let M_{AB} be such an operator and define the state

$$|\Psi_M\rangle_{AA'BB'} = M_{AB} |\Phi_2\rangle_{AA'} \otimes |\Phi_2\rangle_{BB'} \quad (3.13)$$

This is an isomorphism, because it is possible to recover the operator M_{AB} from the state $|\Psi_M\rangle_{AA'BB'}$ via

$$M_{AB} |\psi\rangle_{AB} = 4 \langle \Phi_2 |_{A'A''} \langle \Phi_2 |_{B'B''} |\Psi_M\rangle_{AA'BB'} |\psi\rangle_{A''B''} \quad (3.14)$$

where $|\psi\rangle$ is an arbitrary input state. This can be checked by expanding the operator and the states in the computational basis. Local equivalence of operators can now be reformulated in terms of states. This can be done by making use of the fact that for any single qubit operator W the following holds

$$W_A \otimes I_B |\Phi_2\rangle_{AB} = I_A \otimes W_B^T |\Phi_2\rangle_{AB} \quad (3.15)$$

which can again be checked by expanding in the computational basis. This shows that under a local transformation of an operator $M_{AB} \rightarrow V_A^{(1)} \otimes V_B^{(2)} M_{AB} W_A^{(1)} \otimes W_B^{(2)}$, the corresponding state $|\Psi_M\rangle_{AA'BB'}$ transforms as

$$\begin{aligned} |\Psi_M\rangle &\rightarrow V_A^{(1)} \otimes V_B^{(2)} M_{AB} W_A^{(1)} \otimes W_B^{(2)} |\Phi_2\rangle_{AA'} \otimes |\Phi_2\rangle_{BB'} \\ &= V_A^{(1)} \otimes \left(W_{A'}^{(1)}\right)^T \otimes V_B^{(2)} \otimes \left(W_{B'}^{(2)}\right)^T M_{AB} |\Phi_2\rangle_{AA'} \otimes |\Phi_2\rangle_{BB'} \\ &= V_A^{(1)} \otimes \left(W_{A'}^{(1)}\right)^T \otimes V_B^{(2)} \otimes \left(W_{B'}^{(2)}\right)^T |\Psi_M\rangle_{AA'BB'} \end{aligned} \quad (3.16)$$

Thus, local equivalence of operators M_{AB} induces an equivalence under $U(2)^4$ for the states $|\Psi_M\rangle_{AA'BB'}$. Now, consider the case where $M_{AB} = U_{AB}$ is a unitary operator. Without loss of generality, assume that $U_{AB} \in SU(4)$ and that the local unitaries are in $SU(2)$ since global phases of the local unitaries only affect the global phase of U_{AB} , which may be absorbed by setting the global phase of any of the local unitaries appropriately. The state $|\Psi_U\rangle_{AA'BB'}$ can then be written as

$$|\Psi_U\rangle_{AA'BB'} = \sum_{j,k=1}^4 N_{jk} |\Phi_j\rangle_{AA'} |\Phi_j\rangle_{BB'} \quad (3.17)$$

where N_{jk} are the components of an $SU(4)$ matrix. From (3.13) and (3.14) it can be shown that N_{jk} are the components of U_{AB} when it is expressed in the magic basis. By lemma 3.2, this matrix transforms under (3.16) as

$$N \rightarrow PNQ \quad (3.18)$$

where $P, Q \in SO(4, \mathbb{R})$. Thus, according to lemma 3.4, N can be transformed to a diagonal matrix with diagonal elements given by the eigenvalues of $\sqrt{N^T N}$. □

The fact that the eigenvalues of $\sqrt{N^T N}$ are local invariants of U_{AB} was first noted in [71] The method given above for finding these eigenvalues of requires U_{AB} to be expressed in the magic basis. However, there is an alternative that works in the computational basis. Firstly, define the unitary operator analog of the spin-flipped density operator which was defined in eq. (1.41).

$$\tilde{U} = \sigma_2 \otimes \sigma_2 U^T \sigma_2 \otimes \sigma_2 \quad (3.19)$$

where T indicates the transpose in the computational basis. The eigenvalues of $\tilde{U}U$ are local invariants of U and, from, eq.(3.3) one can see that these invariants are in fact squares of the eigenvalues of U_d , which are the same as the eigenvalues of $N^T N$.

3.3 Single Copy Entangling Capacity

3.3.1 Purity of States in the Optimal Protocol

In this section we determine whether optimal protocols can be found for generating entanglement using one application of U_{AB} that only involve pure states at every stage. An argument of [12] can be used to establish that this is the case. Further, this argument can be extended to show that optimal pure state protocols can be found that start with the minimum possible amount of initial entanglement. Thus, all the important details of the single-copy entangling capacity of U_{AB} can be established by considering pure states only.

Making a suitable definition of the entangling capacity over mixed states is not quite as straightforward as the pure state case. In particular, the choice of entanglement measure for the initial and final states may be different. For the initial state, it seems natural to use a measure of the minimum average amount of entanglement required to generate it (i.e. the entanglement of formation²). However, for the final state it makes more sense to measure the maximum amount of entanglement that can be extracted from it (i.e. the distillable entanglement).

To make this more specific, consider an initial mixed state ρ_0 . Let $\rho_0 = \sum_j p_j |\psi_j\rangle \langle \psi_j|$ be the decomposition of ρ_0 with minimal ensemble average entanglement. To generate an ensemble of n states described by ρ_0 , the state $|\psi_j\rangle$ may be prepared with probability p_j and then the information about which state was prepared may be discarded. As $n \rightarrow \infty$, the amount of entanglement per state used in this procedure will be $E_F(\rho_0)$, where E_F is the entanglement of formation (1.38). The operation U_{AB} can then be applied to each state individually yielding n copies of the state $\rho_1 = U_{AB}\rho_0U_{AB}^\dagger$. These states can then be distilled to singlets by LOCC and as $n \rightarrow \infty$ the yield of singlets per copy of ρ_1 will be $E_D(\rho_1)$, where E_D is the distillable entanglement. Note that, although this protocol involves collective processing of the states, the

²The entanglement cost is also considered in this context in [12]

fact that U_{AB} is applied to each copy of ρ_0 individually means that it can still be regarded as a single-copy protocol with respect to the unitary operator.

With this in mind, the mixed state single-copy entangling capacity, C_E^{mixed} , is defined as

$$C_E^{mixed} = \max_{\rho_0} (E_D(\rho_1) - E_F(\rho_0)) \quad (3.20)$$

Then

$$\begin{aligned} E_D(\rho_1) - E_F(\rho_0) &\leq E_F(\rho_1) - E_F(\rho_0) \\ &= E_F \left(\sum_j p_j U_{AB} |\psi_j\rangle \langle \psi_j| U_{AB}^\dagger \right) \\ &\quad - \sum_j p_j E_F(|\psi_j\rangle \langle \psi_j|) \end{aligned} \quad (3.21)$$

$$\begin{aligned} &\leq \sum_j p_j \left[E_F(U_{AB} |\psi_j\rangle \langle \psi_j| U_{AB}^\dagger) \right. \\ &\quad \left. - E_F(|\psi_j\rangle \langle \psi_j|) \right] \end{aligned} \quad (3.22)$$

$$\leq \max_{\psi_j} (E_F(U_{AB} |\psi_j\rangle \langle \psi_j|) - E_F(|\psi_j\rangle \langle \psi_j|)) \quad (3.23)$$

Here, (3.21) follows because $\rho_0 = \sum_j p_j |\psi_j\rangle \langle \psi_j|$ is an optimal decomposition of ρ_0 , (3.22) follows from the convexity of E_F (1.39) and (3.23) follows because (3.22) is a convex sum. This demonstrates that for every mixed state, there is a pure state for which the action of U_{AB} generates at least as much entanglement.

Next we show that for any mixed state that achieves the entangling capacity there is a pure state that achieves the entangling capacity with entanglement not greater than the entanglement of formation of the mixed state. Let $|\psi\rangle$ be a pure state that achieves the entangling capacity with the minimal possible initial entanglement. Let ρ be a mixed state that also achieves the entangling capacity. From eqs.(3.22) and (3.23), by convexity, it is clear that the optimal decomposition of ρ must be a mixture of pure states that achieve the entangling capacity. Since this is the optimal decomposition of ρ , $E_F(\rho)$ is just the weighted average of the entanglements of these pure states. Thus, $E_F(\rho) \geq E_F(|\psi\rangle \langle \psi|)$ because $|\psi\rangle$ has the minimal entanglement of any possible state in this ensemble.

3.3.2 Single Application with no ancillas

In this section, the entangling capacity of two-qubit unitaries of the form of eq.(3.3) when no ancillas are allowed is determined. This depends on the entanglement measure that is chosen for the optimization. In §3.3.2 the square of concurrence is used and then in §3.3.2 these results are extended to other measures of entanglement.

Square of concurrence

One entanglement measure that is particularly convenient to optimize is the square of the concurrence [101], C , defined by

$$C(|\psi\rangle) = |\langle\psi|\sigma_2 \otimes \sigma_2|\psi^*\rangle| \quad (3.24)$$

where $|\psi^*\rangle$ is the state vector obtained by taking the complex conjugates of the components of $|\psi\rangle$ in the computational basis. An argument from [60] can be adapted to perform the optimization here.

Writing $|\psi\rangle = \sum_j b_j |\Phi_j\rangle$ gives

$$\Delta C^2 = C_f^2 - C_0^2 = \left| \sum_j e^{2i\lambda_j} b_j^2 \right|^2 - \left| \sum_j b_j^2 \right|^2 = \sum_{j,k} (e^{2i(\lambda_j - \lambda_k)} - 1) b_j^2 b_k^{*2} \quad (3.25)$$

where C_0 is the initial concurrence and C_f is the final concurrence after applying U_{AB} .

This can be optimized by imposing the normalization condition $\sum_j |b_j|^2 = 1$ with a Lagrange multiplier, 2μ , i.e. the expression to be maximized is

$$L = \sum_{j,k} (e^{2i(\lambda_j - \lambda_k)} - 1) b_j^2 b_k^{*2} - 2\mu \left(\sum_j b_j b_j^* - 1 \right) \quad (3.26)$$

Differentiating gives

$$\frac{\partial L}{\partial b_j} = 2b_j e^{2i\lambda_j} \sum_k e^{-2i\lambda_k} b_k^{*2} - 2b_j \sum_k b_k^{*2} - 2\mu b_j^* = 0 \quad (3.27)$$

multiplying by b_j and summing over j gives

$$\sum_{j,k} (e^{2i(\lambda_j - \lambda_k)} - 1) b_j^2 b_k^{*2} - \mu \sum_j |b_j|^2 = 0 \quad (3.28)$$

which yields

$$\mu = C_f^2 - C_0^2 \quad (3.29)$$

Substituting eqs.(3.29) and (3.25) into eq.(3.27) gives

$$b_j e^{2i\lambda_j} e^{2i\eta} C_f - b_j e^{2i\epsilon} C_0 - C_f^2 b_j^* + C_0^2 b_j^* = 0 \quad (3.30)$$

where ϵ, η are phases depending on all of the b_j 's. One possible solution is $b_j = 0$. To find the other solutions, write $b_j = \beta_j e^{i\gamma_j}$ where $\beta_j, \gamma_j \in \mathbb{R}$. These solutions must have $\beta_j \neq 0$ and so eq.(3.30) reduces to

$$C_f^2 - e^{2i(\lambda_j + \gamma_j + \eta)} C_f - C_0^2 + e^{2i(\gamma_j + \epsilon)} C_0 = 0 \quad (3.31)$$

There are as many equations (3.31) as there are non-zero b_j 's. For generic λ_j 's, at most two of these equations can be satisfied simultaneously.

To see this, firstly consider the case when the optimal starting state has $C_0 = 0$. Then,

$$C_f (C_f - e^{2i(\lambda_j + \gamma_j + \eta)}) = 0 \quad (3.32)$$

Since C_f is real and the maximum is required, it must be the case that $C_f = 1$. This shows that it is only best to start in a product state if U_{AB} can generate one e-bit of entanglement when no ancillas are present. The conditions for this were found in [60] to be

$$\alpha_1 + \alpha_2 \geq \frac{\pi}{4} \text{ and } \alpha_2 + \alpha_3 \leq \frac{\pi}{4} \quad (3.33)$$

so the focus here will be on the cases where (3.33) is violated and the optimal starting state must have non-zero C_0 .

Subtracting any two of eqs.(3.31) gives

$$\sin(\lambda_j - \lambda_k + \gamma_j - \gamma_k) C_f = e^{i(2\epsilon - 2\eta - \lambda_j - \lambda_k)} \sin(\gamma_j - \gamma_k) C_0 \quad (3.34)$$

This gives consistency conditions for the simultaneous solution of any pair of eqs.(3.31). In particular, since C_f and C_0 are both real,

$$2(\epsilon - \eta) - \lambda_j - \lambda_k = n\pi, n \in \mathbb{Z} \quad (3.35)$$

For generic λ_j 's this condition cannot be satisfied for more than one pair of equations in (3.31). Thus, at most two b_j 's can be non-zero³. This means that the optimal starting state will always be in a subspace spanned by two of the eigenvectors of U_{AB} . The maximum will be given by choosing the two eigenvectors and the coefficients b_j that maximize ΔC^2 . Re-expressing eq.(3.25) in terms of β_j, γ_j gives

$$\Delta C^2 = 4 \sum_{j < k} \beta_j^2 \beta_k^2 [\sin(2(\gamma_j - \gamma_k) + \lambda_j - \lambda_k) \sin(\lambda_k - \lambda_j)] \quad (3.36)$$

Only one term in this sum can be non-zero and for this term γ_j, γ_k may be chosen so that $\Delta C^2 = 4\beta_j^2 \beta_k^2 |\sin(\lambda_k - \lambda_j)|$. This is maximized by $\beta_j = \beta_k = \frac{1}{\sqrt{2}}$. Thus, the entangling capacity is given by

$$\mathcal{E}C_{C^2} = \max_{j < k} |\sin(\lambda_k - \lambda_j)| \quad (3.37)$$

Note that this is greater than the corresponding result of $\max_{j < k} |\sin(\lambda_k - \lambda_j)|^2$ found in [60] when the starting state is restricted to be a product. This shows that when (3.33) is violated, initial entanglement is always required to achieve the optimal capacity when no ancillas are allowed. There are two parameter regions where (3.33) does not hold.

1. $\alpha_1 + \alpha_2 < \frac{\pi}{4}, \alpha_2 + \alpha_3 < \frac{\pi}{4}$. In this region, the maximum is given by making the $j = 3, k = 4$ term nonzero. $\mathcal{E}C_{C^2} = \sin(2(\alpha_1 + \alpha_2))$ and the optimal starting state is $|\psi\rangle = \left(\sin\left(\frac{\alpha_1 + \alpha_2}{2} - \frac{\pi}{8}\right) |01\rangle - i \cos\left(\frac{\alpha_1 + \alpha_2}{2} - \frac{\pi}{8}\right) |10\rangle\right)$. This gives an optimal initial entanglement of $C_0^2 = \frac{1}{2} (1 - \sin 2(\alpha_1 + \alpha_2))$
2. $\alpha_1 + \alpha_2 > \frac{\pi}{4}, \alpha_2 + \alpha_3 > \frac{\pi}{4}$. In this region, the maximum is given by making the $j = 1, k = 4$ term nonzero. $\mathcal{E}C_{C^2} = \sin(2(\alpha_2 + \alpha_3))$ and the optimal starting state is $|\psi\rangle = \frac{1}{\sqrt{2}} (|\Phi_1\rangle + e^{i(\frac{\pi}{4} + \alpha_2 + \alpha_3)} |\Phi_4\rangle)$. This gives an optimal initial entanglement of $C_0^2 = \frac{1}{2} (1 - \sin 2(\alpha_2 + \alpha_3))$.

³This result can be extended to all possible λ_j 's by noting that eq.(3.34) can only be satisfied for more than one pair if some of the eigenvalues are degenerate. Further, it can be shown that one can choose only one of the corresponding b_j 's to be non-zero.

Note that the entangling capacity is always found to be a function of $\alpha_1 + \alpha_2$ or $\alpha_2 + \alpha_3$, i.e. a sum of only two of the parameters of the unitary. The value of the third parameter does not affect the entangling capacity at all when no ancillas are allowed.

Other entanglement measures

All two-qubit pure-state entanglement measures, E , are monotonic functions of one another and in particular of the concurrence squared (i.e. $E = E(C^2)$). Generalizing the strategy of eqs. (3.25-3.34) to an arbitrary entanglement measure, E , by making use of $\frac{\partial E}{\partial b_j} = \frac{\partial E}{\partial(C^2)} \frac{\partial(C^2)}{\partial b_j}$ gives

$$\sin(\lambda_j - \lambda_k + \gamma_j - \gamma_k) C_f \frac{dE_f}{d(C_f^2)} = e^{i(2\epsilon - 2\eta - \lambda_j - \lambda_k)} \sin(\gamma_j - \gamma_k) C_0 \frac{dE}{d(C_0^2)} \quad (3.38)$$

This gives the same consistency conditions as eq.(3.35) so we still have that at most two b_j 's can be non-zero. The only exception is when $\frac{dE}{d(C^2)} \propto \frac{1}{C}$, which occurs when the entanglement measure is the concurrence itself. In this case, the analog of eq. (3.30) is

$$\frac{1}{C_0} (b_j^* C_0^2 + b_j e^{2i\epsilon} C_0) - \frac{1}{C_f} (b_j^* C_f^2 + b_j e^{2*i(\lambda_j + \eta)} C_f) = 0 \quad (3.39)$$

This equation is only valid when $C_0, C_f \neq 0$, which is expected because the concurrence does not have a well defined derivative at 0. This leaves open the possibility that $C_0 = 0$ is the optimal initial concurrence and indeed this is the only consistent solution.

To see this, first assume that $C_0, C_f \neq 0$. Then eq. (3.39) again has $b_j = 0$ as a possible solution. If $b_j \neq 0$ then substituting $b_j = \beta_j e^{i\gamma_j}$ gives

$$C_f - C_0 = e^{2i\gamma_j} (e^{2i(\lambda_j + \eta)} - e^{2i\epsilon}) \quad (3.40)$$

Since, $C_f - C_0$ is real,

$$C_f - C_0 = |e^{2i(\lambda_j + \eta)} - e^{2i\epsilon}| \quad (3.41)$$

This implies that for each pair of non-zero b_j, b_k

$$\lambda_j = \lambda_k + n\pi \quad n \in \mathbb{Z} \quad (3.42)$$

For generic λ_j 's, this equation is not satisfied by any pair λ_j, λ_k when $j \neq k$. Thus, the only possibility is to have just a single non-zero b_j . However, since the basis $|\Phi_j\rangle$ is an eigenbasis for U_{AB} , this solution gives $C_f - C_0 = 0$ and is not a maximum. Therefore, choosing $C_0 = 0$ is the only remaining possibility to achieve a maximum⁴.

For all other entanglement measures we focus on the case where $\alpha_1 + \alpha_2 < \frac{\pi}{4}, \alpha_2 + \alpha_3 < \frac{\pi}{4}$. If b_j and b_k are chosen to be non-zero for some choice of $j \neq k = 1, 2, 3, 4$ then the resulting optimal ΔE is always a function of the corresponding λ_j and λ_k only. In fact, it must be the same function of λ_j and λ_k for all choices of j and k . For all the measures considered below the optimal ΔE is always a monotonically increasing function of $|\lambda_j - \lambda_k|$ ⁵. As with the square of concurrence, the j and k that give the largest value of $|\lambda_j - \lambda_k|$ are chosen, namely $j = 3, k = 4$. Thus, the optimal starting state can be written in its Schmidt decomposition as

$$|\psi\rangle = \cos(\theta) |01\rangle + e^{i\phi} \sin(\theta) |10\rangle \quad (3.43)$$

and ΔE simply has to be optimized over the Schmidt parameter θ and relative phase ϕ . This gives the following results.

1. Concurrence: $C = |\langle \psi | \sigma_2 \otimes \sigma_2 | \psi^* \rangle|$. As discussed above, this measure is unusual in that the optimal state is always a product state. Thus, $\mathcal{E}C_C = \sin(2(\alpha_1 + \alpha_2))$, which coincides with the result of [60].
2. Entropy of entanglement: $E = -\text{Tr}(\rho_A \log_2 \rho_A)$, where ρ_A is Alice's reduced density matrix. This gives a transcendental equation in θ , which can be optimized numerically for each $\alpha_1 + \alpha_2$. For results see Fig. 3.1.

⁴In the non-generic case, where eq. (3.42) is satisfied for some pairs $j \neq k$ it is possible to have more than one non-zero coefficient, but it is simple to show that $C_f - C_0$ is still zero.

⁵Similarly, when $\alpha_1 + \alpha_2 > \frac{\pi}{4}, \alpha_2 + \alpha_3 > \frac{\pi}{4}$, for any choice of j and k , the optimal ΔE is always a monotonically decreasing function of $|\lambda_j - \lambda_k|$.

3. Linearized entropy: $R = 1 - \text{Tr}(\rho_A^2)$. This gives $\mathcal{EC}_R = \sin(2(\alpha_1 + \alpha_2))$.

3.3.3 Ancillas

Next we consider whether adding ancillas can increase the entangling capacity. This problem has not been solved analytically yet, but some numerical optimizations are presented here, using entropy of entanglement as the measure. Specifically, the following definition of entangling capacity is used when ancillas are present.

$$\mathcal{EC}_E = \max_{|\psi\rangle \in \mathcal{H}_{AA'BB'}} \left[S \left(\text{Tr}_{BB'} \left(U_{AB} |\psi\rangle \langle\psi| U_{AB}^\dagger \right) \right) - S \left(\text{Tr}_{BB'} (|\psi\rangle \langle\psi|) \right) \right] \quad (3.44)$$

where S is the von Neumann entropy, \mathcal{H}_A (\mathcal{H}_B) is the Hilbert space of the qubit that Alice (Bob) acts on with U_{AB} and $\mathcal{H}_{A'}$ ($\mathcal{H}_{B'}$) is a finite dimensional ancillary Hilbert space for Alice (Bob). Only pure states over the Hilbert space $\mathcal{H}_{AA'BB'} = \mathcal{H}_A \otimes \mathcal{H}_{A'} \otimes \mathcal{H}_B \otimes \mathcal{H}_{B'}$ need to be considered because the argument of §3.3.1 implies that they are optimal.

Note that, here we are only concerned with the extent to which interaction between Alice and Bob, represented by U_{AB} , can generate entanglement between Alice and Bob. Thus, only initial and final entanglement between Alice and Bob are relevant and the entanglement of Alice or Bob with their local ancillas is not counted as part of this entanglement.

The optimization was performed for three different families of operations:

- The CNOT family $e^{i\alpha\sigma_1^A \otimes \sigma_1^B}$.
- The double CNOT (DCNOT) family $e^{i\alpha(\sigma_1^A \otimes \sigma_1^B + \sigma_2^A \otimes \sigma_2^B)}$.
- The SWAP family $e^{i\alpha(\sigma_1^A \otimes \sigma_1^B + \sigma_2^A \otimes \sigma_2^B + \sigma_3^A \otimes \sigma_3^B)}$.

The families are so named because setting $\alpha = \frac{\pi}{4}$ gives operations that are locally equivalent to the CNOT, DCNOT and SWAP operations⁶.

⁶The DCNOT is defined as $CNOT_{B \rightarrow A} CNOT_{A \rightarrow B}$. Similarly, SWAP can be defined as $CNOT_{A \rightarrow B} CNOT_{B \rightarrow A} CNOT_{A \rightarrow B}$. These operations, together with the identity are the extreme points of the parametrization (3.3).

The simulations were run with both one and two ancillary qubits on each side (i.e. with dimension 2 and 4 for $\mathcal{H}_{A'}$ and $\mathcal{H}_{B'}$). Adding 1 ancillary qubit on each side increased the entangling capacity for the DCNOT and SWAP families (see figs. 3.3 and 3.4), but there was no further increase on adding more ancillary qubits. We conjecture that one ancillary qubit on each side is the most general system required to optimize single-copy entangling capacity. Note that, for every α , the SWAP family has a higher entangling capacity than the DCNOT family. This shows that the entangling capacity is generally a function of all three parameters $(\alpha_1, \alpha_2, \alpha_3)$ of the unitary, in contrast to the case considered above where no ancillas are allowed.

For the CNOT family, adding ancillas had no effect at all (see fig. 3.2). In [60], the entangling capacity for the CNOT family starting from a product state with ancillas was found to be $H(\cos^2 \alpha) = -\cos^2(\alpha) \log_2[\cos^2(\alpha)] - \sin^2(\alpha) \log_2[\sin^2(\alpha)]$. No ancillas were required to achieve this capacity. The numerical results with initial entanglement exceed this capacity, which demonstrates that allowing initial entanglement can still increase the entangling capacity even if ancillas are present. In other words, it is not possible to achieve the entangling capacity by trading some of the initial entanglement required when no ancillas are present with entanglement between the system and ancillas.

3.4 Collective Processing

We now turn to the question of whether the entangling capacity is increased by applying n copies of a unitary operation to pairs of qubits in the most general initial state which may be entangled and may contain ancillas. The n -copy entangling capacity is then defined to be the optimal increase in entanglement over Alice and Bob's entire Hilbert space per application of the unitary. In this definition, Alice and Bob are again allowed to have arbitrarily large, but

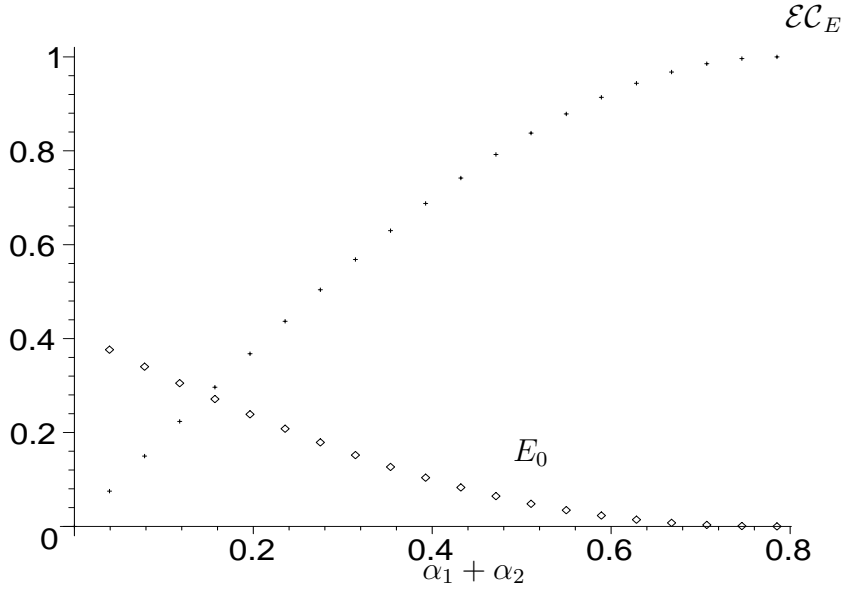


Figure 3.1: Single-copy entangling capacity and optimal initial entanglement for a general two-qubit unitary of the form of eq.(3.3) when no ancillas are allowed. Crosses show the entangling capacity and diamonds show the minimum initial entanglement of a state that achieves the capacity.

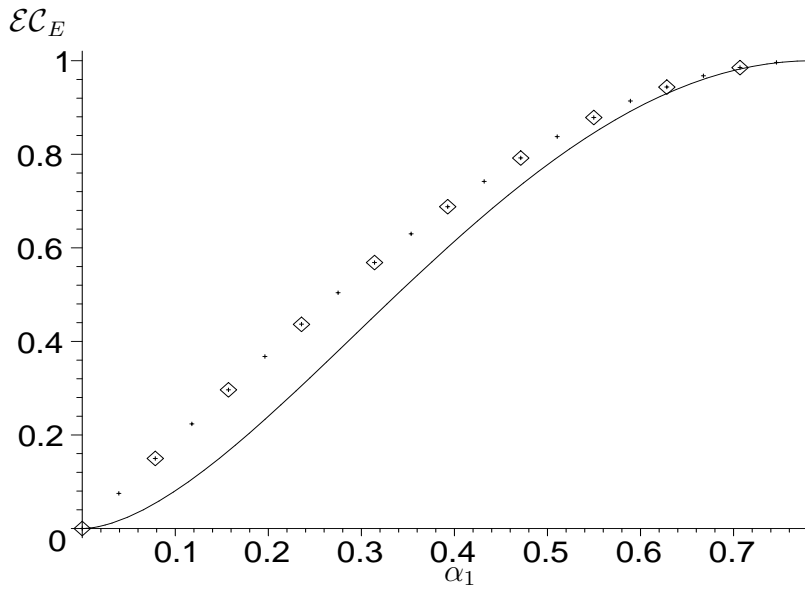


Figure 3.2: Single-copy entangling capacity for the CNOT family. Crosses are for no ancillas, diamonds are for one ancilla on each side and the line shows the equivalent result from [60] when the starting state is restricted to be a product between Alice and Bob

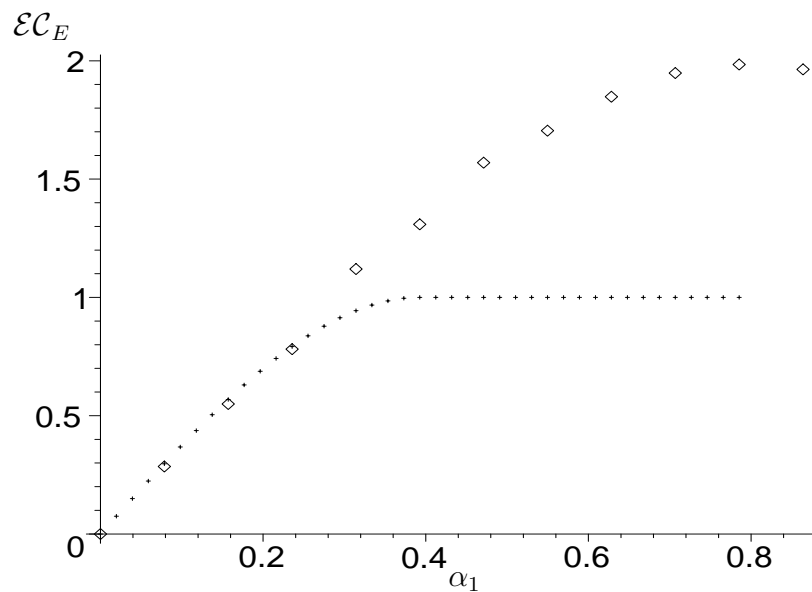


Figure 3.3: Single-copy entangling capacity for the DCNOT family. Crosses are for no ancillas and diamonds are for one ancilla on each side.

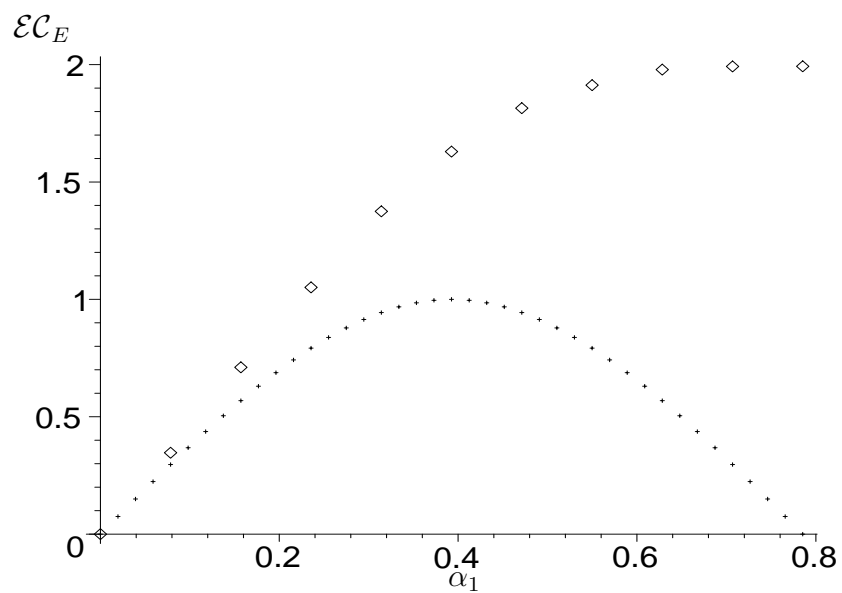


Figure 3.4: Single-copy entangling capacity for the SWAP family. Crosses are for no ancillas and diamonds are for one ancilla on each side.

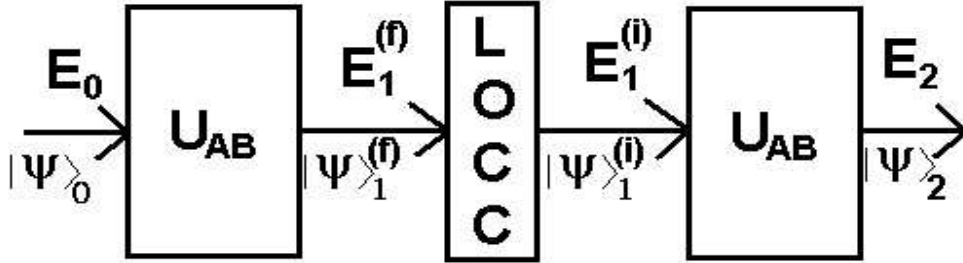


Figure 3.5: A general entanglement generation protocol when two applications of the unitary U_{AB} are available.

finite dimensional ancillary Hilbert spaces. We restrict our attention to the case where the state in the entire Hilbert space is pure at every stage of the protocol, but note that the results also hold for the case where mixed states are allowed [12]. In this setting, the unitaries may be applied simultaneously or one after another. Collective LOCC may be performed on all the qubits between applications and each unitary may be applied to an arbitrarily chosen pairs of qubits. However, all protocols of this form can be reduced to simpler protocols, which yield the same amount of entanglement.

First, observe that the effect of applying unitaries simultaneously can be achieved by applying them one after the other and doing nothing in between. Second, because local unitary operations (e.g. local SWAP operations) can be applied as part of the LOCC, all the unitaries can be applied to the same pair of qubits. Thus the problem reduces to a sequence of single-copy problems, where all the qubits that U_{AB} does not act on can be regarded as ancillas.

This is illustrated for the two-copy case in fig. (3.5). We start with a state $|\psi\rangle_0$ containing entanglement E_0 according to an entanglement measure E and apply U_{AB} to the chosen pair of qubits. This results in a state $|\psi\rangle_1^{(f)}$, with entanglement $E_1^{(f)}$ according to the same entanglement measure. This state can then be manipulated by LOCC to obtain a state $|\psi\rangle_1^{(i)}$ with entanglement

$E_1^{(i)}$ and then U_{AB} is applied to the chosen pair of qubits a second time yielding the state $|\psi\rangle_2$ with entanglement E_2 . The 2-copy entangling capacity will be defined as

$$\mathcal{EC}_E^{(2)}(U_{AB}) = \frac{1}{2} \max(E_2 - E_0) \quad (3.45)$$

where the maximum is taken over all possible starting states $|\psi\rangle_0$ and all possible LOCC operations that convert $|\psi\rangle_1^{(f)}$ to $|\psi\rangle_1^{(i)}$ with certainty. This definition generalizes to the n -copy entangling capacity $\mathcal{EC}_E^{(n)}$ in the obvious manner.

An upper bound for the entanglement generated by any 2-copy protocol can be found as follows

$$\begin{aligned} \Delta E &= E_2 - E_0 \\ &= (E_2 - E_1^{(i)}) + (E_1^{(i)} - E_1^{(f)}) + (E_1^{(f)} - E_0) \\ &\leq 2\mathcal{EC}_E(U_{AB}) \end{aligned} \quad (3.46)$$

The second line follows by adding and subtracting terms $E_1^{(f)}$ and $E_1^{(i)}$. The third line follows by noting that the terms $E_2 - E_1^{(i)}$ and $E_1^{(f)} - E_0$ are bounded from above by the single copy entangling capacity \mathcal{EC}_E and that $E_1^{(i)} - E_1^{(f)}$ can only be zero or negative due to the fact that entanglement cannot be increased by LOCC. This argument can easily be extended to non-deterministic protocols by considering the average entanglement generated by the protocol.

In this argument it was assumed that the 2-copy entangling capacity was defined in terms of the difference between the values of the same entanglement measure of the state at the output and input of the protocol. This is appropriate for the pure-state case as discussed here, but different entanglement measures may be appropriate if mixed states are allowed, as in §3.3.1. This issue is addressed in a generalization of this argument given in [12], which shows that the maximum of $2\mathcal{EC}_E^{(1)}$ still holds for the mixed states as well.

It is straightforward to see that this maximum can be achieved by acting with U_{AB} on 2 completely separate copies of the optimal single-copy input state, where each separate state contains the necessary number of ancillas.

The generalization to $n > 2$ copies of U_{AB} is also straightforward. Thus, $\mathcal{EC}_E^{(n)}(U_{AB}) = \mathcal{EC}_E^{(1)}(U_{AB})$ for all entanglement measures E and all bipartite unitary operators U_{AB} . This holds in the case where the most general initial resources are available, i.e. entanglement and ancillas.

If initial entanglement is not available then collective processing can do better per use of the unitary, since the first few copies of the unitary can be used to generate entanglement, which can then be used to make a state with optimal initial entanglement. This can then be used as the starting state for the subsequent copies, which can then be used to generate more entanglement than they could if the collective processing were not available.

Protocols that start with initial entanglement can outperform protocols that start with product states for all finite n . However, the asymptotic case, where $n \rightarrow \infty$, is more subtle. In the case where the starting state is a product, some of the first few operations can be used to generate the entanglement required for the optimal initial state. Then the entanglement of the states at each stage can be diluted so that the unitary always acts on the best initial state⁷. The number of operations required for the first stage of this protocol is fixed and finite, so as $n \rightarrow \infty$ the same entangling capacity will be achieved as if initial entanglement had been present. This means that asymptotic entangling capacity of a unitary starting with a product state is the same as the capacity that would be obtained starting with initial entanglement.

3.5 Conclusions

The results of this chapter show that, for all finite numbers of copies of U_{AB} , initial entanglement is required to achieve the optimal entangling capacity. If this initial entanglement and ancillas are available, then collective processing does not help to achieve this maximum.

⁷To achieve this it is actually necessary to take two limits. The number of product states needed at the beginning of the protocol must be large enough so that the dilution is efficient.

These results have implications for the asymptotic interconvertibility of bipartite unitary operations. For example, it is known that one can reversibly convert between a CNOT and a singlet state via LOCC [29, 40]. Thus, one can asymptotically simulate the action of $n\mathcal{EC}_E(U_{AB})$ CNOTs using n copies of U_{AB} and LOCC by generating entanglement and then distilling or diluting it to singlets. Further, it is impossible to generate more CNOTs than this, since otherwise one could generate more than $\mathcal{EC}_E(U_{AB})$ e-bits per application of U_{AB} by first converting to CNOTs and then using them to generate singlet states. More generally, it is not known whether converting between any unitary operation and entanglement via LOCC is reversible (i.e. whether one can asymptotically generate n copies of U_{AB} acting on an arbitrary input state given $n\mathcal{EC}_E(U_{AB})$ e-bits). However, $n\mathcal{EC}_E(U_{AB})$ is a lower bound on how much entanglement is needed to generate n copies of U_{AB} . Also, $\frac{\mathcal{EC}_E(U_1)}{\mathcal{EC}_E(U_2)}$ is an upper bound on how many copies of a bipartite unitary U_2 can be generated asymptotically per application of another bipartite unitary U_1 . Whether these bounds can be achieved remains an open question.

3.6 Related Work

Entanglement generation was first discussed in the context of the optimal rate of entanglement generated by an two-qubit interaction Hamiltonian [38]. It has also been discussed for two-qubit unitaries in the single copy case where no initial entanglement is allowed [60]. The work in this chapter can be viewed as a generalization of both these approaches. Since the publication of these results, a formula for the entanglement capacities of self-inverse [97] and product Hamiltonians in arbitrary dimensions [26] have been found in terms of the entangling capacity of $\sigma_1 \otimes \sigma_1$.

As indicated in the conclusion, entanglement generation from quantum operations is closely connected to other problems to do with non-local operations. One of the most interesting of these is the connection between the ability of

operations to generate entanglement and their ability to facilitate classical communication between the parties. There is a strong relationship between the general problem of entanglement generation with some initial entanglement and free classical communication and the problem of classical communication with some initial classical communication and free entanglement. This was first discussed in [12] and then in [16, 15].

Conditions under which a single-copy of a unitary operation may be used to simulate another with non-zero probability of success have been found [37, 35, 45]. Also, the inverse of entanglement generation, i.e. generating non-local unitaries from entanglement and LOCC has been investigated. This was first discussed in [29, 40]. The most general known protocol is given in [27]. It is a single-copy protocol, but it is not known whether it is optimal or whether it can be improved upon by collective processing.

Chapter 4

Measuring Polynomial Invariants of Multi-Party Quantum States

4.1 Introduction

The previous two chapters have focussed on the entanglement properties of Hamiltonians and unitary operators where both the operators and states involved are fully known to all the participants. In this chapter, a different connection between the entanglement of quantum states and quantum operations, specifically measurements, is investigated. The question of how, given several copies of an unknown but identically prepared quantum state, the entanglement properties of the state can be efficiently inferred is addressed. The focus here will be on multi-party ($n > 2$) states for which there is as yet no general scheme for the classification and quantification of entanglement. Nevertheless, as discussed in §1.3.3, progress can be made by considering invariants under local transformations such as Local Unitary (LU) transformations and Stochastic Local Operations and Classical Communication (SLOCC).

Invariants are rather abstract mathematical objects and it is natural to

ask whether any physical meaning can be given to them. One way of doing this is to investigate how these quantities might be measured. This could be done by simply measuring the coefficients of the state and then calculating the invariants. However, finding procedures to measure the invariants directly may be more efficient and also lends the invariants a physical interpretation as “collective observables” of the state.

For bipartite pure states, the Schmidt coefficients are a complete set of LU invariants and optimal protocols for measuring them were given in [3]. Also, in [51] a method was given for estimating the polynomial SLOCC invariants of a general two-qubit state.

In this chapter, networks for estimating two classes of polynomial invariants for multi-party states are presented: the LU invariants for multi-party states with arbitrary local Hilbert space dimension and the SLOCC invariants for multi-qubit states. In both cases, the protocol works for both pure and mixed states. In particular, the structure of the networks reflects the structure of the invariants in a very simple way.

In §4.2, the construction of local invariants under LU^1 in [18, 80]. In §4.3, the networks for measuring these invariants are presented. The next two sections address the same issues for invariants under SLOCC transformations, reviewing their construction in §4.4 and presenting networks to measure them in §4.5. In order to construct the networks for SLOCC invariants, the Structural Physical Approximation (SPA) to non-physical maps is used, which was introduced in [50]. The relevant details of this are presented in §4.6. In §4.7, estimation protocols based on the networks are evaluated by comparing them to simple techniques based on estimating the state coefficients. Some of the results from statistical inference used in §4.7 are reviewed in appendix §4.A and the integrals that arise in the same section are computed in appendix §4.B.

¹However, since this work first appeared it has been shown [21] that the SLOCC invariants can be estimated without the need for introducing noise and circuits have been explicitly constructed for the concurrence and 3-tangle. The statistical efficiency of these circuits has not yet been analyzed.

4.2 Polynomial Invariants under LU transformations

4.2.1 Pure states

Recall that two n -party pure states $|\psi\rangle, |\psi'\rangle \in \bigotimes_{j=1}^n \mathbb{C}^{d_j}$ are equivalent under LU transformations if

$$|\psi'\rangle = U_1 \otimes U_2 \otimes \dots \otimes U_n |\psi\rangle \quad (4.1)$$

where $U_j \in U(d_j)$ is a unitary operation acting on the Hilbert space of the j th party. States on the same orbit under this action have the same entanglement properties. Given a particular state, we might be interested in determining which orbit it belongs to. This can be done by establishing a canonical point on each orbit, such as the Schmidt form for bipartite states. However, canonical forms rapidly become more complicated as the number of parties is increased [1, 22]. Alternatively, polynomial functions of the state coefficients that are invariant on each orbit can be constructed. Theorems from invariant theory guarantee that a finite set of such polynomials is enough to distinguish the generic orbits under this action [76]. The construction of such a set is given in the following sections.

One party

Consider the state $|\psi\rangle = \sum_{i=1}^d \alpha^i |i\rangle$ in a single party Hilbert space \mathbb{C}^d , where $\{|i\rangle\}$ is an orthonormal basis. The only independent invariant under unitary transformations of this state is the norm $\langle\psi|\psi\rangle$. This may be written as

$$\langle\psi|\psi\rangle = \sum_i \alpha^i \alpha_i^* = \sum_{i,j} \alpha^i \delta_i^j \alpha_j^* \quad (4.2)$$

where δ_i^j is the Kronecker delta. δ_i^j is the $U(d)$ invariant tensor and invariants for larger numbers of parties are formed by similar contractions of the state coefficients with their complex conjugates.

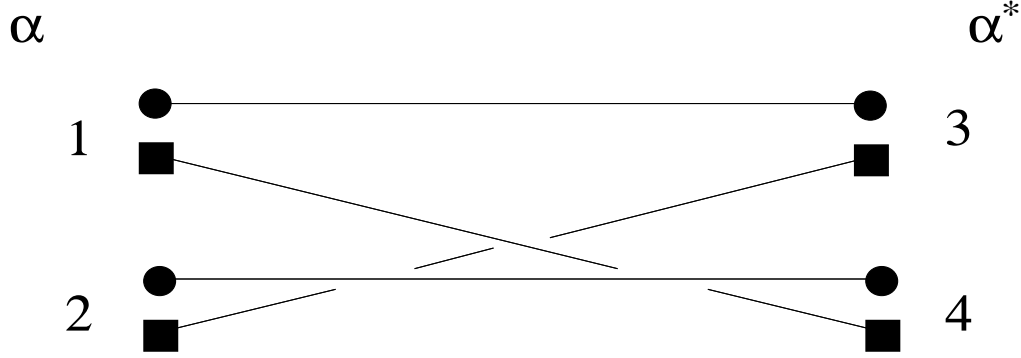


Figure 4.1: Diagrammatic representation of the quartic two-qubit LU invariant $J_{(\sigma, \tau)}$, given in eq.(4.3). The first index of each term is represented by a circle and the second by a square. A line joins indices that are contracted with a δ .

Two qubits

As an example, consider a two-qubit state $|\psi\rangle = \sum_{i,j=0}^1 \alpha^{ij} |ij\rangle$. There is only one independent quadratic invariant, which is simply the norm of the state. However, at quartic order there is another invariant, which is functionally independent of the norm given by

$$\begin{aligned} J &= \sum \alpha^{i_1 j_1} \alpha^{i_2 j_2} \delta_{i_1}^{i_3} \delta_{i_2}^{i_4} \delta_{j_1}^{j_3} \delta_{j_2}^{j_4} \alpha_{i_3 j_3}^* \alpha_{i_4 j_4}^* \\ &= \sum \alpha^{i_1 j_1} \alpha^{i_2 j_2} \alpha_{i_1 j_2}^* \alpha_{i_2 j_1}^* \end{aligned}$$

For two qubits, this is the only other independent invariant because every state has a canonical Schmidt form $|\psi\rangle = \sqrt{p} |00\rangle + \sqrt{1-p} |11\rangle$, with $1/2 \leq p \leq 1$ and $J = 2(p^2 - p) + 1$ determines p uniquely.

Another useful way of representing the invariant is to define two permutations σ, τ on the set $\{1, 2\}$ where σ is the identity permutation and $\tau(1) = 2, \tau(2) = 1$. Then

$$J_{(\sigma, \tau)} = \sum \alpha^{i_1 j_1} \alpha^{i_2 j_2} \alpha_{i_{\sigma(1)} j_{\tau(1)}}^* \alpha_{i_{\sigma(2)} j_{\tau(2)}}^* \quad (4.3)$$

This also suggests a diagrammatic way of representing the invariant (see fig. 4.1).

General case

A multipartite pure state can be written in terms of an orthonormal basis as follows

$$|\psi\rangle = \sum_{i,j,k,\dots} \alpha^{ijk\dots} |ijk\dots\rangle \quad (4.4)$$

A general polynomial function of the state coefficients and their complex conjugates can be written as

$$\sum c_{i_1 j_1 k_1 \dots i_2 j_2 k_2 \dots}^{i_r j_r k_r \dots} \alpha^{i_1 j_1 k_1 \dots} \alpha^{i_2 j_2 k_2 \dots} \dots \alpha_{i_r j_r k_r \dots}^* \dots \quad (4.5)$$

If the polynomial (4.5) has equal numbers of α 's and α^* 's and all the indices of the α 's are contracted using the invariant tensor δ with those of the α^* 's, each index being contracted with an index corresponding to the same party then the polynomial is manifestly invariant under LU transformations.

Such polynomials can be written in terms of permutations on the indices. Let r be the degree of the polynomial in α (and hence also the degree in α^*). Let σ, τ, μ, \dots be permutations acting on the set $\{1, 2, \dots, r\}$ and let $\vec{\sigma} = (\sigma, \tau, \mu, \dots)$. Then the invariants can be written as:

$$J_{\vec{\sigma}} = \sum \prod_{s=1}^r \alpha^{i_s j_s k_s \dots} \alpha_{i_{\sigma(s)} j_{\tau(s)} k_{\mu(s)} \dots}^* \quad (4.6)$$

In fact, σ can always be chosen to be the identity permutation by permuting the α terms in this expression. Additionally, each $J_{\vec{\sigma}}$ can be associated with a diagram constructed in the same way as fig.4.1.

The invariants $J_{\vec{\sigma}}$ are enough to completely distinguish the generic orbits under LU transformations. In fact, invariant theory guarantees that only a finite collection of them are needed to do this. However, except in a few simple cases, it is unknown which $J_{\vec{\sigma}}$ invariants form minimal complete sets.

4.2.2 Mixed states

Two mixed states ρ, ρ' are equivalent under LU transformations if

$$\rho' = U_1 \otimes U_2 \otimes \dots \otimes U_n \rho U_1^\dagger \otimes U_2^\dagger \otimes \dots \otimes U_n^\dagger \quad (4.7)$$

The LU invariants for mixed states can be derived by rewriting the pure state invariants (4.6) in terms of the density matrix $\rho = |\psi\rangle\langle\psi|$ and noting that the resulting expressions are still invariant under LU transformations for general density matrices. This can be done by noting that terms such as $\alpha^{i_1 j_1 \dots} \alpha_{i_2 j_2 \dots}^*$ are elements of the density matrix. A general density matrix may be written in terms of an orthonormal basis as

$$\rho = \sum \rho_{xyz\dots}^{ijk\dots} |ijk\dots\rangle\langle xyz\dots| \quad (4.8)$$

and the corresponding expression for an LU invariant is

$$J_{\vec{\sigma}} = \sum \prod_{s=1}^r \rho_{i_{\sigma(s)} j_{\tau(s)} k_{\mu(s)} \dots}^{i_s j_s k_s \dots} \quad (4.9)$$

4.3 Measuring Invariants under LU transformations

4.3.1 Network construction

The general construction of the network used to measure the LU invariants is shown in fig.4.2. It generalizes networks for estimating functionals of bipartite states given in [52, 41, 51]. Measuring an LU invariant of degree r in α (and also degree r in α^*) requires the collective processing of batches of r copies of the unknown state ρ . In addition, a Hadamard rotation H is applied to a single qubit in the state $|0\rangle$ to transform it to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. The next step consists of applying a unitary operation U on the r copies of ρ controlled by the Hadamard rotated qubit. Finally a measurement is performed on the single qubit in the $\{|0\rangle, |1\rangle\}$ basis. The expectation value of this measurement will be

$$\langle Z \rangle = \text{Re} (\text{Tr} (U \rho^{\otimes r})) \quad (4.10)$$

When $\rho = |\psi\rangle\langle\psi|$ is a pure state then this is equivalent to

$$\langle Z \rangle = \text{Re} \langle \psi |^{\otimes r} U | \psi \rangle^{\otimes r} \quad (4.11)$$

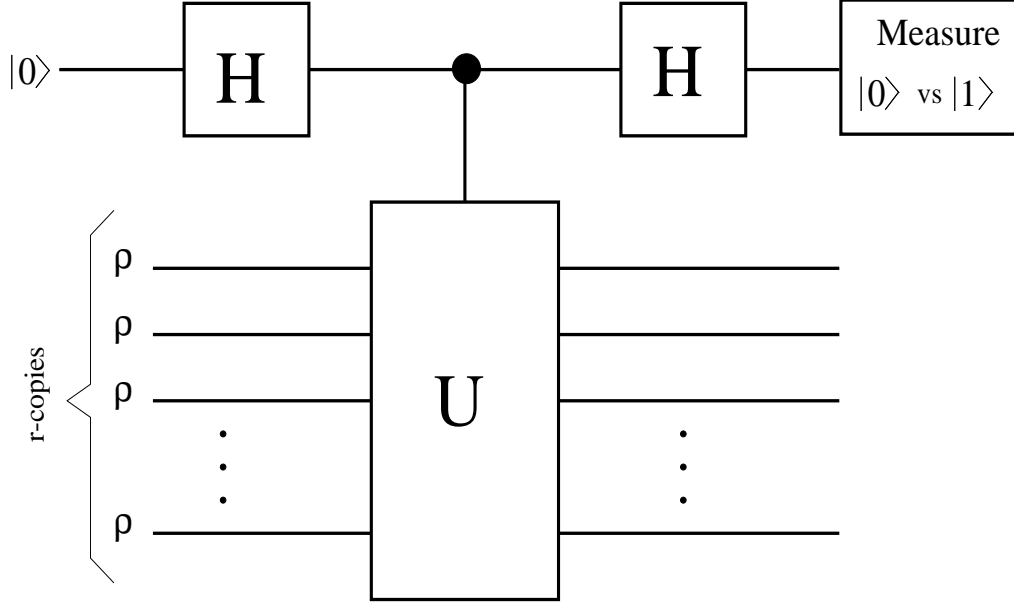


Figure 4.2: General construction of network to measure polynomial LU invariants.

In order to determine networks for measuring the LU invariants, it only remains to show that there is a U such that the invariants can be expressed in the form (4.10).

To do this for pure states, polynomials of the form (4.6) must be expressed in the form of (4.11). Firstly, note that $\alpha^{ijk\dots} = \langle ijk\dots | \psi \rangle$, $\alpha_{ijk\dots}^* = \langle \psi | ijk\dots \rangle$ and each permutation σ in (4.6) can be associated to a permutation matrix

$$P_\sigma = \sum_{i_1, i_2, \dots, i_r=1}^d |i_{\sigma(1)} i_{\sigma(2)} \dots i_{\sigma(r)}\rangle \langle i_1 i_2 \dots i_r| \quad (4.12)$$

where P_σ acts on the Hilbert space of the same party for each of the r copies of the state $|\psi\rangle$. Then to each $\vec{\sigma}$ the permutation matrix

$$P_{\vec{\sigma}} = P_\sigma \otimes P_\tau \otimes P_\mu \otimes \dots \quad (4.13)$$

is associated, where $P_\sigma, P_\tau, P_\mu, \dots$ act on the Hilbert space of the same party as σ, τ, μ, \dots in (4.6) on each of the r copies of the state. Then it will be shown that (4.6) can be written as

$$J_{\vec{\sigma}} = \langle \psi |^{\otimes r} P_{\vec{\sigma}} | \psi \rangle^{\otimes r} \quad (4.14)$$

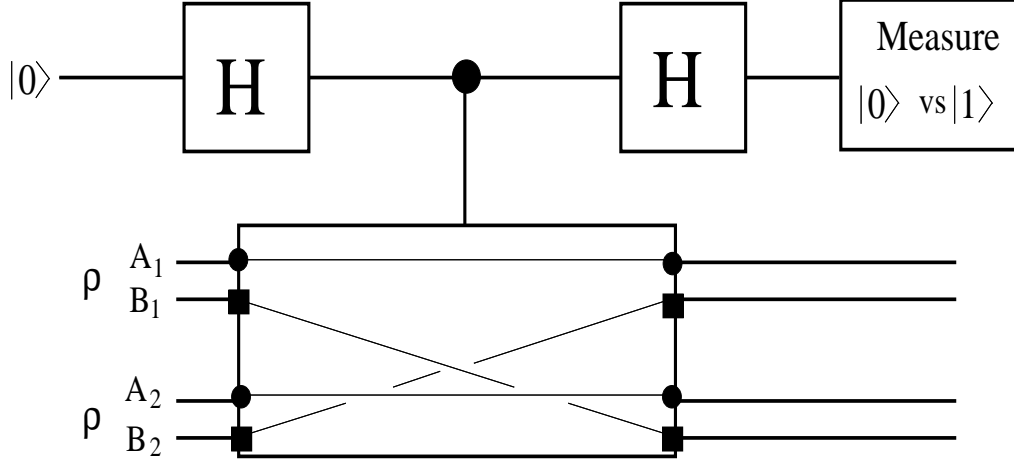


Figure 4.3: Network for measuring the 2-qubit quartic invariant.

Since $P_{\vec{\sigma}}$ is unitary these invariants can be estimated with the network in fig.4.2 by setting $U = P_{\vec{\sigma}}$ to obtain the real part and $U = iP_{\vec{\sigma}}$ to obtain the imaginary part. For the specific example of the 2-qubit invariant (4.3) this construction gives

$$J_{(\sigma,\tau)} = \langle \psi |_{A_1 B_1} \langle \psi |_{A_2 B_2} I_{A_1 A_2} \otimes SWAP_{B_1 B_2} | \psi \rangle_{A_1 B_1} | \psi \rangle_{A_2 B_2} \quad (4.15)$$

Note also that the physical construction of $P_{\vec{\sigma}}$ is closely related to the diagram associated with $J_{\vec{\sigma}}$ (compare figs. 4.1 and 4.3 for example).

If ρ is a mixed state then applying the same procedure without modification will give the invariants of eq. (4.9). These results can be summarized by the following theorem.

Theorem 4.1.

$$\text{Tr}(P_{\vec{\sigma}} \rho^{\otimes r}) = J_{\vec{\sigma}} \quad (4.16)$$

where $P_{\vec{\sigma}}$ is defined in (4.13) and $J_{\vec{\sigma}}$ is given by (4.9).

Proof. Firstly, consider the action of $P_{\vec{\sigma}}$ on the operator basis

$$\bigotimes_{s=1}^r (|i_s j_s k_s \dots\rangle \langle x_s y_s z_s \dots|).$$

$$\begin{aligned} P_{\vec{\sigma}} \left(\bigotimes_{s=1}^r |i_s j_s k_s \dots\rangle \langle x_s y_s z_s \dots| \right) \\ = \bigotimes_{s=1}^r (|i_{\sigma(s)} j_{\tau(s)} k_{\mu(s)} \dots\rangle \langle x_s y_s z_s \dots|) \end{aligned} \quad (4.17)$$

then using the relation $\text{Tr}(|\psi\rangle\langle\phi|) = \langle\phi|\psi\rangle$ gives

$$\begin{aligned} \text{Tr}(P_{\vec{\sigma}}(\bigotimes_{s=1}^r|i_s j_s k_s \dots\rangle\langle x_s y_s z_s \dots|)) \\ = \prod_{s=1}^r \left(\delta_{i_{\sigma(s)}}^{x_s} \delta_{j_{\tau(s)}}^{y_s} \delta_{k_{\mu(s)}}^{z_s} \dots \right) \end{aligned} \quad (4.18)$$

$\rho^{\otimes r}$ may be expanded in this operator basis as follows

$$\rho^{\otimes r} = \bigotimes_{s=1}^r \sum_{i_s, j_s, k_s, \dots, x_s, y_s, z_s, \dots} \rho_{x_s y_s z_s \dots}^{i_s j_s \dots} |i_s j_s k_s \dots\rangle\langle x_s y_s z_s \dots| \quad (4.19)$$

In this expression, all the sums may be taken in front of the tensor product and this results in a linear combination of the operator basis elements, with coefficients given by the corresponding density matrix elements. Then $\text{Tr}(P_{\vec{\sigma}}\rho^{\otimes r})$ is found by multiplying (4.18) by the coefficients of this linear combination and summing. Removing the contracted δ 's then gives (4.9). \square

It has previously been noted [44] that all homogeneous polynomial LU invariants are determined by the expectation values of two observables on r copies of a state. Here, an explicit network for measuring these observables has been given. Also, similar constructions can be made to estimate other polynomial functionals of quantum states [41] and these can be modified to enable the estimation to proceed by LOCC [4], i.e. with no collective operations over the n -parties. A similar modification would enable the LU invariants to be estimated by LOCC. However, these modifications are not considered here because they would affect the efficiency of the estimation to be discussed in §4.7.

4.4 Polynomial invariants under SLOCC

When attempting to classify entanglement, it is often useful to consider invariants under local transformations that are more general than unitary transformations. For this purpose, invariants under SLOCC have been introduced [13]. In §4.5, a network to measure the modulus squared of these invariants is constructed, for the case where each party has a single qubit (i.e. the Hilbert space is $(\mathbb{C}^2)^{\otimes n}$).

4.4.1 Pure states

Two n -party pure states $|\psi\rangle$ and $|\psi'\rangle$ are equivalent under SLOCC if it is possible to obtain $|\psi'\rangle$ with non-zero probability via a sequence of Local Operations and Classical Communication (LOCC) starting from a single copy of $|\psi\rangle$ and vice-versa. In [36], this criterion was shown to be equivalent to

$$|\psi'\rangle = M_1 \otimes M_2 \otimes \dots \otimes M_n |\psi\rangle \quad (4.20)$$

where $M_j \in GL(d_j)$ is an invertible linear transformation acting on the d_j -dimensional Hilbert space of the j th party.

In what follows, polynomial invariants for the special case where $M_j \in SL(2)$ are found, i.e. the transformation has unit determinant and each party has a single qubit. Networks to determine the modulus squared of these invariants will be given in §4.5. Note that it is not possible to measure $SL(2)^n$ invariants directly because they are not necessarily invariant under global phase transformations $|\psi\rangle \rightarrow e^{i\theta} |\psi\rangle$, which have no physical significance. It is for this reason that the modulus squared is measured, because this is invariant under these phase transformations.

Under general $GL(2)^n$ transformations, the polynomial $SL(2)^n$ invariants are still invariant up to a multiplicative factor, which is just some power of the determinant of $M_1 \otimes M_2 \otimes \dots \otimes M_n$. Thus, ratios of appropriate powers of these polynomials will be invariants under $GL(2)^n$.

Two qubits

In order to illustrate the polynomial invariants under $SL(2)^n$, first consider the case where $n = 2$. Two states $|\psi\rangle = \sum_{j,k=1}^2 \alpha^{jk} |jk\rangle$ and $|\psi'\rangle = \sum_{j,k=1}^2 \alpha'^{jk} |jk\rangle$ satisfy (4.20) if

$$\alpha' = M_1 \alpha M_2^T \quad (4.21)$$

This means that $\det(\alpha) = \det(\alpha')$ is an $SL(2) \times SL(2)$ invariant, since $\det(M_1) = \det(M_2) = 1$. This may be written as

$$\det \alpha = \sum \epsilon_{i_1 i_2} \epsilon_{j_1 j_2} \alpha^{i_1 j_1} \alpha^{i_2 j_2} \quad (4.22)$$

where the totally antisymmetric tensor ϵ_{ij} is the $SL(2)$ invariant tensor. For two qubit pure states, this is the only independent $SL(2) \times SL(2)$ invariant.

General case

The $SL(2)^n$ invariants can be constructed in a similar way to the LU invariants except the invariant tensor is now ϵ_{ij} , and α 's are contracted with α 's instead of α^* 's. Thus, polynomials of the form

$$K_{\vec{\sigma}} = \sum_1^2 \prod_{s=1}^{r/2} \epsilon_{i_{2s-1} i_{2s}} \epsilon_{j_{2s-1} j_{2s}} \epsilon_{k_{2s-1} k_{2s}} \cdots \alpha^{i_{\sigma(2s-1)} j_{\tau(2s-1)} k_{\mu(2s-1)} \cdots} \alpha^{i_{\sigma(2s)} j_{\tau(2s)} k_{\mu(2s)} \cdots} \quad (4.23)$$

are manifestly invariant. Note that it is straightforward to generalize this construction to the case where each party has a d -dimensional Hilbert space by contracting with the $SL(d)^n$ invariant tensor $\epsilon_{i_1 i_2 \dots i_d}$ instead of ϵ_{ij} . However, it is not yet clear how to measure these invariants because the effect of the higher rank ϵ tensors cannot be physically implemented by linear transformations on states.

4.4.2 Mixed states

In general, two mixed states ρ, ρ' are equivalent under SLOCC if there exists two completely positive maps $\mathcal{E}_1, \mathcal{E}_2$ which are implementable via LOCC with non-zero probability of success such that $\rho' = \mathcal{E}_1(\rho)$ and $\rho = \mathcal{E}_2(\rho')$. In order to derive invariants using the expressions from the previous section, only the case where ρ and ρ' are related by

$$\rho' = M_1 \otimes M_2 \otimes \dots \otimes M_n \rho M_1^\dagger \otimes M_2^\dagger \otimes \dots \otimes M_n^\dagger \quad (4.24)$$

with $M_j \in SL(2)$ is considered here. The resulting expressions may not be invariant under more general SLOCC transformations, but are related to important quantities in entanglement theory as described in §4.4.3

Unlike the LU invariants, it is not clear that (4.23) can be written simply in terms of the coefficients of the density matrix $\rho = |\psi\rangle\langle\psi|$. However, $|K_{\vec{\sigma}}|^2$ can be written as follows

$$|K_{\vec{\sigma}}|^2 = \sum_1^2 \prod_{s=1}^{r/2} \epsilon_{i_{2s-1}i_{2s}} \epsilon_{j_{2s-1}j_{2s}} \epsilon_{k_{2s-1}k_{2s}} \dots \epsilon^{x_{2s-1}x_{2s}} \epsilon^{y_{2s-1}y_{2s}} \epsilon^{z_{2s-1}z_{2s}} \rho_{x_{\sigma(2s-1)}y_{\tau(2s-1)}z_{\mu(2s-1)}} \dots \rho_{x_{\sigma(2s)}y_{\tau(2s)}z_{\mu(2s)}} \dots \quad (4.25)$$

and these will also be $SL(2)^n$ invariants for mixed states.

4.4.3 Examples of $SL(2)^n$ invariants

The $K_{\vec{\sigma}}$ invariants are especially interesting in entanglement theory because many important entanglement measures can be easily calculated from them. For example, in the case of two-qubits, the concurrence is a simple function of the eigenvalues of $\rho\tilde{\rho}$, where

$$\tilde{\rho} = \sigma_y \otimes \sigma_y \rho^T \sigma_y \otimes \sigma_y \quad (4.26)$$

and T stands for transpose in the computational basis. These eigenvalues can be calculated from $\text{Tr}((\rho\tilde{\rho})^m)$ for $m = 1, 2, 3, 4$, which are simply the moduli squared of $K_{\vec{\sigma}}$ invariants. In [51], networks were constructed to estimate these invariants for two qubits and this construction is generalized here to $K_{\vec{\sigma}}$ invariants for larger numbers of parties.

Another interesting example is the 3-tangle [102, 28], which is defined for pure states as the modulus of the following 3-qubit $K_{\vec{\sigma}}$ invariant.

$$\tau_3 = \sum_1^2 \alpha^{i_1j_1k_1} \alpha^{i_2j_2k_2} \epsilon_{i_1i_3} \epsilon_{j_1j_3} \epsilon_{k_1k_4} \epsilon_{i_2i_4} \epsilon_{j_2j_4} \epsilon_{k_2k_3} \alpha^{i_3j_3k_3} \alpha^{i_4j_4k_4} \quad (4.27)$$

The 3-tangle gives information about the genuine 3-party entanglement between the qubits.

Finally, note that the $K_{\vec{\sigma}}$ invariants can be given similar diagrammatic representations to the $J_{\vec{\sigma}}$ invariants. This is illustrated for the 3-tangle in fig.4.4.

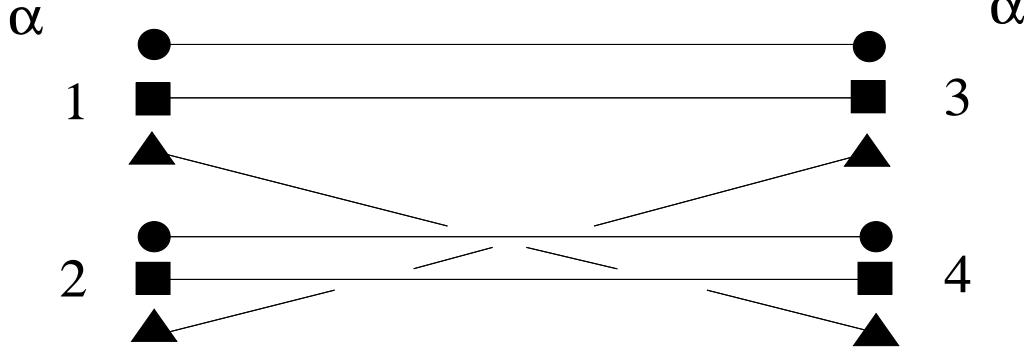


Figure 4.4: Diagrammatic representation of the 3-tangle. The first index of each term is represented by a circle, the second by a square and the third by a triangle. A line joins indices that are contracted with an ϵ .

4.5 Measuring SLOCC invariants

The modulus squared of the SLOCC invariants can be measured using a network similar to fig. 4.2 except that the unknown states ρ must be preprocessed prior to the controlled- U operation. If $K_{\vec{\sigma}}$ is of degree r in α then batches of r copies of ρ are collectively processed. The preprocessing stage will consist of collective unitary operations and completely positive maps that act on the entire Hilbert space of the r copies of ρ . The resulting state ρ' , will yield the expectation value

$$\langle Z \rangle = \text{Re}(\text{Tr}(U\rho')) \quad (4.28)$$

for the measurement at the end of the network. In this section, the preprocessing operations and unitary operations U that enable the modulus squared of the SLOCC invariants to be written in this form are described.

First, the inverse of the permutation matrix associated with $\vec{\sigma}$ is applied to the r copies of ρ to obtain $P_{\vec{\sigma}}^\dagger \rho^{\otimes r} P_{\vec{\sigma}}$.

The second, and final, part of the preprocessing stage is to apply a completely positive map $\bar{\Lambda}$ to the state. To describe $\bar{\Lambda}$, first define the multi-party analogue of eq. (4.26).

$$\tilde{\rho} = \sigma_y \otimes \sigma_y \otimes \dots \otimes \sigma_y \rho^T \sigma_y \otimes \sigma_y \otimes \dots \otimes \sigma_y \quad (4.29)$$

Next, define a map Λ that acts on a product of r states by applying the tilde

operation to the even numbered states as follows

$$\Lambda(\rho_1 \otimes \rho_2 \otimes \dots \otimes \rho_r) = \rho_1 \otimes \tilde{\rho}_2 \otimes \rho_3 \otimes \dots \otimes \tilde{\rho}_r \quad (4.30)$$

where each ρ_j is an n -party state.

Unfortunately, Λ cannot be physically implemented, since it is not a completely positive map. This can be dealt with by using the Structural Physical Approximation (SPA) to Λ , denoted by $\bar{\Lambda}$. $\bar{\Lambda}$ is the “closest” physical map to Λ . This is discussed in §4.6, but for now the network is constructed as if Λ could be implemented perfectly.

The final pre-processed state ρ' will be

$$\rho' = \Lambda(P_{\bar{\sigma}}^\dagger \rho^{\otimes r} P_{\bar{\sigma}}) \quad (4.31)$$

Next, the controlled- U operation in the network must be chosen such that $\langle Z \rangle = |K_{\bar{\sigma}}|^2$ when ρ' is used as the input. The pairwise SWAP gate, defined by

$$\begin{aligned} U |\phi_1\rangle \otimes |\phi_2\rangle \otimes \dots \otimes |\phi_{r-1}\rangle \otimes |\phi_r\rangle = \\ |\phi_2\rangle \otimes |\phi_1\rangle \otimes \dots \otimes |\phi_r\rangle \otimes |\phi_{r-1}\rangle \end{aligned} \quad (4.32)$$

where $|\phi_j\rangle$ is an n -party state fulfils this condition. These results can be summarized by the following theorem.

Theorem 4.2.

$$\text{Tr} \left(U \Lambda \left(P_{\bar{\sigma}}^\dagger \rho^{\otimes r} P_{\bar{\sigma}} \right) \right) = |K_{\bar{\sigma}}|^2 \quad (4.33)$$

where U is the pairwise SWAP operation (4.32), Λ is given by (4.30) and $|K_{\bar{\sigma}}|^2$ is given by (4.25).

Proof. The proof proceeds in a similar fashion to the proof of theorem 4.1 by considering the action on $\bigotimes_{s=1}^r |i_s j_s k_s \dots\rangle \langle x_s y_s z_s \dots|$ and then obtaining the expression for a general density matrix using the expansion (4.19) and the linearity of the operations in (4.33).

The action of the permutation operators is given by

$$\begin{aligned} P_{\bar{\sigma}}^\dagger \left(\bigotimes_{s=1}^r |i_s j_s k_s \dots\rangle \langle x_s y_s z_s \dots| \right) P_{\bar{\sigma}} \\ = \bigotimes_{s=1}^r |i_{\bar{\sigma}(s)} j_{\bar{\tau}(s)} k_{\bar{\mu}(s)} \dots\rangle \langle x_{\bar{\sigma}(s)} y_{\bar{\tau}(s)} z_{\bar{\mu}(s)} \dots| \end{aligned} \quad (4.34)$$

where $\bar{\cdot}$ denotes the inverse of a permutation.

The map Λ consists of two parts. Firstly, every second n -party state is transposed.

$$\begin{aligned} & \bigotimes_{s=1}^{r/2} \left(\left| i_{\bar{\sigma}(2s-1)} j_{\bar{\tau}(2s-1)} k_{\bar{\mu}(2s-1)} \dots \right\rangle \left\langle x_{\bar{\sigma}(2s-1)} y_{\bar{\tau}(2s-1)} z_{\bar{\mu}(2s-1)} \dots \right| \right. \\ & \quad \left. \otimes \left| x_{\bar{\sigma}(s)} y_{\bar{\tau}(s)} z_{\bar{\mu}(s)} \dots \right\rangle \left\langle i_{\bar{\sigma}(s)} j_{\bar{\tau}(s)} k_{\bar{\mu}(s)} \dots \right| \right) \end{aligned} \quad (4.35)$$

Next, every second n -party state is conjugated by $\sigma_2 \otimes \sigma_2 \otimes \dots \otimes \sigma_2$. Note that $(\sigma_2)_{ij} = -i\epsilon_{ij}$, which gives

$$\begin{aligned} & (-1)^{nr/2} \sum_{a,b,c,\dots,d,e,f,\dots} \bigotimes_{s=1}^{r/2} \epsilon_{x_{\bar{\sigma}(2s)} a_s} \epsilon_{y_{\bar{\tau}(2s)} b_s} \epsilon_{z_{\bar{\mu}(2s)} c_s} \dots \epsilon_{d_s i_{\bar{\sigma}(2s)}} \epsilon_{e_s j_{\bar{\tau}(2s)}} \epsilon_{f_s k_{\bar{\mu}(2s)}} \dots \\ & \quad \left(\left| i_{\bar{\sigma}(2s-1)} j_{\bar{\tau}(2s-1)} k_{\bar{\mu}(2s-1)} \dots \right\rangle \left\langle x_{\bar{\sigma}(2s-1)} y_{\bar{\tau}(2s-1)} z_{\bar{\mu}(2s-1)} \dots \right| \right. \\ & \quad \left. \otimes \left| a_s b_s c_s \dots \right\rangle \left\langle d_s e_s f_s \dots \right| \right) \end{aligned} \quad (4.36)$$

where the sum is taken over all subscripted indices labelled by $a, b, c, \dots, d, e, f, \dots$

The action of the pairwise SWAP (4.32) transforms the bracketed term into

$$\begin{aligned} & \left| a_s b_s c_s \dots \right\rangle \left\langle x_{\bar{\sigma}(2s-1)} y_{\bar{\tau}(2s-1)} z_{\bar{\mu}(2s-1)} \dots \right| \\ & \quad \otimes \left| i_{\bar{\sigma}(2s-1)} j_{\bar{\tau}(2s-1)} k_{\bar{\mu}(2s-1)} \dots \right\rangle \left\langle d_s e_s f_s \dots \right| \end{aligned} \quad (4.37)$$

Taking the trace then gives

$$\begin{aligned} & (-1)^{nr/2} \prod_{s=1}^{r/2} \epsilon_{x_{\bar{\sigma}(2s)} x_{\bar{\sigma}(2s-1)}} \epsilon_{y_{\bar{\tau}(2s)} y_{\bar{\tau}(2s-1)}} \epsilon_{z_{\bar{\mu}(2s)} z_{\bar{\mu}(2s-1)}} \dots \\ & \quad \epsilon_{i_{\bar{\sigma}(2s-1)} i_{\bar{\sigma}(2s)}} \epsilon_{j_{\bar{\tau}(2s-1)} j_{\bar{\tau}(2s)}} \epsilon_{k_{\bar{\mu}(2s-1)} k_{\bar{\mu}(2s)}} \dots \end{aligned} \quad (4.38)$$

Swapping the indices of the ϵ terms with x, y, z, \dots indices produces another factor $(-1)^{nr/2}$. Since r is even, this gives a factor of $(-1)^{nr} = +1$ at the front. Multiplying the resulting expression by the density matrix coefficients and then summing gives eq.(4.25). \square

4.6 The Structural Physical Approximation

The Λ operation encountered in the previous section is an example of a positive, but not completely positive map. These cannot be implemented exactly, but instead an approximation can be applied. Suppose $\Theta : \mathfrak{B}(\mathbb{C}^d) \rightarrow \mathfrak{B}(\mathbb{C}^d)$ is a

trace-preserving, positive map. Instead of Θ , a map of the following form can be applied.

$$\bar{\Theta}(\rho) = p \frac{I}{d} + (1-p)\Theta(\rho) \quad (4.39)$$

where I is the identity operator and p is a constant ($0 \leq p \leq 1$) chosen such that $\bar{\Theta}$ is completely positive. The optimal map of this form (i.e. the map with the smallest value of p) was found in [50] to be

$$\bar{\Theta}_{\text{opt}}(\rho) = \frac{\lambda d^2}{\lambda d^2 + 1} \frac{I}{d} + \frac{1}{\lambda d^2 + 1} \Theta(\rho) \quad (4.40)$$

where $\lambda = \max(0, -\lambda')$ and λ' is the smallest eigenvalue of the operator

$$\mathbb{I} \otimes \Theta(\Pi_+) \in \mathfrak{B}(\mathbb{C}^d \otimes \mathbb{C}^d) \quad (4.41)$$

Here, $\Pi_+ = |\phi^+\rangle\langle\phi^+|$ is the projector onto the maximally entangled state $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^d |jj\rangle$ and \mathbb{I} acts as the identity on $\mathfrak{B}(\mathbb{C}^d)$.

In order to apply this result to the map Λ , it is useful to decompose it into two parts $\Lambda = \Lambda_2 \circ \Lambda_1$. Λ_1 acts as the transpose on every second n -party state and the identity on the remainder. Λ_2 is given by conjugating every second n -party state with $\sigma_2 \otimes \sigma_2 \otimes \dots \otimes \sigma_2$. Since Λ_2 is unitary, it is Λ_1 that is relevant for calculating the optimal p . The smallest eigenvalue of the operator (4.41) when $\Theta = \Lambda_1$ is $2^{-nr/2}$. Applying the formalism, one finds that the optimal approximation to Λ is given by

$$\bar{\Lambda}(\rho) = \frac{2^{\frac{3}{2}nr}}{2^{\frac{3}{2}nr} + 1} \frac{I}{2^{nr}} + \frac{1}{2^{\frac{3}{2}nr} + 1} \Lambda(\rho) \quad (4.42)$$

where n is the number of qubits in each copy of the state and r is the degree of the $K_{\vec{\sigma}}$ for which the modulus squared is being estimated.

On replacing Λ with $\bar{\Lambda}$ in the network, the expectation value of the Z measurement still allows the modulus squared of the $K_{\vec{\sigma}}$ invariant to be determined via

$$|K_{\vec{\sigma}}|^2 = \left(2^{\frac{3}{2}nr} + 1\right) \langle Z \rangle - 2^{nr} \quad (4.43)$$

However, the SPA does affect the accuracy to which the invariant is determined. This is discussed further in the next section. Additionally, in [4], it is

shown that this sort of SPA can be replaced by one that can be implemented via LOCC. Thus, the SLOCC invariants could be estimated by LOCC, but the estimation efficiency discussed in §4.7 would be affected.

4.7 Evaluation

The main aim of the protocols presented in §4.3 and §4.5 is to provide a physical interpretation for the polynomial invariants. However, the question of how efficient these measurement protocols are has not yet been addressed. In this section, the efficiency of these protocols is compared to the efficiency of protocols based on simply measuring the state coefficients and calculating the invariants. Unbiased estimators based on counting [85, 64, 32] are used to estimate the invariants based on the data provided by a finite number of uses of the networks. Also, the same type of estimators can be used to find the state coefficients and then these estimates can be used to calculate the invariants. The estimators for the invariants arising from this second procedure will generally be biased. However, the analysis is performed in the limit where a large number of copies of the state have been measured, so that the bias is negligible and the variances of the estimates are small and can be treated to first order in all subsequent calculations. The techniques used are standard in experimental error analysis and are reviewed in appendix 4.A. Note that more sophisticated estimation procedures are also possible [42], but the purpose here is to compare the networks to methods that are easily accessible experimentally.

Measuring the state coefficients would clearly be a more straightforward procedure to perform experimentally than using the network. Although more parameters have to be determined, this does not necessarily mean that it is a less efficient method for estimating the invariants than using the networks. There are several quite general reasons why this might be the case.

Firstly, suppose that we are interested in measuring a complete set of polynomial LU invariants for some unknown state of n parties, where each party

has a d -dimensional Hilbert space. In general, it is not known how many would need to be measured, but parameter counting arguments [68, 23, 70] show that the number of local degrees of freedom is linear in n whereas the total number of degrees of freedom is exponential in n . Thus, for large n almost all the degrees of freedom are non-local. Even for moderately sized n , there are nearly as many invariants as there are state coefficients. In addition, the invariants are typically highly non-linear functions of the state coefficients. For these reasons, we expect that measuring a complete set of invariants directly will generally not be more efficient than measuring the state coefficients for large n . Similar considerations also apply to the SLOCC invariants.

Despite these considerations, it may be the case that the networks are more efficient if only a small incomplete subset of the invariants is measured. Also, they may be more efficient for estimating complete sets when n is small. For this reason, and for simplicity, the focus is on estimating two qubit invariants in this section.

There are also other reasons why the networks may not be efficient. For example, they only employ a two-outcome measurement for each r copies of the state whereas estimating the state coefficients uses a two-outcome measurement on each copy. Also, for the $K_{\vec{\sigma}}$ invariants, it is shown that using the SPA introduces a lot of noise into the measurement. Nonetheless, there are still some cases where using the networks is more efficient than estimating the state coefficients.

4.7.1 Statistical analysis of the network

For a particular setup in the network, repeated measurements of an observable Z , with expectation value $F = \text{Tr}(U\rho')$ are made. Z is a random variable²

²The statistical inference theory used in this section can be found in many statistics textbooks, such as [31].

with distribution

$$\begin{aligned} p(Z = +1) &= \frac{1}{2}(1 + F) \\ p(Z = -1) &= \frac{1}{2}(1 - F) \end{aligned} \tag{4.44}$$

If the event $Z = +1$ is defined as a success and if $p = P(Z = +1)$ then repeating the network N times is equivalent to performing N Bernoulli trials with parameter p . The number of successes N_s is a random variable with a binomial distribution and its expectation value is $\langle N_s \rangle = Np = \frac{N}{2}(1 + F)$. In an actual experiment, the observed number of successes \hat{N}_s can be used to compute an unbiased estimator for F , given by

$$\hat{F} = 2\frac{\hat{N}_s}{N} - 1 \tag{4.45}$$

with variance

$$\text{var}(\hat{F}) = \frac{1}{N}(1 - F^2) \tag{4.46}$$

We are interested in determining how many trials are needed in order for the estimate \hat{F} to be reasonably accurate. Specifically, we would like to quantify how many trials are needed to make $\text{var}(\hat{F}) \leq \epsilon$ for some $\epsilon > 0$. In an experimental situation, $\text{var}(\hat{F})$ could not be calculated from the data, so it would have to be estimated using the sample variance, $\hat{\text{var}}(\hat{F})$. However, in the limit $N \rightarrow \infty$ the fact that $\text{var}(\hat{F}) = O(N^{-1})$ and $\text{var}(\hat{\text{var}}(\hat{F})) = O(N^{-4})$ can be used, i.e. $\hat{\text{var}}(\hat{F})$ converges to the true variance much faster than \hat{F} converges to F so $\hat{\text{var}}(\hat{F}) \approx \text{var}(\hat{F})$. Thus, in this limit

$$N \gtrsim \frac{1}{\epsilon}(1 - F^2) \tag{4.47}$$

Recall that for the LU invariants, the real and imaginary parts of the invariant are estimated independently and that each use of the network requires r copies of the state, where r is the degree of the invariant in α . If the same number of samples are used for estimating both the real and imaginary parts then the total number of copies required is

$$M \gtrsim \frac{r}{\epsilon}(2 - |J_{\vec{\sigma}}|^2) \tag{4.48}$$

In some cases, it is known a priori that the invariant is always real or always imaginary. If this is the case, then the same accuracy can be achieved with

$$M \gtrsim \frac{r}{\epsilon} (1 - |J_{\bar{\sigma}}|^2) \quad (4.49)$$

For the SLOCC invariants, each use of the network requires r copies of the state, where r is the degree of the invariant in α . Also the estimate of the invariant must take into account the use of the SPA via (4.43). In this case, the total number of copies required is

$$M \gtrsim \frac{r}{\epsilon} \left[\left(2^{\frac{3}{2}nr} + 1 \right)^2 - (|K_{\bar{\sigma}}|^2 + 2^{nr})^2 \right] \quad (4.50)$$

Notice that the 2^{3nr} term will dominate the term in the square bracket for large n and r . This is due to the noise introduced into the measurement by the SPA.

4.7.2 Comparison to methods based on state estimation

In order to evaluate estimation protocols based on the networks, they are compared to methods based on estimating the density matrix of the state and then calculating the invariants. This can be done by estimating each state coefficient using observations on single copies of the state. This is known as quantum state tomography (see [42] for an overview and also [85, 64, 32]). This is not the optimal way of reconstructing the state in general [6], but it will greatly simplify the analysis.

Example: Two-qubit LU invariants

A general two-qubit density matrix can be written as

$$\rho = \frac{1}{4} \left(I_2 \otimes I_2 + \sum_j a_j \sigma_j \otimes I_2 + \sum_j b_j I_2 \otimes \sigma_j + \sum_{j,k} R_{jk} \sigma_j \otimes \sigma_k \right) \quad (4.51)$$

The two-qubit LU invariant (4.3) can be written in terms of these coefficients as

$$J = \text{Tr}(\rho_B^2) = \frac{1}{2} \left(1 + \sum_j b_j^2 \right) \quad (4.52)$$

Each b_j can be determined by simply performing a σ_j measurement on N_j copies of Bob's half of the state. The probability distributions of the associated random variables are given by

$$\begin{aligned} p(\sigma_j = +1) &= \frac{1}{2}(1 + b_j) \\ p(\sigma_j = -1) &= \frac{1}{2}(1 - b_j) \end{aligned} \quad (4.53)$$

Thus, each b_j can be estimated in the same way as F in (4.45) and

$$\text{var}(\hat{b}_j) = \frac{1 - b_j^2}{N_j} \quad (4.54)$$

Then, an estimator for J can be constructed, which is given by

$$\hat{J} = \frac{1}{2} \left(1 + \sum_j \hat{b}_j^2 \right) \quad (4.55)$$

which will be biased, but in the large N_j limit

$$\text{var}(\hat{J}) \approx \sum_j b_j^2 \left(\frac{1 - b_j^2}{N_j} \right) \quad (4.56)$$

to first order in $\text{var}(b_j)$.

If the additional restriction that each observable σ_j is sampled the same number of times (i.e. $N_j = \frac{N}{3}$) is made, then

$$N \gtrsim \frac{3}{\epsilon} \sum_j b_j^2 (1 - b_j^2) \quad (4.57)$$

for the estimate to have variance $\lesssim \epsilon$.

One way to compare this to the result for the network is to take an average over all pure states. If all pure states are equally likely then this amounts to integrating (4.7.2) and (4.49) using Haar measure (for details see appendix 4.B). This shows that on average 3/2 times as many copies of the state are needed if the coefficient estimation method is used. This is half of what one might expect from parameter counting alone, since three times as many parameters are estimated in the state coefficient method. The factor of two is explained by the fact that each use of the network uses two copies of the state³.

³and by the fact that the terms $1 - |J_{\sigma}|^2$ in (4.49) and $\sum_j b_j(1 - b_j^2)$ in (4.57) happen to have the same average when integrated over Haar measure. The result will not necessarily be so simple for other invariants.

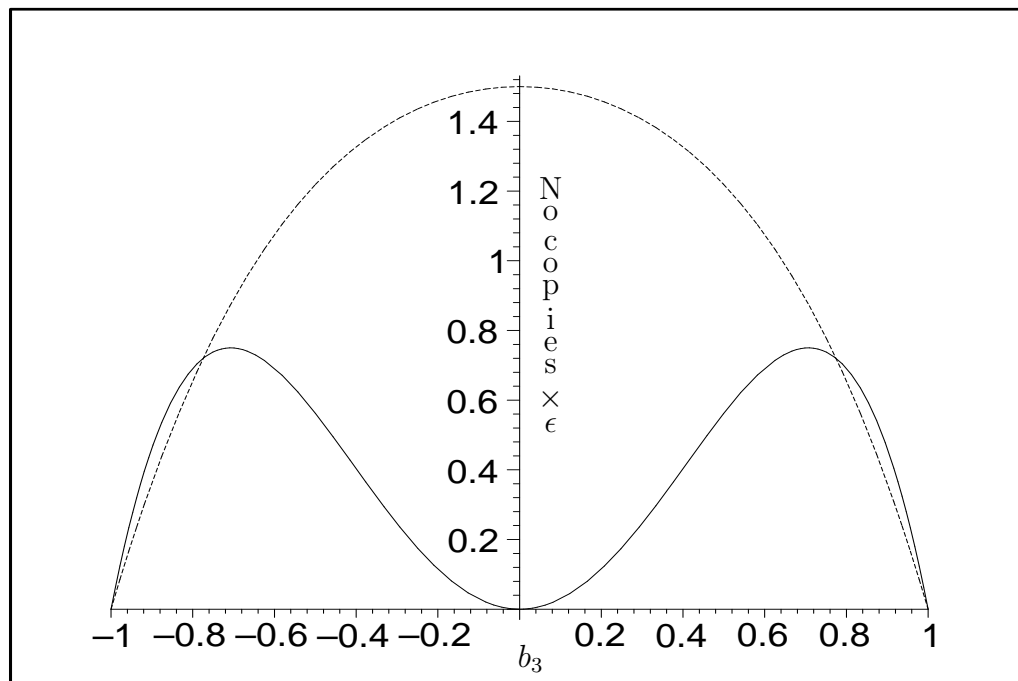


Figure 4.5: No. copies required $\times \epsilon$ in the asymptotic limit for the case where $b_1 = b_2 = 0$. The solid line is for the method based on estimating the state coefficients and the dashed line if for the network.

However, it is possible to find parameter ranges in which the state coefficient method performs better than the networks. One such range is given by setting $b_1 = b_2 = 0$ and then solving the RHS (4.49) $>$ RHS (4.57), which gives

$$\begin{aligned}
 2 \left(1 - \left(\frac{1}{2} (1 + b_3^2) \right)^2 \right) &> 3b_3^2 (1 - b_3^2) \\
 \Rightarrow 5b_3^4 - 8b_3^2 + 3 &> 0
 \end{aligned} \tag{4.58}$$

This is satisfied when either $b_3^2 > 1$ or $b_3^2 < \frac{3}{5}$. The first option violates the normalization condition, so $-\sqrt{\frac{3}{5}} < b_3 < \sqrt{\frac{3}{5}}$ is the only possible solution (see fig. 4.5). This illustrates the fact that parameter counting does not always reflect the statistical efficiency of a given protocol. Any partial information that available about the type of states being measured might change the judgement of which protocol is more efficient.

Example: Two-qubit SLOCC invariants

For the two-qubit SLOCC invariants the quadratic invariant (4.22) can be taken as an example. In terms of the decomposition (4.51) this can be written as

$$|K|^2 = \frac{1}{4} \left[1 - \sum_j (a_j^2 + b_j^2) + \sum_{jk} R_{jk}^2 \right] \quad (4.59)$$

If this is estimated by measuring all 15 of the state coefficients an equal number of times then, by a similar analysis to the LU case, at least

$$N \gtrsim \frac{15}{4\epsilon} \left[\sum_j [a_j^2(1 - a_j^2) + b_j(1 - b_j^2)] + \sum_{jk} R_{jk}^2(1 - R_{jk}^2) \right] \quad (4.60)$$

copies of the state are needed to get a variance $\lesssim \epsilon$.

Taking averages, one finds that fewer copies are needed in the state coefficient protocol by a factor $\approx 5 \times 10^3$ despite the fact that many more parameters have to be estimated in this protocol than when using the network. This is largely due to the factor 2^{12} that appears in (4.50), which arises from the noise introduced by the SPA. This suggests that other estimation and detection protocols based on the SPA [52, 51] may be less efficient than parameter counting arguments would imply. In fact, there are no states for which the network performs better than the coefficient estimation method. Even in the best possible case for the network, the state coefficient method requires fewer states by about 3 orders of magnitude.

4.8 Conclusions

In this chapter, networks for measuring the polynomial invariants of quantum states under LU and SLOCC transformations have been presented. The structure of these networks is closely related to the structure of the invariants themselves and thus gives the invariants a physical interpretation. Comparison of these networks with methods based on estimating the state coefficients indicate that the networks are of limited practical use for estimating complete

sets of invariants, although the LU networks may be useful for estimating small subsets of the invariants. Indeed, our results suggest that any estimation procedure that employs the SPA is statistically inefficient even when the number of parties is small⁴.

No procedure for estimating a complete set of invariants directly can outperform protocols based on estimating the state coefficients as the number of parties is increased. For small numbers of parties it seems that there can be some increase in efficiency, but the optimal protocol is not known in general.

4.A Statistical Inference

In this section, some details of statistical parameter estimation that are used in §4.7.1 are reviewed. Suppose an experiment is performed that has a set of possible outcomes Ω and that the outcome obtained depends in some way on an unknown parameter θ . By performing the experiment, some data $\mathbf{X} \in \Omega$ is collected that arises from a random process that depends on the parameter θ . Assuming that the distribution $p(\mathbf{X}|\theta)$ as a function of θ is known a-priori, an estimator $\hat{\theta}(\mathbf{X})$ can be constructed for θ that represents a guess of the value of θ given the data \mathbf{X} . The estimator is called unbiased if $\langle \hat{\theta} \rangle = \sum_{\mathbf{X} \in \Omega} p(\mathbf{X}|\theta) \hat{\theta} = \theta$. Generally, the bias of an estimator is defined to be

$$B(\hat{\theta}) = \langle \hat{\theta} \rangle - \theta \quad (4.61)$$

The central quantity of interest for deciding the quality of an estimator is the mean square error (MSE), defined as

$$\text{MSE}(\hat{\theta}) = \langle (\hat{\theta} - \theta)^2 \rangle \quad (4.62)$$

For an unbiased estimator, $\text{MSE}(\hat{\theta}) = \text{var}(\hat{\theta})$, so the mean square error and

⁴However, since this work first appeared it has been shown [21] that the SLOCC invariants can be estimated without the need for introducing noise and circuits have been explicitly constructed for the concurrence and 3-tangle. The statistical efficiency of these circuits has not yet been analyzed.

variance may be used interchangeably. In the general case,

$$\begin{aligned}
 \text{MSE}(\hat{\theta}) &= \langle (\hat{\theta} - \theta)^2 \rangle \\
 &= \langle \hat{\theta}^2 \rangle - 2\theta \langle \hat{\theta} \rangle + \theta^2 \\
 &= \langle \hat{\theta}^2 \rangle - \langle \hat{\theta} \rangle^2 + \langle \hat{\theta} \rangle^2 - 2\theta \langle \hat{\theta} \rangle + \theta^2 \\
 &= \text{var}(\hat{\theta}) + (\langle \hat{\theta} \rangle - \theta)^2 \\
 &= \text{var}(\hat{\theta}) + B^2(\hat{\theta})
 \end{aligned} \tag{4.63}$$

For the binomial distribution used in §4.7.1 $\hat{p} = \frac{\hat{N}_s}{N}$ is an unbiased estimator for the probability of success p . The estimators used to infer the invariants from the data produced by the networks are linear functions of an estimator of this sort and therefore they are also unbiased. The same applies to the estimators of the state coefficients. However, the estimators used to infer the invariants in the state coefficient method are non-linear functions of these estimators and will typically be biased.

Since asymptotic $N \rightarrow \infty$ properties of estimators are the main focus of interest, only terms to lowest order in N^{-1} will be retained. Suppose an estimator $\hat{\theta}$ is constructed as a function of some unbiased estimators, denoted by a vector $\hat{\mathbf{b}}$, i.e. $\hat{\theta} = g(\hat{\mathbf{b}})$. If g can be Taylor expanded about the true value \mathbf{b} then

$$\hat{\theta}(\hat{\mathbf{b}}) = g(\mathbf{b}) + \sum_j \frac{\partial g}{\partial b_j} (\hat{b}_j - b_j) + \frac{1}{2} \sum_{jk} \frac{\partial^2 g}{\partial b_j \partial b_k} (\hat{b}_j - b_j) (\hat{b}_k - b_k) + \dots \tag{4.64}$$

For all the estimators used in §4.7.1, only terms up to quadratic order in this expansion need to be retained because this is enough to determine the leading terms of the mean squared errors, variances and biases. The Taylor expansion of linear functions of binomial estimators will terminate after the linear term in any case. For functions of the state coefficient estimators $\text{var}(\hat{b}_j) = O(N^{-1})$. Also, any correlation terms, such as $\langle (\hat{b}_j - b_j) (\hat{b}_k - b_k) \rangle$ are zero, since the estimators of the state coefficients are unbiased and independent. Higher order moments about the mean are all $O(N^{-2})$ or smaller and so can be neglected.

Using (4.64) the mean square error is given to leading order by

$$\begin{aligned} \text{MSE}(\hat{\theta}) &= \sum_{jk} \frac{\partial g}{\partial b_j} \frac{\partial g}{\partial b_k} \langle (\hat{b}_j - b_j) (\hat{b}_k - b_k) \rangle + O(N^{-2}) \\ &= \sum_j \left(\frac{\partial g}{\partial b_j} \right)^2 \text{var}(\hat{b}_j) + O(N^{-2}) \end{aligned} \quad (4.65)$$

However, in §4.7.1 variances of estimators are used rather than mean square errors. This can be done because they are the same to leading order. To show this, (4.64) can be used to calculate the bias as follows.

$$\begin{aligned} B(\hat{\theta}) &= \langle \hat{\theta} - \theta \rangle \\ &= \sum_j \frac{\partial g}{\partial b_j} \langle \hat{b}_j - b_j \rangle + \frac{1}{2} \sum_{jk} \frac{\partial^2 g}{\partial b_j \partial b_k} \langle (\hat{b}_j - b_j) (\hat{b}_k - b_k) \rangle + O(N^{-2}) \\ &= \frac{1}{2} \sum_j \frac{\partial^2 g}{\partial b_j^2} \text{var}(\hat{b}_j) + O(N^{-2}) \end{aligned} \quad (4.66)$$

Thus, the B^2 term in (4.63) is $O(N^{-2})$ whereas the mean square error is $O(N^{-1})$. Thus, the variance must also be $O(N^{-1})$ and have the same leading term as the mean square error. This fact, combined with the result (4.65) is used to calculate all the asymptotic variances in §4.7.1.

4.B Integrals over Haar measure for two-qubits

To perform the integrals over Haar measure mentioned in §4.7.2 it is convenient to use the following parametrization of pure two-qubit density matrices [20].

$$\begin{aligned} \rho &= \frac{1}{4} \left(I \otimes I + \cos \alpha [\vec{r}_1 \cdot \vec{\sigma} \otimes I + I \otimes \vec{r}_2 \cdot \vec{\sigma}] \right. \\ &\quad \left. + \sin \alpha \cos \psi \left[\vec{k}_1 \cdot \vec{\sigma} \otimes \vec{k}_2 \cdot \vec{\sigma} - \vec{l}_1 \cdot \vec{\sigma} \otimes \vec{l}_2 \cdot \vec{\sigma} \right] \right. \\ &\quad \left. - \sin \alpha \sin \psi \left[\vec{k}_1 \cdot \vec{\sigma} \otimes \vec{l}_2 \cdot \vec{\sigma} - \vec{l}_1 \cdot \vec{\sigma} \otimes \vec{k}_2 \cdot \vec{\sigma} \right] \right) \end{aligned} \quad (4.67)$$

where $0 \leq \psi, \phi_1, \phi_2 < 2\pi$, $0 \leq \alpha, \theta_1, \theta_2 \leq \pi$ and

$$\begin{aligned} \vec{k}_j &= (\sin \phi_j, -\cos \phi_j, 0) \\ \vec{l}_j &= (\cos \theta_j \cos \phi_j, \cos \theta_j \sin \phi_j, -\sin \theta_j) \\ \vec{r}_j &= (\sin \theta_j \cos \phi_j, \sin \theta_j \sin \phi_j, \cos \theta_j) \end{aligned} \quad (4.68)$$

In this parametrization, the unitarily invariant integration measure is given by

$$\cos^2 \alpha \sin \alpha \sin \theta_1 \sin \theta_2 d\alpha d\psi d\phi_1 d\theta_1 d\phi_2 d\theta_2 \quad (4.69)$$

Details of how to find this measure are given in [20].

The integrals from §4.7.2 can be computed quite easily using this parametrization. For example, the number of copies required by the network to measure the two-qubit LU (4.49) becomes

$$M \gtrsim \frac{1}{2\epsilon} (3 - 2 \cos^2 \alpha - \cos^4 \alpha) \quad (4.70)$$

and the Haar integral of the right hand side is $\frac{24}{35\epsilon}$.

The number of copies required by the state coefficient method (4.57) is

$$N \gtrsim \frac{3}{\epsilon} \cos^2 \alpha (1 - \cos^2 \alpha (\sin^4 \theta_2 \cos^4 \phi_2 + \sin^4 \theta_2 \sin^4 \phi_2 + \cos^4 \theta_2)) \quad (4.71)$$

and this integrates to $\frac{36}{35\epsilon}$. The ratio N/M is $3/2$, as stated in §4.7.2.

In the parametrization (4.67), the two qubit SLOCC invariant takes the particularly simple form $K = 1 - \cos^2 \alpha$ and thus the number of copies required by the network (4.50) is

$$M \gtrsim \frac{2}{\epsilon} (16384 - (17 - \cos^2 \alpha)^2) \quad (4.72)$$

and the right hand side integrates to $\frac{1128048}{35\epsilon}$. Notice also that the minimum of (4.72) is $32190/\epsilon$.

The expression for the number of copies required in the state coefficient method (4.60) is rather complicated, but it integrates to $\frac{179}{28\epsilon}$. Also, (4.60) is upper bounded by $9/\epsilon$, which can be shown by imposing the condition $\text{Tr}(\rho^2) \leq 1$ and optimizing. Thus, there are no states for which our network requires less copies than the state coefficient method.

Chapter 5

Conclusions

In this thesis, I have investigated the fundamental limits to the processing of the entanglement properties of quantum operations and ways in which the entanglement of quantum states can be inferred by measurements. These fundamental limits are important because practical protocols for achieving these tasks can be evaluated by comparison to them. This area of research is still very much in its early stages of development and many open questions remain. Three of the most important ones are:

- Which classes of quantum operations can be reversibly converted into entangled states by LOCC? If all operations can be reversibly converted into entangled states, then many questions about the limits to processing entanglement in operations could be answered by the theory of entanglement in states. Conversely, any operations that cannot be reversibly converted into entangled states would represent a fundamentally new type of entanglement.
- Which protocols involving the manipulation of entanglement in states and operations are improved by collective processing? Given that collective processing is so important in the theory of entangled states, it would be surprising if some protocols for manipulating entanglement in operations could not be improved by using it.

- What is the optimal way of inferring the entanglement properties of unknown states and operations by measurements? In particular, detecting and measuring multi-particle entanglement will become increasingly important as experimentalists work towards large-scale quantum computers.

Additionally, it would be interesting to generalize simulation protocols and the conversion between operations and entanglement to more general operations, such as completely positive maps and POVMs. Some authors have begun work in these directions [27, 34, 81, 57].

I conclude with a few remarks on the possible application of the theory of entanglement in quantum operations.

The theory of quantifying entanglement in quantum states is ideally suited to determining the fundamental limits on quantum communication tasks, such as teleportation, superdense coding and cryptography. In these protocols, an entangled state is often being used directly as a resource.

In contrast, the fundamental resource in quantum control theory and computing is an operation, such as a Hamiltonian or a quantum gate. It is natural, therefore, to speculate that the quantification of entanglement in operations will have its most direct application in these areas. In particular, there are still many unanswered questions about which features of quantum mechanics are responsible for the power of quantum computing. Indeed, it can be argued that providing an explanation of this power is *the* fundamental problem in quantum computing theory, since it would lead to new insights into how to design algorithms that cannot be simulated efficiently on classical computers.

It is often stated that multi-party entanglement is one of the main features that is required in a quantum algorithm. However, at least in pure state models, it is known that multi-party entanglement in quantum states is only a necessary, but not sufficient requirement for algorithms to exhibit exponential speedup over their classical counterparts [58]. Since quantum gates are the fundamental resource used in quantum computing, it is possible that a neces-

sary and sufficient condition for exponential speedup could be formulated in terms of the entanglement properties of the gates involved in the algorithm as well as the states. Consequently, I believe that it is well worth pursuing further problems in this area.

Appendix A

Notation and Conventions

A few of the notations used throughout the thesis are summarized here.

A.1 Hilbert Spaces

- \mathcal{H} - A Hilbert space, usually of arbitrary dimension.
- \mathbb{C}^d - A finite dimensional, complex Hilbert space of dimension d .
- $\mathfrak{B}(\mathcal{H})$ - The space of linear operators on a Hilbert space \mathcal{H} .

Vectors in a Hilbert space are indicated with the Dirac notation $|\cdot\rangle$, their duals with $\langle\cdot|$ and inner products by $\langle\cdot|\cdot\rangle$. Subsystems are usually denoted on Hilbert spaces, vectors and operators by subscripts A,B,C,\dots . The label is often omitted if the subsystem referred to is clear from the context.

A.2 Bases

- Computational basis - The standard basis for \mathbb{C}^d , with basis vectors denoted by $|j\rangle$ for $j \in \mathbb{Z}$, $1 \leq j \leq d$. For \mathbb{C}^2 , $|0\rangle, |1\rangle$ is often used instead.

- Bell basis - A maximally entangled basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ given by

$$\begin{aligned}
 |\phi^-\rangle &= \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) \\
 |\phi^+\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \\
 |\psi^+\rangle &= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \\
 |\psi^-\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)
 \end{aligned} \tag{A.1}$$

$|\phi^+\rangle$ is also used to denote $\frac{1}{\sqrt{d}} \sum_{j=1}^d |jj\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$.

- Magic basis - Another maximally entangled basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$, often used to simplify calculations. It is given by

$$\begin{aligned}
 |\Phi_1\rangle &= \frac{-i}{\sqrt{2}} (|00\rangle - |11\rangle) \\
 |\Phi_2\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \\
 |\Phi_3\rangle &= \frac{-i}{\sqrt{2}} (|01\rangle + |10\rangle) \\
 |\Phi_4\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)
 \end{aligned} \tag{A.2}$$

A.3 Operators

- T - The transpose of an operator, usually taken in the computational basis.
- $*$ - The complex conjugate of an operator, usually taken in the computational basis.
- \dagger - The adjoint operator $A^\dagger = (A^*)^T$.
- ρ - A positive, trace class operator on \mathcal{H} with trace one. Usually referred to as a density operator.
- I - The identity operator on a Hilbert space \mathcal{H} , sometimes denoted I_N where N is the dimension of the space I_N acts on.
- σ_0 - An alternative notation for I_2

- $\sigma_1, \sigma_2, \sigma_3$ - The Pauli matrices, given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{A.3})$$

- $\vec{\sigma}$ - A vector consisting of the three Pauli matrices $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^T$.
- \mathbb{I} - The identity operator on $\mathfrak{B}(\mathcal{H})$.

Bibliography

- [1] A. Acín, A. Andrianov, E. Jane, and R. Tarrach. Three-qubit pure state canonical forms. *J. Phys. A*, 34:6725, 2001. quant-ph/0009107.
- [2] A. Acín, G. Vidal, and J. I. Cirac. On the structure of a reversible entanglement generating set for three-partite states. *Quant. Inf. Comp.*, 3:55, 2003. quant-ph/0202056.
- [3] Antonio Acín, Rolf Tarrach, and Guifré Vidal. Optimal estimation of two-qubit pure-state entanglement. *Phys. Rev. A*, 61:62307, 2000. quant-ph/9911008.
- [4] C. M. Alves, P. Horodecki, D. K. L. Oi, L. C. Kwek, and A. K. Ekert. Direct estimation of functionals of density operators by local operations and classical communication. quant-ph/0304123, 2003.
- [5] A. Barenco et al. Elementary gates for quantum computation. *Phys. Rev. A*, 52:3457, 1995. quant-ph/9503016.
- [6] O. E. Barndorff-Nielsen, R. Gill, and P. E. Jupp. On quantum statistical inference. to appear in *J. Royal Stat. Soc. B*, 2003.
- [7] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.
- [8] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 54:4707–4710, 1996. quant-ph/9511030.
- [9] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.
- [10] C. H. Bennett, J. I. Cirac, M. S. Leifer, D. W. Leung, N. Linden, S. Popescu, and G. Vidal. Optimal simulation of two-qubit Hamiltonians using general local operations. *Phys. Rev. A*, 66:012305, 2002. quant-ph/0107035.

-
- [11] C. H. Bennett, D. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824, 1996. quant-ph/9604024.
- [12] C. H. Bennett, A. Harrow, D. W. Leung, and J. A. Smolin. On the capacities of bipartite Hamiltonians and unitary gates. quant-ph/0205057, 2002.
- [13] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63:012307, 2001. quant-ph/9908073.
- [14] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881, 1992.
- [15] D. W. Berry and B. C. Sanders. Relation between classical communication capacity and entanglement capability for two-qubit operations. submitted to *Phys. Rev. A*, quant-ph/0207065, 2002.
- [16] D. W. Berry and B. C. Sanders. Relations for classical communication capacity and entanglement capability of two-qubit operations. *Phys. Rev. A*, 67:040302, 2003. quant-ph/0205181.
- [17] R. Bhatia. *Matrix Analysis*. Number 169 in Graduate Texts in Mathematics. Springer, 1997.
- [18] R. Brauer. On algebras which are connected with the semisimple continuous groups. *Annals of Mathematics*, 38:857, 1937.
- [19] M. J. Bremner, C. M. Dawson, J. L. Dodd, A. Gilchrist, A. W. Harrow, D. Mortimer, M. A. Nielsen, and T. J. Osborne. Practical scheme for quantum computation with any two-qubit entangling gate. *Phys. Rev. Lett.*, 89:247902, 2002. quant-ph/0207072.
- [20] V. Bužek et al. Quantum state reconstruction from incomplete data. *Chaos, Solitons and Fractals*, 10:981–1074, 1999. quant-ph/9805020.
- [21] H. A. Carteret. Noiseless circuits for the concurrence and residual 3-tangle. quant-ph/0309212, 2003.
- [22] H.A. Carteret, A. Higuchi, and A. Sudbery. Multipartite generalisation of the Schmidt decomposition. *J. Math. Phys.*, 41:7932–7939, 2000. quant-ph/0006125.
- [23] H.A. Carteret, N. Linden, S Popescu, and A. Sudbery. Multiparticle entanglement. *Foundations of Physics*, 29(4):527–552, 1999.

- [24] H.A. Carteret and A. Sudbery. Local symmetry properties of pure 3-qubit states. *J. Phys. A*, 33:4981–5002, 2000. quant-ph/0001091.
- [25] H. Chen. Necessary conditions for efficient simulation of Hamiltonians using local unitary transformations. *Quantum Information and Computation*, 3:249, 2003. quant-ph/0109115.
- [26] A. M. Childs, D. W. Leung, and G. Vidal. Reversible simulation of bipartite product Hamiltonians. quant-ph/0303097, 2003.
- [27] J.I. Cirac, W. Dür, B. Kraus, and M. Lewenstein. Entangling operations and their implementation using a small amount of entanglement. *Phys. Rev. Lett.*, 86:544, 2001. quant-ph/0007057.
- [28] Valerie Coffman, Joydip Kundu, and William K. Wootters. Distributed entanglement. *Phys. Rev. A*, 61:052306, 2000. quant-ph/990747.
- [29] D. Collins, N. Linden, and S. Popescu. Non-local content of quantum operations. *Phys. Rev. A*, 64, 2001. quant-ph/0005102.
- [30] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [31] D. R. Cox and D. V. Hinkley. *Theoretical Statistics*. Chapman and Hall, 1974.
- [32] G. M. d’Ariano, C. Macchiavello, and M. G. A. Paris. Detection of the density matrix through optical homodyne tomography without filtered back projection. *Phys. Rev. A*, 50:4298–4302, 1994.
- [33] J.L. Dodd, M.A. Nielsen, M.J. Bremner, and R.T. Thew. Universal quantum computation and simulation using any entangling Hamiltonian and local unitaries. *Phys. Rev. A*, 65:040301, 2001. quant-ph/0106064.
- [34] W. Dür and J. I. Cirac. Nonlocal operations: Purification, storage, compression, tomography, and probabilistic implementation. *Phys. Rev. A*, 64:012317, 2001.
- [35] W. Dür and J.I. Cirac. Equivalence classes of non-local unitary operations. *Quantum Information and Computation*, 2:240–254, 2002. quant-ph/0201112.
- [36] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, 2000. quant-ph/0005115.
- [37] W. Dür, G. Vidal, and J.I. Cirac. Optimal conversion of non-local unitary operations. *Phys. Rev. Lett.*, 89:057901, 2001. quant-ph/0112124.

-
- [38] W. Dür, G. Vidal, J.I. Cirac, N. Linden, and S. Popescu. Entanglement capabilities of non-local Hamiltonians. *Phys. Rev. Lett.*, 87:137901, 2001. quant-ph/0006034.
- [39] B. Einstein, A. Podolsky and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.
- [40] J Eisert, K. A. Jacobs, P. Papadopoulos, and M. B. Plenio. Optimal local implementation of nonlocal quantum gates. *Phys. Rev. A*, 62:052317, 2000. quant-ph/0005101.
- [41] Artur K. Ekert et al. Direct estimations of linear and nonlinear functionals of a quantum state. *Phys. Rev. Lett.*, 88:217901, 2002. quant-ph/0203016.
- [42] R. Gill and M. I. Guta. An invitation to quantum tomography. submitted to J. Royal Stat. Soc. B, quant-ph/0303020, 2003.
- [43] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, CalTech, 2000. quant-ph/975052.
- [44] Markus Grassl, Martin Rötteler, and Thomas Beth. Computing local invariants of qubit systems. *Phys. Rev. A*, 58:1833–1839, 1998. quant-ph/9712040.
- [45] K. Hammerer, G. Vidal, and J. I. Cirac. Characterization of non-local gates. *Phys. Rev. A*, 66:062321, 2002. quant-ph/0205100.
- [46] H. L. Haselgrove, M. A. Nielsen, and T. J. Osborne. On the practicality of time-optimal two-qubit Hamiltonian simulation. quant-ph/0303070, 2003.
- [47] S. Hill and W. K. Wootters. Entanglement of a pair of quantum bits. *Phys. Rev. Lett.*, 78:5022, 1997. quant-ph/9703041.
- [48] M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-state entanglement and distillation: is there a “bound” entanglement in nature? *Phys. Rev. Lett.*, 80:5239, 1998. quant-ph/9801069.
- [49] M. Horodecki, P. Horodecki, and R. Horodecki. Limits for entanglement measures. *Phys. Rev. Lett.*, 84(9):2014–2016, 2000.
- [50] Pawel Horodecki. From limits of quantum nonlinear operations to multicopy entanglement witnesses and state spectrum estimation. quant-ph/0111036, 2001.
- [51] Pawel Horodecki. Measuring quantum entanglement without prior state reconstruction. quant-ph/0111064, 2001.

- [52] Pawel Horodecki and Artur Ekert. Method for direct detection of quantum entanglement. *Phys. Rev. Lett.*, 89, 2002. quant-ph/0111064.
- [53] G. Jaeger, A. V. Sergienko, B. E. A. Saleh, and M. C. Teich. Entanglement, mixedness, and spin-flip symmetry in multiple-qubit systems. 2003.
- [54] G. Jaeger, A. Teodorescu-Frumosu, M. Sergienko, B. A. E. Saleh, and M. C. Teich. Multiphoton Stokes-parameter invariant for entangled states. *Phys. Rev. A*, 67:032307, 2003. quant-ph/0301128.
- [55] D. Jozsa, P. Wocjan, and Th. Beth. Complexity of decoupling and time-reversal for n spins with pair-interactions. *Phys. Rev. A*, 66:042311, 2002. quant-ph/0106085.
- [56] D Jonathan and M. B. Plenio. Entanglement-assisted local manipulation of pure quantum states. *Phys. Rev. Lett.*, 83:3566–3569, 1999.
- [57] R. Jozsa et al. Entanglement cost of generalised measurements. quant-ph/0303167, 2003.
- [58] R. Jozsa and N. Linden. On the role of entanglement in quantum computational speed-up. quant-ph/0201143, 2002.
- [59] N. Khaneja, R. Brockett, and S. J. Glaser. Time optimal control in spin systems. *Phys. Rev. A*, 63:032308 1–13, Feb. 2001.
- [60] B. Kraus and J.I. Cirac. Optimal creation of entanglement using a two-qubit gate. *Phys. Rev. A*, 63:062309, 2001. quant-ph/0011050.
- [61] B Kraus, K. Hammerer, G. Giedke, and J. I. Cirac. Entanglement generation and Hamiltonian simulation in continuous variable systems. *Phys. Rev. A*, 67:042314, 2003. quant-ph/0210136.
- [62] M. S. Leifer, L. Henderson, and N. Linden. Optimal entanglement generation from quantum operations. *Phys. Rev. A*, 67:012306, 2002. quant-ph/0205055.
- [63] M. S. Leifer, N. Linden, and A. Winter. Measuring polynomial invariants of multi-party quantum states. quant-ph/0308008, submitted to *Phys. Rev. A*, 2003.
- [64] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, 1997.
- [65] D. Leung. Simulation and reversal of n-qubit Hamiltonians using Hadamard matrices. *Journal of Modern Optics*, 49:1199–1217, 2002. quant-ph/0107041.

-
- [66] D. Leung and C. H. Bennett. private communication, 2002.
- [67] N. Linden, S. Massar, and S. Popescu. Purifying noisy entanglement requires collective measurements. *Phys. Rev. Lett.*, 81:3279, 1998. quant-ph/9805001.
- [68] N. Linden and S. Popescu. On multi-particle entanglement. *Fortsch. Phys.*, 46:567–578, 1998. quant-ph/9711016.
- [69] N Linden, S Popescu, B Schumacher, and M Westmoreland. Reversibility of local transformations of multiparticle entanglement. quant-ph/9912039, 1999.
- [70] N. Linden, S. Popescu, and A. Sudbery. Non-local properties of multiparticle density matrices. *Phys. Rev. Lett.*, 83:243–247, 1999. quant-ph/9801076.
- [71] Y. Makhlin. Nonlocal properties of two-qubit gates and mixed states and optimization of quantum computations. *Quantum Information Processing*, 1:243–252, 2002. quant-ph/0002045.
- [72] S. Massar and S. Popescu. How much information can be obtained by a quantum measurement? *Phys. Rev. A*, 61:062303, 2000. quant-ph/9907066.
- [73] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83(2):436–439, 1999.
- [74] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [75] M.A. Nielsen, M.J. Bremner, J.L. Dodd, C.M. Childs, and C.M. Dawson. Universal simulation of Hamiltonian dynamics for quantum systems with finite-dimensional state spaces. *Phys. Rev. A*, 66:022317, 2002. quant-ph/0109064.
- [76] P. J. Olver. *Equivalence, Invariants and Symmetry*. Cambridge University Press, 1995.
- [77] S. Popescu and D. Rohrlich. Thermodynamics and the measure of entanglement. *Phys. Rev. A*, 56:R3319, 1997. quant-ph/9610044.
- [78] J. Preskill. Lecture notes on quantum information and computation. Course notes for Caltech lecture course <http://www.theory.caltech.edu/people/preskill/ph229/#lecture>.
- [79] E. M. Rains. Rigorous treatment of distillable entanglement. *Phys. Rev. A*, 60:173–178, 1999. quant-ph/9809078.

- [80] E. M. Rains. Polynomial invariants of quantum codes. *IEEE Trans. Info. Theory*, 46:54, 2000. quant-ph/9704042.
- [81] Benni Reznik. Remote generalized measurements (POVMs) require non-maximal entanglement. quant-ph/0203055, 2002.
- [82] E. Schmidt. Zur theorie der linearen und nichtlinearen integralgleichungen. i. teil: Entwicklung willkürlicher funktionen nach systemen vorgeschriebener. *Math. Ann.*, 63:433, 1907.
- [83] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, 1995.
- [84] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26:1484–1509, 1997.
- [85] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography. *Phys. Rev. Lett.*, 70:1244–1247, 1993.
- [86] A. Sudbery. On local invariants of pure three-qubit states. *J. Phys. A*, 34:643–652, 2001. quant-ph/0001116.
- [87] M. Teodorescu-Frumosu and G. Jaeger. Quantum Lorentz-group invariants of n -qubit systems. *Phys. Rev. A*, 67:052305, 2003.
- [88] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying entanglement. *Phys. Rev. Lett.*, 78:2275–2279, 1997. quant-ph/9702027.
- [89] F. Verstraete, J. Dehaene, and B. De Moor. Normal forms, entanglement monotones and optimal filtration of multiparticle quantum systems. quant-ph/0105090, 2001.
- [90] F. Verstraete, J. Dehaene, and B. De Moor. The Lorentz singular value decomposition and its applications to pure states of 3 qubits. *Phys. Rev. A*, 65:032308, 2002. quant-ph/0108043.
- [91] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde. Four qubits can be entangled in nine different ways. *Phys. Rev. A*, 65:052112, 2002. quant-ph/0109033.
- [92] G. Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.*, 83:1046–1049, 1999. quant-ph/9902033.
- [93] G. Vidal and J. Cirac. Catalysis in non-local quantum operations. *Phys. Rev. Lett.*, 88:167903, 2001. quant-ph/0108077.
- [94] G. Vidal and J. Cirac. Optimal simulation of nonlocal Hamiltonians using local operations and classical communication. *Phys. Rev. A*, 66:022315, 2002. quant-ph/0108076.

- [95] G. Vidal, K. Hammerer, and J.I. Cirac. Interaction cost of non-local gates. *Phys. Rev. Lett.*, page 237902, 2002. quant-ph/0112168.
- [96] Guifré Vidal. Entanglement monotones. *Journal of Modern Optics*, 47:355–376, 2000.
- [97] X. Wang and Sanders B. C. Entanglement capability of self-inverse Hamiltonians. quant-ph/0212035, 2002.
- [98] P. Wocjan, D. Janzing, and Th. Beth. Simulating arbitrary pair-interactions by a given Hamiltonian. *Quantum Information and Computation*, 2:117, 2002. quant-ph/0106077.
- [99] P. Wocjan, M. Roetteler, D. Janzing, and Th. Beth. Simulating Hamiltonians in quantum networks. *Phys. Rev. A*, 65:042309, 2002. quant-ph/0109088.
- [100] P. Wocjan, M. Roetteler, D. Janzing, and Th. Beth. Universal simulation of Hamiltonians using a finite set of control operations. *Quantum Information and Computation*, 2:133, 2002. quant-ph/0109063.
- [101] W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245, 1998. quant-ph/9709029.
- [102] W. K. Wootters. Quantum entanglement as a quantifiable resource. *Phil. Trans. of the Royal Society*, 356(1743):1717–1731, 1998.
- [103] P. Zanardi. Entanglement of quantum evolutions. *Phys. Rev. A*, 63:040304(R), 2001. quant-ph/0010074.
- [104] G. M. Ziegler. *Lectures on polytopes*. Graduate Texts in Mathematics. Springer, 1994.