Entanglement Witnesses with Tensor Networks

Characterizing entanglement in large systems

Aluno:
Lucas Vieira Barbosa

Orientador:
Prof. Dr. Reinaldo O. Vianna

Co-orientador:
Thiago O. Maciel

Dissertação apresentada à
UNIVERSIDADE FEDERAL DE MINAS GERAIS - UFMG
como requisito parcial para a obtenção do grau de
MESTRE EM FÍSICA

Belo Horizonte, MG, Brasil
2019–11–13

Dados Internacionais de Catalogação na Publicação (CIP)

B238e Barbosa, Lucas Vieira.
Entanglement witnesses with tensor networks: characterizing entanglement in large systems / Lucas Vieira Barbosa. – 2019.
81f., enc.: il.

Orientador: Reinaldo Oliveira Vianna.
Coorientador: Thiago Oliveira Maciel
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Departamento de Física.
Bibliografia: f. 74-81.

1. Informação quântica. 2. Teoria quântica. 3. Emaranhamento, teses. I. Título. II. Vianna, Reinaldo Oliveira. III. Universidade Federal de Minas Gerais, Departamento de Física.

CDU – 530.145(043)

Elaborada pela Biblioteca Professor Manoel Lopes de Siqueira da UFMG.

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Programa de Pós-Graduação em Física
Caixa Postal 702
30.123-970 Belo Horizonte - MG - Brasil

Telefone (xx) (31) 3499 5637
(xx) (31) 3499 5633
Fax (xx) (31) 3499 5688
(xx) (31) 3499 5600
e-mail pgfisica@fisica.ufmg.br

ATA DA SESSÃO DE ARGUIÇÃO DA 625ª DISSERTAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA DEFENDIDA POR LUCAS VIEIRA BARBOSA, orientado pelo professor Reinaldo Oliveira Vianna e coorientado pelo doutor Thiago Oliveira Maciel para obtenção do grau de **MESTRE EM FÍSICA**. Às 09:00 horas de dois de agosto de 2019, na sala 4129 do Departamento de Física da UFMG, reuniu-se a Comissão Examinadora, composta pelos professores **Reinaldo Oliveira Vianna** (Orientador - DF/UFMG), **Carlos Henrique Monken** (Departamento de Física/UFMG) e **Walber Hugo de Brito** (Departamento de Física/UFMG) para dar cumprimento ao Artigo 37 do Regimento Geral da UFMG, submetendo o bacharel **LUCAS VIEIRA BARBOSA** à arguição de seu trabalho de dissertação, que recebeu o título de **"Entanglement Witnesses with Tensor Networks: Characterizing entanglement in large systems"**. Às 14:00 horas do mesmo dia o candidato fez uma exposição oral de seu trabalho durante aproximadamente 50 minutos. Após esta, os membros da comissão prosseguiram com a sua arguição e apresentaram seus pareceres individuais sobre o trabalho, concluindo pela aprovação do candidato.

Belo Horizonte, 02 de agosto de 2019.

Prof. Reinaldo Oliveira Vianna
Orientador do estudante
Departamento de Física/UFMG

Prof. Carlos Henrique Monken
Departamento de Física/UFMG

Prof. Walber Hugo de Brito
Departamento de Física/ UFMG

Candidato

Entanglement Witnesses with Tensor Networks

Characterizing entanglement in large systems

Lucas Vieira Barbosa

2019–11–13

Resumo

Emaranhamento é certamente um dos fenômenos mais fascinantes observados na Natureza, e mesmo após de décadas de pesquisas na teoria de emaranhamento ainda há muito a ser descoberto e entendido. Em particular, ainda há poucos resultados detalhando e caracterizando a estrutura do emaranhamento em sistemas fortemente correlacionados em larga escala, envolvendo um número grande de subsistemas.

A forma mais comum de se estudar detalhadamente a estrutura de emaranhamento em estados quânticos tem sido através da otimização de Testemunhas de Emaranhamento (*Entanglement Witnesses*), otimizadas utilizando-se técnicas algorítmicas como Programação Semidefinida (*Semidefinite Programming*) aplicada à matrizes. No entanto, esta descrição matricial torna a técnica computacionalmente inviável para sistemas em larga escala devido ao crescimento exponencial da dimensão do espaço de parâmetros a serem otimizados. Todas as técnicas se tornam inviáveis para um número relativamente pequeno de partículas (na ordem de $n \approx 10^1$).

Por estes motivos, Redes Tensoriais (*Tensor Networks*) têm atraído a atenção de pesquisadores nas últimas décadas por serem formas eficientes de descrever e simular sistemas quânticos compostos de muitas partes, pois são descrições mais naturais e eficientes das correlações quânticas entre os subsistemas.

No entanto, pouco tem sido feito para caracterização detalhada do emaranhamento em tais sistemas, e a literatura ainda carece de uma ponte entre os dois formalismos. Neste trabalho, propomos um primeiro pequeno passo rumo ao objetivo de adaptar o formalismo de testemunhas de emaranhamento de forma compatível com a descrição de sistemas quânticos através de redes tensoriais, tornando possível a caracterização de emaranhamento em sistemas de larga escala.

Palavras-chave

emaranhamento, emaranhamento em estados puros, testemunhas de emaranhamento, testemunhas decomponíveis, transposição parcial negativa, NPT, programação semidefinida, SDP, redes tensoriais, Matrix Product State, MPS, Matrix Product Operator, MPO, Density Matrix Renormalization Group, DMRG, transposição parcial, entropia de emaranhamento

Abstract

Entanglement is certainly one of the most fascinating phenomena observed in Nature, and even after decades of research in entanglement theory there is still much to be discovered and understood. In particular, there are still few results detailing and characterizing the entanglement structure in strongly correlated systems on a large scale, involving a large number of subsystems.

The most typical approach for studying in detail the entanglement structure in quantum states has been through the optimization of Entanglement Witnesses, optimized by means of algorithmic techniques such as Semidefinite Programming applied to matrices. However, this matrix description renders the technique computationally nonviable for large scale systems due to the exponential growth of the dimension of the optimization parameters. All these techniques are nonviable for a relatively small number of particles (on the order of $n \approx 10^1$).

For these reasons, Tensor Networks have attracted the attention of researchers over the last decades for being more efficient ways to describe and simulate quantum systems composed of many parts, as they are more natural and efficient descriptions of the quantum correlations between subsystems.

However, few progress has been made for the detailed characterization of entanglement in such systems, and the literature still hasn't fully bridged the gap between the two formalisms. In this work, we propose a small step towards the goal of adapting the formalism of entanglement witnesses in a way which is compatible with the description of quantum systems by means of tensor networks, making it possible the characterization of entanglement in large scale systems.

Keywords

entanglement, entanglement in pure states, entanglement witnesses, decomposable witnesses, negative partial transpose, NPT, semidefinite programming, SDP, tensor networks, Matrix Product State, MPS, Matrix Product Operator, MPO, Density Matrix Renormalization Group, DMRG, partial transpose, entanglement entropy

Agradecimentos

Minha família, com muito amor: à minha mãe, Maria Lúcia, que me ensinou, desde pequeno, a *entender e apreciar* o mundo ao meu redor através da matemática e da física; ao meu pai, Armando, que me ensinou, desde pequeno, a *observar e apreciar* tanto o mundo exterior quanto o interior através dos sentidos, pensamentos e reflexão; à minha irmã Candida, pelo infinito carinho, e por ser uma fonte de inspiração e sabedoria imensa; à minha irmã Isis, pelo carinho e apoio constantes; ao meu irmão, João, pelo companheirismo durante a vida.

Colegas do grupo InfoQuant: ao Reinaldo, pela sabedoria, encorajamento, discussões produtivas e grande paciência com minhas dúvidas. Sua paixão e curiosidade por informação quântica são absolutamente contagiantes, e serei eternamente grato por ter compartilhado disso; ao Thiago, pelo conhecimento descomunal, colaborações constantes, muita paciência, companheirismo e excelentes sessões de *Planescape*; ao Marcello, pela calorosa amizade e encorajamento nas horas difíceis; ao Diego, pela amizade e paciência em tirar minhas dúvidas sobre tensor networks; à Ludmila, pelas longas conversas sobre um pouco de tudo; ao André Tanus, pelas boas conversas e por compartilhar algumas loucuras acadêmicas.

Amigos da UFMG: à Anna, pela profunda amizade, carinho, encorajamento, entendimento, e por compartilhar da enorme empolgação com a física e o Universo, bem como a apreciação da beleza nas pequenas coisas; ao André, pelos calorosos abraços, sorrisos e senso de humor inigualável; à Stela e Bonnie, pela amizade e apoio especialmente nos momentos difíceis. Não teria chegado aqui hoje sem vocês; ao Rodolfo, pelo companheirismo, as diversas horas no grupo de estudos e as conversas bizarras.

Professores: ao Elmo, pela grande amizade, apoio e encorajamento constantes; ao Paulo Sérgio, pela sua sabedoria em momentos que precisei encontrar um rumo acadêmico: você estava certíssimo; ao Jim Freericks da *Georgetown University*, pelo apoio e confiança na minha capacidade e as colaborações acadêmicas.

Amigos: à Ria, por todo carinho, apoio, paciência, e encorajamento, mesmo quando não consegui ser o melhor amigo que poderia; à Amber, pelo apoio e amizade incondicionais, e por compartilhar da loucura e empolgação com esse Universo fantástico e louco; à Renata, pelo carinho, compaixão e todas as bizarrices compartilhadas; à Lélia, por me ensinar a ser gentil comigo mesmo e acreditar mais no meu potencial; à Lorena e ao Bruno, por terem me dado a auto-confiança para dar os primeiros grandes passos nessa jornada.

"Get busy living, or get busy dying."
Andy Dufresne

Este trabalho só foi possível graças ao apoio financeiro, direto e indireto, das instituições:

- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (**CAPES**)
- Conselho Nacional de Desenvolvimento Científico e Tecnológico (**CNPq**)

Contents

1 Quantum states and quantum operations **9**
1.1 Mathematical foundations of quantum mechanics . . . 9
1.1.1 Operators, observables, and projectors . . . 10
1.2 Quantum states as a convex set . . . 10
1.2.1 The set of density matrices . . . 10
1.2.2 The maximally-mixed state as a special reference point . . . 12
1.3 Distance and similarity measures . . . 13
1.3.1 Norms of operators . . . 14
1.3.2 Fidelity . . . 15
1.4 Linear maps . . . 16
1.4.1 Positive and completely positive maps . . . 16
1.4.2 The operator-vector correspondence . . . 17
1.4.3 Representations of linear maps . . . 17

2 Entanglement and separability **19**
2.1 Qubits: our quantum system of choice . . . 20
2.1.1 Qutrits and qudits . . . 21
2.2 Entanglement and composite systems . . . 21
2.2.1 Defining separability and entanglement . . . 22
2.3 Separability criteria . . . 23
2.3.1 Peres criterion . . . 24
2.3.2 Peres-Horodecki criterion . . . 25
2.3.3 Other criteria . . . 25
2.4 Multipartite separability and entanglement . . . 25
2.4.1 Hierarchies of partitions . . . 27
2.4.2 Genuine multipartite entanglement . . . 27

3 Entanglement witnesses **29**
3.1 The space of separable and entangled states . . . 29
3.1.1 Separable mixed states . . . 29
3.1.2 PPT bound entanglement mixed states . . . 30
3.1.3 NPT entangled mixed states . . . 30
3.1.4 NPT bound entanglement mixed states . . . 30
3.1.5 NPT entangled pure states . . . 31
3.1.6 Separable pure states . . . 31
3.1.7 Some remarks . . . 31
3.2 Entanglement witnesses . . . 32
3.2.1 The Hahn-Banach separation theorem . . . 32
3.2.2 Defining entanglement witnesses . . . 33

3.2.3 Optimal entanglement witnesses . . . 33
3.2.4 Decomposable and non-decomposable witnesses . . . 34
3.3 Entanglement characterization via entanglement measures . . . 37
3.3.1 Hilbert-Schmidt distance . . . 38
3.3.2 Entanglement entropy . . . 38
3.3.3 Negativity . . . 38
3.3.4 Genuine Multipartite Negativity . . . 39
3.3.5 Robustness . . . 39
3.4 Semidefinite programming . . . 40
3.4.1 Convex optimization and Semidefinite Programming (SDP) . . . 40
3.4.2 Duality . . . 41
3.5 Finding entanglement witnesses with Robust SDPs . . . 42
3.6 Final remarks . . . 43

4 Tensor networks **44**
4.1 A quick introduction to tensors . . . 44
4.2 Tensor network diagrams . . . 46
4.2.1 The basics . . . 46
4.2.2 Index operations . . . 47
4.2.3 Additional features . . . 48
4.3 Tensor networks . . . 48
4.3.1 Tensor networks *vs.* the Hilbert space . . . 49
4.3.2 Types of indices . . . 50
4.3.3 Tensor network contraction . . . 51
4.4 Singular Value Decomposition . . . 51
4.5 Matrix Product States . . . 52
4.5.1 Special representations . . . 54
4.5.2 Benefits of the MPS representation . . . 56
4.6 Matrix Product Operators . . . 56
4.7 Other representations with tensor networks . . . 57
4.8 Tensor network algorithms . . . 57
4.8.1 Density Matrix Renormalization Group (DMRG) . . . 57

5 Entanglement witnesses with tensor networks **60**
5.1 Software used . . . 60
5.2 The tensor network witness model . . . 61
5.3 Why the negativity and decomposable witnesses? . . . 61
5.3.1 Representative entanglement states for 4 and 5 qubits . . . 62
5.3.2 Conclusions . . . 65
5.4 Approaches and results . . . 65
5.4.1 Optimal decomposable witnesses via projectors . . . 65
5.4.2 Hybrid tensor network with SDP blocks . . . 70
5.4.3 Locally-optimized witness MPOs for GMN . . . 71
5.5 Conclusions . . . 71
5.6 Future opportunities for research . . . 72
5.6.1 Detailed k-entanglement survey of large pure states . . . 72
5.6.2 Optimizing GMN witness MPOs via tensor networks . . . 72
5.6.3 SDP for tensor optimization . . . 72
5.6.4 Phase transitions . . . 72

5.6.5 MPO-MPO optimization for mixed states 72
5.6.6 SDPs *via* tensor networks 72
5.6.7 Multiple dimensions 73
5.6.8 Generalized NPT k-partite entanglement 73

List of Figures

1.1 A visualization of Carathéodory's theorem 12

2.1 Visual representation of α_k and k separable sets 27
2.2 The hierarchical structure of k-separable sets. 27

3.1 The Hahn-Banach separation theorem in 2-dimensional Euclidean space. 33
3.2 Illustration of entanglement witnesses 34
3.3 Illustration of non-decomposable *vs.* decomposable witnesses 36
3.4 Geometric interpretation of random robustness 40
3.5 Geometric interpretation of RSDP 43

4.1 Basics of tensor network diagrams 46
4.2 Tensor contractions 47
4.3 Tensor index operations 47
4.4 Tensor trace and directed indices 48
4.5 Accessible states *vs.* space of states 50
4.6 A rank-4 tensor representing a 4-partite state 52
4.7 First step of converting a rank-4 tensor to a MPS 53
4.8 Fully decomposed rank-4 tensor 53
4.9 The final representation of an MPS for a state 54
4.10 Left-canonical orthogonalized form of a MPS 55
4.11 Graphical representation of a Matrix Product Operator 56
4.12 Graphical representation of a PEPS state 57
4.13 Overview of the DMRG algorithm 58

5.1 General tensor network witness model 61
5.2 Entanglement entropy and negativity for an ancilla+4 qubits system 64
5.3 Entanglement entropy and negativity for an ancilla+4 qubits system (sorted) . . 64
5.4 MPO projector constructed from MPS 67
5.5 Entanglement entropy and negativity for the Ising model 69
5.6 Tensor network with witness blocks 71

Introduction

The theory of quantum mechanics has come a long way since its first steps in the early 20th century. Still, even after all these developments quantum entanglement has kept—if not *increased*—its reputation as one of the most fascinating and mysterious of all physical phenomena.

A great deal is understood today about entanglement and the structure of entanglement in the space of quantum states, but most of what it is known has been restricted to very small systems and carefully constructed states. Large-scale physical systems, with even a modest number of particles ($n \approx 10^1$), have mostly eluded our theoretical and experimental grasp so far. Our typical approaches are simply not cut out for tackling these large-scale problems, and as a result, our understanding of large-scale multipartite entanglement and its physical effects is still relatively poor.

Significant progress has been made in these fronts in the last decades, with the invention and application of tensor networks and tensor network algorithms. Today, tensor networks have proven themselves to be powerful, reliable and some would say even more fundamental descriptions of quantum systems. Even so, tensor network techniques are still in their infancy and much work is yet to be done.

In particular, there hasn't been enough efforts in bridging the gap between the use of entanglement witnesses, a powerful tool for characterizing and quantifying multipartite entanglement, and the tensor network formalism. In fact, a survey of the literature reveals that while tensor networks are increasingly being used to study large-scale quantum systems, the analysis of their entanglement structures has largely been based on the same techniques which have been used for over half a century. A modern, bolder, and more complete analysis of large-scale entanglement is warranted, and our research hopes to make progress in this front.

In Chapter 1, we will briefly cover the basic mathematical formalism for representing quantum states and operations under the quantum information tradition. Chapter 2 introduces entanglement, the structures and properties that emerge once systems are put together, from the perspective of the space of states itself, agnostic to any physical processes. Chapter 3 discusses entanglement witnesses and semidefinite programming, a powerful and general tool to study entanglement of all types. Chapter 4 introduces tensor networks, their graphical notation and explains their physical motivation and relevance. Finally, Chapter 5 puts together some of these pieces into a few working examples, and presents a few proof of concepts and working ideas for incorporating entanglement witnesses within tensor networks.

The reader should be familiar with the mathematics of graduate-level quantum mechanics and have a basic understanding of qubits, partial transposes, partial traces and their role in quantum information theory, all of which is covered in most introductory material on the subject.

Chapter 1

Quantum states and quantum operations

"Let no one uninterested in geometry enter here."

Attributed to Pythagoras, at the entrance of his temple

This chapter briefly reviews the basic mathematical structure used in quantum mechanics and quantum information. It is by no means intended to be an introduction to the topic, but merely a quick review for contextualizing the rest of this document and introducing the notational conventions that will be used.

Contrary to how it is typically done, additional results and theorems will be introduced within the context of their usage, as it better highlights their motivation and usefulness as part of the mathematical toolset of quantum mechanics.

1.1 Mathematical foundations of quantum mechanics

The modern foundations of quantum mechanics were first laid out by John von Neumann in the late 1920s [1, 2], placing quantum mechanics in a rigorous, axiomatic mathematical framework. Previous work by Schrödinger, Heisenberg, Dirac and others were in a lot of ways clumsy and *ad hoc*.

In von Neumann's formalism, quantum mechanics is described in terms of vectors and linear operators of a complex N-dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^N$. In mathematics, Hilbert spaces arose as a generalization of Euclidean spaces [3], with a notion of a metric[1], completeness[2], inner products[3], and, importantly, the possibility of arbitrary dimensions [4] – even the infinite ones which are required for describing continuous physical systems[4].

An important note: In this document we will restrict our attention to finite-dimensional Hilbert spaces with bounded operators. Keep in mind, however, that many of the earlier results that will be presented have infinite-dimensional and unbounded generalizations.

In a complex Hilbert space, each distinguished physical state is represented by a family of

[1] A metric defines the notion of "distance" between elements of a space.

[2] A space is complete if all Cauchy sequences converge to an element of the space. Intuitively, there's no point "missing" in the set. Completeness is a requirement for the techniques of analysis and calculus to be applicable to the space.

[3] Inner products define the notion of "similarity" or "angles" between two elements of the space.

[4] Of particular interest is the necessity of well-defined inner products for square-integrable functions.

state vectors, all (complex) scalar multiples of one another[5]. For each possible physical state we can choose one of its vectors, the unit vector with respect to the inner product, to be part of an orthonormal basis for the entire space. Such vectors can be represented by $N \times 1$ column matrices, or symbolically in Dirac notation as kets.

For example, the z-component of the angular momentum of a spin-1/2 particle has two distinguishable states: $+\hbar/2$ (up) or $-\hbar/2$ (down). The states of this system are represented by the Hilbert space $\mathcal{H} \cong \mathbb{C}^2$, and we define two orthonormal column vectors as the basis for this space, $|0\rangle = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $|1\rangle = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, corresponding to spin up and spin down states.

Physical processes and observable quantities are incorporated into the formalism through the set $\mathcal{L}(\mathcal{H})$ of linear operators acting[6] on $\mathcal{H}$, which is itself a vector space that can be represented by $N \times N$ complex matrices in the finite-dimensional case. In finite dimensions we are further restricted to the set of bounded Hermitian operators acting on $\mathcal{H}$, which in this text will be denoted by $\mathcal{B}(\mathcal{H})$. Hermitian (or self-adjoint) matrices are used due to their many useful properties, such as having real eigenvalues and a spectral decomposition [6].

1.1.1 Operators, observables, and projectors

The space $\mathcal{L}(\mathcal{H})$ of linear operators of a Hilbert space onto itself can be rigorously identified as the composite space given by the tensor product of $\mathcal{H}$ and its dual: $\mathcal{L}(\mathcal{H}) = \mathcal{H} \otimes \mathcal{H}^*$. Elements of such space, in Dirac's notation, are typically represented in terms of tensor products of a ket with a bra (or equivalently, a vector outer product), which are notated as $|\psi\rangle \otimes \langle\phi| = |\psi\rangle\langle\phi|$.

This result generalizes to linear maps L between arbitrary spaces as

$$L : \mathcal{X} \mapsto \mathcal{Y} \qquad \to \qquad L \in \mathcal{L}(\mathcal{X}, \mathcal{Y}) = \mathcal{Y} \otimes \mathcal{X}^*$$

which makes explicit the connection to linear maps between spaces through the Riesz theorem [4]. Although a familiar result, we highlight this structure as it will be exploited in chapter 4 regarding tensor networks, where the tensors are interpreted as linear maps connecting Hilbert spaces in this fashion.

1.2 Quantum states as a convex set

1.2.1 The set of density matrices

By construction, a state $|\psi\rangle$ is a mathematical object which specifies all available information about a system, which is the reason why it is referred to as a *pure state*.

In practice, however, complete information is not obtainable and the specification of physical states relies on the preparation of ensembles [4]. This notion of *statistical mixtures* of pure states was also addressed by von Neumann and Lev Landau, who realized that pure states and such mixtures can all be expressed by a *density operator* or density matrix[7]. In order for an operator to represent a state, it must match the following

Definition 1. *A* ***density operator*** *or* ***density matrix*** *is a bounded operator $\rho \in \mathcal{B}(\mathcal{H})$ which satisfies the following three properties:*

1. *Hermitian:* $\qquad \rho^\dagger = \rho$

[5]Therefore, the state $|\psi\rangle$ is taken as physically equal to $e^{i\theta}|\psi\rangle$ for $\theta \in \mathbb{R}$. Global phases are irrelevant.

[6]Linear operators may act on different spaces, such that $A \in \mathcal{L}(X, Y)$ represents the linear mapping $A : \mathcal{X} \mapsto \mathcal{Y}$. The notation $\mathcal{L}(X)$ is to be understood as $\mathcal{L}(X, X)$. This space has special properties [5] which will be covered at a later moment.

[7]We will adhere to "density matrix" as our focus will be on the numerical properties of such objects, not their algebraic properties.

2. *Normalized:* $\mathrm{Tr}[\rho] = \|\rho\|_1 = 1$

3. *Positive semi-definite:* $\langle \psi \,|\, \rho \,|\, \psi \rangle \geq 0$, *for all* $|\psi\rangle$, *or simply* $\rho \succeq 0$

Property 1 ensures there is a spectral decomposition with orthonormal eigenvectors and real eigenvalues, which for ρ specify a probability distribution. Property 2 ensures the total probability adds up to unity: it is a *complete* description of the system. Property 3 ensures that the system is physical: there are no negative probabilities. The set of operators obeying the above properties will be denoted by $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$.

Given a pure quantum state $|\psi\rangle$, the density matrix representing it is given by the outer product

$$\rho = |\psi\rangle\langle\psi| \,. \tag{1.1}$$

which automatically satisfies all properties listed previously, provided the state is properly normalized. The resulting matrix has *rank* 1, that is, it has a single non-zero eigenvalue. We now turn to the idea of convexity.

Definition 2. *A **convex combination** B of N elements $\{A_i\}$ is any linear combination of the form:*

$$B = \sum_{i=1}^{N} \alpha_i A_i \qquad \textit{with } \alpha \in [0,1] \quad \textit{and} \quad \sum_{i=1}^{N} \alpha_i = 1.$$

A statistical mixture of arbitrary states from a basis $\{|\psi_i\rangle\}$ is a convex combination of pure states, which is itself also a state:

$$\rho = \sum_i \lambda_i \, |\psi_i\rangle\langle\psi_i| \,, \quad \rho\, |\psi_i\rangle = \lambda_i \, |\psi_i\rangle \quad \text{with } \lambda_i \in [0,1] \quad \text{and} \quad \sum_i \lambda_i = 1. \tag{1.2}$$

This defines a set of diagonal operators ρ for the given basis $\{|\psi_i\rangle\}$, referred to as the *eigensemble* or *eigenvalue simplex*, and every density matrix in $\mathcal{M}$ is an element of some eigenvalue simplex. These sets correspond to $(N-1)$-dimensional cuts through $\mathcal{M}$ [7].

The combinations will not result in a pure state if any $\lambda_i \notin \{0,1\}$, in which case the state is said to be *mixed*[8]. In turn, mixed states themselves may be used in convex combinations to obtain other states. Therefore, combinations leading to a given mixed state are not unique.

We now address the following.

Definition 3. *A **convex set** is the set of all elements which are a convex combination of elements of another set.*

Definition 4. *A **convex hull** of a given set S is the smallest convex set which contains S.*

The above definitions also mean the set of density matrices $\mathcal{M}$ is a *convex set*, and that the pure states define the vertices (or *pure points*) of a polytope[9] which defines a *convex hull* within the space $\mathcal{B}(\mathcal{H})$. Two important theorems arise from this structure [7, Ch.1]:

Theorem 1. *(Minkowski's theorem) Any convex set is the convex hull of its pure points.*

Theorem 2. *(Carathéodory's theorem) If X is a subset of $\mathbb{R}^N$, then any point in the convex hull of X can be expressed as a convex combination of at most N+1 points in X.*

[8]Or more accurately, non-pure.

[9]A polytope is the multidimensional generalization of polygons and polyhedra.

Theorem 1 tells us the intimate connection between convex hulls of pure points and convex sets created from combinations of the pure points. On the other hand, theorem 2 tells us that all points inside (and including) the convex hull defining a convex set can be specified by convex combinations of the pure points. In effect, the two theorems tell us the pure points are all the information we need to fully assemble the convex set, although the number of such pure points may be infinite in some cases.

What this means is that a basis of N pure states $|\psi_i\rangle$ is sufficient to fully specify any element of $\mathcal{M}$ (after an appropriate diagonalization of the operator), even though a naive application of Carathéodory's theorem to the space of complex matrices $\mathbb{C}^{N^2} = \mathbb{R}^{2N^2}$ would suggest a much larger number of elements required. This structure of the space of density matrices is therefore a fortunate result.

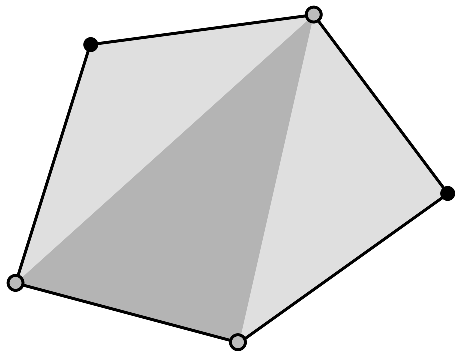


Figure 1.1: A visualization of an application of Carathéodory's theorem in two dimensions ($\mathbb{R}^2$). Any point inside the pentagon lies inside a triangle formed by some set of three vertices. The region inside the triangle is the convex set defined by its vertices, as per Minkowski's theorem. This is always a N-dimensional simplex, since it is the simplest polytope with dimension equal the dimension in which the convex set is embedded.

The rank of ρ immediately gives us the number of pure points necessary to represent it as a convex combination. If ρ has at least one zero eigenvalue, it lies on the boundary of the convex set. The density matrices of mixed states have rank greater than 1, indicating that there are more than one pure state present in a convex mixture.

This space has many deep geometric structures which are beyond the scope of this text. More information may be found in Życzkowski and Bengtsson [7, Ch.8]. We will concern ourselves with only a few of its structures (chapter 2 section 3.1).

This brief discussion focused on describing single quantum systems, but the richness of quantum mechanics comes to the surface once we consider composite quantum systems. That will be the topic of the next chapter.

1.2.2 The maximally-mixed state as a special reference point

As it is, there is no special reference state in $\mathcal{M}$, since the requirements imposed for states excluded the origin of the Hilbert space and its associated operator in the Hilbert-Schmidt space (a trace zero operator is not a state). However, this restricted space still has a linear structure, which means it is now an *affine space* characterized by invariance of straight lines under affine transformations [8]. These transformations are ubiquitous in quantum mechanics.

This fact ties to a geometric meaning of convex mixtures, which is useful to expound upon. Suppose we have two states ρ_A and ρ_B. The family of states

$$\rho(t) = (1-t)\rho_A + t\rho_B, \quad 0 \le t \le 1,$$

are convex combinations of both states parameterized by the real number t. This family traces a straight line in $\mathcal{M}$, in the sense that the convex combination structure of this family is preserved under affine transformations. Therefore, per Minkowski's and Carathéodory's theorems, a convex set remains convex under affine transformations.

Affine spaces allow us to pick an arbitrary reference point to specify our elements. The nature of density matrices as (effectively) probability distributions induces us to select a particular state for this task. A well-known result in probability and information theory [9] is that the discrete distribution which maximizes the Shannon entropy

$$H = -\sum_{i=1}^{N} p_i \ln p_i, \tag{1.3}$$

is given by the uniform distribution $p_i = 1/N$. The uniform distribution can be intuitively understood as the distribution of *maximum ignorance*. The analogue for density matrices is the state which is an equal mixture of all states $|\psi_i\rangle$ in a basis, which gives rise to the following

Definition 5. *The* ***maximally mixed state*** $\rho_\star$ *is the state which is composed of an equal convex mixture of all pure states* $|\psi_i\rangle\langle\psi_i|$ *in any basis* $\{|\psi_i\rangle\}$ *for* $\mathcal{H}$*:*

$$\rho_\star = \frac{1}{N}\sum_{i=1}^{N} |\psi_i\rangle\langle\psi_i| = \frac{\mathbb{I}}{N}. \tag{1.4}$$

While density matrices are in general not uniquely defined as convex combinations, the spectral decomposition of a given state ρ provides us with a special convex combination in which the eigenvalues (acting as weights in the combination) correspond to probabilities. This fact lead us to the natural generalization of the Shannon entropy [5, 7, 10].

Definition 6. *The* ***von Neumann entropy*** *of a state* ρ *is defined as the quantity*

$$S(\rho) = -\sum_{i=1}^{N} p_i \ln p_i = -\operatorname{Tr}[\rho \ln \rho]\,. \tag{1.5}$$

where p_i *are the eigenvalues of the spectral decomposition of* ρ*.*

By extension, the maximally mixed state can then be seen as the state which maximizes von Neumann entropy. By construction, this state lies at the center of any eigenvalue simplex, which highlights its special role in $\mathcal{M}$. The maximally mixed state, therefore, is *full rank* (rank N), and all of its eigenvalues are equal to $1/N$, indicating that we need to consider all states in a basis equally to represent it.

1.3 Distance and similarity measures

With the space of operators and density matrices defined, it is now important to look at the notion of distance and similarity between elements of the set.

1.3.1 Norms of operators

Operator norms, denoted by $\|\cdot\|$, are functions acting on elements of a linear space $\mathcal{L}(X)$ of some generic space X, which map operators in that space to scalars: $\|\cdot\| : \mathcal{L}(X) \mapsto \mathbb{R}$. In order for a function to be a norm, it must obey three properties [5]:

1. Positive semi-definite: $\|A\| \geq 0$, for all $A \in \mathcal{L}(X)$ and $\|A\| = 0$ *iff* $A = 0$
2. Positive scalability: $\|\alpha A\| = |\alpha|\|A\|$ for all $A \in \mathcal{L}(X)$ and $\alpha \in \mathbb{C}$
3. Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$, for all $A, B \in \mathcal{L}(X)$

Such norms appear naturally in the structure of the spaces we have defined so far. The Hilbert-Schmidt space of bounded operators acting on $\mathcal{H}$, which we previously denoted by $\mathcal{B}(\mathcal{H}) \subset \mathcal{H} \otimes \mathcal{H}^*$, is itself also a Hilbert space on its own right. This means it also admits an inner product [5, 7], which in this case is given by

$$\langle A, B\rangle = c\,\mathrm{Tr}\left[A^\dagger B\right],$$

where c is some arbitrary constant which sets the scale. We will assume c = 1. The scalar product of an operator with itself defines an Euclidean norm, which in $\mathcal{B}(\mathcal{H})$ suggests

Definition 7. *The* ***Frobenius norm*** *of an operator A is defined as*

$$\|A\|_2 = \sqrt{\mathrm{Tr}[A^\dagger A]}.$$

We can use this norm to define a generalized Euclidean distance called the *Hilbert-Schmidt distance*:

$$D^2_{HS}(A, B) = \|A - B\|_2^2 \equiv \langle A - B, A - B\rangle \ = \mathrm{Tr}\left[(A - B)^\dagger (A - B)\right]. \tag{1.6}$$

This is the distance between operators in the space of matrices defined by $\mathcal{L}(\mathcal{H})$, treating them as vectors. Other norms may also be defined by generalizing the Euclidean norm in terms of a real parameter $p \geq 1$, giving

Definition 8. *The* ***Schatten p-norms*** *[5] of an operator A are defined as*

$$\|A\|_p = \left(\mathrm{Tr}\left[A^\dagger A\right]^{p/2}\right)^{1/p} \qquad p \geq 1.$$

These norms can be equivalently defined [5] in terms of a vector $s(A)$ with the singular values of A, where the Schatten p-norms become equivalent to the ordinary vector p-norm of $s(A)$:

$$\|A\|_p = \|s(A)\|_p.$$

For $p = 2$, we have the Frobenius norm. We also have

Definition 9. *The* ***trace norm*** *of an operator A is defined as the Schatten-p norm of A with $p = 1$:*

$$\|A\|_1 = \mathrm{Tr}\left[\sqrt{A^\dagger A}\right].$$

The trace norm has useful properties for characterizing entanglement criteria [11], as we will briefly see in chapter 2. With the trace norm, one may also define the *trace distance*:

$$D_{Tr}(A, B) = \|A - B\|_1. \tag{1.7}$$

Another important norm is given by

Definition 10. *The* ***operator norm***[10] *of an operator A is defined as*

$$\|A\|_\infty = \max\{\|Au\| : u \in \mathcal{H}, \|u\| = 1\} = \lim_{p\to\infty} \|A\|_p, \tag{1.8}$$

where here $\|\cdot\|$ is the typical vector Euclidean norm.

This norm is special because it is the norm induced by the Euclidean norm of elements of $\mathcal{H}$, as expressed in equation 1.8. It has the property

$$\|A^\dagger A\| = \|A\|^2,$$

which is, in a sense, the proper generalization of Euclidean norms of vectors to operators. In terms of the singular values $s(A)$, we have $\|A\|_\infty = \max\{s(A)\}$.

1.3.2 Fidelity

Beyond these norms, there are other ways we can characterize the similarity between density matrices. One of the most important ones is the fidelity, which we define next.

Definition 11. *Given two positive semidefinite operators $P, Q \in \mathrm{Pos}(\mathcal{B}(\mathcal{H}))$, we define the* ***fidelity*** *[5, 10] between the two as*

$$F(P, Q) = \|\sqrt{P}\sqrt{Q}\|_1 = \mathrm{Tr}\left[\sqrt{\sqrt{P}Q\sqrt{P}}\right] \tag{1.9}$$

Typically, the fidelity is used to measure similarity between density matrices. However, there's no reason to not encompass positive semidefinite operators in general.

The fidelity function has some properties worth mentioning. From the definition, it is clear that it is symmetric:

$$F(P, Q) = F(Q, P).$$

Given a pure state $\rho = |\psi\rangle\langle\psi| \in \mathcal{M}$ and a positive semidefinite operator $Q \in \mathrm{Pos}(\mathcal{B}(\mathcal{H}))$, the following holds

$$F(\rho, Q) = \sqrt{\langle\psi\,|\,Q\,|\,\psi\rangle} = \sqrt{\mathrm{Tr}[Q\rho]},$$

or more generally

$$F(\rho_1, \rho_2) = |\,\langle\rho_1, \rho_2\rangle\,|.$$

Another particularly important property of the fidelity is that it is multiplicative with respect to tensor products

$$F(P_1 \otimes \cdots \otimes P_N, Q_1 \otimes \cdots \otimes Q_N) = \prod_{i=1}^{N} F(P_i, Q_i).$$

[10] Also referred to as *spectral norm*.

1.4 Linear maps

Definition 12. *A **linear map** is a linear function Φ which maps elements X from a set $\mathcal{X}$ onto elements Y of another set $\mathcal{Y}$:*

$$\Phi : \mathcal{X} \mapsto \mathcal{Y} \qquad \Phi(X) = Y \quad \textit{where} \quad X \in \mathcal{X}, Y \in \mathcal{Y}$$

Under a natural definition of scalar multiplication and addition of linear maps, the maps themselves may also form a linear space $\mathrm{M}(\mathcal{X}, \mathcal{Y})$ [5]. By virtue of being a linear function, linear maps enjoy the following properties:

- $\Phi(\alpha X_1 + \beta X_2) = \alpha\Phi(X_1) + \beta\Phi(X_2)$, *for* $\alpha, \beta \in \mathbb{C}$.
- $(\Phi + \Psi)(\alpha X) = \alpha\Phi(X) + \alpha\Psi(X)$, *where Ψ is another linear map.*
- $(\alpha\Phi)(X) = \alpha(\Phi(X))$.

The transpose $\mathrm{T} : \mathcal{X} \mapsto \mathcal{X} : \mathrm{T}(X) = X^{\mathrm{T}}$ and the trace $\mathrm{Tr} : \mathcal{X} \mapsto \mathbb{C}$ are two important examples of linear maps which we use extensively.

The subject of quantum operations as linear maps on the space of quantum states[11] is *vast*, and there is no reason to go through these details here. The interested reader may acquaint themselves by checking the typical literature on quantum information theory [5, 10, 12]. We will simply highlight a few key results needed for our work.

1.4.1 Positive and completely positive maps

Linear maps acting on operators may be classified in various ways. Two classifications of note are positiveness and complete-positiveness. We will only briefly define them here to contextualize our usage.

Definition 13. *A linear map $\Phi : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H})$ acting on a positive operator $A \in \mathcal{B}(\mathcal{H})$ is a **positive map** if $\Phi(A) \succeq 0$ for all $A \succeq 0$ [5, 7].*

An example of a positive map is the matrix transpose.

Definition 14. *A linear map $\Phi : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H})$ acting on a positive operator $A \in \mathcal{B}(\mathcal{H})$ is a **completely positive map** if the extended map $\Phi \otimes \mathbb{I}_K$ acting on the extended space $\mathcal{B}(\mathcal{H} \otimes \mathcal{H}_K)$ is positive for all K [5, 7].*

Completely positive maps are important in quantum mechanics and quantum information as the only maps which can describe physical processes [5, 7, 10]: intuitively, a linear map representing a physical process cannot map density matrices, which represent physical states, into something which is not a density matrix.

The definition of completely positive maps in terms of extensions of the Hilbert space arises from our freedom to attach other interacting quantum systems, such as the environment or some *ancilla*, to the system of interest. In order for this to hold in general, any such extended system should remain positive under the action of the extended map acting on the original system, otherwise it cannot represent a natural process [5, 7].

The matrix transpose is an example of a positive map which is not completely positive. This is a key result underlying our work (see chapter 2 section 2.3.1).

[11] Or even the space of linear maps $\mathrm{M}(\mathcal{X}, \mathcal{Y})$ itself.

1.4.2 The operator-vector correspondence

There is a bijection between the space $\mathcal{L}(\mathcal{X},\mathcal{Y})$ of operators $A : \mathcal{X} \mapsto \mathcal{Y}$ and vectors in the space $\mathcal{Y} \otimes \mathcal{X}$. This is typically referred to as the *vectorization* of an operator, and will be useful in later chapters.

Definition 15. *The* ***vectorization*** *map [5]* $\mathrm{vec} : \mathcal{L}(\mathcal{X},\mathcal{Y}) \mapsto \mathcal{Y} \otimes \mathcal{X}$ *is the linear map that changes from the canonical basis* $\{E_{ij}\}$ *for* $\mathcal{L}(\mathcal{X},\mathcal{Y})$ *to the canonical basis* $\{e_i\}$ *for* $\mathcal{Y} \otimes \mathcal{X}$*. In Dirac's bra-ket notation, this is equivalent to*

$$\mathrm{vec}[\,E_{ij}\,] = e_j \otimes e_i \qquad \rightarrow \qquad \mathrm{vec}[\,|j\rangle\langle i|\,] = |j\rangle \otimes |i\rangle = |j\rangle\,|i\rangle\,,$$

where E_{ij} *is the matrix with elements* $\delta_{i,j}$ *and* e_i *a vector with zeros except for a* 1 *at the* i*-th position. The inverse map will be denoted by* $\mathrm{vec}^{-1}[\,\cdot\,]$*, which we might casually refer to "unvectorization".*

In more concrete terms, this map is equivalent to stacking the columns of a $m \times n$ matrix on top of one another, creating a $mn \times 1$ column vector.

1.4.3 Representations of linear maps

Linear maps may be represented in multiple ways, and each way has its own benefits. We'll briefly define two useful representations.

The natural representation

Perhaps the most straightforward way to specify a linear map is to simply write explicitly the connection between input and output spaces [5].

Suppose we have a linear map $\Phi : \mathcal{X} \mapsto \mathcal{Y}$ and let $X \in \mathcal{X}$. It is clear that the map $\mathrm{vec}[\,\mathcal{X}\,] \mapsto \mathrm{vec}[\,\Phi(\mathcal{X})\,]$ is also linear, as it can be represented as a composition of linear map. Therefore, this map can be represented by a linear operator $K(\Phi)$ living on a higher space $\mathcal{L}(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y})$. This leads us to

Definition 16. *An operator* $K(\Phi) \in \mathcal{L}(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y})$ *is the* ***natural representation*** *[5] of a linear map* $\Phi : \mathcal{X} \mapsto \mathcal{Y}$ *if*

$$K(\Phi)\mathrm{vec}[\,X\,] = \mathrm{vec}[\,\Phi(X)\,]\,, \qquad \textit{for all } \; X \in \mathcal{L}(X). \tag{1.10}$$

In simpler terms, by unraveling the space of operators we may treat them as vectors, in which case any linear map between the original operators may be represented as an operator $K(\Phi)$ acting in the unraveled space.

This technique is particularly useful when we are dealing with numerical matrices, and will be used in later chapters.

In concrete terms, for a space of $N \times N$ matrices, one may write $K(\Phi)$ explicitly as

$$K(\Phi) = \sum_{\substack{i=1\\j=1}}^{N} \sum_{\substack{k=1\\l=1}}^{N} \langle E_{kl}, \Phi(E_{ij})\rangle\, E_{ki} \otimes E_{lj}. \tag{1.11}$$

The Choi-Jamiołkowski representation

Another important representation of linear maps is the Choi-Jamiołkowski representation, which also takes us to a higher-dimensional space but in a more interesting way.

Definition 17. *A linear map* $\Phi : \mathcal{X} \mapsto \mathcal{Y}$ *can be represented by the linear operator* $J(\Phi)$: $\mathrm{M}(\mathcal{X}, \mathcal{Y}) \mapsto \mathcal{L}(\mathcal{Y} \otimes \mathcal{X})$ *defined as*

$$J(\Phi) = \sum_{i,j} \Phi(E_{ij}) \otimes E_{ij}, \tag{1.12}$$

which is known as the ***Choi-Jamiołkowski representation*** *of the map* Φ.

The action of the original map Φ can be recovered from $J(\Phi)$ through

$$\Phi(X) = \mathrm{Tr}_{\mathcal{X}}\big[J(\Phi) \cdot (\mathbb{I}_{\mathcal{Y}} \otimes X^{\mathrm{T}})\big] \tag{1.13}$$

The Choi-Jamiołkowski representation is useful because it can be used to characterize the positivity of linear maps, among many other important properties [5]. For example, the following theorem is quite useful:

Theorem 3. *For every linear map* $\Phi : \mathcal{X} \mapsto \mathcal{Y}$ *with* $\Phi \in \mathrm{M}(\mathcal{X}, \mathcal{Y})$, Φ *is completely positive iff* $J(\Phi) \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$.

Chapter 2

Entanglement and separability

"Concordia res parvae crescunt."
(*"Small things thrive in harmony."*)

Gaius Sallustius Crispus

The most striking feature of quantum mechanics is arguably the phenomenon of quantum *entanglement*[1]. First considered by Albert Einstein, Boris Podolsky and Nathan Rosen in 1935 as a thought experiment [14], the now known as the "*EPR Paradox*" postulates a situation wherein a pair of correlated particles originating from a common source are physically separated by a large distance. An experimenter on one end may measure the state of their particle by some locally randomized process, and in doing so, would seem to immediately affect the results of measurements done to another particle, violating special relativity.

Since this behavior seemed absurd[2], it was claimed quantum theory must be incomplete. In defense, others such as John von Neumann [15, p.144-145] and David Bohm [16] suggested that more information must be carried by the particles which were not originally being accounted for, what is now known as the *hidden variables* hypothesis. This notion of hidden variables later sparked deep discussions by John S. Bell [17, 18], Andrew M. Gleason [19], Simon B. Kochen and Ernst Specker [20] on the physical and mathematical nature of the these quantum correlations, which have laid the foundations of modern entanglement theory.

Ultimately, the existence of quantum entanglement as a legitimate phenomenon was vindicated by experiments [21], and has since been observed in many physical systems such as photons [22], electrons [23], large molecules [24], and even macroscopic crystals [25]. The lesson learned, as Richard Feynman famously stated, is that

"*The imagination of Nature is far, far greater than the imagination of man.*"
— Richard Feynman

We shall not attempt to cover the basic details of entanglement theory or its history. One may refer to typical graduate-level introductory literature [4, 10] on quantum mechanics. Extra knowledge out of the typical curriculum will be introduced as necessary.

[1]The term *entanglement* was first suggested by Erwin Schrödinger in a letter to Einstein, as his own translation of his German term *Verschränkung* [13].

[2]Being derogatorily referred to as *spooky action at a distance.*

2.1 Qubits: our quantum system of choice

To begin developing our main work, we must pick a building block system to work with more concretely. The smallest quantum system of interest for this purpose has two states. There are many examples of such systems in Nature[3], but two of the most important ones in experimental quantum physics are spin-1/2 particles and photon states[4]. We will be using spin-1/2 particles in most of our models, as our work will focus on one-dimensional spin-1/2 chains.

If we measure a spin-1/2 particle's angular momentum component in any specific direction we always measure either $+\hbar/2$ or $-\hbar/2$. This is the case even for iterated measurements along orthogonal directions, a striking result completely incompatible with any classical explanation [4, 26].

If we arbitrarily take this direction to be the z-axis, we can define the $+\hbar/2$ state as $|0\rangle$ and the $-\hbar/2$ state as $|1\rangle$, which together form a complete basis for the 2-dimensional *complex* Hilbert space of the particle's spin state, $\mathcal{H} \cong \mathbb{C}^2 \cong \mathbb{R}^4$. A general pure state for this system may then be written as

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle \,, \qquad \text{with} \quad |\alpha_0|^2 + |\alpha_1|^2 = 1,$$

or, in the density matrix formalism, as

$$\rho = \begin{bmatrix} |\alpha_0|^2 & {\alpha_0}^* \alpha_1 \\ \alpha_1 {\alpha_0}^* & |\alpha_1|^2 \end{bmatrix}.$$

These may also be parameterized in more useful ways, such as the Bloch sphere [10], where pure and mixed states have a natural interpretation in terms of the norm of their Bloch vectors. We will not need more refined descriptions than this, however, as our ultimate goal will involve systems with many particles where such parameterizations are not only unavailable, but *useless*.

Numerically, we may represent the basis kets as the column vectors

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

This representation will be used throughout the rest of this text.

The above system and description defines a *quantum bit*, *qubit* for short, which is the quantum analogue of a classical bit of information. Just as in the classical bit, the qubit is defined in terms of two distinct states, 0 and 1. The difference in the quantum case is the existence of superposition states, which lead to a variety of interesting applications in quantum computing [10].

As we know, quantum systems may be joined into larger composite systems forming a larger Hilbert space. We can have n qubits as a single 2^n-dimensional quantum system, in which case we need 2^n possible basis states. Conventionally, we order them in ascending binary order. For $n = 2$, we have

$$\{\ |00\rangle\,, |01\rangle\, |10\rangle\, |11\rangle\ \},$$

for $n = 3$,

$$\{\ |000\rangle\,, |001\rangle\,, |010\rangle\,, |011\rangle\,, |100\rangle\,, |101\rangle\, |110\rangle\,, |111\rangle\ \},$$

and so forth. We refer to these bases as the *computational basis* for a given system of many qubits, and take it as the implied *canonical basis* for their state space.

[3] All of them are mathematically equivalent, despite the vastly different physical realizations.

[4] Polarization states, photon counts, and photon paths are typical implementations [10].

Definition 18. *For a set of n qubits, the **computational basis** is defined as the set of kets*

$$\{|b_n, \cdots, b_i, \cdots, b_1\rangle\}, \quad b_i \in \{0, 1\}$$

ordered in ascending binary order. The least significant qubit may be on the left or the right (as is above). The latter will be the convention we will use.

Since the number of available states in the composite system grows exponentially with the number of qubits, this can very quickly render any naive numerical description of these systems computationally intractable.

For example, if we wish to represent the full density matrix for a system of n qubits, accounting to all possible pure and mixed states, we need a matrix with $(2^n)^2 = 2^{2n} = 4^n$ complex values. The full density matrix of a system with a mere 35 qubits stored in a typical 128-bit precision complex numbers would require 2^{77} bits of storage, which is on the order of the amount of *all* digitally stored data in the world as of 2018 [27]. As Carlton Caves famously said,

> "*Hilbert space is a big place.*"
>
> — Carlton Caves

It seems that we must be a lot more clever if we wish to entertain the idea of studying such large systems numerically. The main subject of this dissertation is precisely on how to study the properties of such large composite systems in an efficient manner. This will be the topic of chapter 4.

2.1.1 Qutrits and qudits

Higher dimensional systems analogous to the qubit are also possible. For $d = 3$, these are referred to as *qutrits*, and for a general dimension d, *qudits*. The formalism is exactly the same, however, except we now include more kets in our basis, $\{|0\rangle, |1\rangle, \cdots, |d-1\rangle\}$ for a general qudit.

We will not concern ourselves with these systems much in this text, but it is worth saying that many of the techniques that will be discussed are readily generalized to arbitrary assemblies of qudits.

2.2 Entanglement and composite systems

Entanglement becomes possible as soon as we create a composite system out of more than one subsystem. This immediately presents us with the daunting challenge of understanding the additional structure in these composite Hilbert spaces.

Given two systems, $\mathcal{A}$ and $\mathcal{B}$ described by the Hilbert spaces $\mathcal{H}_A$ with dimension d_A and $\mathcal{B}$ with dimension d_B, the space of a composite system constructed by uniting the two is given by the tensor product of the two spaces $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. This is a $d_{AB} = d_A d_B$ dimensional Hilbert space which has a *much* richer structure than that of the individual systems in isolation[5]. In quantum mechanics, the whole is greater than the sum of its parts[6] [7].

Two-qubit entanglement has been the subject of most of the initial research on the subject of entanglement [10, 17, 19, 20], and is to a large extent a *solved* problem [5, 10, 28, 29].

[5]The Hilbert space of a single d_{AB}-dimensional system is isomorphic to that of two d_A and d_B-dimensional systems put together, but it is the way in which the space is partitioned and interpreted as subsystems that gives rise to richer structures. For example a 4×2 system is fundamentally different than a $2 \times 2 \times 2$ system, even though both Hilbert spaces are isomorphic.

[6]After all, $d_A d_B \geq d_A + d_B$ for $d_A, d_B \geq 2$.

Beyond qubits, the Hilbert spaces of small-dimensional two-part (or *bipartite*) quantum systems are quite well understood [7,11,29], as well as their entanglement structures. It is the larger systems which pose a significant challenge, and these will be our main concern.

But first, in order to get there, we will briefly introduce some concepts regarding entanglement and multipartite quantum systems. We will begin with bipartite systems and generalize from there.

2.2.1 Defining separability and entanglement

"The ships hung in the sky in much the same way that bricks don't."

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*
on the odd behavior of brick-shaped Vogon starships

As with many other odd things in life, entanglement is typically defined in terms of what it is ***not***. Let us consider two density matrices $\rho_A \in \mathcal{M}_A$ and $\rho_B \in \mathcal{M}_B$ for two systems $\mathcal{A}$ and $\mathcal{B}$ with arbitrary dimensions d_A and d_B, and the density matrix of the composite $(d_A d_B)$-dimensional system $\rho_{AB} \in \mathcal{M}_{AB}$.

Prejudiced by our classical intuitions, one may be fooled into thinking that any state of ρ_{AB} can be perfectly described by a complete simultaneous description of ρ_A and ρ_B. In mathematical terms, we would be claiming that

$$\rho_{AB} = \rho_A \otimes \rho_B.$$

The above is referred to as a *product state*, and it indicates that the systems $\mathcal{A}$ and $\mathcal{B}$ are completely independent of one another. In more general terms, we have

Definition 19. *A quantum state $\rho \in \mathcal{M}_S$ of a composite system $\mathcal{S}$ with n parts $\rho_i \in \mathcal{M}_i$ (a n-partite system) is referred to as a* ***product state*** *if it can be written as*

$$\rho = \bigotimes_{i=1}^{n} \rho_i = \rho_1 \otimes \cdots \otimes \rho_n. \tag{2.1}$$

Given a general state ρ, we may test if it is a product state by checking whether the following equation involving the partial traces of ρ holds:

$$\rho = \mathrm{Tr}_{\{n\}\backslash 1}[\rho] \otimes \mathrm{Tr}_{\{n\}\backslash 2}[\rho] \otimes \cdots \otimes \mathrm{Tr}_{\{n\}\backslash (n-1)}[\rho] \otimes \mathrm{Tr}_{\{n\}\backslash n}[\rho]\,,$$

that is, whether we can re-assemble the original state piece by piece by throwing away everything but one part at a time.

Here, it will be convenient to also establish the following definition.

Definition 20. *The total dimension D of the Hilbert space $\mathcal{H}_S$ of a n-partite system, each part with dimension d_i, is given by the product of the dimensions of each part:*

$$D = \prod_{i=1}^{n} d_i.$$

We will be using the capital letter D to refer to the total dimension of general n-partite systems.

Product states represent systems in which there are no correlations whatsoever between the subsystems [4, 7]. But correlations can and do exist, and it is important to understand if they are classical or quantum in nature. As we have seen in chapter 1, the space of density matrices forms a convex set, so any convex combination of product states will also be a valid quantum state (although not necessarily a product state). This leads us to

Definition 21. *A quantum state $\sigma \in \mathcal{M}_S$ of a n-partite composite system $\mathcal{S}$, with states $\rho_{k,i} \in \mathcal{M}_i$ for each part, is referred to as a* ***separable state*** *[5, 7, 30] if it can be written as a convex combination of a set $\{k\}$ of product states*

$$\sigma = \sum_{k=1}^{M} p_k \, \rho_{k,1} \otimes \cdots \otimes \rho_{k,n} \qquad \sum_{k=1}^{M} p_k = 1, \quad p_k \in [0,1]. \tag{2.2}$$

The smallest number of terms $M = |\{k\}|$ needed is referred to as the cardinality of the state σ.

Here, we note our use of the Greek letter sigma (σ) to denote the separable state. Henceforth, a state denoted by σ is to be understood as a separable state, whereas ρ will be a general state. As usual, the decomposition in equation 2.2 is not unique.

Separable states are ensembles of product states, and may give rise to classical correlations. For example, with the following separable state

$$\rho = \frac{1}{2} \left(|00\rangle\langle 00| + |11\rangle\langle 11| \right),$$

a measurement (in the canonical basis) of the first system will correlate with the second, but there is nothing fundamentally quantum about this correlation. It is as prosaic as sending out party invitations to multiple guests, forgetting which date the party was supposed to be, and later having them all show up on time.

A special subset of the separable states $\mathrm{Sep}(\mathcal{M}_S) \subset \mathcal{M}_S$, defined in equation 2.2, is fully contained within a ball centered around the maximally mixed state $\rho_\star$ [7, Ch.15§5], which means there are always some separable states near $\rho_\star$. As the number of parts increases, the set of all separable states $\mathrm{Sep}(\mathcal{M}_S)$ contains a *vanishingly small* fraction of the total Hilbert space $\mathcal{H}$ [31, 32], which goes to show just how much richer the structure of quantum states are compared with their classically-describable counterparts[7].

At last, it is with these realizations that we come to the (seemingly anticlimactic) definition of entanglement:

Definition 22. *A quantum state $\rho \in \mathcal{M}$ is said to be* ***entangled*** *if it is* ***not separable****.*

We will further refine these definitions of separable and entangled states in section 2.4.

2.3 Separability criteria

The definition of entanglement logically induces us to ask the following question:

If we are given a state ρ, how do we tell if it is separable or not?

Answering this question lies at the center of all of our present research.

Unfortunately[8] in 2002 it was proven by Leonid Gurvits that the separability problem belongs to the complexity class[9] of NP−HARD[10] problems [34], even in the bipartite case for sufficiently large-dimensional parts. This difficulty is made *much, much worse* in our case, as we are interested in systems with *many* parts!

[7] Indeed, one of the most intriguing features of quantum physics is that of how our everyday experience is dominated by seemingly separable states.

[8] Or rather fortunately, as it gives us an interesting problem to work on!

[9] The notion of a *complexity class* arises in computational complexity theory as a way to classify problems in terms of their inherent difficulty, that is, how much computational resources are required to solve them [33].

[10] NP−HARD stands for *non-deterministic polynomial-time hardness*, which refers to a class of problems that are, in a sense, at least as hard as the hardest problems in NP, the non-deterministic polynomial time class of problems whose answers can be verified in polynomial time [33]. Putting it bluntly, *very hard* problems!

Is there any hope for our pursuit?

Fortunately, there is! In special circumstances, it is possible to determine whether a system is separable or not simply by testing if we can manipulate it *as if it were* separable. If we *break*[11] the state by fiddling with it, then we can say with certainty that the state is not separable and therefore entangled. In certain cases, a measure of "how badly" we broke the state may even be used to quantify that entanglement, as we shall see later (chapter 3 section 3.3.4).

In effect, what we have is the following:

Definition 23. *A* ***separability criterion*** *is a procedure which can determine if a given class of states is separable or not.*

We will now very briefly discuss some of these criteria.

2.3.1 Peres criterion

One of the earliest, simplest and most important separability criteria was first suggested by Asher Peres [35] in 1996. The criterion is quite simply stated:

Theorem 4. *(Peres criterion or PPT criterion) A bi-partite quantum state ρ_{AB} is not separable if its partial transposition with respect to either one of its parts results in a negative operator.*

The criterion is typically stated only in terms of bipartite systems, as done above, but one should be careful to note there are many subtleties involved when we are dealing with larger systems, which are obscured in this approach. We shall cover these details explicitly later in section 2.4. For now, let us offer the simple proof.

Proof. The proof is by contradiction. Suppose ρ_{AB} *is* indeed separable. Then, it may be written as in definition 21:

$$\rho_{AB} = \sum_k p_k \, \rho_{A,k} \otimes \rho_{B,k}, \qquad p_k \in [0,1], \quad \sum_k p_k = 1.$$

By linearity of the partial transpose and its distributivity through tensor products, we have:

$$\rho^{\mathrm{T}_A} = \sum_k p_k \, \rho_{A,k}^{\mathrm{T}} \otimes \rho_{B,k}, \quad \text{or} \quad \rho^{\mathrm{T}_B} = \sum_k p_k \, \rho_{A,k} \otimes \rho_{B,k}^{\mathrm{T}}, \quad \text{since } \rho_{AB}^{\mathrm{T}_B} = (\rho_{AB}^{\mathrm{T}_A})^{\mathrm{T}}.$$

The matrix transpose is a positive map, therefore, if the state is separable both ρ_A^{T} and ρ_B^{T} are still positive semidefinite matrices, with non-negative eigenvalues. Since the eigenvalues of a tensor product are the products of the eigenvalues of each term, if either partial transpose is not positive semidefinite (if it has even a single negative eigenvalue), then ρ_{AB} could not have been separable, a contradiction.

□

The Peres criterion tells us a *necessary* condition that must be obeyed by all bi-separable states: they must be positive over all of their partial transpositions. For this reason, the criterion also goes by the name *"PPT criterion"*, standing for *Positive Partial Transpose.* The criterion is powerful and quite effective for small Hilbert spaces, although it relies on verifying the positivity of the partially transposed matrix, which for larger systems is not by itself an easy task!

[11] If we violate some property that all separable states must obey, or if the modified state violates the definition of a valid density matrix, as in definition 1.

The Peres criterion is not perfect, however. While a partial transposition returning a negative matrix guarantees that the state is entangled, a positive partial transposition does not guarantee separability, as we shall see next.

Our work will largely hinge on—and even extend upon—this criterion, so we will not focus on others as much. We will periodically return to it in the following chapters.

2.3.2 Peres-Horodecki criterion

Shortly after Peres' criterion was introduced, the Horodecki brothers[12] proved that checking for positive partial transpositions to identify separable states was not only necessary, but also sufficient for 2×2 and 2×3 dimensional systems [29].

In the same article, it was also proved that the criterion ceases to be sufficient for any higher dimensional case. This was done by demonstrating the existence of entangled states which have positive partial transpose, a type of entanglement now referred to as *bound entanglement* or *PPT entanglement*. We will discuss these further in chapter 3 section 3.1.

Woronowicz-Størmer theorem

The Horodecki brothers' result hinged on earlier theorems by Stanisław Woronowicz [36] and Erling Størmer [37], which applied to positive and unital linear maps in broader terms in the 2×2 and 2×3 cases. What the Horodeckis proved was that the matrix transpose map was the *only* such map for those spaces.

2.3.3 Other criteria

Over the years many other criteria were conceived, all with their own strengths and weaknesses. Typically, a class of states detected by one are not detected by others. Due to the $\mathtt{NP}-\mathrm{HARD}$-ness of the problem, no simple general technique is likely to exist[13].

An important class of criteria, more useful in higher dimensional systems[14], is the realignment or reshuffling criterion proposed by Oliver Rudolph [38] as well as Kai Chen and Ling-An Wu [39]. The criterion states that if a system is bi-separable, then reshuffling its matrix elements does not increase its trace norm. This is ultimately based on the observation that for any separable state, the sum of the singular values of the matrix should be less than or equal to 1.

While Rudolph, Chen and Wu's result was restricted to bipartite systems, it was later shown by the Horodecki brothers [11] that such a technique is available to multipartite systems, and can even detect higher-order entanglements. It was shown that any linear map which does not increase the trace norm on separable states may be used as a necessary condition for detecting entanglement on a subset of states, which is known as the contraction criterion [7]. We will not make use of these other criteria.

2.4 Multipartite separability and entanglement

In section 2.2 we defined entanglement and separability in broad terms, which we used in section 2.3 to talk about separability criteria for two-part systems.

Our work involves systems with many parts, however, so we must be more careful when defining separability: separability with respect to *which partition* of the system?

[12] Michał, Paweł, and Ryszard Horodecki.

[13] Assuming $\mathtt{P} \neq \mathtt{NP}$, which seems to be the case.

[14] For lower dimensional systems, the criterion can fail to detect entanglement in even the simplest cases. For example, there are classes of two-qubit states in which this criterion fails [7].

The issue arises because given a system of n parts, there are multiple ways to split it into k non-empty disjoint groups. For example, for $n = 4$ and $k = 2$, there are 7 unique partitions available,

$$1|234 \quad 2|134 \quad 3|124 \quad 4|123 \quad 12|34 \quad 13|24 \quad 14|23,$$

all of which are potentially separable or not. Therefore, the task of defining separability and entanglement in multipartite systems must take into account the inherent structures implied by such partitions, which cannot be assumed *a priori*. This leads us to a more refined definition of separability and entanglement, which we will establish now. We will largely be following the definitions given by Jungnitsch *et al.* [40], Dür *et al.* [41], and Seevinck *et al.* [42].

Consider a composite system of n parts described by the state $\rho_{(n)} \in \mathcal{M}_{(n)}$. Let $\alpha_k = \{S_1, \cdots, S_k\}$ denote a partition of the system into $k \leq n$ disjoint non-empty subsets S_i.

The set α_k has been referred to as a *k-partite split* [41] of the composite system. In order to avoid the present overuse of the term "partite", we shall stick to the following terminology instead:

Definition 24. *A **k-partition** α_k of a n-partite composite system is a set of k disjoint non-empty subsets $\{S_1, \cdots, S_k\}$ of the composite system.*

With this, we have

Definition 25. *A n-partite quantum state $\rho_{(n)}$ is said to be α_k-**separable**[15] under a specific k-partition α_k iff it is separable in terms of the k subsystems in this partition [41, 43–45],* i.e., iff *it can be written as a convex combination of product states under α_k:*

$$\rho_{(n)} = \sum_i p_i \bigotimes_{j=1}^{k} \rho_i^{S_j}, \qquad p_i \in [0,1], \quad \sum_i p_i = 1, \tag{2.3}$$

where $\rho_i^{S_j}$ is the state of i-th subsystem according to the subset S_j under the partition α_k.

By the construction given in equation 2.3, the set of all such α_k-separable states (for a fixed α_k) forms a convex set $\mathcal{M}_{(n)}^{\alpha_k\text{-sep}} \subset \mathcal{M}_{(n)}$. Any state $\rho_{(n)} \notin \mathcal{M}_{(n)}^{\alpha_k\text{-sep}}$ is said to be *α_k-inseparable* or *α_k-entangled.*

We can further generalize this by considering all such sets α_k, giving

Definition 26. *A state $\rho_{(n)}$ is **k-separable**[16] iff it can be written as a convex combination of α_k-separable states $\rho_i^{\alpha_k}$*

$$\rho_{(n)} = \sum_i p_i \, \rho_i^{\alpha_k}, \qquad p_i \in [0,1], \quad \sum_i p_i = 1. \tag{2.4}$$

Again, by construction this forms a convex set $\mathcal{M}_{(n)}^{k\text{-sep}} \in \mathcal{M}_{(n)}$. Any state $\rho_{(n)} \notin \mathcal{M}_{(n)}^{k\text{-sep}}$ is said to be *k-inseparable* or *k-entangled.*

[15] Also referred to as *partially separable* or *semi-separable*, but we will refrain from using that term as it can be ambiguously understood.

[16] Also referred to as *fully separable* or *fully k-separable*, but we will also refrain from using it as the word "full" can only result in confusion due to its connotations with product states or n-separability (as opposed to k-separability).

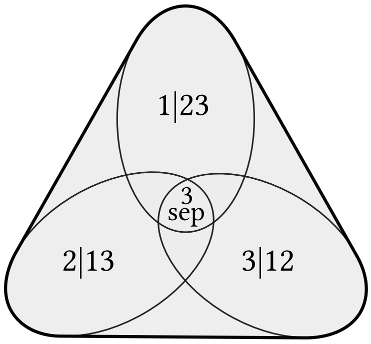


Figure 2.1: A visual representation of the α_k and k separable sets for $k = 2$ and $n = 3$, giving the three 2-partitions: $\alpha_2^{1|23}$, $\alpha_2^{2|13}$, and $\alpha_2^{3|12}$. The boundary of each ellipse represents the product states of their partition, the convex combinations of which results in their own α_2-separable set. At the intersection of all three ellipses, we have the convex set of 3-separable states. The convex hull (thick line) of the three α_2 sets forms the set of 2-separable states.

As detailed by Seevinck & Uffink [42] and Dür & Cirac [41, 45], the properties of convex combinations of different α_k or k-separable or α_k or k-entangled systems is very subtle and hard to predict. In figure 2.1 above, the three intersections between pairs of ellipses correspond to such sub-classes of 2-separable states. We shall not go into further details here, however.

2.4.1 Hierarchies of partitions

It is worth noting that a state which is separable for $\ell = k$ will also be separable for any $\ell < k$, since we may always treat arbitrary unions of subsets S_j of a given α_k k-partition as a single subsystem in a lesser-k partition. Thus, the notion of k-separability logically induces a hierarchy of k-separable convex sets:

$$\mathcal{M}_{(n)}^{n\text{-sep}} \subset \mathcal{M}_{(n)}^{(n-1)\text{-sep}} \subset \cdots \subset \mathcal{M}_{(n)}^{1\text{-sep}}.$$

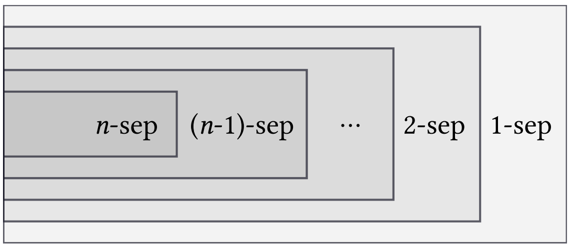


Figure 2.2: A visual representation of the hierarchical structure of the convex sets $\mathcal{M}_{(n)}^{k\text{-sep}}$.

2.4.2 Genuine multipartite entanglement

The concept of *genuine multipartite entanglement* initially arose with the study of 3-partite states, such as the GHZ state[17], in which the entanglement is observed to be intrinsically related to more than two subsystems [46–48].

With definition 26, we can propose the following

[17] Greenberger-Horne-Zeilinger state, first studied in detail 1989 [18].

Definition 27. *A state $\rho_{(n)}$ is said to be* ***genuinely***[18] ***k-partite entangled*** *(or simply* ***genuinely k-entangled****) when it is ℓ-entangled for all $2 \leq \ell \leq k$.*

These definitions allow us to assign a degree of k-entanglement and k-separability to any n-partite system, defining a hierarchical scale of degrees of entanglement and separability.

- At one extreme we have the n-entangled or *fully entangled* states (*e.g.* the GHZ states), in which all particles are mutually entangled and cannot be separated using any k-partition. Conversely, this can also be understood as a degenerate "1-separable" state or an *atomic state*, which can only be "separable" into a single part.

- At the other extreme we have the n-separable or *fully separable* states (*e.g.* product states, the maximally mixed state $\rho_\star$), in which all particles are uncorrelated. Conversely, this can also be understood as a degenerate "1-entangled" state, in which each particle is only "correlated to itself".

Regardless of which is the case, a genuinely entangled system is in general difficult to identify, as the k-separability problem falls on the NP−HARD-ness we have mentioned earlier just the same.

In practice, genuine k-entanglement may be identified by carefully studying the overlaps of all possible $(k-1)$-partitions of the composite system and excluding them. If all of the partitions are entangled, the system must necessary be at least k-entangled.

[18]Traditionally, 2-entanglement is not referred to as "genuine", although there is really no special reason to single-out 2-entanglement in this matter.

Chapter 3

Entanglement witnesses

While no operationally feasible and analytical technique exists to determine if a general state is entangled, there are numerical techniques which are capable of detecting *any* type of entangled state in any number of dimensions[1].

From these, of particular interest is the technique of constructing entanglement witness operators. This will be covered in more detail in section 3.2, but first we must take a small detour to get there.

3.1 The space of separable and entangled states

"It seems very pretty, but it's rather hard to understand!
Somehow it seems to fill my head with ideas
— only I don't exactly know what they are!" – Alice

Lewis Carroll, *Alice's Adventures in Wonderland*

As we have seen, Hilbert spaces are incomprehensibly vast. We may attempt to visualize it in terms of geometry we are familiar with, but this is quite challenging: the space of states for n-partite systems, $\mathcal{H}^D$, resides in a $(2^D - 1)$-dimensional[2] *real* space [7]. The simplest quantum system possible featuring entanglement, with 2 qubits, is already ***15-dimensional!***

The notion of k-separable and k-entangled states points to the existence of a detailed structure for separability and entanglement within the space of pure states $\mathcal{H}$ and of density matrices $\mathcal{M}$. We will not concern ourselves with a full or detailed structure of this space, as ultimately we will focus our attention on a very small part of it. Nonetheless, it is important to establish the basic landscape where our musings are taking place.

We will proceed to analyze $\mathcal{M}$ "from the inside out" in our description. More details on the geometry of quantum states may be found in Bengtsson and Życzkowski [7, Ch.15§5].

3.1.1 Separable mixed states

Definition 28. *If a mixed n-partite state $\rho \in \mathcal{M}$ is k-separable (definition 26), then it is an element of the set of k-separable states: $\rho \in \mathrm{Sep}_k(\mathcal{M}) \subset \mathcal{M}$. For $k = n$, we will use simply $\mathrm{Sep}(\mathcal{M})$.*

At the literal center of the space of density matrices $\mathcal{M}(\mathcal{H}^D)$ lies the maximally mixed state, $\rho_\star = \mathbb{I}/D$ [7], which as we have seen is n-separable. Surrounding this there several convex

[1] Provided if you have enough computational power to spare

[2] A dimension D system requires D complex coefficients, each of which may be regarded as subsets of $\mathbb{R}^2$, leading to 2^D dimensions. The requirement of normalization takes off one of the dimensions.

sets of the k-separable *mixed* states, one containing the other (as in figure 2.2), which rapidly becomes vanishingly small as D increases [31, 34].

3.1.2 PPT bound entanglement mixed states

Definition 29. *If a state $\rho \in \mathcal{M}$ is entangled (definition 22) and has a positive partial transpose in all of its bipartitions (as in theorem 4), then it is an element of the set of entangled **PPT states** $\rho \in \mathrm{PPTES}(\mathcal{M}) \subset \mathcal{M}$. The states in this set are said to contain **bound** or **non-distillable entanglement**[3] [5, 49].*

Moving outward, for $D > 6$, numerical evidence shows [50] there is a thin, non-uniform[4], non-zero volume shell of non-separable PPT states surrounding the separable states. These states cannot be identified by the Peres criterion (section 2.3.1, but other techniques covered in this chapter are applicable to detect them.

This shell disappears for $D \leq 6$, since the Peres-Horodecki criterion (section 2.3.2) establishes that *all* entangled states in these dimensions have a negative partial transpose, and are thus detectable by the Peres criterion.

Originally, it was an open question whether such states *actually* existed in Nature, due to poorly understood connections between partial transposes, local time reversal symmetries and non-locality [49, 51]. Some experimental procedures have since been suggested [52] and reported as successful. Some of these results have been either disputed [53] or merely pseudo-bound entangled states [54], but the phenomenon of bound entanglement is now established as experimentally feasible [55, 56].

As of yet, it is unclear whether this set is convex [57].

3.1.3 NPT entangled mixed states

Definition 30. *If a mixed state $\rho \in \mathcal{M}$ is entangled and has a negative partial transpose with respect to any of its 2-partitions, then it is an element of the set of **NPT states** $\rho \in \mathrm{NPT}(\mathcal{M}) \subset \mathcal{M}$. The states in this set are said to contain **free** or **distillable entanglement**.*

Outside the separable and PPT states lies the majority of the volume of $\mathcal{M}$ [7, 58]. A random state ρ uniformly sampled from $\mathcal{M}$ (with an appropriate metric [58, 59]) will almost certainly be of this type [7]. These states feature entanglement which may be distilled to be used as a resource in LOCC protocols [5], and are of great interest in quantum computation and quantum information. LOCC (Local Operations and Classical Communication) protocols are a class of restricted protocols in quantum information, wherein multiple parties sharing quantum states are restricted to classical communication for mutual coordination, while performing only local operations on their respective parts of the state. Such protocols are important as they represent physically realizable quantum protocols obeying relativistic causality contraints. A detailed discussion about such protocols is not relevant to our present work, but may be found elsewhere [5, 7].

3.1.4 NPT bound entanglement mixed states

It is an open question whether bound entanglement also exists for negative partial transpose mixed states [49]. In recent years this question has attracted the attention of a few researchers [60],

[3] In rough terms, entanglement which cannot be used as a resource (say, to obtain a maximally entangled pure state such as $|\Psi^-\rangle$ to be used for quantum teleportation) by two parties performing local operations and classical communication (LOCC) protocols [5]. This subject is out of the scope of this text, however.

[4] The set of entangled NPT states shares a border with set of separable states, so things become hard to visualize.

and limited evidence has been found that such states may exist [61]. No example seems to have been found as of the time this text was written.

3.1.5 NPT entangled pure states

At the boundaries of the convex set $\mathcal{M}$ we have many pure states [7], some of which are entangled. All of these states have a negative partial transpose and will be detectable by the Peres criterion under at least one 2-partition.

All of our present work is focused on properly handling this class of states, as they are mainly the ones currently being studied within the tensor network formalism[5].

3.1.6 Separable pure states

Pure states, which lie at the outer boundary of $\mathcal{M}$ may also be separable (in the k-separable sense). These states form the *pure points* of their respective k-separable convex sets [7].

3.1.7 Some remarks

As we shall see in more details in the next chapter, in our work we will be focusing on studying pure states, which only form a very small fraction of all the possible quantum states describable by density matrices.

Nonetheless, the subset of all pure states is arguably the most important set in practice. Pure separable and entangled states, which we shall collectively call by the name $\mathrm{Pure}\,(\mathcal{M})$[6], are the building blocks of all other mixed states, and are the states experimental physicists are the most interested in practice to understand quantum systems.

Even quantum information seeks to obtain pure states in practice, as they offer more entanglement to act as an informational resource to be used for various tasks[7]. Modern simulations of many-body quantum systems largely focus on searching for pure states of systems of interest, such as the ground state of important Hamiltonians or the time evolution of systems under action of external perturbations.

But do not be fooled: our focus on pure states does not mean our task will be simple or that the challenge is not great. As we will see next, even within the restricted set $\mathrm{Pure}\,(\mathcal{M})$ there is an immense richness and complexity which is highly non-trivial to analyze and study, and the full description of these pure states for large n-partite systems is still *monstrous.* We are still confronting a fierce and powerful adversary!

It is now time for us to get ready for our struggle in taming quantum entanglement.

[5]However, we are already looking into extending our approach to mixed states.

[6]Our own notation, for the sake of terseness.

[7]This is what entanglement distillation aims to accomplish: obtain maximally entangled states from several copies of mixed states. Once a maximally-entangled state is distilled, it may be used in various interesting experiments.

3.2 Entanglement witnesses

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the mome raths outgrabe.

"Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
The frumious Bandersnatch!"

He took his vorpal sword in hand:
Long time the manxome foe he sought –
So rested he by the Tumtum tree,
And stood awhile in thought.

And, as in uffish thought he stood,
The Jabberwock, with eyes of flame,
Came whiffling through the tulgey wood,
And burbled as it came!

One, two! One, two! And through and through
The vorpal blade went snicker-snack!
He left it dead, and with its head
He went galumphing back.

"And, has thou slain the Jabberwock?
Come to my arms, my beamish boy!
O frabjous day! Callooh! Callay!"
He chortled in his joy.

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the mome raths outgrabe.

Lewis Carroll
"Jabberwocky", from *Through the Looking Glass*

As we have seen, characterizing entanglement precisely is a tricky task. Since the separability problem is in general NP−HARD, we may instead resort to sufficiently efficient iterative numerical algorithms to solve[8] the problem in a case-by-case basis.

Fortunately, in the case of entanglement, there is a powerful and general numeric approach to this task, which unites a simple but important theorem in functional analysis with a powerful class of linear optimization algorithms.

And just as the illustriously sharp *vorpal sword* which slays the *Jabberwock*, our weapon for the task will *cut through* our foe with deadly precision—provided we also "stay awhile in thought" long enough to slay it!

3.2.1 The Hahn-Banach separation theorem

An important theorem arises in functional and convex analysis which allows for classification of points in a space in terms of linear functionals [7, 8]. Intuitively, such functionals may be thought of as hyperplanes which partition the space into two sets: a set in which the functional is negative and the other in which it is positive[9]. Stated formally:

Theorem 5. *(Hahn-Banach separation theorem [7, 8, 62]) Given a convex set $\mathcal{S}$, an element $x \notin \mathcal{S}$ and an arbitrary element $y \in \mathcal{S}$, then one can find a linear function f such that:*

$$f(y) > 0 \qquad \textit{and} \qquad f(x) < 0$$

The Hahn-Banach separation theorem[10] tells us that such functional *always* exists, and thus we can freely choose any element of the space to find such functionals, which will separate the point from the convex set. Due to the nature of convex sets, such a procedure guarantees that

[8] Up to any desirable accuracy, at least.
[9] The functional is defined so as to be zero on the hyperplane itself.
[10] Originally simply *hyperplane separation theorem*, a result in Euclidean geometry. United with the Hahn-Banach theorem, its results are extended to topological spaces [62].

there is always a (possibly infinite) family of functionals which collectively separate the convex set from the rest of the space.

In effect, together with the definition of convexity, this theorem tells us we can find a functional that reveals to us if a given element is contained within a convex set or not, by sharply cutting the space into two parts. In fact, due to the nature of linear functionals, it can also *quantify* that information, as we shall soon see.

The importance of this theorem for us is made evident when we recall the convex nature of all the sets we are interested in, in particular the convex set of separable states. This key insight and its applications are the main subject of this chapter.

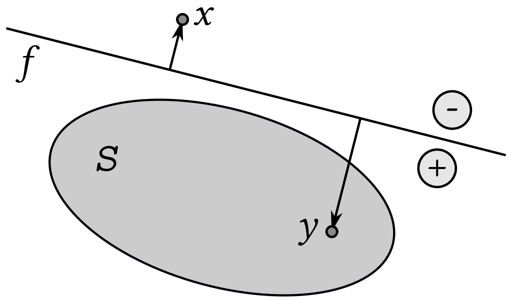


Figure 3.1: The Hahn-Banach separation theorem in 2-dimensional Euclidean space. The linear functional f (a linear equation in this case) separates the plane into two regions, a positive (on the side of the set $\mathcal{S}$ and y) and a negative (on the side of x). The functional also quantifies the distances of x and y to its cutting plane, represented by the arrows.

3.2.2 Defining entanglement witnesses

Using the Hahn-Banach separation theorem, we now want to define linear functionals $W \in \mathcal{B}(\mathcal{H})$ acting on the space $\mathcal{M}$ that can distinguish between separable and non-separable states. This leads to the following theorem:

Theorem 6. *(Entanglement Witness Theorem [7, 29, 63, 64]) A n-partite state $\rho \in \mathcal{M}$ is entangled iff there exists a Hermitian operator $W \in \mathcal{B}(\mathcal{H})$ such that $\mathrm{Tr}[\rho W] < 0$ and $\mathrm{Tr}[\sigma W] \geq 0$ for all separable states $\sigma \in \mathrm{Sep}\,(\mathcal{M})$.*

Definition 31. *The operator W is referred to as an **entanglement witness (EW)** for the state ρ.*

Proof. The proof follows immediately from the fact that the set of separable states, $\mathrm{Sep}\,(\mathcal{M})$, is a convex set. By the Hahn-Banach theorem, a functional with such properties exists and so we can distinguish ρ from any $\sigma \in \mathrm{Sep}\,(\mathcal{M})$. □

We will assume all entanglement witnesses are normalized, that is, $\mathrm{Tr}[W] = 1$.

3.2.3 Optimal entanglement witnesses

The numerical value of the functional, $\mathrm{Tr}[\rho W]$, gives information about the distance ρ is relative to the cutting hyperplane. But as it is, this information is not *useful* to us, as the functional is largely arbitrary. In order to quantify this distance in a meaningful way, we need a unique, privileged reference scale. To establish it, we first begin with

Definition 32. *Let $\Omega(W)$ be the set of entangled states detected by a witness W, and let there be two witness operators W_1 and W_2. We say witness W_2 is **finer** [65] than W_1 if every entangled state detected by W_1 can also be detected by W_2 that is, if $\Omega(W_1) \subset \Omega(W_2)$.*

This allows us to establish the following

Definition 33. *An entanglement witness W_{opt} is said to be an **optimal entanglement witness** [65] (**OEW**) if for an entangled state ρ, $\mathrm{Tr}\left[\rho W_{opt}\right] \leq \mathrm{Tr}[\rho W]$ for all witness W, that is, if there are no witnesses finer than W_{opt}. Equivalently [63, 66] a witness W_{opt} is optimal iff there is a separable state σ_0 for which $\mathrm{Tr}\left[\sigma_0 W_{opt}\right] = 0$.*

In intuitive terms, the optimal entanglement witness provides us with the *closest cut* to the convex set in which the witness is positive, as shown in figure 3.2 below.

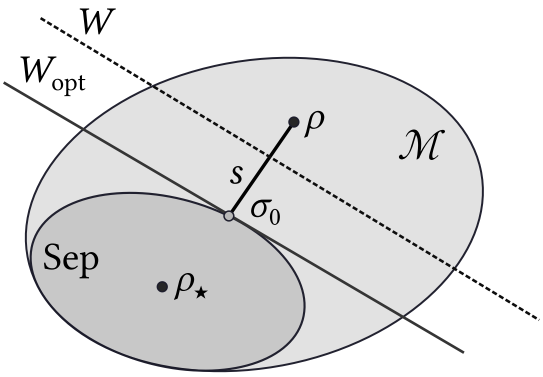


Figure 3.2: Conceptual illustration of the difference between two witness operators W and W_{opt} for a state ρ. The optimal witness' functional gives the best separation between the convex set and ρ. The hyperplane where $\mathrm{Tr}\left[\rho W_{\mathrm{opt}}\right] = 0$ is tangent to the convex set of separable states at the state σ_0, and the distance s can be related to the Hilbert-Schmidt distance (equation 1.6) between ρ and σ_0 [63, 66, 67]. There is no entangled state $\rho \in \mathcal{M}$ detected by W which is also not detected by W_{opt}.

Given these properties, we shall henceforth restrict our attention to optimal witnesses.

3.2.4 Decomposable and non-decomposable witnesses

Following results by Woronowicz [36], witnesses of 2-entanglement may be non-decomposable or decomposable operators, and this significantly changes their behavior [65]. As discussed at length by the Horodeckis [29,48,49,68], Kraus *et al.* [69], Clarisse [70,71], Bruß *et al.* [72], Lewenstein *et al.* [65], and Jungnitsch *et al.* [40], by construction decomposable witnesses cannot detect PPTES entanglement, whereas non-decomposable witnesses can. We will not get sidetracked into the details of their proofs, but the basic ideas are intuitive and deserve a mention.

Decomposable witnesses

Definition 34. *A witness operator W for a n-partite system is said to be a **decomposable witness** [69, 72] (**d-EW**) iff it can be written in the form:*

$$W = P + \sum_{\alpha} Q_{\alpha}^{\mathrm{T}_{\alpha}} \qquad Q_{\alpha}, P \succeq 0, \quad \textit{and} \quad Q_{\alpha}^{\mathrm{T}_{\alpha}} \preceq 0 \tag{3.1}$$

where α is a 2-partition from the set of all 2-partitions of the system, and T_α the partial transpose with respect to either one of the two parts in the 2-partition α.

As shown by the preceding authors, if W_2 is finer than W_1, then they differ from one another by a positive operator P. Intuitively, what this means is that the positive operator is a *perpendicular offset* on the hyperplane defined by the witness, which "pushes it away" from the convex set. By removing it, we bring the hyperplane to be tangent to the set, which is optimal [63].

This suggests that we may refine the entanglement witness above, as in

Theorem 7. *(Optimal Decomposable Witness [65]) Given a decomposable witness W_d as in equation 3.2, we can construct a finer and optimal decomposable witness via*

$$W_{d\text{-}opt} = (W_d - P) = \sum_\alpha Q_\alpha^{\mathrm{T}_\alpha}, \tag{3.2}$$

with Q_α and P as before, provided that the Q_α have no product states in their range.

Proof. The positive offset follows from the previous argument, but the requirement for no product states in the range requires further comment. If a 2-partite product state $|a, b\rangle$ is in the rage of Q_α, then we may write $W' \propto (Q_\alpha - \lambda |a, b\rangle\langle a, b|)^{\mathrm{T}_\alpha}$ for some $\lambda > 0$. This can be used to define a further refinement on W. Therefore, if we cannot do it, the witness is optimal. □

An important advantage of optimal decomposable witnesses is that they can be constructed directly with existing optimization algorithms, by finding the projectors for the negative subspace of the terms in $W_{\text{d-opt}}$ directly. We will discuss this approach in detail in section 5.4.1.

Due to their construction, the value of the functional $\mathrm{Tr}\left[\rho W_{\text{d-opt}}\right]$ is the sum of the negative eigenvalues of all negative projectors for the terms $Q_\alpha^{\mathrm{T}_\alpha}$. This quantity is referred to as *negativity*, and was shown by Vidal *et al.* and Jungnitsch *et al.* to be a computable entanglement measure [28, 40]. We will discuss this further shortly.

Decomposable witnesses are limited to detecting NPT entanglement via 2-partitions, as they are based on the same underlying principles as the Peres criterion (section 2.3.1). Therefore, they cannot see entanglement in states with a positive partial transposition (the PPTES states). For this task, we need non-decomposable witnesses, to be described next.

Non-decomposable witnesses

Definition 35. *A witness operator W is said to be **non-decomposable (nd-EW)** iff it can be written in the form [65, 70]:*

$$W = P + \sum_\alpha Q_\alpha^{\mathrm{T}_\alpha} - \epsilon \mathbb{I}, \qquad P \succeq 0, \quad \textit{and} \quad Q_\alpha^{\mathrm{T}_\alpha} \preceq 0 \tag{3.3}$$

with

$$0 < \epsilon < \inf_{|a,b\rangle} \left\{ \langle a, b| \left(P + \sum_\alpha Q_\alpha^{\mathrm{T}_\alpha} \right) |a, b\rangle \right\}.$$

If this witness is optimal, we refer to it simply as W_{opt}, and it is always finer than $W_{d\text{-}opt}$.

Proof of this result can be found in the references given. These witnesses are much more general than decomposable ones, and are capable of detecting any type of entanglement, including the PPTES states. However, their construction and refinement is highly non-trivial in general, as it was in the case of decomposable witnesses, and require the use of iterative algorithms [70]. This difficult optimization procedure is suggested by the *infimum* condition, specified in the definition. The problem in general is NP−HARD.

In our current work we will not be addressing optimization of these witnesses.

Fully decomposable witnesses

There is a subclass of much stricter decomposable witnesses, referred to as *fully decomposable witnesses* [40, 73, 74], which are not as trivial to obtain as the ones described previously. In fact, just as the non-decomposable ones, their optimization is NP−HARD [73].

Whereas the decomposable witnesses just described are based on the union of sets of projectors for the negative spaces of each 2-partition, the fully decomposable witnesses require the *intersection* of these sets. Mathematically, we have

Definition 36. *A witness operator W_G for a n-partite system is said to be an (optimal)* ***fully decomposable witness (fd-EW)*** *iff it can be written in the form:*

$$W_G = Q_\alpha^{\mathrm{T}_\alpha} \qquad \textit{for all} \quad Q_\alpha \succeq 0. \tag{3.4}$$

where α is a 2-partition from the set of all 2-partitions of the system, and T_α the partial transpose with respect to either one of the two parts in the 2-partition α.

In other words, we need an entire set of Q_α operators which are interrelated via partial transpositions. Naturally, this type of witness *still* cannot detect PPTES states. However, it can detect genuine 3-partite NPT entanglement or higher [40, 73], which is why we will opt for the notation W_G in describing it.

The value of the functional $\mathrm{Tr}[\rho W_\mathrm{G}]$ is also related to the negativity of ρ. This quantity is referred to as *Genuine Multipartite Negativity* [40, 73], and was also shown by Jungnitsch *et al.* to be an entanglement measure [40].

Conclusions on entanglement witnesses

Figure 3.3 visually explains the difference between the two main types of optimal entanglement witnesses.

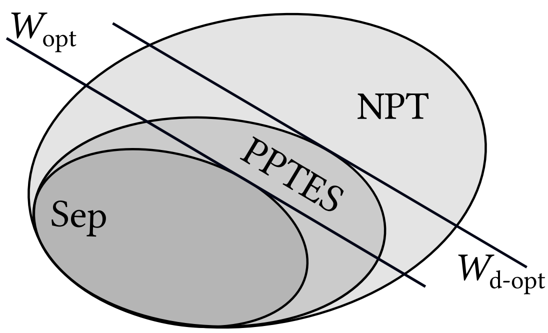


Figure 3.3: Conceptual illustration of the difference between optimal decomposable and a non-decomposable entanglement witness. By construction, decomposable witnesses cannot detect entanglement within the set PPTES.

As we can see, while entanglement witnesses are extremely precise tools to discern entangled and separable states, in practice their implementation can be quite laborious. Even so, there are efficient polynomial-time algorithms to approach this problem, which we will cover soon.

3.3 Entanglement characterization via entanglement measures

Now that have defined our witnesses we must interpret *what* they are measuring in order to use them effectively. As we have seen, entanglement measures need to have certain desirable properties.

There is a vast number of different *measures of entanglement* which can be used to quantify and characterize it [7, Ch.15 §6], and for mixed states the details can get very complex [75–78]. At the very minimum [7], an entanglement measure satisfies the following

Definition 37. *A (minimally) useful* ***entanglement measure*** *is a function $E : \rho \mapsto \mathbb{R}_+$ obeying*

- (I) ***Weak discriminance****: if ρ is separable, then $E(\rho) = 0$.*
- (II) ***Monotonicity****: invariance under unitary transformations*

$$E(\rho) = E(U\rho U^\dagger),$$

and non-increasing under deterministic LOCC maps Φ_{LOCC}:

$$E(\rho) \geq E\left(\Phi_{LOCC}(\rho)\right).$$

- (III) ***Convexity****:*

$$E\left[t\rho_1 + (1-t)\rho_2\right] \leq tE(\rho_1) + (1-t)E(\rho_2), \quad \textit{with} \quad t \in [0,1].$$

Under ideal conditions we would have **strong discriminance** instead of (I), where $E(\rho) = 0$ *iff* ρ is separable. This restriction turns out to be too strong in general [7], so we settle for less. Property (II) ensures the measure *makes sense* in the picture of entanglement as a resource [5] and invariance under *local* unitaries [7]. Property (III) ensures overall consistency with the fact convex combinations of states are also states. Together, properties (II)+(III) define *entanglement monotones* [7] which have numerous important properties.

There are many other desirable properties proposed in the literature [7, Ch.15 §6], although there are certain inconsistencies on which are deemed more important. Of particular note to our cause is that we want a measure to be *efficiently computable*, otherwise it will be of no use considering the large systems we will deal with. Some other desirable properties are

- (IV) **Continuity**: If $\|\rho_1 - \rho_2\|_1 \to 0$, then $|E(\rho_1) - E(\rho_2)| \to 0$.
- (V) **Additivity**: $E(\rho_2 \otimes \rho_2) = E(\rho_1) + E(\rho_2)$.

Property (IV) ensures our function is well-behaved under continuous state changes. Property (V) would hold for an optimal entanglement measure, although it is usually hard to achieve. A more lax version of (V) is *extensivity*, which is achievable if the following limit E_R (the *regularized* measure) exists:

$$E(\rho^{\otimes n}) = nE(\rho) \quad \leftarrow \quad E_R(\rho) = \lim_{n\to\infty} \frac{1}{n} E(\rho^{\otimes n})$$

Keeping our specific requirements in mind, four measures come to our attention: the Hilbert-Schmidt distance, the entanglement entropy, negativity and robustness.

3.3.1 Hilbert-Schmidt distance

For bipartite states, the Bertlmann-Narnhofer-Thirring Theorem [63] establishes that the maximum violation of the entanglement witness inequality

$$\mathrm{Tr}[\rho_{\mathrm{ref}} W] - \mathrm{Tr}[\rho W] \geq 0, \qquad \text{for all} \quad \rho,$$

is given precisely by the Hilbert-Schmidt distance between the maximally entangled state $\rho = \rho_{\max}$ (a Bell state) and the tangent state $\rho_{\mathrm{ref}} = \sigma_0$ for the optimal witness. This establishes the geometric meaning behind optimal entanglement witnesses, as the value of the functional $\mathrm{Tr}\left[\rho W_{\mathrm{opt}}\right]$ can be interpreted as the Hilbert-Schmidt distance to the closest point in the hyperplane tangent to the convex set [63,66,67]. With this, we may define he following entanglement measure [63]:

$$E_{\mathrm{HS}}(\rho) = \min_{\sigma \in \mathrm{Sep}} \|\rho - \sigma\|_2,$$

which is evidently greater than zero for all non-separable states and zero only if $\rho \in \mathrm{Sep}$, as desirable.

3.3.2 Entanglement entropy

A ubiquitous measure for pure states is the *entanglement entropy*, given by

Definition 38. *The **entanglement entropy** of a n-partite state $\rho \in \mathcal{M}(\mathcal{H}_1 \otimes \cdots \mathcal{H}_n)$ relative to a subsystem $\mathcal{H}_k$ is given by the von Neumann entropy of the reduced state ρ_k:*

$$E_E(\rho, k) = S(\rho_k) \qquad \textit{with} \quad \rho_k = \mathrm{Tr}_{\{1,\cdots,n\}\setminus k}[\rho]\,. \tag{3.5}$$

Here, $\mathcal{H}_k$ may be a subset of the n systems.

This measure of entanglement is easy to compute, but it relies on *discarding* part of the entire system and looking at only the remaining part of it. This can be problematic, as this procedure may destroy some of the information we want to quantify. The von Neumann entropy is also not appropriate for characterizing entanglement in mixed states [5, 7], although this will be of no concern to us.

3.3.3 Negativity

In the case of decomposable witnesses we have that all the contributions come from the projectors in the subspace of negative eigenvalues of ρ^{T_α}. This suggests the *negativity measure*:

$$\mathcal{N}(\rho^{\mathrm{T}_\alpha}) = \sum_{\lambda_i < 0} |\lambda_i|$$

The negativity is a convex function of ρ, so it satisfies property (III) in definition 37. Negativity has been shown to be an entanglement monotone [28,79], and therefore it is non-increasing under LOCC.

Naturally, from the definition the negativity cannot say anything about PPTES states, but as we will be focusing on pure states this will be of no concern to us.

It is also worth noticing that the negativity is fundamentally a measure of 2-inseparability only, and cannot say anything about the order of the entanglement detected. For that, we need to rely on the Genuine Multipartite Negativity, which we will introduce shortly.

Logarithmic Negativity

A particularly useful extension of the negativity is defined simply as

$$E_N(\rho) = \log_2(2\mathcal{N}(\rho) + 1) \tag{3.6}$$

It has many useful properties, of which we highlight:

- Does not reduce to the entropy of entanglement on pure states, like other measures. We will make great use of this property.
- It is additive (property (V) in definition 37).

3.3.4 Genuine Multipartite Negativity

The negativity is computable, which means it is directly accessible given a state by means of partial transposition and eigendecomposition. As we have pointed out, it can only tells us about the 2-inseparability of states, and it is not obvious how to interpret its value in terms of the k-entanglement order we have previously established.

But even NPT states may have genuine k-entanglement of higher orders, and it is important to be able to characterize them. To do this, Jungnitsch *et al.* [40] introduced the *Genuine Multipartite Negativity* as an entanglement monotone, computed using the fully decomposable witnesses introduced in section 3.2.4.

$$\mathcal{N}_{\mathrm{G}}(\rho) = -\operatorname{Tr}[\rho W_{\mathrm{G}}]\,, \tag{3.7}$$

with W_{G} defined as in 36. This is also a valid entanglement measure, which was extensively characterized by Hofmann [73, 74], and also has its own logarithmic version with similar properties.

3.3.5 Robustness

Another useful entanglement measure is the notion of *robustness*, which intuitively determines how much an entangled state can be perturbed (mixed with another) before it becomes separable. All measures of robustness below have been shown to be proper entanglement measures [80,81].

The *robustness of entanglement* was originally introduced by Vidal and Tarrach [80], given by the following

Definition 39. *Given a state $\rho \in \mathcal{M}$ and a separable state $\sigma \in \mathrm{Sep}$, we call the* ***robustness*** *of ρ relative to σ, $R(\rho|\sigma)$, the minimal $s \geq 0$ such that*

$$\frac{\rho + s\sigma}{1+s} \in \mathrm{Sep}\,.$$

Given this definition, the robustness might be infinite [80].

Random robustness

Vidal *et al.* go on to define the *random robustness*.

Definition 40. *The* ***random robustness*** *of ρ is its robustness relative to the maximally mixed state, $\rho_\star$, that is:*

$$R_R(\rho) = R(\rho|\rho_\star) = \min_s \left\{ \frac{\rho + s\mathbb{I}/D}{1+s} \in \mathrm{Sep} \right\}.$$

With this definition, it is clear that the random robustness is zero *iff* ρ is separable. A geometric representation of the random robustness is given in figure 3.4.

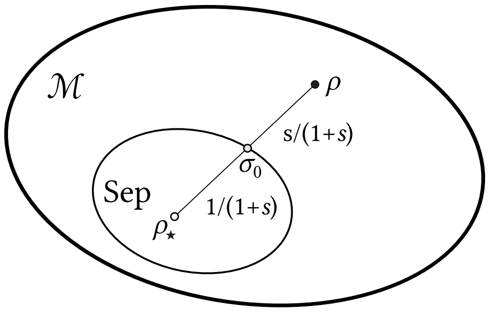


Figure 3.4: Geometric interpretation for the random robustness. The separable state σ_0 is the state which minimizes s, and defines the random robustness.

The random robustness has proven itself to be very useful experimentally, when one considers how stable an entangled state is when subject to experimental noise. The noise is typically modeled as a mixture with the maximally mixed state, which introduces uncertainty into the system [10, 80, 81].

Generalized robustness

Later, Steiner [81] highlighted the fact that convex mixtures of two entangled states may themselves result in separable states. This lead him to define

Definition 41. *The **generalized robustness** of ρ is given by*

$$R_G(\rho) = \min_{s|\tau\in\mathcal{M}} \left\{ \frac{\rho + s\tau}{1+s} \in \mathrm{Sep} \right\},$$

that is, the minimum value of s for which there exists a state τ that can be mixed with ρ to make it separable.

3.4 Semidefinite programming

"The purpose of computing is insight, not numbers."

Richard Hamming

Entanglement witnesses are linear Hermitian operators living in $\mathcal{B}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$, and the functionals they represent are also linear and bound by linear constraints. We can use techniques from convex optimization to tackle such problems, and this section will cover one such approach applied to the entanglement witnesses we have just introduced.

3.4.1 Convex optimization and Semidefinite Programming (SDP)

Convex optimization refers to a class of problems (and related algorithms) involving finding extrema (minima our maxima) of convex functions constrained to convex sets. Such type of problems are bountiful in mathematics, physics, engineering, economics and many other fields [62].

In a convex optimization problem, the function to be extremized is referred to as the *objective function*[11], and the set defined by the constraints as the *feasible set* [62].

Semidefinite programs (SDPs) are a subclass of convex optimization problems which can be written as a minimization of a *linear objective function* subject to *semidefinite constraints*, that is, the optimization variables are confined to the feasible set of positive semidefinite matrices [62].

These constraints can be expressed in terms of matrix inequalities by virtue of the space of semidefinite matrices being a *convex cone* [5, 7], which allow for the construction of a partial ordering between matrices [6, 62].

There are many equivalent different ways to specify a semidefinite program [62, Ch.4 §6.2], but in general, a typical form will be[12]:

$$\begin{aligned} &\textit{minimize}: && f(x) = c^\dagger x \\ &\textit{optimizing}: && x \\ &\textit{subject to}: && F(x) = F_0 + \sum_{i=1}^{m} x_i F_i \succeq 0 \end{aligned}$$

Here, x is a variable vector to be optimized and c a vector defining the objective function $f(x)$. The Hermitian matrices F_0, F_i are given and define a positive semidefinite constraint. This type of problem is called *convex conical optimization* since the feasible region, the values of x which satisfy the constraint, is a convex set, and the objective function is linear, which together with the convex set defines a convex cone. In short, if x and y are feasible, so is a convex combination of them. The solution to the problem, $p_\star = c^{\mathrm{T}} x_\star$, is referred to as the *objective*.

Semidefinite programs benefit from having no local minima: if there are more than one minima, then all of them are global minima. When the unique solution cannot be found analytically it can be estimated numerically using efficient polynomial time algorithms [62, 82, 83] for approximating solutions to arbitrary precision.

3.4.2 Duality

Another important benefit of SDPs is that each problem (referred to as *primal*) has a *dual problem*, defined as [62]:

$$\begin{aligned} &\textit{maximize}: && -\operatorname{Tr}[F_0 Z] \\ &\textit{optimizing}: && Z \\ &\textit{subject to}: && \operatorname{Tr}[F_i Z] = c_i, \qquad \text{for all} \quad i = 1, \cdots, m \\ &&& Z \succeq 0 \end{aligned}$$

The solution to a dual problem, the objective $d_\star = -\operatorname{Tr}[F_0 Z_\star]$, defines a lower bound to the primal problem. Both problems converge towards one another, but may not necessarily match.

The difference between the two, $p_\star - d_\star$ is referred to as the *duality gap*, and it converges to a finite value. This leads to the question of whether this gap converges to zero, and this is answered by

[11]Depending on the context, also a *loss function*

[12]For clarity, we prefer to include an explicit definition of *what* is being optimized. Most authors simply include the objective function and constraints, or place the optimizing variable as a subscript. This is unnecessarily obtuse, in my opinion.

Theorem 8. *(Strong Duality Theorem [8,62,74]) The primal and dual objectives are equal, $p_\star = d_\star$, if either:*

- *The primal problem is strictly feasible,* i.e.*, there exists an x with $F(x) \succ 0$*
- *The dual problem is strictly feasible,* i.e.*, there exists a Hermitian $Z \succ 0$ and $\mathrm{Tr}[F_i Z] = c_i$ for all $i = 1, \cdots, m$.*

In most cases, strong duality is used simply to establish good heuristics for stopping conditions for algorithms [62]. For most practical purposes, the stopping condition is merely the desired precision.

3.5 Finding entanglement witnesses with Robust SDPs

So how can we use SDPs to obtain entanglement witnesses? In the simplest case as in definition 31, we would need to impose an infinite number of constraints for separable states, since we need the witness to be positive for all of them.

This is easy to do in theory, but in practice that definition is not useful, as if we had such a precise understanding of the separable set we wouldn't have to worry about numerical optimizations in the first place: we would have in our hands an operational sufficient and necessary separability criterion, which we know is a NP−HARD problem. It cannot be that easy.

How can we proceed? As pointed out by Vianna and Brandão [84, 85] and Brandão [86], what can be done instead is to *approximate* the witness by constructing a *pseudo*-entanglement witness via *Robust SDPs*, which use many constraints for an appropriate optimization of the witness.

Theorem 9. *(RSDP Witness [84]) A n-partite state $\rho \in \mathcal{M}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$ is entangled (violates definition 21) iff the optimal value of the following RSDP is negative:*

$$\begin{aligned} \text{minimize:} \quad & \mathrm{Tr}[\rho W] \\ \text{optimizing:} \quad & W \\ \text{subject to:} \quad & \sum_{i_1=1}^{d_n} \sum_{j_1=1}^{d_n} \cdots \sum_{i_{n-1}=1}^{d_n} \sum_{j_{n-1}=1}^{d_n} \left(a_{i_1}^* \cdots a_{i_{n-1}}^* a_{j_1} \cdots a_{j_{n-1}} W_{i_1 \cdots i_{n-1} j_1 \cdots j_{n-1}} \right) \succeq 0 \\ & \mathrm{Tr}[W] = 1, W = W^\dagger \\ & \forall\, a_{i_k} \in \mathbb{C}, \quad 1 \le k \le n \end{aligned}$$

where d_n is the dimension of $\mathcal{H}_n$, $W_{i_1 \cdots i_{n-1} j_1 \cdots j_{n-1}} = \langle i_1, \cdots, i_{n-1} | W | j_1, \cdots, j_{n-1} \rangle \in \mathcal{B}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_{n-1})$ and $|j\rangle_k$ is an orthonormal basis for $\mathcal{H}_k$. If ρ is entangled, the solution matrix W which minimizes $\mathrm{Tr}[\rho W]$ is the optimal witness for ρ.

The proof and more details may be found in Brandão and Vianna [84]. We include it here due to its brevity and important as a context.

Proof. As previously defined, ρ is entangled *iff* there is an operator W such that $\mathrm{Tr}[\rho W] \le 0$ and $\mathrm{Tr}[\sigma W] > 0$ for all separable states σ. The latter condition can be expressed by requiring

$$\langle \psi_1 | \otimes \cdots \otimes \langle \psi_n | W | \psi_n \rangle \otimes \cdots \otimes | \psi_1 \rangle \ge 0, \text{for all} \quad |\psi_k\rangle \in \mathcal{H}_k.$$

Additionally, the matrix for each subsystem k must also be positive semidefinite. We can single out $k = n$ for our target system, so that the matrix:

$$\langle\psi_1|\otimes\cdots\otimes\langle\psi_{n-1}|\,W\,|\psi_{n-1}\rangle\otimes\cdots\otimes|\psi_1\rangle\geq 0,$$

for all $|\psi_n\rangle \in \mathcal{H}_n$. If we set $|\psi_n\rangle = \sum_j a_{j_n}|j_n\rangle$, with $\{|j_n\rangle\}$ an orthonormal basis for $\mathcal{H}_k$, we obtain the SDP in 9. □

It is important to understand what this theorem is telling us. A robust SDP is a way to turn such infinitely-constrained optimization problems into a typical SDP by imposing the separability in the rest of the space as a constraint. This separability condition is the requirement for positive semidefiniteness, and we can enforce it approximately.

The idea is quite intuitive: since we know the general form of separable states (definition 21) which go into the constraint, we can simply generate[13] many such states covering the separable space for the $n-1$ subsystems, leaving just one for a typical SDP optimization. Then, we optimize our pseudo-witness relative to the approximation of the set of separable states provided by the constraints, as in figure 3.5.

Our pseudo-witness will be optimized relative to the convex set defined by the constraints, which for a sufficiently large sample of separable states will be a good approximation of the entire set Sep, and thus, of our optimal witness.

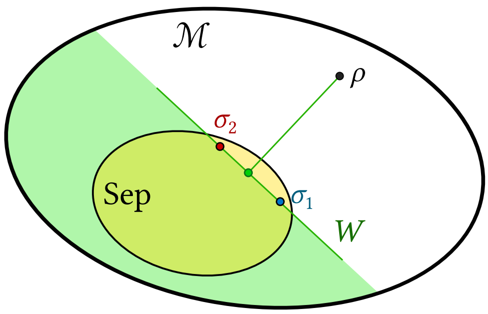


Figure 3.5: Geometric interpretation of the robust SDP. Two randomized separable states σ_1 and σ_2 are generated. The simultaneous constraint that the witness must be positive on both defines a convex region (in green) which is a rough approximation of the convex set of separable states. As more and more separable states are added as constraints, the better this approximation will be.

3.6 Final remarks

In this chapter we have seen that the separability problem is approachable with the right mathematical and computational toolsets. However, all of these implementations still are limited to the study to small-dimensional systems, as the optimizations rely on the density matrix formalism to represent quantum states and the witness operators.

As we have seen in chapter 2, this approach is intractable if we ever hope to look into large systems ($n > 10$). In the next chapter, we will be looking at how tensor networks allow us to break free from these limitations, adding to our extremely versatile toolset.

[13]There are techniques to generate separable edge states nearer to the region of interest.

Chapter 4

Tensor networks

Come, let us hasten to a higher plane
Where dyads tread the fairy fields of Venn,
Their indices bedecked from one to n
Commingled in an endless Markov chain!

I'll grant thee random access to my heart,
Thou'lt tell me all the constants of thy love;
And so we two shall all love's lemmas prove,
And in our bound partition never part.

Cancel me not – for what then shall remain?
abscissas some mantissas, modules, modes,
A root or two, a torus and a node:
The inverse of my verse, a null domain.

"Love and Tensor Algebra"
Stanisław Lem, *The Cyberiad*

The study of large-scale composite quantum systems with strong correlations has been met with challenging limitations in the past. Most algebraic and numeric approaches restricted themselves to small-dimensional cases, as these are easier to consider and simulate, or to very rough approximated models (*e.g.* perturbation theory, mean field theory, *etc.*) which can be well studied analytically and numerically, but which have limited applications due to their discarding of important properties related to quantum entanglement.

Therefore, many important phenomena in physics, in particular condensed matter physics, are still poorly understood due to these limitations. For example, the behavior of certain large-scale systems under phase transitions [87, 88] is still largely unexplored.

Significant advancements have been made in the past two decades with the introduction and development of tensor network algorithms for the study of such large-scale systems [87]. Since then, tensor networks have proven themselves to be *extremely* versatile tools [87, 88].

Inspired by this pioneering field, our research turned into studying precisely that which has been discarded in previous approaches: the entanglement properties of strongly-correlated large-scale quantum systems. But before we get to that, we shall present a brief overview of tensor networks and key algorithms in order to establish the inner workings of this novel and extremely powerful tool.

4.1 A quick introduction to tensors

Tensors are the natural generalization of the notion of scalars, vectors and matrices which greatly simplify the representation of multilinear maps between spaces [89].

Given a field $\mathcal{F}$, a scalar is taken to be a rank-0 tensor[1], a vector a rank-1 tensor and a matrix

[1]Some authors choose to use the term *order* instead of *rank*, as in certain abstract definitions the term rank is reserved for the multilinear generalization of matrix ranks. This will not be the case here.

a rank-2 tensor, with higher-rank tensors being referred to simply as tensors [89, 90]. For our limited purposes, however, we can mostly consider rank-d tensors to simply be d-dimensional arrays of complex numbers encoding a linear map between spaces and dual spaces[2].

Tensors are typically notated in terms of their indices, making explicit the structure of its elements. For example, the vector $x = (x_1, x_2, x_3)$ would instead be written as x_i, with the index $i = 1, 2, 3$ being referred to as having dimension $|i| = 3$. A $m \times n$ matrix M would be denoted as M_{ij}, with the indices i and j having dimensions m and n, respectively. The total number of elements in a tensor is obtained by the product of the dimensions of its indices, as expected.

Typically, with respect to tensors, indices may appear as subscripts or superscripts which carry additional meaning. Subscripts typically indicate covariance, and superscripts contravariance[3]. This notational convention is exceptionally useful when the tensors are being acted upon directly by linear transformations [89, 90], but since this is not going to be our case we will not use it. We will, however, make use of lower and upper indices in a different convention we will establish soon.

As with scalars, vectors, and matrices, tensors can be multiplied in much the same way. A typical complex vector inner-product can be written as

$$u^\dagger v = \sum_i u_i^* v_i,$$

and a typical matrix product as

$$AB = \sum_j A_{ij} B_{jk} = C_{ik},$$

both of which reveal the general structure of tensor products as a multilinear map on their respective spaces. A tensor can then be understood as an object which maps tensors into other tensors in a multilinear way.

In multilinear maps specified by tensor multiplication the sums are always performed on the common indices, as notated, and are referred to as a *tensor index contraction*. The common index disappears during the sum, and any remaining indices specify the structure of the resulting tensor.

Since this operation always involves a sum on repeated indices of the same dimension, a very useful convention introduced by Albert Einstein allows us to omit the summation symbol altogether, having it be understood implicitly by merely observing any indices repeated exactly twice. If available in $\mathcal{F}$, as is the case in this text where $\mathcal{F} = \mathbb{C}$, the commutative nature of scalar products and additions allow us to specify multiple contractions simultaneously, and the tensors in any order we desire:

$$A_{ijk} B_{klm} C_{il} = \sum_{ilk} A_{ijk} B_{klm} C_{il} = D_{jm}. \tag{4.1}$$

This notation also proves convenient to specify many other operations in linear algebra. The trace of a matrix M can be stated simply as $\mathrm{Tr}[M] = M_{ii}$, and a tensor product[4] $A \otimes B = C$ as $A_{ij} B_{kl} = C_{ijkl}$. A well known property of tensor products is that while they are not commutative, that is, $A \otimes B \neq B \otimes A$ in general, both products are similar up to a change of basis.

[2] We justify this concrete representation since our tensors will always be explicitly base-dependent, as we will see soon.

[3] Covariance and contravariance relate to how the tensors change under linear transformations.

[4] Also referred to as Kronecker product or outer product

This becomes evident in this notation, as this change of basis results in a simple permutation of pairs of indices.

We will not go into much more details on tensors and this brief introduction is all that is necessary to understand in what follows. As in the style of this text so far, additional details will be introduced as they become necessary.

4.2 Tensor network diagrams

A powerful way to gain intuition for tensor operations is by using a visual notation originally proposed by Roger Penrose [91]. This type of notation has since been informally extended significantly [87, 88, 92, 93] and is now typically referred to as *tensor network diagrams*.

We will very briefly discuss some of the main features of the notation and their meaning in terms of operations on tensors.

4.2.1 The basics

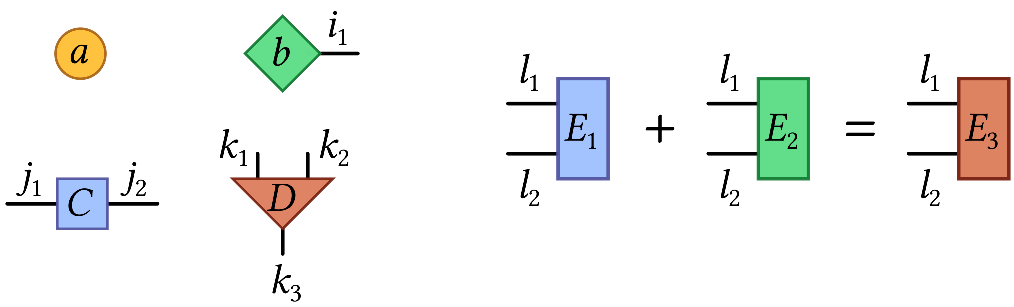


Figure 4.1: Basics of tensor network diagrams. Each shape represents a tensor, and each line segment or leg an index: a scalar a represented as a yellow circle, a vector b as a green diamond, a matrix C as a blue square and a rank-3 tensor D as a red triangle. On the right, two tensors with the same index structure E_1 and E_2 are added to obtain E_3.

In tensor network diagrams, individual rank-n tensors are represented by geometric shapes (*e.g.* circles, rectangles, triangles, *etc.*) with n lines coming out of them. There is no standard meaning behind shapes, and different authors and contexts define their own conventions. We will stick with squares and rectangles for general tensors, circles for unitary matrices, and the occasional diamond for diagonal matrices. Colors will be used to highlight different objects.

The lines coming out of the shapes are typically referred to as *legs* or *wires* [93], and each represents an index of their tensor (figure 4.1). A scalar is a tensor with no legs, a vector a tensor with one leg, and so forth. Tensors with the same index structure may be added or subtracted together, and this is represented by writing plus or minus signs between multiple tensors (figure 4.1).

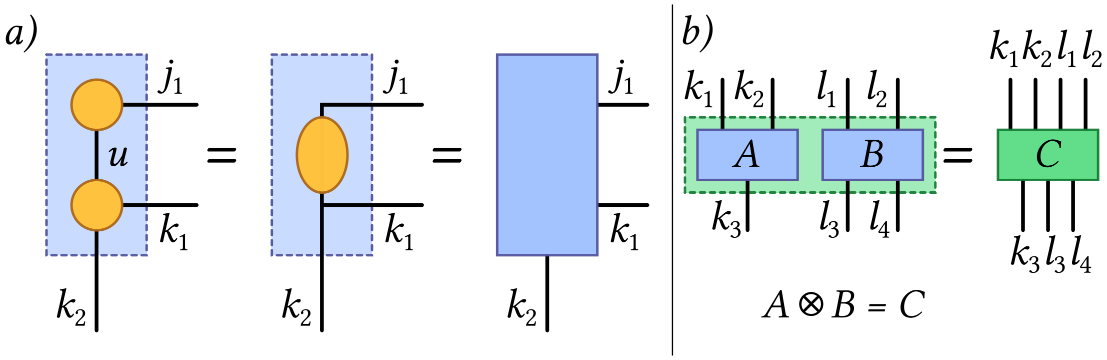


Figure 4.2: (a) Tensors which share legs (*e.g.* the two yellow circles sharing u) can be contracted into other tensors. By drawing shapes around multiple tensors (blue rectangle), we can infer the structure of the tensor contractions. (b) The tensor product between two tensors $A \otimes B$ is just a juxtaposition of the two tensors. The result is a single tensor C.

Two tensors sharing a leg may be contracted by summing over the common index, resulting in a new tensor. The total tensor of a contracted network may be seen by drawing a bounding shape around the network and seeing which indices come out of it (figure 4.2a). Therefore, the contraction of a network is itself also a tensor.

A tensor product (alternatively, outer product or Kronecker product) between two tensors $A \otimes B$ is automatically represented by simply putting the two tensors adjacent to one another. The bounding shape around them reveals the structure of the final tensor (4.2b).

4.2.2 Index operations

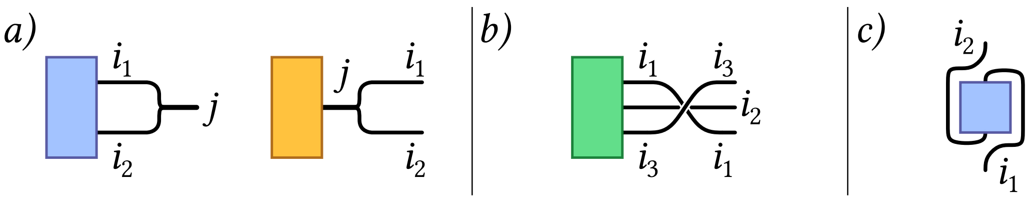


Figure 4.3: (a) Indices can be fused $((i_1, i_2) \mapsto j)$ or split $(j \mapsto (i_1, i_2))$, which reshapes the tensor. (b) Indices can be reshuffled by twisting the legs of a tensor. (c) A matrix transpose swaps the two indices.

There is no meaning attached to the direction in which the legs are pointing. All that matters is the connectivity implied by the diagram.

Fusing and splitting

Indices may be fused together or split into separate indices. The total dimension of a fused index is the product of dimensions from the original indices, and the converse for the split case (figure 4.3a). It is common to denote fused indices with a thicker line.

Reshuffling

Index reshuffling is illustrated by simply twisting the indices around, swapping their location (figure 4.3b)

Transpose

A matrix transpose is merely a type of reshuffling, and consists of swapping the two indices of a rank-2 tensor (a matrix) (figure 4.3c)

4.2.3 Additional features

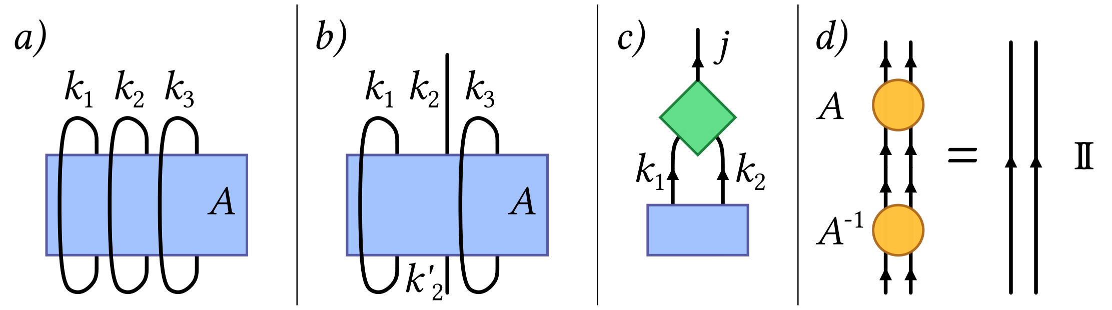


Figure 4.4: (a) The total trace of a tensor may be calculated by joining all indices with their duals, creating loops. The result is a scalar. (b) A partial trace may be performed by leaving some free indices. (c) A tensor with directed indices, representing a linear map from the space of (k_1, k_2) space to the space of j. (d) If a tensor (directed or not) is contract with its inverse, the result is the identity tensor which is represented by lines.

Trace

Traces are represented by joining a tensor's legs into loops. For a total trace $\mathrm{Tr}[A]$, the resulting contraction of the tensor has no external legs, and is therefore a scalar (figure 4.4a). A partial trace $\mathrm{Tr}_{k_1 k_3}[A] = A_{k_2}$ can be performed by only looping some of the indices (figure 4.4b), returning a tensor with lower rank.

Directed indices

Indices may be directed, in which case the direction is represented by arrows over the legs (figure 4.4c). Directed indices are used to represent specific applications of linear maps, and are particularly useful when we represent operators and their inverses in symmetrized tensors [94–96].

Identity

The identity tensor is represented by wires (directed or not), and may be obtained by contracting a tensor with its inverse (figure 4.4d).

4.3 Tensor networks

The structure of tensor diagrams suggest a novel way to look at multilinear maps, by considering the topological structure induced by tensor contractions when tensors are assembled into a network. It turns out that this has *deep* implications in quantum mechanics [87, 97].

As we have seen so far, the mathematical formalism of quantum mechanics hinges on the incredibly rich structure of Hilbert spaces. The number of available states in these spaces grows exponentially with the number of subsystems, as we have seen in previous chapters 1 and 2, which has been the main drawback in dealing with larger composite systems.

Or so it seemed. It turns out that there are very strong mathematical and physical arguments to be made that the vast majority of the Hilbert space is not only unnecessary to be considered, but outright *unphysical*, as we shall see next.

4.3.1 Tensor networks *vs.* the Hilbert space

Handling the vastness of Hilbert space may seem incredibly daunting, but a careful look at the underlying physics will be very illuminating.

Imagine a physical composite system, with multiple interacting parts. At a glance, typical descriptions of systems using kets and density matrices reveals nothing about their physical structure, geometry, length scales, or relative proximity of these parts. Since this type of relevant information is not at all explicit in these generic descriptions their generality actually *obfuscates* important information [87].

Therefore, more efficient, customized and even fundamental representations of different quantum systems must exist that reflect their properties explicitly. As we shall see, the flexible framework of tensor networks offer us precisely a method to specify such descriptions.

A tensor network is, fundamentally, a *network of correlations*. The topological structure of a tensor network tells us fundamental facts about the structure of these correlations. We can think of states represented as tensor networks as a direct representation of the state's entanglement structure itself, which is why tensor networks can efficiently and naturally represent 1, 2, and 3-dimensional systems [87].

This efficiency arises from an important observation regarding natural phenomena: interactions in Nature seem to be local, that is, all interactions seem to take a minimum finite time to occur. We will take this claim as a postulate, in the absence of experimental evidence of the contrary.

This property implies that, given certain initial conditions and a finite elapsed time, only a subset of all possible states could have been accessible by physical processes.

Poulain *et al.* have rigorously quantified this analysis [97], and have shown that, in general terms, the number of physically accessible states in polynomial time is exponentially small with respect to the number of subsystems. In their own words:

> "*Thus, we conclude that the overwhelming majority of states in the Hilbert space of a quantum many-body system can only be reached after a time scaling exponentially with the number of particles. For this reason, even for moderate size systems, these states are not physically accessible so they do not represent any state occurring in Nature.* "
>
> — David Poulain, Angie Qarry, R.D. Somma, and Frank Verstraete [97]

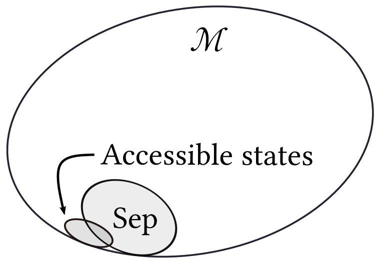


Figure 4.5: Conceptual illustration of the set of accessible states with respect to other known sets. Local states accessible in polynomial time (with respect to number of subsystems) form an exponentially small subset of the entire space of states, $\mathcal{M}$. Therefore, the *vast* majority of the space is completely irrelevant for all practical purposes as they are *unphysical.*

The authors go on to posit that Hilbert is, perhaps, nothing more than a "convenient illusion" [97]:

> "*This raises the question whether it makes sense to describe many-body quantum systems as vectors in a linear Hilbert space.*"
>
> — Poulain *et al.*

As Poulain *et al.* [97] and Orús [87] point out (and this author agrees), tensor networks seem to be not just convenient, but also more *fundamental* descriptions of quantum systems which make more efficient and representative use of the underlying properties of Hilbert spaces. Thus, the full structure of the Hilbert space may be restricted to the foundational framework of our physical theories, on top of which we build this new tensor network formalism[5].

On an intimately related idea, it is also conjectured[6] that the ground states of gapped local Hamiltonians[7] obey what is referred to as *area laws* [87, 88, 93, 101, 102]: the entanglement entropy of the ground states ρ_0 with respect to a subsystem A grows linearly with the area of the subsystem's *boundary*, ∂A, as opposed to its volume:

$$S(\rho_0^A) \approx \mathrm{Area}(\partial A).$$

It appears evident that many important and known features of quantum systems seem to be extremely restrictive, creating a hierarchy of accessibility for physical states. Such restrictions can be used to more efficiently represent quantum systems.

Thus, our goals of studying entanglement structures in large-scale composite systems now seems much more feasible than ever before. This makes our endeavors with this line of research interesting and quite exciting, and many opportunities for further research—as well as theoretical and experimental applications—are now in our horizon.

4.3.2 Types of indices

A tensor-network representation of physical systems may have two types of indices [87,88,103]:

[5]It is uncertain, at this point, if a foundational framework for quantum mechanics may be built solely on such correlation network structures, separately from the concept of Hilbert spaces. As it stands it is still, ultimately, an alternate description built on top of the traditional mathematical formulations.

[6]Indeed, it has been proven in a few cases [98–100].

[7]Hamiltonians including only terms with short-range interactions and featuring an energy gap between their ground state and first excited state. Such Hamiltonians are ubiquitous in condensed-matter physics.

- **Physical indices**, which are "visible" from outside the tensor network. These correspond to indices in the original representation of a quantum object, and typically relate to individual parts of a composite system (*e.g.* specific spins in different sites on a chain) or some input parameter (*e.g.* total charge). Sometimes these are referred to as *site-type indices.*

- **Auxiliary indices**, which are "invisible" from outside the tensor network, and are responsible for connecting the tensors with one another. These indices are responsible for carrying quantum correlations in the networks we will be using. Sometimes, these are referred to as *link-type indices*, *bonds* or *virtual indices.*

In this text we will adopt the following convention: physical indices will be denoted by subscripts such as $A_{ijk\ell}$, and auxiliary indices with superscripts such as B^{uv}. Therefore, a tensor C^{uv}_{ijk} has two auxiliary and three physical indices. In chapter 5, physical indices referring to the canonical spin-1/2 basis will be referred to by s_i and internal indices will be arbitrary.

4.3.3 Tensor network contraction

Ultimately, we will want to contract entire tensor networks to obtain numbers. This process is more subtle than it may seem at first, since there is a significant amount of freedom involved in which order such a contraction can happen, and this choice has significant consequences to the efficiency of the algorithms.

The task of finding out the most efficient way of contracting general tensor networks is a NP−HARD problem, and the specifics vary significantly on a case-by-case basis [87, 88]. There is a sort of *art* involved in developing efficient algorithms for tensor network contractions [104, 105], and some of these techniques are treated almost as trade secrets.

Ultimately, tensor networks provide an exceptionally useful tool for scientific research, but its usage and application seems to be more of an art. As expertly put by Rick Sanchez in the TV series *Rick and Morty*:

> "*Well, sometimes science is more art than science, Morty.*
> *A lot of people don't get that.*"
>
> — Rick Sanchez, *Rick and Morty*

4.4 Singular Value Decomposition

Singular value decompositions are an extremely useful way to represent arbitrary matrices. The user should be familiar with it, but in any case, details and proofs about their existence and properties may be found elsewhere [5–7]. Here, we will just recall the basic result:

Theorem 10. *(Singular Value Decomposition [5, 6]) Given any arbitrary matrix M, it can be factorized as the product*

$$M = USV^\dagger, \tag{4.2}$$

with U and $V^\dagger$ containing the left and right ***singular vectors*** *of M, and S a diagonal matrix with the* ***singular values*** *of M in its diagonal. If the singular values are chosen to be in descending order, then S is unique. If the M is a square matrix, U and $V^\dagger$ are also unique.*

There are two equivalent forms of specifying a SVD regarding the dimensions of the matrices.

If $M \in \mathbb{C}^{m\times n}$, then:

- Full SVD: $U \in \mathbb{C}^{m\times m}$, $S \in \mathbb{C}^{m\times n}$, $V^\dagger \in \mathbb{C}^{n\times n}$.

- Thin SVD: $U \in \mathbb{C}^{m\times\ell}$, $S \in \mathbb{C}^{\ell\times\ell}$, $V^\dagger \in \mathbb{C}^{\ell\times n}$.

We will opt for the thin SVD as it generalizes better for reshaped tensors and it leads itself better for truncations.

4.5 Matrix Product States

Matrix Product States (MPS) are an efficient way to represent pure states using tensor networks, and will be used as our main representation of quantum systems. Consider a state composed of n qubits given in the computational basis,

$$|\Psi\rangle = \sum_{i_1\cdots i_n=0}^{1} A_{i_1\cdots i_n} |i_1 \cdots i_n\rangle ,$$

where A is a collection of the 2^n complex coefficients constrained by $\sum |A_{i_1,\cdots,i_n}|^2 = 1$. We can write A as a rank-n tensor which contains all the coefficients which specify the state.

In terms of tensor network diagrams, such a tensor A for $n = 4$ would be a rank-4 tensor with $2^4 = 16$ elements, as depicted in figure 4.6.

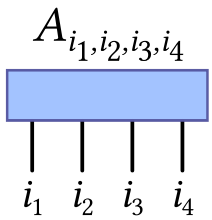


Figure 4.6: A rank-4 tensor representing the state $|\Psi\rangle$ of a 4-partite system.

This representation of a tensor network always assumes an underlying standard basis, which is what allows us to specify the indices freely. In this work and all the following examples, we will always assume the computational basis (definition 18).

Note: In what follows, we will give a generic outline of the procedure to generate a MPS decomposition of a given state, without concerns for computational efficiency or properties of the state. The idea here is simply to show how the familiar representation is connected to the MPS representation, revealing their equivalence. The following exposition should be taken as merely a didactic presentation.

In practice, one does not construct a MPS from a state directly, as that would require us to already know the full representation of the state. That misses the entire point! Instead, what we do in practice is construct a MPS representation (*e.g.* by randomizing it), then optimize the MPS using specialized algorithms so that it converges to the desired state [87, 88]. There are many algorithms available to perform state decompositions more efficiently in different scenarios [87, 88, 103]. In special cases, one may manually fill in the values of the MPS, if the representation is known in advance. We have made use of both techniques in our research.

As we have seen, we may reshape tensors by fusing and splitting their indices. We will fuse indices i_2 to i_4 into a single index $i_{(234)}$ with dimension $2^3 = 8$, and in doing so we reshape the rank-4 tensor A into a rank-2 tensor $A_{i_1,i_{(234)}}$, which is a regular 2×8 matrix.

$$A_{i_1,\cdots,i_4} \rightarrow \text{fusing indices 2 to 4} \rightarrow A_{i_1,i_{(234)}} \qquad \rightarrow \qquad A_{i_1,\cdots,i_4} \cong A_{i_1,i_{(234)}}.$$

By applying the singular value decomposition on this matrix, we obtain

$$A_{i_1,i_{(234)}} = [A^{(1)}]^{u_1}_{i_1}[S]^{u_1,v_1}[A^{(234)}]^{v_1}_{i_{(234)}},$$

where se used square brackets to better highlight the index structure. Following the notation specified earlier, we keep physical indices below and auxiliary indices on top.

We can now split the index $i_{(234)}$ back into 3 separate indices, and get:

$$A_{i_1,i_{(234)}} = [A^{(1)}]^{u_1}_{i_1}[S]^{u_1,v_1}[A^{(234)}]^{v_1}_{i_2,i_3,i_4}.$$

The net result of this operations is that we have effectively separated index i_1 from the others, and placed it in its own smaller tensor as in figure 4.7.

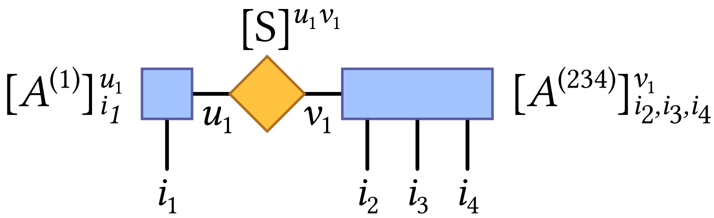


Figure 4.7: By applying the singular value decomposition on the tensor A, carefully reshaped in a specific way, we can separate the index i_1 from the others.

The diagonal matrix of singular values S (in yellow) contains structural information about the tensor. We can continue with this procedure until we have completely decomposed the rank-4 tensor into a network of rank-2 and rank-3 tensors, as in 4.8.

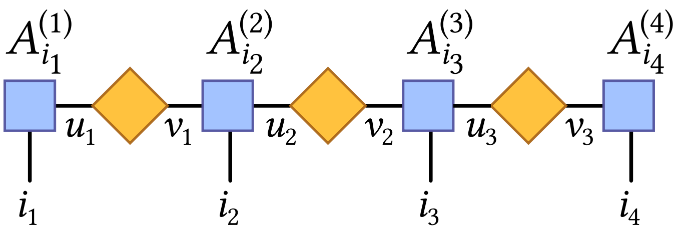


Figure 4.8: Repeating the process of isolating an index using SVDs, we obtain a tensor network structure with several auxiliary indices u_i, v_i.

At this step we have many freedoms with respect to the singular value matrices [87, 93]. We can choose to truncate the size of the tensors based on the magnitude of singular values, therefore disposing of correlated information [87, 93] between two neighboring tensors $A^{(k)}_{i_k}$ and $A^{(k+1)}_{i_{(k+1)}}$. After truncation, we may also choose to contract each of the singular value matrices either to left or right, which lead to different properties in the final representation [93]. Therefore, there is some degree of ambiguity involved as it is, which relates to the gauge freedoms in this type of representation[8].

[8] We can always insert an identify $\mathbb{I} = UU^\dagger$ between any two tensors, contracting to each side and obtaining two new tensors with the same contraction.

This ambiguity has been analyzed in detail by Perez-Garcia *et al.* [106]. For open boundary conditions, which is the case considered here, there is a canonical choice unique up to degeneracies in the spectrum of local reduced density operators, which is what is shown in figure 4.8. This canonical form is intimately related to the Schmidt decomposition between each neighboring site [87, 106], as the Schmidt coefficients are given by the singular values at each bond.

If we proceed by contracting the singular value tensors to the right (contracting the indices v_i), we obtain what is shown in figure 4.9.

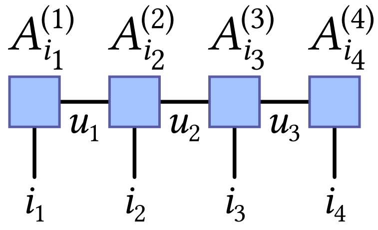


Figure 4.9: Repeating the process to all physical indices, we isolate each one of them into their own rank-2 (edges) and rank-3 (middle) tensors.

This final form of the rank-4 tensor is referred to as the *Matrix Product State* representation of the original state, $|\Psi\rangle$. Using the tensor contraction rules we have seen previously (equation 4.1) and remembering that we embedded the singular-value matrices within each tensor, we can easily see that if no truncation was performed this is an equivalent representation of the original state:

$$\sum_{i_1,i_2,i_3,i_4} [A^{(1)}]^{u_1}_{i_1}[A^{(2)}]^{u_1,u_2}_{i_2}[A^{(3)}]^{u_2,u_3}_{i_3}[A^{(4)}]^{u_3}_{i_4} |i_1,i_2,i_3,i_4\rangle = \sum_{i_1,i_2,i_3,i_4} [A^{(1234)}]_{i_1,i_2,i_3,i_4} |i_1,i_2,i_3,i_4\rangle = |\Psi\rangle\,. \tag{4.3}$$

If, however, we truncate the bond dimension of the auxiliary indices, not all states in the Hilbert space can be represented exactly in this form [87, 93].

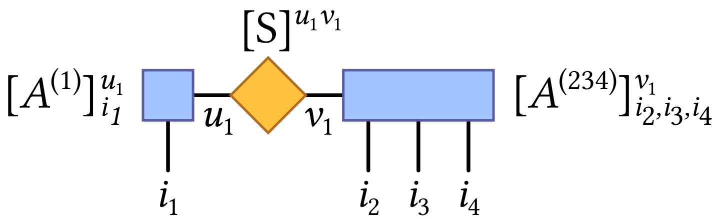
In the process of forming the Matrix Product State, we have created several additional auxiliary indices u_i, v_i. These internal indices are invisible to anything outside the network, and their presence is what makes the tensor network representation of the state efficient and, in special cases, exact.

Each one of these auxiliary indices u_i have their own independent dimension, known as the *bond dimension.* If all bond dimensions of a rank-n MPS are 1, the state is a product state and n-separable, as the contractions as described in equation 4.3 reduce to outer products between the individual tensors. All the information about correlations are regulated by bond dimensions larger than 1, and the number of describable states grows with the bond dimension [87]. Therefore, by fine-tuning the maximum bond dimension we may restrict our attention to a sufficiently small neighborhood of the entire space of states, focusing all of our computational resources precisely on the states we are interested in [87, 93].

4.5.1 Special representations

The form of the MPS as in figure 4.8, where all singular value matrices are exposed, is related to the unique representation for the MPS as described by Perez-Garcia *et al.* [106], in which the matrices S contain the Schmidt coefficients of each pair of sites [9]. This representation, often referred to as "Vidal form", is largely unique [106] and can be extremely useful in many situations, as one can opt to truncate the bond dimension in terms of these coefficients. This reveals how the bond dimension is intimately related to bipartite correlations.

There is also another special form for the MPS worth mentioning, which are the orthogonalized canonical forms [103]. Let us recall figure 4.7, in which we have only applied the SVD to the first tensor.



Since the two matrices U and $V^\dagger$ of the singular value decomposition (as in theorem 10) are unitary matrices ($UU^\dagger = \mathbb{I}$, $VV^\dagger = \mathbb{I}$) containing the singular vectors of the original tensor, tensors on the left ($A^{(1)}$) and on the right ($A^{(234)}$) are both obtained from these matrices, and are therefore representing local bijections (*i.e.* changes of basis) acting on their respective physical indices. Consequently, if we always contract S to the right and continue with the procedure, we obtain a chain of local unitaries $U^{(k)}_{k_i}$ on the left, with a general tensor on the rightmost end, as in figure 4.10a. This is known as the *left-canonical representation* of the MPS [103]. If we started from the right towards the left, we would have a chain of $V^{(k)\dagger}_{k_i}$ and the *right-canonical representation* [103]. If we had both and stopped at some location in the middle, we could keep the list of singular values separated and two large unitaries on either side. This is, effectively, the Schmidt decomposition of the entire MPS over the bipartition specified by that bond [103].

This form of representing the MPS is very computationally efficient, as if we know it is in orthogonal form and the location of its orthogonality center, then certain contractions and operations can be significantly optimized by exploiting the identity $UU^\dagger = U^\dagger U = \mathbb{I}$, as in figure 4.10b.

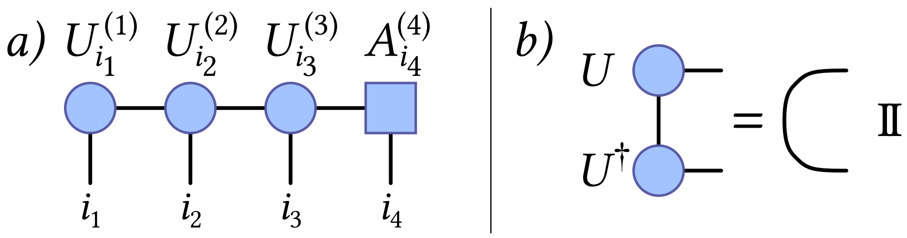


Figure 4.10: (a) In the left-canonical orthogonalized form of a MPS, all tensors on the left but the last one are local isometries U. In this form, the lone non-unitary tensor (here $A^{(4)}_{i_4}$) is referred to as the *orthogonality center* of the MPS, and contains the weights of the state in the representation defined by the change of basis of all unitaries on the left. (b) If we know the tensors are unitaries, we can immediately treat certain contractions as identity operations, massively reducing the number of contractions (and thus, computations) that need to be performed.

[9] After all, the Schmidt decomposition is simply a restatement of the singular value decomposition theorem.

4.5.2 Benefits of the MPS representation

In order to appreciate the benefits of this representation, let us analyze in detail what this approach has brought us [87, 107].

Suppose we have a 1-dimensional chain n of qudits with dimension d. The complete description of a pure state for this system requires d^n parameters (complex numbers). To compare, take the MPS representation of this state, and let us assume all the bond dimensions are m. If densely stored, each tensor[10] requires dm^2 parameters, and so the entire MPS requires ndm^2. In other words, the memory storage required for the MPS representation grows *linearly* with the number of subsystems and qudit dimension, which is a *massive* improvement from what we have seen in section 2.1. This can be improved even further by imposing symmetries and other constraints in the representation [94–96].

Storage also grows quadratically with the bond dimension, which is quite reasonable. However, the question remains of what bond dimension m is necessary to represent our state satisfactorily. It is clear that a state can always be represented exactly if $m = d^{n/2}$ [87, 108], as we would have as many parameters as in the traditional vectorial description of the state. Therefore, the exponential dependency on the number of particles has moved to the bond dimensions.

Luckily for us, this is precisely where we can make sacrifices. By truncating the bond dimension and keeping only the most significant correlations between adjacent sites, we can get an arbitrarily accurate representation of our states by simply fine-tuning the maximum value for m. Increasing m merely allows us to describe more states in the Hilbert space, but as we have seen (section 4.3.1), most of the states we will ever be interested in will be described by small bond dimensions.

Therefore, tensor networks provide an immensely efficient representation of physical states.

4.6 Matrix Product Operators

The ideas behind MPS representations can largely be generalized to operators [87, 109, 110], where similar freedoms due to gauge symmetries arise. The result is a description of operators in terms of *Matrix Product Operators* (MPOs).

A MPO can be illustrated in a similar fashion as a MPS, as shown in figure 4.11. Traditionally, the indices are *primed* to represent the dual space of the *unprimed* indices.

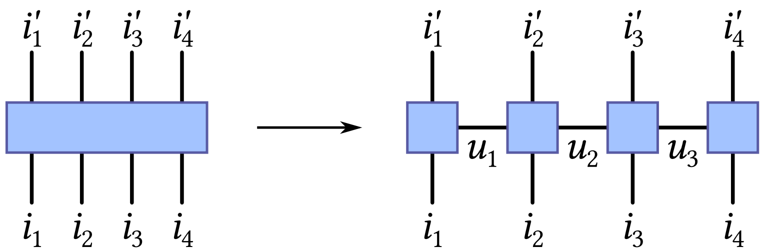


Figure 4.11: Graphical representation of a Matrix Product Operator. On the left, the operator as a single rank-8 tensor and on the right its MPO decomposition.

Just as in the case of MPS representations, by fine-tuning the bond dimensions we may adjust how precise our description of a quantum operator may be. In certain cases, operators may even be represented exactly.

[10] Ignoring the two rank-2 tensors on the edges, which contribute insignificantly as n increases.

4.7 Other representations with tensor networks

While we are restricting ourselves to 1-dimensional spin chains, which can be efficiently represented by an MPS, it is important to highlight that larger-dimensional systems can also benefit from alternative representations. For example, the *Projected Entangled Pair States* (PEPS) formalism [87, 88, 93, 107] extends the concepts developed for MPS to 2-dimensions (figure 4.12), where many extra subtleties appear. For higher dimensional systems, the *Multi-scale Entanglement Renormalization Ansatz* (MERA) approach has proven exceptionally useful [87, 88, 104].

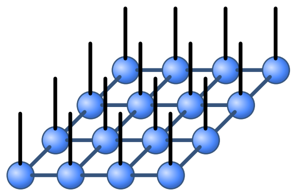


Figure 4.12: The tensor network in a *Projected Entangled Pair States* (PEPS) representation of a 2-dimensional system of 4×4 spins, which highlights how the geometric structure and topology of the tensor network is constructed as to highlight the correlations and interactions within a physical system. A PEPS representation is ideal for describing systems under nearest-neighbor interaction Hamiltonians. Periodic boundary conditions may be created by connecting the tensors on opposite edges with one another.

These are by no means the only representations possible, and the tensor networks may be constructed to specially suit several types of situations, such as specific lattices, continuous systems, infinite systems, varying boundary conditions, *etc.*

All of this is to say that tensor networks have proven themselves as an *extremely* versatile approach to tackling problems in quantum mechanics. An exciting amount of possibilities lie ahead of us in this field.

4.8 Tensor network algorithms

There is a vast and increasing literature on tensor network algorithms, so we will not attempt to cover any of them in detail. Interested readers may refer to the excellent introductions and reviews by Orús [87], Silvani *et al.* [88], Biamonte *et al.* [93], Bridgeman *et al.* [111].

We will only cover in some detail one of the earliest and most famous of the tensor network algorithms, which we use as a platform for most of our present research.

4.8.1 Density Matrix Renormalization Group (DMRG)

The most famous tensor network algorithm is arguably the *Density Matrix Renormalization Group* (DMRG), initially introduced by Steve White in 1992 [103, 112–116]. The algorithm was developed to find both the energy and ground state of Hamiltonians in 1-dimension spin chains.

DMRG is actually not just one, but an entire class of variational algorithms based on local optimizations of tensors in a MPS. The many variants on the algorithm rely on different properties of the Hamiltonians, boundary conditions, convergence rates, and approaches to handling local minima [109, 113, 115, 115, 117, 118]. We will only briefly explain the *gist* of how the algorithm works in a finite closed boundary case. A very detailed account can be found in the review by Schollwock [103].

Suppose we have a MPS representation of n qudits in a 1-dimensional chain, each representing a *physical site*, and some model Hamiltonian specifying mutual interactions (*e.g.* Ising model). The DMRG algorithm works by locally optimizing the tensors of each site based on an "effective Hamiltonian" acting locally, constructed by contracting the entire network. This can be efficiently performed by pre-computing and reusing partial contractions.

Figure 4.13 shows an illustrative example of a two-site DMRG procedure.

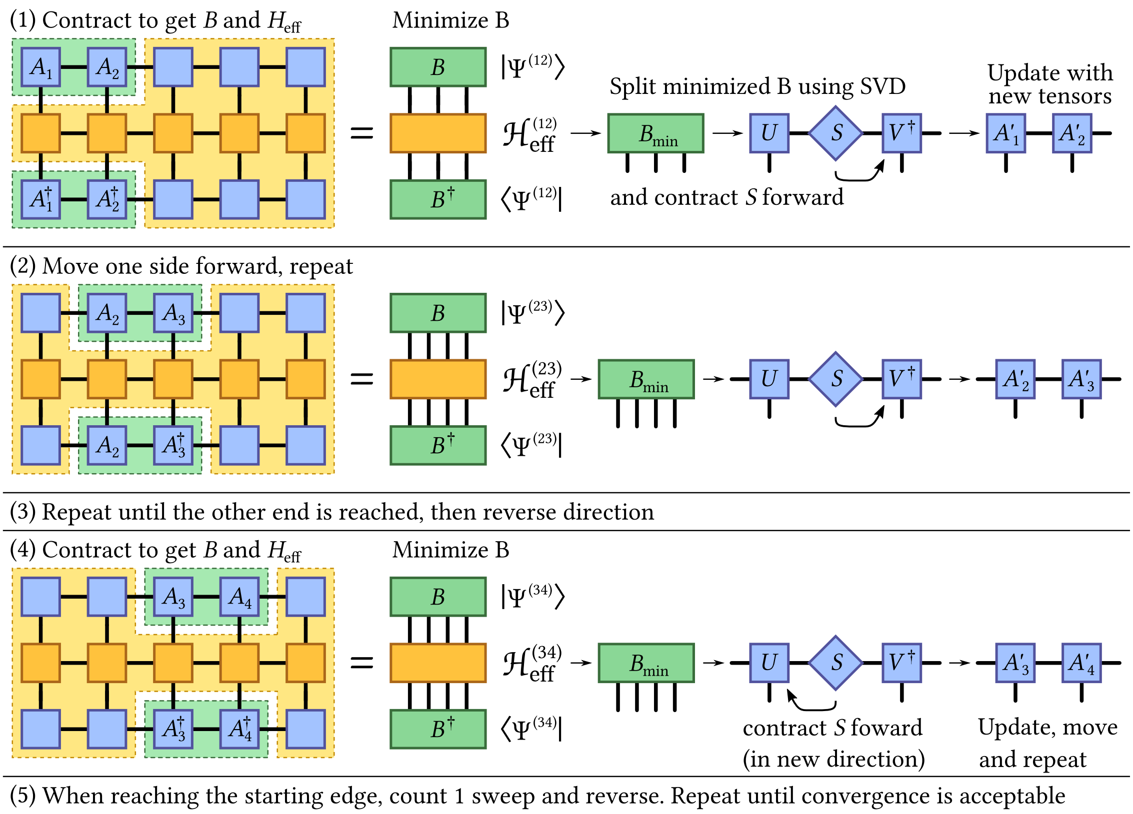


Figure 4.13: The basics of a 2-site DMRG algorithm, which finds the ground state MPS of a Hamiltonian as an MPO.

Typically, one starts with a randomized MPS in canonical form, with the starting position being the orthogonality center.

In every step of the algorithm, two site tensors are contracted into a bond tensor B and the rest of the network is contracted into an effective Hamiltonian $\mathcal{H}_{\text{eff}}$. This significantly reduces the overall problem to a smaller scale, determined by the bond dimensions. A specialized eigenvalue algorithm (*e.g.* Lanczos, Davidson) is then used to find the eigenvector for the smallest eigenvalue, which is B_{min}.

This minimized tensor is split back using a SVD, and the matrix with singular values is contracted towards the next site forward, leaving a unitary matrix behind. This effectively shifts the orthogonality center one step forward, which keeps the representation efficient.

The two tensors obtained are then updated in the original MPS, and the process repeats until the edge is reached, where the direction is reversed and the process repeats in a new direction.

Once the original starting point is reached, we have performed a single DMRG *sweep*. Sweeps are repeated as many times as necessary for convergence within desired accuracy. For first and second neighbor Hamiltonians and a 2-site DMRG, 3-5 sweeps is typically enough to obtain precision down to 5 significant figures, which shows how fast the DMRG algorithm converges.

There are variants of DMRG which optimize one site at a time and are more memory and computationally efficient [103, 116, 117], but these are typically less numerically stable and often get stuck in local minima. By optimizing two sites at the same time the algorithm ensures there's always enough local perturbations to get out of such local minima, at the cost of reduced performance. In our work we have opted to use two-site DMRGs.

Chapter 5

Entanglement witnesses with tensor networks

*"Science is what we understand well enough to explain to a computer.
Art is everything else we do."*

Donald Knuth

At the beginning we set out to solve a very daunting task, and in the first chapters we have collected several tools that allow us to perform that task, provided we look at sufficiently small systems and run computations for a sufficiently long time.

We have also seen that, through the use of tensor networks, we may tame the incomprehensible vastness of Hilbert space by restricting our attention to a tiny neighborhood of states of interest, which is exactly as large as we need it to be.

It is now time to gather all the tools we have collected thus far and use them on a few key problems to see what we can achieve and what paths forward are revealed to us.

5.1 Software used

For implementing our algorithms we have made use of existing software and libraries, which we will briefly describe next.

ITensor

Most of the code was developed using C++. For handling the tensor networks, we have opted to use the open source [119] C++ library ITensor [120] (*"Intelligent Tensor"*). The library is quite versatile and easy to use, and has many excellent features out of the box. More importantly, ITensor offers a high level of abstraction for tensor operations.

The "Intelligent" in its name comes from the fact tensor indices and their numerical representation are handled abstractly, so that tensor contractions and outer products can be performed automatically by matching common indices. This significantly reduces the mental effort required to keep track of which elements to deal with.

ITensor also uses sparse tensors and specialized algorithms by default, which means there are significant automatic optimizations running in the back end. Overall, our experience with this library has been excellent, and we recommend it to newcomers interested in using tensor networks for their research.

Also of note, during the development of this research the present author has also contributed to this software's code base [121].

MathWorks MATLAB and MOSEK

For computing SDPs and to explore the structure of entanglement witnesses, we have used MOSEK [82] running under MATLAB.

Wolfram Mathematica

For data analysis, analytical computation of matrix operations and graph plotting.

5.2 The tensor network witness model

In many tensor network algorithms such as DMRG (section 4.8.1), we are given a MPO (*e.g.* the interaction Hamiltonian for many particles) and we wish to optimize a state represented as a MPS. Typically, we randomize an initial MPS which will be locally optimized to the ground state for that Hamiltonian [87].

Our approach, however, is the opposite: we will be given the state as a MPS, which is fixed, for which we want to *optimize the (effective) MPO* representing the witness.

Figure 5.1 shows the general tensor network model we will work on. The potential methods for optimizing the individual W_i tensors (or their effective representation) is what will be our focus.

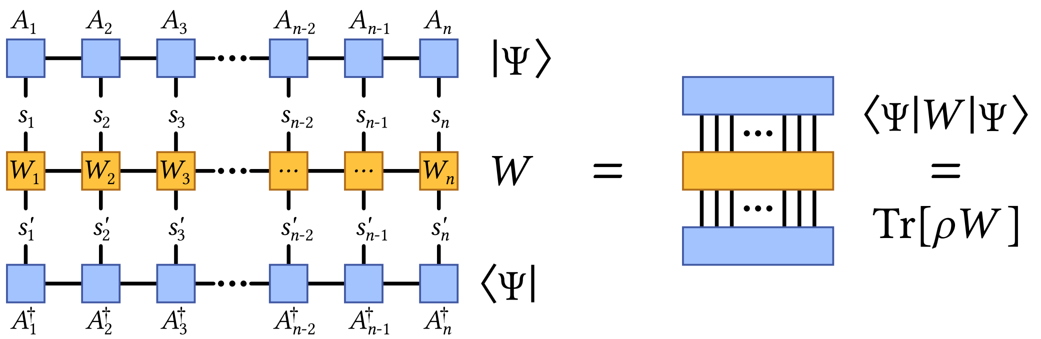


Figure 5.1: The general tensor network witness model we will be using. The pure n-partite state $|\Psi\rangle$ will be represented as a fixed MPS, which we mirror with complex conjugation ($\langle\Psi|$) to compute the contraction of the network obtaining the expectation value $\langle\Psi \mid W \mid \Psi\rangle = \mathrm{Tr}[\rho W]$. Our algorithms will attempt to find the optimal tensors W_i which minimize this value.

5.3 Why the negativity and decomposable witnesses?

In most traditional analyses of entanglement in large systems, including those using tensor networks, what it is typically done with a given (pure) state is to look at the reduced state for a subset of systems[1], *discarding* all others. Usually, one obtains the entanglement entropy (definition 38) for this reduced state, or in more careful analyses, optimize a witness for a subset of subsystems [73,74,87,122]. It is not uncommon for authors to look at the entanglement entropy between a single pair of subsystems while ignoring all the rest [123]. For pure states, it has also been shown that single-particle reduced states can impose significant constraints about the

[1]The subsystem is usually chosen to be a contiguous set of subsystems on the same length scale as the correlation length.

multi-partite entanglement structure of a large system [102] and even act a measure of global entanglement when taken together [124–126], albeit in a non-discriminatory way [125].

In light of this, it is important to point out the differences and what benefit, if any, our approach offers by looking at *the entire system at once* using a decomposable witness.

The decomposable witness is particularly useful because while it can only detect NPT entanglement, all entangled pure states are of the NPT type [7, 49]. Thus, for most practical purposes which involve pure states this computable witness and the entanglement measure it provides are sufficient, well-understood, and act as a solid initial first step.

To understand this more closely, we may look at small systems with exact solutions.

5.3.1 Representative entanglement states for 4 and 5 qubits

Our first task is to, at least qualitatively, discern the different types of information which are revealed by decomposable witnesses and the traditional approach of using the entanglement entropy.

Ancilla + 3 qubits: Bell-W-GHZ hybrid state

Two important classes of states, available for any n-partite qubit system, are the GHZ and W states. These completely symmetric states exhibit remarkable entanglement properties, and form the only two distinct classes of entangled states for 3-qubits, as shown by Dür *et al.* [127].

The n-partite GHZ pure state [18] is defined as

$$|\mathrm{GHZ}_n\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}\right)$$

GHZ states are very useful because they feature genuine n-partite entanglement in any dimension, and are often defined to be[2] the maximally-entangled states under various measures. This means any measurement in any of its parts—performed by applying partial traces—destroys all the entanglement in the state, leaving behind a fully separable state. The technique of measuring entanglement entropy via reduced density matrices (definition 38) doesn't quite work for characterizing the detailed entanglement structure of the state, as any partial trace of any kind leads to an identical maximally mixed state.

The n-partite W pure state [127] is defined as

$$|\mathrm{W}_n\rangle = \frac{1}{\sqrt{n}}\left(|10\cdots0\rangle + |01\cdots0\rangle + \cdots + |0\cdots10\rangle + |0\cdots01\rangle\right).$$

W states are also remarkable for having genuine n-partite entanglement, but also genuine k-partite entanglement for all $k \leq n$ upon tracing out $n-k$ subsystems.

With these two states and a 2-dimensional ancilla we can create an odd type of 4-qubit system, which we found useful to test some hypotheses. This is the following hybrid "Bell-W-GHZ" state:

$$|\mathrm{BWG}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \otimes |\mathrm{GHZ}_3\rangle + |1\rangle \otimes |\mathrm{W}_3\rangle\right), \tag{5.1}$$

which is, in a sense, a mixture of a Bell state with the W and GHZ states acting as the second system. This is a 4-qubit system, for which we may construct $2^{4-1} - 1 = 7$ bipartitions via partial transposition. A useful way to enumerate the bipartitions is by writing a n-bit binary

[2]For $n > 3$, there is no uniquely defined notion of a maximally entangled state.

string, using 1 for systems which are transposed and 0 for those that are not. In this way, our 7 bipartitions are given by the set

$$\{0001, 0010, 0011, 0100, 0101, 0110, 0111\},$$

where we note that inverting all the bits results in an equivalent bipartition due to $(A^{\mathrm{T}})^{\mathrm{T}} = A$, and the omission of 0000 and 1111 as they are not valid bipartitions (definition 24).

In principle, one can apply the same logic to the entanglement entropy by choosing which systems to trace out (1) and which systems to leave behind (0). Most authors seem unconcerned with this distinction[3], but for the sake of comparative thoroughness we will make it.

Using the state $\rho = |\mathrm{BWG}\rangle\langle\mathrm{BWG}|$ as specified in equation 5.1, we looked at the negativity and entanglement entropy for all the 7 bipartitions. The results are shown below in figure 5.1.

Partition	S	$\mathcal{N}$
0001, 0010, 0100	0.97987...	0.49301...
0011, 0101, 0110	1.55459...	0.98410...
0111	1	1/2

Table 5.1: Entanglement entropy and negativity for the "Bell-W-GHZ" state. We see that both detect the different mixtures of entanglement similarly, both quantitatively and qualitatively, which depends on the bipartition. **Note**: No direct comparison should be made between the two numerical quantities, as these are non-commensurable measures. Values shown are obtained analytically, but decimal approximations were used for clarity.

The results shown in table 5.1 suggest that a proper entanglement analysis *should* take into account how the system is being partitioned, which is something inherent to our approach and something which was already pointed out by others [124–126, 128]. In this case both measures had similar results, but as we will see next, there are subtleties involved.

Here, it is interesting to note that the entanglement measured is related to how a bipartition, whether by partial trace or transpose, separates two subsystems which are entangled. This suggests that the typical approach of selecting a contiguous block of subsystems would obscure information about the entanglement within that block, which turns out to be a crucial aspect of these entanglement measures for our purposes. It is also one of the reasons why we have been looking at alternative approaches as in section 5.4.2.

Ancilla + 4 qubits

As described by Verstraete *et al.* [129], entanglement in four qubits splits into nine different classes, as opposed to just two in the case for three qubits. Following the approach in the previous section, we now attach a 9-dimensional ancilla and assemble the following state:

$$|\Psi\rangle = \sum_{i=0}^{8} |i\rangle \otimes |E_i\rangle\,, \tag{5.2}$$

where $\{|E_i\rangle\}$ are the nine representative states as described in [129][4], where we assumed that all the (arbitrary) complex weights $a, b, c, d = 1$. Comparative results are shown in figure 5.2.

[3]We couldn't find any specific discussion on why this is the case, as the existence of genuine multipartite entanglement suggests a more careful analysis is warranted.

[4]In order: G_{abcd}, L_{abc_2}, $L_{a_2b_2}$, L_{ab_3}, L_{a_4}, $L_{a_20_{3\oplus\bar{1}}}$, $L_{0_{5\oplus\bar{3}}}$, $L_{0_{7\oplus\bar{1}}}$, and $L_{0_{3\oplus\bar{1}}0_{3\oplus\bar{1}}}$

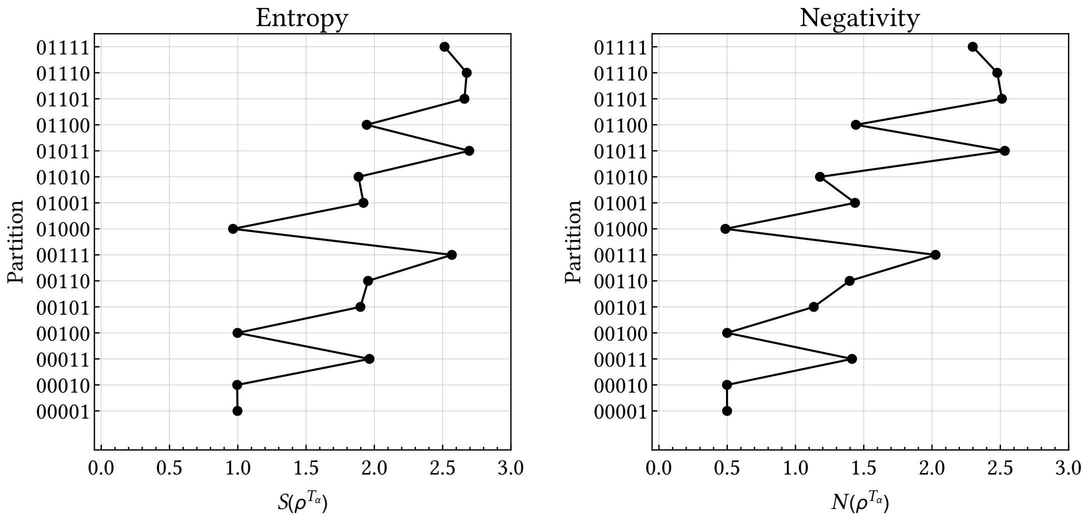


Figure 5.2: Entanglement entropy S and negativity $\mathcal{N}$ for an ancilla+4 qubits system, under different bipartitions. Points were joined with lines as a visual aid only. Here, we also see qualitative similarities between the two measures for each partition, but the two measures seem more or less sensitive to entanglement in different partitions, which suggests they are not entirely equivalent. Different "flavors" of entanglement are detected more or less by the two measures.

A careful analysis of figure 5.2 shows that there are roughly 3 classes of values being measured (low, medium and high), which correlate to the number of subsystems traced out or transposed much more than their continuity, as predicted earlier. In other words, contiguous blocks do not reveal the full picture. This is shown more clearly in figure 5.3

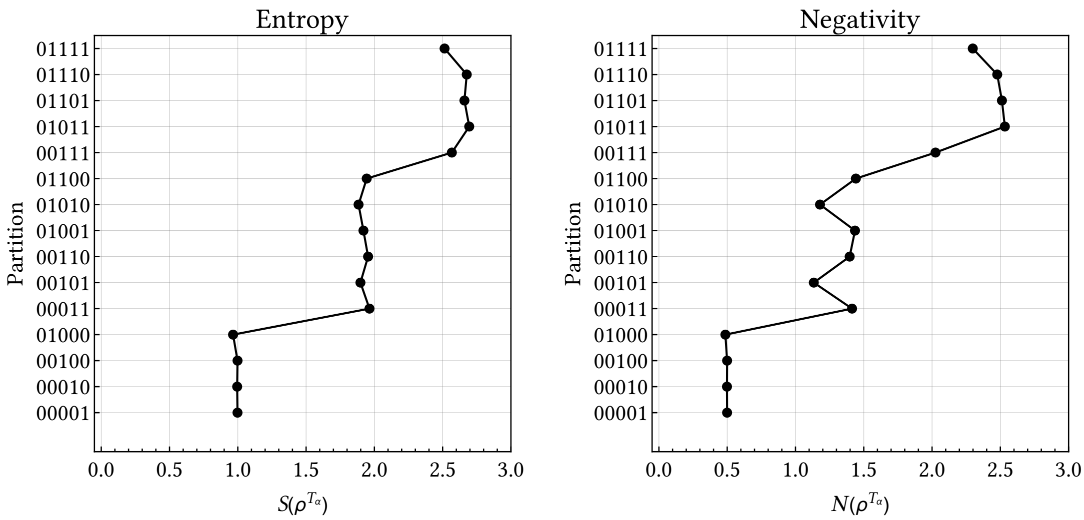


Figure 5.3: Same data as in figure 5.2, but now with the partitions sorted by number of subsystems traced out or transposed (number of 1's). As before, lines are drawn only for visual aid in qualitatively comparing the behavior of the two measures.

5.3.2 Conclusions

This brief analysis shows that there is some detailed entanglement structure within even the simplest states, for which measurement is sensitive to the way the system is partitioned. In particular, the negativity has shown itself to be more sensitive to this structure–*i.e.* displays larger relative changes–than the entanglement entropy via partial traces.

Since both operations are trivially performed under the tensor network formalism, it may be beneficial to look further into this detailed structure in the future.

5.4 Approaches and results

5.4.1 Optimal decomposable witnesses via projectors

As seen previously in section 3.2.4 (via definition 34 and theorem 7), we can construct an optimal decomposable witness by finding

$$W_{\text{d-opt}} = \sum_{\alpha} Q_{\alpha}^{\mathrm{T}_\alpha}, \quad Q_\alpha \succeq 0, \qquad \text{for all partitions } \alpha$$

and then use it to compute $\mathrm{Tr}\left[\rho W_{\text{d-opt}}\right]$. With the identity[5] $\mathrm{Tr}\left[AB^{\mathrm{T}}\right] = \mathrm{Tr}\left[A^{\mathrm{T}}B\right]$ and the linearity of the trace, we can rewrite

$$\mathrm{Tr}\left[\rho W_{\text{d-opt}}\right] = \mathrm{Tr}\left[\rho \sum_{\alpha} Q_{\alpha}^{\mathrm{T}_\alpha}\right] = \sum_{\alpha} \mathrm{Tr}\left[\rho Q_{\alpha}^{\mathrm{T}_\alpha}\right] = \sum_{\alpha} \mathrm{Tr}\left[\rho^{\mathrm{T}_\alpha} Q_\alpha\right]. \tag{5.3}$$

In other words, we can exploit the NPT property of the state to obtain Q_α directly. To figure out how to do this, we rewrite ρ^{T_α} in its spectral decomposition in terms of the eigenbasis $\{|P_{\alpha,j}\rangle\}$ and eigenvalues $\lambda_{\alpha,j}$ *in increasing order*,

$$\rho^{\mathrm{T}_\alpha} = \sum_{j} \lambda_{\alpha,j} \, |P_{\alpha,j}\rangle\langle P_{\alpha,j}| \,, \qquad (\text{where } \lambda_{\alpha,j} \text{ may be negative if } \rho \in \text{NPT})$$

such that each term in equation 5.3 becomes

$$\mathrm{Tr}\left[\rho^{\mathrm{T}_\alpha} Q_\alpha\right] = \mathrm{Tr}\left[\sum_{j} \lambda_{\alpha,j} \, |P_{\alpha,j}\rangle\langle P_{\alpha,j}| \, Q_\alpha\right] = \sum_{j} \lambda_{\alpha,j} \, \mathrm{Tr}[|P_{\alpha,j}\rangle\langle P_{\alpha,j}| \, Q_\alpha]\,. \tag{5.4}$$

From theorem 7, for this witness to be optimal we must keep only the terms such that

$$\left\langle P_{\alpha,j} \,\middle|\, \rho^{\mathrm{T}_\alpha} \,\middle|\, P_{\alpha,j} \right\rangle < 0,$$

which means we are looking only for the projectors $|P_{\alpha,j}\rangle\langle P_{\alpha,j}|$ that project to the negative space of ρ^{T_α}, that is, where $\lambda_{\alpha,j} < 0$. Let us say there are m_α such projectors. We may then write Q_α as

$$Q_\alpha = \sum_{j=1}^{m_\alpha} Q_{\alpha,j} = \sum_{j=1}^{m_\alpha} |P_{\alpha,j}\rangle\langle P_{\alpha,j}| \,,$$

so that, by orthogonality, the trace in equation 5.4 reduces to:

[5] Since the transpose is trace-preserving and cyclic, $\mathrm{Tr}\left[AB^{\mathrm{T}}\right] = \mathrm{Tr}\left[(BA^{\mathrm{T}})^{\mathrm{T}}\right] = \mathrm{Tr}\left[BA^{\mathrm{T}}\right] = \mathrm{Tr}\left[A^{\mathrm{T}}B\right]$.

$$\mathrm{Tr}[|P_{\alpha,j}\rangle\langle P_{\alpha,j}|\,Q_\alpha] = \begin{cases} 1 & j \leq m_\alpha \\ 0 & j > m_\alpha \end{cases} .$$

Finally, we obtain from equation 5.4:

$$\mathrm{Tr}\left[\rho^{\mathrm{T}_\alpha} Q_\alpha\right] = \sum_{j=1}^{m_\alpha} \lambda_{\alpha,j} = -\mathcal{N}(\rho^{\mathrm{T}_\alpha}),$$

which is merely the sum of the m_α negative eigenvalues of ρ^{T_α}. This relates directly to the negativity with respect to the 2-partition α, as seen in section 3.3.3, with only a change of sign.

Therefore, our task is to find the m_α vectors $|P_{\alpha,j}\rangle$ for each 2-partition α. This is an optimization problem we can tackle with tensor networks. In fact, this task is mathematically equivalent to a typical minimization problem to obtain the j-th lowest energy eigenstate $|\psi_j\rangle$ of a Hamiltonian $\mathcal{H}$, except we have rewritten

$$\mathcal{H} \mapsto \rho^{\mathrm{T}_\alpha} \qquad \text{and} \qquad |\psi_j\rangle \mapsto |P_{\alpha,j}\rangle .$$

As we have seen from section 4.8.1, this is precisely what the DMRG algorithm is designed to do. Therefore, we can optimize the projectors $|P_{\alpha,j}\rangle\langle P_{\alpha,j}|$ using a typical DMRG implementation, provided we ensure the $(j+1)$-th projector is orthogonal to the j-th one,

$$\langle P_{\alpha,j+1} \,|\, P_{\alpha,j}\rangle = 0,$$

which is a common feature available from the minimum eigenvalue algorithms underlying DMRG, typically used to find excited states [87].

We can look for such states $|P_{\alpha,j}\rangle$ from the bottom up, stopping at the $(m_\alpha + 1)$-th which results in $\left\langle P_{\alpha,m_\alpha+1} \,\middle|\, \rho^{\mathrm{T}_\alpha} \,\middle|\, P_{\alpha,m_\alpha+1}\right\rangle \geq 0$. Effectively, we will have at the end of this procedure

$$W_{\text{d-opt}} = \sum_\alpha Q_\alpha^{\mathrm{T}_\alpha} = \sum_\alpha \sum_{j=1}^{m_\alpha} \left(|P_{\alpha,j}\rangle\langle P_{\alpha,j}|\right)^{\mathrm{T}_\alpha} . \tag{5.5}$$

Naturally, we need not store these projectors in practice unless we wish to assemble $W_{\text{d-opt}}$ explicitly as a MPO (figure 5.4). We may choose simply to keep track of the negative values of $\lambda_{\alpha,j}$ obtained in the optimization process, which is all we need to find $\mathrm{Tr}\left[\rho W_{\text{d-opt}}\right]$.

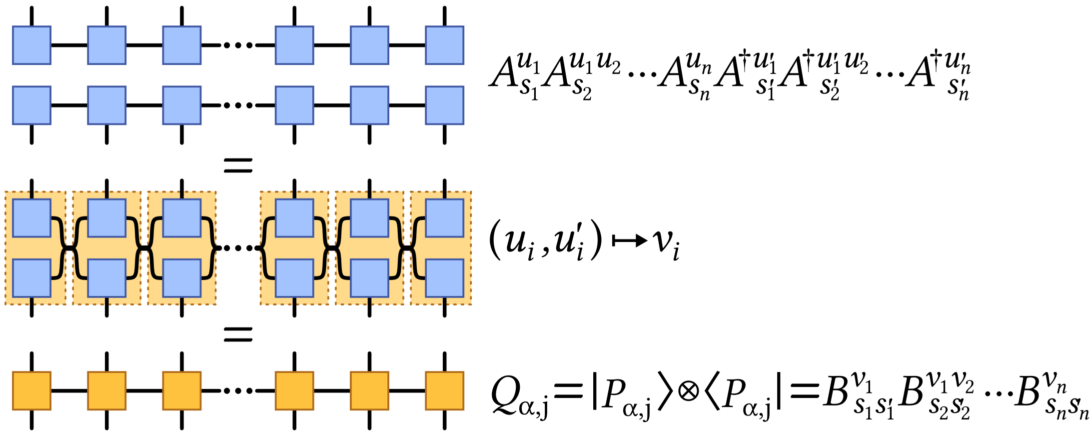


Figure 5.4: Construction of a MPO term from a MPS. For each 2-partition α, we find all orthogonal eigenvectors $|P_{\alpha,j}\rangle$ having negative eigenvalues. These are used to construct the j-th projector term for a partition α's contribution to the decomposable witness, $Q_{\alpha,j} = |P_{\alpha_j}\rangle\langle P_{\alpha_j}|$. The sum of all such terms for the partition gives us Q_α, and the sum of these for all 2-partitions gives us the optimal decomposable witness, as in theorem 7, described as a MPO.

This is, of course, a very brute force method to obtain $W_{\text{d-opt}}$ and even $\text{Tr}\left[\rho W_{\text{d-opt}}\right]$, which doesn't seem very useful at first glance. But there is more here than meets the eye.

Our first reaction is to notice that there are too many partitions to check and that too many runs of DMRG must be performed to obtain the projectors, which there may be many as well. What's the point of this ordeal, then?

It is true, the number of unique 2-partitions α which need to be checked grows exponentially as $|\{\alpha\}| = 2^{n-1} - 1$, for a n-partite system [6]. However, not all partitions are equivalent, and different heuristics will be available depending on the system [7].

One also may think that it would be better to use diagonalization algorithms for matrices and compute the negative eigenvalues of ρ^{T_α} directly. But this becomes intractable for even modest-sized systems, as we have pointed out on sections 2.1 and 4.3.1. A mere $n = 20$ system already requires diagonalization of a $1\,048\,576 \times 1\,048\,576$ matrix, which is completely unfeasible in general [8].

Still, there are clearly alternative approaches to this problem and many possible venues for optimization. We present it here only as a *proof of concept* for using tensor networks as means to obtain entanglement witnesses, and a very first step towards that line of research.

Application to the Ising model ground state

We have applied this technique on the transverse Ising model with open boundary conditions, which has well-known properties and is efficiently handled by tensor networks and DMRG. We chose a modest system with $n = 16$, which is trivial for tensor network algorithms but which would already require a $65\,536 \times 65\,536$ matrix and $32\,767$ 2-partitions.

The Hamiltonian used was,

[6] Each system may or may not be transposed, so we have 2^n possible subsets. However, $\text{Tr}[\rho W] = \text{Tr}\left[(\rho W)^{\text{T}}\right]$, and thus we only need to check half of them. And finally, we take one out to exclude the case where no system is transposed.

[7] We presently have some interesting preliminary results in this front, based on a statistical analysis of the partitions and projectors, which we hope to analyze fully in the next couple of months.

[8] For $n = 20$, a dense complex matrix would require 16 TB of storage.

$$\mathcal{H} = -\sum_{j=1}^{n-1} \sigma_{\mathbf{z}j}\sigma_{\mathbf{z}j+1} - g\sum_{j=1}^{n} \sigma_{\mathbf{x}j}, \tag{5.6}$$

where $\sigma_{\mathbf{x}}, \sigma_{\mathbf{z}}$ are Pauli matrices and we assumed $\hbar = 1$. This system displays a phase-transition in the thermodynamic limit at $g = 1.0$ [130, 131].

We sampled 21 uniformly-spaced states ranging from $g \in [0, 2]$ in total, and from each we sampled the same 10 random partitions. The algorithm took about 1 hour to run, obtaining all the projectors of the 10 partitions in the process.

As seen in figure 5.5, the negativity obtained by our method was capable of detecting the known phase transition of the Ising model, which can also be easily detected by the two-particle entanglement entropy. These results verify that our technique and algorithm work for entanglement detection, albeit still quite inefficiently.

The negativity also seemed to be significantly more sensitive to changes in the $0.0 \leq g \leq 0.6$ range, which could be related to the sensitivity to different bipartitions as shown in figure 5.3.

Conclusion

While the entropy can be trivially obtained from any tensor network, compared to the negativity, it is only a single number. And as have seen, the choice of subsystems to trace out matters, so one should also have the same concerns about which bipartition to verify as with the negativity.

However, the amount of data our approach collects is much richer, as we obtain the negative projections and projectors for the state under every partition we verify. Our preliminary results, along with known approaches to the genuine multipartite negativity [28, 73, 74] and with known results for classifying pure states [102], suggest that this information could be used to systematically unravel a detailed k-entanglement structure of general, large-scale quantum states.

Since entanglement structures of large systems are still poorly understood [7, 87, 88, 102, 132–134], such detailed surveys may prove valuable in the future. Further research is necessary in this front and we have a few paths going forward.

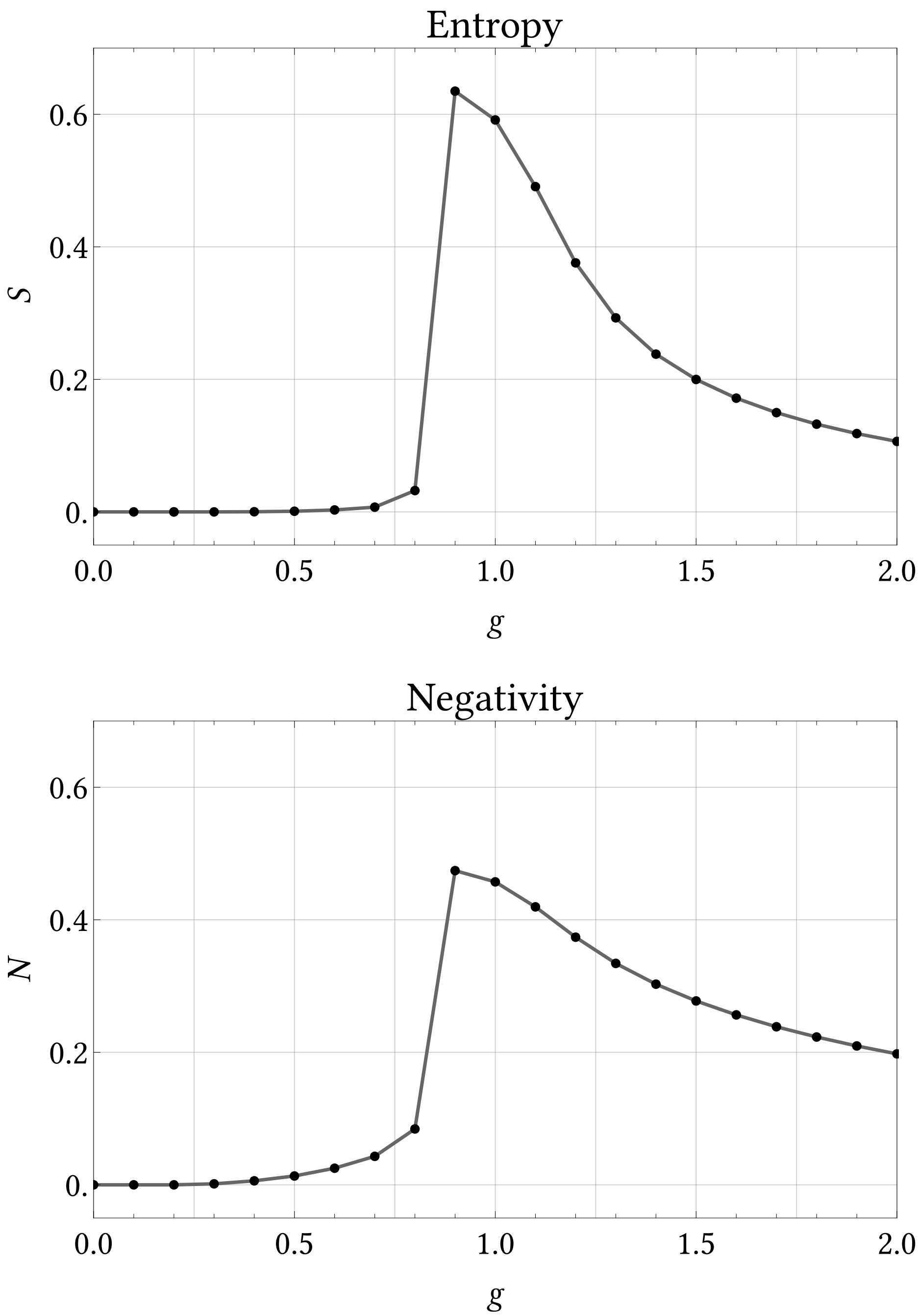


Figure 5.5: Entanglement entropy and negativity for the Ising model ground state at various values for g, where the known phase transition around $g = 1.0$ can be observed. Due to the small system, the transition appears to occur at $g < 1.0$. A system with $n = 16$ spins was simulated as a tensor network using DMRG to obtain these results. The entanglement entropy was obtained by analyzing a single site and its bond to a neighboring site. The negativity was obtained by our tensor network algorithm averaging over 10 random 2-partitions, although all of these 10 displayed similar results.

5.4.2 Hybrid tensor network with SDP blocks

Another possibility is to unite the power of SDPs and tensor networks by embedding the SDP algorithm within the local optimization for blocks in the tensor network. In this manner, it may be possible to *locally* detect genuine k-partite entanglement directly, while still considering the entire state.

While the general idea of mixing SDPs and tensor networks has been attempted before [74], such attempts have been based on explicitly discarding the rest of the system and finding the witness for the resulting reduced state. Our approach is fundamentally different, as we want to abide by our goals of looking at entanglement on the entire state directly.

As a modest first step, we propose a procedure for detecting 3-entanglement locally on the subset of three spins $\{i_\ell, j_\ell, k_\ell\}$ in the 3-partite ℓ-th block in a chain of $n = 3m$ spins ($m = 2, 3, \cdots$), as illustrated in figure 5.6. A SDP which finds the optimal fully decomposable witness (definition 36) detecting genuine 3-entanglement is given by:

$$\begin{aligned}
&\textit{minimize}: && \mathrm{Tr}[\rho W] \\
&\textit{optimizing}: && W \\
&\textit{subject to}: && \langle \Psi_{i_\ell j_\ell} \,|\, W \,|\, \Psi_{i_\ell j_\ell} \rangle \geq 0, \quad \langle \Psi_{i_\ell k_\ell} \,|\, W \,|\, \Psi_{i_\ell k_\ell} \rangle \geq 0, \quad \langle \Psi_{j_\ell k_\ell} \,|\, W \,|\, \Psi_{j_\ell k_\ell} \rangle \geq 0 \\
& && \mathrm{Tr}[W] = 1.
\end{aligned}$$

Here, the three constraints for *block-positiveness* are included to ensure the witness is *not* detecting 2-entanglement. If the result of this SDP is negative, then we can be certain our state contains genuine 3-entanglement [40, 73, 135].

In order for this setup to be locally optimizable in a stable fashion W must have translation symmetry, that is, all Q_ℓ must be equal. In practice, we may assume the optimization is iterative, and the Q_ℓ are taken as the previous iteration while we currently optimize a new one.

These constraints may also be stated explicitly as

$$[R^{(\ell)}]_{i_\ell, i'_\ell, j_\ell, j'_\ell} = \left\langle \Psi_{n \setminus i_\ell j_\ell} \,\middle|\, (Q^{u_m, v_1}_{i_1, i'_1, j_1, j'_1, k_1, k'_1} \cdots Q^{u_{m-1}, v_m}_{i_m, i'_m, j_m, j'_m, k_m, k'_m}) \,\middle|\, \Psi_{n \setminus i_\ell j_\ell} \right\rangle \succeq 0, \tag{5.7}$$

with similar constraints for (i_ℓ, k_ℓ) and (j_ℓ, k_ℓ), that is, the *matrix* obtained by fixing a block ℓ and contracting its u, v bond indices must be positive with respect to states for *all other spins* except i_ℓ, j_ℓ. This is a similar approach to the robust SDP described in section 3.5. The constraint specified in equation 5.7 can easily be incorporated into a typical SDP program, as it reduces to a simple linear relation between elements of the MPS tensors and Q_ℓ.

Due to the cyclic nature of the witness and the iterative approach to the optimization, one may write the constraint explicitly for the $(t+1)$-th iteration by referring to the values of the Q_ℓ tensors in the t-th iteration, that is, the tensor contraction in the constraint is with respect to itself, and once the SDP converges this tensor is updated with the results. We would then move to the next block and repeat the process until convergence is achieved. We also note that due to translational symmetry, $|u| = |v|$, and thus, each Q_ℓ would consist of $|u|^2$ SDP matrices for each block. This may be a bottleneck for the technique which requires further study.

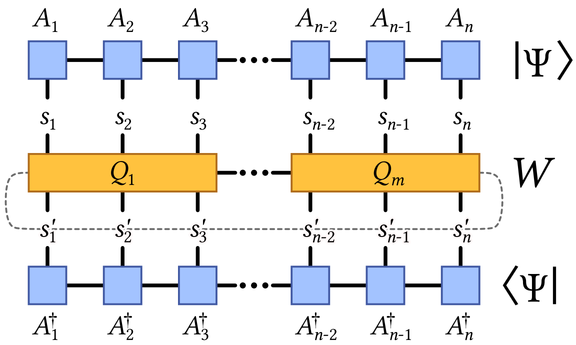


Figure 5.6: A tensor network with entanglement witness blocks. Each Q_ℓ is a rank-8 tensor made out of several local entanglement witnesses optimized by SDPs and local effective states. The dotted line represents a cyclic bond between the first and last block, referring to the translational symmetry of W. The state needs no such restriction.

Unfortunately, there are a few details that need to be sorted out at this time, and therefore we have not implemented this algorithm yet.

5.4.3 Locally-optimized witness MPOs for GMN

We have also attempted to locally optimize the MPO tensors for a fully decomposable witness (definition 36). This was, in fact, our original goal with our approach at inserting entanglement witnesses within tensor networks, which turned out to be more difficult than anticipated.

The idea was to optimize the witness MPO tensors locally, as is done with two-site DMRG on the MPS tensors, while preserving block-positiveness. Our initial algorithm relied on a tensor network adaptation of gradient descent and conjugate gradient descent methods to optimize the two-site bond (MPO) tensor.

Complications

However, we soon realized this approach did not converge even for simple states, such as a 3-qubit W state.

We identified that the issue arises because the local optimization process is incompatible with the block-positiveness constraint imposed on the witness, since this is a global property. Having to verify the global block-positiveness while performing local optimizations is already a lost cause, in terms of efficiency, so we have turned to looking into alternative and more refined approaches to the problem.

Future

We have not given up on this approach, however, and we intend on carefully studying how to best impose these types of constraints from within the tensor network formalism. This is still a work in progress.

5.5 Conclusions

While our results so far were very modest, we have successfully identified many of the caveats and complications which may arise when one attempts to adapt known techniques into a tensor network formalism. We are confident that with a few more months of work we will be able to refine our methods to something viable.

5.6 Future opportunities for research

Many alternative paths have been found on the course of our research. Naturally, we do not claim that we will pursue or achieve all of them, but merely that these ideas have been presented and set aside for further analysis. We will briefly discuss them now.

5.6.1 Detailed k-entanglement survey of large pure states

Borrowing from many methods already used in the literature for characterizing multipartite entanglement [124–126, 136, 137], it seems possible that the tensor network formalism may be used to characterize the detailed k-entanglement structure of large, strongly-correlated systems. This, however, is not related to entanglement witnesses directly.

5.6.2 Optimizing GMN witness MPOs via tensor networks

We are still pursuing our goal of creating an efficient tensor network algorithm to apply block-positiveness constraints during local optimizations. This would enable us to obtain the genuine n-entanglement (not $k < n$) witness MPO directly via DMRG-like local optimizations.

5.6.3 SDP for tensor optimization

There are many opportunities for including semidefinite programming and other matrix optimization algorithms as substeps in the optimization of tensor networks. Alternate tensor network structures beyond MPS/MPO may be useful when incorporating such techniques.

5.6.4 Phase transitions

The most immediate and natural application of our present research is to study phase transitions in large-scale spin-chains under different physical models. By studying large-scale entanglement structures in entire systems we may be able to uncover many novel physical properties which so far have eluded us due to a lack of appropriate computational techniques.

5.6.5 MPO-MPO optimization for mixed states

The work presented here focused on pure states described in the Matrix Product State (MPS) formalism. Such states are very commonly studied in large-scale numerical simulations and are typically sufficient to applied quantum mechanics in condensed-matter physics.

On the more theoretical side, we hope to extend our techniques to general mixed states so that we may begin to numerically explore the geometry of entanglement in higher-dimensional Hilbert spaces, something which is completely out of the question today. This would require a working methodology for MPO-MPO optimization.

5.6.6 SDPs *via* tensor networks

Since SDPs are linear convex optimization algorithms, we believe that there may be ways to reinterpret and adapt the known algorithms to work entirely within a tensor network formalism. If such endeavor is successful, it could pave way to an entire new field of research for large-scale optimization of linear systems, with applications far beyond quantum information.

5.6.7 Multiple dimensions

Our present research focused on 1-dimensional spin chains. While this is still a very active field of research and a rich system to explore, we hope that our eventual techniques may be generalized to 2-dimensional and 3-dimensional systems, perhaps through PEPS and MERA representations of states [87].

5.6.8 Generalized NPT k-partite entanglement

Our present research relied heavily on the existence of NPT states and the Peres criterion (theorem 4).

This means the decomposable witnesses we have used in the present research are restricted to classifying parts of composite systems into two groups, which means we are limited to what 2-partitions can reveal to us about certain systems.

The author is presently developing an unorthodox technique to generalize the Peres criterion which may be able to tackle arbitrary k-partitions, and act as an useful operational technique to detect entanglement in novel ways.

Bibliography

[1] J. von Neumann, *The Mathematical Foundations of Quantum Mechanics.* Princeton University Press, 1932.

[2] L. van Hove, "Von Neumann's contributions to quantum theory," *Bull. Amer. Math. Soc.*, vol. 64, pp. 95–100, May 1958.

[3] P. R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity.* Chelsea Pub Co, 1951.

[4] L. Ballentine, *Quantum Mechanics - A Modern Development.* World Scientific Publishing, 2000.

[5] J. Watrous, *The Theory of Quantum Information.* Cambridge University Press, Apr. 2018.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge ; New York: Cambridge University Press, 2nd ed ed., 2012.

[7] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement.* Cambridge; New York: Cambridge University Press, 2006. OCLC: 70190210.

[8] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, 1970.

[9] C. Shannon and W. Weaver, *The Mathematical Theory of Communication.* The University of Illinois Press, 1964.

[10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, 2010.

[11] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed quantum states: Linear contractions approach," *arXiv:quant-ph/0206008*, June 2002.

[12] H. Barnum, S. Wehner, and A. Wilce, "Introduction: Quantum Information Theory and Quantum Foundations," *Found Phys*, vol. 48, pp. 853–856, Aug. 2018.

[13] M. Kumar, *Quantum.* Icon Books, 2009.

[14] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?," *Phys. Rev.*, vol. 47, pp. 777–780, May 1935.

[15] J. Baggott, *Beyond Measure: Modern Physics, Philosophy, and the Meaning of Quantum Theory.* Oxford University Press, 2004.

[16] D. Bohm, *Quantum Theory.* Prentice-Hall, 1951.

[17] J. S. Bell, "On the Einstein Podolsky Rosen paradox," vol. 1, no. 3, p. 6.

[18] D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going Beyond Bell's Theorem," in *Bell's Theorem, Quantum Theory and Conceptions of the Universe* (M. Kafatos, ed.), pp. 69–72, Dordrecht: Springer Netherlands, 1989.

[19] A. M. Gleason, "Measures on the Closed Subspaces of a Hilbert Space," p. 10.

[20] S. Kochen and E. P. Specker, "The problem of hidden variables in quantum mechanics," p. 36.

[21] A. Aspect, J. Dalibard, and G. Roger, "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers," *Phys. Rev. Lett.*, vol. 49, pp. 1804–1807, Dec. 1982.

[22] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, "New High-Intensity Source of Polarization-Entangled Photon Pairs," *Phys. Rev. Lett.*, vol. 75, pp. 4337–4341, Dec. 1995.

[23] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres," *Nature*, vol. 526, pp. 682–686, Oct. 2015.

[24] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger, "Wave–particle duality of C60 molecules," *Nature*, vol. 401, pp. 680–682, Oct. 1999.

[25] K. C. Lee, M. R. Sprague, B. J. Sussman, J. Nunn, N. K. Langford, X.-M. Jin, T. Champion, P. Michelberger, K. F. Reim, D. England, D. Jaksch, and I. A. Walmsley, "Entangling Macroscopic Diamonds at Room Temperature," *Science*, vol. 334, pp. 1253–1256, Dec. 2011.

[26] D. F. Styer, *The Strange World of Quantum Mechanics.* Cambridge University Press, 2000.

[27] D. Reinsel, J. Gantz, and J. Rydning, "The Digitization of the World from Edge to Core," p. 28, 2018.

[28] G. Vidal and R. F. Werner, "A computable measure of entanglement," *Phys. Rev. A*, vol. 65, p. 032314, Feb. 2002.

[29] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: Necessary and sufficient conditions," *Physics Letters A*, vol. 223, pp. 1–8, Nov. 1996.

[30] R. F. Werner, "Quantum states with Einstein-Podolski-Rosen correlations admitting a hidden variable model," *Phys. Rev. A*, vol. 40, no. 4277.

[31] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, "Volume of the set of separable states," *Phys. Rev. A*, vol. 58, pp. 883–892, Aug. 1998.

[32] S. Szarek, "The volume of separable states is super-doubly-exponentially small," *Phys. Rev. A*, vol. 72, p. 032304, Sept. 2005.

[33] J. Kleinberg and É. Tardos, *Algorithm Design.* Addison-Wesley, 2 ed., 2006.

[34] L. Gurvits, "Quantum Matching Theory (with new complexity theoretic, combinatorial and topological insights on the nature of the Quantum Entanglement)," *arXiv:quant-ph/0201022*, Jan. 2002.

[35] A. Peres, "Separability Criterion for Density Matrices," *Phys. Rev. Lett.*, vol. 77, pp. 1413–1415, Aug. 1996.

[36] S. Woronowicz, "Positive maps of low dimensional matrix algebras," *Reports on Mathematical Physics*, vol. 10, pp. 165–183, Oct. 1976.

[37] E. Stormer, "Decomposable Positive Maps on C $_*$ -Algebras," *Proceedings of the American Mathematical Society*, vol. 86, p. 402, Nov. 1982.

[38] O. Rudolph, "Further results on the cross norm criterion for separability," *Quantum Information Processing*, vol. 4, pp. 219–239, Aug. 2005.

[39] K. Chen and L.-A. Wu, "A matrix realignmnet meethod for recognizing entanglement," p. 10.

[40] B. Jungnitsch, T. Moroder, and O. Gühne, "Taming Multiparticle Entanglement," *Phys. Rev. Lett.*, vol. 106, p. 190502, May 2011.

[41] W. Dür and J. I. Cirac, "Classification of multi-qubit mixed states: Separability and distillability properties," *Phys. Rev. A*, vol. 61, p. 042314, Mar. 2000.

[42] M. Seevinck and J. Uffink, "Partial separability and entanglement criteria for multiqubit quantum states," *Physical Review A*, vol. 78, Sept. 2008.

[43] K. Nagata, M. Koashi, and N. Imoto, "Configuration of separability and tests for multipartite entanglement in Bell-type experiments," *Phys. Rev. Lett.*, vol. 89, p. 260401, Dec. 2002.

[44] S. Yu, Z.-B. Chen, J.-W. Pan, and Y.-D. Zhang, "Classifying N-qubit Entanglement via Bell's Inequalities," *Phys. Rev. Lett.*, vol. 90, p. 080401, Feb. 2003.

[45] W. Dür and J. I. Cirac, "Multiparticle entanglement and its experimental detection," *J. Phys. A: Math. Gen.*, vol. 34, pp. 6837–6850, Sept. 2001.

[46] V. Coffman, J. Kundu, and W. K. Wootters, "Distributed Entanglement," *Phys. Rev. A*, vol. 61, p. 052306, Apr. 2000.

[47] R. Horodecki, "Informationally coherent quantum systems," *Physics Letters A*, vol. 187, pp. 145–150, Apr. 1994.

[48] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Rev. Mod. Phys.*, vol. 81, pp. 865–942, June 2009.

[49] M. Horodecki, P. Horodecki, and R. Horodecki, "Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature?," *Physical Review Letters*, vol. 80, pp. 5239–5242, June 1998.

[50] K. Zyczkowski, "On the volume of the set of mixed entangled states II," *Phys. Rev. A*, vol. 60, pp. 3496–3507, Nov. 1999.

[51] A. Sanpera, R. Tarrach, and G. Vidal, "Quantum separability, time reversal and canonical decompositions," *arXiv:quant-ph/9707041*, July 1997.

[52] P. Hyllus, C. M. Alves, D. Bruss, and C. Macchiavello, "Generation and detection of bound entanglement," *Phys. Rev. A*, vol. 70, p. 032316, Sept. 2004.

[53] J. Lavoie, R. Kaltenbaek, M. Piani, and K. J. Resch, “Experimental bound entanglement?,” *Nature Phys*, vol. 6, pp. 827–827, Nov. 2010.

[54] H. Kampermann, X. Peng, D. Bruss, and D. Suter, “Experimental generation of pseudo bound entanglement,” *Phys. Rev. A*, vol. 81, p. 040304, Apr. 2010.

[55] J. Lavoie, R. Kaltenbaek, M. Piani, and K. J. Resch, “Experimental bound entanglement in a four-photon state,” *Phys. Rev. Lett.*, vol. 105, p. 130501, Sept. 2010.

[56] R. Horodecki, “Bound entanglement – mysterious invention of nature,” *Europhysics News*, vol. 41, pp. 21–24, Nov. 2010.

[57] S. Das, T. Chanda, M. Lewenstein, A. Sanpera, A. S. De, and U. Sen, “The separability versus entanglement problem,” *arXiv:1701.02187 [quant-ph]*, Jan. 2017.

[58] K. Zyczkowski, K. A. Penson, I. Nechita, and B. Collins, “Generating random density matrices,” *Journal of Mathematical Physics*, vol. 52, p. 062201, June 2011.

[59] H.-J. Sommers and K. Zyczkowski, “Statistical properties of random density matrices,” *J. Phys. A: Math. Gen.*, vol. 37, pp. 8457–8466, Sept. 2004.

[60] Ł. Pankowski, M. Piani, M. Horodecki, and P. Horodecki, “A few steps more towards NPT bound entanglement,” *IEEE Trans. Inform. Theory*, vol. 56, pp. 4085–4100, Aug. 2010.

[61] D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and A. V. Thapliyal, “Evidence for Bound Entangled States with Negative Partial Transpose,” *Phys. Rev. A*, vol. 61, p. 062312, May 2000.

[62] S. P. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, UK ; New York: Cambridge University Press, 2004.

[63] R. A. Bertlmann, H. Narnhofer, and W. Thirring, “A Geometric Picture of Entanglement and Bell Inequalities,” *Phys. Rev. A*, vol. 66, p. 032319, Sept. 2002.

[64] B. M. Terhal, “Bell Inequalities and the Separability Criterion,” *Physics Letters A*, vol. 271, pp. 319–326, July 2000.

[65] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, “Optimization of entanglement witnesses,” *Phys. Rev. A*, vol. 62, Oct. 2000.

[66] R. Bertlmann, “Bell’s Universe: A Personal Recollection,” in *Quantum [Un]Speakables II* (R. Bertlmann and A. Zeilinger, eds.), pp. 17–80, Cham: Springer International Publishing, 2017.

[67] R. A. Bertlmann, K. Durstberger, B. C. Hiesmayr, and P. Krammer, “Optimal entanglement witnesses for qubits and qutrits,” *Phys. Rev. A*, vol. 72, Nov. 2005.

[68] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of n-particle mixed states: Necessary and sufficient conditions in terms of linear maps,” *Physics Letters A*, vol. 283, pp. 1–7, May 2001.

[69] B. Kraus, M. Lewenstein, and J. I. Cirac, “Characterization of distillable and activatable states using entanglement witnesses,” *Phys. Rev. A*, vol. 65, p. 042327, Apr. 2002.

[70] L. Clarisse, “Entanglement Distillation,” p. 192.

[71] L. Clarisse, "Entanglement Distillation; A Discourse on Bound Entanglement in Quantum Information Theory," *arXiv:quant-ph/0612072*, Dec. 2006.

[72] D. Bruss, J. I. Cirac, P. Horodecki, F. Hulpke, B. Kraus, M. Lewenstein, and A. Sanpera, "Reflections upon separability and distillability," *Journal of Modern Optics*, vol. 49, pp. 1399–1418, July 2002.

[73] M. Hofmann, T. Moroder, and O. Gühne, "Analytical characterization of the genuine multiparticle negativity," *J. Phys. A: Math. Theor.*, vol. 47, p. 155301, Apr. 2014.

[74] M. Hofmann, "Characterizing quantum correlations: the genuine multiparticle negativity as entanglement monotone," p. 103.

[75] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying Entanglement," *Phys. Rev. Lett.*, vol. 78, pp. 2275–2279, Mar. 1997.

[76] D. Bruss, "Characterizing Entanglement," *Journal of Mathematical Physics*, vol. 43, pp. 4237–4251, Sept. 2002.

[77] M. Horodecki, P. Horodecki, and R. Horodecki, "Limits for entanglement measures," *Phys. Rev. Lett.*, vol. 84, pp. 2014–2017, Feb. 2000.

[78] M. J. Donald, M. Horodecki, and O. Rudolph, "The uniqueness theorem for entanglement measures," *Journal of Mathematical Physics*, vol. 43, pp. 4252–4272, Sept. 2002.

[79] J. Eisert, "Entanglement in quantum information theory," *arXiv:quant-ph/0610253*, Oct. 2006.

[80] G. Vidal and R. Tarrach, "Robustness of entanglement," *Phys. Rev. A*, vol. 59, pp. 141–155, Jan. 1999.

[81] M. Steiner, "Generalized robustness of entanglement," *Phys. Rev. A*, vol. 67, May 2003.

[82] "Mosek ApS."

[83] "Cvxopt/cvxopt."

[84] F. G. S. L. Brandão and R. O. Vianna, "Robust semidefinite programming approach to the separability problem," *Phys. Rev. A*, vol. 70, Dec. 2004.

[85] F. G. S. L. Brandão and R. O. Vianna, "Separable Multipartite Mixed States: Operational Asymptotically Necessary and Sufficient Conditions," *Phys. Rev. Lett.*, vol. 93, Nov. 2004.

[86] F. G. S. L. Brandão, "Quantifying entanglement with witness operators," *Phys. Rev. A*, vol. 72, Aug. 2005.

[87] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," *Annals of Physics*, vol. 349, pp. 117–158, Oct. 2014.

[88] P. Silvi, F. Tschirsich, M. Gerster, J. Junemann, D. Jaschke, M. Rizzi, and S. Montangero, "The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems.," p. 120.

[89] P. Renteln, *Manifolds, Tensors, and Forms: An Introduction for Mathematicians and Physicists.* Cambridge: Cambridge University Press, 2013.

[90] M. Itskov, *Tensor Algebra and Tensor Analysis for Engineers.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[91] R. Penrose, "Applications of negative dimensional tensors," in *Combinatorial Mathematics and Its Applications*, Academic Press, 1971.

[92] P. Schmoll, S. Singh, M. Rizzi, and R. Orús, "A programming guide for tensor networks with global SU (2) symmetry," p. 54.

[93] J. D. Biamonte and V. Bergholm, "Quantum Tensor Networks in a Nutshell," p. 34.

[94] S. Singh, R. N. C. Pfeifer, and G. Vidal, "Tensor network states and algorithms in the presence of a global U(1) symmetry," *Phys. Rev. B*, vol. 83, Mar. 2011.

[95] S. Singh, R. N. C. Pfeifer, and G. Vidal, "Tensor network decompositions in the presence of a global symmetry," *Phys. Rev. A*, vol. 82, Nov. 2010.

[96] S. Singh and G. Vidal, "Tensor network states and algorithms in the presence of a global SU(2) symmetry," *Phys. Rev. B*, vol. 86, Nov. 2012.

[97] D. Poulin, A. Qarry, R. D. Somma, and F. Verstraete, "Quantum simulation of time-dependent Hamiltonians and the convenient illusion of Hilbert space," *Phys. Rev. Lett.*, vol. 106, p. 170501, Apr. 2011.

[98] M. B. Hastings, "An area law for one-dimensional quantum systems," *J. Stat. Mech.*, vol. 2007, pp. P08024–P08024, Aug. 2007.

[99] J. Eisert, M. Cramer, and M. B. Plenio, "Area laws for the entanglement entropy – a review," p. 28.

[100] M. Cramer, J. Eisert, M. B. Plenio, and J. Dreißig, "An entanglement-area law for general bosonic harmonic lattice systems," p. 15.

[101] H. He, Y. Zheng, B. A. Bernevig, and N. Regnault, "Entanglement Entropy From Tensor Network States for Stabilizer Codes," *Phys. Rev. B*, vol. 97, p. 125102, Mar. 2018.

[102] M. Walter, B. Doran, D. Gross, and M. Christandl, "Entanglement Polytopes: Multiparticle Entanglement from Single-Particle Information," *Science*, vol. 340, pp. 1205–1208, June 2013.

[103] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," *Annals of Physics*, vol. 326, pp. 96–192, Jan. 2011.

[104] G. Vidal, "Class of Quantum Many-Body States That Can Be Efficiently Simulated," *Phys. Rev. Lett.*, vol. 101, Sept. 2008.

[105] G. Vidal, "Entanglement Renormalization," *Phys. Rev. Lett.*, vol. 99, Nov. 2007.

[106] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, "Matrix Product State Representations," *arXiv:quant-ph/0608197*, Aug. 2006.

[107] F. Verstraete, V. Murg, and J. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," *Advances in Physics*, vol. 57, pp. 143–224, Mar. 2008.

[108] "Matrix Product State / Tensor Train - TensorNetwork.org."

[109] G. K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, and S. R. White, “Matrix product operators, matrix product states, and ab initio density matrix renormalization group algorithms,” *J. Chem. Phys.*, vol. 145, p. 014102, July 2016.

[110] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, “Matrix product operator representations,” *New J. Phys.*, vol. 12, p. 025012, Feb. 2010.

[111] J. C. Bridgeman and C. T. Chubb, “Hand-waving and interpretive dance: An introductory course on tensor networks,” *J. Phys. A: Math. Theor.*, vol. 50, p. 223001, May 2017.

[112] S. R. White, “Density matrix formulation for quantum renormalization groups,” *Phys. Rev. Lett.*, vol. 69, pp. 2863–2866, Nov. 1992.

[113] S. R. White, “Density matrix renormalization group algorithms with a single center site,” *Phys. Rev. B*, vol. 72, Nov. 2005.

[114] A. J. Daley, C. Kollath, U. Schollwoeck, and G. Vidal, “Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces,” *J. Stat. Mech.: Theor. Exp.*, vol. 2004, p. P04005, Apr. 2004.

[115] F. Verstraete, D. Porras, and J. I. Cirac, “DMRG and periodic boundary conditions: A quantum information perspective,” *Phys. Rev. Lett.*, vol. 93, p. 227205, Nov. 2004.

[116] U. Schollwöck, “The density-matrix renormalization group,” *Rev. Mod. Phys.*, vol. 77, no. 1, p. 57, 2005.

[117] C. Hubig, I. P. McCulloch, U. Schollwöck, and F. A. Wolf, “Strictly single-site DMRG algorithm with subspace expansion,” *Phys. Rev. B*, vol. 91, Apr. 2015.

[118] S. R. White, “Density-matrix algorithms for quantum renormalization groups,” *Phys. Rev. B*, vol. 48, no. 14, p. 10345, 1993.

[119] “ITensor - A C++ library for efficient tensor network calculations (GitHub),” June 2019.

[120] “ITensor - Intelligent Tensor Library.”

[121] “About ITensor - Collaborators and Contributors.”

[122] M. V. den Nest, W. Dür, and H. J. Briegel, “Classical spin models and the quantum stabilizer formalism,” *Phys. Rev. Lett.*, vol. 98, p. 117207, Mar. 2007.

[123] G. Vidal, “Entanglement Renormalization: An Introduction,” pp. 115–138, CRC Press, Nov. 2010.

[124] D. A. Meyer and N. R. Wallach, “Global entanglement in multiparticle systems,” *Journal of Mathematical Physics*, vol. 43, pp. 4273–4278, Sept. 2002.

[125] P. J. Love, A. M. van den Brink, A. Y. Smirnov, M. H. S. Amin, M. Grajcar, E. Il’ichev, A. Izmalkov, and A. M. Zagoskin, “A Characterization of Global Entanglement,” *Quantum Inf Process*, vol. 6, pp. 187–195, June 2007.

[126] G. Rigolin, T. R. de Oliveira, and M. C. de Oliveira, “Operational Classification and Quantification of Multipartite Entangled States,” *Phys. Rev. A*, vol. 74, p. 022314, Aug. 2006.

[127] W. Dür, G. Vidal, and J. I. Cirac, “Three qubits can be entangled in two inequivalent ways,” *Phys. Rev. A*, vol. 62, Nov. 2000.

[128] T.-C. Wei, D. Das, S. Mukhopadyay, S. Vishveshwara, and P. M. Goldbart, “Global entanglement and quantum criticality in spin chains,” *Phys. Rev. A*, vol. 71, p. 060305, June 2005.

[129] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, “Four qubits can be entangled in nine different ways,” *Phys. Rev. A*, vol. 65, p. 052112, Apr. 2002.

[130] A. H. Kole, “Density Matrix Renormalization Group calculations for the Ising Model with a Transverse Field,” p. 41, 2018.

[131] P. Pfeuty, “The one-dimensional Ising model with a transverse field,” *Annals of Physics*, vol. 57, no. 1, pp. 79–90, 1970.

[132] K. Zyczkowski and I. Bengtsson, “An Introduction to Quantum Entanglement: A Geometric Approach,” *arXiv:quant-ph/0606228*, June 2006.

[133] G. Gour and N. R. Wallach, “Classification of multipartite entanglement in all dimensions,” *Phys. Rev. Lett.*, vol. 111, p. 060502, Aug. 2013.

[134] M. Walter, D. Gross, and J. Eisert, “Multi-partite entanglement,” *arXiv:1612.02437 [cond-mat, physics:math-ph, physics:quant-ph]*, Dec. 2016.

[135] F. G. S. L. Brandão and R. O. Vianna, “Witnessed Entanglement,” *Int. J. Quantum Inform.*, vol. 04, pp. 331–340, Apr. 2006.

[136] C.-s. Yu and H.-s. Song, “Global Entanglement for Multipartite Quantum States,” *Phys. Rev. A*, vol. 73, p. 022325, Feb. 2006.

[137] G. K. Brennen, “An observable measure of entanglement for pure states of multi-qubit systems,” *arXiv:quant-ph/0305094*, May 2003.